

Bounds on Absolute Positiveness of Multivariate Polynomials

Prashant Batra

Institute for Computer Technology, Hamburg University of Technology, 21071 Hamburg, Germany

Vikram Sharma

Max Planck Institute for Informatics, 66123 Saarbrücken, Germany

Abstract

We propose a general framework for obtaining bounds on absolute positiveness of multivariate polynomials. We show that a known bound by Hong is a nearly optimal bound within this framework. We also derive quality results for any bound in this framework, and then propose a general approach to improve the quality of any such bound.

Key words: Root bounds, multivariate polynomials, absolute positiveness.

1. Introduction

Root bounds are functions that operate on univariate polynomials with complex coefficients and compute an upper bound on the absolute value of its roots. The literature contains many root bounds; see, e.g., [16, Chap. 6]. A subclass of root bounds which is often useful in practice is the class of **absolute root bounds**, i.e. root bounds that depend only upon the absolute value of the coefficients of the polynomial. Van der Sluis [15] studied these special root bounds for their effectiveness and quality, and showed that the root bound by Fujiwara [7] is nearly optimal among all absolute root bounds, and is also tight relative to the largest absolute value among all roots of the polynomial. Often, however, one is interested in the special case of upper bounds on just the positive real roots of a polynomial with real coefficients; for instance, in the continued fraction based algorithms for real root isolation [2]. For this special case, Kioustelidis [11] did a study similar to Van der Sluis', and showed that the Knuth-Johnson bound [10] is nearly

Email addresses: batra@tuhh.de (Prashant Batra), vsharma@mpi.de (Vikram Sharma).

optimal among all root bounds that depend only on the negative coefficients of the polynomial. Recently Akritas et al. [3], generalizing a result of Ştefănescu [13], proposed a broader framework for obtaining bounds on the positive real roots of univariate polynomials. We remark here that the above bounds on positive real roots are not known to be tight relative to the largest positive root of the polynomial, whereas in the complex case we have the quality result for Fujiwara’s bound.

In fact, Hong [8] showed that most of the above root bounds are bounds for *absolute positiveness* of a polynomial, i.e., a real number such that the polynomial and all its non-vanishing derivatives are positive for any value greater than this real number. His results, however, are derived in the more general setting of multivariate polynomials, where the notion of absolute positiveness is the following: A multivariate polynomial $P(x_1, \dots, x_n)$ with real coefficients is said to be **absolutely positive** from a real value B iff P and all its non-zero partial derivatives of arbitrary order are positive for $x_1 \geq B, \dots, x_n \geq B$. The infimum of all such bounds for a polynomial $P(x_1, \dots, x_n)$ is called its **threshold of absolute positiveness**, and is denoted by A_P . In *loc. cit.*, Hong derived a bound on absolute positiveness that is essentially tight relative to the threshold of absolute positiveness of a polynomial.

In this paper, we propose a general scheme (see Theorem 2.2) for obtaining bounds on absolute positiveness, and show that the bound by Hong can be obtained as a special case. Moreover, we derive optimality results for any bound within this general scheme (see Theorem 2.4), and show that the bound by Hong is nearly optimal within this framework (see Theorem 2.3). This paper thus generalizes the investigations into the complex one-dimensional case by Van der Sluis. We also propose ways to improve upon the bounds obtained within this general scheme, and suggest an independent possible improvement for the specific bound by Hong (see §2.2). For the special case of univariate polynomials, we show that any bound (of bounded overestimation) within the framework has to be of comparable algebraic complexity as Hong’s bound (see Section 3).

2. A Framework for Bounds on Absolute Positiveness

We begin with some useful notations (we tacitly assume that $0 \in \mathbb{N}$).

- For $\pi := (\pi_1, \dots, \pi_n) \in \mathbb{N}^n$ let $|\pi| := \pi_1 + \dots + \pi_n$, and $\pi! := \pi_1! \dots \pi_n!$.
- For $\pi, \nu \in \mathbb{R}^n$, define $\pi - \nu := (\pi_1 - \nu_1, \dots, \pi_n - \nu_n)$.
- For $\pi, \nu \in \mathbb{R}^n$, we introduce the partial ordering ‘ \geq' ’ and write $\pi \geq \nu$ if for all $i = 1, \dots, n$, $\pi_i \geq \nu_i$, and write $\pi > \nu$ if $\pi \geq \nu$ and $\pi \neq \nu$.
- Let $X := (x_1, \dots, x_n)$, and for $\pi \in \mathbb{N}^n$ we denote by X^π the monomial $x_1^{\pi_1} \dots x_n^{\pi_n}$.
- Define

$$\Omega_n := 1 - \sqrt[n]{\frac{1}{2}}. \tag{1}$$

This quantity plays an important role in the following results. We remark that $1/\Omega_n = \Theta(n)$ (see [8, p. 4]).

Let $P(X) \in \mathbb{R}[X]$ be a multivariate polynomial. We say that a **monomial** $a_\pi X^\pi$, $a_\pi \neq 0$, **dominates the monomial** $a_\nu X^\nu$, $a_\nu \neq 0$, iff $\pi > \nu$. A monomial $a_\pi X^\pi$, $a_\pi \neq 0$, is said to be a **dominating monomial** for $P(X)$ iff no other monomial in $P(X)$ dominates it; since the ordering of the monomials is a partial ordering, there may be several dominating monomials for $P(X)$. We introduce the sets

$$\nu(P) := \{\nu \in \mathbb{N}^n \mid a_\nu X^\nu \text{ is a monomial in } P(X) \text{ and } a_\nu < 0\}$$

and

$$\pi(P) := \{\pi \in \mathbb{N}^n \mid a_\pi X^\pi \text{ is a monomial in } P(X) \text{ and } a_\pi > 0\}.$$

Hong and Jakus [9] showed that a bound for absolute positiveness for $P(X)$ exists iff every dominating monomial of $P(X)$ has a positive coefficient. To avoid trivialities, we assume the following in the remainder of the paper.

(*) The polynomial $P(X) \in \mathbb{R}[X]$ is such that $|\nu(P)| > 0$, $|\pi(P)| > 0$, and all its dominating monomials are positive.

In the rest of the section we will introduce a family of bounds on absolute positiveness that encompasses the bound by Hong. We begin with the following special case.

Lemma 2.1. Let $P(X)$ be of the form

$$\sum_{\pi \in I} a_\pi X^\pi + a_\nu X^\nu$$

with $a_\nu < 0$, $a_\pi > 0$ and $\pi > \nu$ for all $\pi \in I \subset \mathbb{N}^n$. Then every number larger than

$$B := \min_{\pi \in I} \left(\frac{|a_\nu|}{a_\pi} \right)^{1/|\pi-\nu|} \quad (2)$$

is a bound on absolute positiveness of $P(X)$. Moreover, $B < A_P/\Omega_n$, where A_P denotes the threshold of absolute positivity.

Proof. For $X > (B, \dots, B)$ (i.e. for all $x_i > B$, $i = 1, \dots, n$) it is evident that $P(X) > 0$. Thus we only have to show that the same holds true for all the non-trivial partial derivatives $P^{(\lambda)}(X)$ of $P(X)$ (i.e. those $\lambda \in \mathbb{N}^n$ such that $0 < \lambda < \nu$).

We know that for $X > (0, \dots, 0)$

$$\begin{aligned} P^{(\lambda)}(X) &= \sum_{\pi \in I} \frac{\pi!}{(\pi-\lambda)!} a_\pi X^{\pi-\lambda} + \frac{\nu!}{(\nu-\lambda)!} a_\nu X^{\nu-\lambda} \\ &= \frac{\nu!}{(\nu-\lambda)!} X^{-\lambda} \left(\sum_{\pi \in I} \frac{\frac{\pi!}{(\pi-\lambda)!}}{\frac{\nu!}{(\nu-\lambda)!}} a_\pi X^\pi + a_\nu X^\nu \right) \\ &> \frac{\nu!}{(\nu-\lambda)!} X^{-\lambda} \left(\sum_{\pi \in I} a_\pi X^\pi + a_\nu X^\nu \right) \\ &= \frac{\nu!}{(\nu-\lambda)!} X^{-\lambda} P(X), \end{aligned}$$

where the penultimate step follows from the observation that $\frac{\nu!}{(\nu-\lambda)!} < \frac{\pi!}{(\pi-\lambda)!}$. Thus, for all $X > (B, \dots, B)$ we have $P^{(\lambda)}(X) > 0$. Hence B , given by formula (2), is a bound on absolute positiveness of $P(X)$.

We now proceed to show the quality statement. From (2) it is clear that

$$a_\pi \leq \frac{|a_\nu|}{B^{|\pi-\nu|}}.$$

Thus $\forall X > (0, \dots, 0)$

$$P(X) \leq |a_\nu| X^\nu \left(\sum_{\pi \in I} \frac{X^{\pi-\nu}}{B^{|\pi-\nu|}} - 1 \right).$$

In particular, for all $t > 0$,

$$P(t, \dots, t) \leq |a_\nu| t^{|\nu|} \left(\sum_{\pi \in I} \frac{t^{|\pi-\nu|}}{B^{|\pi-\nu|}} - 1 \right).$$

Thus, the positive root of the polynomial inside the brackets in RHS is bounded by A_P . Hence, the positive root α of the polynomial

$$Q(u) := \sum_{\pi \in I} u^{|\pi-\nu|} - 1$$

is such that $\alpha < A_P/B < 1$. But for $u \in (0, 1)$ we have

$$\begin{aligned} Q(u) &= \sum_{\pi > \nu} u^{|\pi-\nu|} - 1 \\ &\leq \sum_{\pi \geq \nu} u^{|\pi-\nu|} - 2 \\ &= \sum_{d_1 \geq \pi_1 \geq \nu_1} u^{\pi_1 - \nu_1} \dots \sum_{d_n \geq \pi_n \geq \nu_n} u^{\pi_n - \nu_n} - 2 \\ &< \sum_{\pi_1 \geq \nu_1} u^{\pi_1 - \nu_1} \dots \sum_{\pi_n \geq \nu_n} u^{\pi_n - \nu_n} - 2 \\ &\leq \sum_{\pi_1 \geq 0} u^{\pi_1} \dots \sum_{\pi_n \geq 0} u^{\pi_n} - 2 \\ &= \frac{1}{(1-u)^n} - 2. \end{aligned}$$

Thus, as $Q(\alpha) = 0$ we have $\alpha > \Omega_n$, which gives us the desired inequality $B < A_P/\Omega_n$.
□

We next generalize the above result to any multivariate polynomial $P(X)$.

Theorem 2.2. Let $P(X) = \sum_{\iota \in I} a_\iota X^\iota \in \mathbb{R}[X]$, $I \subset \mathbb{N}^n$, satisfy (*). Let $\Delta := [\delta_{\pi, \nu}]$, $\delta_{\pi, \nu} \in [0, 1]$, be a coefficient weight matrix such that

$$\text{for any } \pi \in \pi(P) \quad \sum_{\nu \in \nu(P): \nu < \pi} \delta_{\pi, \nu} \leq 1 \text{ and for any } \nu \in \nu(P) \quad \sum_{\pi \in \pi(P): \pi > \nu} \delta_{\pi, \nu} > 0 \quad (3)$$

(and all other entries zero). Then every number larger than the value

$$B_\Delta := \max_{\nu \in \nu(P)} \min_{\pi > \nu, a_\pi > 0} \left(\frac{|a_\nu|}{\delta_{\pi, \nu} a_\pi} \right)^{1/|\pi-\nu|} \quad (4)$$

is a bound on absolute positiveness of $P(X)$.

Proof. Given the conditions on the weights $\delta_{\pi,\nu}$ it is clear that we can write $P(X)$ as

$$P(X) = \sum_{\nu \in \nu(P)} P_\nu(X) + R(X)$$

where

$$P_\nu(X) := \sum_{\pi > \nu, a_\pi > 0} \delta_{\pi,\nu} a_\pi X^\pi + a_\nu X^\nu$$

and $R(X)$ is a polynomial with non-negative coefficients. Using Lemma 2.1 we get that

$$\min_{\pi > \nu, a_\pi > 0} \left(\frac{|a_\nu|}{\delta_{\pi,\nu} a_\pi} \right)^{1/|\pi-\nu|}$$

is a bound on the absolute positiveness of $P_\nu(X)$. Clearly the maximum among these minima over all $\nu \in \nu(P)$ is a bound on the absolute positiveness of $P(X)$. \square

Different choices of $\delta_{\pi,\nu}$ yield different bounds on absolute positiveness. In particular, by using the matrix $H := [h_{\pi,\nu}]$ where

$$h_{\pi,\nu} := \Omega_n^{|\pi-\nu|} \text{ if } \nu \in \nu(P) \text{ and } \pi > \nu, \quad h_{\pi,\nu} := 0 \text{ otherwise,} \quad (5)$$

we obtain the absolute positivity bound in [8], namely

$$B_H = \frac{1}{\Omega_n} \max_{\nu \in \nu(P)} \min_{\pi > \nu, a_\pi > 0} \left(\frac{|a_\nu|}{a_\pi} \right)^{1/|\pi-\nu|}. \quad (6)$$

To show that the choice of weights in (5) satisfies the requirements in (3) we only need to show that for all $\pi \in \pi(P)$, $\sum_{\nu: \pi > \nu} \Omega_n^{|\pi-\nu|} \leq 1$. This is indeed correct because

$$\begin{aligned} \sum_{\nu: \pi > \nu} \Omega_n^{|\pi-\nu|} &= \sum_{\nu: \pi \geq \nu} \Omega_n^{|\pi-\nu|} - 1 \\ &= \sum_{\pi_1 \geq \nu_1 \geq 0} \Omega_n^{\pi_1 - \nu_1} \dots \sum_{\pi_n \geq \nu_n \geq 0} \Omega_n^{\pi_n - \nu_n} - 1 \\ &= \sum_{\pi_1 \geq \nu_1 \geq 0} \Omega_n^{\nu_1} \dots \sum_{\pi_n \geq \nu_n \geq 0} \Omega_n^{\nu_n} - 1 \\ &< \sum_{\nu_1 \geq 0} \Omega_n^{\nu_1} \dots \sum_{\nu_n \geq 0} \Omega_n^{\nu_n} - 1 \\ &= \left(\frac{1}{1 - \Omega_n} \right)^n - 1 \\ &= 1, \end{aligned}$$

where the last step follows from the definition of Ω_n (see (1)).

We next show that the coefficient weight matrix H (specified above by (5)) is a reasonable choice among all valid choices of $\delta_{\pi,\nu}$.

Theorem 2.3. Given a polynomial $P(X)$ satisfying (*), for all choices of a coefficient weight matrix Δ satisfying the conditions in Theorem 2.2 we have

$$\frac{B_H}{B_\Delta} \leq \frac{1}{\Omega_n}.$$

Proof. Suppose the bound B_H is obtained at a certain index $\nu \in \nu(P)$. Then

$$B_H = \frac{1}{\Omega_n} \min_{\pi > \nu, a_\pi > 0} \left(\frac{|a_\nu|}{a_\pi} \right)^{1/|\pi-\nu|}.$$

Moreover, from the definition of B_Δ in Theorem 2.2 we know that

$$B_\Delta \geq \min_{\pi > \nu, a_\pi > 0} \left(\frac{|a_\nu|}{\delta_{\pi,\nu} a_\pi} \right)^{1/|\pi-\nu|}.$$

Thus,

$$\frac{B_H}{B_\Delta} \leq \frac{1}{\Omega_n} \min_{\pi > \nu, a_\pi > 0} \delta_{\pi,\nu}^{1/|\pi-\nu|} \leq \frac{1}{\Omega_n}.$$

□

Hong showed [8, p. 4] that $\frac{1}{\Omega_n} \leq \frac{n}{\ln 2} + \frac{1}{2} + \frac{\ln 2}{12n}$, so that Theorem 2.3 tells us that B_H is close to any good bound.

2.1. A General Quality Bound

Let B_Δ be the bound defined in Theorem 2.2. How does B_Δ compare with the actual threshold of absolute positiveness A_P of $P(X)$? We show the following.

Theorem 2.4. Assumptions and Definitions as in Theorem 2.2. Let B_Δ be the bound defined in (4). Let $d_i := \deg_{x_i}(P)$, and $D := d_1 + \dots + d_n$. We claim that

$$B_\Delta < \frac{A_P D}{\Gamma_\Delta \ln 2}, \quad (7)$$

where

$$\Gamma_\Delta := \min_{\nu \in \nu(P)} \min_{\pi > \nu, a_\pi > 0} \delta_{\pi,\nu}^{1/|\pi-\nu|},$$

for non-vanishing $\delta_{\pi,\nu}$'s.

Proof. Suppose the bound B_Δ is obtained for a certain ν . Consider the ν th derivative of $P(X)$

$$P^{(\nu)}(X) = \nu! \left(\sum_{\pi > \nu} \binom{\pi}{\nu} a_\pi X^{\pi-\nu} + a_\nu \right). \quad (8)$$

In particular, for $t > 0$,

$$\begin{aligned} P^{(\nu)}(t, \dots, t) &= \nu! \left(\sum_{\pi > \nu} \binom{\pi}{\nu} a_\pi t^{|\pi-\nu|} + a_\nu \right) \\ &\leq \nu! \left(\sum_{\pi > \nu, a_\pi > 0} \binom{\pi}{\nu} a_\pi t^{|\pi-\nu|} + a_\nu \right) \\ &\leq |a_\nu| \nu! \left[\sum_{\pi > \nu, a_\pi > 0} \binom{\pi}{\nu} \left(\frac{t}{B_\Delta \Gamma_\Delta} \right)^{|\pi-\nu|} - 1 \right], \end{aligned}$$

because from the definition of B_Δ we know that $a_\pi \leq |a_\nu|/(\delta_{\pi,\nu} B_\Delta^{|\pi-\nu|})$ and from the definition of Γ_Δ that $\delta_{\pi,\nu} \geq \Gamma_\Delta^{|\pi-\nu|}$. The positive root α of the polynomial

$$Q(u) := \sum_{\pi > \nu, a_\pi > 0} \binom{\pi}{\nu} u^{|\pi-\nu|} - 1$$

satisfies $\alpha < A_P/(B_\Delta \Gamma_\Delta)$, because the polynomial $P^{(\nu)}(t, \dots, t)$ is non-positive for $t = \alpha B_\Delta \Gamma_\Delta$. However, we can rewrite $Q(u)$ as

$$\begin{aligned} Q(u) &= \sum_{\pi > \nu, a_\pi > 0} \binom{\pi}{\nu} u^{|\pi-\nu|} - 1 \\ &= \sum_{i=1}^D \left(\sum_{\pi > \nu, |\pi-\nu|=i} \binom{\pi}{\nu} \right) u^i - 1 \\ &\leq \sum_{i=1}^D \frac{1}{i!} (Du)^i - 1, \end{aligned}$$

where the last step follows from the claim that $\sum_{\pi:|\pi-\nu|=i} \binom{\pi}{\nu} \leq D^i/i!$ (which we prove later). Thus,

$$Q(u) \leq \sum_{i=1}^D \frac{1}{i!} (Du)^i - 1 < \sum_{i=0}^{\infty} \frac{1}{i!} (Du)^i - 2 = \exp(Du) - 2.$$

The RHS of this inequality vanishes at $\ln 2/D$. Thus, we have shown that $\ln 2/D < \alpha \leq A_P/(B_\Delta \Gamma_\Delta)$, which implies our desired inequality (7).

To complete the proof we only need to prove the claim that $\sum_{\pi:|\pi-\nu|=i} \binom{\pi}{\nu} \leq D^i/i!$. But this follows from the observation that

$$\begin{aligned} \sum_{\pi:|\pi-\nu|=i} \binom{\pi}{\nu} &= \sum_{\pi:|\pi-\nu|=i} \prod_{j=1}^n \binom{\pi_j}{\nu_j} \\ &= \sum_{\pi:|\pi-\nu|=i} \prod_{j=1}^n \frac{\pi_j(\pi_j-1)\cdots(\nu_j+1)}{(\pi_j-\nu_j)!} \\ &\leq \sum_{\pi:|\pi-\nu|=i} \prod_{j=1}^n \frac{\pi_j^{\pi_j-\nu_j}}{(\pi_j-\nu_j)!} \\ &\leq \sum_{\pi:|\pi-\nu|=i} \prod_{j=1}^n \frac{d_j^{\pi_j-\nu_j}}{(\pi_j-\nu_j)!} \\ &= \frac{D^i}{i!}, \end{aligned}$$

where the last step follows from the multinomial theorem. \square

In particular, for Hong's bound B_H the quality statement yields $B_H \leq A_P D / (\Omega_n \ln 2)$,

since from the definition (5) of the weights it follows that $\Gamma_H = \Omega_n$; this quality estimate was first obtained in [8], where a potential improvement was also mentioned.

2.2. Possible Improvements

Given a general bound B_Δ , obtained from Theorem 2.2, on the threshold of absolute positivity, we may ask whether it is possible to improve upon the quality statement in Theorem 2.4. Here we give an approach that can improve the quality statement by bringing the factor of $\ln 2$ in the denominator of (7) arbitrarily close to one.

The bound B_Δ is obtained for a certain $\nu \in \nu(P)$, and is an upper bound on the threshold of absolute positiveness of the special polynomial $P_\nu(X) := \sum_{\pi > \nu, a_\pi > 0} \delta_{\pi, \nu} a_\pi X^\pi + a_\nu X^\nu$ where $\Delta := [\delta_{\pi, \nu}]$. But the threshold of absolute positiveness A_ν of $P_\nu(X)$ is just the positive real root of the univariate polynomial $P_\nu(t, \dots, t)$. This univariate polynomial belongs to a class of polynomials called Cauchy polynomials [12]. Based upon the special structure of these polynomials, we can improve upon the bound B_Δ .

In a more general sense, we can restate our goal: Given a Cauchy polynomial $f(x) = \sum_{\pi: \pi > \nu} a_\pi x^\pi - a_\nu$, where $a_\pi, a_\nu > 0$, and an $\epsilon \in (0, 1]$, get an upper bound B on its positive root, say t^* , such that $B \leq t^*(1 + \epsilon)$. A straightforward approach to obtain such a bound is to start a Newton iteration from any point $t > t^*$. Since both $f(x)$ and its derivative are positive for $x > t^*$ we know that the Newton iteration will converge to t^* ; there are approaches for improving the convergence, e.g. see [14], or [12]. The problem with this approach is that we cannot give a bound on the number of iterations required because we have no guarantee on the *initial* rate of convergence. Instead we use procedures such as the Dekker-Brent algorithm [6] or the Quadratic Interval Refinement (QIR) by John Abbott [1]. These methods take as input an interval containing a unique root of a continuous function and return an interval of desired width that contains the root. Moreover, these latter procedures guarantee at least linear convergence from the starting, because in the worst case the methods perform only bisections.

Given such procedures we can improve the bound B_Δ as follows: Given a polynomial $P(X)$, for each $\nu \in \nu(P)$, we construct the polynomial $P_\nu(X)$ and the bound B_ν defined as in (2); then for each $\nu \in \nu(P)$ we call one of the refining procedures (say QIR) on the univariate polynomial $t^{-|\nu|} P_\nu(t, \dots, t)$ and the interval $(B_\nu \Omega_n, B_\nu)$; we stop the iteration when we reach an interval whose width is smaller than ϵ times the left endpoint; for each $\nu \in \nu(P)$, let B'_ν be the right endpoint of the terminating interval; we output $\max_{\nu \in \nu(P)} B'_\nu$. The terminating criterion ensures that $B'_\nu \leq A_\nu(1 + \epsilon)$. We need to consider all $\nu \in \nu(P)$, instead of a single ν' for which $B_{\nu'} = B_\Delta$, because even though $B_\Delta = \max_{\nu \in \nu(P)} B_\nu$, we cannot be certain that $A_{\nu'} = \max_{\nu \in \nu(P)} A_\nu$; for instance, there might be two indices ν, ν' such that $B_\nu = B_{\nu'}$ but $A_\nu > A_{\nu'}$. What is the quality result for the improved bound?

Suppose the bound is obtained for some index ν . Then we know that

$$P_\nu(X) = \sum_{\pi > \nu, a_\pi > 0} \delta_{\pi, \nu} a_\pi X^\pi + a_\nu X^\nu$$

and the ν th derivative of $P(t, \dots, t)$ is

$$\begin{aligned}
P^{(\nu)}(t, \dots, t) &\leq \nu! \left(\sum_{\pi > \nu, a_\pi > 0} \binom{\pi}{\nu} a_\pi t^{|\pi - \nu|} + a_\nu \right) \\
&\leq \nu! \left(\sum_{\pi > \nu, a_\pi > 0} \binom{\pi}{\nu} \delta_{\pi, \nu} a_\pi \left(\frac{t}{\Gamma_\Delta} \right)^{|\pi - \nu|} + a_\nu \right) \\
&\leq \nu! \left(\sum_{\pi > \nu, a_\pi > 0} \delta_{\pi, \nu} a_\pi \left(\frac{Dt}{\Gamma_\Delta} \right)^{|\pi - \nu|} + a_\nu \right)
\end{aligned}$$

because

$$\binom{\pi}{\nu} = \prod_{i=1}^n \binom{\pi_i}{\nu_i} \leq \prod_{i=1}^n \frac{\pi_i!}{\nu_i!} \leq \prod_{i=1}^n \pi_i^{\pi_i - \nu_i} \leq \prod_{i=1}^n d_i^{\pi_i - \nu_i} < (d_1 + \dots + d_n)^{|\pi - \nu|} = D^{|\pi - \nu|}.$$

Thus $P^{(\nu)}(t) \leq \nu! t^{-|\nu|} P_\nu(Dt/\Gamma_\Delta)$, which implies that $A_\nu \leq DA_P/\Gamma_\Delta$, and hence $B'_\nu \leq DA_P(1 + \epsilon)/\Gamma_\Delta$, where ϵ can be made as small as desired. Thus we can improve upon the quality bound in (7) by a constant multiplicative factor.

Now we address the algebraic cost entailed by this improvement as compared to the algebraic cost of computing B_Δ alone. It is clear that the additional cost accrued is in calling QIR for obtaining the improvement. The univariate polynomial $P_\nu(t, \dots, t)$ can have at most D coefficients. Thus each iteration of QIR entails $O(D)$ algebraic operations. The number of iterations, for a given ν , is in the worst case (i.e. when we only do bisections) bounded by $\log(B_\nu(1 - \Omega_n)/(B_\nu\Omega_n\epsilon))$, because the width of the initial interval is $B_\nu(1 - \Omega_n)$ and the width of the terminating interval is at least $B_\nu\Omega_n\epsilon$; from the observation that $1/\Omega_n = \Theta(n)$ it follows that the number of iterations is bounded by $O(\log(n/\epsilon))$. Thus the additional cost is bounded by $O(ND \log(n/\epsilon))$, where N is the number of terms in $P(X)$. The cost of computing B_Δ on the other hand is $O(N^2)$. In the worst case $N = (d_1 + 1) \dots (d_n + 1)$, and hence asymptotically speaking the cost of computing B_Δ dominates the cost of the proposed improvement.

The above approach applies to all bounds within the framework proposed in Theorem 2.2. However, for the important case of Hong's bound B_H we give a special procedure that *potentially* improves upon the bound, but in the worst case might be same as B_H . The procedure is a generalization of the procedure proposed in [4] for the univariate case.

The key idea behind Hong's bound is that we can write the monomial $a_\pi x^\pi$, $\pi \in \pi(P)$, as $\sum_{k > 0} \Omega_1^k a_\pi x^\pi + (1 - \sum_{k > 0} \Omega_1^k) a_\pi x^\pi$, and *associate* with $\nu \in \nu(P)$ the term $\Omega_1^{\pi - \nu} a_\pi x^\pi$, i.e., choose $\delta_{\pi, \nu} := \Omega_1^{\pi - \nu}$. In [4] the authors observed the following: suppose we consider the negative coefficients in decreasing order of degree; now if there is some $k < \pi - \nu$ such that $\Omega_1^k a_\pi x^\pi$ has not been associated with any index $\nu' \in \nu(P)$, $\nu' > \nu$, then we can associate $\Omega_1^k a_\pi x^\pi$ with ν , i.e., choose $\delta_{\pi, \nu} := \Omega_1^k > \Omega_1^{\pi - \nu}$; since we have increased $\delta_{\pi, \nu}$ we may have improved upon B_H , but in the worst case there is no such k , and $\delta_{\pi, \nu}$ remains unchanged as does B_H . There is a minor obstacle in generalizing this approach: in the univariate case, for a given $k \in \mathbb{N}_{>0}$ there is exactly one monomial of the form $\Omega_1^k a_\pi x^\pi$ and hence at most one index $\nu \in \nu(P)$ with which it can be associated, but in the multivariate case this is not the situation; this is because the monomial $a_\pi X^\pi$ is now written as $\sum_{\iota > 0} \Omega_n^{|\iota|} a_\pi X^\pi + (1 - \sum_{\iota > 0} \Omega_n^{|\iota|}) a_\pi X^\pi$, and clearly, for a given $k \in \mathbb{N}_{>0}$, there is more than one index ι such that $|\iota| = k$, and hence possibly more than one

index ν with which $\Omega_n^{|\nu|} a_\pi X^\pi$ can be associated. If $C(n, k)$ is the number of ways of writing k as the sum of n non-negative integers, where ordering of the integers matters, then $\sum_{\iota > 0} \Omega_n^{|\iota|} = \sum_{k > 0} C(n, k) \Omega_n^k$, and hence $\Omega_n^k a_\pi X^\pi$ can be associated with at most $C(n, k)$ indices $\nu \in \nu(P)$, $\nu < \pi$. Thus the improvement given below has to keep track of these indices to avoid associating more than we can afford.

INPUT: A polynomial $P(X) = \sum_{\iota \in I} a_\iota X^\iota$, $I \subset \mathbb{N}^n$, satisfying (*).
 OUTPUT: An upper bound on the threshold of absolute positiveness of $P(X)$.

1. Let $B := 0$, and $D := d_1 + \dots + d_n$, where $d_i := \deg_{x_i}(P)$.
3. Construct an array L of D linked lists.
 - ◁ The list $L[i]$ will contain all monomials $a_\iota X^\iota$ such that $|\iota| = i$.
4. Traverse the polynomial $P(X)$ and insert the monomial $a_\iota X^\iota$ in the list $L[|\iota|]$.
4. If D' is the largest index for which $L[D']$ is not empty then
 - For $k = 1, \dots, D'$, compute $C(n, k)$.
5. For every $\pi \in \pi(P)$ do the following:
 - 5.a. Assign an array C_π of size $|\pi|$ initialized to zero, and a counter $k_\pi := 1$.
 - ◁ $C_\pi[k_\pi]$, $0 \leq k_\pi \leq |\pi|$, will be the number of indices $\nu \in \nu(P)$
 - ◁ such that $\nu < \pi$ and $\delta_{\pi, \nu} = \Omega_n^{k_\pi}$.
 6. For $i = D, \dots, 0$ do:
 - 6.a. Let $\nu \in \nu(P)$ be an entry in $L[i]$; $B' := \infty$.
 - For $j = D, \dots, i$ do:
 - Let $a_\pi X^\pi$ be an entry in $L[j]$ such that $\pi \in \pi(P)$ and $\pi > \nu$.
 - If $C_\pi[k_\pi] \geq C(n, k)$ then $k_\pi := k_\pi + 1$.
 - Increment $C_\pi[k_\pi]$ by one.
 - Define $k_{\pi, \nu} := k_\pi$.
 - $B' := \min(B', (|a_\nu| / (\Omega_n^{k_{\pi, \nu}} a_\pi))^{1/|\pi - \nu|})$.
 - $B := \max(B, B')$.
7. Output B .

Given two indices π, ν , such that $\pi \in \pi(P)$, $\nu \in \nu(P)$ and $\pi > \nu$, we claim that $k_{\pi, \nu} \leq |\pi - \nu|$. The proof is by induction. Initially $k_{\pi, \nu} = 1$ and hence is obviously no larger than $|\pi - \nu|$. Let $\nu' \in \nu(P)$ be the index considered just before ν in step 6.a., thus $|\nu'| \geq |\nu|$, and suppose that $\pi > \nu'$; if the supposition is false then $k_{\pi, \nu} = 1$ and the claim follows immediately. Suppose the claim holds for π, ν' , i.e., $k_{\pi, \nu'} \leq |\pi - \nu'|$. Then since ν immediately follows ν' we know that $k_{\pi, \nu} \leq k_{\pi, \nu'} + 1 \leq |\pi - \nu'| + 1 \leq |\pi - \nu|$. Thus, the following inequality holds for weights

$$\delta_{\pi, \nu} := \Omega_n^{k_{\pi, \nu}} \geq h_{\pi, \nu} = \Omega_n^{|\pi - \nu|}.$$

Moreover, the new weights satisfy the constraints in Theorem 2.2, because $\sum_{\iota > 0} \Omega_n^{|\iota|} < 1$.

From the complexity perspective, the extra overhead introduced in the above procedure is in computing $C(n, k)$, for $0 < k \leq D$. Let $C'(n, k)$ be the number of ways of writing k as a sum of *exactly* n non-zero numbers, or in terms of the theory of partitions, the compositions of k into exactly n parts [5, Chap. 4]; if $k \geq n$ then $C'(n, k) = \binom{k-1}{n-1}$, otherwise it is zero. Then the number of ways of writing k into n parts where exactly i

of the parts are zeros is $\binom{n}{i}C'(n-i, k)$. Summing this for $i = 0, \dots, n-1$ we obtain,

$$C(n, k) = \sum_{i=0}^{n-1} C'(n-i, k) \binom{n}{i}.$$

Thus the algebraic complexity of computing $C(n, k)$, for $k = 1, \dots, D$, is $O(Dn)$. In the worst case, this is dominated by the cost of computing the actual bound in step 6 of the procedure.

3. Computational Optimality of Bounds

As Theorem 2.3 suggests, in some sense the bound by Hong is close to being optimal in terms of quality. But is it an optimal bound in terms of complexity? Is it possible that there is a bound B_Δ requiring fewer operations than B_H and yet giving us a similar optimality result? In this section we show that this does not even hold true for the special case of univariate polynomials.

3.1. Number of Weights

Is it possible to use less weights than used in the bound B_H to obtain a reasonable bound on absolute positiveness? Precisely, suppose a bound uses only the sign of a coefficient $a_\pi \in \mathbb{R}$, and the relative position to other coefficients (of the same or *opposite* sign). Can any set of $\delta_{\pi, \nu}$'s smaller than the number of non-zero entries in $H = h_{\pi, \nu}$ (defined in (5)) yield a reasonable bound on the set of all polynomials?

Suppose we do not use all of the coefficient ratios $|a_\nu|/a_\pi$, $\nu \in \nu(P)$, i.e. we assign for one combination of indices with $h_{\hat{\pi}, \hat{\nu}} \neq 0$ (implying $\hat{\pi} > \hat{\nu}$) the weight $\delta_{\hat{\pi}, \hat{\nu}} = 0$.

Let us consider then a polynomial with the following sign distribution:

$$\underbrace{+, +, \dots, +, +}_{P_2}, \underbrace{-, -, \dots, -, -}_{N_2}, \underbrace{+, +, \dots, +, +}_{P_1}, \underbrace{-, -, \dots, -, -}_{N_1}, \quad (9)$$

i.e. a polynomial that has $4d$ non-vanishing coefficients in four blocks P_2, N_2, P_1, N_1 each of size d such that all indices in P_2 are greater in order than those in N_2 and N_1 , while all indices in P_1 are greater than those in N_1 .

If the left-out combination $(\hat{\pi}, \hat{\nu})$, $\hat{\pi} > \hat{\nu}$, is from index blocks P_2, N_2 then we consider a family of polynomials with the following coefficients:

$$a_{\hat{\pi}} := 1, a_i := t^3 \forall i \in P_2 \setminus \{\hat{\pi}\}, a_{\hat{\nu}} := 1, a_j := t^4 \forall j \in N_2 \setminus \{\hat{\nu}\}, \quad (10)$$

$$a_k := 1, \forall k \in P_1, a_l := t^4 \forall l \in N_1, \quad (11)$$

From these coefficients it follows that $B_H = 1/\Omega_n$.

A bound B_Δ not considering $(\hat{\pi}, \hat{\nu})$ but assigning positive weight for some other pair $\delta_{i, \hat{\nu}}$ will yield

$$\max \left\{ \min_{i \in P_2 \setminus \{\hat{\pi}\}} \left\{ i^{-\hat{\nu}} \sqrt{1/(t^3 \delta_{i, \hat{\nu}})} \right\}; \min_{i \in P_2 \setminus \{\hat{\pi}\}} \left\{ i^{-j} \sqrt{t^4/(t^3 \delta_{i, j})} \right\}, \min_{i \in P_1, j \in N_1} \left\{ i^{-j} \sqrt{t^4/\delta_{i, j}} \right\} \right\}. \quad (12)$$

As $t \rightarrow 0$ the bound B_Δ tends to infinity while B_H is $1/\Omega_n$.

Similar examples can be constructed when $\hat{\pi} \in P_2$ and $\hat{\nu} \in N_1$, or when $\hat{\pi} \in P_1$ and $\hat{\nu} \in N_1$.

Thus, a bound not using all possible coefficient ratios will be arbitrarily bad for an infinite number of polynomials.

3.2. Bounds with comparisons

The above example works under the assumption that the bound considers only the sign and the relative positions of the coefficients while constructing the coefficient weight matrix. The next question we ask is can a bound do better than Hong's bound if in addition to the options mentioned above we are allowed to do $O(d)$ comparisons among the coefficients while constructing the coefficient weight matrix. We show that a concrete instance of such a bound, namely the "local max" bound by Akritas et al. [3], can be much worse than Hong's bound for an infinite family of polynomials.

Before we proceed, let us recall the "local max" bound of Akritas et al.

INPUT: A univariate polynomial $P(x) = \sum_{i=0}^d a_i x^i$, $a_i \in \mathbb{R}$, $a_d > 0$.
 OUTPUT: An upper bound on the threshold of absolute positiveness of $P(X)$.

1. Let $S[1, \dots, d]$ be an array initially assigned to zero; $B := 0$.
2. For every $\nu \in \nu(P)$ do the following:
 - Let a_π , $\pi \in \pi(P)$, be the largest coefficient such that $\pi > \nu$.
 - Increment $S[\pi]$ by one. $\triangleleft 1 \leq S[\pi] \leq |\pi - \nu|$
 - $B := \max(B, (|a_\nu| / (\Omega_1^{S[\pi]} a_\pi))^{1/|\pi - \nu|})$.
3. Output B .

Now we construct a family of polynomials for which this bound can get much worse compared to Hong's bound. Let us consider the polynomial with the sign distribution shown in (9). Since we are allowed only $O(d)$ comparisons, we know that there must be a pair $(\hat{\pi}, \hat{\nu})$ of indices such that $\hat{\pi} \in [3d/2, d+1]$, $\hat{\nu} \in [d/2, 0]$, and the two are not compared. Let us assign the values to the coefficients as follows: $a_{4d} := 1/t$, $a_{\hat{\pi}} := 1/\sqrt{t}$, $a_{\hat{\nu}} := 1$, all other positive coefficients are t , and all the remaining negative coefficients are t^3 ; it is clear that for t sufficiently small the largest positive coefficient is a_{4d} . Then the value computed by the "local max" bound is the maximum between

$$\left(\frac{t}{\delta_{4d, \hat{\nu}}}\right)^{1/(4d - \hat{\nu})} \quad \text{and} \quad \max_{\nu \in \nu(P) \setminus \{\hat{\nu}\}} \left(\frac{t^4}{\delta_{4d, \nu}}\right)^{1/(4d - \nu)},$$

where $\delta_{4d, \hat{\nu}}$ and $\delta_{4d, \nu}$'s are some powers of $\Omega_1 = 1/2$, and since they are smaller than 1 we know that the "local max" bound is at least the maximum between

$$t^{1/(4d - \hat{\nu})} \quad \text{and} \quad \max_{\nu \in \nu(P) \setminus \{\hat{\nu}\}} t^{4/(4d - \nu)}.$$

For $t < 1$ the first term is the largest, because for all values of ν , $16d - 4\hat{\nu} > 4d - \nu$ as $\hat{\nu} \in [d/2, 0]$.

The bound B_H , on the other hand, is

$$2 \max \begin{cases} \min\{t^{1/4d - \hat{\nu}}, \sqrt{t}^{1/\hat{\pi} - \hat{\nu}}, t^{-1/\pi - \hat{\nu}}\} \\ \min\{t^{4/4d - \nu}, t^{2/\pi - \nu}\} \text{ for } \pi \in (4d, 3d) \text{ and } \nu \in [3d, 2d) \\ \min\{t^{4/4d - \nu}, t^{4.5/\hat{\pi} - \nu}, t^{2/\pi - \nu}\} \text{ for } \pi \in [2d, d), \pi \neq \hat{\pi} \text{ and } \nu \in [d, 0], \nu \neq \hat{\nu}. \end{cases}$$

We claim that $B_H = 2\sqrt{t}^{1/\hat{\pi} - \hat{\nu}}$ for $t < 1$:

- (1) $t^{1/4d - \hat{\nu}} > \sqrt{t}^{1/\hat{\pi} - \hat{\nu}}$, because $2\hat{\pi} < 4d + \hat{\nu}$ and $\hat{\pi} \leq 3d/2$.
- (2) $\sqrt{t}^{1/\hat{\pi} - \hat{\nu}} > t^{4/4d - \nu}$, $\nu \in [3d, 2d)$, because $8(\hat{\pi} - \hat{\nu}) > 4d > 4d - \nu$; $\sqrt{t}^{1/\hat{\pi} - \hat{\nu}} > t^{2/\pi - \nu}$, $\pi \in (4d, 3d)$, because $4(\hat{\pi} - \hat{\nu}) > 2d > \pi - \nu$.

- (3) The minimum in the third “minimum term” is either $t^{4.5/\hat{\pi}-\nu}$ or $t^{2/\pi-\nu}$, $\pi \in [2d, d)$ and $\nu \in [d, 0]$, because

$$4d - \nu > 2(\pi - \nu) \text{ and } 18d - 0.5\nu > 4\hat{\pi}.$$

Now $\sqrt{t}^{1/\hat{\pi}-\hat{\nu}} > t^{4.5/\hat{\pi}-\nu}$, because $9(\hat{\pi} - \hat{\nu}) > 4d > \hat{\pi} - \nu$; and $\sqrt{t}^{1/\hat{\pi}-\hat{\nu}} > t^{2/\pi-\nu}$, because $4(\hat{\pi} - \hat{\nu}) > 2d > \pi - \nu$.

Thus the ratio of the “local max” bound and the bound B_H is at least

$$\frac{t^{1/(4d-\hat{\nu})}}{2t^{1/2(\hat{\pi}-\hat{\nu})}} = \frac{1}{2}t^{(2\hat{\pi}-4d-\hat{\nu})/2(4d-\hat{\nu})(\hat{\pi}-\hat{\nu})} = \left(\frac{1}{2t}\right)^{(4d+\hat{\nu}-2\hat{\pi})/2(4d-\hat{\nu})(\hat{\pi}-\hat{\nu})}$$

which tends to infinity as t approaches zero because $4d + \hat{\nu} > 2\hat{\pi}$ as $\hat{\pi} \in [3d/2, d + 1]$.

4. Conclusion

We presented a general scheme for obtaining bounds on absolute positiveness of a multivariate polynomial $P(X)$. The choice of the bound is governed by the sign of the coefficients and a special coefficient weight matrix with certain constraints on its entries. The bound by Hong is obtained (in the scheme) for a special choice of the coefficient weight matrix, and it was shown to be at most a multiplicative factor of $1/\Omega_n$ from the best possible bound within this framework. Moreover, we derived an optimality result giving the relation between any bound in this framework and the threshold of absolute positiveness of $P(X)$. An additional iterative approach was suggested to improve upon the quality of any bound within the framework. For the special case of Hong’s bound, we suggested an improvement, which generalizes a similar improvement for the case of univariate polynomials by [4], without accruing any substantial overhead.

Though we have seen possible ways of obtaining bounds on absolute positiveness of $P(X)$, the question remains of getting similar bounds on its positiveness. The utility of such bounds cannot be overstated, as most of the known bounds on absolute positiveness are in practice used as a substitute for bounds on positiveness. Thus, the search of similar bounds on positivity remains an open problem.

Acknowledgements

The first author is grateful to the Algorithms and Complexity group (AG1) at the Max Planck Institute for Informatics, Saarbrücken, and to its director Kurt Mehlhorn for their warm hospitality during his visits. Both the authors are indebted to the AG1 group for an atmosphere that fostered the research presented here.

References

- [1] Abbott, J., 2006. Quadratic Interval Refinement, presented as a poster in ISSAC’06. Available at <http://www.dima.unige.it/~abbott>.
- [2] Akritas, A. G., 1978. Vincent’s theorem in algebraic manipulation. Ph.D. thesis, Operations Research Program, North Carolina State University, Raleigh, North Carolina.

- [3] Akritas, A. G., Strzeboński, A., Vigklas, P., 2006. Implementations of a New Theorem for Computing Bounds for Positive Roots of Polynomials. *Computing* 78, 355–367.
- [4] Akritas, A. G., Strzeboński, A., Vigklas, P., 2008. Improving the performance of the continued fractions method using new bounds of positive roots. *Nonlinear Analysis: Modelling and Control* 13 (3), 265–279.
- [5] Andrews, G. E., 1984. *The Theory of Partitions*. Cambridge University Press.
- [6] Brent, R. P., 1973. *Algorithms for minimization without derivatives*. Prentice Hall, Englewood Cliffs, NJ.
- [7] Fujiwara, M., 1916. Über die obere Schranke des absoluten Betrages der Wurzeln einer algebraischen Gleichung. *Tôhoku Math. J.* 10, 167–171.
- [8] Hong, H., 1998. Bounds for absolute positiveness of multivariate polynomials. *J. of Symb. Comp.* 25 (5), 571–585.
- [9] Hong, H., Jakus, D., 1998. Testing Positiveness of Polynomials. *J. of Automated Reasoning* 21, 23–38.
- [10] Johnson, J. R., 1991. *Algorithms for polynomial real root isolation*. Ph.D. thesis, Ohio State University, Columbus, OHIO, USA.
- [11] Kioustelidis, J., 1986. Bounds for the positive roots of the polynomials. *Journal of Computational and Applied Mathematics* 16, 241–244.
- [12] Schmeisser, G., 2007. *Theory of Polynomials and Applications, Recent Advances (dedicated to the memory of M.A. Malik)*. Deep & Deep Publications, New Delhi, Ch. Cauchy Polynomials, pp. 149–164, preprint available at <http://www.mi.uni-erlangen.de/~schmeis/publiste/cauchy02.pdf>.
- [13] Ștefănescu, D., 2005. New bounds for the positive roots of polynomials. *Journal of Universal Computer Science* 11 (12), 2132–2141.
- [14] Stoer, J., R., B., 1992. *Introduction to Numerical Analysis*. Springer-Verlag, New York, 2nd.
- [15] van der Sluis, A., 1970. Upper bounds for roots of polynomials. *Numer. Math.* 15, 250–262.
- [16] Yap, C. K., 2000. *Fundamental Problems of Algorithmic Algebra*. Oxford University Press.