Pólya's Theorem with Zeros

Mari Castle * Victoria Powers † Bruce Reznick ‡

February 5, 2011

1 Introduction

Let $\mathbb{R}[X] := \mathbb{R}[X_1, \ldots, X_n]$ and let $\mathbb{R}^+[X]$ denote the set of polynomials in $\mathbb{R}[X]$ with nonnegative coefficients. We write Δ_n for the standard n-simplex

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i \geq 0, \sum x_i = 1\}.$$

Pólya's Theorem [8] says that if p is a homogeneous polynomial in n variables which is positive on Δ_n, then for a sufficiently large exponent $N \in \mathbb{N}$, all of the coefficients of $(X_1 + \cdots + X_n)^N p$ are strictly positive. This elegant and beautiful result has many applications, both in pure and applied mathematics.

In [9], the second and third authors gave an explicit bound for the exponent N in terms of the degree, the size of the coefficients, and the minimum of p on the simplex. The current paper is the culmination of a project, begun in [3] and [10], to characterize forms, possibly with zeros on Δ_n, which satisfy a slightly relaxed version of Pólya's Theorem (in which the condition of "strictly positive" is replaced by "nonnegative") and to give a bound for the N needed. In this paper we give such a characterization along with a bound. This is a broad generalization of the results in [3] and [10].

There are recent results by other authors related to the work in this paper. Recently, H.-M. Mok and W.-K. To [7] gave a sufficient condition for a form to satisfy the relaxed version of Pólya's Theorem, along with a bound in this case. In [1], S. Burgdorf, C. Scheiderer, and M. Schweighofer look at more general questions on polynomial identities certifying strict or non-strict positivity of a polynomial on a closed set in $\mathbb{R}^n$. As a corollary to one of their results, they give a sufficient condition for the relaxed Pólya's Theorem to hold for a form, involving the positivity of the partial derivatives of a form on faces of the simplex. For both of these results,

*Department of Mathematics and Statistics, Kennesaw State University, Kennesaw, GA 30144. Email: mfc7379@kennesaw.edu

†Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322. Email: vicki@mathcs.emory.edu.

‡Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801. Email: reznick@math.uiuc.edu

the condition given is sufficient but not necessary; they can be deduced from our results.

The original Pólya's Theorem with bound from [9] has been used by other authors in applications. For example, in [11] it is used to give an algorithmic proof of Schmüdgen's Positivstellensatz, and in [4] it is used to give results on approximating the stability number of a graph. Also, in [6], an easy generalization of Pólya's Theorem and the bound to a noncommutative setting is given and used to construct relaxations for some semidefinite programming problems which arise in control theory. We believe that the results in this paper should have broad application to these and other areas.

2 Preliminaries

Let $Po(n,d)$ be the set of forms of degree d in n variables for which there exists an $N \in \mathbb{N}$ such that $(X_1 + \cdots + X_n)^N p \in \mathbb{R}^+[X]$. In other words, $Po(n,d)$ are the forms which satisfy the conclusion of Pólya's Theorem, with "positive coefficients" replaced by "nonnegative coefficients."

For $I \subseteq \{1,\dots,n\}$, let $F(I)$ denote the face of Δ_n given by

$$\{(x_1,\dots,x_n) \in \Delta_n \mid x_i = 0 \text{ for } i \in I\}.$$

Note that the relative interior of the face $F(I)$ is the set

$$\{(x_1,\dots,x_n) \in F(I) \mid x_j > 0 \text{ for all } j \notin I\}.$$

For $f(X) \in \mathbb{R}[X]$, $Z(f)$ denotes the real zeros of f.

Given $f = \sum_{\alpha \in \mathbb{N}^n} a_\alpha X^\alpha \in \mathbb{R}[X]$, let $\operatorname{supp}(f)$ denote $\{\alpha \in \mathbb{N}^n \mid a_\alpha \neq 0\}$ and define

$$\Lambda^+(f) := \{\alpha \in \operatorname{supp}(f) \mid a_\alpha > 0\}, \quad \Lambda^-(f) := \{\beta \in \operatorname{supp}(f) \mid a_\beta < 0\}.$$

For $\alpha = (\alpha_1,\dots,\alpha_n)$, $\beta = (\beta_1,\dots,\beta_n)$ in $\mathbb{N}^n$, we write $\alpha \preceq \beta$ if $\alpha_i \leq \beta_i$ for all i, and $\alpha \prec \beta$ if $\alpha \preceq \beta$ and $\alpha \neq \beta$.

For $\alpha = (\alpha_1,\dots,\alpha_n) \in \mathbb{N}^n$ and a face $F = F(I)$ of Δ_n, it will be convenient to use the notation α_F for $(\tilde{\alpha}_1,\dots,\tilde{\alpha}_n) \in \mathbb{N}^n$, where $\tilde{\alpha}_i = \alpha_i$ for $i \in I$ and $\tilde{\alpha}_j = 0$ for $j \notin I$. Then $\alpha_F \preceq \beta_F$ iff $\alpha_i \leq \beta_i$ for all $i \in I$. (This is denoted $\alpha \preceq_F \beta$ in [2] and [7].)

For a form $p = \sum a_\alpha X^\alpha \in \mathbb{R}[X]$, we can write $p = p^+ - p^-$ for uniquely determined $p^+, p^- \in \mathbb{R}^+[X]$. Then for $N \in \mathbb{N}$ and $d = \deg p$, we have

$$\begin{aligned}(X_1 + \cdots + X_n)^N p =& (X_1 + \cdots + X_n)^N (p^+ - p^-) \\ =& (X_1 + \cdots + X_n)^N p^+ - (X_1 + \cdots + X_n)^N p^- \\ =& \sum_{|\gamma| = N+d} A_\gamma X^\gamma - \sum_{|\gamma| = N+d} B_\gamma X^\gamma,\end{aligned}$$

where $|\gamma|$ denotes $\gamma_1 + \cdots + \gamma_n$.

We call A_γ the positive part and B_γ the negative part of the coefficient of X^γ (although $B_\gamma \geq 0$). From calculations given in [9], we have

$$A_\gamma = \sum_{\substack{\alpha \in \Lambda^+(p) \\ \alpha \preceq \gamma}} \frac{N!}{(\gamma_1 - \alpha_1)! \cdots (\gamma_n - \alpha_n)!} \cdot a_\alpha \tag{1}$$

$$B_\gamma = \sum_{\substack{\beta \in \Lambda^-(p) \\ \beta \preceq \gamma}} \frac{N!}{(\gamma_1 - \beta_1)! \cdots (\gamma_n - \beta_n)!} \cdot (-a_\beta). \tag{2}$$

We begin with some simple observations about forms in $Po(n, d)$.

Proposition 1. *Suppose $p \in Po(n, d)$.*

a. *If u is a point in the relative interior of a face F of Δ_n and $p(u) = 0$, then p vanishes everywhere in F. In particular, if $p(u) = 0$ for u an interior point of Δ_n, then p is the zero form.*

b. *$Z(p) \cap \Delta_n$ is a union of faces of Δ_n.*

c. *If $\beta \in \Lambda^-(p)$, then for every face F of Δ_n, there is $\alpha \in \Lambda^+(p)$ such that $\alpha_F \preceq \beta_F$.*

Proof. We note that (a) is easy (a proof is given in [10] and [2, Prop. 2]) and (b) follows immediately from (a).

For (c), without loss of generality we can assume $F = F(\{1, \ldots, r\})$ with $0 \leq r < n$. We have $N \in \mathbb{N}$ with $(\sum X_i)^N p \in \mathbb{R}^+[X]$. Let $\gamma = (\beta_1, \ldots, \beta_{n-1}, \beta_n + N) \in \mathbb{N}^n$; then $|\gamma| = N + d$ and $\beta \preceq \gamma$. Write the coefficient of X^γ in $(\sum X_i)^N p$ as $A_\gamma - B_\gamma$ as above; then since $\beta \preceq \gamma$, by equation (2), $B_\gamma > 0$. Since the coefficient of X^γ in $(\sum X_i)^N p$ must be nonnegative, this implies $A_\gamma > 0$ and hence, by equation (1), there is $\alpha \in \Lambda^+(p)$ with $\alpha \preceq \gamma$. This in turn implies $\alpha_F \preceq \beta_F$, which proves (c). □

Remarks 1. The conditions in the proposition are necessary but not sufficient conditions for a real form p of degree d that is nonnegative on Δ_n to be in $Po(n, d)$. This follows from [2, Example 2] or [7, Example 5.1]. Condition (c) was noticed independently in [7].

In [7], H.-N. Mok and W.-K. To give a sufficient condition for a real form p of degree d that is nonnegative on Δ_n to be in $Po(n, d)$. This condition is related to (c) in Proposition 1.

Theorem 1 ([7], Theorem 2)**.** *Suppose p is a form of degree d such that $p \geq 0$ on Δ_n, $Z(p) \cap \Delta_n$ is a union of faces of Δ_n, and p satisfies the following property: For every face F of Δ_n with $F \subseteq Z(p)$ and each $\beta \in \Lambda^-(p)$, there is $\alpha \in \Lambda^+(p)$ such that $\alpha_F \prec \beta_F$. Then $p \in Po(n, d)$.*

This theorem will follow easily from our main theorem. The new property in the above theorem is not always necessary, as shown by a simple example in [7, Example 5.2], and also by the following 1-parameter family of examples based on an example in [5].

Example. Let $p_a(X_1, X_2, X_3, X_4) = X_1^4 + X_2^4 + X_1^2(X_3^2 - aX_3X_4 + X_4^2)$, where $0 < a \le 2$. Then $p \ge 0$ on Δ_4, $Z(p) \cap \Delta_4 = F(\{1,2\})$, $\Lambda^-(p) = \{(2,0,1,1)\}$, and $\Lambda^+(p) = \{(4,0,0,0),(0,4,0,0),(2,0,0,2),(2,0,2,0)\}$. Hence all conditions of Proposition 1 hold. Note that there is no $\alpha \in \Lambda^+(p)$ with $\alpha_F \prec (2,0,1,1)_F$, where $F = F(\{1,2\})$, so that the new property in Theorem 1 doesn't hold.

If $\sum_{j=1}^4 \gamma_j = N + 4$, then by (1), the coefficient of $\frac{N!}{\gamma_1!\gamma_2!\gamma_3!\gamma_4!} X_1^{\gamma_1} X_2^{\gamma_2} X_3^{\gamma_3} X_4^{\gamma_4}$ in $(\sum_j X_j)^N p_a$ is

$$\begin{aligned} f_a(\gamma_1,\gamma_2,\gamma_3,\gamma_4) := {} & \gamma_1(\gamma_1-1)(\gamma_1-2)(\gamma_1-3) + \gamma_2(\gamma_2-1)(\gamma_2-2)(\gamma_2-3) \\ & +\gamma_1(\gamma_1-1)(\gamma_3(\gamma_3-1) + \gamma_4(\gamma_4-1) - a\gamma_3\gamma_4). \end{aligned} \tag{3}$$

We want to determine the smallest N so that for all such γ, $f_a(\gamma) \ge 0$. We first observe that

$$\begin{aligned} f_a(2,3,k,k) &= 2(2k(k-1) - ak^2) = 2k((2-a)k-2), \\ f_a(3,3,k,k) &= 6(2k(k-1) - ak^2) = 6k((2-a)k-2). \end{aligned}$$

If $a = 2$, then $f_a(3,3,k,k) < 0$, so no N will ever work. So we must take $a < 2$, whence $0 < a < 2$, and observe that $f_a(3,3,k,k) < 0$ if

$$k < \frac{2}{2-a}.$$

Thus, if $N = 2M$ and all coefficients are nonnegative, we have

$$2M + 4 = 6 + 2k \implies N = 2k + 2 \ge 2 + 2\left\lceil \frac{2}{2-a} \right\rceil,$$

and if $N = 2M + 1$ and all coefficients are nonnegative, we have

$$2M + 5 = 6 + 2k \implies N = 2k + 1 \ge 1 + 2\left\lceil \frac{2}{2-a} \right\rceil.$$

Thus the smallest N satisfies the equation

$$N \ge 1 + 2\left\lceil \frac{2}{2-a} \right\rceil.$$

A messy calculation, which we omit, shows that $p_a \in Po(n,d)$ for all $a \in (0,2)$, and there is a smallest such $N = 4(2-a)^{-1} + \mathcal{O}(1)$. In the last section, we will see that the first statement follows from our main theorem.

3 Local Versions of Pólya's Theorem

For $\alpha \in \mathbb{N}^n \setminus \{(0,\ldots,0)\}$, let $\|\alpha\|$ denote the unit vector $\frac{\alpha}{|\alpha|}$ and note that $\|\alpha\| \in \Delta_n$. The original proof of Pólya's Theorem is "coefficient by coefficient": For $p > 0$ on Δ_n and $N = 0, 1, \ldots$, let $\epsilon = \frac{1}{N+d}$; then a sequence of real polynomials p_ϵ is constructed which converge uniformly to p on Δ_n as $N \to \infty$, such that the coefficient of X^α in $(\sum X_i)^N p$ is a constant positive multiple (depending only on n and d) of $p_\epsilon(\|\alpha\|)$. Using this technique, we can obtain "local" versions of the theorem, by which we mean the result for coefficients which correspond to exponents α such that $\|\alpha\|$ lies in a given closed subset of Δ_n. To prove our main theorem, we will write Δ_n as a union of closed subsets so that we can apply one of the local versions to each of the subsets.

The key to our local versions of Pólya's Theorem and the bounds we obtain is the simple observation that the main theorem in [9] generalizes immediately to subsets of Δ_n on which the form is positive.

If $|\alpha| = d$, define $c(\alpha) := \frac{d!}{\alpha_1! \cdots \alpha_n!}$. Suppose $p \in \mathbb{R}[X]$ is homogeneous of degree d, then write

$$p(X) = \sum_{|\alpha|=d} a_\alpha X^\alpha = \sum_{|\alpha|=d} c(\alpha) b_\alpha X^\alpha,$$

and let $L(p) := \max_{|\alpha|=d} |b_\alpha|$. The following local result, which is in [3], is immediate from the proof of Theorem 1 in [9]:

Proposition 2. *Suppose $S \subseteq \Delta_n$ is nonempty and closed and $p \in \mathbb{R}[X]$ is homogeneous of degree d such that $p(x) > 0$ for all $x \in S$. Let λ be the minimum of p on S. Then for*

$$N > \frac{d(d-1)}{2} \frac{L(p)}{\lambda} - d$$

and $\alpha \in \mathbb{N}^n$ such that $|\alpha| = N+d$ and $\|\alpha\| \in S$, the coefficient of X^α in $(\sum_{i=1}^n X_i)^N p$ is nonnegative.

The above theorem will give us an N with a bound for the region of the simplex away from the zeros. Then we will apply local results which work for certain closed subsets of Δ_n whose union contains the zero set of the form. We start with notation for certain closed subsets of Δ_n containing subsets of a given face.

Definition 1. Let $F = F(I)$ be a face of Δ_n.

1. For $0 < \epsilon < 1$, let $\Delta(F, \epsilon)$ denote the following closed subset of Δ_n containing F:

$$\Big\{(x_1,\ldots,x_n) \in \Delta_n \mid \sum_{i\in I} x_i \le \epsilon\Big\} = \Big\{(x_1,\ldots,x_n) \in \Delta_n \mid \sum_{j\notin I} x_j \ge 1-\epsilon\Big\}.$$

2. We need notation for certain closed subsets of Δ_n containing the "middle" of F, i.e., the part of F away from the lower-dimensional subfaces. Given $0 < \epsilon < \tau < 1$, let
$$C(F, \epsilon, \tau) := \{(x_1, \ldots, x_n) \in \Delta(F, \epsilon) \mid x_j \geq \tau - \epsilon \text{ for } j \notin I\}.$$

3. Given $0 < \tau < 1$, define the following closed subset of the relative interior of F:
$$W(F, \tau) := \{(x_1, \ldots, x_n) \in F \mid x_j \geq \tau \text{ for } j \notin I\}.$$

Remark 2. It is easy to check that if F is a face of dimension k, and $F_1, \ldots, F_{k+1}$ are the subfaces of F of dimension $k-1$, then
$$\Delta(F, \epsilon) \subseteq C(F, \epsilon, \tau) \cup \Delta(F_1, \tau) \cup \cdots \cup \Delta(F_{k+1}, \tau).$$

The following proposition is a local result for closed neighborhoods of vertices of the simplex and is similar to Proposition 2 in [3]. For $1 \leq i \leq n$, let v_i denote the i-th vertex of Δ_n, i.e., $v_i = F(I)$, where $I = \{1, \ldots, i-1, i+1, \ldots, n\}$.

Proposition 3. *Suppose $p = \sum a_\alpha X^\alpha$ is a form of degree d such that $p \geq 0$ on Δ_n. Let $F = v_i$ for some i, and suppose that for every $\beta \in \Lambda^-(p)$ there is some $\alpha \in \Lambda^+(p)$ such that $\alpha_F \preceq \beta_F$. Let $c_{\min}$ be the minimum and $c_{\max}$ the maximum of $\{a_\alpha \mid \alpha \in \Lambda^+(p)\}$, and U the sum of the absolute values of the coefficients of p. Now define the following constants:*
$$\epsilon = \frac{c_{\min}}{c_{\max} + 2U}, \qquad s = \frac{c_{\min}}{2}\left(\frac{2c_{\min}}{c_{\max} + 2U}\right)^d.$$
Then for any $N \in \mathbb{N}$ such that
$$N > \frac{d(d-1)}{2}\frac{L(p)}{s},$$
and every $\gamma \in \mathbb{N}^n$ with $|\gamma| = N + d$ and $\|\gamma\| \in \Delta(F, \epsilon)$, the coefficient of X^γ in $(\sum X_j)^N p$ is nonnegative.

Proof. For ease of exposition assume $F = v_1$. Suppose $\beta \in \Lambda^-(p)$ and $\alpha \in \Lambda^+(p)$ such that $\alpha_F \preceq \beta_F$. This means that $\alpha_i \leq \beta_i$ for $i = 2, \ldots, n$ and hence, since p is homogeneous and $\alpha \neq \beta$, it follows that $\alpha_1 > \beta_1$.

For each $\beta \in \Lambda^-(p)$, pick one $\alpha \in \Lambda^+(p)$ such that $\alpha_F \preceq \beta_F$ and denote this α by $m(\beta)$. Then for each $\alpha \in \Lambda^+(p)$, define f_α as follows:
$$f_\alpha = a_\alpha X^\alpha + \sum_{\substack{\gamma \in \Lambda^-(p) \\ m(\gamma) = \alpha}} a_\gamma X^\gamma.$$

By the previous remark, we have
$$f_\alpha = X_2^{\alpha_2} \cdots X_n^{\alpha_n}\left(a_\alpha X_1^{\alpha_1} + q(X)\right),$$

where every monomial in $q(X)$ has X_1-degree $< \alpha_1$ and coefficient < 0. (We allow the possibility that $q(X) = 0$.)

By construction, for every $\beta \in \Lambda^-(p)$, the term $a_\beta X^\beta$ occurs in one and only one f_α. This implies that p is precisely the sum of the f_α's. It therefore suffices to show that the proposition holds for the f_α's and hence for a form of the type $f = aX_1^e + q(X)$, where $a > 0$, $0 < e \le d$, every monomial in $q(X)$ has degree in X_1 less than e, and the set of coefficients of f is a subset of the set of coefficients of p.

Let $f = aX_1^e + q(X)$ be as above and let $\tilde{U}$ be the sum of the absolute value of the coefficients of f. By [3, Lemma 2], $f \ge u$ on $\Delta_n(F, r)$, where

$$r = \frac{a}{a + 2\tilde{U}}, \quad u = \frac{a}{2}\left(\frac{2\tilde{U}}{a + 2\tilde{U}}\right)^e.$$

Since the set of coefficients of f is a subset of the set of coefficients of p and a is the coefficient of some X^α with $\alpha \in \Lambda^+(p)$, we have $\tilde{U} \le U$, $\tilde{U} \ge c_{\min}$, $a \ge c_{\min}$ and $a \le c_{\max}$. Then $a + 2\tilde{U} \le c_{\max} + 2U$, and hence,

$$r = \frac{a}{a + 2\tilde{U}} \ge \frac{c_{\min}}{c_{\max} + 2U} = \epsilon,$$

and, noting that $\frac{2c_{\min}}{c_{\max}+2U} \le 1$,

$$u = \frac{a}{2}\left(\frac{2\tilde{U}}{a + 2\tilde{U}}\right)^e \ge \frac{c_{\min}}{2}\left(\frac{2c_{\min}}{c_{\max} + 2U}\right)^d = s.$$

Since $\epsilon \le r$, $\Delta_n(F, \epsilon) \subseteq \Delta_n(F, r)$. Since $f \ge u$ on $\Delta_n(F, r)$ and $s \le u$, it follows that $f \ge s$ on $\Delta_n(F, \epsilon)$. Finally, we apply Proposition 2 to f with $S = \Delta_n(F, \epsilon)$. $\square$

We need a localized Pólya's Theorem which holds on the closed subsets $C(F, \epsilon, \tau)$ defined above. This result, without the explicit bound, is a special case of [3, Proposition 1].

Lemma 1. *Suppose $F = F(I)$ is a nonempty face of Δ_n (so that $I \ne \{1, \dots, n\}$), $p \in \mathbb{R}[X]$ is homogeneous of degree d, and we can write $p = \phi + \psi$ for forms $\phi, \psi \in \mathbb{R}[X]$ so that*

1. *$\phi > 0$ on the relative interior of F,*
2. *for $i \in I$, X_i does not occur in ϕ, and*
3. *every monomial in ψ contains at least one factor X_i for some $i \in I$.*

Given $\tau \in \mathbb{R}$ with $0 < \tau < 2/(n - |I|)$, define λ as follows: Let m be the minimum of ϕ on the closed subset $W(F, \tau/2)$ of the relative interior of F and let

$\lambda = \min(1, m)$. *Let U be the sum of the absolute values of the coefficients of p. Then for any $N \in \mathbb{N}$ with*

$$N > d(d-1)\frac{L(p)}{\lambda} - d,$$

any ϵ with

$$\epsilon < \min\left\{\frac{\lambda}{2d+2U}, \tau/2\right\},$$

and any $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$ such that $\beta_j = 0$ for all $j \notin I$, the coefficient of X^α in $(\sum X_i)^N X^\beta p$ is nonnegative for any α such that $\|\alpha\| \in C(F, \epsilon, \tau)$.

Proof. Claim: For ϵ and λ as given, $p \geq \lambda/2$ on $C(F, \epsilon, \tau)$.

Proof of claim: For ease of exposition, let $C = C(F, \epsilon, \tau)$. We want to bound ϕ and ψ on C. Given $x = (x_1, \ldots, x_n) \in \Delta(F, \epsilon)$, suppose $\gamma \in \mathrm{supp}(\psi)$. Then $x_i \leq \epsilon$ for $i \in I$ and X^γ contains a factor X_i for some $i \in I$, hence X^γ evaluated at x is $\leq \epsilon$. It follows that $|\psi(x)| \leq U\epsilon$ on $\Delta(F, \epsilon)$. Since $C \subseteq \Delta(F, \epsilon)$, we have $\psi \geq -U\epsilon$ on C.

Given $x \in C$, we have $\sum\limits_{j \notin I} x_j = 1 - t$ for $t = \sum\limits_{i \in I} x_i \leq \epsilon$. Define $a = (a_1, \ldots, a_n)$ by setting $a_j = \frac{x_j}{1-t}$ for $j \notin I$ and $a_i = 0$ for $i \in I$. Then $a \in F$ and for $j \notin I$ we have $a_j = \frac{x_j}{1-t} \geq x_j \geq \tau - \epsilon \geq \tau/2$, hence $a \in W(F, \tau/2)$. Then, since ϕ is a form of degree d containing none of the variables X_i for $i \in I$, $\phi(x) = (1-t)^d \phi(a) \geq (1-\epsilon)^d \lambda$.

Putting together the two bounds and noting that $\lambda \leq 1$, we have for $x \in C$,

$$p(x) \geq (1-\epsilon)^d \lambda - U\epsilon \geq (1 - d\epsilon)\lambda - U\epsilon = \lambda - (d\lambda + U)\epsilon \geq \lambda - (d+U)\epsilon.$$

Hence with ϵ as given we have $p(x) \geq \lambda - \lambda/2 = \lambda/2$, and the claim is proven.

Now we apply Proposition 2 to p with $S = C$. Let λ' be the minimum of p on C; then by the claim, $\lambda' \geq \lambda/2$. It follows that $N \geq \frac{d(d-1)}{2}\frac{L(p)}{\lambda'} - d$ and hence, by Proposition 2, the coefficient of X^α in $(\sum X_i)^N p$ is nonnegative for $\|\alpha\| \in C$. It remains to show that this still holds if we replace p by $X^\beta p$.

Suppose $\gamma \in \mathrm{supp}((\sum X_i)^N X^\beta p)$ with $\|\gamma\| \in C$; then $\gamma = \alpha + \beta$ for some $\alpha \in \mathrm{supp}((\sum X_i)^N p)$. Since $\beta_j = 0$ for $j \notin I$, we have $\gamma_j = \alpha_j$ for $j \notin I$. Then since $|\gamma| \geq |\alpha|$, it is easy to see that $\|\gamma\| \in C$ implies $\|\alpha\| \in C$. It follows that the coefficient of X^α in $(\sum X_i)^N p$ is nonnegative and therefore the coefficient of X^γ in $(\sum X_i)^N X^\beta p$ is nonnegative, completing the proof. $\square$

Remark 3. The dependence of ϵ on τ is due to the fact that λ depends on τ.

4 Pólya's Theorem with Zeros

In this section we give necessary and sufficient conditions for a form to be in $Po(n, d)$ and a bound on the exponent N needed. One condition is the necessary condition (3) from Proposition 1 for the faces of Δ_n in $Z(p)$. The second condition involves the positivity of certain forms related to p on the relative interior of the faces in $Z(p)$.

For a form p and a face F of Δ_n, we say that $\alpha \in \Lambda^+(p)$ is *minimal with respect to* F if there is no $\gamma \in \Lambda^+(p)$ such that $\gamma_F \prec \alpha_F$. Note that it is possible to have $\alpha_F \prec \beta_F$ and $\beta_F \prec \alpha_F$ for distinct α and $\beta \in \Lambda^+(p)$. We start with notation for certain subforms of p related to the elements minimal with respect to a face.

Definition 2. Suppose $p = \sum a_\alpha X^\alpha \in \mathbb{R}[X]$ is homogeneous.

1. For $\Gamma \subseteq \operatorname{supp}(p)$, let $p(\Gamma)$ denote the form $\sum_{\gamma \in \Gamma} a_\gamma X^\gamma$.
2. For $\alpha \in \operatorname{supp}(p)$ and a face F of Δ_n, define

$$p(\alpha, F) := p(\{\gamma \in \operatorname{supp}(p) \mid \gamma_F = \alpha_F\})/X^{\alpha_F}.$$

   Note that $p(\alpha, F)$ is a nonzero form in the variables $\{X_j \mid j \notin I\}$, where $F = F(I)$, and $X^{\alpha_F} p(\alpha, F)$ is a subform of p.

Lemma 2. *Suppose $p \in Po(n, d)$ and F is a face of Δ_n. Then for every $\alpha \in \Lambda^+(p)$ which is minimal with respect to F, the form $p(\alpha, F)$ must be strictly positive on the relative interior of F.*

Proof. Suppose $F = F(I)$ and set $q := p(\alpha, F)$. Since $p \in Po(n, d)$, there exists $N \in \mathbb{N}$ such that $(\sum_{i=1}^n X_i)^N p \in \mathbb{R}^+[X]$. We claim that $(\sum_{j \notin I} X_j)^N X^{\alpha_F} q \in \mathbb{R}^+[X]$.

Suppose γ is in the support of $(\sum_{j \notin I} X_j)^N X^{\alpha_F} q$; then by the definition of q, $\gamma_F = \alpha_F$. Consider the coefficient of X^γ in $(\sum_{i=1}^n X_i)^N p$ and let A_γ be the positive part and B_γ the negative part, as in (1) and (2) in §2. Contributions to A_γ come from $\delta \in \Lambda^+(p)$ with $\delta \preceq \gamma$, which implies $\delta_F \preceq \gamma_F = \alpha_F$. Since α is minimal with respect to F, it follows that the only contributions to A_γ come from $\delta \in \Lambda^+(p)$ with $\delta_F = \alpha_F$. Since all such δ are in $\operatorname{supp}(X^{\alpha_F} q)$, it follows that A_γ is also the positive part of the coefficient of X^γ in $(\sum_{j \notin I} X_j)^N X^{\alpha_F} q$. Since $X^{\alpha_F} q$ is a subform of p, the negative part of the coefficient of $(\sum_{j \notin I} X_j)^N X^{\alpha_F} q$ is clearly $\leq B_\gamma$ and it follows that the coefficient of X^γ in $(\sum_{j \notin I} X_j)^N X^{\alpha_F} q \geq A_\gamma - B_\gamma \in \mathbb{R}^+$.

From the claim it follows that $(\sum_{j \notin I} X_j)^N q \in \mathbb{R}^+[X]$. Since q is a form in $\{X_j \mid j \notin I\}$, this means that q satisfies Pólya's Theorem on the simplex $F(I)$. Hence, by Proposition 1, q is strictly positive on the relative interior of F. $\square$

Theorem 2. *Given p, a nonzero form of degree d, such that $p \geq 0$ on Δ_n and $Z(p) \cap \Delta_n$ is a union of faces. Then $p \in Po(n, d)$ if and only if for every face $F \subseteq Z(p)$ the following two conditions hold:*

1. *For every $\beta \in \Lambda^-(p)$, there is $\alpha \in \Lambda^+(p)$ so that $\alpha_F \preceq \beta_F$.*
2. *For every $\alpha \in \Lambda^+(p)$ which is minimal with respect to F, the form $p(\alpha, F)$ is strictly positive on the relative interior of F.*

Proof. Condition (1) is necessary by Proposition 1 and condition (2) is necessary by Lemma 2.

Suppose the conditions hold for $p = \sum a_\alpha X^\alpha \in \mathbb{R}[X]$. Since p is not identically zero, $\Delta_n \not\subseteq Z(p)$ (since p is homogeneous), so any face in $Z(p) \cap \Delta_n$ must be proper, i.e., must have dimension $\leq n-2$. For a closed subset of Δ_n which does not intersect the zero set, we can apply Proposition 2, hence the theorem will follow easily from the following claim:

Claim: For every face F contained in $Z(p) \cap \Delta_n$ we can find ϵ such that $0 < \epsilon < 1$ and $N \in \mathbb{N}$ so that for any $\theta \in \mathbb{N}^n$ with $|\theta| = N + d$ and $\|\theta\| \in \Delta(F, \epsilon)$, the coefficient of X^θ in $(X_1 + \cdots + X_n)^N p$ is nonnegative.

Proof of claim: We prove the claim by induction on the dimension of F. If the dimension is 0, then we are done by Proposition 3.

Now suppose F has dimension k, $1 \leq k \leq n-2$, and the claim is true for all subfaces of F of dimension $k-1$. Let τ be less than the minimum of $2/(1+k)$ and the ϵ's that occur among these subfaces and $\tilde{N}$ the maximum of the N's.

By assumption, for each $\beta \in \Lambda^-(p)$, there exists at least one $\alpha \in \Lambda^+(p)$ such that $\alpha_F \preceq \beta_F$. Let $\alpha = \alpha(\beta)$ be such a vector which is minimal with respect to F. Construct a subset Ω of $\Lambda^+(p)$ by the following procedure: Let $\Lambda^-(p) = \{\beta^{(1)}, \ldots, \beta^{(m)}\}$ and set $\Omega = \emptyset$. Now for $i = 1, 2, \ldots, m$, if there is no $\gamma \in \Omega$ with $\gamma_F \preceq \beta^{(i)}_F$, add $\alpha(\beta^{(i)})$ to Ω. Suppose α, α' are distinct elements of Ω and that in the above constuction of Ω, α was added to Ω before α' was added. Then there must be some $\beta \in \Lambda^-(p)$ such that $\alpha'_F \preceq \beta_F$ but $\alpha_F \not\preceq \beta_F$. Then the transitivity of $\preceq$ implies that $\alpha_F \not\preceq \alpha'_F$, and this last implies $\alpha'_F \not\preceq \alpha_F$ by the minimality of α with respect to F.

To summarize, we have constructed a subset Ω of the set of $\alpha \in \Lambda^+(p)$ which are minimal with respect to F such that for every $\beta \in \Lambda^-(p)$, there is $\alpha \in \Omega$ such that $\alpha_F \preceq \beta_F$. Further, for any two $\alpha, \alpha' \in \Omega$, we have that $\alpha_F \not\preceq \alpha'_F$ and $\alpha'_F \not\preceq \alpha_F$. Order the set Ω in some way and, one at a time, for each $\alpha \in \Omega$ define a form ψ_α as follows: Let Γ_α be the set of $\beta \in \Lambda^-(p)$ such that $\alpha_F \prec \beta_F$ and β is not contained in any previously defined Γ_α. Now let $\psi_\alpha = p(\Gamma_\alpha)/X^{\alpha_F}$. Then ψ_α is a form (possibly the zero form) and every monomial of ψ_α contains at least one variable X_i for $i \in I$. Furthermore, by construction, if $\alpha, \gamma \in \Omega$ and $\alpha \neq \gamma$, then $\mathrm{supp}(\psi_\alpha) \cap \mathrm{supp}(\psi_\gamma) = \emptyset$.

Now, for each $\alpha \in \Omega$, let $\phi_\alpha = p(\alpha, F)$ and consider the subform $X^{\alpha_F}(\phi_\alpha + \psi_\alpha)$ of p. By definition of Ω, $\mathrm{supp}(\phi_\alpha) \cap \mathrm{supp}(\phi_\gamma) = \emptyset$ if $\alpha \neq \gamma$. By assumption, $p(\alpha, F)$ is strictly positive on the relative interior of F. Hence ϕ_α, ψ_α, and α_F satisfy the conditions of Lemma 1 and thus we may apply the lemma taking the β to be α_F; we conclude that there is some $N_\alpha \in \mathbb{N}$ and $\epsilon_\alpha > 0$ such that the coefficient of X^γ in $(X_1 + \cdots + X_n)^{N_\alpha} X^{\alpha_F}(\phi_\alpha + \psi_\alpha)$ is nonnegative for all $\gamma \in \mathbb{N}^n$ with $\|\gamma\| \in C(F, \epsilon_\alpha, \tau)$.

By construction, for every $\beta \in \Lambda^-(p)$, the term $a_\beta X^\beta$ in p occurs in $X^{\alpha_F}(\phi_\alpha + \psi_\alpha)$ for some unique $\alpha \in \Omega$. Hence we can write

$$p = \sum_{\alpha \in \Omega} X^{\alpha_F}(\phi_\alpha + \psi_\alpha) + \tilde{p}, \tag{4}$$

where $\tilde{p}$ has only positive coefficients. Let $\epsilon > 0$ be less than the minimum of the ϵ_α's, $\tau/2$, and $\lambda/(2d+2U)$, where λ and U are defined as in Lemma 1, and let M be the maximum of the N_α's; then for any $\gamma \in \operatorname{supp}((X_1 + \cdots + X_n)^M p)$ with $\|\gamma\| \in C(F, \epsilon, \tau)$, the coefficient of X^γ is nonnegative. Let $F_1, \ldots F_{k+1}$ denote the subfaces of F of dimension $k-1$. By the inductive hypothesis, we have $\tilde{N}$ such that for $i = 1, \ldots, k+1$ and any $\gamma \in \operatorname{supp}((X_1 + \cdots + X_n)^{\tilde{N}} p)$ with $\|\gamma\| \in \Delta(F_i, \tau)$, the coefficient of X^γ is nonnegative. Since $\Delta(F, \epsilon) \subseteq C(F, \epsilon, \tau) \cup \Delta(F_1, \tau) \cup \cdots \cup \Delta(F_{k+1}, \tau)$, the claim now follows.

Now write $Z(p) \cap \Delta_n$ as a union of faces $G_1 \cup \cdots \cup G_l$, where $G_i \not\subseteq G_j$ for any $i \neq j$, and apply the claim to each G_i; say we have that the claim holds with ϵ_i and N_i. Let S be the closure of $\Delta_n \setminus \bigcup_{i=1}^{l} \Delta(G_i, \epsilon_i)$; then $p > 0$ on S. By Proposition 2, there is M such that for every $\theta \in \mathbb{N}^n$ with $\|\theta\| \in S$ the coefficient of X^θ in $(X_1 + \cdots + X_n)^M p$ is nonnegative. Taking N to be the maximum of M and the N_i's, we are done.

□

Remark 4. The sufficient condition for p to be in $Po(n,d)$ given in [7] implies the sufficient condition given in Theorem 2 above. Hence [7, Theorem 2] follows from Theorem 2.

Corollary 1. *Suppose p is a form of degree d with $p \geq 0$ on Δ_n and $Z(p) \cap \Delta_n$ is a union of faces. Suppose further that for every face $F \subseteq Z(p)$ and every $\beta \in \Lambda^-(p)$, there exists $\alpha \in \Lambda^+$ such that $\alpha_F \prec \beta_F$. Then $p \in Po(n,d)$.*

Proof. If the given condition holds for p, then the first condition of Theorem 2 holds trivially. For every α which is minimal with respect to F, by the given condition, there is no $\beta \in \Lambda^-(p)$ such that $\beta_F = \alpha_F$. Hence every $p(\alpha, F)$ has only positive coefficients and thus must be strictly positive on the relative interior of F. By Theorem 2, this implies $p \in Po(n,d)$. □

We now give a bound on the exponent N needed in Theorem 2. The bound will depend on the degree of p, the size of the coefficients, and constants which are defined recursively in terms of minimums of the $p(\alpha, F)$'s on a certain closed subset of the relative interior of F. We begin with the definition of these constants.

Definition 3. Suppose $p \in \sum a_\alpha X^\alpha$ is a form of degree d and F is a face of Δ_n such that either F is a vertex or p satisfies (2) of Theorem 2 on every subface G of F (including $G = F$), i.e., for every $\alpha \in \Lambda^+(p)$ which is minimal with respect to G, the form $p(\alpha, G)$ is strictly positive on the relative interior of G. Let U be the sum of the absolute value of the the coefficients of p, and let $c_{\min}$ (respectively, $c_{\max}$) denote the mimimum (respectively, maximum) of $\{a_\alpha \mid \alpha \in \Lambda^+(p)\}$.

We will recursively define constants $N(F)$ and $\epsilon(F)$ which will correspond to the N and ϵ of the claim in the proof of Theorem 2. First we define ϵ_0, λ_0, and N_0 as the ϵ, s and N in Proposition 3:

$$\epsilon_0 = \min\left\{\frac{c_{\min}}{c_{\max} + 2U}, \frac{1}{n}\right\}, \quad \lambda_0 = \frac{c_{\min}}{2}\left(\frac{2c_{\min}}{c_{\max} + 2U}\right)^d, \quad N_0 = \frac{d(d-1)}{2}\frac{L(p)}{\lambda_0}.$$

Suppose $\dim F = k$; then we define ϵ_i, λ_i, and N_i for $i = 1, \ldots, k$ recursively as follows: Suppose G is a subface of F of dimension i. By hypothesis, if $\alpha \in \Lambda^+(p)$ is minimal with respect to G, then $p(\alpha, G)$ is strictly positive on the relative interior of G and hence $p(\alpha, G)$ has a positive minimum m_α on $W(G, \epsilon_{i-1}/2)$. Let $\lambda(G)$ be the minimum of the m_α's, taken over the set of α's which are minimal with respect to G, or 1 if this minimum is larger than 1. Now let λ_i be the minimum over all G's of the $\lambda(G)$'s, and choose any positive ϵ_i with

$$\epsilon_i < \min\left\{\frac{\lambda_i}{2d + 2U}, \epsilon_{i-1}\right\}, \tag{5}$$

and define N_i as

$$\max\left\{N_{i-1}, \frac{d(d-1)}{2}\frac{L(p)}{\lambda_i}\right\}.$$

Finally, set $N(F)$ to be N_k and $\epsilon(F)$ to be ϵ_k.

Theorem 3. *Suppose $p \in Po(n, d)$, so that $Z(p) \cap \Delta_n$ is a union of faces of Δ_n. Write $Z(p) \cap \Delta_n = G_1 \cup \cdots \cup G_l$, where each G_i is a face and $G_i \not\subseteq G_j$ for all $i \neq j$. Let $M = \max\{N(G_i) \mid 0 \leq i \leq l\}$, $\epsilon = \min\{\epsilon(G_i) \mid 0 \leq i \leq l\}$, and define τ as the minimum of p on the closure of $\Delta_n \setminus (\Delta(G_1, \epsilon) \cup \cdots \cup \Delta(G_l, \epsilon))$. Then $(X_1 + \cdots + X_n)^N p \in \mathbb{R}^+[X]$ for*

$$N > \max\left\{M, \frac{d(d-1)}{2}\frac{L(p)}{\tau}\right\}.$$

Proof. Suppose N satisfies the inequality. For each G_i, following the proof of the claim in Theorem 2, we see that for any $\theta \in \mathbb{N}^n$ with $|\theta| = N(G_i) + d$ and $\|\theta\| \in \Delta(G_i, \epsilon(G_i))$, the coefficient of X^θ in $(X_1 + \cdots + X_n)^{N(G_i)} p$ is nonnegative. Since $N \geq N(G_i)$ for all i and $\epsilon \leq \epsilon(G_i)$, this holds with $N(G_i)$ replaced by N and $\epsilon(G_i)$ replaced by ϵ. By Proposition 2, for any $\theta \in \mathbb{N}^n$ with $|\theta| = N + d$ and $\|\theta\|$ in the closure of $\Delta_n \setminus (\Delta(G_1, \epsilon) \cup \cdots \cup \Delta(G_l, \epsilon))$, the coefficient of X^θ in $(X_1 + \cdots + X_n)^N p$ is nonnegative. Since Δ_n is the union of these sets, the theorem follows. $\square$

Example. We continue with our example from Section 2. For $0 < a \leq 2$, we have

$$p = X_1^4 + X_2^4 + X_1^2 X_3^2 + X_1^2 X_4^2 - a X_1^2 X_3 X_4.$$

Recall $Z(p) \cap \Delta_n$ is the face $F = F(\{1, 2\})$. There are two elements α of $\Lambda^+(p)$ which are minimal with respect to F: $(2, 0, 2, 0)$ and $(2, 0, 0, 2)$. In both cases, the form $p(\alpha, F) = q := X_3^2 + X_4^2 - a X_3 X_4$. Note that $q > 0$ on the relative interior of F iff $a < 2$ and hence Theorem 2 says that $p \in Po(n, d)$ iff $a < 2$, as claimed in Section 2.

We now compute the bound from Theorem 3. We are interested in the behavior as $a \to 2$; hence there is no harm in assuming $a \geq 1$. The first step is to compute the constants $\epsilon = \epsilon(F)$ and $M = N(F)$ from Definition 3. We have $d = n = 4$,

$L(p) = 1$, $c_{\min} = c_{\max} = 1$, and $U = 4 + a$, hence $\epsilon_0 = \frac{1}{9+2a}$, $\lambda_0 = \frac{1}{2}\left(\frac{2}{9+2a}\right)^4$, and $N_0 = \frac{3}{4}(9+2a)^4$.

Next we need to find λ_1, which is the minimum of q on $W(F, \epsilon_0/2)$. It's easy to check that, on F, $q(0, 0, t, 1-t) = 1 - (2+a)(t - t^2)$, which achieves its minimum $\frac{2-a}{4}$ at $(0, 0, 1/2, 1/2)$ and that this is the global minimum on $W(F, \epsilon_0/2)$. Hence $\lambda(F) = \lambda_1 = \frac{2-a}{4}$ and $N_1(F) = \max\{N_0, 6\frac{1}{\lambda_1}\} = \max\{\frac{3}{4}(9+2a)^4, \frac{24}{2-a}\}$. This means that for a very close to 2, the M in the statement of Theorem 3 is $\frac{24}{2-a}$.

Finally, we need to estimate ϵ_1 and then estimate τ, the minimum of p on S, the closure of $\Delta_4 \setminus \Delta(F, \epsilon_1)$. By definition, we need

$$\epsilon_1 < \min\left\{\frac{\lambda_1}{2d + 2U}, \epsilon_0\right\} = \min\left\{\frac{2-a}{64+8a}, \frac{1}{9+2a}\right\} = \frac{2-a}{64+8a}.$$

There are no critical points in the relative interior of Δ_n, whence the minimum occurs on the boundary of S. Observe that $p(e_1, 0, (1-e_1)/2, (1-e_1)/2) = e_1^4$, hence $\tau \le e_1^4$ and $\frac{d(d-1)}{2}\frac{L(p)}{\tau} > 6\frac{(64+8a)^4}{(2-a)^4}$. This means that the estimate from the theorem is several orders of magnitude worse than the true value computed earlier.

Question: The previous computation shows that the bound from the theorem is not sharp for our example. Is there a non-trivial example for which the bound from the theorem is sharp? If not, is a better bound possible?"

References

[1] S. Burgdorf, C. Scheiderer, and M. Schweighofer, *Pure states, nonnegative polynomials and sums of squares*, preprint.

[2] M. Castle, *Pólya's Theorem with Zeros*, Ph.D. thesis, Emory University, 2008.

[3] M. Castle, V. Powers, and B. Reznick, *A quantitative Pólya's Theorem with zeros*, Journal of Symbolic Computation **44** (2009), no. 9, 1285 – 1290, Effective Methods in Algebraic Geometry.

[4] E. de Klerk and D. V. Pasechnik, *Approximation of the stability number of a graph via copositive programming*, SIAM Journal on Optimization **12** (2002), no. 4, 875–892.

[5] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 2nd ed., Cambridge Univ. Press, 1952.

[6] C.W.J. Hol and C.W. Scherer, *Matrix sum-of-squares relaxations for robust semidefinite programs*, Math. Programming **107** (2006), 198–211.

[7] H.-N. Mok and W.-K. To, *Effective Pólya semi-positivity for non-negative polynomials on the simplex*, Journal of Complexity **24** (2008), no. 4, 524 – 544.

[8] G. Pólya, *Über positive Darstellung von Polynomen*, Vierteljahrschrift Naturforsch. Ges. Zürich **73** (1928), 141–145, in *Collected Papers* **2** (1974), MIT Press, 309-313.

[9] V. Powers and B. Reznick, *A new bound for Pólya's Theorem with applications to polynomials positive on polyhedra*, Journal of Pure and Applied Algebra **164** (2001), no. 1-2, 221 – 229.

[10] V. Powers and B. Reznick, *A quantitative Pólya's Theorem with corner zeros*, Proceedings of the 2006 International Symposium on Symbolic and Algebraic Computation (J.-G. Dumas, ed.), ACM Press, 2006, pp. 285–290.

[11] M. Schweighofer, *An algorithmic approach to Schmüdgen's Positivstellensatz*, Journal of Pure and Applied Algebra **166** (2002), no. 3, 307 – 319.