PROBABILISTIC CONDITION NUMBER ESTIMATES FOR REAL POLYNOMIAL SYSTEMS I: A BROADER FAMILY OF DISTRIBUTIONS

ALPEREN ERGÜR, GRIGORIS PAOURIS, AND J. MAURICE ROJAS

Abstract. We consider the sensitivity of real roots of polynomial systems with respect to perturbations of the coefficients. In particular — for a version of the condition number defined by Cucker and used later by Cucker, Krick, Malajovich, and Wschebor — we establish new probabilistic estimates that allow a much broader family of measures than considered earlier. We also generalize further by allowing over-determined systems.

In Part II, we study smoothed complexity and how sparsity (in the sense of restricting which terms can appear) can help further improve earlier condition number estimates.

Communicated by Felipe Cucker

1. Introduction

When designing algorithms for polynomial system solving, it quickly becomes clear that complexity is governed by more than simply the number of variables and degrees of the equations. Numerical solutions are meaningless without further information on the spacing of the roots, not to mention their sensitivity to perturbation. A mathematically elegant means of capturing this sensitivity is the notion of *condition number* (see, e.g., [3, 6] and our discussion below).

A subtlety behind complexity bounds incorporating the condition number is that computing the condition number, even within a large multiplicative error, is provably as hard as computing the numerical solution one seeks in the first place (see, e.g., [17] for a precise statement in the linear case). However, it is now known that the condition number admits *probabilistic* bounds, thus enabling its use in average-case analysis, high probability analysis, and smoothed analysis of the complexity of numerical algorithms. This probabilistic approach has revealed (see, e.g., [2, 5, 23]) that, in certain settings, numerical solving can be done in polynomial-time on average, even though numerical solving has exponential worst-case complexity. More recently, the condition number has also proved to be a central quantity in the algorithmic complexity of deeper geometric problems such as the computation of the homology groups of semi-algebraic sets (see, e.g., [14, 8]).

The numerical approximation of complex roots provides an instructive example of how one can profit from randomization.

First, there are classical reductions showing that deciding the existence of complex roots for systems of polynomials in $\bigcup_{m,n\in\mathbb{N}}(\mathbb{Z}[x_1,\ldots,x_n])^m$ is already **NP**-hard. However, classical algebraic geometry (e.g., Bertini's Theorem and Bézout's Theorem [32]) tells us that, with probability 1, the number of complex roots of a *random* system of homogeneous polynomials, $P:=(p_1,\ldots,p_m)\in\mathbb{C}[x_1,\ldots,x_n]$ (with each p_i having fixed positive degree d_i), is 0, $\prod_{i=1}^n d_i$, or infinite, according as $m>n-1$, $m=n-1$, or $m<n-1$. (Any probability measure on the coefficient space, absolutely continuous with respect to Lebesgue measure, will do in the preceding statement.)

2010 *Mathematics Subject Classification.* Primary 65Y20; Secondary 51F99, 68Q25.

Key words and phrases. Condition number, epsilon net, probabilistic bound, kappa, real-solving, over-determined, sub-gaussian .

A.E. was partially supported by NSF grant CCF-1409020 and NSF CAREER grant DMS-1151711. G.P. was partially supported by BSF grant 2010288 and NSF CAREER grant DMS-1151711. J.M.R. was partially supported by NSF grants CCF-1409020 and DMS-1460766.

Secondly, examples like $P := (x_1 - x_2^2, x_2 - x_3^2, \ldots, x_{n-1} - x_n^2, (2x_n - 1)(3x_n - 1))$, which has affine roots $\left(2^{-2^{n-1}}, \ldots, 2^{-2^0}\right)$ and $\left(3^{-2^{n-1}}, \ldots, 3^{-2^0}\right)$, reveal that the number of digits of accuracy necessary to distinguish the coordinates of roots of P may be exponential in n (among other parameters). However, it is now known via earlier work on discriminants and random polynomial systems (see, e.g., [9, Thm. 5]) that the number of digits needed to separate roots of P is polynomial in n *with high probability*, assuming the coefficients are rational, and the polynomial degrees and coefficient heights are bounded. More simply, a classical observation from the theory of resultants (see, e.g., [7]) is that, for any positive continuous probability measure on the coefficients, P having a root with Jacobian matrix possessing small determinant is a rare event. So, with high probability, small perturbations of a P with no degenerate roots should still have no degenerate roots. More precisely, we review below a version of the condition number used in [33, 2, 23]. Recall that the *singular values* of a matrix $T \in \mathbb{R}^{k\times(n-1)}$ are the (nonnegative) square roots of the eigenvalues of $T^\top T$, where $T^\top$ denotes the transpose of T.

Definition 1.1. *Given $n, d_1, \ldots, d_m \in \mathbb{N}$ and $i \in \{1, \ldots, m\}$, let $p_i \in \mathbb{R}[x_1, \ldots, x_n]$ be homogeneous polynomials with $\deg p_i = d_i$, and let $P := (p_1, \ldots, p_m)$ be the corresponding polynomial system. We set $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ where $\alpha := (\alpha_1, \ldots, \alpha_n)$, and let $c_{i,\alpha}$ denote the coefficient of x^α in p_i. We define the* Weyl-Bombieri norms *of p_i and P to be, respectively,*

$$\|p_i\|_W := \sqrt{\sum_{\alpha_1+\cdots+\alpha_n=d_i} \frac{|c_{i,\alpha}|^2}{\binom{d_i}{\alpha}}} \text{ and } \|P\|_W := \sqrt{\sum_{i=1}^m \|p_i\|_W^2}.$$

Let $\Delta_m \in \mathbb{R}^{m\times m}$ be the diagonal matrix with diagonal entries $\sqrt{d_1}, \ldots, \sqrt{d_m}$ and let $DP(x)|_{T_xS^{n-1}} : T_xS^{n-1} \longrightarrow \mathbb{R}^m$ denote the linear map between tangent spaces induced by the Jacobian matrix of the polynomial system P evaluated at the point x. Finally, when $m=n-1$, we define the (normalized) local condition number *(for solving $P = \mathbf{O}$) to be $\tilde{\mu}_{\text{norm}}(P, x) := \|P\|_W \sigma_{\max}\left(DP(x)|_{T_xS^{n-1}}^{-1} \Delta_{n-1}\right)$ or $\tilde{\mu}_{\text{norm}}(P, x) := \infty$, according as $DP(x)|_{T_xS^{n-1}}$ is full rank or not, where $\sigma_{\max}(A)$ is the largest singular value of a matrix A.* $\diamond$

Clearly, $\tilde{\mu}_{\text{norm}}(P, x) \to \infty$ as P approaches a system possessing a degenerate root $\zeta \in \mathbb{P}_{\mathbb{C}}^{n-1}$ and x approaches ζ. The intermediate normalizations in the definition are useful for geometric interpretations of $\tilde{\mu}_{\text{norm}}$: There is in fact a simple and elegant algebraic relation between $\|P\|_W$, $\sup_{x\in S^{n-1}} \tilde{\mu}_{\text{norm}}(P, x)$, and the distance of P to a certain discriminant variety (reviewed in Section 2 and Theorem 2.1 below, see also [12]). But even more importantly, the preceding condition number (in the special case $m = n - 1$) was a central ingredient in the recent positive solution to *Smale's 17th Problem* [2, 23]: For the problem of numerically approximating a *single* complex root of a polynomial system, a particular randomization model (independent complex Gaussian coefficients with specially chosen variances) enables *polynomial-time* average-case complexity, in the face of exponential deterministic complexity.[1]

1.1. **From Complex Roots to Real Roots.**

It is natural to seek similar average-case speed-ups for the harder problem of numerically approximating real roots of real polynomial systems. However, an important subtlety one must consider is that the number of real roots of $n - 1$ homogeneous polynomials in n variables (of fixed degree) is no longer constant with

[1]Here, "complexity" simply means the total number of field operations over $\mathbb{C}$ needed to find a start point x_0 for Newton iteration, such that the sequence of Newton iterates $(x_n)_{n\in\mathbb{N}}$ converges to a true root ζ of P (see, e.g., [3, Ch. 8]) at the rate of $|x_n - \zeta| \leq (1/2)^{2^{n-1}} |x_0 - \zeta|$ or faster.

probability 1, even if the probability measure for the coefficients is continuous and positive. Also, small perturbations can make the number of real roots of a polynomial system go from positive to zero or even infinity. A condition number for real solving that takes all these subtleties into account was developed in [10] and applied in the series of papers [11, 12, 13]. In these papers, the authors performed a probabilistic analysis assuming the coefficients were independent real Gaussians with mean 0 and very specially chosen variances.

Definition 1.2. [10] *Let* $\tilde{\kappa}(P,x) := \frac{\|P\|_W}{\sqrt{\|P\|_W^2 \tilde{\mu}_{\mathrm{norm}}(P,x)^{-2} + \|P(x)\|_2^2}}$ *and* $\tilde{\kappa}(P) := \sup_{x \in S^{n-1}} \tilde{\kappa}(P,x)$. *We respectively call* $\tilde{\kappa}(P,x)$ *and* $\tilde{\kappa}(P)$ *the* local *and* global *condition numbers for real solving.* ⋄

Note that a large condition number for real solving can be caused not only by a root with small Jacobian determinant, but also by the existence of a critical point for P with small corresponding critical value. So a large $\tilde{\kappa}$ is meant to detect the spontaneous creation of real roots, as well as the bifurcation of a single degenerate root into multiple distinct real roots, arising from small perturbations of the coefficients.

Our main results, Theorems 3.9 and 3.10 in Section 3.4 below, show that useful condition number estimates can be derived for a much broader class of probability measures than considered earlier: Our theorems allow non-Gaussian distributions, dependence between certain coefficients, and, unlike the existing literature, our methods do not use any additional algebraic structure, e.g., invariance under the unitary group acting linearly on the variables (as in [33, 11, 12, 13]). This aspect also allows us to begin to address sparse polynomials (in the sequel to this paper), where linear changes of variables would destroy sparsity. Our framework also allows over-determined systems ($m > n-1$). We leave the under-determined case ($m < n-1$) for future work.

To compare our results with earlier estimates, let us first recall a central estimate from [13].

Theorem 1.3. [13, Thm. 1.2] *Let* $P := (p_1, \ldots, p_{n-1})$ *be a random system of homogenous* n*-variate polynomials where* $n \geq 3$ *and* $p_i(x) := \sum_{\alpha_1+\cdots+\alpha_n=d_i} \sqrt{\binom{d_i}{\alpha}} c_{i,\alpha} x^\alpha$ *where the* $c_{i,\alpha}$ *are independent real Gaussian random variables having mean* 0 *and variance* 1. *Then, letting* $N := \sum_{i=1}^{n-1} \binom{n+d_i-1}{d_i}$, $d := \max_i d_i$, $M' := 1 + 8d^2 \sqrt{(n-1)^5 N \prod_{i=1}^{n-1} d_i}$, *and* $t \geq \sqrt{\frac{n-1}{4\prod_{i=1}^{n-1} d_i}}$, *we have:*

1. $\mathrm{Prob}(\tilde{\kappa}(P) \geq tM') \leq \frac{\sqrt{1+\log(tM')}}{t}$
2. $\mathbb{E}(\log(\tilde{\kappa}(P))) \leq \log(M') + \sqrt{\log M'} + \frac{1}{\sqrt{\log M'}}$.

The expanded class of distributions we allow for the coefficients of P satisfy the following more flexible hypotheses:

Notation 1.4. *For any* $d_1, \ldots, d_m \in \mathbb{N}$ *and* $i \in \{1, \ldots, m\}$, *let* $d := \max_i d_i$, $N_i := \binom{n+d_i-1}{d_i}$, *and assume* $C_i = (c_{i,\alpha})_{\alpha_1+\cdots+\alpha_n=d_i}$ *are independent random vectors in* $\mathbb{R}^{N_i}$ *with probability distributions satisfying:*

1. *(Centering) For any* $\theta \in S^{N_i-1}$ *we have* $\mathbb{E}\langle C_i, \theta\rangle = 0$.
2. *(Subgaussian) There is a* $K > 0$ *such that for every* $\theta \in S^{N_i-1}$ *we have*
$$\mathrm{Prob}\left(|\langle C_i, \theta\rangle| \geq t\right) \leq 2e^{-t^2/K^2} \text{ for all } t > 0.$$
3. *(Small Ball) There is a* $c_0 > 0$ *such that for every vector* $a \in \mathbb{R}^{N_i}$ *we have*
$$\mathrm{Prob}\left(|\langle a, C_i\rangle| \leq \varepsilon \|a\|_2\right) \leq c_0\varepsilon \text{ for all } \varepsilon > 0.$$ ⋄

By the vectors C_i being independent we simply mean that the probability density function for the longer vector $C_1\times\cdots\times C_m$ can be expressed as a product of the form $\prod_{i=1}^m f_i(\ldots,c_{i,\alpha},\ldots)$. This is a much weaker assumption than having *all* the $c_{i,\alpha}$ be independent, as is usually done in the literature on random polynomial systems.

The standard Gaussian distribution is a typical example of a collection of random vectors satisfying our assumptions with universal constants. This easily follows from the fact that for a standard Gaussian random vector $C\in\mathbb{R}^{N_i}$, and for any $\theta\in S^{N_i-1}$, the one dimensional marginal $\langle\theta,C\rangle$ is a standard Gaussian random variable.

Another example of a collection of random vectors satisfying the 3 assumptions above can be obtained by letting $p>2$ and letting C_i have the uniform distribution on $B_p^{N_i}:=\left\{x\in\mathbb{R}^{N_i}\mid\sum_{j=1}^{N_i}x_j^p\leq 1\right\}$ for all i: In this case the subgaussian assumption follows from [1, Sec. 6] and the small-ball assumption is a direct consequence of the fact that $B_p^{N_i}$ satisfies Bourgain's Hyperplane Conjecture (see, e.g., [21]). Yet another important example (easier to verify) is to let the C_i have the uniform distribution on ℓ_2 unit-spheres of varying dimension.

The subgaussian and small-ball assumptions are standard assumptions in modern non-asymptotic theory of random matrices and in general in the applications of high-dimensional probability to Data Sciences (See [30], [31]). One of the reasons that these assumptions are so popular is that these properties "tensorize nicely": In particular a standard application of Bernstein's inequality shows that if $X_i, i\in\{1,2,\ldots,N_i\}$ are independent centered random variables that are all subgaussian with constant K then the random vector $X=(X_1,\cdots,X_{N_i})$ is also subgaussian with constant CK, where C is an absolute universal constant. Also a recent result of Rudelson and Vershynin ([29]) states that if all the X_i have the small ball property with constant c_0 then the random vector $X=(X_1,\cdots,X_{N_i})$ has the small ball property with constant C_1c_0, where C_1 is a universal constant. The best possible constant in this case is known (see [24] or [25]). This "tensorization property" also shows that there are numerous examples of random vectors that satisfy our assumptions. Examples of subgaussian random variables that satisfy the small ball assumption are the random variables X_p, $p\geq 2$ that have densities $f(t):=c_pe^{-|t|^p}$, $t\in\mathbb{R}$, where c_p is a constant depending on p such that $\int f=1$. (In this case the subgaussian constant and the small ball constants are universal constants, independent of p).

Other examples of random variables that have the two properties are the random variables that has a bounded density f with a bounded support. (In these cases the subgaussian constant depends on the size of the support of the density and the small ball constant depends on the "infinity norm" of the density).

A simplified summary of our main results (Theorems 3.9 and 3.10 from Section 3.4), in the special case of square dense systems, is the following:

Corollary 1.5. *There is an absolute constant $A>0$ with the following property. Let $P:=(p_1,\ldots,p_{n-1})$ be a random system of homogenous n-variate polynomials where $p_i(x):=\sum_{\alpha_1+\cdots+\alpha_n=d_i}\sqrt{\binom{d_i}{\alpha}}c_{i,\alpha}x^\alpha$ and $C_i=(c_{i,\alpha})_{\alpha_1+\cdots+\alpha_n=d_i}$ are independent random vectors satisfying the centering, subgaussian and small-ball assumptions, with underlying constants c_0 and K. Then, for $n\geq 3$, $d:=\max_i d_i$, $d\geq 2$, $N:=\sum_{i=1}^{n-1}\binom{n+d_i-1}{d_i}$, and $M:=A\sqrt{N}(Kc_0)^{2(n-1)}(3d^2\log(ed))^{2n-3}\sqrt{n}$ the following bounds hold:*

1. $\mathrm{Prob}(\tilde{\kappa}(P) \geq tM) \leq \begin{cases} 3t^{-\frac{1}{2}} & ;\ \textit{if}\ 1 \leq t \leq (ed)^{2(n-1)} \\ 3t^{-\frac{1}{2}} \left(\frac{t}{(ed)^{2(n-1)}}\right)^{\frac{1}{4\log(ed)}} & ;\ \textit{if}\ t \geq (ed)^{2(n-1)} \end{cases}$

2. $\mathbb{E}(\log \tilde{\kappa}(P)) \leq 1 + \log M$.

Corollary 1.5 is proved in Section 3.4. Theorems 3.9 and 3.10 in Section 3.4 below in fact state much stronger estimates than our simplified summary above.

Note that, for fixed d and n, the bound from Assertion (1) of Corollary 1.5 shows a somewhat slower rate of decay for the probability of a large condition number than the older bound from Assertion (1) of Theorem 1.3: $O(1/t^{0.3523})$ vs. $O(\sqrt{\log t}/t)$. However, the older $O(\sqrt{\log t}/t)$ bound was restricted to a special family of Gaussian distributions (satisfying invariance with respect to a natural $O(n)$-action on the root space $\mathbb{P}^{n-1}_{\mathbb{R}}$) and assumes $m = n - 1$. Our techniques come from geometric functional analysis, work for a broader family of distributions, and we make no group-invariance assumptions.

Furthermore, our techniques allow condition number bounds in a new setting: over-determined systems, i.e., $m \times n$ systems with $m > n - 1$. See the next section for the definition of a condition number enabling $m > n - 1$, and the statements of Theorems 3.9 and 3.10 for our most general condition number bounds. The over-determined case occurs in many important applications involving large data, where one may make multiple redundant measurements of some physical phenomenon, e.g., image reconstruction from multiple projections. There appear to have been no probabilistic condition number estimates for the case $m > n - 1$ until now. In particular, for m proportional to n, we will see at the end of this paper how our condition number estimates are close to optimal.

To the best of our knowledge, the only other result toward estimating condition numbers of non-Gaussian random polynomial systems is due to Nguyen [26]. However, in [26] the degrees of the polynomials are assumed to be bounded by a small fraction of the number of variables, $m = n - 1$, and the quantity analyzed in [26] is not the condition number considered in [33] or [11, 12, 13].

The precise asymptotics of the decay rate for the probability of having a large condition number remain unknown, even in the restricted Gaussian case considered by Cucker, Krick, Malajovich and Wschebor. So we also prove *lower* bounds for the condition number of a random polynomial system. To establish these bounds, we will need one more assumption on the randomness.

Notation 1.6. *For any $d_1, \ldots, d_m \in \mathbb{N}$ and $i \in \{1, \ldots, m\}$, let $d := \max_i d_i$, $N_i := \binom{n+d_i-1}{d_i}$, and assume $C_i = (c_{i,\alpha})_{\alpha_1+\cdots+\alpha_n=d_i}$ is an independent random vector in $\mathbb{R}^{N_i}$ with probability distribution satisfying:*

4. (Euclidean Small Ball) There is a constant $\tilde{c}_0 > 0$ such that for every $\varepsilon > 0$ we have

$$\mathrm{Prob}\left(\|C_i\|_2 \leq \varepsilon\sqrt{N_i}\right) \leq (\tilde{c}_0\varepsilon)^{N_i}. \ \diamond$$

Remark 1.7. *If the vectors C_i have independent coordinates satisfying the centering and small-ball Assumptions, then Lemma 3.4 from Section 3.3 implies that the Euclidean small ball assumption holds as well. Moreover, if the C_i are each uniformly distributed on a convex body X and satisfy our Centering and Subgaussian assumptions, then a result of Jean Bourgain* [4] *(see also* [15] *or* [20] *for alternative proofs) implies that both the small ball and Euclidean small ball assumptions hold, and with $\tilde{c}_0$ depending only on the subgaussian constant K (not the convex body X).* $\diamond$

Corollary 1.8. *Suppose $n, d \geq 3$, $m = n-1$, and $d_j = d$ for all $j \in \{1, \ldots, n-1\}$. Also let $P := (p_1, \ldots, p_m)$ be a random polynomial system satisfying our centering, subgaussian, small ball, and Euclidean small ball assumptions, with respective underlying constants K and $\tilde{c}_0$. Then, there are constants $A_2 \geq A_1 > 0$ depending only on c_0 and K (i.e., independent of n and d), such that*

$$A_1(n\log(d) + d\log(n)) \leq \mathbb{E}(\log \tilde{\kappa}(P)) \leq A_2(n\log(d) + d\log(n)). \quad \blacksquare$$

Corollary 1.8 follows immediately from a more general estimate: Lemma 3.13 from Section 3.3. It would certainly be more desirable to know bounds within a constant multiple of $\tilde{\kappa}(P)$ instead. We discuss more refined estimates of the latter kind in Section 3.5, after the proof of Lemma 3.13.

As we close our introduction, we point out that one of the tools we developed to prove our main theorems may be of independent interest: Theorem 2.4 of the next section extends, to polynomial systems, an earlier estimate of Kellog [19] on the norm of the derivative of a single multivariate polynomial.

2. Technical Background

We start by defining an inner product structure on spaces of polynomial systems. For n-variate degree d homogenous polynomials $f(x) := \sum_{|\alpha|=d} b_\alpha x^\alpha, g(x) := \sum_{|\alpha|=d} c_\alpha x^\alpha \in \mathbb{R}[x_1, \ldots, x_n]$, their *Weyl-Bombieri inner product* is defined as

$$\langle f, g \rangle_W := \sum_{|\alpha|=d} \frac{b_\alpha c_\alpha}{\binom{d}{\alpha}}.$$

It is known (see, e.g., [22, Thm. 4.1]) that for any $U \in O(n)$ we have

$$\langle f \circ U, g \circ U \rangle_W = \langle f, g \rangle_W.$$

Let $D := (d_1, \ldots, d_m)$ and let H_D denote the space of (real) $m \times n$ systems of homogenous n-variate polynomials with respective degrees $d_1, \ldots, d_m$. Then for $F := (f_1, \ldots, f_m) \in H_D$ and $G := (g_1, \ldots, g_m) \in H_D$ we define the Weyl-Bombieri inner product *for two polynomial systems* to be $\langle F, G \rangle_W := \sum_{i=1}^m \langle f_i, g_i \rangle_W$. We also let $\|F\|_W := \sqrt{\langle F, F \rangle}$.

A geometric justification for the definition of the condition number $\tilde{\kappa}$ can then be derived as follows: First, for $x \in S^{n-1}$, we abuse notation slightly by also letting $DP(x)$ denote the $m \times n$ Jacobian matrix of P, evaluated at the point x. For $m = n-1$ we denote the set of polynomial systems with singularity at x by

$$\Sigma_{\mathbb{R}}(x) := \{P \in H_D \mid x \text{ is a multiple root of } P\}$$

and we then define $\Sigma_{\mathbb{R}}$ (the *real part of the disciminant variety for H_D*) to be:

$$\Sigma_{\mathbb{R}} := \{P \in H_D \mid P \text{ has a multiple root in } S^{n-1}\} = \bigcup_{x \in S^{n-1}} \Sigma_{\mathbb{R}}(x).$$

Using the Weyl-Bombieri inner-product to define the underlying distance, we point out the following important geometric characterization of $\tilde{\kappa}$:

Theorem 2.1. [12, Prop. 3.1] *When $m = n-1$ we have $\tilde{\kappa}(P) = \frac{\|P\|_W}{\mathrm{Dist}(P, \Sigma_{\mathbb{R}})}$ for all $P \in H_D$.* $\blacksquare$

We call a polynomial system $P = (p_1, \ldots, p_m)$ with $m = n-1$ (resp. $m \geq n$) *square* (resp. *over-determined*). Newton's method for over-determined systems was studied in [16]. So now that we have a geometric characterization of the condition number for square systems it will be useful to also have one for over-determined systems.

Definition 2.2. *Let $\sigma_{\min}(A)$ denote the smallest singular value of a matrix A. For any system of homogeneous polynomials $P\in(\mathbb{R}[x_1,\ldots,x_n])^m$ set*
$$L(P,x) := \sqrt{\sigma_{\min}\left(\Delta_m^{-1}DP(x)|_{T_xS^{n-1}}\right)^2 + \|P(x)\|_2^2}$$
For notational convenience, we also set
$$L(P) = \min_{x\in S^{n-1}} L(P,x)$$
we then define
$$\tilde{\kappa}(P,x) = \frac{\|P\|_W}{L(P,x)}$$
and
$$\tilde{\kappa}(P) = \sup_{x\in S^{n-1}} \kappa(P,x) = \frac{\|P\|_W}{L(P)}$$
The quantity $\min\limits_{x\in S^{n-1}} L(P,x)$ thus plays the role of $\mathrm{Dist}(P,\Sigma_{\mathbb{R}})$ in the more general setting of $m \geq n-1$. We now recall an important observation from [12, Sec. 2]: Setting $D_x(P) := DP(x)|_{T_xS^{n-1}}$ we have $\sigma_{\min}(\Delta_{n-1}^{-1}D_x(P)) = \sigma_{\max}\left(D_x(P)^{-1}\Delta_{n-1}\right)^{-1}$, when $m=n-1$ and $D_x(P)$ is invertible. So by the definition of $\tilde{\mu}_{\mathrm{norm}}(P,x)$ we have
$$L(P,x) = \sqrt{\sigma_{\max}\left(D_x(P)^{-1}\Delta_{n-1}\right)^{-2} + \|P(x)\|_2^2} = \sqrt{\|P\|_W^2\tilde{\mu}_{\mathrm{norm}}(P,x)^{-2} + \|P(x)\|_2^2}$$
and thus our more general definition agrees with the classical definition in the square case.

Since the Bombeiri-Weyl norm of a random polynomial system has strong concentration properties for a broad variety of distributions (see, e.g., [34]), we will be interested in the behavior of $L(P,x)$. So let us define the related quantity
$$\mathcal{L}(x,y) := \sqrt{\|\Delta_m^{-1}D^{(1)}P(x)(y)\|_2^2 + \|P(x)\|_2^2}$$
For $m\geq n-1$, it follows directly that $L(P,x) = \inf\limits_{\substack{y\perp x\\ y\in S^{n-1}}} \mathcal{L}(x,y)$.

We now recall a classical result of O. D. Kellog. The theorem below is a summary of [19, Thms. 4–6].

Theorem 2.3. [19] *Let $p\in\mathbb{R}[x_1,\ldots,x_n]$ have degree d and set $\|p\|_\infty := \sup_{x\in S^{n-1}}|p(x)|$ and $\|D^{(1)}p\|_\infty := \max_{x,u\in S^{n-1}}|D^{(1)}p(x)(u)|$. Then:*

1. *We have $\|D^{(1)}p\|_\infty \leq d^2\|p\|_\infty$ and, for any mutually orthogonal $x,y\in S^{n-1}$, we also have $|D^{(1)}p(x)(y)| \leq d\|p\|_\infty$.*
2. *If p is homogenous then we also have $\|D^{(1)}p\|_\infty \leq d\|p\|_\infty$.* ■

For any system of homogeneous polynomials $P := (p_1,\ldots,p_m)\in(\mathbb{R}[x_1,\ldots,x_n])^m$ define $\|P\|_\infty := \sup_{x\in S^{n-1}}\sqrt{\sum_{i=1}^m p_i(x)^2}$. Let $DP(x)(u)$ denote the image of the vector u under the linear operator $DP(x)$, and set
$$\left\|D^{(1)}P\right\|_\infty := \sup_{x,u\in S^{n-1}} \|DP(x)(u)\|_2 = \sup_{x,u\in S^{n-1}} \sqrt{\sum_{i=1}^m \langle\nabla p_i(x),u\rangle^2}.$$

Theorem 2.4. *Let $P := (p_1,\ldots,p_m)\in(\mathbb{R}[x_1,\ldots,x_n])^m$ be a polynomial system with p_i homogeneous of degree d_i for each i and set $d:=\max_i d_i$. Then:*

(1) *We have* $\|D^{(1)}P\|_\infty \leq d^2\|P\|_\infty$ *and, for any mutually orthogonal* $x, y \in S^{n-1}$*, we also have* $\|DP(x)(y)\|_2 \leq d\|P\|_\infty$.
(2) *If* $\deg(p_i) = d$ *for all* $i \in \{1,\ldots,m\}$ *then we also have* $\|D^{(1)}P\|_\infty \leq d\|P\|_\infty$.

Proof. Let (x_0, u_0) be such that $\|D^{(1)}P\|_\infty = \|DP(x_0)(u_0)\|_2$ and let $\alpha := (\alpha_1,\ldots,\alpha_m)$ where $\alpha_i := \frac{\langle \nabla p_i(x_0), u_0\rangle}{\|D^{(1)}P\|_\infty}$. Note that $\|\alpha\|_2 = 1$. Now define a polynomial $q \in \mathbb{R}[x_1,\ldots,x_n]$ of degree d via $q(x) := \alpha_1 p_1(x) + \cdots + \alpha_m p_m(x)$ and observe that

$$\nabla q(x) = \left(\alpha_1 \frac{\partial p_1}{\partial x_1} + \cdots + \alpha_m \frac{\partial p_m}{\partial x_1}, \ldots, \alpha_1 \frac{\partial p_1}{\partial x_n} + \cdots + \alpha_m \frac{\partial p_m}{\partial x_n}\right),$$

$$\langle \nabla q, u\rangle = u_1\left(\alpha_1 \frac{\partial p_1}{\partial x_1} + \cdots + \alpha_m \frac{\partial p_m}{\partial x_1}\right) + \cdots + u_n\left(\alpha_1 \frac{\partial p_1}{\partial x_n} + \cdots + \alpha_m \frac{\partial p_m}{\partial x_n}\right),$$

and $\langle \nabla q(x), u\rangle = \sum_{i=1}^m \alpha_i \langle \nabla p_i(x), u\rangle$. In particular, for our chosen x_0 and u_0, we have

$$\langle \nabla q(x_0), u_0\rangle = \sum_{i=1}^m \alpha_i \langle \nabla p_i(x_0), u_0\rangle = \sum_{i=1}^m \frac{\langle \nabla p_i(x_0), u_0\rangle^2}{\|D^{(1)}P\|_\infty} = \left\|D^{(1)}P_\infty\right\|.$$

Using the first part of Kellog's Theorem we have

$$\|D^{(1)}P\|_\infty \leq \sup_{x,u \in S^{n-1}} |\langle \nabla q(x), u\rangle| \leq d^2\|q\|_\infty.$$

Now we observe by the Cauchy-Schwarz Inequality that

$$\|q\|_\infty = \sup_{x \in S^{n-1}} \left|\sum_{i=1}^m \alpha_i p_i(x)\right| \leq \sup_{x \in S^{n-1}} \sqrt{\sum_{i=1}^m p_i(x)^2}.$$

So we conclude that $\|D^{(1)}P\|_\infty \leq d^2\|q\|_\infty \leq d^2 \sup_{x \in S^{n-1}} \sqrt{\sum_{i=1}^m p_i(x)^2} = d^2\|P\|_\infty$. We also note that when $\deg(p_i) = d$ for all i, the polynomial q is homogenous of degree d. So for this special case, the second part of Kellog's Theorem directly implies $\|D^{(1)}P\|_\infty \leq d\|P\|_\infty$.

For the proof of the first part of Assertion (1) we define $\alpha_i = \frac{\langle \nabla p_i(x), y\rangle}{\|DP(x)(y)\|_2}$ and $q(x) = \alpha_1 p_1 + \cdots + \alpha_n p_n$. Then $\langle \nabla q(x), y\rangle = \sum_i \alpha_i \langle \nabla p_i(x), y\rangle = \|DP(x)(y)\|_2$. By applying Kellog's Theorem on the orthogonal direction y we then obtain

$$\|DP(x)(y)\|_2 = \langle \nabla q(x), y\rangle \leq d\|q\|_\infty \leq d\|P\|_\infty. \qquad \blacksquare$$

Using our extension of Kellog's Theorem to polynomial systems, we develop useful estimates for $\|P\|_\infty$ and $\|D^{(i)}P\|_\infty$. In what follows, we call a subset $\mathcal{N}$ of a metric space X a δ-*net on* X if and only if the every point of X is within distance δ of some point of $\mathcal{N}$. A basic fact we'll use repeatedly is that, for any $\delta > 0$ and *compact* X, one can always find a finite δ-net for X.

Lemma 2.5. *Let* $P := (p_1,\ldots,p_m) \in (\mathbb{C}[x_1,\ldots,x_n])^m$ *be a system of homogenous polynomials,* $\mathcal{N}$ *a* δ*-net on* S^{n-1}*, and set* $d := \max_i d_i$*. Let* $\max_{\mathcal{N}}(P) := \sup_{y \in \mathcal{N}} \|P(y)\|_2$*. Similarly let us define* $\max_{\mathcal{N}^{k+1}}(D^{(k)}P) := \sup_{x,u_1,\ldots,u_k \in \mathcal{N}} \|D^{(k)}P(x)(u_1,\ldots,u_k)\|_2$*, and set* $\left\|D^{(k)}P\right\|_\infty := \sup_{x,u_1,\ldots,u_k \in S^{n-1}} \left\|D^{(k)}P(x)(u_1,\ldots,u_k)\right\|_2$*. Then:*

(1) $\|P\|_\infty \leq \frac{\max_{\mathcal{N}}(P)}{1-\delta d^2}$ *and* $\|D^{(k)}P\|_\infty \leq \frac{\max_{\mathcal{N}^{k+1}}(D^{(k)}P)}{1-\delta d^2\sqrt{k+1}}$.

(2) *If* $\deg(p_i) = d$ *for each* $i \in \{1, \ldots, m\}$ *then we have*

$$\|P\|_\infty \leq \frac{\max_{\mathcal{N}}(P)}{1 - \delta d} \text{ and } \|D^{(k)}P\|_\infty \leq \frac{\max_{\mathcal{N}^{k+1}}(D^{(k)}P)}{1 - \delta d\sqrt{k+1}}.$$

Proof. We first prove Assertion (2). Observe that the Lipschitz constant of P on S^{n-1} is bounded from above by $\|D^{(1)}p\|_\infty$: This can be seen by taking $x, y \in S^{n-1}$ and considering the integral $P(x) - P(y) = \int_0^1 DP(y + t(x-y))(x-y)\, dt$.

Since $\|y + t \cdot (x-y)\|_2 \leq 1$ for all $t \in [0,1]$, the homogeneity of the system P implies

$$\|DP(y + t(x-y))(x-y)\|_2 \leq \|D^{(1)}P\|_\infty \|x-y\|_2$$

Using our earlier integral formula, we conclude that $\|P(x) - P(y)\|_2 \leq \|D^{(1)}P\|_\infty \|x-y\|_2$.

Now, when the degrees of the p_i are identical, let the Lipschitz constant of P be M. By Assertion (2) of Theorem 2.4 we have $M \leq \|D^{(1)}P\|_\infty \leq d\|P\|_\infty$. Let $x_0 \in S^{n-1}$ be such that $\|P(x_0)\|_2 = \|P\|_\infty$ and let $y \in \mathcal{N}$ satisfy $|x_0 - y| \leq \delta$. Then $\|P\|_\infty = \|P(x_0)\|_2 \leq \|P(y)\|_2 + \|x_0 - y\|_2 M \leq \max_{\mathcal{N}}(P) + \delta d\|P\|_\infty$, and thus

$$(\star) \qquad \|P\|_\infty (1 - d\delta) \leq \max_{x \in \mathcal{N}} P(x).$$

To bound the norm of $D^{(k)}P(x)(u_1, \ldots, u_k)$ let us consider the net defined by $\mathcal{N} \times \cdots \times \mathcal{N} = \mathcal{N}^{k+1}$ on $S^{n-1} \times \cdots \times S^{n-1}$. Let $x := (x_1, \ldots, x_{k+1}) \in S^{n-1} \times \cdots \times S^{n-1}$ and $y := (y_1, \ldots, y_{k+1}) \in \mathcal{N}^{k+1}$ be such that $\|x_i - y_i\|_2 \leq \delta$ for all i. Clearly, $\|x - y\|_2 \leq \delta\sqrt{k+1}$. Since x was arbitrary, this argument proves that $\mathcal{N}^{k+1}$ is a $\delta\sqrt{k+1}$-net. Note also that $D^{(k)}P(x)(u_1, \ldots, u_k)$ is a homogenous polynomial system with $(k+1)n$ variables and degree d. The desired bound then follows from Inequality $(\star)$ obtained above.

To prove Assertion (1) of our current lemma, the preceding proof carries over verbatim, simply employing Assertion (1), instead of Assertion (2), from Theorem 2.4. ■

3. Condition Number of Random Polynomial Systems

3.1. Introducing Randomness. Now let $P := (p_1, \ldots, p_m)$ be a *random* polynomial system where $p_j(x) := \sum_{|\alpha|=d_j} c_{j,\alpha} \sqrt{\binom{d_j}{\alpha}} x^\alpha$. In particular, recall that $N_j = \binom{n+d_j-1}{d_j}$ and we let $C_j = (c_{j,\alpha})_{|\alpha|=d_j}$ be a random vector in $\mathbb{R}^{N_j}$ satisfying the centering, subgaussian, and small ball assumptions from the introduction. Letting $\mathcal{X}_j := \left(\sqrt{\binom{d_j}{\alpha}} x^\alpha\right)_{|\alpha|=d_j}$ we then have $p_j(x) = \langle C_j, \mathcal{X}_j \rangle$. In particular, recall that the subgaussian assumption is that there is a $K > 0$ such that for each $\theta \in S^{N_j - 1}$ and $t > 0$ we have $\text{Prob}\left(|\langle C_j, \theta\rangle| \geq t\right) \leq 2e^{-t^2/K^2}$. Recall also that the small ball assumption is that there is a $c_0 > 0$ such that for every vector $a \in \mathbb{R}^{N_i}$ and $\varepsilon > 0$ we have $\text{Prob}\left(|\langle a, C_j\rangle| \leq \varepsilon\|a\|_2\right) \leq c_0\varepsilon$. In what follows, several of our bounds will depend on the parameters K and c_0 underlying the random variable being subgaussian and having the small ball property.

For any random variable ξ on $\mathbb{R}$ we denote its median by $\text{Med}(\xi)$. Now, if $\xi := |\langle C_j, \theta\rangle|$, then setting $t := 2K$ in the subgaussian assumption for C_j yields $\text{Prob}(\xi \geq 2K) \leq \frac{1}{2}$, i.e., $\text{Med}(\xi) \leq 2K$. On the other hand, setting $\varepsilon := \frac{1}{2c_0}$ in the small ball assumption for C_j yields $\text{Prob}(\xi \leq \frac{1}{2c_0}) \leq \frac{1}{2}$, i.e., $\text{Med}(\xi) \geq \frac{1}{2c_0}$. Writing $1 = \text{Med}(\xi) \cdot \frac{1}{\text{Med}(\xi)}$ we then easily obtain

$$(1) \qquad Kc_0 \;\geq\; \frac{1}{4}.$$

In what follows we will use Inequality (1) several times.

3.2. **The Subgaussian Assumption and Bounds Related to Operator Norms.** . We will need the following inequality, reminiscent of Hoeffding's classical inequality [18].

Theorem 3.1. [34, Prop. 5.10] *There is an absolute constant $c>0$ with the following property: If $X_1,\ldots,X_n$ are subgaussian random variables with mean zero and underlying constant K, and $a=(a_1,\ldots,a_n)\in\mathbb{R}^n$ and $t\geq 0$, then*

$$\mathrm{Prob}\left(\left|\sum_i a_iX_i\right|\geq t\right)\leq 2\exp\left(\frac{-ct^2}{K^2\|a\|_2^2}\right). \qquad \blacksquare$$

Lemma 3.2. *Let $P:=(p_1,\ldots,p_m)$ be a random polynomial system where, as before, $p_j(x)=\sum_{|\alpha|=d_j}c_{j,\alpha}\sqrt{\binom{d_j}{\alpha}}x^\alpha$ and the the coefficient vectors C_j are independent random vectors satisfying the centering, subgaussian, and small ball assumptions from the introduction, with underlying constants K and c_0, and $m\geq n-1$. Then, for $\mathcal{N}$ a δ-net over S^{n-1} and $t\geq 2$, we have the following inequalities:*

1. *If $\deg(p_j)=d$ for all $j\in\{1,\ldots,m\}$ then*
$$\mathrm{Prob}\left(\|P\|_\infty\leq\frac{2tK\sqrt{m}}{1-d\delta}\right)\geq 1-2|\mathcal{N}|e^{-O(t^2m)}$$
In particular, there is a constant $c_1\geq 1$ such that for $\delta=\frac{1}{3d}$ and $t=s\log(ed)$ with $s\geq 1$ we have $\mathrm{Prob}\left(\|P\|_\infty\leq 3sK\sqrt{m}\log(ed)\right)\geq 1-e^{-c_1s^2m\log(ed)}$.
2. *If $d:=\max_j\deg p_j$ then*
$$\mathrm{Prob}\left(\|P\|_\infty\leq\frac{2tK\sqrt{m}}{1-d^2\delta}\right)\geq 1-2|\mathcal{N}|e^{-O(tm)}$$
In particular, there is a constant $c_2\geq 1$ such that for $\delta=\frac{1}{3d^2}$, $t=s\log(ed)$ with $s\geq 1$, we have $\mathrm{Prob}\left(\|P\|_\infty\leq 3sK\sqrt{m}\log(ed)\right)\geq 1-e^{-c_2s^2m\log(ed)}$.

Proof. We prove Assertion (2) since the proofs of the two assertions are virtually identical. First observe that the identity $(x_1^2+\cdots+x_n^2)^d=\sum_{|\alpha|=d}\binom{d}{\alpha}x^{2\alpha}$ implies $\|\mathcal{X}_j\|_2=1$ for all $j\leq m$. Using our subgaussian assumption on the random vectors C_j, and the fact that $p_j(x)=\langle C_j,\mathcal{X}_j\rangle$, we obtain that $\mathrm{Prob}\left(|p_j(x)|\geq t\right)\leq 2e^{-t^2/K}$ for every $x\in S^{n-1}$.

Now we need to tensorize the preceding inequality. By Theorem 3.1, we have for all $a\in S^{m-1}$ that $\mathrm{Prob}\left(|\langle a,P(x)\rangle|\geq t\right)\leq 2e^{-ct^2/K^2}$. Letting $\mathcal{M}$ be a δ-net on S^{m-1} we then have $\mathrm{Prob}\left(\max_{a\in\mathcal{M}}|\langle a,P(x)\rangle|\geq t\right)\leq 2|\mathcal{M}|e^{-ct^2/K^2}$, where we have used the classical union bound for the multiple events defined by the (finite) δ-net $\mathcal{M}$. Since $\|P(x)\|_2=\max_{\theta\in S^{m-1}}|\langle\theta,P(x)\rangle|$, an application of Lemma 2.5 for the linear polynomial $\langle\,\cdot\,,P(x)\rangle$ gives us $\mathrm{Prob}\left(\|P(x)\|_2\geq\frac{t\sqrt{m}K}{1-\delta}\right)\leq 2|\mathcal{M}|e^{-ct^2m}$.

It is known that for any $\delta>0$, S^{m-1} admits a δ-net $\mathcal{M}$ such that $|\mathcal{M}|\leq\left(\frac{3}{\delta}\right)^m$ (see, e.g, [34, Lemma 5.2]). So for $t\geq 1$ and $\delta=\frac{1}{2}$, using a union bound over the δ-net we have

$$\mathrm{Prob}\left(\|P(x)\|_2\geq 2t\sqrt{m}K\right)\leq 2e^{-c_2t^2m}$$

for some suitable constant $c_2\leq c$. We have thus arrived at a point-wise estimate on $\|P(x)\|_2$. Doing a union bound on a δ-net $\mathcal{N}$ now on S^{n-1} we then obtain:

$$\text{Prob}\left(\max_{x\in\mathcal{N}}\|P(x)\|_2 \geq 2t\sqrt{m}K\right) \leq 2|\mathcal{N}|e^{-c_1t^2m}.$$

Using Lemma 2.5 once again completes our proof. ■

Theorem 2.4 and Lemma 3.2 then directly imply the following:

Corollary 3.3. *Let P be a random polynomial system as in Lemma 3.2. Then there are constants $c_1, c_2 \geq 1$ such that the following inequalities hold for $s \geq 1$:*

1. *If $\deg(p_j) = d$ for all $j \in \{1,\ldots,m\}$ then both $\text{Prob}\left(\|D^{(1)}P\|_\infty \leq 3sK\sqrt{m}d\log(ed)\right)$ and $\text{Prob}\left(\|D^{(2)}P\|_\infty \leq 3sK\sqrt{m}d^2\log(ed)\right)$ are bounded from below by $1-2e^{-c_1s^2m\log(ed)}$.*
2. *If $d := \max_j \deg p_j$ then both $\text{Prob}\left(\|D^{(1)}P\|_\infty \leq 3sK\sqrt{m}d^2\log(ed)\right)$ and $\text{Prob}\left(\|D^{(2)}P\|_\infty \leq 3sK\sqrt{m}d^4\log(ed)\right)$ are bounded from below by $1-2e^{-c_2s^2m\log(ed)}$.* ■

3.3. **The Small Ball Assumption and Bounds for $L(P)$.** We will need the following standard lemma (see, e.g., [27, Lemma 2.2] or [35]).

Lemma 3.4. *Let $\xi_1,\ldots,\xi_m$ be independent random variables such that, for every $\varepsilon > 0$, we have $\text{Prob}(|\xi_i| \leq \varepsilon) \leq c_0\varepsilon$. Then there is a universal constant $\tilde{c} > 0$ such that for every $\varepsilon > 0$ we have $\text{Prob}\left(\sqrt{\xi_1^2+\cdots+\xi_m^2} \leq \varepsilon\sqrt{m}\right) \leq (\tilde{c}c_0\varepsilon)^m$.* ■

We can then derive the following result:

Lemma 3.5. *Let $P=(p_1,\ldots,p_m)$ be a random polynomial system, satisfying the small ball assumption with underlying constant c_0. Then there is a universal constant $\tilde{c} > 0$ such that for every $\varepsilon > 0$ and $x \in S^{n-1}$ we have $\text{Prob}(\|P(x)\|_2 \leq \varepsilon\sqrt{m}) \leq (\tilde{c}c_0\varepsilon)^m$.*

Proof. By the small ball assumption on the random vectors C_i, and observing that $p_i(x) = \langle C_i, \mathcal{X}_i\rangle$ and $\|\mathcal{X}_i\|_2 = 1$ for all $x \in S^{n-1}$, we have $\text{Prob}(|p_i(x)| \leq \varepsilon) \leq c_0\varepsilon$. By Lemma 3.4 we are done. ■

The next lemma is a variant of [26, Claim 2.4]. The motivation for the technical statement below, which introduces new parameters α, β, γ, is that it is the crucial covering estimate needed to prove a central probability bound we'll need later: Theorem 3.7.

Lemma 3.6. *Let $n \geq 2$, let $P := (p_1,\ldots,p_m)$ be a system of n-variate homogenous polynomials, and assume $\|P\|_\infty \leq \gamma$. Let $x, y \in S^{n-1}$ be mutually orthogonal vectors with $\mathcal{L}(x,y) \leq \alpha$, and let $r \in [-1,1]$. Then for every w with $w = x + \beta ry + \beta^2 z$ for some $z \in B_2^n$, we have the following inequalities:*

1. *If $d := \max_i d_i$ and $0 < \beta \leq d^{-4}$ then $\|P(w)\|_2^2 \leq 8(\alpha^2 + (2+e^4)\beta^4d^4\gamma^2)$.*
2. *If $\deg(p_i) = d$ for all $i \in [m]$, and $0 < \beta \leq d^{-2}$ then $\|P(w)\|_2^2 \leq 8(\alpha^2 + (2+e^4)\beta^4d^4\gamma^2)$.*

Proof. We will prove just Assertion (1) since the proof of Assertion (2) is almost the same. We start with some auxiliary observations on $\|P\|_\infty$: First note that Theorem 2.4 tells us that $\|P\|_\infty \leq \gamma$ implies $\|D^{(1)}P\|_\infty \leq d^2\gamma$ and, similarly, $\|D^{(k)}P\|_\infty \leq d^{2k}\gamma$ for every $k \geq 1$. Also, for any w and $u_i \in S^{n-1}$ with $i \in \{1,\ldots,k\}$, $\|P\|_\infty \leq \gamma$ and the homogeneity of the p_i implies $\sup_{u_1,\ldots,u_k}\|D^{(k)}P(w)(u_1,\ldots,u_k)\|_2 \leq \|w\|_2^{d-k}d^{2k}\gamma$. These observations then yield the following inequality for $w = x + \beta ry + \beta^2 z$ with $z \in B_2^n$, $|r| \leq 1$, $\beta \leq d^{-1}$, $k = 3$, and $u_1, u_2, u_3 \in S^{n-1}$:

$$\|D^{(3)}P(w)(u_1,u_2,u_3)\|_2 \leq \|w\|_2^{d-3}d^6\gamma \leq (1+2\beta)^{d-3}d^6\gamma$$

Now, by Taylor expansion, we have the following equality:

$$p_j(w) = p_j(x) + \langle \nabla p_j(x), \beta ry + \beta^2 z\rangle + \frac{1}{2}(\beta ry + \beta^2 z)^T D^{(2)} p_j(x)(\beta ry + \beta^2 z) + (1+\beta)^3 \beta^3 A_j(x),$$

where $A_j(x) := \int_0^1 D^{(3)} p_j(x + t\|v\|_2 v)(v,v,v)dt$ and $v = \frac{\beta ry + \beta^2 z}{\|\beta ry + \beta^2 z\|}$.

Breaking the second and third order terms of the expansion of $p_j(w)$ into pieces, we then have the following inequality:

$$|p_j(w)| \leq |p_j(x)| + \beta|\langle \nabla p_j(x), y\rangle| + \beta^2|\langle \nabla p_j(x), z\rangle| + \frac{1}{2}\beta^2|D^{(2)}p_j(x)(y,y)| + \frac{1}{2}\beta^3|D^{(2)}p_j(x)(y,z)|$$

$$+\frac{1}{2}\beta^3|D^{(2)}p_j(x)(z,y)| + \frac{1}{2}\beta^4|D^{(2)}p_j(x)(z,z)| + (1+\beta)^3\beta^3\,|A_j(x)|\,.$$

Applying the Cauchy-Schwarz Inequality to the vectors $(1, \beta d_j^{\frac{1}{2}}, 1, 1, 1, 1, 1, 1)$ and $(|p_j(x)|, d_j^{-\frac{1}{2}}|\langle \nabla p_j(x), y\rangle|, \ldots, (1+\beta)^3\beta^3\,|A_j(x)|)$ then implies the following inequality:

$$p_j(w)^2 \leq (7 + \beta^2 d_j)(p_j(x)^2 + d_j^{-1}\langle \nabla p_j(x), y\rangle^2 + \beta^4\langle \nabla p_j(x), z\rangle^2 + \frac{1}{4}\beta^4(D_j^{(2)}p_j(x)(y,y))^2$$

$$+\frac{1}{4}\beta^6|D^{(2)}p_j(x)(y,z)|^2 + \frac{1}{4}\beta^6|D^{(2)}p_j(x)(z,y)|^2 + \frac{1}{4}\beta^8|D^{(2)}p_j(x)(z,z)|^2 + \beta^6(1+\beta)^6 A_j(x)^2)$$

We sum all these inequalities for $j \in \{1, \ldots, m\}$. On the left-hand side we have $\|P(w)\|_2^2$. On the right-hand side, the summation of the terms $p_j(x)^2 + d_j^{-1}\langle \nabla p_j(x), y\rangle^2$ is $\|P(x)\|_2^2 + \|M^{-1}D^{(1)}P(x)(y)\|_2^2$, and its magnitude is controlled by the assumption $\mathcal{L}(x,y) \leq \alpha$. The summations of the other terms are controlled by the assumption $\|P\|_\infty \leq \gamma$ and Theorem 2.4. Summing all the inequalities for $j \in \{1, \ldots, m\}$, we have

$$\|P(w)\|_2^2 \leq (7 + \beta^2 d)(\|P(x)\|_2^2 + \|M^{-1}D^{(1)}P(x)(y)\|_2^2 + \beta^4 d^4\gamma^2 + \frac{1}{4}\beta^4 d^4\gamma^2$$

$$+\frac{1}{4}\beta^6 d^6\gamma^2 + \frac{1}{4}\beta^6 d^6\gamma^2 + \frac{1}{4}\beta^8 d^8\gamma^2 + \beta^6(1+\beta)^6\sum_j A_j(x)^2)$$

The assumption $\beta \leq d^{-4}$ implies that $\beta^8 d^8 \leq \beta^4 d^4$ and $\beta^6 d^6 \leq \beta^4 d^4$. Therefore,

$$\|P(w)\|_2^2 \leq (7+\beta^2 d)(\|P(x)\|_2^2 + \|M^{-1}D^{(1)}P(x)(y)\|_2^2 + \beta^4 d^4\gamma^2 + \beta^4 d^4\gamma^2 + \beta^6(1+\beta)^6\sum_j A_j(x)^2).$$

Clearly $\sum_{j\leq m} A_j(x)^2 \leq \max_{w\in V_{x,y}}\|D^{(3)}P(w)(u_1,u_2,u_3)\|_2^2 \leq (1+2\beta)^{2d-6}d^{12}\gamma^2$. Hence we have $\|P(w)\|_2^2 \leq (7+\beta^2 d)(\alpha^2 + \beta^4 d^4\gamma^2 + \beta^4 d^4\gamma^2 + (1+2\beta)^{2d}\beta^6 d^{12}\gamma^2)$. Since $\beta \leq d^{-4}$, we finally get $\|P(w)\|_2^2 \leq (7+\beta^2 d)(\alpha^2 + (2+e^4)\beta^4 d^4\gamma^2) \leq 8(\alpha^2 + (2+e^4)\beta^4 d^4\gamma^2)$. ■

Lemma 3.6 controls the growth of the norm of the polynomial system $P = (p_1, \ldots, p_m)$ over the region $\{w \in \mathbb{R}^n : w = x + \beta ry + \beta^2 z, |r| \leq 1, y \in S^{n-1}, y \perp x, z \in B_2^n\}$. Note in particular that we are using *cylindrical neighborhoods* instead of ball neighborhoods. This is because we have found that (a) our approach truly requires us to go to order 3 in the underlying Taylor expansion and (b) cylindrical neighborhoods allow us to properly take contributions from tangential directions, and thus higher derivatives, into account.

We already had a probabilistic estimate in Lemma 3.5 that said that for any w with $\|w\|_2 \geq 1$, the probability of $\|P(w)\|_2$ being smaller than $\varepsilon\sqrt{m}$ is less than ε^m up to some universal constants. The controlled growth provided by Lemma 3.6 holds for a region with a certain volume, which will ultimately contradict the probabilistic estimates provided by Lemma 3.5. This will be the main trick behind the proof of the following theorem.

Theorem 3.7. *Let $m \geq n-1 \geq 1$ and let $P := (p_1, \ldots, p_m)$ be a system of random homogenous n-variate polynomials such that $p_j(x) = \sum_{|a|=d_j} c_{j,a}\sqrt{\binom{d_i}{a}}x^a$ where $C_j = (c_{j,a})_{|a|=d_j}$ are random vectors satisfying the small ball assumption with underlying constant c_0. Let $\alpha, \gamma > 0$, $d := \max_i d_i$, and assume $\alpha \leq \gamma \min\{d^{-6}, d^2/n\}$. Then*

$$\text{Prob}(L(P) \leq \alpha) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right) + \alpha^{\frac{3}{2}+m-n}\sqrt{n}(\gamma d^2)^{n-\frac{3}{2}}\left(\frac{Cc_0}{\sqrt{m}}\right)^m$$

where C is a universal constant.

Proof. We assume the hypotheses of Assertion (1): Let $\alpha, \gamma > 0$ and $\beta \leq d^{-4}$. Let $\mathbf{B} := \{P \mid \|P\|_\infty \leq \gamma\}$ and let

$$\mathbf{L} := \{P \mid L(P) \leq \alpha\} = \{P \mid \text{There exist } x, y \in S^{n-1} \text{ with } x \perp y \text{ and } \mathcal{L}(x,y) \leq \alpha\}.$$

Let $\Gamma := 8(\alpha^2 + (2+e^4)\beta^4 d^4\gamma^2)$ and let B_2^n denote the unit ℓ_2-ball in $\mathbb{R}^n$. Lemma 3.6 implies that if the event $\mathbf{B} \cap \mathbf{L}$ occurs then there exists a non-empty set

$$V_{x,y} := \{w \in \mathbb{R}^n : w = x + \beta r y + \beta^2 z, x \perp y, |r| \leq 1, z \perp y, z \in B_2^n\} \setminus B_2^n$$

such that $\|P(w)\|_2^2 \leq \Gamma$ for every w in this set. Let $V := \text{Vol}(V_{x,y})$. Note that for $w \in V_{x,y}$ we have $\|w\|_2^2 = \|x + \beta^2 z\|_2^2 + \|\beta y\|_2^2 \leq 1 + 4\beta^2$. Hence we have $\|w\|_2 \leq 1 + 2\beta^2$. Since $V_{x,y} \subseteq (1+2\beta^2)B_2^n \setminus B_2^n$, we have showed that

$$\mathbf{B} \cap \mathbf{L} \subseteq \{P \mid \text{Vol}(\{x \in (1+2\beta^2)B_2^n \setminus B_2^n \mid \|P(x)\|_2^2 \leq \Gamma\}) \geq V\}.$$

Using Markov's Inequality, Fubini's Theorem, and Lemma 3.5, we can estimate the probability of this event. Indeed,

$$\begin{aligned}
&\text{Prob}\left(\text{Vol}(\{x \in (1+2\beta^2)B_2^n \setminus B_2^n : \|P(x)\|_2^2 \leq \Gamma\}) \geq V\right)\\
&\quad\leq \frac{1}{V}\mathbb{E}\text{Vol}\big(\{x \in (1+2\beta^2)B_2^n \setminus B_2^n : \|P(x)\|_2^2 \leq \Gamma\}\big)\\
&\quad\leq \frac{1}{V}\int_{(1+2\beta^2)B_2^n \setminus B_2^n} \text{Prob}\left(\|P(x)\|_2^2 \leq \Gamma\right) dx\\
&\quad\leq \frac{\text{Vol}((1+2\beta^2)B_2^n \setminus B_2^n)}{V} \max_{x \in (1+2\beta^2)B_2^n \setminus B_2^n} \text{Prob}\left(\|P(x)\|_2^2 \leq \Gamma\right).
\end{aligned}$$

Now recall that $\text{Vol}(B_2^n) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}$. Then $\frac{\text{Vol}(B_2^n)}{\text{Vol}(B_2^{n-1})} \leq \frac{c'}{\sqrt{n}}$ for some constant $c' > 0$. If we assume that that $\beta^2 \leq \frac{1}{n}$, then we obtain $(1+2\beta^2)^n \leq 1 + 4n\beta^2$, and we have that

$$\frac{\text{Vol}((1+2\beta^2)B_2^n \setminus B_2^n)}{V} \leq \frac{\text{Vol}(B_2^n)\left((1+2\beta^2)^n - 1\right)}{\beta(\beta^2)^{n-1}\text{Vol}(B_2^{n-1})} \leq c\sqrt{n}\beta\beta^{2-2n},$$

for some absolute constant $c > 0$. Note that here, for a lower bound on V, we used the fact that $V_{x,y}$ contains more than half of a cylinder with base having radius β^2 and height 2β.

Writing $\tilde{x} := \frac{x}{\|x\|_2}$ for any $x \neq 0$ we then obtain, for $z \notin B_2^n$, that

$$\|P(z)\|_2^2 = \sum_{j=1}^m |p_j(z)|^2 = \sum_{j=1}^m |p_j(\tilde{z})|^2 \|z\|_2^{2d_j} \geq \sum_{j=1}^m |p_j(\tilde{z})|^2 = \|P(\tilde{z})\|_2^2.$$

This implies, via Lemma 3.5, that for every $w \in (1+2\beta^2)B_2^n \setminus B_2^n$ we have

$$\text{Prob}\left(\|P(w)\|_2^2 \leq \Gamma\right) \leq \text{Prob}\left(\|P(\tilde{w})\|_2^2 \leq \Gamma\right) \leq \left(cc_0\sqrt{\frac{\Gamma}{m}}\right)^m.$$

So we conclude that

$\text{Prob}(L(P) \leq \alpha) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right) + \text{Prob}\left(\mathbf{B} \cap \mathbf{L}\right) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right) + c\sqrt{n}\beta\beta^{2-2n}\left(cc_0\sqrt{\frac{\Gamma}{m}}\right)^m.$

Recall that $\Gamma = 8(\alpha^2+(5+e^4)\beta^4 d^4\gamma^2)$. Setting $\beta^2 := \frac{\alpha}{\gamma d^2}$, our assumption $\alpha \leq \gamma \min\{d^{-6}, d^2/n\}$ and our choice of β then imply that $\Gamma = C\alpha^2$ for some constant C. So we obtain

$$\text{Prob}(L(P) \leq \alpha) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right) + c\sqrt{n}\left(\frac{\alpha}{\gamma d^2}\right)^{\frac{3}{2}-n}\left(\frac{Cc_0\alpha}{\sqrt{m}}\right)^m$$

and our proof is complete. ■

3.4. **The Condition Number Theorem and its Consequences.** We will now need bounds for the Weyl-Bombieri norms of polynomial systems. Note that, with

$$p_j(x) = \sum_{\alpha_1+\cdots+\alpha_n=d_j} \sqrt{\binom{d_j}{\alpha}} c_{j,\alpha} x^\alpha,$$

we have $\|p_j\|_W := \|(c_{j,\alpha})_\alpha\|_2$ for $j \in \{1,\ldots,m\}$. The following lemma, providing large deviation estimates for the Euclidean norm, is standard and follows, for instance, from Theorem 3.1.

Lemma 3.8. *There is a universal constant $c' > 0$ such that for any random n-variate polynomial system $P = (p_1,\ldots,p_m)$ satisfying the Centering and Subgaussian assumptions, with underlying constant K, $j \in \{1,\ldots,m\}$, $N_j := \binom{n+d_j-1}{d_j}$, $N := \sum_{j=1}^m N_j$, $m \geq n-1$, and $t \geq 1$, we have*

1. Prob $\left(\|p_j\|_W \geq c' t K \sqrt{N_j}\right) \leq e^{-t^2 N_j}$
2. Prob $\left(\|P\|_W \geq c' t K \sqrt{N}\right) \leq e^{-t^2 N}$.

We are now ready to prove our main theorem on the condition number of random polynomial systems.

Theorem 3.9. *There are universal constants $A, C > 0$ such that the following hold: Assume $m \geq n-1 \geq 1$, let $P = (p_1,\ldots,p_m)$ be a system of homogenous random polynomials with $p_j(x) = \sum_{|\alpha|=d_j} c_{j,\alpha}\sqrt{\binom{d_j}{\alpha}} x^\alpha$ where $C_j = (c_{j,\alpha})_{|\alpha|=d_j} \in \mathbb{R}^{N_i}$ are independent random vectors satisfying the subgaussian and small ball assumptions, with respective underlying constants K and c_0. Let $d := \max_i \deg(p_i)$. Then, setting*

$$M := \sqrt{\frac{N}{m}}(Kc_0C)^{\frac{m}{m-n+\frac{3}{2}}}\left(3d^2\log(ed)\right)^{\frac{n-\frac{3}{2}}{m-n+\frac{3}{2}}} n^{\frac{1}{2m-2n+3}} \max\left\{d^6, \frac{n}{d^2}\right\}$$

we have two cases:

(1) *If* $N \geq m\log(ed)$ *then* $\text{Prob}(\tilde{\kappa}(P) \geq tM)$ *is bounded from above by*

$$\begin{cases} \frac{3}{t^{m-n+\frac{3}{2}}} & \textit{if } 1 \leq t \leq e^{\frac{m\log(ed)}{m-n+\frac{3}{2}}} \\ \frac{3}{t^{m-n+\frac{3}{2}}}\left(\frac{(m-n+\frac{3}{2})\log t}{m\log(ed)}\right)^{\frac{n-\frac{3}{2}}{2}} & \textit{if } e^{\frac{m\log(ed)}{m-n+\frac{3}{2}}} \leq t \leq e^{\frac{N}{m-n+\frac{3}{2}}} \\ \frac{3}{t^{m-n+\frac{3}{2}}}\left(\frac{(m-n+\frac{3}{2})\log t}{N}\right)^{\frac{m}{2}}\left(\frac{N}{m\log(ed)}\right)^{\frac{n-\frac{3}{2}}{2}} & \textit{if } e^{\frac{N}{m-n+\frac{3}{2}}} \leq t \end{cases}$$

(2) *If* $N \leq m\log(ed)$ *then* $\text{Prob}(\tilde{\kappa}(P) \geq tM)$ *is bounded from above by*

$$\begin{cases} \frac{3}{t^{m-n+\frac{3}{2}}} & \textit{if } 1 \leq t \leq e^{\frac{N}{m-n+\frac{3}{2}}} \\ \frac{3}{t^{m-n+\frac{3}{2}}}\left(\frac{(m-n+\frac{3}{2})\log t}{N}\right)^{\frac{m}{2}} & \textit{if } e^{\frac{N}{m-n+\frac{3}{2}}} \leq t \end{cases}$$

Proof. Recall that $\tilde{\kappa}(P) = \frac{\|P\|_W}{L(P)}$. Note that if $u > 0$, and the inequalities $\|P\|_W \leq ucK\sqrt{N}$ and $L(P) \geq \frac{ucK\sqrt{N}}{tM}$ hold, then we clearly have $\tilde{\kappa}(P) \leq tM$. In particular, $u > 0$ implies that

$$\text{Prob}\left(\tilde{\kappa}(P) \geq tM\right) \leq \text{Prob}\left(\|P\|_W \geq ucK\sqrt{N}\right) + \text{Prob}\left(L(P) \leq \frac{ucK\sqrt{N}}{tM}\right).$$

Our proof will then reduce to optimizing u over the various domains of t.

Toward this end, note that Lemma 3.8 provides a large deviation estimate for the Weyl norm of our polynomial system. So, to bound $\text{Prob}\left(\|P\|_W \geq ucK\sqrt{N}\right)$ from above, we need to use Lemma 3.8 with the parameter u. As for the other summand in the upper bound for $\text{Prob}\left(\tilde{\kappa}(P) \geq tM\right)$, Theorem 3.7 provides an upper bound for $\text{Prob}\left(L(P) \leq \frac{ucK\sqrt{N}}{tM}\right)$.

However, the upper bound provided by Theorem 3.7 involves the quantity $\text{Prob}\left(\|P\|_\infty \geq \gamma\right)$. Therefore, in order to bound $\text{Prob}\left(L(P) \leq \frac{ucK\sqrt{N}}{tM}\right)$, we will need to use Theorem 3.7 together with Lemma 3.2. In particular, we will set $\alpha := \frac{ucK\sqrt{N}}{tM}$ and $\gamma := 3sK\sqrt{m}\log(ed)$ in Theorem 3.7 and Lemma 3.2, and then optimize the parameters u, s, and t at the final step of the proof.

Now let us check if the assumptions of Theorem 3.7 are satisfied: We have that $s \geq 1$, $u \geq 1$, and (since $\alpha \leq \min\left\{d^{-6}, \frac{d^2}{n}\right\}\gamma$) we have

$$\frac{ucK\sqrt{N}}{tM} \leq 3sK\sqrt{m}\log(ed)\min\left\{d^{-6}, \frac{d^2}{n}\right\}.$$

So $\frac{uc\sqrt{N}}{\sqrt{m}\log(ed)} \leq 3stM\min\left\{d^{-6}, \frac{d^2}{n}\right\}$ and we thus obtain

$$(*) \qquad \frac{uc}{\log(ed)}\sqrt{\frac{N}{m}} \leq 3st\sqrt{\frac{N}{m}}(Kc_0C)^{\frac{m}{m-n+\frac{3}{2}}}(3d^2\log(ed))^{\frac{n-\frac{3}{2}}{m-n+\frac{3}{2}}}n^{\frac{1}{2m-2n+3}}.$$

Since $Kc_o \geq \frac{1}{4}$, the inequality $(*)$ holds if $u \leq s$, $t \geq 1$, and we take the constant C from Theorem 3.7 to be at least 4. Under the preceding restrictions we then have that $Q := \text{Prob}(\tilde{\kappa}(p) \geq tM)$ implies

$$Q \leq \left(\frac{ucK\sqrt{N}}{tM}\right)^{\frac{3}{2}+m-n}\sqrt{n}(3sK\sqrt{m}\log(ed)d^2)^{n-\frac{3}{2}}\left(\frac{Cc_0}{\sqrt{m}}\right)^m + e^{-c_2s^2m\log(ed)} + e^{-u^2N}$$

Note that we set $M:=\sqrt{\frac{N}{m}}(Kc_0C)^{\frac{m}{m-n+\frac{3}{2}}}(3d^2\log(ed))^{\frac{n-\frac{3}{2}}{m-n+\frac{3}{2}}}n^{\frac{1}{2m-2n+3}}\max\left\{d^6,\frac{n}{d^2}\right\}$, therefore we have

$$Q \leq \frac{u^{m-n+\frac{3}{2}}s^{n-\frac{3}{2}}}{t^{m-n+\frac{3}{2}}} + e^{-c_2 s^2 m\log(ed)} + e^{-u^2N}$$

for some suitable $c_2>0$.

We now consider the case where $N \geq m\log(ed)$. If $1 \leq t \leq e^{\frac{c_1 m\log(ed)}{m-n+\frac{3}{2}}}$ then we set $u=s=1$, noting that $(*)$ is satisfied. We then obtain

$$Q \leq \frac{1}{t^{m-n+\frac{3}{2}}} + e^{-c_2 m\log(ed)} + e^{-N} \leq \frac{3}{t^{m-n+\frac{3}{2}}}$$

provided $c_2 \geq 1$.

In the case where $e^{\frac{m\log(ed)}{m-n+\frac{3}{2}}} \leq t \leq e^{\frac{N}{m-n+\frac{3}{2}}}$, we choose $u=1$ and $s:=\sqrt{\frac{(m-n+\frac{3}{2})\log t}{m\log(ed)}} \geq 1$. (Note that $u \leq s$). These choices then yield

$$Q \leq \frac{1}{t^{m-n+\frac{3}{2}}}\left(\frac{(m-n+\frac{3}{2})\log t}{m\log(ed)}\right)^{n-\frac{3}{2}} + \frac{1}{t^{c_2(m-n+\frac{3}{2})}} + e^{-N} \leq \frac{3}{t^{m-n+\frac{3}{2}}}\left(\frac{(m-n+\frac{3}{2})\log t}{m\log(ed)}\right)^{\frac{n-\frac{3}{2}}{2}}.$$

In the case where $e^{\frac{N}{m-n+\frac{3}{2}}} \leq t$, we choose $s:=\sqrt{\frac{(\log t)(m-n+\frac{3}{2})}{m\log(ed)}}$ and $u:=\sqrt{\frac{(m-n+\frac{3}{2})\log t}{N}}$. (Note that $u \leq s$ also in this case). So we get

$$\begin{aligned}Q &\leq \frac{1}{t^{m-n+\frac{3}{2}}}\left(\frac{(m-n+\frac{3}{2})\log t}{N}\right)^{\frac{m}{2}}\left(\frac{N}{m\log(ed)}\right)^{\frac{n-\frac{3}{2}}{2}} + \frac{1}{t^{c_2(m-n+\frac{3}{2})}} + \frac{1}{t^{m-n+\frac{3}{2}}}\\ &\leq \frac{3}{t^{m-n+\frac{3}{2}}}\left(\frac{(m-n+\frac{3}{2})\log t}{N}\right)^{\frac{m}{2}}\left(\frac{N}{m\log(ed)}\right)^{\frac{n-\frac{3}{2}}{2}}.\end{aligned}$$

We consider now the case where $N \leq m\log(ed)$. When $1 \leq t \leq e^{\frac{N}{m-n+\frac{3}{2}}}$ we choose $s=1$ and $u=1$ to obtain $Q \leq \frac{1}{t^{m-n+\frac{3}{2}}} + e^{-c_2 m\log(ed)} + e^{-N} \leq \frac{3}{t^{m-n+\frac{3}{2}}}$ as before. In the case $t \geq e^{\frac{N}{m-n+\frac{3}{2}}}$, we choose $s=u:=\sqrt{\frac{(m-n+\frac{3}{2})\log t}{N}}$. Note that again $(*)$ is satisfied and, with these choices, we get

$$\begin{aligned}Q &\leq \frac{1}{t^{m-n+\frac{3}{2}}}\left(\frac{(m-n+\frac{3}{2})\log t}{N}\right)^{\frac{m}{2}} + \frac{1}{t^{c_2(m-n+\frac{3}{2})m\log(ed)/N}} + \frac{1}{t^{m-n+\frac{3}{2}}}\\ &\leq \frac{3}{t^{m-n+\frac{3}{2}}}\left(\frac{(\log t)(m-n+\frac{3}{2})}{N}\right)^{\frac{m}{2}}.\end{aligned}$$

■

Theorem 3.10. *Let P be a random polynomial system as in Theorem 3.9, let $d:=\max_j \deg p_j$, and let M be as defined in Theorem 3.9. Set*

$$\delta_1 := \frac{q\sqrt{\pi n}}{m-n+\frac{3}{2}}\left(\frac{n-\frac{3}{2}}{2em\log(ed)}\right)^{\frac{n-\frac{3}{2}}{2}}\frac{1}{\left(1-\frac{q}{m-n+\frac{3}{2}}\right)^{\frac{n}{2}}} \qquad \textit{and}$$

$$\delta_2 := \left(\frac{m}{N}\right)^{\frac{m-n+\frac{3}{2}}{2}} \frac{q\sqrt{\pi m}e^{-\frac{m}{2}}}{\left(m-n+\frac{3}{2}-q\right)\left(1-\frac{q}{m-n+\frac{3}{2}}\right)^{\frac{m}{2}}(\log(ed))^{\frac{n}{2}-1}}.$$

We then have the following estimates:

(1) *If $N \geq m\log(ed)$ and $q \in (0, m-n+\frac{3}{2})$ then*

$$(\mathbb{E}(\tilde{\kappa}(P)^q))^{\frac{1}{q}} \leq M\left(1+\frac{q}{m-n-q+2}+\delta_1+\delta_2\right)^{\frac{1}{q}}.$$

In particular, $q \in \left(0, (m-n+\frac{3}{2})\left(1-\frac{1}{2\log(ed)}\right)\right] \Longrightarrow (\mathbb{E}(\tilde{\kappa}(P)^q))^{\frac{1}{q}} \leq M\left(\frac{3m\log(ed)}{n}\right)^{\frac{1}{q}}$, *and* $q \in \left(0, \frac{m-n+\frac{3}{2}}{2}\right] \Longrightarrow (\mathbb{E}(\tilde{\kappa}(P)^q))^{\frac{1}{q}} \leq 4^{1/q}M$.
Furthermore, $\mathbb{E}(\log\tilde{\kappa}(P)) \leq 1+\log M$.

(2) *If $N \leq m\log(ed)$, then $(\mathbb{E}(\tilde{\kappa}(P)^q))^{\frac{1}{q}} \leq M\left(1+\frac{q}{m-n-q+\frac{3}{2}}+\delta_2\right)^{\frac{1}{q}}$.*
In particular, $q \in \left(0, (m-n+\frac{3}{2})\left(1-\frac{m}{eN}\right)\right] \Longrightarrow (\mathbb{E}(\tilde{\kappa}(P)^q))^{\frac{1}{q}} \leq M\left(\frac{3m\log(ed)}{n}\right)^{\frac{1}{q}}$ *and* $q \in \left(0, \frac{m-n+\frac{3}{2}}{2}\right] \Longrightarrow (\mathbb{E}(\tilde{\kappa}(P)^q))^{\frac{1}{q}} \leq 4^{1/q}M$.
Furthermore, $\mathbb{E}(\log\tilde{\kappa}(P)) \leq 1+\log M$.

Proof. Set $\Lambda_1 := \left(\frac{m-n+\frac{3}{2}}{m\log ed}\right)^{\frac{n-\frac{3}{2}}{2}}$, $\Lambda_2 := \left(\frac{m-n+\frac{3}{2}}{N}\right)^{\frac{m}{2}}\left(\frac{N}{m\log ed}\right)^{\frac{n-\frac{3}{2}}{2}}$,

$$r := m-n-q+\frac{5}{2},\ a_1 := \frac{m\log ed}{m-n+\frac{3}{2}},\ \text{and } a_2 := \frac{N}{m-n+\frac{3}{2}}.$$

Note that we have $r \geq 1$ by construction. Using Theorem 3.9 and the formula

$$\mathbb{E}((\tilde{\kappa}(P))^q) = q\int_0^\infty t^{q-1}\text{Prob}\left(\tilde{\kappa}(p) \geq t\right)dt$$

(which follows from the definition of expectation), we have that

$$\mathbb{E}((\tilde{\kappa}(P))^q) \leq M^q\left(1+q\int_1^\infty t^{q-1}\text{Prob}\left(\tilde{\kappa}(p) \geq tM\right)dt\right),$$

or $\frac{\mathbb{E}((\tilde{\kappa}(P))^q)}{M^q} \leq 1+q\int_1^{e^{a_1}}\frac{1}{t^r}dt+q\Lambda_1\int_{e^{a_1}}^{e^{a_2}}\frac{(\log t)^{\frac{n-\frac{3}{2}}{2}}}{t^r}dt+q\Lambda_2\int_{e^{a_2}}^\infty\frac{(\log t)^{\frac{m}{2}}}{t^r}dt$. We will give upper bounds for the last three integrals. First note that

$$q\int_1^{e^{a_1}}\frac{1}{t^r}dt = \frac{q}{r-1}\left(1-e^{(r-1)a_1}\right) \leq \frac{q}{r-1}.$$

Also, we have that

$$q\Lambda_1\int_{e^{a_1}}^{e^{a_2}}\frac{(\log t)^{\frac{n-\frac{3}{2}}{2}}}{t^r}dt = q\Lambda_1\int_{a_1}^{a_2}t^{\frac{n-\frac{3}{2}}{2}}e^{(r-1)t}dt = \frac{q\Lambda_1}{(r-1)^{\frac{n}{2}}}\int_{a_1(r-1)}^{a_2(r-1)}t^{\frac{n-\frac{3}{2}}{2}}e^{-t}dt$$

$$\leq \frac{q\Lambda_1}{(r-1)^{\frac{n}{2}-\frac{1}{4}}}\Gamma\left(\frac{n}{2}-\frac{1}{4}\right) \leq \frac{q\sqrt{\pi n}}{m-n+\frac{3}{2}}\left(\frac{n-\frac{3}{2}}{2em\log(ed)}\right)^{\frac{n}{2}-\frac{3}{4}}\frac{1}{\left(1-\frac{q}{m-n+\frac{3}{2}}\right)^{\frac{n}{2}-\frac{1}{4}}}.$$

Finally, we check that

$$q\Lambda_2\int_{e^{a_2}}^{\infty}\frac{(\log t)^{\frac{m}{2}}}{t^r}dt = q\Lambda_2\int_{a_2}^{\infty}t^{\frac{m}{2}}e^{(r-1)t}dt = \frac{q\Lambda_2}{(r-1)^{\frac{m}{2}+1}}\int_{a_2(r-1)}^{\infty}t^{\frac{m}{2}}e^{-t}dt$$

$$\leq \frac{q\Lambda_2}{(r-1)^{\frac{m}{2}+1}}\Gamma\left(\frac{m}{2}+1\right) \leq \frac{q\sqrt{\pi m}}{(m-n-q+\frac{3}{2})^{\frac{m}{2}+1}}\left(\frac{m(m-n+\frac{3}{2})}{eN}\right)^{\frac{m}{2}}\left(\frac{N}{m\log ed}\right)^{\frac{n}{2}-\frac{3}{4}}$$

$$= \left(\frac{m}{N}\right)^{\frac{m}{2}-\frac{n}{2}+\frac{3}{4}}\frac{1}{\left(1-\frac{q}{m-n+\frac{3}{2}}\right)^{\frac{m}{2}}}\cdot\frac{qe^{-m/2}\sqrt{\pi m}}{m-n+\frac{3}{2}-q}\cdot\frac{1}{(\log(ed))^{\frac{n}{2}-\frac{3}{4}}}.$$

Note that if $q \leq (m-n+\frac{3}{2})\left(1-\frac{1}{2\log(ed)}\right)$ then $\delta_1, \delta_2 \leq 1$.

For the case $N \leq m\log(ed)$, working as before, we get that

$$\frac{\mathbb{E}((\tilde{\kappa}(P))^q)}{M^q} \leq 1 + q\int_1^{e^{a_2}}\frac{1}{t^r}dt + q\Lambda_2\int_{e^{a_2}}^{\infty}\frac{(\log t)^{\frac{m}{2}}}{t^r}dt \leq 1 + \frac{q}{r-1} + \delta_2.$$

In the case $N \leq m\log(ed)$ we have $\delta_2 \leq \frac{\sqrt{\pi m}q}{m-n+\frac{3}{2}}\left(\frac{m}{eN}\right)^{\frac{m}{2}}\frac{1}{\left(1-\frac{q}{m-n+\frac{3}{2}}\right)^{\frac{m}{2}+1}}$. In particular, for this case, it easily follows that $q \leq (m-n+\frac{3}{2})\left(1-\frac{m}{N}\right)$ implies $\delta_2 \leq 1$. ■

Note that if $m = n-1$, $n \geq 3$, and $d \geq 2$, then $N \geq m\log(ed)$ and, in this case, it is easy to check that $(*)$ still holds even if we reduce M by deleting its factor of $\max\left\{d^6, \frac{n}{d^2}\right\}$. So then, for the important case $m = n-1$, our main theorems immediately admit the following refined form:

Corollary 3.11. *There are universal constants $A, c > 0$ such that if P is any random polynomial system as in Theorem 3.9, but with $m = n-1$, $n \geq 3$, $d := \max_j \deg p_j$, $d \geq 2$, and $M := \sqrt{N}(Kc_0C)^{2(n-1)}(3d^2\log(ed))^{2n-3}\sqrt{n}$ instead, then we have:*

$$\mathrm{Prob}(\tilde{\kappa}(P) \geq tM) \leq \begin{cases} 3t^{-\frac{1}{2}} & \text{if } 1 \leq t \leq e^{2(n-1)\log(ed)} \\ 3t^{-\frac{1}{2}}\left(\frac{\log t}{2(n-1)\log(ed)}\right)^{\frac{n-\frac{3}{2}}{2}} & \text{if } e^{2(n-1)\log(ed)} \leq t \leq e^{2N} \\ 3t^{-\frac{1}{2}}\left(\frac{\log t}{2N}\right)^{\frac{1}{4}}\left(\frac{\log t}{2(n-1)\log(ed)}\right)^{\frac{n-\frac{3}{2}}{2}} & \text{if } e^{2N} \leq t \end{cases},$$

and, for all $q \in \left(0, \frac{1}{2} - \frac{1}{4\log(ed)}\right]$, we have $(\mathbb{E}(\tilde{\kappa}(P)^q))^{\frac{1}{q}} \leq Me^{\frac{1}{q}}$.
Furthermore, $\mathbb{E}(\log\tilde{\kappa}(P)) \leq 1 + \log M$. ■

We are now ready to prove Corollary 1.5 from the introduction.

Proof of Corollary 1.5: From Corollary 3.11, Bound (2) follows immediately, and Bound (1) is clearly true for the smaller domain of t. So let us now consider $t = xe^{2(n-1)\log(ed)}$ with $x \geq 1$. Clearly, $\left(\frac{\log t}{2(n-1)\log(ed)}\right)^{\frac{n}{2}-\frac{3}{4}} = \left(1+\frac{\log x}{2(n-1)\log(ed)}\right)^{\frac{n}{2}-\frac{3}{4}}$, and thus $\left(\frac{\log t}{2(n-1)\log(ed)}\right)^{\frac{n}{2}-\frac{3}{4}} < e^{\frac{\log x}{4\log(ed)}} = x^{\frac{1}{4\log(ed)}}$.

Since $x = \frac{t}{e^{2(n-1)\log(ed)}}$ we thus obtain $3t^{-\frac{1}{2}}\left(\frac{\log t}{2(n-1)\log(ed)}\right)^{\frac{n-\frac{3}{2}}{2}} \leq 3t^{-\frac{1}{2}}\left(\frac{t}{e^{2(n-1)\log(ed)}}\right)^{\frac{1}{4\log(ed)}}$. Renormalizing the pair (M, t) (since the M from Corollary 3.11 is larger than the M from Corollary 1.5 by a factor of A), we are done. ■

3.5. **On the Optimality of Condition Number Estimates.** As mentioned in the introduction, to establish a lower bound we need one more assumption on the randomness. For the convenience of the reader, we recall our earlier Euclidean small ball assumption.

(**Euclidean Small Ball**) *There is a constant* $\tilde{c}_0 > 0$ *such that for each* $j \in \{1, \ldots, m\}$ *and* $\varepsilon > 0$ *we have* $\text{Prob}\left(\|C_j\|_2 \leq \varepsilon\sqrt{N_j}\right) \leq (\tilde{c}_0\varepsilon)^{N_j}$.

We will need an extension of Lemma 3.4: Lemma 3.12 below (see also [29, Thm. 1.5 & Cor. 8.6]). Toward this end, for any matrix $T := (t_{i,j})_{1\leq i,j\leq m}$, write $\|T\|_{HS}$ for the *Hilbert-Schmidt norm* of T and $\|T\|_{op}$ for the *operator norm* of T, i.e.,

$$\|T\|_{HS} := \left(\sum_{i,j=1}^{m} t_{i,j}^2\right)^{\frac{1}{2}} \quad \text{and} \quad \|T\|_{op} := \max_{\theta \in S^{n-1}} \|T\theta\|_2.$$

Lemma 3.12. *Let* $\xi_1, \ldots, \xi_m$ *be independent random variables satisfying* $\text{Prob}(\xi_i \leq \varepsilon) \leq c_0\varepsilon$ *for all* $i \in \{1, \ldots, m\}$ *and* $\varepsilon > 0$. *Let* $\xi := (\xi_1, \ldots, \xi_m)$. *Then there is a constant* $c > 0$ *such that for any* $m \times m$ *matrix* T *and* $\varepsilon > 0$ *we have* $\text{Prob}(\|T\xi\|_2 \leq \varepsilon\|T\|_{HS}) \leq (cc_0\varepsilon)^{c\frac{\|T\|_{HS}^2}{\|T\|_{op}^2}}$. ■

Our main lower bound for the condition number is then the following:

Lemma 3.13. *Let* $P = (p_1, \ldots, p_m)$ *be a homogeneous* n*-variate polynomial system with* $d_j = \deg p_j$ *for all* j. *Then* $\tilde{\kappa}(P) \geq \frac{\|P\|_W}{\|P\|_\infty\sqrt{m+1}}$. *Moreover, if* $P := (p_1, \ldots, p_m)$ *is a random polynomial system satisfying our Subgaussian and Euclidean small ball assumptions, with respective underlying constants* K *and* $\tilde{c}_0$, *then we have*

$$\text{Prob}\left(\tilde{\kappa}(P) \leq \varepsilon \frac{\sqrt{N}}{Kmd\log(ed)}\right) \leq (c\tilde{c}_0\varepsilon)^{c'\min\left\{N\frac{\min_j N_j}{\max_j N_j}, md\log(ed)\right\}} \quad \text{and}$$

$$\text{Prob}\left(\tilde{\kappa}(P) \leq \varepsilon \frac{\sqrt{N}}{Km\log(ed)}\right) \leq (c\tilde{c}_0\varepsilon)^{c'm\log(ed)}, \text{ if } d_j = d \text{ for all } j \in \{1, \ldots, m\},$$

where $c, c' > 0$ *are absolute constants. In particular when* $d = d_j$ *for all* $j \in \{1, \ldots, m\}$, *we have* $\mathbb{E}(\tilde{\kappa}(P)) \geq c\frac{\sqrt{N}}{m\log(ed)}$.

Proof. First note that Theorem 2.3 implies that for every $x, y \in S^{n-1}$ we have

$$\|d_j^{-1}D^{(1)}p_j(x)y\|_2^2 \leq \|p_j\|_\infty^2.$$

So we have $\|M^{-1}D^{(1)}P(x)(y)\|_2^2 \leq \sum_{j=1}^m \|p_j\|_\infty^2 \leq m\|P\|_\infty^2$. Now recall that

$$\mathcal{L}^2(x,y) := \|M^{-1}D^{(1)}P(x)(y)\|_2^2 + \|p(x)\|_2^2.$$

So we get $L^2(P) := \min_{x\perp y} \mathcal{L}^2(x,y) \leq (m+1)\|P\|_\infty^2$, which in turn implies that

$$\tilde{\kappa}(P) \geq \frac{\|P\|_W}{L(P)} \geq \frac{\|P\|_W}{\|P\|_\infty\sqrt{m+1}}.$$

The proof for the case where $d_j = d$ for all $j \in \{1, \ldots, m\}$ is identical.

We now show that, under our Euclidean small ball Assumption, we have that $\text{Prob}\left(\|P\|_W \leq \varepsilon\sqrt{N}\right) \leq (c\tilde{c}_0\varepsilon)^{cN\frac{\min_j N_j}{\max_j N_j}}$ for every $\varepsilon \in (0,1)$. Indeed, recall that $\|p_j\|_W = \|C_j\|_{\ell_2^{N_j}}$. Then $\text{Prob}\left(\|p_j\|_W \leq \varepsilon\sqrt{N_j}\right) \leq (\tilde{c}_0\varepsilon)^{N_j} \leq (\tilde{c}_0\varepsilon)^{N_{j_0}}$ for any fixed $\varepsilon \in (0,1)$, where $j_0 \in \{1, \ldots, m\}$ satisfies $N_{j_0} := \min_j N_j$. Let $\xi_j := \frac{\|p_j\|_W}{\sqrt{N_j}}$ for any $j \in \{1, \ldots, m\}$. Set $\xi := (\xi_1, \cdots, \xi_m)$

and $T := \operatorname{diag}(\sqrt{N_1}, \cdots, \sqrt{N_m})$. Note that $\|P\|_W = \|T\xi\|_2$, $\|T\|_{HS} = \sqrt{\sum_{j=1}^m N_j} = \sqrt{N}$, and $\|T\|_{op} := \max_{1\leq j\leq m} \sqrt{N_j}$. Then Lemma 3.12 implies $\operatorname{Prob}\left(\|P\|_W \leq \varepsilon\sqrt{N}\right) \leq (c\tilde{c_0}\varepsilon)^{cN\frac{\min_j N_j}{\max_j N_j}}$. Recall that Lemma 3.2 implies that for every $t \geq 1$ we have

$$\operatorname{Prob}\left(\|p\|_\infty \geq ctK\sqrt{m}\log(ed)\right) \leq e^{-t^2 m \log(ed)}.$$

So using our lower bound estimate for the condition number, we get

$$\operatorname{Prob}\left(\frac{\|P\|_W}{\|P\|_\infty} \geq \frac{c'\varepsilon\sqrt{N}}{tK\sqrt{m}\log(ed)}\right) \leq \operatorname{Prob}\left(\tilde{\kappa}(P) \geq \frac{c\varepsilon\sqrt{N}}{tKmd\log(ed)}\right),$$

$$\operatorname{Prob}\left(\{\|P\|_W \geq c'\varepsilon\sqrt{N}\} \cap \{\|P\|_\infty \leq ctK\sqrt{m}\log(ed)\}\right) \leq \operatorname{Prob}\left(\tilde{\kappa}(P) \geq \frac{c\varepsilon\sqrt{N}}{tKmd\log(ed)}\right),$$

and

$$\operatorname{Prob}\left(\{\|P\|_W \geq c'\varepsilon\sqrt{N}\} \cap \{\|P\|_\infty \leq ctK\sqrt{m}\log(ed)\}\right) \geq 1 - (c\tilde{c_0}\varepsilon)^{cN\frac{\min_j N_j}{\max_j N_j}} - e^{-t^2 m \log(ed)}$$

We may choose $t := \sqrt{\log \frac{1}{\varepsilon}}$ and, by adjusting constants, we get our result. The case where $d_j = d$ for all $j \in \{1, \ldots, m\}$ is similar. The bounds for the expectation follow by integration. ■

Observe that the dominant factor in the very last estimate of Lemma 3.13 is $\sqrt{N}$, which is the normalization coming from the Weyl-Bombieri norm of the polynomial system. So it makes sense to seek the asymptotic behavior of $\frac{\tilde{\kappa}(P)}{\sqrt{N}}$. When $m = n-1$, the upper bounds we get are exponential with respect to n, while the lower bounds are not. But when $m = 2n-3$ and $d = d_j$ for all $j \in \{1, \ldots, m\}$, we have the following upper bound (by Theorem 3.10) and lower bound (by Theorem 3.13):

$$\frac{A_1}{nd\log(ed)} \leq \frac{\mathbb{E}(\tilde{\kappa}(P))}{\sqrt{N}} \leq \frac{A_2 \log ed \max\{d^8, n\}}{\sqrt{n}},$$

where A_1, A_2 are constants depending on (K, c_0). This suggests that our estimates are closer to optimality when m is a constant multiple of n.

Remark 3.14. *There are similarities between our probability tail estimates and the older estimates in the linear case studied in* [28]. *In particular our estimates in the quadratic case $d = 2$, when m is a constant multiple of n, are quite similar to the optimal result (for the linear case) appearing in* [28]. ◇

4. Acknowledgements

We would like to thank the anonymous referees for detailed remarks that greatly helped clarify our paper.

References

[1] Franck Barthe and Alexander Koldobsky, *"Extremal slabs in the cube and the Laplace transform,"* Adv. Math. 174 (2003), pp. 89–114.

[2] Carlos Beltrán and Luis-Miguel 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.

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

[4] Jean Bourgain, *"On the isotropy-constant problem for ψ_2-bodies,"* in *Geometric Aspects of Functional Analysis*, Lecture Notes in Mathematics, vol. 1807, pp. 114–121, Springer Berlin Heidielberg, 2003.

[5] Peter Bürgisser and Felipe Cucker, *"On a problem posed by Steve Smale,"* Annals of Mathematics, pp. 1785–1836, Vol. 174 (2011), no. 3.

[6] Peter Bürgisser and Felipe Cucker, *Condition,* Grundlehren der mathematischen Wissenschaften, no. 349, Springer-Verlag, 2013.

[7] John F. Canny, *The Complexity of Robot Motion Planning,* ACM Doctoral Dissertation Award Series, MIT Press, 1987.

[8] Peter Bürgisser; Felipe Cucker; and Pierre Lairez, *"Computing the homology of basic semialgebraic sets in weakly exponential time,"* Math ArXiV preprint `arXiv:1706.07473` .

[9] D. Castro, Juan San Martín, Luis M. Pardo, *"Systems of Rational Polynomial Equations have Polynomial Size Approximate Zeros on the Average,"* Journal of Complexity 19 (2003), pp. 161–209.

[10] Felipe Cucker, *"Approximate zeros and condition numbers,"* Journal of Complexity 15 (1999), no. 2, pp. 214–226.

[11] Felipe Cucker, Teresa Krick, Gregorio Malajovich, and Mario Wschebor, *"A numerical algorithm for zero counting I. Complexity and accuracy,"* J. Complexity 24 (2008), no. 5–6, pp. 582–605

[12] Felipe Cucker, Teresa Krick, Gregorio Malajovich, and Mario Wschebor, *"A numerical algorithm for zero counting II. Distance to ill-posedness and smoothed analysis,"* J. Fixed Point Theory Appl. 6 (2009), no. 2, pp. 285–294.

[13] Felipe Cucker, Teresa Krick, Gregorio Malajovich, and Mario Wschebor, *"A numerical algorithm for zero counting III: Randomization and condition,"* Adv. in Appl. Math. 48 (2012), no. 1, pp. 215–248.

[14] Felipe Cucker; Teresa Krick; and Mike Shub, *"Computing the Homology of Real Projective Sets,"* Found. Comp. Math., to appear. (Earlier version available as Math ArXiV preprint `arXiv:1602.02094` .)

[15] Nikos Dafnis and Grigoris Paouris, *"Small ball probability estimates, Ψ_2-behavior and the hyperplane conjecture,"* Journal of Functional Analysis 258 (2010), pp. 1933–1964.

[16] Jean-Piere Dedieu, Mike Shub, *"Newton's Method for Overdetermined Systems Of Equations,"* Math. Comp. 69 (2000), no. 231, pp. 1099–1115.

[17] James Demmel, Benjamin Diament, and Gregorio Malajovich, *"On the Complexity of Computing Error Bounds,"* Found. Comput. Math. pp. 101–125 (2001).

[18] Wassily Hoeffding, *"Probability inequalities for sums of bounded random variables,"* Journal of the American Statistical Association, 58 (301):13–30, 1963.

[19] O. D. Kellog, *"On bounded polynomials in several variables,"* Mathematische Zeitschrift, December 1928, Volume 27, Issue 1, pp. 55–64.

[20] Bo'az Klartag and Emanuel Milman, *"Centroid bodies and the logarithmic Laplace Transform – a unified approach,"* J. Func. Anal., 262(1):10–34, 2012.

[21] Alexander Koldobsky and Alain Pajor, *"A Remark on Measures of Sections of L_p-balls,"* Geometric Aspects of Functional Analysis, Lecture Notes in Mathematics 2169, pp. 213–220, Springer-Verlag, 2017.

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

[23] Pierre Lairez, *"A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time,"* Foundations of Computational Mathematics, DOI 10.1007/s10208-016-9319-7.

[24] G. Livshyts, Grigoris Paouris and P. Pivovarov, *"Sharp bounds for marginal densities of product measures,"* Israel Journal of Mathematics, Vol. 216, Issue 2, pp. 877–889.

[25] G. Paouris and P. Pivovarov, *"Randomized Isoperimetric Inequalities"*, IMA Volume "Discrete Structures: Analysis and Applications" (Springer)

[26] Hoi H. Nguyen, *"On a condition number of general random polynomial systems,"* Mathematics of Computation (2016) 85, pp. 737–757

[27] Mark Rudelson and Roman Vershynin, *"The Littlewood-Offord Problem and Invertibility of Random Matrices,"* Adv. Math. 218 (2008), no. 2, pp. 600–633.

[28] Mark Rudelson and Roman Vershynin, *"The Smallest Singular Value of Rectangular Matrix,"* Communications on Pure and Applied Mathematics 62 (2009), pp. 1707–1739.

[29] Mark Rudelson and Roman Vershynin, *"Small ball Probabilities for Linear Images of High-Dimensional Distributions,"* Int. Math. Res. Not. (2015), no. 19, pp. 9594–9617.

[30] Roman Vershynin , *"High Dimensional Probability: An Introduction with Application in Data Science"*, available at https://www.math.uci.edu/ rvershyn/papers/HDP-book/HDP-book.pdf

[31] Roman Vershynin , *"Four Lectures on Probabilistic Methods for Data Science"*, available at https://www.math.uci.edu/ rvershyn/papers/four-lectures-probability-data.pdf

[32] Igor R. Shafarevich, *Basic Algebraic Geometry 1: Varieties in Projective Space*, 3rd edition, Springer-Verlag (2013).

[33] Mike Shub and Steve Smale, *"Complexity of Bezout's Theorem I. Geometric Aspects,"* J. Amer. Math. Soc. 6 (1993), no. 2, pp. 459–501.

[34] Roman Vershynin, *"Introduction to the Non-Asymptotic Analysis of Random Matrices,"* Compressed sensing, pp. 210–268, Cambridge Univ. Press, Cambridge, 2012.

[35] Assaf Naor and Artem Zvavitch, *"Isomorphic embedding of ℓ_p^n, $1 < p < 2$, into $\ell_1^{(1+\varepsilon)n}$,"* Israel J. Math. **122** (2001), pp. 371–380.

Technische Universität Berlin, Institut für Mathematik, Sekretariat MA 3-2, Strae des 17. Juni 136, 10623, Berlin, Germany

E-mail address: erguer@math.tu-berlin.de

Department of Mathematics, Texas A&M University TAMU 3368, College Station, Texas 77843-3368, USA.

E-mail address: grigoris@math.tamu.edu

Department of Mathematics, Texas A&M University TAMU 3368, College Station, Texas 77843-3368, USA.

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