Tree Breadth of the Continued Fractions Root Finding Method

K. Kalorkoti

Abstract: The continued fractions algorithm for isolating the real roots of polynomials with certainty is one of the most efficient known and widely used. It can be viewed as exploring a tree created by a sequence of simple transformations. In this paper we produce new upper bounds for the breadth of the tree that are significantly smaller than the degree of the input polynomial. We also consider the expected breadth under a reasonable distribution and derive a bound, subject to a plausible assumption, that grows logarithmically with the degree and coefficient size.

Keywords: Polynomial roots, continued fractions, tree breadth.

AMS Subject Classification: 12Y05

1 Introduction

The continued fractions algorithm for isolating the (positive) real roots of polynomials is one of the most efficient known and widely used, excluding algorithms that are subject to numerical instability. Its ultimate source is a result of Vicent [21] which guarantees the termination of a sequence of transformations that enable us to explore the real roots in (0,1) and (1,∞). Alesina and Galuzzi [4] give a modern proof of the result as well as historical information. Rather than reiterate the extensive bibliography we will refer the reader to a few papers that cover it in great detail (in particular Krandick and Mehlhorn [14], Tsigaridas and Emiris [20]).

Krandick and Mehlhorn [14] analyse a version that uses homothetic transformations given by Collins and Akritas [6], obtaining new bounds on the number of recursive subdivisions. They also prove that the breadth of the recursion tree is bounded by the degree of the input square free polynomial.

By contrast, Tsigaridas and Emiris [20] give complexity and implementation results for the method without homothetic transformations (see below for further details on this method). Their analysis relies on a conjecture regarding the continued fractions expansions of non-quadratic algebraic irrationals (see p.161 of [20]). They also discuss an earlier analysis by Akritas [2,3].

The original continued fractions method (without homothetic transformations) can be seen as exploring a tree of transformations with vertices labeled by polynomials and edges by one of two transformations. In this paper we provide bounds for the tree breadth of individual polynomials as well as for the average tree breadth (subject to an assumption) under a reasonable distribution.

For individual polynomials, Lemma 2.4 provides a bound in terms of the number of certain types of roots, which immediately implies that the degree is an upper bound. Theorem 3.2 provides a bound in terms of the degree and the size of coefficients which is good for cases of sequences where the coefficients have sufficiently controlled growth, the bound is even better if the number of non-zero coefficients is bounded.

For the average case we consider drawing uniformly at random square free polynomials of degree n with integer coefficients in [−B,B] where B ≥ 1, we also consider primitive square free polynomials. Theorem 4.1 provides an upper bound that is logarithmic in n and B but this is subject to an assumption on the distribution of roots in the unit disc (discussed at the start of §4.1, see also §4.2).

2 Definitions

Throughout we consider non-zero polynomials in z with coefficients from $\mathbb{R}$ (in practice the coefficients are from $\mathbb{Q}$ or, equivalently for root finding, from $\mathbb{Z}$ and we will assume this in some places). If

$$f = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$$

then we define $\text{vc}(f)$, the variation of coefficients, to be the number of sign changes in the sequence $a_n, a_{n-1}, \ldots, a_0$ of coefficients (ignoring as usual any occurrences of 0). The continued fractions...
method of finding the non-negative real roots of a square free $f$ is based on the transformations

$$f^T = f(z + 1) = a_n(z + 1)^n + a_{n-1}(z + 1)^{n-1} + \cdots + a_0,$$

and

$$f^I = f(1/(z + 1))(z + 1)^n = a_n + a_{n-1}(z + 1) + \cdots + a_0(z + 1)^n,$$

where we have assumed that $n$ is the degree of $f$ (for the sake of completeness we can set $0^T = 0^I = 0$). For simplicity later on, if $f^T(0) = 0$ then we replace $f^T(z)$ by $f^T(z)/z$ and similarly for $f^I(z)$. This ensures that 0 is not a root of $f^T$ of of $f^I$ and so it has non-zero constant coefficient enabling us to express various bounds using this (the alternative is to use the trailing coefficient). For the same reasons we will assume that $a_0 \neq 0$.

In practice transformations of the first type are replaced by $z \mapsto z + c$ where $c$ is a good integer lower bound on the smallest positive root of $f$, see [2, 3]. Without this optimisation the algorithm is provably exponential, see [6]. For the purposes of this paper it makes no difference to the overall situation so we will stay with the definition as given (our analysis holds unchanged no matter which version is used). Vincent’s Theorem [21] shows that if $f$ is square free then for all sufficiently long sequences of transformations involving $T$ and $I$ we produce a polynomial with variation of coefficients either 0 or 1. This is false if the polynomial is not square free, for example if $f = (2z^2 - 1)^2$ then $f^{ITT} = f^I = z^4 + 4z^3 + 2z^2 - 4z + 1$. In connection with non-square free polynomials see the result cited by [20] as Theorem 5 of that paper. A polynomial $f$ is called terminal if and only if $vc(f) \leq 1$.

In order to avoid confusion we note here that we will employ two forms of notation for transformations. The exponent form is compact and makes for ease of readability, but it must be borne in mind that it conforms to the algebraic notation of writing the argument on the left with order of application being as shown. Thus $f^{IT}$ means that we apply $I$ first and then $T$. In some situations it is more convenient to use the standard function notation with the argument on the right so that the order of application is the reverse of the written one. Thus if we wish to represent $z^{IT}$ in standard notation it becomes $T(I(z))$, which can be abbreviated to $TI(z)$: this will be significant when considering sequences of transformations as single Moebius transforms.

Given a polynomial $f$ we associate with it a binary tree, denoted by $tree(f)$, of the possible sequences of transformations of $f$ as follows. If $f$ is terminal the tree is empty otherwise the root is labeled with $f$. At any vertex $v$ labeled with the polynomial $g$ if $g^T$ not terminal then there is a right child with the edge labeled $T$ and the vertex labeled with $g^T$. Similarly if $g^I$ is not terminal there is a left child with the edge labeled $I$ and the vertex labeled with $g^I$.

We define the depth of a vertex to be the number of edges from the unique path to it starting at the root. Thus the root has depth 0 and its children, if any, have depth 1. The breath at depth $d$ of the tree is the number of vertices of depth $d$. The breadth of a tree is 0 if it is empty otherwise it is the maximum breadth over all depths (if this is unbounded we take the breadth to be $\infty$, this does not happen in our case). We denote the breadth of $tree(f)$ by $br(f)$.

We define $R^+(f)$ to be the set of strictly positive real roots of $f$. We will say that $\alpha$ is complex to mean that $\alpha \in \mathbb{C} - \mathbb{R}$. Let $C(f)$ denote the set of complex roots $\alpha$ of $f$ such that $Re(\alpha) > 0$ and $Re(\alpha) - |Re(\alpha)| > |\alpha - |Re(\alpha)||^2$. We also define $c(f)$ to be 0 if all complex roots $\alpha$ of $f$ that do not belong to $C(f)$ have $Re(\alpha) \leq 0$ and to be 1 otherwise. Note that $vc((z - \alpha)(z - \overline{\alpha})) = 0$ if and only if $Re(\alpha) \leq 0$ since $(z - \alpha)(z - \overline{\alpha}) = z^2 - 2Re(\alpha)z + |\alpha|^2$. It follows that if $|R^+(f)| + c(f) = 0$ then $vc(f) = 0$.

**Lemma 2.1** For all polynomials $f$

1. $C(f)$ consists of all complex roots $\alpha$ of $f$ with $Re(\alpha) > 0$ that are in the open disc $|z - [Re(z)] - 1/2| < 1/2$.

2. If $\alpha \not\in C(f)$ is complex then $Re(\alpha) \leq |\alpha|^2$. 
PROOF. Let $\alpha$ be a root of $f$ and set $m = |\text{Re}(\alpha)|$ throughout. For the first part, we have $\alpha$ with Re($\alpha$) > 0 and $\alpha$ is in the given disc if and only if

$$(\alpha - m - 1/2)(\bar{\alpha} - m + 1/2) < 1/4 \iff (\alpha - m)(\bar{\alpha} - m) - (\alpha - m)/2 - (\bar{\alpha} - m)/2 + 1/4 < 1/4$$

$$\iff |\alpha - m|^2 - (\text{Re}(\alpha) - m) < 0$$

which completes the proof.

For the second part, if Re($\alpha$) ≤ 0 the claim is trivial. So assume that Re($\alpha$) > 0. Since $\alpha \not\in C(f)$ we have

$$\text{Re}(\alpha) - m \leq |\alpha - m|^2 = (\alpha - m)(\bar{\alpha} - m) = |\alpha|^2 - 2m\text{Re}(\alpha) + m^2$$

Thus Re($\alpha$) ≤ |$\alpha|^2 - m(2\text{Re}(\alpha) - m - 1)$. If $m = 0$ the claim follows immediately. Otherwise if $m > 0$ then $2\text{Re}(\alpha) - m - 1 \geq 2m - m - 1 = m - 1 \geq 0$ and hence Re($\alpha$) ≤ |$\alpha$|^2.

We note that the preceding lemma remains true for the set $\hat{C}(f)$ that is defined is the same way as $C(f)$ but without the condition that its members must be complex. This will be useful in [41].

Define $C_m(f)$ to be the set of points of $C(f)$ satisfying $|z - m - 1/2| < 1/2$ for $m = 0, 1, 2, \ldots$. Note that $C_0(f)$ is the same as the set $C$ of Definition 19 in [14]. It follows from Lemma 2.1 that $C(f) = \cup_{m=0}^\infty C_m(f)$. We define $\hat{C}_m(f)$ similarly. Clearly the sets $C_i(f)$ are disjoint as are the sets $\hat{C}_i(f)$. The next Lemma follows from Lemma 2.6 the proof given here is more direct.

Lemma 2.2 $C(f^T) = \cup_{m=0}^\infty T^{-1}(C_m(f))$ and $C(f^T) \subseteq I^{-1}(C_0(f))$. In particular $C_m(f^T) = T^{-1}(C_{m+1}(f))$. Furthermore a root $\alpha$ of $f$ cannot both be mapped by $T$ a root in $C(f^T)$ and mapped by $I$ to a root in $C(f^T)$.

PROOF. We have for a complex $\beta$ that

$$\beta \in C_m(f^T) \iff f^T(\beta) = 0 \& \text{Re}(\beta) > 0 \& |\beta - m - 1/2| < 1/2$$

$$\iff f(\alpha) = 0 \& \beta = \alpha - 1 \& \text{Re}(\alpha) > 1 \& |\alpha - 1 - m - 1/2| < 1/2$$

$$\iff \alpha \in C_{m+1}(f) \& \beta = T^{-1}(\alpha).$$

This proves the claims regarding $f^T$. For the claim regarding $f^t$, suppose $\beta \in C(f^t)$. Then $\beta = 1/\alpha - 1 = \bar{\alpha}/|\alpha|^2 - 1$ where $f(\alpha) = 0$. If $\alpha \not\in C_0(f)$ then one of three possibilities holds: (i) Re($\alpha$) ≤ 0, (ii) 0 < Re($\alpha$) < 1 and Re($\alpha$) ≤ |$\alpha$|^2 or (iii) Re($\alpha$) ≥ 1. If any of these conditions hold then Re($\beta$) ≤ 0 which contradicts the assumption that $\beta \in C(f^t)$.

The final claim follows from the description of $C(f^T)$ and $C(f^t)$ given by the first part. □

Note that we cannot strengthen the second containment to an equality, e.g., if $f = 8z^2 - 4z + 1$ then $C_0(f) = \{1/4 + i/4, 1/4 - i/4\}$. However $f^t = z^2 - 2z + 5$ and $C(f^t) = \emptyset$ since the roots are $1 \pm 2i$. Indeed it follows from Lemma 2.6 that $C_m(f^t)$ consists of all $T^{-1}(\alpha)$ where $\alpha$ is a root of $f$ that is in the open disc $|z - (2m + 3)/2(m + 1)(m + 2)| < 1/2(m + 1)(m + 2)$. Just as above, the preceding lemma holds for $\hat{C}$ and $\hat{C}_m$.

Lemma 2.3 For all polynomials $f$ we have

1. If $\alpha \in \mathbb{R}^+(f)$ but $T^{-1}(\alpha) \not\in \mathbb{R}^+(f^T)$ and $I^{-1}(\alpha) \not\in \mathbb{R}^+(f^t)$ then $\alpha = 1$. Moreover $T(\mathbb{R}^+(f^T)) \cap I(\mathbb{R}^+(f^t)) = \emptyset$
2. $|\mathbb{R}^+(f)| \geq |\mathbb{R}^+(f^T)| + |\mathbb{R}^+(f^t)|$
3. $|C(f)| \geq |C(f^T)| + |C(f^t)|$
4. $|C_0(f)| \geq |C(f^T)| + c(f^t)$ and $c(f) \geq c(f^T)$.
PROOF. For the first claim note that \( T^{-1}(1) = I^{-1}(1) = 0 \notin T^{-1}(\mathbb{R}^+(f^T)) \cup I^{-1}(\mathbb{R}^+(f^I)) \). If \( \alpha > 1 \) then \( T^{-1}(\alpha) = \alpha - 1 > 0 \) while \( I^{-1}(\alpha) = 1/\alpha - 1 < 0 \) and so \( T^{-1}(\alpha) \in \mathbb{R}^+(f^T) \) but \( I^{-1}(\alpha) \notin \mathbb{R}^+(f^I) \). If \( \alpha < 1 \) the same argument shows that \( T^{-1}(\alpha) \notin \mathbb{R}^+(f^T) \) but \( I^{-1}(\alpha) \in \mathbb{R}^+(f^I) \).

The second claim follows from the first.

For the third claim, suppose a complex root \( \alpha \notin \mathbb{C}(f) \) then under \( I \) it is mapped to \( 1/\alpha - 1 \equiv \pi/|\alpha|^2 - 1 \). From Lemma 2.1 we have \( \Re(\alpha) \leq |\alpha|^2 \). It follows that \( \Re(1/\alpha - 1) \leq 0 \) and so \( 1/\alpha - 1 \notin \mathbb{C}(f^T) \cup \mathbb{C}(f^I) \). Under \( T \) the root is mapped to \( 1 \) and again this cannot be in \( \mathbb{C}(f^T) \cup \mathbb{C}(f^I) \), e.g., by the second part of the preceding Lemma. It follows from Lemma 2.2 that if \( \alpha \in \mathbb{C}(f) \) then we cannot have both \( 1/\alpha - 1 \in \mathbb{C}(f^I) \) and \( \alpha - 1 \in \mathbb{C}(f^T) \). The claim now follows.

We now deal with the first part of the fourth claim. By Lemma 2.2 we have \( \mathbb{C}(f^I) \subseteq I^{-1}(\mathbb{C}_0(f)) \) and so if \( c(f^I) = 0 \) the claim follows immediately. Suppose now that \( c(f^I) = 1 \) so there is a complex root \( \beta \) of \( f^I \) with \( \Re(\beta) > 0 \) and \( \beta \notin \mathbb{C}(f^I) \). It follows that \( \beta = 1/\alpha - 1 \) for some complex root \( \alpha \) of \( f \). Thus \( \Re(\beta) = \Re(\alpha)/|\alpha|^2 - 1 \) and so \( \Re(\alpha) > |\alpha|^2 \). It follows that \( 0 < \Re(\alpha) < 1 \) and hence \( \alpha \in \mathbb{C}_0(f) \). Since \( \mathbb{C}(f^I) \subseteq I^{-1}(\mathbb{C}_0(f)) \) and \( \beta \notin \mathbb{C}(f^I) \) the containment is strict. The claim now follows. The second part of fourth claim follows immediately from Lemma 2.2.

**Lemma 2.4** \( \text{br}(f) \leq |\mathbb{R}^+(f)| + |\mathbb{C}(f)| + c(f) \) and hence \( \text{br}(f) \leq \text{deg}(f) \).

**Proof.** If \( \text{tree}(f) \) is empty the claim is trivial. We now assume that \( \text{tree}(f) \) is not empty and use induction on the depth \( d \). If \( d = 0 \) then \( |\mathbb{R}^+(f)| + |\mathbb{C}(f)| + c(f) \geq 1 \) since \( f \) is not terminal and so \( \text{br}(f) = 1 \leq |\mathbb{R}^+(f)| + |\mathbb{C}(f)| + c(f) \). Assume now that \( d > 0 \). By induction \( \text{br}(f^T) \leq |\mathbb{R}^+(f^T)| + |\mathbb{C}(f^T)| + c(f^T) \) and \( \text{br}(f^I) \leq |\mathbb{R}^+(f^I)| + |\mathbb{C}(f^I)| + c(f^I) \leq |\mathbb{R}^+(f^I)| + \mathbb{C}_0(f) \) by the fourth part of Lemma 2.3. Now, using Lemma 2.2 in the third line below and Lemma 2.3 in the last line,

\[
\text{br}(f) \leq \text{br}(f^T) + \text{br}(f^I) \\
\leq |\mathbb{R}^+(f^T)| + |\mathbb{C}(f^T)| + c(f^T) + |\mathbb{R}^+(f^I)| + \mathbb{C}_0(f) \\
\leq |\mathbb{R}^+(f^T)| + |\mathbb{R}^+(f^I)| + \sum_{m=1}^{\infty} |\mathbb{C}_m(f)| + |\mathbb{C}_0(f)| + c(f) \\
\leq |\mathbb{R}^+(f)| + |\mathbb{C}(f)| + c(f),
\]

which establishes the main inequality. The consequence is immediate since \( f \) has at least \( |\mathbb{R}^+(f)| + |\mathbb{C}(f)| + c(f) \) distinct roots.

Krandick and Mehlhorn [13] prove that the breadth is bounded by the degree of the input square free polynomial for the variant method that also employs homothetic transformations (see their Theorem 29).

**2.1 Effect of Möbius Transforms**

For the reader’s convenience we collect together some simple results on the effect of Möbius transforms on certain discs.

**Lemma 2.5** Assume that \((ck - cr - a)(ck + cr - a) > 0\). Then the set of values satisfying \(|(az + b)/(cz + d) - k| < r\) is given by the open disc

\[
\left| z - \frac{(ad + bc)k - ab - cd(k^2 - r^2)}{(ck - cr - a)(ck + cr - a)} \right| < \frac{|ad - bc|r}{(ck - cr - a)(ck + cr - a)}
\]

In particular if \( k > r \geq 0 \) then under the transform \( z \mapsto 1/z \) the disc \(|z - k| < r\) goes to \(|z - k/(k^2 - r^2)| < r/(k^2 - r^2)\).
PROOF. The given disc is the same as \( |(az + b) - k(cz + d)| < r|cz + d| \). Setting \( z = u + iv \) and squaring both sides, the disc is given by \((a - kc)u + b - kd)^2 + (a - kc)^2 v^2 - r^2(cu + d)^2 - r^2 c^2 v^2 < 0\), since \((ck - cr - a)(ck + cr - a) > 0\) the derived inequality is equivalent to \((u - C)^2 + v^2 - R^2 < 0\) where

\[
C = \frac{(ad + bc)k - ab - cd(k^2 - r^2)}{(ck - cr - a)(ck + cr - a)} \quad \text{and} \quad R = \frac{|ad - bc|r}{(ck - cr - a)(ck + cr - a)},
\]

which is the claimed disc.

The rest follows by noting that \( a = 0, b = 1, c = 1, d = 0 \) so that the condition \((ck - cr - a)(ck + cr - a) > 0\) reduces to \((k - c)(k + c) > 0\) and since \( k > c \geq 0 \) it is satisfied. Substituting the values of \( a, b, c, d \) into the general derived disc we obtain claimed disc. \( \square \)

We note that if \( M \) consists of a sequence of \( T^{-1} \) and \( I^{-1} \) transformations then \( ad - bc = (-1)^s \) where \( s \) is the number of occurrences of \( I^{-1} \), cf. Theorem 8 of Collins and Krandick [5]. This can be shown by a straightforward induction. Thus, in this situation, the disc in the preceding lemma can be written as

\[
\left| z - \frac{(ad + bc)k - ab - cd(k^2 - r^2)}{(ck - cr - a)(ck + cr - a)} \right| < \frac{r}{(ck - cr - a)(ck + cr - a)}
\]

Lemma 2.6: Let \( M(z) \) be a Möbius transform composed of \( T \) and \( I \) and set \( M^{-1}(z) = (az + b)/(cz + d) \). Then \( C_m(f^M) \) consists of all \( M^{-1}(\alpha) \) where \( \alpha \) is a complex root of \( f \) that is in the open disk

\[
\left| z - \frac{(ad + bc)(m + 1/2) - ab - cd(m + 1)}{(cm - a)(c(m + 1) - a)} \right| < \frac{1}{2(cm - a)(c(m + 1) - a)}
\]

Proof. The roots \( \beta \) of \( f^M \) are precisely all \( \beta = M^{-1}(\alpha) \) where \( \alpha \) is a root of \( f \). Now a root \( \beta \) belongs to \( C_m(f^M) \) if and only \( |\beta - (m + 1/2)| < 1/2 \), i.e., \(|M^{-1}(\alpha) - (m + 1/2)| < 1/2 \). Since \( M \) is a composition of \( T \) and \( I \) we have \( M(z) = (Az + B)/(Cz + D) \) with \( A, B, C, D \geq 0 \) and not both \( C, D \) are 0 (similarly for \( A, B \)). We have \( M^{-1}(z) = (Dz - B)/(-Cz + A) \). Thus the inequality \((ck - cr - a)(ck + cr - a) > 0\) of Lemma 2.5 becomes \((-C(m + 1/2) + C/2 - D)(-C(m + 1/2) - C/2 - D) > 0\) which is equivalent to \((Cm + D)(C(m + 1) + D) > 0\). Since \( C, D \geq 0 \) and at least one is non-zero the inequality holds. The result now follows from Lemma 2.5. \( \square \)

2.2 Bounds on the number of roots

This section summarises some well known results for the reader’s convenience. The Mahler measure of a polynomial \( f = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0 \) with roots \( a_1, \ldots, a_n \) is \( M(f) = |a_n| \prod_{j=1}^{n} \max\{1, |a_j|\} \). As is well known, Jensen’s formula (see, e.g., Ahlfors [1]) yields the bound \( M(f) \leq \sum_{j=0}^{n} |a_j| \). Suppose \( \rho > 1 \) and set \( \epsilon(f, \rho) = \{ |a_j| \mid 1 \leq j \leq n \& |a_j| > \rho \} \). Then

\[
\epsilon(f, \rho) \leq \frac{1}{\log \rho} \left( \log \sum_{j=0}^{n} |a_j| - \log |a_n| \right).
\]

This follows from the simple observation that \( \rho^x(f, \rho) \leq |a_n|^{-1} M(f) \leq |a_n|^{-1} \sum_{j=0}^{n} |a_j| \). Now suppose that \( \rho < 1 \) and set \( \mu(f, \rho) = \{ |a_j| \mid 1 \leq j \leq n \& |a_j| < \rho \} \). It follows from (1) and the transformation \( z \to 1/z \) that

\[
\mu(f, \rho) \leq \frac{1}{\log (1/\rho)} \left( \log \sum_{j=0}^{n} |a_j| - \log |a_0| \right).
\]

Hughes and Nikeghbali [10] give the bounds

\[
\epsilon(f, 1/(1 - \rho)) \leq \frac{1}{\rho} \left( \log \sum_{i=0}^{n} |a_i| - \log |a_n| \right),
\]

\[
\mu(f, 1 - \rho) \leq \frac{1}{\rho} \left( \log \sum_{i=0}^{n} |a_i| - \log |a_0| \right),
\]
These are slightly weaker than the ones above since \( \log(1/(1 - \rho)) = \rho + \rho^2/2 + \rho^3/3 + \cdots \).

It will be helpful to set

\[
L(f) = \frac{1}{\sqrt{|a_0| |a_n|}} \sum_{j=0}^{n} |a_j|.
\]

**Lemma 2.7** Let \( f = a_n z^n + \cdots + a_0 \) where \( a_n a_0 \neq 0 \) and assume that \( \rho > 1 \). Then

\[
\epsilon(f, \rho) + \mu(f, 1/\rho) \leq \frac{2}{\log \rho} L(f).
\]

**Proof.** Straightforward application of (1) and (2). \( \square \)

**Lemma 2.8** Let \( S \) be a set of univariate polynomials of degree \( n \) with non-zero constant term and coefficients bounded in absolute value from above by \( B \). Consider drawing polynomials at random from \( S \) (using any probability distribution). Then

\[
\mathbb{E}[\epsilon(f, \rho)] \leq \frac{\log(n + 1)B}{\log \rho},
\]

when \( \rho > 1 \) and

\[
\mathbb{E}[\mu(f, \rho)] \leq \frac{\log(n + 1)B}{\log(1/\rho)},
\]

when \( \rho < 1 \).

**Proof.** The inequalities are an immediate consequence of (1) and (2). \( \square \)

**Lemma 2.9** Let \( f = a_n z^n + \cdots + a_0 \) where \( f \in \mathbb{Z}[z] \) and \( a_n a_0 \neq 0 \), then the number of positive integer roots of \( f \) (counted with multiplicity) is at most

\[
1 + \log |a_0| / \log 2.
\]

**Proof.** Suppose \( m_1, \ldots, m_r \) are integer roots of \( f \), necessarily non-zero. A simple argument based on Gauss’s Lemma for the content and primitive part of polynomials shows that

\[
m_1 \cdots m_r | a_0.
\]

Consider now the positive integer roots of \( f \) other than 1 and assume there are \( s \) of them. Since each root is an integer which is at least 2, it follows that \( s \leq \log |a_0| / \log 2 \). Taking the possibility that 1 is a root into account now yields the result. \( \square \)

## 3 Bound on the Tree Breadth for a Polynomial

We will use a famous result of Erdős and Turán [8] in the form given by Rahman and Schmeisser [17], Theorem 11.6.4: denote by \( n_f(\theta, \phi) \) the number of zeros in the sector \( \{ z \mid \theta \leq \arg z < \phi \} \), where \( 0 < \phi - \theta < 2\pi \). Then

\[
\left| n_f(\theta, \phi) - \frac{\phi - \theta}{2\pi} n \right| \leq C \sqrt{n \log L(f)},
\]

where \( C = \sqrt{2\pi/G} < 2.62 \) and \( G = \sum_{m=0}^{\infty} (-1)^{m-1} (2m + 1)^{-2} \) is Catalan’s constant. The original paper [8] had \( C = 16 \), the improved constant is due to Ganelius [9].

We are interested in the number of roots \( N_f(\phi) \) within a wedge defined by the angles \( -\phi, \phi \). Note that in this we include those roots \( \alpha \) with \( \arg(\alpha) = \pm \phi \). It follows that

\[
N_f(\phi) < \sqrt{\frac{2\pi}{G}} \sqrt{n \log L(f)} + \frac{\epsilon \phi}{\pi} n \tag{3}
\]

for all \( \epsilon > 1 \).

Suppose now that \( f \) has no more than \( k \) non-zero coefficients. Proposition 11.2.4 of [17] states: for \( 0 < \phi - \theta < 2\pi \), denote by \( n_f(\alpha, \beta) \) the number of zeros of \( f \) in the sector \( \{ z \mid \theta < \arg z < \phi \} \) then

\[
\left| n_f(\theta, \phi) - \frac{\phi - \theta}{2\pi} n \right| \leq k.
\]

It follows that

\[
N_f(\phi) < k + \frac{\epsilon \phi}{\pi} n \tag{4}
\]

for all \( \epsilon > 1 \).
Figure 1: The discs corresponding to $C_m(f^M)$ for $M(z) = II(z) = z^I$ (left of unit circle) and $M(z) = IT(z) = z^TI$ (right of unit circle) and $m = 0, 1, 2, 3, 4$. Also shown is the wedge enclosing the discs corresponding to $C_m(f^M)$ for $m \geq 2$. 


Theorem 3.1 Let $f$ be any polynomial of degree $n$ with real coefficients and non-zero constant term. For $m \geq 0$ define $\phi = \arctan\left(\frac{1}{2(m+2)\sqrt{(m+1)(m+3)}}\right)$. Then

1. the breadth of $\text{tree}(f)$ satisfies

$$br(f) \leq \frac{2}{\log\left(\frac{m+2}{m+1}\right)} \log L(f) + \sqrt{\frac{2\pi}{G}} \sqrt{n \log L(f)} + \frac{\phi}{\pi} n + 4.$$  

Suppose $a$ satisfies $0 < a < 4$ and $b \geq 0$. Choose $m_0$ such that $a(m+b)^4 < 4(m+1)(m+2)^2(m+3)$ for all $m \geq m_0$. If $\sqrt{a}(m_0+b)^2 \geq 2/\pi$, then

$$br(f) < \frac{1}{\sqrt{a} \log^2(2) \sqrt{a}} \log L(f) + \frac{\phi}{\pi} n + \sqrt{\frac{2\pi}{G}} \sqrt{n \log L(f)} - \frac{1}{\log 2} \left(\frac{b}{\log 2} - 2\right) \log L(f) + 4.$$  

2. Suppose that the number of non-zero coefficients of $f$ is no more than $k$. Then

$$br(f) \leq \frac{2}{\log\left(\frac{m+2}{m+1}\right)} \log L(f) + \frac{\epsilon \phi}{\pi} n + k + 4.$$  

Let $a$, $b$, $\phi$ and $m_0$ be as above. If $\sqrt{a}(m_0+b)^2 \geq 2/\pi$, then

$$br(f) < \frac{1}{\sqrt{a} \log^2(2) \sqrt{a}} \log L(f) + \frac{\epsilon \phi}{\pi} n + k - \frac{1}{\log 2} \left(\frac{b}{\log 2} - 2\right) \log L(f) + 4.$$  

Proof. We look first at item 1. Consider $\text{tree}(f)$. If it is not complete at depth 2 (i.e., one of $f^I$, $f^T$, $f^{II}$, $f^{IT}$, $f^{TT}$ is terminal) we may extend it artificially by adding the appropriate vertices and paths. It follows from Lemma 2.4 that $br(f)$ is bounded from above by the maximum of 2 and

$$br(f^{II}) + br(f^{IT}) + br(f^{TT}) \leq |C(f^{II})| + |C(f^{IT})| + |C(f^{TT})| + R^+(f^{II}) + R^+(f^{IT}) + R^+(f^{TT}) + c(f^{II}) + c(f^{IT}) + c(f^{TT}) \leq |C(f^{II})| + |C(f^{IT})| + |C(f^{TT})| + \left| C(f^{II}) \right| + \left| C(f^{IT}) \right| + \left| C(f^{TT}) \right| + R^+(f) + 4$$

where the final line is justified by the second part of Lemma 2.3. Thus the preceding expression acts as an upper bound for $br(f)$ in any case.

Using Lemmas 2.2 and 2.6 we have the following cases.

1. $C_m(f^{II})$ is a subset of the set of all $I^{-1}I^{-1}(\alpha)$ where $\alpha$ ranges over all roots $\alpha$ of $f$ in $C(f)$ that lie in the open disc

$$\left| z - \frac{2m^2 + 8m + 7}{2(m+2)(m+3)} \right| < \frac{1}{2(m+2)(m+3)}.$$  

2. $C_m(f^{IT})$ is a subset of the set of all $I^{-1}T^{-1}(\alpha)$ where $\alpha$ ranges over all roots $\alpha$ of $f$ in $C(f)$ that lie in the open disc

$$\left| z - \frac{2m + 5}{2(m+2)(m+3)} \right| < \frac{1}{2(m+2)(m+3)}.$$  

3. $C_m(f^{TT})$ is a subset of the set consists of all $T^{-1}I^{-1}(\alpha)$ where $\alpha$ ranges over all roots $\alpha$ of $f$ in $C(f)$ that lie in the open disc

$$\left| z - \frac{2m^2 + 8m + 7}{2(m+1)(m+2)} \right| < \frac{1}{2(m+1)(m+2)}.$$
4. \(C(f^{TT})\) consists of all \(T^{-1}T^{-1}(\alpha)\) where \(\alpha\) ranges over all roots \(\alpha\) of \(f\) in \(\cup_{m=2}^{\infty}C_m(f)\) that lie in any of the open discs

\[
\left| z - \frac{2m + 5}{2} \right| < \frac{1}{2}.
\]

Note that, by Lemma 2.5, cases 1 and 3 are dual to each other via the transformation \(z \mapsto 1/z\) and similarly for 2 and 4; however this observation does not lead to any advantage over the proof below. We can deal with the second and fourth cases easily. For the second case the disc for \(m\) is contained in a disc \(D(0,1/(m+2))\) and hence \(C(f^{TT})\) is contained in the disc \(D(0,1/2)\). For the fourth case, all the roots \(\alpha\) in question satisfy \(|\alpha| > 2\). Thus, by Lemma 2.7,

\[
|C(f^{TT})| + |C(f^{TT})| \leq (2/\log 2) \log L(f).
\]

For the other two cases we consider a wedge that encloses \(C_m(f^{II})\). Suppose the lines \(y = \pm \tan(\phi)x\), where \(\phi > 0\), are tangent to the boundary of \(C_m(f^{II})\) with centre \(c\) and radius \(r\), see Figure 1. It follows easily that

\[
\tan^2(\phi) = \frac{r^2}{c^2 - r^2} = \frac{1}{4(m+1)(m+2)^2(m+3)}.
\]

Clearly such a wedge will enclose \(C_s(f^{II})\) for all \(s \geq m\). Furthermore it does the same for \(C_s(f^{TT})\), since the ratio of the centre to the radius of the disc for \(C_m(f^{TT})\) is the same as that for \(C_m(f^{II})\) and \(\tan^2(\phi) = 1/(c^2/r^2 - 1)\).

The boundary of the disc corresponding to \(C_m(f^{II})\) crosses the \(x\)-axis at \(x = (m+1)/(m+2)\), nearest to the origin, while the boundary for the disc corresponding to \(C_m(f^{TT})\) crosses the \(x\)-axis at \(x = (m+2)/(m+1)\), furthest from the origin. Thus, \(|\cup_{m=0}^{m-1}C_m(f^{II})|\) is bounded from above by the number of roots of \(f\) in the interior of the disc with radius \(\sigma = (m+1)/(m+2)\). Similarly \(|\cup_{m=0}^{m-1}C_m(f^{TT})|\) is bounded from above by the number of roots of \(f\) in the exterior of the disc with radius \(1/\sigma\). By Lemma 2.7, the number of such roots is bounded from above by \(2 \log L(f)/\log (m+2)/(m+1)\).

Since the wedge includes \(R^+(f)\) it follows from (3) that for all \(\epsilon > 1\),

\[
|R^+(f)| + |C(f^{II})| + |C(f^{TT})| + |C(f^{TT})| < \frac{2}{\log (m+2)/(m+1)} \log L(f) + \sqrt{\frac{2\pi}{G}} \sqrt{n \log L(f)} + \frac{\epsilon \phi}{\pi} n.
\]

The first claim of item 1 now follows from (5) and the observation regarding \(br(f)\) at the start of this proof by taking \(\epsilon\) arbitrarily close to 1.

For the second claim, it is easily seen that \(\tan(\phi) > \phi\) for \(0 < \phi < \pi/2\) and so, by (5), we have \(4(m+1)(m+2)^2(m+3) < \phi^2\). For all large enough \(m\), say \(m \geq m_0 \geq 0\), we have \(a(m+b)^4 \leq 4(m+1)(m+2)^2(m+3)\). Thus it suffices to have \(m < 1/\sqrt{a}\sqrt{b} - b\) in order to ensure the condition \(4(m+1)(m+2)^2(m+3) < 1/\phi^2\). So for a given \(\phi > 0\) the corresponding value of \(m\) satisfies \(m < 1/\sqrt{a}\sqrt{b} - b\). In order to have \(m \geq m_0\) we need \(\phi < 1/\sqrt{a}(m_0 + b) \leq \pi/2\).

We claim that \(\log((m+2)/(m+1)) > 2\log^2 2/(m+2\log 2)\). To see this consider the function \(h(x) = \log((x+2)/(x+1)) - 2\log^2 2/(x+2\log 2)\) defined over the non-negative reals. The denominator of \(dh/dx\) is positive and the numerator is \((c_1 x + c_0)x\) where \(c_1 = 2\log^2 2 - 1 = -0.390\ldots\) and \(c_0 = 6\log^2 2 - 4\log 2 = 0.110\ldots\). Hence \(h\) is increasing for \(0 \leq x \leq -c_0/c_1 \approx 2.817\) and decreasing for \(x \geq -c_0/c_1\). Now \(h(0) = 0\) and then becomes positive while \(\lim_{x \to \infty} h(x) = 0\). Hence \(h(x) \geq 0\) for all \(x \geq 0\) and the claim is established. By the first part of item 1 we have

\[
br(f) \leq 2 \left( \frac{1}{\log 2} + \frac{m}{2 \log^2 2} \right) \log L(f) + \sqrt{\frac{2\pi}{G}} \sqrt{n \log L(f)} + \frac{\phi}{\pi} n + 4
\]

\[
< \frac{2}{\log 2} \log L(f) + \frac{1}{\log^2 2} \left( \frac{1}{\sqrt{a}\sqrt{b}} - b \right) \log L(f) + \sqrt{\frac{2\pi}{G}} \sqrt{n \log L(f)} + \frac{\phi}{\pi} n + 4
\]

\[
= \frac{1}{\sqrt{a}\log^2(2)\sqrt{\phi}} \log L(f) + \frac{\phi}{\pi} n + \sqrt{\frac{2\pi}{G}} \sqrt{n \log L(f)} - \frac{1}{\log 2} \left( \frac{b}{\log 2} - 2 \right) \log L(f) + 4
\]
The claims of item 2 follow by using \([4]\) instead of \([3]\) in the derivation above. \(\square\)

**Theorem 3.2** Let \(f\) be any polynomial of degree \(n\) with real coefficients and non-zero constant term. Suppose \(a\) satisfies \(0 < a < 4\) and \(b \geq 0\). Choose \(m_0\) such that \(a(m+b)^4 \leq 4(m+1)(m+2)^2(m+3)\) and set \(c = (b/\log 2 - 2)/\log 2\). If \(\sqrt{a}(m_0+b)^2 \geq 2/\pi\) and \(\log L(f) \leq 2n \log^2(2)/\pi \sqrt{a}(m_0+b)^3\) then

\[
br(f) < \frac{3}{\sqrt[3]{4\pi \sqrt{a} \log^4 2}} \sqrt[3]{n \log^2 L(f)} + \sqrt[3]{\frac{2\pi}{\sqrt{a}} \sqrt{n \log L(f)} - c \log L(f)} + 4
\]

\[
< \frac{2.11}{\sqrt{a}} \sqrt[3]{n \log^2 L(f)} + 2.62 \sqrt[3]{\sqrt{n \log L(f)} - c \log L(f)} + 4
\]

Suppose that the number of non-zero coefficients of \(f\) is no more than \(k\). Then

\[
br(f) < \frac{3}{\sqrt[3]{4\pi \sqrt{a} \log^4 2}} \sqrt[3]{n \log^2 L(f)} - c \log L(f) + k + 4
\]

\[
< \frac{2.11}{\sqrt{a}} \sqrt[3]{n \log^2 L(f)} - c \log L(f) + k + 4
\]

**Proof.** The function \(g(\phi) = u/\sqrt{a} + v\phi\) has its minimum at \(\phi = (u/2v)^{2/3}\). As seen in the proof of Theorem 3.1 the r.h.s. needs to be bounded from above by \(1/\sqrt{a}(m_0+b)^2\) and this holds provided that \(u < 2v/\sqrt{a}(m_0+b)^3\). Thus the minimum value of \(g(\phi)\) is \((2u^2v)^{1/3} + (u^2v/4)^{1/3} = 3(u^2v/4)^{1/3}\). Setting \(u = \log L(f)/\sqrt{a} \log 2\) and \(v = n/\pi\) the inequality is \(\log L(f) < 2n \log^2 2/\pi \sqrt{a}(m_0+b)^3\) and the value of \(g(\phi)\) is \(3(n \log^2 L(f))/4\pi \sqrt{a} \log 2\) \(1/3\). Substituting into the second inequality of item 1 of Theorem 3.1 yields the desired bound. The second bound follows by substituting into the second inequality of item 2 of Theorem 3.1. \(\square\)

The assumption that \(\log L(f) \leq 2n \log^2(2)/\pi \sqrt{a}(m_0+b)^3\) ensures that the bounds of the preceding lemmas are strictly positive; a simple argument shows that under the assumption \(c \log L(f) < 3 \sqrt[3]{n \log^2 L(f)}/\sqrt[3]{4\pi \sqrt{a} \log^4 2}\). If the condition on \(L(f)\) is satisfied then, since \(\sqrt{a}(m_0+b)^2 \geq 2/\pi\), we have \(\log L(f) \leq n \log^2(2)/(m_0+b)\). Thus \(L(f) \leq 2n \log^2(2)/(m_0+b)\). If \(f\) has coefficients bounded in absolute value by \(B\) (the situation of the next section) then \(\log L(f) \leq \log(n+1)B\) and the condition on \(L(f)\) is satisfied if \(B \leq 2n \log^2(2)/(m_0+b)/(n+1)\), i.e., for all polynomials of sufficiently high degree. As a simple illustration we may take \(a = 9/8\), \(b = 3\) and \(m_0 = 1\). The condition \(\sqrt{a}(m_0+b)^2 \geq 2/\pi\) is satisfied while the condition on \(L(f)\) is satisfied if \(\log L(f) \leq 0.05n\). The second bound of the preceding lemma becomes

\[
L(f) < 2.07 \sqrt[3]{n \log^2 L(f)} - 3.35 \log L(f) + k + 4.
\]

**4 Bound on the expected breadth of the transformation tree**

Let \(B > 0\) be an integer and let \(P(n, B)\) be the set of non-zero polynomials from \(\mathbb{Z}[z]\) of degree \(n\) with non-zero constant term and coefficients from \([-B, B]\). It will be convenient to set \(M = 2B+1\) in various places below. We set

\[
P^+(n, B) = \{ f \in P(n, B) \mid f \text{ is square free and has positive leading coefficient}\},
\]

\[
P^o(n, B) = \{ f \in P^+(n, B) \mid f \text{ is primitive}\},
\]

\[
sq(n, B) = |P^+(n, B)|,
\]

\[
psq(n, B) = |P^o(n, B)|.
\]

This situation is typical of experiments to measure statistics for algorithms on polynomials, we generate them by choosing the coefficients uniformly at random from the chosen range \([-B, B]\)
and accept the polynomial if it satisfies any further conditions (for efficiency we draw the leading coefficient from \([1, B]\)). The requirement in \(P^+(n, B)\) that \(f\) has positive leading coefficient and \(f(0) \neq 0\) is clearly reasonable in the context of root finding. Insisting that such a polynomial is primitive is also well motivated, but note that the procedure of first generating coefficients does not guarantee uniform sampling, for example all 8 members of \(P^+(1,2)\) are square free and all but 2\( \pm \)2 are primitive. We can correct this as follows. For a polynomial \(f\) define \(|f|\) to be the maximum absolute value of its coefficients. A generated polynomial \(f\) is accepted with probability \(1/c\) where \(c = |B/|p(p)|\) and \(p(p)\) is the primitive part of \(f\). The reason for this is that a primitive polynomial \(f\) of degree \(n\) and non-zero constant term accounts for \(|B/|f|\) members of \(P^+(n, B)\).

**Lemma 4.1** Let \(p\) be any prime that divides \(M = 2B + 1\). Then

1. \(sq(n, B) \geq (1 - 1/p)^2 B(2B + 1)^n\).

2. \(psq(n, B) \geq (1 - 1/p)^2 (2B + 1)^n\).

**Proof.** It suffices to show that \(|P(n, B)| \geq (1 - 1/p)^2 (M-1)^n\) since clearly \(sq(n, B) = |P(n, B)|/2\).

Define \(\phi : P(n, B) \to \mathbb{F}_p[z]\) by \(f \mapsto f \mod p\), where \(\mathbb{F}_p\) is the field of integers modulo \(p\). If \(g \in \mathbb{F}_p[z]\) is square free and has degree \(n\) then the same holds for every \(f \in \phi^{-1}(g)\). For if \(f = h^2 q\) where \(h\) is not a constant then \(p \nmid |c(h)|\) since \(p \nmid |c(f)|\) and so \(\phi(h)\) is not a constant. Hence \(\phi(f) = \phi(h)^2 |\phi(q)|\) is not square free, which is a contradiction.

Clearly the image of \(\phi\) contains all polynomials in \(\mathbb{F}_p[z]\) of degree \(n\). Furthermore \(|\phi^{-1}(g)| = (M/p)^{n+1}\) for all \(g \in \mathbb{F}_p[z]\) with \(n = \deg g\) and \(g(0) \neq 0\). For if \(g(0) \neq 0\) then \(f(0) \neq 0\), for all \(f \in \phi^{-1}(g)\). The cardinality claim follows from the fact that if \(r\) is a residue modulo \(p\) then there are \(M/p\) numbers in \([-B, B]\), equivalently in \([0, 2B]\), with residue \(r\); namely \(r + qp\) for \(q = 0, 1, \ldots, M/p - 1\). Similarly \(|\phi^{-1}(g)| = (M/p)^n (M/p - 1)\) for all \(g \in \mathbb{F}_p[z]\) with \(n = \deg g\) and \(g(0) = 0\). For if \(f \in \phi^{-1}(g)\) then \(f(0) \neq 0\) provided the constant term is non-zero and this is so for \((M/p)^n (M/p - 1)\) members of \(\phi^{-1}(g)\).

As is well known, the number of monic square free polynomials in \(\mathbb{F}_p[z]\) of degree \(n \geq 2\) is \(p^n - p^{n-1}\), e.g., see Mignotte [15]. The number of these that have 0 as a root is the number of square free monic polynomials of degree \(n - 1\). Hence, for \(n \geq 3\), the number of monic square free polynomials that do not have 0 as a root is \((p^n - p^{n-1}) - (p^{n-1} - p^{n-2}) = (1 - 1/p)^2 p^n\). For \(n = 2\) the number that have 0 as a root is the number of those of form \(z^2 + az\) with \(a \neq 0\), i.e., there are \(p - 1\) of them. Hence the number of monic square free polynomials of degree 2 that do not have 0 as a root is \((p^2 - p) - (p - 1) = (1 - 1/p)^2 p^2\). It follows that the formula \((1 - 1/p)^2 p^n\) applies to the case \(n = 2\) as well. Hence the number of square free such polynomials is \((p - 1)(1 - 1/p)^2 p^n\). These correspond to \((p - 1)(1 - 1/p)^2 p^n (M/p)^{n+1} = (1 - 1/p)^3 M^{n+1}\) members of \(P(n, B)\). Similarly the number of square free polynomials in \(\mathbb{F}_p[z]\) of degree \(n \geq 2\) with 0 as a root is \((p - 1)(p^{n-1} - p^{n-2}) (M/p)^n (M/p - 1) = (1 - 1/p)^2 M^n (M/p - 1)\) members of \(P(n, B)\). Thus the number polynomial in \(P(n, B)\) with \(n \geq 2\) is at least

\[
(1 - 1/p)^3 M^{n+1} + (1 - 1/p)^2 M^n (M/p - 1) = (1 - 1/p)^2 (M - 1) M^n.
\]

For \(n = 1\) we have \(|P(1, B)| = (M - 1)^2\). Since \(p \leq M\) we have \((1 - 1/p)^2 (M - 1) M < (M - 1)^2\).

For \(n = 0\) we have \(|P(0, B)| = M - 1 > (1 - 1/p)^2 (M - 1)\).

For the second claim suppose that \(f \in P^+(n, B)\). This corresponds to all \(c f \in P^+(n, B)\) for all \(c\) with \(1 \leq c \leq [B/|f|]\), i.e., to \([B/|f|] \leq B\) distinct members of \(P^+(n, B)\). Since different \(f\) give rise to different sets of members of \(P^+(n, B)\) it follows that \(|P^+(n, B)| \geq |P^+(n, B)|/B\) and the claim follows from part 1.

We note that the assumption that \(p\) divides \(M\) is for the sake of simplicity, if \(p \leq M\) is any prime then the preceding argument shows that \(sq(n, B) \geq (1 - 1/p)^2 p^n |M/n|^p |M/p - 1|\).

An alternative approach to a (weaker) lower bound for \(sq(n, B)\) is to observe that \(f \in P(n, B)\) is square free provided that the leading coefficient or the trailing coefficient is a square free integer.
Using \( Q(B) \) to denote the number of square free integers between 1 and \( B \) we then have that

\[
|P(n, B)| \geq 2Q(B)M^n + 4Q(B)(B - Q(B))M^{n-1}
\]

Now \( 6B/\pi^2 \leq Q(B) < 6B/\pi^2 + \sqrt{B} \), see Moser and McLeod [16], so that

\[
|P(n, B)| > \frac{12}{\pi^2} BM^n + \frac{24}{\pi^2} B^2 \left( 1 - \frac{6}{\pi^2} - \frac{1}{\sqrt{B}} \right) M^{n-1}
\]

\[
= \frac{6}{\pi^2} (M - 1)M^n + \frac{6}{\pi^2} \left( 1 - \frac{6}{\pi^2} - \frac{1}{\sqrt{B}} \right) (M - 1)^2 M^{n-1}
\]

Since \( \text{sq}(n, B) = |P(n, B)|/2 \) we have

\[
\text{sq}(n, B) \geq \frac{6}{\pi^2} B(2B + 1)^n + \frac{12}{\pi^2} \left( 1 - \frac{6}{\pi^2} - \frac{1}{\sqrt{B}} \right) B^2(2B + 1)^{n-1}.
\]

### 4.1 Expected breadth

We define two open discs parametrised by \( m \). \( S_m \) is the disc given by

\[
|z| < \frac{m + 2}{m + 3}.
\]

Note that \( L_m \) is the smallest disc centred at the origin that encloses \( S_m \).

**Root Distribution Assumption:** Let \( N_m \) denote the expected number of roots in \( L_m \) of members of \( P^+(n, B) \) drawn uniformly at random. There is a constant \( m_0 \) such that for all \( m \geq m_0 \) the expected number of roots in \( S_m \) is at most \( N_m(s_m/l_m)^2 \) where \( s_m \) is the radius of \( S_m \) and \( l_m \) is the radius of \( L_m \). The same applies to the situation in which members of \( P^c(n, B) \) are drawn uniformly at random.

It follows from the second bound in Lemma 2.8 that \( N_m \leq \log((n + 1)/B)/\log((m + 3)/(m + 2)) \).

Shepp and Vanderbei [18] study the expected number of complex roots in a region for the case of coefficients drawn from a normal distribution while Ibragimov and Zeitouni [11] prove a more general result. Unfortunately it does not seem possible to deduce the assumption stated above from these results. An intuitive justification is that roots cluster around the unit disc and are distributed fairly uniformly by angle, see the experimental evidence presented in [4.4].

**Lemma 4.2** Assume that members of \( P^+(n, B) \) are drawn uniformly at random and that the root distribution assumption holds. Then

\[
E[||\hat{C}(f^{IT})||] + E[||\hat{C}(f^{TT})||] + E[||\hat{C}(f^{TT})||] + E[||\hat{C}(f^{TT})||] < (4.44 + m_0) \log(n + 1)/B.
\]

The same inequality holds for \( P^0(f) \).

**Proof.** As in Theorem 3.1 we consider the four cases of \( \hat{C}_m(f^{IT}), \hat{C}_m(f^{TT}), \hat{C}_m(f^{TT}) \) and \( \hat{C}_m(f^{TT}) \), but with \( \hat{C} \) instead of \( C \), using the same enumeration as there. It is shown there that \( ||\hat{C}(f^{IT})|| + ||\hat{C}(f^{TT})|| \leq (2/\log 2) \log L(f) \); the inequality there is observed for \( C \) but it is clearly valid for \( \hat{C} \).

For the first case, the disc for \( m \) is contained in the disc \( L_m \) of the root distribution assumption. The discs for \( m < m_0 \) are all contained in the disc centred at the origin with radius \((m_0 + 1)/(m_0 + 2)\). Thus the total number of roots of any polynomial is no more than \( \log(n + 1)/B/\log((m_0 + 2)/(m_0 + 1)) \). The transformation \( f(z) \mapsto (-1)^{\deg f}f(-z) \) preserves \( P^+(n, B) \) as well as \( P^0(n, B) \).
and shows that the number of roots with strictly positive real part is no more than half the total number of roots. Thus \( \sum_{0 \leq i < m_0} |\tilde{C}_m(f^{II})| \leq \log(n+1)B/2 \log((m_0+2)/(m_0+1)) \). Assuming that \( m_0 > 0 \),

\[
\mathbb{E}[|\tilde{C}(f^{II})|] = \mathbb{E}\left[ \sum_{m=0}^{m_0-1} |\tilde{C}_m(f)| \right] + \mathbb{E}\left[ \sum_{m=m_0}^{\infty} |\tilde{C}_m(f)| \right] \\
\leq \frac{\log(n+1)B}{2\log \left( \frac{m_0+2}{m_0+1} \right)} + \log(n+1)B \sum_{m=m_0}^{\infty} \frac{1}{4(m+2)^2} \left( \frac{m+3}{m+2} \right)^2 \frac{1}{\log \left( \frac{m+3}{m+2} \right)}.
\]

A simple argument shows that \( \log(1+x) > 2x/(x+2) \) for all \( x > 0 \), hence \( 1/\log(1+x) < 1/(x+1/2) \). Now

\[
\mathbb{E}[|\tilde{C}(f^{II})|] < \log(n+1)B \left( \frac{2m_0+3}{4} + \sum_{m=m_0}^{\infty} \frac{m+5/2}{4(m+2)^4} \right) \\
\leq \log(n+1)B \left( \frac{2m_0+3}{4} + \frac{m_0+5/2}{4(m_0+2)^4} + \int_{m_0}^{\infty} \frac{m+5/2}{4(m+2)^4} dm \right) \\
= \log(n+1)B \left( \frac{2m_0+3}{4} + \frac{m_0+5/2}{4(m_0+2)^4} + \left[ -\frac{3m-7}{24(m+2)^3} \right]_{m_0}^{\infty} \right) \\
= \left( \frac{m_0+2}{2} - \frac{1}{4} + \frac{m_0+2}{8(m_0+2)^2} + \frac{1}{24(m_0+2)^3} + \frac{1}{8(m_0+2)^4} \right) \log(n+1)B
\]

As pointed out in the proof of Theorem 3.1, the first and third cases transform to each other by the transform \( f(z) \mapsto (-z)^{\deg(f)}f(1/z) \). Hence \( |\tilde{C}(f^{II})| \leq |\tilde{C}(f^{IT})| \) and \( |\tilde{C}(f^{TI})| \leq |\tilde{C}(f^{II})| \), hence \( |\tilde{C}(f^{II})| = |\tilde{C}(f^{II})| \). Thus bound above applies to \( \mathbb{E}[|\tilde{C}(f^{II})|] \) and so

\[
\mathbb{E}[|\tilde{C}(f^{II})|] + \mathbb{E}[|\tilde{C}(f^{IT})|] + \mathbb{E}[|\tilde{C}(f^{TI})|] + \mathbb{E}[|\tilde{C}(f^{TT})|] \\
\leq 2 \left( \frac{1}{\log 2} - \frac{1}{4} + \frac{m_0+2}{2} + \frac{1}{8(m_0+2)^2} + \frac{1}{24(m_0+2)^3} + \frac{1}{8(m_0+2)^4} \right) \log(n+1)B \\
\leq 2 \left( \frac{1}{\log 2} - \frac{1}{4} + \frac{m_0+2}{2} + \frac{1}{8 \cdot 3^2} + \frac{1}{24 \cdot 3^3} + \frac{1}{8 \cdot 3^4} \right) \log(n+1)B \\
= \left( \frac{2 \log 2 + 503}{324} + m_0 \right) \log(n+1)B
\]

If \( m_0 = 0 \) the argument above becomes simpler and leads to

\[
\mathbb{E}[|\tilde{C}(f^{II})|] + \mathbb{E}[|\tilde{C}(f^{IT})|] + \mathbb{E}[|\tilde{C}(f^{TI})|] + \mathbb{E}[|\tilde{C}(f^{TT})|] \\
\leq 2 \left( \frac{1}{\log 2} + \frac{29}{384} \right) \log(n+1)B < 3.037 \log(n+1)B.
\]

**Theorem 4.1** Assume that members \( f \) of \( P^+(n, B) \) or of \( P^o(f) \) are drawn uniformly at random and that the root distribution assumption holds. Then

\[
\mathbb{E}[br(f)] < (4.44 + m_0) \log(n+1)B + 1.45 \log B + 5.
\]

**Proof.** As in Theorem 3.1 we have

\[
\mathbb{E}[br(f)] \leq \mathbb{E}[br(f^{II})] + \mathbb{E}[br(f^{IT})] + \mathbb{E}[br(f^{TI})] + \mathbb{E}[br(f^{TT})] + \mathbb{E}[br(f^{TT})].
\]
It follows from the first part of Lemma 2.3 that 
\[ \mathbb{E}[|\hat{C}(f^{1/2})|] + \mathbb{E}[|\hat{C}(f^{1/2})|] + \mathbb{E}[|\hat{C}(f^{1/2})|] + \mathbb{E}[|\hat{C}(f^{1/2})|] \]
accounts for all the real roots except for the integer roots of \( f \). By Lemmas 2.4 and 2.9 we now have
\[ \mathbb{E}[\text{br}(f)] \leq \mathbb{E}[|\hat{C}(f^{1/2})|] + \mathbb{E}[|\hat{C}(f^{1/2})|] + \mathbb{E}[|\hat{C}(f^{1/2})|] + \mathbb{E}[|\hat{C}(f^{1/2})|] + 1 + \log B/\log 2 + 4. \]
The result follows by Lemma 4.2

4.2 ALTERNATIVE HYPOTHESIS

The experimental evidence of [14] suggests that the Root Distribution Assumption is true, with one exception, for all \( m \geq 4 \) provided we replace the upper bound estimate \( N_m(s_m/l_m)^2 \) by \( N_m(s_m/l_m)^{1+\epsilon} \) for some \( \epsilon > 0 \). Under this assumption the first bound in Lemma 4.2 becomes
\[ \mathbb{E}[|\hat{C}(f^{1/2})|] \leq \log(n + 1)B \left( \sum_{m=0}^{\infty} \frac{m + 5/2}{21+\epsilon(m + 2)^2} \right) \]
\[ \leq \log(n + 1)B \left( \sum_{m=0}^{4} \frac{m + 5/2}{2(m + 2)^2} + \int_{m=4}^{\infty} \frac{m + 5/2}{21+\epsilon(m + 2)^2} \, dm \right) \]
\[ = \log(n + 1)B \left( \frac{12209}{14400} + \frac{-(1 + 2\epsilon)m + 2 + 5\epsilon}{\epsilon(1 + 2\epsilon)2^{2+\epsilon}(m + 2)^{1+2\epsilon}} \right) \]
\[ < \log(n + 1)B \left( \frac{12209}{14400} + \frac{6 + 13\epsilon}{\epsilon(1 + 2\epsilon)2^{2+\epsilon}(6)^{1+2\epsilon}} \right) \]
A simple argument shows that \( (6 + 13\epsilon)/\epsilon(1 + 2\epsilon)2^{2+\epsilon}6^{1+2\epsilon} \leq 1/2\log 2 \) and so
\[ \mathbb{E}[|\hat{C}(f^{1/2})|] < \log(n + 1)B \left( \frac{12209}{14400} + \frac{1}{2\epsilon\log 2} \right). \]

Thus
\[ \mathbb{E}[|\hat{C}(f^{1/2})|] + \mathbb{E}[|\hat{C}(f^{1/2})|] + \mathbb{E}[|\hat{C}(f^{1/2})|] + \mathbb{E}[|\hat{C}(f^{1/2})|] \]
\[ < 2 \left( \frac{1}{\log 2} + \frac{12209}{14400} + \frac{1}{2\epsilon\log 2} \right) \log(n + 1)B. \]

The bound of Theorem 4.1 now becomes
\[ \mathbb{E}[\text{br}(f)] < 2 \left( \frac{1}{\log 2} + \frac{12209}{14400} + \frac{1}{2\epsilon\log 2} \right) \log(n + 1)B + 1.45\log B + 5 \]
\[ < \left( \frac{2}{\pi} \log n + o(\log n) \right) \log(n + 1)B + 1.45\log B + 5. \]

4.3 DEPENDENCE OF EXPECTED BREATH ON THE REAL ROOTS

Lemma 2.4 shows that \( \mathbb{E}[\text{br}(f)] \leq \mathbb{E}[R^+(f)] + \mathbb{E}[C(f)] + 1. \) In this section we show that \( \mathbb{E}[R^+(f)] \) is in line with the bound of Theorem 4.1 but without any assumptions.

Consider the situation in which we draw the coefficients of non-zero polynomials of degree at most \( n \) uniformly at random from \([-B,B] \), where \( B \geq 1 \). Thus the probability of drawing any given polynomial is \( 1/(M^{n+1} - 1) \) since the zero polynomial is excluded. Let \( N(n,B) \) be the expected number of real roots. It is shown by Ibragimov and Maslova [12] that if we draw the coefficients independently from the same distribution that has mean zero and which belongs to the domain of attraction of the normal law then the expected number of real roots is equal to
\[ \frac{2}{\pi} \log n + o(\log n). \]

This generalises the corresponding result of Kac [13] for the standard normal distribution. The situation of [12] covers our case, as observed by Cucker and Roy [7]. Since the transformation
\( z \mapsto (-1)^{\deg(f)} z \) swaps positive and negative roots while keeping the class of polynomials invariant. It follows that the number of expected positive real roots, \( N^+(n, B) \), is equal to
\[
\frac{1}{\pi} \log n + o(\log n).
\]

Consider now drawing the members of \( P^+(n, B) \), respectively \( P^o(n, B) \), uniformly at random and denote the expected number of real roots by \( N^+(n, B) \), respectively \( N^o(n, B) \).

**Lemma 4.3** Let \( p \) be a prime that divides \( M = 2B + 1 \). Then
\[
N^+(n, B) < \frac{(2 + 1/B)}{\pi(1 - 1/p)^2} \log n + o(\log n) < 2.13 \log n + o(\log n),
\]
and
\[
N^o(n, B) < \frac{(2B + 1)}{\pi(1 - 1/p)^2} (\log n + o(\log n)) < 0.72(2B + 1) (\log n + o(\log n)).
\]

**Proof.** Using the first part of Lemma 4.1 in the fourth line below, we have
\[
N^+(n, B) = \sum_{f \in P^+(n, B)} \frac{R^+(f)}{sq(n, B)}
\]
\[
= \frac{M^{n+1} - 1}{sq(n, B)} \sum_{f \in P^+(n, B)} \frac{R^+(f)}{M^{n+1} - 1}
\]
\[
\leq \frac{M^{n+1} - 1}{sq(n, B)} \sum_{f \in P(n, B)} \frac{R^+(f)}{M^{n+1} - 1}
\]
\[
< \frac{M^{n+1}}{(1 - 1/p)^2 BM} N^+(n, B)
\]
\[
= \frac{(2 + 1/B)}{\pi(1 - 1/p)^2} \log n + o(\log n).
\]

The numerical constant follows from the fact that \( M \geq 3 \) is odd and so \( p \geq 3 \). The second inequality follows similarly by using the second part of Lemma 4.1. \( \square \)

### 4.4 Experimental evidence

The results presented here are based on data produced by Zhao Zheng, a student at the School of Informatics in Edinburgh working under the supervision of the author. The root finding package used from [19] is based on Laguerre’s method (org.apache.commons.math3.analysis.solvers.LaguerreSolver).

In the tables we present the expected number of roots found in the discs \( S_m \) and \( L_m \) of the Root Distribution Assumption (denoted by \( E[S_m] \) and \( E[L_m] \) respectively) in the first two columns. The third column gives the estimated upper bound on \( \frac{E[S_m]}{E[L_m]} \) represented by \( u_m = \frac{E[L_m]}{E[S_m]}(s_m/l_m)^2 \). The fourth column gives \( \epsilon_m = (E[S_m]/E[L_m])/\log(s_m/l_m) \) from which we may derive an approximate value of \( \epsilon \) in the modified conjecture of §4.2. The polynomials generated were both square free and primitive. A comparison was made with purely square free polynomials (for two cases), these showed no significant difference with the corresponding cases of square free and primitive polynomials, we present only the data for the square free primitive case in Tables 1–9.

In each experiment polynomials with the appropriate degree and coefficient bound were generated uniformly at random then accepted with appropriate probability if they were square free and primitive. Each experiment was run for a maximum time, this is the reason that the sample sizes vary somewhat. Once the roots were found for each polynomial the polynomial was evaluated at each root and this sample was rejected if any value was not close to 0 (within \( 10^{-8} \)). As can be seen from the tables the two hypotheses hold after an initial period and stay this way. One
exception is the case for $n = 30$ and $B = 100$ (Table 9) where the result for $m = 1024$ is somewhat anomalous. It is not clear if this is a genuine trend or a software issue, unfortunately further investigation was not possible due to time constraints but it will be investigated in the future.

The bounds obtained in §§4.1, 4.2 are dominated by $c \log(n + 1)B$ for some constant $c$. By Lemmas 2.3 and 2.4 $E[|\hat{C}|]$ is dominated by $E[|C|]$. Figure 2 shows the ratio of $E[|\hat{C}|]$ to $\log(n + 1)B$. In all cases this is well below the theoretical constant, however there is an increase as $n$ increases for any given $B$.

Acknowledgement. The author is grateful to an anonymous referee for helpful comments.

References


<table>
<thead>
<tr>
<th>$m$</th>
<th>$E[S_m]$</th>
<th>$E[L_m]$</th>
<th>$u_m$</th>
<th>$e_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.700392</td>
<td>0.056823</td>
<td>0.002162</td>
<td>0.868987</td>
</tr>
<tr>
<td>2</td>
<td>1.183154</td>
<td>0.038298</td>
<td>0.001155</td>
<td>0.989846</td>
</tr>
<tr>
<td>4</td>
<td>1.953484</td>
<td>0.023188</td>
<td>3.768294E-4</td>
<td>1.036732</td>
</tr>
<tr>
<td>8</td>
<td>2.887957</td>
<td>0.010172</td>
<td>7.219982E-5</td>
<td>1.066128</td>
</tr>
<tr>
<td>16</td>
<td>3.715613</td>
<td>0.002733</td>
<td>8.848721E-6</td>
<td>1.114440</td>
</tr>
<tr>
<td>32</td>
<td>4.209781</td>
<td>0.000000</td>
<td>7.875604E-7</td>
<td>—</td>
</tr>
<tr>
<td>64</td>
<td>4.465587</td>
<td>0.000000</td>
<td>5.883596E-8</td>
<td>—</td>
</tr>
<tr>
<td>128</td>
<td>4.595769</td>
<td>0.000000</td>
<td>4.022766E-9</td>
<td>—</td>
</tr>
<tr>
<td>256</td>
<td>4.652807</td>
<td>0.000000</td>
<td>2.625284E-10</td>
<td>—</td>
</tr>
<tr>
<td>512</td>
<td>4.681456</td>
<td>0.000000</td>
<td>1.676756E-11</td>
<td>—</td>
</tr>
<tr>
<td>1024</td>
<td>4.695251</td>
<td>0.000000</td>
<td>1.059277E-12</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 1: Results for $n = 10$, $B = 1$ based on 162803 square free and primitive polynomials generated at random. $E[|\hat{C}|] = 2.511434$.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$E[S_m]$</th>
<th>$E[L_m]$</th>
<th>$u_m$</th>
<th>$e_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.213440</td>
<td>0.054137</td>
<td>0.003745</td>
<td>1.075881</td>
</tr>
<tr>
<td>2</td>
<td>1.649975</td>
<td>0.039199</td>
<td>0.001611</td>
<td>1.079093</td>
</tr>
<tr>
<td>4</td>
<td>2.343683</td>
<td>0.021803</td>
<td>4.520993E-4</td>
<td>1.093707</td>
</tr>
<tr>
<td>8</td>
<td>3.193095</td>
<td>0.009437</td>
<td>7.982736E-5</td>
<td>1.099237</td>
</tr>
<tr>
<td>16</td>
<td>3.920581</td>
<td>0.003261</td>
<td>9.336850E-6</td>
<td>1.095493</td>
</tr>
<tr>
<td>32</td>
<td>4.412287</td>
<td>8.946619E-4</td>
<td>8.254449E-7</td>
<td>1.097806</td>
</tr>
<tr>
<td>64</td>
<td>4.691315</td>
<td>2.783392E-4</td>
<td>6.181002E-8</td>
<td>1.072740</td>
</tr>
<tr>
<td>128</td>
<td>4.838881</td>
<td>4.638987E-5</td>
<td>4.235567E-9</td>
<td>1.108062</td>
</tr>
<tr>
<td>256</td>
<td>4.909407</td>
<td>0.000000</td>
<td>2.770676E-10</td>
<td>—</td>
</tr>
<tr>
<td>512</td>
<td>4.943935</td>
<td>0.000000</td>
<td>1.770762E-11</td>
<td>—</td>
</tr>
<tr>
<td>1024</td>
<td>4.960595</td>
<td>0.000000</td>
<td>1.119140E-12</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 2: Results for $n = 10$, $B = 10$ based on 150895 square free and primitive polynomials generated at random. $E[|\hat{C}|] = 3.021147$.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$E[S_m]$</th>
<th>$E[L_m]$</th>
<th>$u_m$</th>
<th>$e_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.258061</td>
<td>0.054619</td>
<td>0.003883</td>
<td>1.085311</td>
</tr>
<tr>
<td>2</td>
<td>1.689504</td>
<td>0.038297</td>
<td>0.001650</td>
<td>1.092643</td>
</tr>
<tr>
<td>4</td>
<td>2.387732</td>
<td>0.021852</td>
<td>4.605965E-4</td>
<td>1.097535</td>
</tr>
<tr>
<td>8</td>
<td>3.229419</td>
<td>0.009366</td>
<td>8.073548E-5</td>
<td>1.102794</td>
</tr>
<tr>
<td>16</td>
<td>3.955142</td>
<td>0.003108</td>
<td>9.419158E-6</td>
<td>1.104266</td>
</tr>
<tr>
<td>32</td>
<td>4.444238</td>
<td>8.447456E-4</td>
<td>8.314222E-7</td>
<td>1.106149</td>
</tr>
<tr>
<td>64</td>
<td>4.722547</td>
<td>2.034789E-4</td>
<td>6.222152E-8</td>
<td>1.108002</td>
</tr>
<tr>
<td>128</td>
<td>4.871740</td>
<td>5.549424E-5</td>
<td>4.264329E-9</td>
<td>1.091527</td>
</tr>
<tr>
<td>256</td>
<td>4.949247</td>
<td>6.166462E-6</td>
<td>2.792546E-10</td>
<td>1.152268</td>
</tr>
<tr>
<td>512</td>
<td>4.988075</td>
<td>0.000000</td>
<td>1.786572E-11</td>
<td>—</td>
</tr>
<tr>
<td>1024</td>
<td>5.007467</td>
<td>0.000000</td>
<td>1.129714E-12</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 3: Results for $n = 10$, $B = 100$ based on 162179 square free and primitive polynomials generated at random. $E[|\hat{C}|] = 3.051264$. 
Table 4: Results for \( n = 20, \ B = 1 \) based on 161519 square free and primitive polynomials generated at random. \( \mathbb{E}[\hat{C}] = 4.063813 \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \mathbb{E}[S_m] )</th>
<th>( \mathbb{E}[L_m] )</th>
<th>( u_m )</th>
<th>( \varepsilon_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.709483</td>
<td>0.058185</td>
<td>0.002190</td>
<td>0.865255</td>
</tr>
<tr>
<td>2</td>
<td>1.238127</td>
<td>0.040825</td>
<td>0.001209</td>
<td>0.984513</td>
</tr>
<tr>
<td>4</td>
<td>2.231433</td>
<td>0.026331</td>
<td>4.303903E-4</td>
<td>1.038078</td>
</tr>
<tr>
<td>8</td>
<td>3.910605</td>
<td>0.013800</td>
<td>9.776512E-5</td>
<td>1.065765</td>
</tr>
<tr>
<td>16</td>
<td>5.989258</td>
<td>0.005232</td>
<td>1.426340E-6</td>
<td>1.087910</td>
</tr>
<tr>
<td>32</td>
<td>7.700692</td>
<td>0.001387</td>
<td>1.440633E-7</td>
<td>1.113115</td>
</tr>
<tr>
<td>64</td>
<td>8.732849</td>
<td>7.429467E-5</td>
<td>1.150589E-7</td>
<td>1.286814</td>
</tr>
<tr>
<td>128</td>
<td>9.267578</td>
<td>0.000000</td>
<td>8.112092E-9</td>
<td>—</td>
</tr>
<tr>
<td>256</td>
<td>9.538061</td>
<td>0.000000</td>
<td>5.381722E-10</td>
<td>—</td>
</tr>
<tr>
<td>512</td>
<td>9.673401</td>
<td>0.000000</td>
<td>3.464709E-11</td>
<td>—</td>
</tr>
<tr>
<td>1024</td>
<td>9.741851</td>
<td>0.000000</td>
<td>2.197820E-12</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 5: Results for \( n = 20, \ B = 10 \) based on 155512 square free and primitive polynomials generated at random. \( \mathbb{E}[\hat{C}] = 4.459321 \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \mathbb{E}[S_m] )</th>
<th>( \mathbb{E}[L_m] )</th>
<th>( u_m )</th>
<th>( \varepsilon_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.245698</td>
<td>0.055668</td>
<td>0.003845</td>
<td>1.075312</td>
</tr>
<tr>
<td>2</td>
<td>1.742496</td>
<td>0.040453</td>
<td>0.001702</td>
<td>1.085750</td>
</tr>
<tr>
<td>4</td>
<td>2.710376</td>
<td>0.025895</td>
<td>5.228349E-4</td>
<td>1.087480</td>
</tr>
<tr>
<td>8</td>
<td>4.341311</td>
<td>0.014121</td>
<td>1.085328E-4</td>
<td>1.081147</td>
</tr>
<tr>
<td>16</td>
<td>6.326965</td>
<td>0.005408</td>
<td>1.506765E-5</td>
<td>1.091261</td>
</tr>
<tr>
<td>32</td>
<td>7.928642</td>
<td>0.001556</td>
<td>1.483286E-6</td>
<td>1.102010</td>
</tr>
<tr>
<td>64</td>
<td>8.922128</td>
<td>4.694172E-4</td>
<td>1.175527E-7</td>
<td>1.085985</td>
</tr>
<tr>
<td>128</td>
<td>9.465688</td>
<td>8.359484E-5</td>
<td>8.285501E-9</td>
<td>1.115934</td>
</tr>
<tr>
<td>256</td>
<td>9.747974</td>
<td>2.572149E-5</td>
<td>5.500601E-10</td>
<td>1.088665</td>
</tr>
<tr>
<td>512</td>
<td>9.888832</td>
<td>0.000000</td>
<td>3.541869E-11</td>
<td>—</td>
</tr>
<tr>
<td>1024</td>
<td>9.958228</td>
<td>0.000000</td>
<td>2.246636E-12</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 6: Results for \( n = 20, \ B = 100 \) based on 154807 square free and primitive polynomials generated at random. \( \mathbb{E}[\hat{C}] = 4.483273 \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \mathbb{E}[S_m] )</th>
<th>( \mathbb{E}[L_m] )</th>
<th>( u_m )</th>
<th>( \varepsilon_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.057730</td>
<td>0.057730</td>
<td>0.004015</td>
<td>1.077719</td>
</tr>
<tr>
<td>2</td>
<td>1.797625</td>
<td>0.041200</td>
<td>0.001755</td>
<td>1.089463</td>
</tr>
<tr>
<td>4</td>
<td>2.763467</td>
<td>0.026917</td>
<td>5.330762E-4</td>
<td>1.082962</td>
</tr>
<tr>
<td>8</td>
<td>4.399762</td>
<td>0.013158</td>
<td>1.099941E-4</td>
<td>1.097000</td>
</tr>
<tr>
<td>16</td>
<td>6.368575</td>
<td>0.005174</td>
<td>1.516674E-5</td>
<td>1.099099</td>
</tr>
<tr>
<td>32</td>
<td>7.962050</td>
<td>0.001518</td>
<td>1.489530E-6</td>
<td>1.105755</td>
</tr>
<tr>
<td>64</td>
<td>8.948142</td>
<td>4.457163E-4</td>
<td>1.178955E-7</td>
<td>1.092017</td>
</tr>
<tr>
<td>128</td>
<td>9.487465</td>
<td>8.397553E-5</td>
<td>8.304563E-9</td>
<td>1.115719</td>
</tr>
<tr>
<td>256</td>
<td>9.771858</td>
<td>1.291931E-5</td>
<td>5.513639E-10</td>
<td>1.147234</td>
</tr>
<tr>
<td>512</td>
<td>9.915928</td>
<td>0.000000</td>
<td>3.551574E-11</td>
<td>—</td>
</tr>
<tr>
<td>1024</td>
<td>9.987991</td>
<td>0.000000</td>
<td>2.253350E-12</td>
<td>—</td>
</tr>
</tbody>
</table>
Table 7: Results for \( n = 30, B = 1 \) based on 152639 square free and primitive polynomials generated at random. \( \mathbb{E}[|\hat{C}|] = 5.313\,858. \)

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \mathbb{E}[S_m] )</th>
<th>( \mathbb{E}[L_m] )</th>
<th>( u_m )</th>
<th>( \epsilon_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.712,845</td>
<td>0.062,848</td>
<td>0.002,200</td>
<td>0.840,221</td>
</tr>
<tr>
<td>2</td>
<td>1.236,715</td>
<td>0.043,455</td>
<td>0.001,208</td>
<td>0.966,166</td>
</tr>
<tr>
<td>4</td>
<td>2.237,816</td>
<td>0.027,267</td>
<td>4.316,775,E,-4</td>
<td>1.030,611</td>
</tr>
<tr>
<td>8</td>
<td>4.165,344</td>
<td>0.015,311</td>
<td>1.041,336,E,-4</td>
<td>1.058,073</td>
</tr>
<tr>
<td>16</td>
<td>7.188,359</td>
<td>0.006,984</td>
<td>1.711,905,E,-5</td>
<td>1.071,477</td>
</tr>
<tr>
<td>32</td>
<td>10.283,021</td>
<td>0.002,031</td>
<td>1.923,734,E,-6</td>
<td>1.101,200</td>
</tr>
<tr>
<td>64</td>
<td>12.413,197</td>
<td>4.651,498,E,-4</td>
<td>1.635,490,E,-7</td>
<td>1.123,390</td>
</tr>
<tr>
<td>128</td>
<td>13.588,310</td>
<td>2.620,562,E,-5</td>
<td>1.189,411,E,-8</td>
<td>1.261,841</td>
</tr>
<tr>
<td>256</td>
<td>14.189,264</td>
<td>0.000,000</td>
<td>8.006,101,E,-10</td>
<td>—</td>
</tr>
<tr>
<td>512</td>
<td>14.489,239</td>
<td>0.000,000</td>
<td>5.189,591,E,-11</td>
<td>—</td>
</tr>
<tr>
<td>1024</td>
<td>14.638,729</td>
<td>0.000,000</td>
<td>3.302,585,E,-12</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 8: Results for \( n = 30, B = 10 \) based on 152852 square free and primitive polynomials generated at random. \( \mathbb{E}[|\hat{C}|] = 5.832\,158. \)

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \mathbb{E}[S_m] )</th>
<th>( \mathbb{E}[L_m] )</th>
<th>( u_m )</th>
<th>( \epsilon_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.278,439</td>
<td>0.064,402</td>
<td>0.003,946</td>
<td>1.033,863</td>
</tr>
<tr>
<td>2</td>
<td>1.779,597</td>
<td>0.044,520</td>
<td>0.001,738</td>
<td>1.064,189</td>
</tr>
<tr>
<td>4</td>
<td>2.778,407</td>
<td>0.027,307</td>
<td>5.359,581,E,-4</td>
<td>1.080,859</td>
</tr>
<tr>
<td>8</td>
<td>4.693,874</td>
<td>0.015,662</td>
<td>1.173,468,E,-4</td>
<td>1.076,335</td>
</tr>
<tr>
<td>16</td>
<td>7.673,410</td>
<td>0.006,568</td>
<td>1.827,420,E,-5</td>
<td>1.091,035</td>
</tr>
<tr>
<td>32</td>
<td>10.658,205</td>
<td>0.002,218</td>
<td>1.993,923,E,-6</td>
<td>1.094,462</td>
</tr>
<tr>
<td>64</td>
<td>12.713,808</td>
<td>5.888,049,E,-4</td>
<td>1.675,097,E,-7</td>
<td>1.100,044</td>
</tr>
<tr>
<td>128</td>
<td>13.879,733</td>
<td>1.766,415,E,-4</td>
<td>1.214,920,E,-8</td>
<td>1.080,896</td>
</tr>
<tr>
<td>256</td>
<td>14.492,444</td>
<td>2.616,910,E,-5</td>
<td>8.177,166,E,-10</td>
<td>1.120,814</td>
</tr>
<tr>
<td>512</td>
<td>14.799,885</td>
<td>0.000,000</td>
<td>5.300,854,E,-11</td>
<td>—</td>
</tr>
<tr>
<td>1024</td>
<td>14.951,463</td>
<td>0.000,000</td>
<td>3.373,139,E,-12</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 9: Results for \( n = 30, B = 100 \) based on 143495 square free and primitive polynomials generated at random. \( \mathbb{E}[|\hat{C}|] = 5.949\,692. \)

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \mathbb{E}[S_m] )</th>
<th>( \mathbb{E}[L_m] )</th>
<th>( u_m )</th>
<th>( \epsilon_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.349,239</td>
<td>0.084,909</td>
<td>0.004,164</td>
<td>0.956,872</td>
</tr>
<tr>
<td>2</td>
<td>1.859,514</td>
<td>0.054,657</td>
<td>0.001,816</td>
<td>1.017,675</td>
</tr>
<tr>
<td>4</td>
<td>2.849,841</td>
<td>0.027,416</td>
<td>5.497,379,E,-4</td>
<td>1.085,871</td>
</tr>
<tr>
<td>8</td>
<td>4.753,587</td>
<td>0.015,555</td>
<td>1.188,397,E,-4</td>
<td>1.080,022</td>
</tr>
<tr>
<td>16</td>
<td>7.703,725</td>
<td>0.007,004</td>
<td>1.834,640,E,-5</td>
<td>1.081,732</td>
</tr>
<tr>
<td>32</td>
<td>10.677,306</td>
<td>0.002,077</td>
<td>1.997,496,E,-6</td>
<td>1.103,179</td>
</tr>
<tr>
<td>64</td>
<td>12.735,517</td>
<td>5.296,352,E,-4</td>
<td>1.677,957,E,-7</td>
<td>1.111,906</td>
</tr>
<tr>
<td>128</td>
<td>13.895,209</td>
<td>1.533,154,E,-4</td>
<td>1.216,275,E,-8</td>
<td>1.094,584</td>
</tr>
<tr>
<td>256</td>
<td>14.506,910</td>
<td>0.000,000</td>
<td>8.185,328,E,-10</td>
<td>—</td>
</tr>
<tr>
<td>512</td>
<td>14.822,266</td>
<td>0.000,000</td>
<td>5.308,870,E,-11</td>
<td>—</td>
</tr>
<tr>
<td>1024</td>
<td>14.980,327</td>
<td>6.968,884,E,-6</td>
<td>3.379,651,E,-12</td>
<td>1.001,429</td>
</tr>
</tbody>
</table>