A Faster Solution to Smale's $17^{\underline{\text{th}}}$ Problem I: Real Binomial Systems

Grigoris Paouris*
grigoris@math.tamu.edu
Texas A&M University
College Station, Texas

Kaitlyn Phillipson†
kphillip@stedwards.edu
St. Edwards University
Austin, Texas

J. Maurice Rojas‡
rojas@math.tamu.edu
Texas A&M University
College Station, Texas

ABSTRACT

Suppose $F := (f_1, \ldots, f_n)$ is a system of random n-variate polynomials with f_i having degree $\leq d_i$ and the coefficient of $x_1^{a_1} \cdots x_n^{a_n}$ in f_i being an independent complex Gaussian of mean 0 and variance $\frac{d_i!}{a_1! \cdots a_n! \left(d_i - \sum_{j=1}^n a_j\right)!}$. Recent progress on Smale's $17^{\underline{\text{th}}}$ Problem by Lairez — building upon seminal work of Shub, Smale, Beltrán, Pardo, Bürgisser, and Cucker — has resulted in a deterministic algorithm that finds a single (complex) approximate root of F using just $N^{O(1)}$ arithmetic operations on average, where $N := \sum_{i=1}^n \frac{(n+d_i)!}{n! d_i!}$ $(= n(n + \max_i d_i)^{O(\min\{n, \max_i d_i\})})$ is the maximum possible total number of monomial terms for such an F. However, can one go faster when the number of terms is smaller, and we restrict to real coefficient and real roots? And can one still maintain average-case polynomial-time with more general probability measures?

We show that the answer is yes when F is instead a binomial system — a case whose numerical solution is a key step in polyhedral homotopy algorithms for solving arbitrary polynomial systems. We give a deterministic algorithm that finds a real approximate root, or correctly decides there are none, using just $O(n^3 \log^2(n \max_i d_i))$ arithmetic operations on average. Furthermore, our approach allows Gaussians with arbitrary variance. We also discuss briefly the obstructions to maintaining average-case time polynomial in $n \log \max_i d_i$ when F has more terms.

*Partially supported by NSF grant DMS-1812240.
†Partially supported by NSF REU grant DMS-1460766 and NSF grant CCF-1409020.
‡Partially supported by NSF REU grant DMS-1757872, and NSF grants CCF-1409020 and CCF-1900881.

ISSAC '19, July 15–18, 2019, Beijing, China

ACM ISBN 978-1-4503-6084-5/19/07…$15.00
https://doi.org/10.1145/3326229.3326267

CCS CONCEPTS

• **Theory of computation** → **Numeric approximation algorithms**; • **Mathematics of computing** → *Probabilistic algorithms*.

KEYWORDS

Smale's 17th Problem, real roots, sparse polynomial, average-case complexity, Newton iteration, approximate root

ACM Reference Format:
Grigoris Paouris, Kaitlyn Phillipson, and J. Maurice Rojas. 2019. A Faster Solution to Smale's $17^{\underline{\text{th}}}$ Problem I: Real Binomial Systems. In *International Symposium on Symbolic and Algebraic Computation (ISSAC '19), July 15–18, 2019, Beijing, China.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3326229.3326267

1 INTRODUCTION

Polynomial system solving has occupied a good portion of research in algebraic geometry for centuries, and inspired numerous algorithms in engineering and optimization. In recent years, *homotopy continuation* (see, e.g., [5, 33, 34, 38, 55]) has emerged as one of the most practical and efficient approaches to leverage high performance computing for the approximation of roots of large polynomial systems.

A refinement particularly useful for sparse systems is *polyhedral homotopy* [25, 30, 60]. To be brutally concise, polyhedral homotopy reduces the solution of an arbitrary polynomial system to (a) solving a finite collection of *binomial* systems to high precision and then (b) iterating a finite collection of rational functions. A complete, average-case complexity analysis of polyhedral homotopy thus implies an average-case complexity upper bound on solving binomial systems. (See also [18] for a more in-depth discussion on the importance of binomial systems.)

A geometric aspect common to both polyhedral homotopy and older homotopy methods is the deformation of a *start* system G (or a collection of start systems $\{G_i\}$), with known roots, into the system F one is trying to solve. Put another way, homotopy algorithms approximate the motion of the roots of a one-parameter family of polynomial systems (called a *homotopy path*), where the resulting path has end-points G and F. As studied in [6, 7, 9, 17, 29, 47, 48], average-case complexity analysis for classical homotopy algorithms hinges on careful probabilistic condition number estimates for the start system G, followed by further probabilistic condition number estimates for systems along the

homotopy path.[1] Since binomial systems play the role of start systems in polyhedral homotopy, it is thus crucial to know the probability that a binomial system is easy to solve. Our main theorem, on average-case complexity, is a step in this direction.

Since solving arbitrary polynomial systems is a numerical problem involving solutions of unknown minimal spacing, we will need to incorporate the cost of approximating well enough to distinguish distinct solutions. A recent and elegant way to handle this is via the notion of *approximate root in the sense of Smale*. In what follows, we use $|\cdot|$ for the standard ℓ_2-norm on $\mathbb{C}^n$.

DEFINITION 1.1. *[13, 50] Given any analytic function $F : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, we define the* Newton endomorphism of F *to be $N_F(z) := z - F'(z)^{-1}F(z)$, where we think of $F(z)$ as a column vector and we identify the derivative $F'(z)$ with the matrix of partial derivatives $\left[\frac{\partial f_i}{\partial x_j}\right]\Big|_{x=z}$. We call $\zeta \in \mathbb{C}^n$ a* non-degenerate *root of F if and only if $F'(\zeta)$ is invertible. Given $z_0 \in \mathbb{C}^n$, we then define its* sequence of Newton iterates $(z_n)_{n\in\mathbb{N}\cup\{0\}}$ *via the recurrence $z_{n+1} := N_F(z_n)$ (for all $n \geq 0$). We then call z_0* an approximate root of F in the sense of Smale (with associated true root ζ) *if and only if F has a non-degenerate root $\zeta \in \mathbb{C}^n$ satisfying $|z_n - \zeta| \leq \left(\frac{1}{2}\right)^{2^{n-1}} |z_0 - \zeta|$ for all $n \geq 1$.* $\diamond$

In essence, once one has an approximate root in the sense above, one can easily compute coordinates within any desired $\varepsilon > 0$ of the coordinates of a *true* root, simply by iterating Newton's method $O\left(\log\log\frac{1}{\varepsilon}\right)$ many times. The special case $F(z_1) := z_1^2 - 2$ already shows that one needs $\Omega\left(\log\log\frac{1}{\varepsilon}\right)$ arithmetic operations to compute $\sqrt{2}$ within ε [16]. So one can arguably consider an approximate root to be the gold standard for specifying a true root. In particular, one no longer has to worry about finding the minimal root spacing of F (to find the right ε to separate distinct roots), since an approximate root in the sense of Smale is guaranteed to converge almost optimally fast to a unique true root.

Of course, this begs the question of how one can possibly find an approximate root. This is the crux of Smale's $17^{\underline{\text{th}}}$ Problem (see [51, 52] and Section 1.1 below), which was recently positively solved by Lairez [29]. (See also the seminal work of Shub and Smale [48], Beltrán and Shub [9, 47], Beltrán and Pardo [6–8], and Bürgisser and Cucker [17].) Roughly, Lairez's discovery was an algorithm that, for a certain class of *random* polynomial systems, finds a single (complex) approximate root in polynomial-time on average. We now introduce some more terminology to be precise:

DEFINITION 1.2. *Suppose $\mathcal{A}_1, \ldots, \mathcal{A}_n \subset \mathbb{Z}^n$ are finite subsets and $\{c_{i,a} \mid i \in \{1, \ldots, n\}$ and $a \in \mathcal{A}_i$ for all $i\}$ is a collection of independent complex (resp. real) Gaussians with mean 0 and the variance of $c_{i,a}$ equal to $w_{i,a}^2$. Letting $a := (a_1, \ldots, a_n)$, $x^a := x_1^{a_1} \cdots x_n^{a_n}$, and $f_i(x) := \sum_{a\in\mathcal{A}_i} c_{i,a} x^a$, we call $F := (f_1, \ldots, f_n)$ an $n \times n$* complex (resp. real) random polynomial system with support $(\mathcal{A}_1, \ldots, \mathcal{A}_n)$. $\diamond$

LAIREZ'S THEOREM. *[29, Thm. 23][2] Following the notation above, let $d_1, \ldots, d_n \in \mathbb{N}$,*

$\mathcal{A}_i := \left\{(a_1, \ldots, a_n) \in (\mathbb{N} \cup \{0\})^n \;\middle|\; \sum_{j=1}^n a_j \leq d_i\right\}$ *for all i, and $w_{i,a}^2 := \frac{d_i!}{a_1!\cdots a_n!(d_i - \sum_{j=1}^n a_j)!}$. Then one can find a (complex) approximate root of a complex random F using just $O(nd^{3/2}N(N+n^3))$ arithmetic operations on average, where $N := \sum_{i=1}^n \frac{(d_i+n)!}{d_i!n!}$ and $d := \max_i d_i$.* ■

Note that restricting the support $(\mathcal{A}_1, \ldots, \mathcal{A}_n)$ is a way to consider *sparsity* for one's polynomial system. In particular, one can think of Lairez's Theorem as solving Smale's $17^{\underline{\text{th}}}$ Problem in the "dense" case, since Lairez assumes that *all* monomial terms up to a given degree appear (with probability 1) in each polynomial f_i. Indeed, one should note that Smale never specified what kind of probability measure one should use in his $17^{\underline{\text{th}}}$ Problem [51, 52]. So Smale's $17^{\underline{\text{th}}}$ Problem actually includes sparse systems if some of the random coefficients have mean, and all higher moments, equal to 0. Smale also observed in [51, 52] that one can pose a more difficult analogue of his $17^{\underline{\text{th}}}$ problem over the real numbers.

REMARK 1.3. *It is worth noting that the number of nonzero complex roots of a complex (or even real) random polynomial system as above attains a* unique *value with probability 1, once $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are fixed. (This follows easily from a classical result of Bernstein [10, 44], relating the mixed volume of Newton polytopes with counting complex roots.) Counting* real *roots for real random systems is more subtle however: One can easily show that, for any continuous positive probability measure on the coefficients, having at least one $d_i \geq 2$ (in the setting of Lairez's Theorem) implies that at least two different possible real root counts can occur with positive probabilities (see, e.g., [20]). For instance, if c_0, c_1, and c_2 are independent real Gaussians of any positive variance, the probability that $c_0 + c_1 x + c_2 x^2$ has exactly k real roots is positive for each $k \in \{0, 2\}$.* $\diamond$

Observe that $\sum_{i=1}^n \frac{(d_i+n)!}{d_i!n!}$ is exactly the maximal possible total number of monomial terms in an $n \times n$ polynomial system where f_i has degree d_i. Note also that just evaluating a monomial of degree d takes $\Omega(\log d)$ arithmetic operations: Simply consider the straight-line program complexity of the integer 2^d (see, e.g., [15, 36, 37]). One should pay attention to the evaluation complexity of F since Lairez's algorithm uses Newton iteration, which in turn requires evaluating F (and its Jacobian) many times. So one can then naturally ask, in the spirit of real fewnomial theory [26]: Can one find a real approximate root of F (or decide whether there are no real roots) using, say, $(t \log d)^{O(1)}$ arithmetic operations on average, when t is the total number of monomial terms of F and $d := \max_i d_i$? (See also [45] for an earlier statement of

[1] The condition number is a measure of the sensitivity of the roots of a system to perturbation of its coefficients. The condition number is thus another important measure of the complexity of numerical solving.

[2] We have paraphrased a bit: Lairez's main theorem is stated in terms of homogeneous polynomials, and he counts square roots as arithmetic operations as well. Via the techniques of, say, [7], one can easily derive our affine statement.

this problem.) This would be a significant new speed-up. For instance, the special case $t = O(n)$ is already quite non-trivial since there are standard algebraic tricks (e.g., the bottom of the first page of [21]) to reduce arbitrary polynomial systems to trinomial systems.

Our main theorem thus solves a special case of a refined version of Smale's $17^{\underline{\text{th}}}$ Problem, and serves as a starting point for a deeper study of the randomized complexity of solving arbitrary real sparse polynomial systems.

THEOREM 1.4. *Suppose $A = [a_{i,j}] \in \mathbb{Z}^{n\times n}$ and all the entries of A have absolute value at most d. Suppose also that $c_{i,j}$ is an independent real Gaussian with mean 0 and variance $w_{i,j}^2$, for each $(i,j) \in \{1,\ldots,n\} \times \{0,1\}$. Let $F := (f_1,\ldots,f_n)$ with $f_i(x) := c_{i,0} + c_{i,1} \cdot x_1^{a_{1,i}} \cdots x_n^{a_{n,i}}$, and set $r := \max_i \max\left\{\left|\frac{w_{i,0}}{w_{i,1}}\right|, \left|\frac{w_{i,1}}{w_{i,0}}\right|\right\}$. Then, on average, one can find a real approximate root of F (or correctly determine there are no real roots) using just $O\big(n^2 \log^2(nd)[1 + n \log\log(er)]\big)$ arithmetic operations and $O(n^{\omega+1}\log^2(dn))$ bit operations, where ω is any upper bound on the matrix multiplication exponent.*

We prove Theorem 1.4 in Section 3. The best current upper bound on ω, as of May 2019, is Legall's estimate 2.3728639 [31] (see also [2, 59]).

At a high level, the algorithm underlying our main theorem has three phases: (I) perform integer linear algebra on the exponent vectors to find a monomial change of variables reducing the input system F to the *diagonal* form $(x_1^{d_1} - c_1, \ldots, x_n^{d_n} - c_n)$, (II) decide if there are real roots, (III) if there are real roots, solve each binomial by a combination of bisection and Newton iteration (based on [61]), paying close attention to how each pair (d_i, c_i) affects the required accuracy. Phases (I) and (II) are well-known in the computational toric geometry community (see, e.g., [18, 21, 25]). Although [18] contains many useful algorithmic details on solving binomial systems, including a discussion of numerical implementations, the computational complexity of binomial system solving does not appear to have been analyzed yet from the point of view of average-case complexity or approximate roots in the sense of Smale. Our primary contribution is thus a new analysis of Phase (III), particularly with respect to average-case complexity in the Gaussian setting.

The complexity of Phase (I) accounts for the bit complexity estimate in Theorem 1.4, thanks to earlier work of Storjohann on fast linear algebra over $\mathbb{Z}$ (see [54, 56] and Section 2.1 below). Phase (II) is an elementary algebra exercise and actually has negligible (deterministic) complexity compared to our main bound. Step (III) is accomplished by a hybrid algorithm of Ye that allows quick approximation of rational powers of a real number [61]. The final key ingredient to establishing the average-case complexity of Phase (III) is estimating the expected value of linear combinations of logarithms of absolute values of standard real Gaussians (see Propositions 2.8 and 2.9 in Section 2.3 below). We were unable to find any explicit asymptotics for such expectations, so we derive these in the latter half of Section 2.3.

We will explain some of the subtleties behind extending Theorem 1.4 to systems with arbitrary supports in Section 1.2 below. First, however, let us briefly review the original statement of Smale's $17^{\underline{\text{th}}}$ Problem.

1.1 Quick Review of Smale's $17^{\underline{\text{th}}}$ Problem

Smale's $17^{\underline{\text{th}}}$ Problem [51, 52] elegantly summarizes the subtleties behind polynomial system solving:

> Can **a zero** of n complex polynomial equations in n unknowns be **found approximately**, **on the average**, in polynomial-time with a uniform algorithm?

[Emphases added.] We clarify the notion of "polynomial-time" below. As motivation, let us first see how the emphasized terms highlight fundamental difficulties in polynomial system solving:

"a zero": We can not expect a fast algorithm approximating *all* the roots since, for $n \geq 2$, there may be infinitely many. In which case, for $d_1 \geq 3$ (e.g., the case of elliptic curves [53]), the roots will likely not admit a rational parametrization. When there are only finitely many roots, systems like $(x_1^2 - 1, \ldots, x_n^2 - 1)$ show that the number of roots can be exponential in n.

"found approximately": Even restricting to integer coefficients, the number of digits of accuracy needed to separate distinct roots can be exponential in n, e.g., $((2x_1 - 1)(3x_1 - 1), x_2 - x_1^2, \ldots, x_n - x_{n-1}^2)$ has roots with $n^{\underline{\text{th}}}$ coordinates $\frac{1}{2^{2^{n-1}}}$ and $\frac{1}{3^{2^{n-1}}}$. So, especially for irrational coefficients, we need a more robust notion of approximation than digits of accuracy. (Hence's Smale's definition of approximate root from [50].)

"on the average": Restricting to integer coefficients, distinguishing between a system having finitely many or infinitely many roots is **NP**-hard (see, e.g., [27, 41]). Furthermore, as already long known in the numerical linear algebra community (e.g., results on the distribution of eigenvalues of random matrices [19, 57]), even if the number of roots is finite, the accuracy needed to separate distinct roots can vary wildly as a function of the coefficients. So averaging over all inputs allows us to amortize the complexity of potentially intractable instances.

The original statement of Smale's $17^{\underline{\text{th}}}$ Problem measures *time* (or *complexity*) as the total number of (a) (exact) field operations over $\mathbb{C}$, (b) comparisons over $\mathbb{R}$, and (c) bit operations [51]. (The underlying computational model is a *BSS machine over* $\mathbb{R}$ [13], which is essentially a classical *Turing* machine [3, 39, 49], augmented so that it can perform any field operation or comparison over $\mathbb{R}$ in one time step.) *Polynomial-time* was then meant as polynomial in the number of (nonzero) coefficients of F. Smale thus interpreted the

number of coefficients (which can be as high as $\sum_{i=1}^{n}\binom{d_i+n}{n}$ for F as specified above) as the *input size*.

REMARK 1.5. *The precise probability distribution over which one averages was never specified in Smale's original statement [51, 52]. In all the literature so far on the problem (see, e.g., [6–9, 17, 29, 47, 48]), the* Bombieri-Weyl measure *was used: This is the choice of variances involving multinomial coefficients written earlier.* ⋄

While the Bombieri-Weyl measure satisfies some very nice group invariance properties (see, e.g., [12, 23, 28, 48]), there is currently no widely-accepted notion of a "natural" probability distribution for a random polynomial. For instance, there are several different distributions of interest already for the matrix eigenvalue problem (see, e.g., [1, 19, 43]). More to the point, much work has gone into finding useful properties of the roots of random polynomials that are distribution independent (see, e.g., [11, 22, 58]).

The meaning of *uniform algorithm* is more technical and is formalized in [13] (see also [3, 39, 49] for the classical Turing case). Roughly, uniformity refers to having a single implementation that can handle all input sizes, as opposed to having different implementations for each input size.

1.2 Current Obstructions to Fully Incorporating Sparsity

As we'll see from the proof of our main theorem, solving an $n \times n$ system of Gaussian random binomials of degree d can be reduced to solving n univariate binomials of degree $(nd)^{O(n)}$, where the underlying coefficients are no longer Gaussian but have reasonably estimable means. Algebraically, this will imply that the underlying field extension (where one adjoins the coordinates of the solutions to the field generated by the coefficients) is always a radical extension.

A natural next step then is to consider $n \times n$ *unmixed* $(n+1)$*-nomial systems*:

$$(c_{1,0}+c_{1,1}x^{a_1}+\cdots+c_{1,n}x^{a_n},\ldots,c_{n,0}+c_{n,1}x^{a_1}+\cdots+c_{n,n}x^{a_n}),$$

where $a_i := (a_{1,i},\ldots,a_{n,i})$ for all i. Via Gauss-Jordan Elimination, one can reduce such a system to a binomial system without affecting the roots. Unfortunately, if one starts with a system of the form above, with Gaussian $c_{i,j}$, the resulting binomial system no longer has Gaussian coefficients. So one needs to consider binomial systems with coefficient distributions more general than Gaussian, and we do this in a sequel to this paper.

Going a bit farther, $n \times n$ unmixed $(n+2)$-nomial systems yield an interesting complication: The underlying field extensions need no longer be radical, even if $n = 1$. A simple example is $x_1^5 - 2x_1 + 10$, which has Galois group S_5 over $\mathbb{Q}$. However, earlier results from [45] indicate that it should be possible to find real approximate roots quickly on average, at least for univariate trinomials. (One should also observe Sagraloff's recent dramatic speed-ups for the worst-case arithmetic complexity of ε-approximating real roots of univariate sparse polynomials [46].) We conjecture that finding a real approximate root (or determining that there are no real roots) for a real Gaussian $n \times n$ unmixed $(n+2)$-nomial system is still possible in time $(n \log d)^{O(1)}$ on average, and hope to address this problem in the future. An interesting intermediate complication is that just counting the real roots within average-case time $(n \log d)^{O(1)}$ is already an open question for $t \geq n+3$ (see, e.g., [4]).

2 BACKGROUND

In what follows, for any $n \times n$ matrix $A \in \mathbb{Z}^{n\times n}$, we define x^A to be the vector of monomials

$$\left(x_1^{a_{1,1}}\cdots x_n^{a_{n,1}},\ldots,x_1^{a_{1,n}}\cdots x_n^{a_{n,n}}\right).$$

We call the substitution $x = z^A$ a *monomial change of variables*. The following proposition is elementary.

PROPOSITION 2.1. *We have that $x^{AB} = (x^A)^B$ for any $A, B \in \mathbb{Z}^{n\times n}$. Also, for any field K, the map defined by $m(x) = x^U$, for any unimodular matrix $U \in \mathbb{Z}^{n\times n}$, is an automorphism of $(K^*)^n := (K \setminus \{0\})^n$.* ■

Our main approach to solving binomial systems is to reduce them to systems of the form $(x_1^{d_1} - c_1,\ldots,x_n^{d_n} - c_n)$ via a monomial change of variables, and then prove that the distortion of the c_i resulting from perturbing the original coefficients is controllable. Later on, we will also detail how a Gaussian distribution on the original coefficients implies that the c_i still have well-behaved distributions. But now we will focus on quantifying our monomial changes of variables.

2.1 Linear Algebra Over $\mathbb{Z}$

DEFINITION 2.2. *Let $\mathbb{GL}_n(\mathbb{Z})$ denote the set of all matrices in $\mathbb{Z}^{n\times n}$ with determinant ± 1 (the set of* unimodular *matrices). Given any $M \in \mathbb{Z}^{n\times n}$, we call any identity of the form $UMV = S$ with $U, V \in \mathbb{GL}_n(\mathbb{Z})$ and S diagonal a* Smith factorization. *In particular, if $S = [s_{i,j}]$ and we require additionally that $s_{i,i} \geq 0$ and $s_{i,i} | s_{i+1,i+1}$ for all $i \in \{1,\ldots,n\}$ (setting $s_{n+1,n+1} := 0$), then S is uniquely determined and is called* the Smith normal form *of M.* ⋄

REMARK 2.3. *Although the Smith normal form is unique, the Smith* factorization *certainly need not be unique. For instance, $\begin{bmatrix}0&0\\0&0\end{bmatrix} = \begin{bmatrix}1&u\\0&1\end{bmatrix}\begin{bmatrix}0&0\\0&0\end{bmatrix}\begin{bmatrix}1&v\\0&1\end{bmatrix}$ for all $u, v \in \mathbb{Z}$. Note, however, that this need not contradict there being* some *factorization with small entries.* ⋄

THEOREM 2.4. *[56, Ch. 6 & 8, pg. 128] For any $A = [a_{i,j}] \in \mathbb{Z}^{n\times n}$, a Smith factorization of A yielding the Smith normal form of A can be computed within*

$$O\left(n^{\omega+1}\log^2(n\max_{i,j}|a_{i,j}|)\right)$$

bit operations. Furthermore, the entries of all matrices in this factorization have bit size $O(n\log(n\max_{i,j}|a_{i,j}|))$. ■

2.2 From Approximate Roots of Univariate Binomials to Systems

We begin with an important observation from the middle author's doctoral dissertation, building upon earlier work of Smale [50] and Ye [61].

LEMMA 2.5. *[40, Thm. 4.10] Let $d\in\mathbb{N}$ satisfy $d\geq 2$, $c>0$, and $f(x_1):=x_1^d-c$. Then we can find an approximate root of f using $O\big((\log d)(\log\log(de\max\{c,c^{-1}\}))\big)$ field operations over $\mathbb{R}$.* ■

Since a monomial change of variables enables us to replace an arbitrary binomial system by a simpler, *diagonal* system of univariate binomials, it's enough to bound how the roots are distorted under such a change of variables. The following lemma gives us the bounds we need.

LEMMA 2.6. *Suppose $c_1,\ldots,c_n\in\mathbb{C}^*$ and $A\in\mathbb{Z}^{n\times n}$ has columns $a_1,\ldots,a_n$ and entries of absolute value at most d. Also let $\sigma:=\max_i\{|\log|c_i||\}$, let $UAV=S$ be the Smith Factorization of A, and let $(\gamma_1,\ldots,\gamma_n):=(c_1,\ldots,c_n)^V$. Then the following bounds hold:*

1. *$\max_i|\log|\gamma_i||\leq n^{4+3n/2}d^{3n}\sigma$.*
2. *If $\zeta=(\zeta_1,\ldots,\zeta_n)\in(\mathbb{C}^*)^n$ is a root of F then $\max_i|\log|\zeta_i||\leq n^{O(n)}d^{O(n)}\sigma$.* ■

Lemma 2.6 follows easily from the second bound of Theorem 2.4, upon observing that $x^A=c$ implies that $z^S=(\gamma_1,\ldots,\gamma_n)$ where $x=z^U$. By combining Lemma 2.6 with Theorem 2.4 and multivariate Taylor's Theorem with Remainder (see, e.g., [24]), we then easily obtain the following estimate:

PROPOSITION 2.7. *Following the notation above, let $\mathbb{R}^n_+$ denote the positive orthant and let $|(y_1,\ldots,y_n)|_\infty$ denote the ℓ_∞-norm $\max_i|y_i|$ of the vector $y=(y_1,\ldots,y_n)$. Suppose also that $\zeta,\mu,x,z\in\mathbb{R}^n_+$ satisfy $\mu^S=\gamma$, $\zeta=\mu^U$, and $x=z^U$. Then $\log|x-\zeta|_\infty=e^{O(n\log(dn))}\sigma+\log|z-\mu|_\infty$.* ■

2.3 A Key Probabilistic Estimate

Let Z be a standard real Gaussian random variable and let $Y:=\log|Z|$. It is not difficult to check that Y has density $\rho_Y(t):=\sqrt{\frac{2}{\pi}}e^{-v(t)}$, $-\infty<t<\infty$, where $v(t):=\frac{e^{2t}}{2}-t$. Indeed, this follows by differentiating the distribution function of Y, $F_Y(t):=\mathbb{P}(-e^t\leq|Z|\leq e^t)=1-2\Phi(-e^t)$. Note that v is a convex function. Let $\alpha:=\mathbb{E}[Y]$ and let τ be the standard deviation of Y. ($\alpha\approx-0.635181...$ and $\tau\approx 1.110720...$, according to the 2018 version of `Maple`.) Consider the centered random variable $W:=Y-\alpha$. Let $a:=(a_1,\ldots,a_k)\in\mathbb{R}^k$, and let $W_a:=a_1W_1+\cdots+a_kW_k$ where W_i are independent copies of W. Let $X_a:=\max\{e^{W_a},e^{-W_a}\}$. We then have the following:

PROPOSITION 2.8. *Let $a=(a_1,\ldots,a_k)\in\mathbb{R}^k$ and assume that $\sum_{i=1}^k a_i=0$. Then W_a is a log-concave random variable with expectation 0 and standard deviation $\gamma:=|a|\tau$. We also have*

$$\mathbb{P}(\log\log(eX_a)\geq t)\leq e^{-\frac{e^t-1}{2\gamma}}\ \text{for } t\geq\log(1+\gamma). \tag{1}$$

Moreover,

$$\mathbb{E}[\log\log(eX_a)]\leq 2+\log(1+\gamma). \tag{2}$$

Proof: Since v is a convex function the density ρ_Y is log-concave and, by a theorem of Borell [14], the law of the random variable Y is log-concave, i.e., for all compact sets A,B and $\lambda\in(0,1)$ one has

$$\mu(\lambda A+(1-\lambda)B)\geq\mu(A)^\lambda\mu(B)^{1-\lambda}, \tag{3}$$

where μ is the measure on $\mathbb{R}$ induced by the density ρ_Y. Also, W is a log-concave random variable and, by the Prékopa-Leindler inequality [32, 42], W_a is also log-concave. We have that $\mathbb{E}[W_a]=\sum_{i=1}^k a_i\mathbb{E}[Y_i]=\alpha\sum_{i=1}^k a_i=0$ and, since the W_i are independent,

$$\mathrm{var}(W_a)=\sum_{i=1}^k a_i^2\mathrm{var}(Y_i)=\tau^2\sum_{i=1}^k a_i^2=\tau^2a^2.$$

Another well-known result of Borell (see e.g., [35]) then states that if μ is a *log*-concave probability measure, K is a symmetric closed convex set in $\mathbb{R}^n$, and $\delta:=\mu(K)\geq\frac{1}{2}$, then for all $t>1$ we have the following:

$$1-\mu(tA)\leq\delta\left(\frac{1-\delta}{\delta}\right)^{\frac{t+1}{2}}. \tag{4}$$

In particular, if X is a *log*-concave random variable with mean 0 and variance γ^2, then we have the following:

$$\mathbb{P}(|X|\geq s)\leq e^{-\frac{s}{2\gamma}},\ \text{for } s\geq\gamma. \tag{5}$$

Indeed, let $A:=\{|x|\leq 2\gamma\}$. Then, by Chebychev's Inequality, we have that $\mathbb{P}(A)=\delta\geq\frac{3}{4}$. Using (4), we obtain:

$$\mathbb{P}(|X|\geq t\gamma)=1-\mathbb{P}(tA)\leq\delta\left(\frac{1-\delta}{\delta}\right)^{\frac{t+1}{2}}\leq\left(\frac{1}{3}\right)^{\frac{t+1}{2}}\leq e^{-\frac{t}{2}}, \tag{6}$$

for $t\geq 1$. So we can estimate as follows:

$$\mathbb{P}(\log\log(eX_a)\geq t)=\mathbb{P}\Big(X_a\geq e^{e^t-1}\Big)$$

$$=\mathbb{P}\Big(\Big\{V_a\geq e^{e^t-1}\Big\}\cup\Big\{V_a\leq e^{-(e^t-1)}\Big\}\Big)$$

$$=\mathbb{P}\Big(V_a\geq e^{e^t-1}\Big)+\mathbb{P}\Big(V_a\leq e^{-(e^t-1)}\Big)$$

$$=\mathbb{P}\big(W_a\geq e^t-1\big)+\mathbb{P}\big(W_a\leq-(e^t-1)\big)$$

$$=\mathbb{P}\big(|W_a|\geq e^t-1\big)\leq e^{-\frac{e^t-1}{2\gamma}},$$

provided $e^t-1\geq\gamma$, where we have also used (6). Finally, since $eX_a\geq e$, we have $\log\log(eX_a)\geq 0$ and thus

$$\mathbb{E}[\log\log(eX_a)]\leq\int_0^\infty\mathbb{P}(\log\log(eX_a)\geq t)\,dt$$

$$\leq\int_0^{\log(1+\gamma)}dt+\int_{\log(1+\gamma)}^\infty e^{-\frac{e^t-1}{2\gamma}}dt$$

$$\leq\log(1+\gamma)+\int_\gamma^\infty\frac{1}{1+s}e^{-\frac{s}{2\gamma}}ds$$

$$=\log(1+\gamma)+\int_{\frac{1}{2}}^\infty\frac{2\gamma}{1+2\gamma x}e^{-x}dx$$

$$\leq\log(1+\gamma)+\frac{2\gamma}{1+\gamma}\int_0^\infty e^{-x}dx\leq 2+\log(1+\gamma).$$ ■

PROPOSITION 2.9. *Let $a \in \mathbb{R}^k$ satisfy $\sum_{i=1}^k a_i = 0$ and assume $t \geq e^2 \approx 7.3890....$ Then*

$$\begin{aligned} \log\log t &\leq \mathbb{E}\log\log\{tX_a\} \\ &\leq \log\log(t/e) + 2 + \log 2 + \log(1+\tau|a|), \end{aligned}$$

where $\tau \approx 1.110720...$ is the standard deviation of a random variable of the form $\log|Z|$, where Z is a standard real Gaussian random variable.

Proof: Note that $a+b \leq 2ab$ for all $a, b \geq 1$. Since $eX_a \geq e$ and $t/e \geq e$, using (2) we get

$$\begin{aligned} \mathbb{E}\log\log(tX_a) &= \mathbb{E}\log(\log(t/e) + \log(eX_a)) \\ &\leq \mathbb{E}\log(2(\log(t/e))\log(eX_a)) \\ &= \log(2) + \log\log(t/e) + E\log\log(eX_a) \\ &\leq \log(2) + \log\log(t/e) + 2 + \log(1+\tau|a|). \end{aligned}$$

Finally, since $X_a \geq 1$ and $t \geq e^2$, we have $\log\log(tX_a) \geq \log\log t$. ■

3 THE PROOF OF THEOREM 1.4

First note that the $c_{i,j}$ are all nonzero with probability 1, so we may assume (since we are considering average-case complexity) that all the $c_{i,j}$ are nonzero. In which case, we can focus solely on roots in $(\mathbb{R}^*)^n$.

Now note that by Proposition 2.1, we can easily decide whether our input binomial system F has a real root: If F is diagonal, i.e., if $F = (c_{1,0} + c_{1,1}x_1^{d_1}, \ldots, c_{n,0} + c_{n,1}x_n^{d_n})$ for some $d_1, \ldots, d_n \in \mathbb{N}$, then F has a real root if and only if the following condition holds: $c_{i,0}c_{i,1} < 0$ for all i with d_i even and nonzero, and $c_{i,0} = -c_{i,1}$ for all i with $d_i = 0$. Should this condition be true, each orthant of $\mathbb{R}^n$ contains at most 1 root of F (if all the d_i are nonzero), or F has infinitely many roots in any orthant where F vanishes (if some d_i is zero). (See [21] or [18, Sec. 3] for a more precise description of the case where F has infinitely many roots in $(\mathbb{R}^*)^n$.) In the latter case, F has free variables that we may set to 1, yielding a $j \times j$ binomial system with $j < n$ and real roots that are coordinate projections of the roots of F.

If F is not diagonal, then after computing a Smith factorization $UAV = S$ (which accounts for our stated bit complexity bound, thanks to Theorem 2.4), we can reduce to the diagonal case and simply check n inequalities and equalities. If there are no real roots, no further work needs to be done.

So let us now assume that there are real roots. Without loss of generality (flipping signs of certain $c_{i,j}$ as needed), we may assume there is a root in the positive orthant $\mathbb{R}^n_+$, and try to approximate a root there. So we may now assume that we are trying to approximate the roots of

$$G := (z_1^{s_{1,1}} - \gamma_1, \ldots, z_n^{s_{n,n}} - \gamma_n)$$

where

$$(\gamma_1, \ldots, \gamma_n) := (-c_{1,0}/c_{1,1}, \ldots, -c_{n,0}/c_{n,1})^V$$

lies in $\mathbb{R}^n_+$, and the $s_{i,i}$ are the diagonal entries of the Smith normal form S of A. In particular, we need to approximate a root μ of G in $\mathbb{R}^n_+$ closely enough so that $\zeta := \mu^U$ is an approximate root of F.

A slight complication arises: Some of the $s_{i,i}$ may be 0, thus making the Jacobian of G have rank too low for Newton iteration to be well-defined. However, this is easily dispensed with by setting $z_i = 1$ for all i with $s_{i,i} = 0$. This has the effect of reducing our problem to solving the $j \times j$ system $G' := (z_1^{s_{1,1}} - \gamma_1, \ldots, z_j^{s_{j,j}} - \gamma_j)$, where $j \leq n$ and $m(x) := (x_1^{s_{1,1}}, \ldots, x_j^{s_{j,j}})$ is a surjective endomorphism on $(\mathbb{C}^*)^j$. So we *can* ultimately obtain approximate roots, simply by applying Newton iteration to G' instead of G. Thus, let us assume without loss of generality that all the $s_{i,i}$ are non-zero (and thus $\det A \neq 0$).

Proposition 2.7 then tells us that to find an approximate root of F, it suffices to find an approximate root of G, but with tighter precision. In particular, the necessary number of additional Newton iterations is $O(n\log(dn))$, and each Newton iteration for G requires $O(n\log d)$ arithmetic operations. So the additional work is bounded from above by our main arithmetic complexity bound. Lemma 2.5 applied to G then implies that to derive our average-case complexity bound, it suffices to compute an upper bound on the expectation of the following quantity:

$$B := \textstyle\sum_{i=1}^n \left[(\log s_{i,i}) \log\log\left(s_{i,i} e \max\left\{|\gamma_i|, \left|\gamma_i^{-1}\right|\right\}\right)\right].$$

We are almost ready to apply our probabilistic estimate Proposition 2.9, save for the fact that the γ_i are monomials in real Gaussians that need *not* have variance 1. However, from the definition of γ, we see that we in fact have $(\gamma_1, \ldots, \gamma_n) := \left(\frac{w_{1,0}}{w_{1,1}}, \ldots, \frac{w_{n,0}}{w_{n,1}}\right)^V \odot \left(\frac{-c'_{1,0}}{c'_{1,1}}, \ldots, \frac{-c'_{n,0}}{c'_{n,1}}\right)^V$, where $\odot$ denotes the natural coordinate-wise multiplication in $(\mathbb{R}^*)^n$, and the $c'_{i,j}$ are real Gaussians with mean 0 and variance 1. Using the inequality $a + b \leq 2ab$ for $a, b \geq 1$, we then see that it is enough to estimate the expectation of B in the special case where all the $c_{i,j}$ have variance 1, provided we *also* add the quantity

$$T := \textstyle\sum_{i=1}^n (\log s_{i,i}) \log\log\left(e \max\left\{w'_i, w'^{-1}_i\right\}\right)$$

to our estimate, where $(w'_1, \ldots, w'_n) := \left(\frac{w_{1,0}}{w_{1,1}}, \ldots, \frac{w_{n,0}}{w_{n,1}}\right)^V$.

We now conclude via Proposition 2.9 and Theorem 2.4: Proposition 2.9 tells us that the expectation of B is no greater than

$$\textstyle\sum_{i=1}^n (\log s_{i,i}) \left[\log\log\left(\max\{s_{i,i}, e\}\right) + 2 + \log(2) + \log\left(1 + \tau|v_i|\right)\right],$$

where v_i is the $i^{\underline{\text{th}}}$ column of V. Theorem 2.4, and the fact that $\sum_{i=1}^n \log|s_{i,i}| = \log|\det A| = O(n\log(dn))$ (thanks to Hadamard's classical inequality on the determinant), imply that the last quantity is no greater than

$$O\left(n\log(dn) \textstyle\sum_{i=1}^n \left(\log(n\log(dn)) + \log\left(1 + \tau\sqrt{n}e^{O(n\log(dn))}\right)\right)\right)$$

$= O(n\log(dn)n\log(n\log(dn)))$. So we obtain that the expectation of B is $O\left(n^2\log^2(dn)\right)$.

Similarly, by Theorem 2.4 and Lemma 2.6, T is no greater than

$$\textstyle\sum_{i=1}^n (\log|s_{i,i}|) \log\log\left(e\left(r^{e^{O(n\log(dn))}}\right)^n\right).$$

So $T = O(n\log(dn))n\log\left(ne^{O(n\log(dn))}\log(er)\right)$

$= O(n\log(dn))n\left[O(n\log(dn)) + \log\log(er)\right]$

$= O\big(n^3 \log^2(dn) \log\log(er)\big)$,
and we are done. ■

ACKNOWLEDGEMENTS

We humbly thank REU students Caleb Bugg and Paula Burkhardt for important discussions on preliminary versions of this work. We also thank the referees for their insightful comments that helped improve this paper. Finally, we thank the NSF for their support through grants CCF-1409020, DMS-1460766, DMS-1757872, DMS-1812240, and CCF-1900881.

REFERENCES

[1] Gernot Akemann; Jinho Baik; Philippe Di Francesco; *The Oxford Handbook of Random Matrix Theory,* Oxford University Press, 2011.

[2] Josh Alman, *"Limits on the Universal Method for Matrix Multiplication,"* proceedings of the 34th Computational Complexity Conference (CCC 2019, July 18-20 in New Brunswick, NJ, USA), to appear. Also available as Math ArXiV preprint 1812.08731 .

[3] Sanjeev Arora and Boaz Barak, *Computational complexity. A modern approach.* Cambridge University Press, Cambridge, 2009.

[4] Osbert Bastani; Chris Hillar, Dimitar Popov, and J. Maurice Rojas, *"Randomization, Sums of Squares, and Faster Real Root Counting for Tetranomials and Beyond,"* Randomization, Relaxation, and Complexity in Polynomial Equation Solving, Contemporary Mathematics, vol. 556, pp. 145–166, AMS Press, 2011.

[5] Daniel J. Bates; Jonathan D. Hauenstein; Andrew J. Sommese; and Charles W. Wampler, *Numerically Solving Polynomial Systems with Bertini,* Software, Environments and Tools series, Society for Industrial and Applied Mathematics, 2013.

[6] Carlos Beltrán and Luis M. Pardo, *"On Smale's 17th Problem: A Probabilistic Positive answer,"* Foundations of Computational Mathematics 8 (1): pp. 1–43.

[7] Carlos Beltrán and Luis M. Pardo, *"Smale's 17th Problem: Average Polynomial Time to compute affine and projective solutions,"* Journal of the American Mathematical Society, 22 (2009), pp. 363–385.

[8] Carlos Beltrán and Luis M. Pardo, *"Efficient Polynomial System Solving by Numerical Methods,"* in Randomization, Relaxation, and Complexity in Polynomial Equation Solving, Contemporary Mathematics, vol. 556, pp. 37–60, AMS Press, 2011.

[9] Carlos Beltrán and Mike Shub, *"Complexity of Bezouts Theorem VII: Distance Estimates in the Condition Metric,"* Foundations of Computational Mathematics, April 2009, Volume 9, Issue 2, pp. 179–195.

[10] David Naumovich Bernstein, *"The Number of Roots of a System of Equations,"* Functional Analysis and its Applications (translated from Russian), Vol. 9, No. 2, (1975), pp. 183–185.

[11] A. T. Bharucha-Reid and M. Sambandham, *Random polynomials,* Academic Press, Orland, 1986.

[12] Pavel Bleher; Bernard Shiffman; and Steve Zelditch; *"Poincare-Lelong approach to universality and scaling of correlations between zeros,"* Commun. Math. Phys. 208 (2000), pp. 771–785.

[13] Lenore Blum; Felipe Cucker; Mike Shub; and Steve Smale; *Complexity and Real Computation,* Springer-Verlag, 1998.

[14] C. Borell *Convex set functions in d-space*, Period. Math. Hungar. 6 (2), pp. 111–136 (1975).

[15] Alfred Brauer, *"On addition chains,"* Bull. Amer. Math. Soc. 45, (1939), pp. 736–739.

[16] Nader H. Bshouty; Yishay Mansour; Baruch Schieber; and Prasoon Tiwari; *"A tight bound for approximating the square root,"* Inform. Process. Lett. 63 (1997), no. 4, pp. 211–213.

[17] Peter Bürgisser and Felipe Cucker, *" On a problem posed by Steve Smale,"* Annals of Mathematics, Vol. 174 (2011), Issue 3, pp. 1785–1836.

[18] Tianran Chen and Tien-Yien Li, *"Solutions to Systems of Binomial Equations,"* Annales Mathematicae Silesianae **28** (2014), pp. 7–34.

[19] Alan Edelman, *"Eigenvalues and condition numbers of random matrices,"* SIAM Journal on Matrix Analysis and Applications 9 (1988), pp. 543–560.

[20] Alan Edelman and Eric Kostlan, *"How Many Zeros of a Random Polynomial are Real?,"* Bull. Amer. Math. Soc., 32, January (1995), pp. 1–37.

[21] David Eisenbud and Bernd Sturmfels, *"Binomial Ideals,"* Duke Mathematical Journal, Vol. 84, No. 1, July 1996.

[22] Alperen Ergür, Grigoris Paouris, and J. Maurice Rojas, *"Probabilistic Condition Number Estimates for Real Polynomial Systems I: A Broader Family of Distributions,"* Foundations of Computational Mathematics, Feb. 2019, Vol. 19, No. 1, pp. 131–157.

[23] Yan V. Fyodorov, Antonio Lerario, and Erik Lundberg, *"On the Number of Connected Components of Random Algebraic Hypersurfaces,"* J. Geom. Phys. (2015), pp. 1–20. DOI: 10.1016/j.geomphys.2015.04.006

[24] John Hubbard and Barbara Burke Hubbard, *Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach,* 5th edition, Matrix Editions, 2015.

[25] Birk Huber and Bernd Sturmfels, *"A polyhedral method for solving sparse polynomial systems,"* Math. Comp. **64** (1995), pp. 1541–1555.

[26] Askold G. Khovanskii, *Fewnomials,* AMS Press, Providence, Rhode Island, 1991.

[27] Pascal Koiran, *"Randomized and Deterministic Algorithms for the Dimension of Algebraic Varieties,"* Proceedings of the $38^{\underline{\text{th}}}$ Annual IEEE Computer Society Conference on Foundations of Computer Science (FOCS), Oct. 20–22, 1997, ACM Press.

[28] Eric Kostlan, *"On the distribution of roots of random polynomials,"* From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990), pp. 419–431, Springer, New York, 1993.

[29] Pierre Lairez, *"A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time,"* Foundations of computational mathematics 17.5 (2017), pp. 1265–1292.

[30] Tsung-Lin Lee and Tien-Yien Li, *"Mixed volume computation in solving polynomial systems,"* in Randomization, Relaxation, and Complexity in Polynomial Equation Solving, Contemporary Mathematics, vol. 556, pp. 97–112, AMS Press, 2011.

[31] Francois Legall, *"Powers of Tensors and Fast Matrix Multiplication,"* Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation (ISSAC 2014), pp. 296–303, 2014.

[32] L. Leindler, *"On a certain converse of Hölder's inequality II,"* Acta Sci. Math. (Szeged) 33 (1972), pp. 217–223.

[33] Tien-Yien Li, *"Numerical solution of multivariate polynomial systems by homotopy continuation methods,"* Acta numerica, 1997, pp. 399–436, Acta Numer., 6, Cambridge Univ. Press, Cambridge, 1997.

[34] Tien-Yien Li and Xiaoshen Wang, *"Solving deficient polynomial systems with homotopies which keep the subschemes at infinity invariant,"* Math. Comp. 56 (1991), no. 194, pp. 693–710.

[35] Vitaly D. Milman and Gideon Schechtman, *"Asymptotic Theory of Finite Dimensional Normed Spaces,"* Springer (Lecture Notes in Mathematics 1200), Berlin, 1986.

[36] Wellington de Melo and Benar Fux Svaiter, *"The cost of computing integers,"* Proc. Amer. Math. Soc. **124** (1996), pp. 1377–1378.

[37] Gustavo T. de Araujo Moreira, *"On asymptotic estimates for arithmetic cost functions,"* Proccedings of the American Mathematical Society, Vol. 125, no. 2, Feb. 1997, pp. 347–353.

[38] Alexander Morgan and Andrew Sommese, *"A homotopy for solving general polynomial systems that respects m-homogeneous structures,"* Appl. Math. Comput. 24 (1987), no. 2, pp. 101–113.

[39] Christos H. Papadimitriou, *Computational Complexity,* Addison-Wesley, 1995.

[40] Kaitlyn Phillipson, *Quantitative Aspects of Sums of Squares and Sparse Polynomial Systems,* doctoral dissertation, Texas A&M University, department of mathematics, 2016.

[41] David A. Plaisted, *"New NP-Hard and NP-Complete Polynomial and Integer Divisibility Problems,"* Theoret. Comput. Sci. 31 (1984), no. 1–2, pp. 125–138.

[42] A. Pr'ekopa, *"On logarithmic concave measures and functions,"* Acta Sci. Math. (Szeged) 34 (1975), pp. 335–343.

[43] J. Maurice Rojas, *"On the Average Number of Real Roots of Certain Random Sparse Polynomial Systems,"* in *The Mathematics of Numerical Analysis*, Lectures in Applied Mathematics, vol. 32, (Jim Renegar, Mike Shub, and Steve Smale eds.), pp. 689–699, American Mathematical Society, 1996.

[44] J. Maurice Rojas, *"Why Polyhedra Matter in Non-Linear Equation Solving,"* in: Contemporary Mathematics, vol. 334, pp. 293–320, American Mathematical Society, 2003.

[45] J. Maurice Rojas and Yinyu Ye, *"On Solving Sparse Polynomials in Logarithmic Time,"* Journal of Complexity, special issue for the 2002 Foundations of Computation Mathematics (FOCM) meeting, February 2005, pp. 87–110.

[46] Michael Sagraloff, *"A near-optimal algorithm for computing real roots of sparse polynomials,"* in Proc. ISSAC 2014 (39th International Symposium on Symbolic and Algebraic Computation), pp. 359–366, ACM Press, 2014.

[47] Mike Shub, *"Complexity of Bezouts Theorem VI: Geodesics in the Condition (Number) Metric,"* Foundations of Computational Mathematics 9(2):171–178, January 2009.

[48] Mike Shub and Steve Smale, *"The Complexity of Bezout's Theorem II: Volumes and Probabilities,"* Computational Algebraic Geometry (F. Eyssette and A. Galligo, Eds.), pp. 267–285, Birkhauser, 1992.

[49] Michael Sipser, *Introduction to the Theory of Computation,* 3rd edition, Cengage Learning, 2012.

[50] Steve Smale, *"Newton's Method Estimates from Data at One Point,"* The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics (Laramie, Wyo., 1985), pp. 185–196, Springer, New York, 1986.

[51] Steve Smale, *"Mathematical Problems for the Next Century,"* Math. Intelligencer 20 (1998), no. 2, pp. 7–15.

[52] Steve Smale, *"Mathematical Problems for the Next Century,"* Mathematics: Frontiers and Perspectives, pp. 271–294, Amer. Math. Soc., Providence, RI, 2000.

[53] Joseph H. Silverman and John Tate, *Rational Points on Elliptic Curves,* Undergraduate Texts in Mathematics, Springer, corrected edition, 1994.

[54] H. J. S. Smith, *"On Systems of Integer Equations and Congruences,"* Philos. Trans. 151, pp. 293–326 (1861).

[55] Andrew J. Sommese and Charles W. Wampler, *"The Numerical Solution to Systems of Polynomials Arising in Engineering and Science,"* World Scientific, Singapore, 2005.

[56] Arne Storjohann, *"Algorithms for matrix canonical forms,"* doctoral dissertation, Swiss Federal Institute of Technology, Zurich, 2000.

[57] Terence Tao and Van Vu, *"From the Littlewood-Offord Problem to the Circular Law: Universality of the Spectral Distribution of Random Matrices,"* Bulletin of the American Mathematical Society, vol. 46, no. 3, July 2009, pp. 377–396.

[58] Terence Tao and Van Vu, *"Local Universality of Zeroes of Random Polynomials,"* IMRN, Vol. 2015, Issue 13, 2015, pp. 5053–5139, DOI: 10.1093/imrn/rnu084

[59] Virginia Vassilevska Williams, *"Multiplying matrices in* $O(n^{2.373})$ *time,"* submitted for publication (earlier version in proceedings of STOC 2012 (ACM Symposium on Theory of Computing, May 19–22, NYU), ACM Press), 2014.

[60] Jan Verschelde, *"Polynomial Homotopy Continuation with PHCpack,"* ACM Communications in Computer Algebra 44(4):217-220, 2010.

[61] Yinyu Ye, *"Combining Binary Search and Newton's Method to Compute Real Roots for a Class of Real Functions,"* J. Complexity 10 (1994), no. 3, 271–280.