Lectures in Applied Mathematics
Volume **00**, 19xx

On the Average Number of Real Roots of Certain Random Sparse Polynomial Systems[1]

J. Maurice Rojas

This paper is dedicated to Sueli Rocha do Nascimento.

ABSTRACT. We derive an explicit formula for the expected number of real roots of certain random sparse polynomial systems. We propose (and use) a probability measure which is natural in the sense that it is invariant under a natural G-action on the space of roots, where G is a product of orthogonal groups. Our formula arose from an effort (now an ongoing project with J.-P. Dedieu) to generalize the recent condition number theorems of Shub and Smale to sparse polynomial systems and compactifications other than projective space.

1. Introduction

One may wonder if the recent advances in toric variety techniques, e.g., convex geometric formulae for the number of complex roots of a polynomial system [**Ber75, LW96, RW96, HS97, Roj96**], can be used to refine Shub and Smale's recent work [**SS2**] on the number of real roots of a random polynomial system. We answer this question in the affirmative for a particular family of sparse polynomial systems and formulate a conjectural answer for more general sparse systems. Our results are formulated in terms of *one* specific probability distribution, but nevertheless present a surprisingly unexplored aspect of random polynomial systems. Also, since real root counting is currently much harder than complex root counting, stochastic real root counts for sparse polynomial systems should be of interest to computational algebraic geometers as well as probabilists.

We can pose our general scenario as follows: Let $F = (f_1, \ldots, f_n)$ be a polynomial system with support $E := (E_1, \ldots, E_n)$ and coefficients which are random variables. This terminology simply means that we identify the monomial

[1] Slightly updated from version which appeared in Lectures in Applied Mathematics, edited by Jim Renegar, Mike Shub, and Steve Smale, American Mathematical Society, 1996.

1991 *Mathematics Subject Classification.* Primary 14M25, 60H25; Secondary 05A10, 12D10, 12Y05, 14N10, 14P99, 14Q99, 15A52, 52A39, 52B20, 52B55, 60D05, 65F50, 65H10, 90D15, 93B27, 93B55.

Key words and phrases. random, equations, polynomial, real, roots, Newton, polyhedra, polytopes, multihomogenous, multinomial, expected, average, eigenvalue, Gaussian.

This research was completed at MSRI and was funded by an NSF Postdoctoral Fellowship.

0075-8485/00 $1.00 + $.25 per page

$x^e := x_1^{e_1} \cdots x_n^{e_n}$ with the point $e := (e_1, \ldots, e_n) \in \mathbb{Z}^n$ and thus let the nonempty finite set $E_i \subset \mathbb{Z}^n$ determine precisely which monomial terms appear in f_i for all i. We call such an F an $n \times n$ *randomized polynomial system with support* E. We will also call E_i the *support* of f_i and the convex hull $\mathrm{Conv}(E_i)$ in $\mathbb{R}^n$ the *Newton polytope* of f_i.

Remark 1. Our notion of sparsity is to specify supports. This idea is natural in the sense that if one specifies the support of a polynomial then one can omit as many monomial terms as one likes. It is also important to note that our approach complements that of Khovanskii [**Kho91**], who has found amazing results and conjectures in terms of solely the *number* of monomial terms in a polynomial.

Remark 2. Working with Newton polytopes is already more precise than working with total degree. For example, an n-variate polynomial has total degree $\leq d$ iff its Newton polytope is contained in $\mathrm{Conv}\{\mathbf{O}, d\hat{e}_1, \ldots, d\hat{e}_n\}$ (the standard n-simplex in $\mathbb{R}^n$ uniformly scaled by a factor of d).

For reasons which will gradually become clearer, it is more convenient to work in terms of Newton polytopes and lattices (subgroups of $\mathbb{Z}^n$) instead of directly with E. Thus let L be a sublattice of $\mathbb{Z}^n$ of finite index. We then say that E is L-*complete* iff for all $i \in [1..n]$ there is an $a_i \in \mathbb{Z}^n$ such that $a_i + E_i = \mathrm{Conv}(a_i + E_i) \cap L$. Another bit of discrete geometry we will need is the following: If $a \in \mathbb{Z}^n$ and Q is an $(n-1)$-dimensional polytope with vertices in $a+L$, then Q naturally determines a (parallel) $(n-1)$-dimensional sublattice M of L. One can then speak of a point $p \in a+L$ being a certain number of *lattice steps*, $d_L(p, Q)$, away from Q simply by considering the 1-dimensional factor lattice L/M. Thinking in these terms, let us henceforth fix the following probability distribution on the coefficients of F.

Definition 1. Let F be an $n \times n$ randomized polynomial system with L-complete support E, and let $\mathcal{C}_E$ denote the vector consisting of all the coefficients of F. Also, for all $i \in [1..n]$ and $e \in E_i$, let $c_{i,e}$ denote the coefficient of the x^e term of f_i and let this coefficient be a (real) Gaussian random variable with mean 0 and variance

$$\prod_{Q \text{ a facet of } \mathrm{Conv}(E_i)} \frac{1}{d_{L_i}(e, Q)!}$$

where L_i is the smallest sublattice of L containing a translate of E_i. This distribution is called the *(real) L-polyhedral distribution* for $\mathcal{C}_E$.

Example 1. Suppose $n = 2$, $L = \mathbb{Z}^2$, and E is the pair of sets

$$(\{(0,0),(1,0),(2,0),(0,1),(1,1),(2,1),(0,2),(1,2)\},\{(0,0),(0,1),(0,2),(1,2)\}).$$

Then, following the notation of the last definition, it is clear that E is L-complete and thus F must be a bivariate polynomial system of the following form:

$$\begin{aligned} f(x,y) &= \alpha_1 + \alpha_2 x + \alpha_3 x^2 + \alpha_4 y + \alpha_5 xy + \alpha_6 x^2 y + \alpha_7 y^2 + \alpha_8 xy^2 \\ g(x,y) &= \beta_1 + \beta_2 y + \beta_3 y^2 + \beta_4 xy^2 \end{aligned}$$

where the α_j's and β_j's are independent real Gaussian random variables with mean 0. (Note that the Newton polygons of f and g are respectively a pentagon and a triangle, and thus $L_1 = L_2 = L$.) Counting lattice steps then simply amounts to partitioning the lattice $\mathbb{Z}^2$ into "slices" parallel to a Newton polygon edge. Using any reasonable drawing of $E_1, E_2 \subset \mathbb{Z}^2$, it is then easy to check that the variances of $\alpha_1, \alpha_2, \alpha_5, \beta_1, \beta_2$ are respectively $\frac{1}{0!2!3!2!0!}, \frac{1}{0!1!2!2!1!}, \frac{1}{1!1!1!1!1!}, \frac{1}{0!2!0!}, \frac{1}{1!1!0!}$. We leave the computation of the remaining variances as an exercise.

Definition 2. Following the notation of the last definition, we say that F is *L-polyhedrally randomized* and let $\mathcal{N}(E)$ denote the expected number of roots of F in $(\mathbb{R}^*)^n$, where $\mathbb{R}^* := \mathbb{R}\setminus\{0\}$.

Remark 3. In general one really wants to count roots in $\mathbb{R}^n$, but restricting to $(\mathbb{R}^*)^n$ is technically convenient. Also, for many cases (e.g., remark 4) it is easily verified that the roots of F in $\mathbb{C}^n$ all avoid the coordinate hyperplanes with probability 1; so sometimes both counts are identical. Nevertheless, there are important subtleties in general and we will leave these for future investigation.

We will soon see that this unusual choice of probability distribution is natural and/or, at very least, convenient. A purely aesthetic reason is that for certain E one can actually derive convex geometric formulae for $\mathcal{N}(E)$. For example, call an r-simplex *L-fundamental* iff its vertices all lie in L and its Euclidean r-volume is minimal over all such (nondegenerate) r-simplices. We then say that a simplex is *L-uniform* iff it is a uniform integral scaling of an L-fundamental simplex. A *product of polytopes of dimensions* $n_1,\ldots,n_k$ is simply a polytope sum $Q = \sum_{j=1}^k Q_j$ such that $\dim Q = \sum_{j=1}^k n_j$ and $\dim Q_j = n_j$ for all j. Letting $\mathrm{Vol}_L(\cdot)$ denote the renormalization of Euclidean n-volume which evaluates to 1 for L-fundamental n-simplices, we have our first main theorem.

Main Theorem 1. Let F be an $n\times n$ L-polyhedrally randomized system with support E. Assume the Newton polytopes of F are integral translates of P, where P is a product of L-uniform simplices of positive dimensions $n_1,\ldots,n_k$. Then

$$\mathcal{N}(E) = \pi^{\frac{k-1}{2}}\Gamma\left(\frac{k+1}{2}\right)\frac{\left(\begin{array}{c}\frac{n-1}{2}\\ \frac{n_1-1}{2},\ldots,\frac{n_k-1}{2},\frac{k-1}{2}\end{array}\right)}{\left(\begin{array}{c}n\\ n_1,\ldots,n_k\end{array}\right)^{\frac{1}{2}}}\cdot\sqrt{\mathrm{Vol}_L(P)}$$

We call the special case where $L=\mathbb{Z}^n$, all translations are $\mathbf{O}$, and P is a product of scaled standard simplices, the *standard case*.

Examples illustrating our main theorems appear at the end of this introduction.

Remark 4. In more concrete terms, the standard case of Main Theorem 1 gives us the expected number of real roots of a random unmixed polynomial system of given multidegree. (Unmixedness simply means that all the polynomials have the same support.) More precisely, if we

1. partition the variables $x_1,\ldots,x_n$ into k vectors $\overline{x}_1,\ldots,\overline{x}_k$ such that $\overline{x}_j$ consists of exactly n_j variables for all j, and
2. assume that $f_{s,1},\ldots,f_{s,n}$ are general ($(\sum n_j)$-variate) polynomials with random variable coefficients and total degree δ_j in $\overline{x}_j$ for all j,

then it is easy to see that the $\mathbb{Z}^n$-polyhedrally randomized systems of the form $F_s := (f_{s,1},\ldots,f_{s,n})$ define the standard case. Also note that with probability 1, such a system has no roots on any coordinate hyperplane.

Remark 5. An alternative characterization of the general case of Main Theorem 1 would be to start with F_s as in the last remark, let $\{b_1,\ldots,b_n\}\subset\mathbb{Z}^n$ be a basis for an n-dimensional lattice L, and let $a_1,\ldots,a_n\in\mathbb{Z}^n$. Then, defining $x^B := (x^{b_1},\ldots,x^{b_n})$, it is clear that any system of the form $(x^{a_1}f_{s,1}(x^B),\ldots,x^{a_n}f_{s,n}(x^B))$ satisfies the hypotheses of Main Theorem 1. Conversely, given L and the Newton polytopes $P_1,\ldots,P_n$ of F, let v be any vertex of P_1. Then by assumption we can pick $a_1 := v$ and $a_2,\ldots,a_n\in\mathbb{Z}^n$ such that $P_i = a_i + P$ for all $i\in[1..n]$. One

can then define $b'_1, \ldots, b'_n$ to be the edges emanating from v and find $\delta_1, \ldots, \delta_k$ (and thus B and F_s) from the Smith factorization [**Jac85**, sec. 3.7] of the matrix $B' := [b'_1, \ldots, b'_n]$. We will see later that this characterization lets us write $\mathrm{Vol}_L(P)$ more explicitly as $\left(\begin{array}{c} n \\ n_1, \ldots, n_k \end{array} \right) \prod \delta_j^{n_j}$.

Remark 6. The multinomial coefficients [**GKP94**] in Main Theorem 1 are defined (in the most obvious way) by using gamma functions so that they are well-defined for half-integral entries. When $\sum n_j < n$ it is clear that $\mathrm{Vol}_L(P) = 0$ and we thus cheat our way out of ambiguity. Note also that the power of π disappears iff there is at most one odd n_j.

Remark 7. Let $\mathcal{M}(E)$ denote the (unnormalized) n-dimensional *mixed volume* [**Ber75, Oda88, Roj94, DGH96, VGC96, EC95**] of the convex hulls of the E_i. Then it easily follows from the results of [**Ber75**] that for *general* E, the expected number of roots of F in $(\mathbb{C}^*)^n$ is exactly $\mathcal{M}(E)$ [**Roj94**]. The connection to real roots is the following: Let $\mathcal{M}_L(\cdot)$ denote the renormalization of $\mathcal{M}(\cdot)$ which evaluates to 1 for n-tuples of L-fundamental n-simplices. Then for E as in Main Theorem 1, $\mathcal{M}_L(E) = \mathrm{Vol}_L(P)$.

Remark 8. Note that the $k=1$ portion of the standard case yields a special case of the following formula due to Shub and Smale [**SS2**]: $\mathcal{N}(E) = \sqrt{\prod d_i}$ when $L = \mathbb{Z}^n$ and, for all i, the Newton polytope of f_i is the standard n-simplex uniformly scaled by a factor of d_i. Main Theorem 1 thus complements their result. (That our distribution specializes to theirs is very easy to verify.) Convex geometry was not mentioned in [**SS2**], so it is even more of a coincidence that $\mathcal{N}(E) = \sqrt{\mathcal{M}_L(E)}$ (by the multilinearity of $\mathcal{M}_L(\cdot)$) in the case they considered.

Given any matrix B with integral entries, let $\mathrm{even}(B)$ (resp. $\mathrm{odd}(B)$) denote the number of nonzero even (resp. odd) entries in its Smith normal form [**Jac85**, sec. 3.7]. Also let $\mathcal{N}_U(E)$ denote the expected number of real roots of F in a region U. Then we can do even better and compute $\mathcal{N}_U(E)$.

Main Theorem 2. Suppose U is a measurable subset of an orthant of $(\mathbb{R}^*)^n$. Then, following the notation of Main Theorem 1 and Remarks 4 and 5,

$$\mathcal{N}_U(E) = \frac{1}{2^{\mathrm{even}(B)}} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{(n+1)}{2}}} \cdot \sqrt{\prod \delta_j^{n_j}} \int_{\exp_B(U)} \prod \frac{d\overline{x}_j}{\sqrt{(1+\overline{x}_j \cdot \overline{x}_j)^{n_j+1}}}$$

where $\exp_B(x) := x^B$.

Remark 9. For the standard case it is clear that B is the identity matrix and thus $\mathrm{even}(B) = 0$ and $\exp_B(U) = U$.

From Main Theorem 2 it is easy to derive that $\mathcal{N}_U(E)$ (for the E specified in Main Theorem 1) is also invariant under a natural group action. If $\sum n_j < n$ this is vacuously true, so let us assume that $\sum n_j = n$. We can then note that $(\mathbb{R}^*)^n$ naturally embeds into the product of projective spaces $X := \mathbb{P}^{n_1}_{\mathbb{R}} \times \cdots \times \mathbb{P}^{n_k}_{\mathbb{R}}$ by homogenizing each group of variables $\overline{x}_j$ with an extra variable. It then follows that under this embedding, the distribution defined in Main Theorem 2 reduces (almost everywhere) to the uniform probability measure on X [**Fol80**]. The product of orthogonal groups $G := O(n_1+1) \times \cdots \times O(n_k+1)$ acts naturally (and isometrically) on X and we thus obtain the following corollary of Main Theorem 2.

Corollary 1. Following the notation of Main Theorems 1 and 2, assume that $n = \sum n_j$. Then there is a natural open embedding $\iota : (\mathbb{R}^*)^n \hookrightarrow \mathbb{P}^{n_1}_{\mathbb{R}} \times \cdots \times \mathbb{P}^{n_k}_{\mathbb{R}}$ making $\mathcal{N}_U(E)$ a well-defined $O(n_1+1)\times\cdots\times O(n_k+1)$-invariant function of $\iota(\exp_B(U))$. □

Remark 10. One can in fact derive the above corollary without resorting to Main Theorem 2. Shub and Smale did this for the case described in Remark 8 and this was how they derived their $O(n+1)$-invariant measure [**SS2**].

For more general E it is somewhat unclear if any sort of natural group invariance can be maintained. However, following the notation of Remarks 4 and 5, the generalization where one replaces δ_j by $d_{ij} \in \mathbb{N}$ preserves G-invariance and is currently under investigation. In particular, the mixed versions of Main Theorems 1 and 2 corresponding to this broader family of E are imminent.

Stretching Main Theorem 1 a little further, we propose the following conjecture.

Square Root Volume Conjecture. Suppose F is an L-polyhedrally randomized system with support E. Then the expected number of roots of F in $(\mathbb{R}^*)^n$ is $K \cdot \sqrt{\mathcal{M}_L(E)}$, where K is a constant depending only on the $\mathbb{GL}_n(\mathbb{Z})$-similarity class of the inner normal fan of the polytope $S := \sum \mathrm{Conv}(E_i)$, modulo $\mathbb{GL}_n(\mathbb{Z})$-similarities of S.

We refer the reader to [**KKMS73, Oda88, Ful93**] for the definitions and properties of normal fans. In particular, the combinatorial condition on K above is equivalent to K depending only on the isomorphism type of the compact *toric variety* over $\mathbb{R}$ associated to S.

Remark 11. As we've already seen, the conjecture is true for those E satisfying the hypotheses of Main Theorem 1. (Simply note that the toric variety associated to P is isomorphic to $\mathbb{P}^{n_1}_{\mathbb{R}} \times \cdots \times \mathbb{P}^{n_k}_{\mathbb{R}}$, following the notation of remark 5.) The conjecture is also clearly true for the case described in remark 8, since the normal fan depends only on n. Also, if $\mathcal{M}_L(E) = 0$ then it easily follows that the expected number of *complex* roots is 0 [**Roj94**]. So the conjecture is true in this degenerate case as well.

Remark 12. If one no longer assumes that E is complete with respect to any lattice, then extending the above conjecture becomes harder. In particular, if one tries to use the $\mathbb{Z}^n$-polyhedral distribution corresponding to the Newton polytopes, then the $n = 1$ case already shows that $\mathcal{N}(E)$ must depend on more information than just Newton polytopes. (For example, one can apply Theorem 1 of the next section to a general trinomial.) One is then faced with an intriguing (but vague) question: Is there a "canonical" probability measure for polynomial systems with incomplete support?

The remainder of our paper is devoted to proving our main theorems. Main Theorem 1 follows from Main Theorem 2 after partitioning $(\mathbb{R}^*)^n$ into orthants and performing a routine multivariable integration. Main Theorem 2, after an algebraic trick, follows from a beautiful integral formula due to Edelman and Kostlan [**EK95**]. Our proofs are detailed in the next section. Also, the references section contains some important additional sources should the reader desire to learn more about random equations, convex geometry, or counting real or complex roots of non-random polynomial systems.

In closing we point out that an important corollary of our L-polyhedral distribution is that it can form the basis for a complexity analysis of solving multihomogeneous polynomial systems. This is ongoing work with J.-P. Dedieu and

initial calculations show that some of the numerical conditioning results of Shub and Smale [**SS1, SS2, SS3, SS4**] — on homotopy methods for solving homogeneous polynomial systems — can be generalized to multihomogeneous systems [**DR96**].

We end this introduction with two examples illustrating our main theorems.

Example 2. (Shifted Powers) Suppose $n=1$, $L:=5\mathbb{Z}$, and $E:=\{-4,1,6\}$. So Main Theorem 1 applies and we are considering a single univariate trinomial of the form $F(x)=\alpha x^{-4}+\beta x+\gamma x^6$, where α,β,γ are independent real Gaussian random variables with mean 0 and respective variances $\frac{1}{2},1,\frac{1}{2}$. Main Theorem 1 tells us that the expected number of roots of F in $\mathbb{R}^*$ is exactly $\sqrt{2}$. Main Theorem 2 tells us that the expected number of positive roots of F is exactly $\frac{1}{\sqrt{2}}$ (since $\exp_B(x)=x^5$).

Example 3. (The Matrix Polynomial Problem) Suppose $L:=\mathbb{Z}^n$ and $E:=(P\cap\mathbb{Z}^n)^n$ where $P:=\mathrm{Conv}\{\mathbf{O},\hat{e}_1,\ldots,\hat{e}_{n-1}\}\times[0,d]$. In particular, this is the standard case with $k=2$ and $(n_1,n_2,\delta_1,\delta_2)=(n-1,1,1,d)$; and $L_1=\cdots=L_n=L$ since $\dim P=n$. Now note that the embedding ι from Corollary 1 can be modified slightly into an embedding $\varphi:\mathbb{R}^n\hookrightarrow\mathbb{P}^{n-1}_{\mathbb{R}}\times\mathbb{R}$. This last embedding can then be used to transform our system into the following randomized version of the $n\times n\times d$ matrix polynomial problem [**GLR82**] (modulo a set of 0 measure): For all $j\in[0..d]$, let A_j be an $n\times n$ matrix consisting of independent real Gaussian random variables with mean 0 and variance $\left(\begin{array}{c}d\\j\end{array}\right)$, and consider the values of λ such that the matrix $A_0+\lambda A_1+\cdots+\lambda^d A_d$ is singular. Letting $x:=(x_1,\ldots,x_n)^{\mathrm{T}}$, Main Theorem 1 (combined with φ) tells us that the expected number of real eigenpairs $(x,\lambda)\in\mathbb{P}^{n-1}_{\mathbb{R}}\times\mathbb{R}$ of

$$A_0x+\lambda A_1x+\cdots+\lambda^dA_dx=0$$

is exactly $\mathcal{N}(E)=\sqrt{\pi}\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\sqrt{d}$. In particular, $\lim_{n\to\infty}\frac{\mathcal{N}(E)}{\sqrt{n}}=\sqrt{\frac{\pi d}{2}}$. Our example complements another result where all the variances are 1 [**EK95**]. When $d=1$, both of these results merge and we have a randomized version of the generalized eigenvalue problem. (This special case was discovered earlier in [**EKS94**].) Finally, for our randomized matrix polynomial problem, Corollary 1 implies that the concentration of real eigenpairs is uniform on $\mathbb{P}^{n-1}_{\mathbb{R}}\times\mathbb{P}^1_{\mathbb{R}}$.

2. Proofs of Our Main Theorems

We will first prove Main Theorem 2. Main Theorem 1 then follows from a simple application of Main Theorem 2 which we will describe at the end of this section.

Let us begin with a theorem which implicitly contains Main Theorem 2, as well as much more.

Theorem 1. [**EK95**, Theorem 7.1] Let $f_0(t),\ldots,f_N(t)$ be any collection of real valued rectifiable functions defined on $\mathbb{R}^m$, let V be a measurable subset of $\mathbb{R}^m$, and let the vectors $a_k:=(a_{k0},\ldots,a_{kN})$, $k\in[1..m]$ be independent and identically distributed. Assume each a_k is a (real) multivariate normal random vector with mean $\mathbf{O}$ and covariance matrix C. Then the expected number of real zeros of the system of equations

$$a_{k0}f_0(t)+\cdots+a_{kN}f_N(t)=0,\ k\in[1..m],$$

that lie in the set V, is

$$\frac{\Gamma\left(\frac{m+1}{2}\right)}{\pi^{\frac{m+1}{2}}}\int_V\left(\det\left[\frac{\partial^2}{\partial x_i\partial y_j}\log(v(x)^TCv(y))\Big|_{y=x=t}\right]_{n\times n}\right)^{\frac{1}{2}}dt,$$

where $v(t)=(f_0(t),\ldots,f_N(t))$. □

Remark 13. It would be a great help to have a mixed version of the above theorem. That is, a version where the vectors $\{a_k\}$ have covariance matrices (and dimensions) depending on k. Such a theorem (which is expected soon) would then immediately give us mixed versions of Main Theorems 1 and 2.

Before beginning our proof of Main Theorem 2 let us define some convenient notation: Let $A_1\oplus\cdots\oplus A_k$ be the block-diagonal matrix (with exactly k blocks) whose $l^{\underline{\text{th}}}$ block (going from northwest to southeast) is the matrix A_l. Also let $(\overline{y}_1,\ldots,\overline{y}_k):=y$, $(\overline{t}_1,\ldots,\overline{t}_k):=t$, and $(\overline{e}_1,\ldots,\overline{e}_k):=e$ be such that $\overline{y}_l$, $\overline{t}_l$, and $\overline{e}_l$ each consist of exactly n_l variables for all $l\in[1..k]$. Finally, let Δ_l be the standard n_l-simplex and define $P_s:=(\delta_1\Delta_1)\times\cdots\times(\delta_k\Delta_k)$.

2.1. Main Theorem 2. First note that the map $\tau\mapsto\tau^n$ defines a conformal automorphism of $\mathbb{R}^*$ (resp. $\mathbb{R}_+$) when n is odd (resp. nonzero and even). Next, note that any $A\in\mathbb{GL}_n(\mathbb{Z})$ defines a permutation of the orthants of $(\mathbb{R}^*)^n$ via $x\mapsto x^A$ (since $A^{-1}\in\mathbb{GL}_n(\mathbb{Z})$ and thus $(y^A)^{A^{-1}}=y^{AA^{-1}}=y$ for any $y\in(\mathbb{R}^*)^n$). Thus, by the Smith factorization, it immediately follows that $\exp_B$ defines a conformal bijection between two subsets $Y,Z\subseteq(\mathbb{R}^*)^n$, each a disjoint union of $2^{\text{odd}(B)}$ orthants. It then becomes clear that $\exp_B:(\mathbb{R}^*)^n\longrightarrow Z$ has exactly $2^{\text{even}(B)}$ analytic inverses, and these inverses differ only by coordinate reflections.

Remark 14. Techniques of this sort are developed further in [**Stu91, PS94, IR95**] in connection with a conjectural combinatorial upper bound on the number of isolated roots in $(\mathbb{R}^*)^n$ of a general non-random sparse polynomial system. The underlying framework is the theory of toric varieties and some excellent references for this theory are [**KKMS73, Oda88, Ful93**].

The preceding facts allow us to simplify the proof of Main Theorem 2 somewhat. In particular, since $x\in U\implies x^B\in\exp_B(U)\implies x\in\exp_B^{-1}(\exp_B(U))$, it is clear from our observations that $2^{\text{even}(B)}\mathcal{N}_U(E)=\mathcal{N}_{\exp_B(U)}(E_s)$, where E_s is the support of F_s (recall remark 5). Also note that multiplying any f_i by a monomial x^{a_i} does not affect the roots of F in $(\mathbb{R}^*)^n$. It thus suffices to work with the standard case.

Now, for the $\mathbb{Z}^n$-polyhedral distribution, it is clear that multiplying the variances of the $c_{i,e}$ by a constant depending only on i does not affect the distribution of the roots of F_s. Also, it is easily checked (since all our simplices our now standard) that $\prod(\delta_l!)$ times the variance of $c_{i,e}$ is actually the following product of multinomial coefficients

$$\prod\binom{\delta_l}{\overline{e}_l}$$

where the $l^{\underline{\text{th}}}$ term is an (n_l+1)-nomial coefficient. We will thus assume in our calculations that the variance of $c_{i,e}$ *is* the above product.

We now invoke Theorem 1 to compute $\mathcal{N}_{\exp_B(U)}(E_s)$ by setting $m:=n$, $V:=\exp_B(U)$, and letting the vector $v(t)$ be $(t^e)_{e\in P_s\cap\mathbb{Z}^n}$. Note that the covariance matrix, in this case, is diagonal. Since the Γ-function factors in front of the corresponding integrals from Main Theorem 2 and Theorem 1 fortuitously match, it

suffices to show that the two integrands are identical, up to the remaining factor of $\sqrt{\prod \delta_l^{n_l}}$. So let us examine the square of the integrand coming from our application of Theorem 1:

$$I^2 := \left[\frac{\partial^2}{\partial x_i \partial y_j} \log(v(x)^T C v(y)) \Big|_{x=y=t} \right]_{n \times n}$$

$$= \left[\frac{\partial^2}{\partial x_i \partial y_j} \log \left(\sum_{e \in P_s \cap \mathbb{Z}^n} \left(\prod_l \binom{\delta_l}{\overline{e}_l} \right) x^e y^e \right) \Big|_{x=y=t} \right]_{n \times n}$$

$$= \left[\frac{\partial^2}{\partial x_i \partial y_j} \log \left(\sum_{e \in P_s \cap \mathbb{Z}^n} \prod_l \binom{\delta_l}{\overline{e}_l} \overline{x}_l^{\overline{e}_l} \overline{y}_l^{\overline{e}_l} \right) \Big|_{x=y=t} \right]_{n \times n}$$

$$= \left[\frac{\partial^2}{\partial x_i \partial y_j} \log \left(\sum_{\overline{e}_k \in \delta_k \Delta_k \cap \mathbb{Z}^n} \cdots \sum_{\overline{e}_1 \in \delta_1 \Delta_1 \cap \mathbb{Z}^n} \prod_l \binom{\delta_l}{\overline{e}_l} \overline{x}_l^{\overline{e}_l} \overline{y}_l^{\overline{e}_l} \right) \Big|_{x=y=t} \right]_{n \times n}$$

$$= \left[\frac{\partial^2}{\partial x_i \partial y_j} \log \left(\prod_l (1 + \overline{x}_l \cdot \overline{y}_l)^{\delta_l} \right) \Big|_{x=y=t} \right]_{n \times n}$$

$$= \left[\frac{\partial^2}{\partial x_i \partial y_j} \left(\sum_l \delta_l \log(1 + \overline{x}_l \cdot \overline{y}_l) \right) \Big|_{x=y=t} \right]_{n \times n}$$

$$= \bigoplus_l \left[\frac{\partial^2}{\partial x_i \partial y_j} \left(\delta_l \log(1 + \overline{x}_l \cdot \overline{y}_l) \right) \Big|_{\overline{x}_l = \overline{y}_l = \overline{t}_l} \right]_{n_l \times n_l}$$

where the $l^{\underline{\text{th}}}$ block in the above block-diagonal matrix is indexed by the variables in $\overline{x}_l$ and $\overline{y}_l$. Since the determinant of a block diagonal matrix is the product of the determinants of the blocks, it thus suffices to know the determinants of the blocks to determine I^2.

In particular,

$$\left[\frac{\partial^2}{\partial x_i \partial y_j} \left(\delta_1 \log(1 + \overline{x}_1 \cdot \overline{y}_1) \right) \Big|_{\overline{x}_1 = \overline{y}_1 = \overline{t}_1} \right]_{n_1 \times n_1}$$

$$= \delta_1 \left[\frac{\partial}{\partial x_i} \left(\frac{x_j}{1 + \overline{x}_1 \cdot \overline{y}_1} \right) \Big|_{\overline{x}_1 = \overline{y}_1 = \overline{t}_1} \right]_{n_1 \times n_1}$$

$$= \delta_1 \left[\frac{\delta_{ij}(1 + \overline{x}_1 \cdot \overline{y}_1) - x_j y_i}{(1 + \overline{x}_1 \cdot \overline{y}_1)^2} \Big|_{\overline{x}_1 = \overline{y}_1 = \overline{t}_1} \right]_{n_1 \times n_1}$$

$$= \delta_1 \left[\frac{\delta_{ij}(1 + \overline{t}_1 \cdot \overline{t}_1) - t_j t_i}{(1 + \overline{t}_1 \cdot \overline{t}_1)^2} \right]_{n_1 \times n_1}$$

$$= \frac{\delta_1}{(1+\overline{t}_1\cdot\overline{t}_1)^2}\left[\delta_{ij}(1+\overline{t}_1\cdot\overline{t}_1) - t_j t_i\right]_{n_1\times n_1}$$

where δ_{ij} is the Kronecker delta. Now (by an observation of J.-P. Dedieu which the author humbly thanks him for) one can use multilinearity to derive that the determinant of the last matrix is

$$\frac{\delta_1^{n_1}}{(1+\overline{t}_1\cdot\overline{t}_1)^{2n_1}}\cdot(1+\overline{t}_1\cdot\overline{t}_1)^{n_1-1} = \frac{\delta_1^{n_1}}{(1+\overline{t}_1\cdot\overline{t}_1)^{n_1+1}}$$

Multiplying the corresponding factors for the remaining blocks of our original large matrix then gives

$$I^2 = \prod \frac{\delta_l^{n_l}}{(1+\overline{t}_l\cdot\overline{t}_l)^{n_l+1}}.$$

Main Theorem 2 follows immediately. □

2.2. Main Theorem 1. Following the notation of the proof of Main Theorem 2, note that we can decompose $(\mathbb{R}^*)^n$ into a disjoint union of $2^{\mathrm{even}(B)}$ reflected copies of Y. Further decomposing into orthants and invoking Main Theorem 2, we obtain by symmetry that

$$\mathcal{N}(E) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{(n+1)}{2}}}\cdot\sqrt{\prod\delta_l^{n_l}}\int_{(\mathbb{R}^*)^n}\prod\frac{d\overline{x}_l}{\sqrt{(1+\overline{x}_l\cdot\overline{x}_l)^{n_l+1}}}$$

Thus our proof amounts to evaluating a product of integrals of the following form:

$$\int_{(\mathbb{R}^*)^{n_1}}\frac{d\overline{x}_1}{\sqrt{(1+\overline{x}_1\cdot\overline{x}_1)^{n_1+1}}}$$

Since the integrand is bounded and continuous on $\mathbb{R}^{n_1}$, we can enlarge the domain of integration (by a set of measure 0) to $\mathbb{R}^{n_1}$ without changing the value of the integral. Then, by converting to spherical coordinates, using the standard formula for the Euclidean measure of a hypersphere, and converting to a trigonometric integral by integration by parts, we obtain that the above integral is precisely

$$\frac{\pi^{\frac{n_1+1}{2}}}{\Gamma(\frac{n_1+1}{2})}$$

Multiplying our integrals together and collecting terms, we obtain that

$$\mathcal{N}(E) = \pi^{\frac{k-1}{2}}\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n_1+1}{2})\cdots\Gamma(\frac{n_k+1}{2})}\cdot\sqrt{\prod\delta_l^{n_l}}.$$

Main Theorem 1 then follows from the fact that

$$\mathrm{Vol}_L(P) = \binom{n}{n_1,\ldots,n_k}\prod\delta_l^{n_l}.$$

This final assertion is easily proved: Using the notation of remark 5, it is clear that $P = BP_s$ and $\det B$ is the lattice index $[\mathbb{Z}^n : L]$. So $\mathrm{Vol}_L(P) = (\det B)\mathrm{Vol}_L(P_s) = [\ \mathbb{Z}^n\ :\ L\ ]\frac{1}{[\mathbb{Z}^n:L]}\mathrm{Vol}_{\mathbb{Z}^n}(P_s) = n!\prod\frac{\delta_l^{n_l}}{n_l!}$ (since the Euclidean d-volume of a standard d-simplex is $\frac{1}{d!}$).

□

3. Acknowledgements

This research grew out of the author's final project for a course on the eigenvalues of random matrices taught by Alan Edelman at U. C. Berkeley. The author thanks Edelman for introducing him to random equations.

While working on this paper the author also visited IRMAR and IBM T. J. Watson Research Center. The author thanks, respectively, Marie-Francoise Roy and Mike Shub for their wonderful hospitality.

References

[Ber75] Bernshtein, D. N., *"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.

[BS86] Bharucha-Ried, A. and Sambandham, M., *Random Polynomials*, Academic Press, New York, 1986.

[CR91] Canny, John F. and Rojas, J. Maurice, *"An Optimal Condition for Determining the Exact Number of Roots of a Polynomial System,"* Proceedings of ISSAC '91 (Bonn, Germany), ACM Press (1991), pp. 96–102.

[DGH96] Dyer, M., Gritzmann, P., and Hufnagel, A., *"On the Complexity of Computing Mixed Volumes,"* SIAM J. Comput., to appear (1996).

[DR96] Dedieu, Jean-Pierre and Rojas, J. Maurice, *"A Condition Number Theorem for Multihomogeneous Polynomial Systems,"* manuscript (1996), MIT.

[EC95] Emiris, Ioannis and Canny, John F., *"Efficient Incremental Algorithms for the Sparse Resultant and the Mixed Volume,"* Journal of Symbolic Computation, vol. 20 (1995), pp. 117–149.

[Egg69] Eggleston, H. G., *Convexity*, Cambridge Tracts in Math. and Math. Physics 47, Cambridge University Press, 1969.

[EO56] Erdös, P. and Offord, A., *"On the Number of Real Roots of a Random Algebraic Equation,"* Proc. London Math. Soc., (3), 6 (1956), pp. 139–160.

[EK95] Edelman, Alan and Kostlan, Eric, *"How Many Zeros of a Random Polynomial are Real?,"* Bull. Amer. Math. Soc., 32, January (1995), pp. 1–37.

[EKS94] Edelman, A., Kostlan, E., and Shub, M. *"How many eigenvalues of a random matrix are real?,"* J. Amer. Math. Soc. 7 (1994), pp. 247–267.

[Fol80] Folland, G. B., *Real Analysis: Modern Techniques and Their Applications*, Wiley, 1980.

[Ful93] Fulton, William, *Introduction to Toric Varieties*, Annals of Mathematics Studies, no. 131, Princeton University Press, Princeton, New Jersey, 1993.

[GKP94] Graham, R. L., Knuth, D. E., and Patashnik, O., *Concrete Mathematics: A Foundation for Computer Science*, 2nd edition, Addison-Wesley, 1994.

[GLR82] Gohberg, I., Lancaster, P., and Rodman, L., *Matrix Polynomials*, Academic Press, 1982.

[Grü69] Grünbaum, Branko, *Convex Polytopes*, Interscience, London, New York, Sydney, 1969.

[GS93] Gritzmann, Peter and Sturmfels, Bernd, *"Minkowski Addition of Polytopes: Computational Complexity and Applications to Gröbner Bases,"* SIAM Journal of Discrete Mathematics, vol. 6, no. 2, pp. 246–269, 1993.

[HS97] Huber, Birk and Sturmfels, Bernd, *"Bernshtein's Theorem in Affine Space,"* Discrete and Computational Geometry 17 (1997), 137–141.

[IR95] Itenberg, Ilya and Roy, Marie-Francoise, *"Multivariate Descartes' Rule,"* Preprint (1995), IRMAR, Université de Rennes I, France.

[Jac85] Jacobson, Nathan, *Basic Algebra I*, 2nd edition, W. H. Freeman and Company, 1985.

[Kac43] Kac, M. *"On the Average Number of Real Roots of a Random Algebraic Equation,"* Bull. Amer. Math. Soc. 49 (1943), pp. 314–320.

[Kho78] Khovanskii, A. G., *"Newton Polyhedra and the Genus of Complete Intersections,"* Functional Analysis (translated from Russian), Vol. 12, No. 1, January–March (1978), pp. 51–61.

[Kho91] Khovanskii, A. G., *Fewnomials*, translated from Russian, AMS monographs, No. 78, 1991.

[KKMS73] Kempf, G., Knudsen, F., Mumford, D., Saint-Donat, B., *Toroidal Embeddings I*, Lecture Notes in Mathematics 339, Springer-Verlag, 1973.

[Kos87] Kostlan, Eric, *"Random Polynomials and the Statistical Fundamental Theorem of Algebra,"* unpublished (1987).

[Kos93] Kostlan, Eric, *"On the Distribution of the Roots of Random Polynomials,"* From Topology to Computation: Proceedings of the Smalefest, edited by M. W. Hirsch, J. Marsden, and M. Shub, Springer Verlag, New York, 1993, Ch. 38, pp. 419–431.

[Kus75] Kushnirenko, A. G., *"A Newton Polytope and the Number of Solutions of a System of k Equations in k Unknowns,"* Usp. Matem. Nauk., 30, No. 2, pp. 266–267 (1975).

[Kus76] Kushnirenko, A. G., *"Newton Polytopes and the Bézout Theorem,"* Functional Analysis and its Applications (translated from Russian), vol. 10, no. 3, July–September (1976), pp. 82–83.

[LO38] Littlewood, J. and Offord, A., *"On the Number of Real Roots of a Random Algebraic Equation,"* J. London Math. Soc. 13 (1938), pp. 288–295.

[LW96] Li, T. Y. and Wang, Xiaoshen, *"The BKK Root Count in $\mathbb{C}^n$,"* Mathematics of Computation, October, 1996.

[Oda88] Oda, Tadeo, *Convex Bodies and Algebraic Geometry: an Introduction to the Theory of Toric Varieties*, Springer-Verlag, 1988.

[PRS93] Pedersen, P., Roy, M.-F., and Szpirglas, A., *"Counting Real Roots in the Multivariate Case,"* Computational Algebraic Geometry, edited by Eyssette and Galligo, Progress in Mathematics 109, pp. 203–224, Birkhauser, 1993.

[PS94] Pedersen, Paul and Sturmfels, Bernd, *"Mixed Monomial Bases,"* to appear in the Proceedings of MEGA '94, "Effective Methods in Algebraic Geometry," Santander, 1994.

[Roj94] Rojas, J. Maurice, *"A Convex Geometric Approach to Counting the Roots of a Polynomial Systems,"* Theoretical Computer Science (1994), vol. 133 (1), pp. 105–140.

[Roj96] Rojas, J. Maurice, *"Toric Intersection Theory for Affine Root Counting,"*, Journal of Pure and Applied Algebra, to appear.

[RW96] Rojas, J. M., and Wang, Xiaoshen, *"Counting Affine Roots of Polynomial Systems Via Pointed Newton Polytopes,"* Journal of Complexity, to appear (1996).

[Sch94] Schneider, Rolf, *Convex Bodies: The Brunn-Minkowski Theory,* Encyclopedia of Mathematics and its Applications, v. 44, Cambridge University Press, 1994.

[SS1] Shub, Mike and Smale, Steve, *"The Complexity of Bezout's Theorem I: Geometric Aspects,"* Journal of the American Mathematical Society 6 (1992), pp. 459–501.

[SS2] Shub, Mike and Smale, Steve, *"The Complexity of Bezout's Theorem II: Volumes and Probabilities,"* Computational Algebraic Geometry (F. Eyssette and A. Galligo, Eds.), pp. 267–285, Birkhauser, 1992.

[SS3] Shub, Mike and Smale, Steve, *"The Complexity of Bezout's Theorem III: Condition Number and Packing,"* Journal of Complexity 9 (1993), pp. 4–14.

[SS4] Shub, Mike and Smale, Steve, *"The Complexity of Bezout's Theorem IV: Probability of Success; Extensions,"* SIAM J. Numer. Anal., to appear (1994).

[Stu91] Sturmfels, Bernd, *"On the Number of Real Roots of a Sparse Polynomial System,"* Hamiltonian and Gradient Flows: Algorithms and Control, (ed. A. Bloch), Fields Institute Communications, Vol. 3, American Math. Soc., Providence, RI, 1991, pp. 137–143.

[VGC96] Verschelde, J., Gatermann, K., and Cools, R., *"Mixed Volume Computation by Dynamic Lifting Applied to Polynomial System Solving,"* Discrete and Computational Geometry, submitted (1996).

Massachusetts Institute of Technology, Mathematics Department, 77 Mass. Ave., Cambridge, MA 02139

E-mail address: rojas@math.mit.edu