PROBABILISTIC CONDITION NUMBER ESTIMATES FOR REAL POLYNOMIAL SYSTEMS II: STRUCTURE AND SMOOTHED ANALYSIS

ALPEREN A. ERGÜR, GRIGORIS PAOURIS, AND J. MAURICE ROJAS

ABSTRACT. We consider the sensitivity of real zeros of polynomial systems with respect to perturbation of the coefficients, and extend our earlier probabilistic estimates for the condition number in two directions: (1) We give refined bounds for the condition number of random *structured* polynomial systems, depending on a variant of sparsity and an intrinsic geometric quantity called *dispersion*. (2) Given any structured polynomial system P, we prove the existence of a nearby well-conditioned structured polynomial system Q, with explicit quantitative estimates.

Our underlying notion of structure is to consider a linear subspace E_i of the space H_{d_i} of homogeneous n-variate polynomials of degree d_i, let our polynomial system P be an element of $E := E_1 \times \cdots \times E_{n-1}$, and let $\dim(E) := \dim(E_1) + \cdots + \dim(E_{n-1})$ be our measure of sparsity. The dispersion $\sigma(E)$ provides a rough measure of how suitable the tuple E is for numerical solving.

Part I of this series studied how to extend probabilistic estimates of a condition number defined by Cucker to a family of measures going beyond the weighted Gaussians often considered in the current literature. We continue at this level of generality, using tools from geometric functional analysis.

1. Introduction

Numerous problems in mathematical modeling reduce to finding a distinguished optimum state of a continuously evolving dynamical system. These type of problems can often be translated into finding the roots of a system of polynomial equations in many variables. Just to name a few examples: The computation of steady states in chemical reaction networks [23], Nash equilibria in economics [McL05], and the determination of protein structures from metric data [1] all lie in this framework. Two frequent features of polynomial systems arising this way are (1) the polynomial systems are highly structured and (2) practitioners are mainly interested in the real solutions of these systems instead of the (non-real) complex solutions.

From a computational complexity point of view, deciding existence of a complex root of a system of complex polynomials is already NP-Hard. From the point of view of geometry, Bertini's Theorem and Bezout's Theorem show that a generic (square) system of homogenous polynomials $P = (p_1, \ldots, p_{n-1})$ with degrees $d_1, \ldots, d_{n-1}$ has exactly $d_1 \cdots d_{n-1}$ many complex solutions. This brings us to a problem posed by Steve Smale: The 17th problem in Steve Smale's list for 21st century mathematicians asks for an efficient numerical algorithm to find an approximation of a single complex root [43]. This problem is now solved after two decades of intensive research [3, 6, 10, 29]: there are now algorithms that can find a single complex approximate root of a system of homogenous polynomials $P = (p_1, \ldots, p_{n-1})$ in average-case polynomial time, and the notion of "approximate" implies a kind of guaranteed

2010 *Mathematics Subject Classification.* Primary 65Y20; Secondary 51F99, 68Q25.
Key words and phrases. Condition number, epsilon net, probabilistic bound, kappa, real-solving, structured polynomials, sub-gaussian .
A.E. was partially supported by Einstein Foundation, Berlin. G.P. was partially supported by Simons Foundation Collaboration grant 527498 and NSF grant DMS-1812240. J.M.R. was partially supported by NSF grants CCF-1409020 and DMS-1460766.

fast convergence with respect to Newton iteration. Also, the underlying input size is taken to be $N = \sum_{i=1}^{n-1} \binom{n+d_i-1}{d_i}$ (the underlying number of monomial terms), and average-case polynomial time means a polynomial upper bound on the expectation of the complexity of the algorithm, assuming a particular model of randomness: The random polynomial systems considered have independent Gaussian random coefficients with specially chosen variances. To be more precise, the state of the art is represented by Lairez's recent article [30], which proves an expected complexity bound of $O(n^6 d^4 N)$, under the preceding randomness model, where $d := \max_i d_i$. The only drawback of these elegant results is that, in practice, the input size is often much smaller than N.

In a different direction, there have been polyhedral homotopy algorithms specifically tailored for solving sparse polynomial systems since the early 90's [46, 22]. The polyhedral homotopy method is implemented in PHCpack and Hom4ps-3 and has been tested on many practical problems (see, e.g., [45, 11]). However, to the best of our knowledge, general and explicit average-case complexity bounds do not yet exist for polyhedral homotopy. Recently, Malajovich has developed a mathematically rigorous toric homotopy iteration for sparse polynomial system solving [34], yielding a promising theoretical framework. Unfortunately, his theory does not yet include an algorithm that is both implementable and provably fast.

A central question behind the complexity of all the algorithms mentioned above is estimating the minimal distance between roots of the underlying polynomial system. This root spacing question is now known to be equivalent to the following question: Given a generic input polynomial system (i.e., a polynomial system with $d_1 \cdots d_{n-1}$ many isolated roots), what is the distance of this generic system to the closest degenerate polynomial system? Both of these questions are captured by the mathematically elegant notion of condition number [7].

The numerical methods discussed in the preceding paragraphs aim to solve generic polynomial systems over the field of complex numbers. In the case of real solutions, generic behavior is replaced by multiple possible typical behaviors. For instance, a small perturbation in the coefficients can change a system from having no real root at all to having many real roots. Luckily, a condition number theory that captures these subtleties was developed by Cucker in [12]. Later, this theory was applied in the design and analysis of a numerical algorithm for real root finding [13, 14, 15]. This condition number also appears in recent papers on numerically computing homology groups of semialgebraic sets [9, 16].

The papers that culminated in the solution of Smale's 17th problem [3, 6, 10, 29], and the series of papers [13, 14, 15] on real root solving, analyzed the condition number of random polynomial systems defined by $N = \sum_{i=1}^{n-1} \binom{n+d_i-1}{d_i}$ independent Gaussian coefficients with specially chosen variances. The specially chosen variances induce a unitarily invariant measure on the coefficient space, and this invariance property is heavily used in the complexity analysis. Structured polynomial systems, however, often form a much smaller space which is not closed under a unitary action on the variables. So, to enable the analysis of numerical algorithms for structured polynomial systems one has to drop the unitary group invariance assumption on the underlying probability measure.

We use techniques coming from asymptotic geometric analysis and high-dimensional probability ([47], [49]) that have been applied very successfully to the non-asymptotic theory of random matrices — a "linear version" of our problem. These techniques have the advantage of allowing probability measures much more general than Gaussians: In Part I [20] of our present work, we analyzed the real condition number of Cucker for a broad family of measures, without any invariance assumptions. In this current paper, we apply our techniques to derive condition number estimates for structured random real polynomial systems. We then derive smoothed analysis type estimates analyzing the change in the condition number under structure preserving random perturbations.

1.1. **The Real Condition Number and Analysis of Algorithms.** In this section we present the real condition number of Cucker and comment on it's relation to analysis of numerical algorithms in real algebraic geometry.

Definition 1.1. *Given $n, d_1, \ldots, d_{n-1} \in \mathbb{N}$ and $i \in \{1, \ldots, n-1\}$, let $p_i \in \mathbb{R}[x_1, \ldots, x_n]$ be a homogenous polynomial with* $\deg p_i = d_i$*, and let $P := (p_1, \ldots, p_{n-1})$ 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,*

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

Let Δ_{n-1} be the diagonal matrix with diagonal entries $\sqrt{d_1}, \ldots, \sqrt{d_{n-1}}$ 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, we define the (normalized) local condition number (for solving $P = \mathbf{0}$) *to be $\tilde{\mu}_{\rm norm}(P, x) := \|P\|_W \left\|DP(x)|_{T_xS^{n-1}}^{-1}\Delta_{n-1}\right\|$ or $\tilde{\mu}_{\rm norm}(P, x) := \infty$, according as $DP(x)|_{T_xS^{n-1}}$ has full rank or not, where $\|\cdot\|$ denotes the operator norm of a matrix A.* $\diamond$

Cucker's condition number definition from [12] is the following:

Definition 1.2. [12] *Let* $\tilde{\kappa}(P, x) := \frac{\|P\|_W}{\sqrt{\|P\|_W^2 \tilde{\mu}_{\rm norm}(P,x)^{-2} + \|P(x)\|_2^2}}$ *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.* $\diamond$

It turns out that the worst case analysis of the numerical algorithm for real root counting in [13], and the worst case analysis of numerical algorithms for computing homology groups in [9, 16], can be analyzed using only two parameters: the condition number of the input polynomial system and the evaluation complexity of the input polynomial system. For instance, in [13] the authors assume the input is a sparse polynomial system with total number of monomials S, and they present an analysis of their algorithm based on $\kappa(P)$ (an earlier variant of $\tilde{\kappa}(P)$) and S. Even though the analysis in [9, 16] is conducted for dense polynomials with evaluation complexity linear in $N = \sum_{i=1}^{n-1} \binom{n+d_i-1}{d_i}$, it can also be modified for sparse polynomial systems with lower evaluation complexity.

A first step toward going beyond worst-case complexity analysis is average-case analysis of algorithms. The authors of [13, 14, 15] and [9, 16] conducted an average-case analysis of their algorithms using a specially chosen Gaussian measure which is invariant under an orthogonal group action over the variables. As explained before, this specially chosen Gaussian measure creates an obstacle toward conducting average analysis of algorithms on structured polynomial systems. The main result of this article at last enables a more general average-case analysis of these numerical algorithms, *including structured inputs*.

Smoothed analysis of algorithms, as conceived by Spielman and Teng [44], can be considered as a common generalization of average-case and worst-case complexity analysis. The idea of smoothed analysis is to draw an ε-ball around every point in the input space, then conduct an average analysis inside all ε-balls, and then take the supremum over all these averages as a complexity measure. Our main results also include a smoothed analysis of $\tilde{\kappa}$ for structured polynomial systems (i.e., the balls are drawn inside the structured linear subspace not in the space of dense polynomials). To the best of our knowledge our results are the first ones to address smoothed analysis of condition number for randomly perturbed structured polynomials.

1.2. **Linear Structure in Spaces of Polynomial Systems.** In this article we are concerned with polynomials that have a simple structure: polynomials that satisfy a given set of linear relations. More precisely, let $H_{d,n}$ denote the vector space consisting of all real homogenous degree d polynomials with n variables (along with the zero polynomial), and let $E' \subseteq H_d$ be a linear subspace. When the number of variables is clear from the context, we will often write H_d instead of $H_{d,n}$. We call E' *full* (in $H_{d,n}$) if for every $v \in S^{n-1}$ the pointwise evaluation map $l_v : E' \longrightarrow \mathbb{R}$, $l_v(p) = p(v)$ is not identically zero. For instance, if $E' \subsetneqq H_{2,3}$ is the subspace of consisting of homogeneous polynomials lying in $\mathbb{R}[x_1, x_2]$, then this particular E' is *not* full in $H_{2,3}$ (since each polynomial in this E' vanishes on $(0, 0, 1)$).

Fullness is a geometric property, and it can be checked by an optimization procedure once an orthonormal basis for the linear space is fixed. We discuss this further in Section 4.

We generalize to the case of polynomial systems in a straightforward manner: Let $D = (d_1, \ldots, d_{n-1})$ be a vector with positive integer coordinates, and let H_D denote the vector space of homogenous real polynomial systems $P = (p_1, \ldots, p_{n-1})$ with $\deg(p_i) = d_i$. Let $E_i \subseteq H_{d_i}$ be linear subspaces and let $E \subseteq H_D$ be defined by $E := (E_1, \ldots, E_{n-1})$. We say E is full if the E_i are full for all i.

One may wonder if the fullness assumption on our linear structures is really necessary for probabilistic analysis. An easy example to consider is the following: Let us pick $u, v \in S^{n-1}$ so that $u \perp v$, and define the following subspaces:

$$E_i := \{p \in H_{d_i} : p(u) = \langle \nabla p(u), v\rangle = 0\}$$

where $\nabla p(u)$ denotes the the gradient of p evaluated at u. Note that these particular E_i are codimension 2 linear subspaces of H_{d_i}. Set $E := (E_1, \ldots, E_{n-1})$. By construction, for any polynomial system $P = (p_1, \ldots, p_{n-1})$ with $p_i \in E_i$, P has a singularity at u. Hence, for all $P \in E$, we have that $\tilde{\kappa}(P)$ is infinite, and a probabilistic analysis of $\tilde{\kappa}$ on this linear space E is not meaningful. On the other hand, we have the following fact on full linear subspaces.

Lemma 1.3. *Let $E \subseteq H_D$ be a full linear subspace. Then for a random polynomial system $P \in E$ with the general model of randomness described in Section 1.3, we have that $\tilde{\kappa}(P) < \infty$ with probability 1.*

We emphasize that the randomness model we consider, detailed fully in the next section, is far more general than the restricted Gaussian models considered earlier, and allows many non-Gaussian probability distributions as well. Lemma 1.3 will in fact be an easy consequence of our main condition number bound: Theorem 1.5 of the next section.

The main results of this paper include two quantities related to the linear structure E: $\dim(E)$ and $\sigma(E)$. The quantity $\dim(E)$ replaces the dimension count N in our earlier bounds from [20], and $\sigma(E)$ is a new quantity — the *dispersion constant for E* (in Section 2 below) — related to the geometry of E. We explore the basic properties of the dispersion constant in Section 4. For instance, our definition will immediately imply $\sigma(E) \geq 1$ and $\sigma(H_D) = 1$. Moreover, we also show for linear subspaces $E \subseteq H_D$ and $F \subseteq H_D$ that, if one can find $T \in O(n)^{n-1}$ such that $T \circ F = E$ (using the obvious $O(n)^{n-1}$ action) then $\sigma(F) = \sigma(E)$. So $\sigma(E)$ is indeed a geometric quantity independent of the underlying basis representation.

Our estimates profit when E is low-dimensional and suffer when $\sigma(E)$ is large. So we investigate how large $\sigma(E)$ can typically be as a function of n and other parameters. Our first theorem below shows that for linear spaces with $\dim(E) > n^2 \log(n) \log(ed)^2$, the quantity $\sigma(E)$ typically admits a *constant* upper bound.

Theorem 1.4. *There are constants $c_1, c_2 > 0$ with the following property: Let $n \geq 3$ and $d_1, \ldots, d_{n-1} \geq 2$ be integers, $E_i \subseteq H_{d_i}$ random linear m dimensional subspaces drawn using Haar measure on the Grassmannian $Gr(m, \dim(H_{d_i}))$ for $1 \leq i \leq n-1$, set $E := (E_1, \ldots, E_{n-1})$, and let $d := \max_i d_i$. Then for any $m \geq c_1 n \log^2(ed) \log n$ and $t \geq 1$ we have*

$$\sigma(E) \leq \frac{\sqrt{m} + c_2 t \sqrt{n \log n} \log(ed)}{\sqrt{m} - c_2 t \sqrt{n \log n} \log(ed)}$$

with probability greater than $1 - \frac{1}{n^{t^2-1}}$.

One should first note that the linear spaces are more general than, say, polynomials with a fixed number of terms in the standard monomial basis: We do *not* restrict to any basis in Theorem 1.4, but rather study linear subspaces of H_D geometrically. Theorem 1.5 is in fact a reflection of a concentration of measure phenomenon on the Grassmanian of linear spaces in H_D (see, e.g., [31]): $\sigma(E)$ is a function of random variables having the property that the graph of σ clusters around the graph of a constant function with high probability.

For numerical algorithms operating over the sphere S^{n-1} involving a structured input from a linear space E, $\sigma(E)$ can be considered as the "condition number of the structure". We discuss this in Section 4, and indicate ways to compute $\sigma(E)$. Theorem 1.4 shows that a linear space $E \subset H_D$ with $\dim(E) = \Omega(n^2 \log^2(ed) \log n)$ is typically well-conditioned for deploying numerical algorithms in a structure-preserving way. Simply put, some families of polynomial systems are better-suited for numerical algorithms than others, and Theorem 1.4 gives us a way to make this precise.

1.3. **Randomness Assumptions.** We say a random vector $X \in \mathbb{R}^n$ satisfies the *Centering*, *Sub-Gaussian*, and *Small Ball* properties, with constants K and c_0, if the following hold true:

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

We note that these three assumptions directly yield a relation between K and c_0: We in fact have $Kc_0 \geq \frac{1}{4}$ (see [20, Inequality (1), just before Sec. 3.2]).

Random vectors that satisfy these three properties form a large family of distributions, including standard Gaussian vectors and uniform measures on a large family of convex bodies called Ψ_2-bodies, such as uniform measures on l_p-balls for all $p \geq 2$. We refer the reader to the upcoming book of Vershynin [47] for more details. Discrete sub-Gaussian distributions, such as the Bernoulli distribution, also satisfy an inequality similar to the small ball inequality in our assumptions. However, the small ball type inequality satisfied by such discrete distributions depends not only on the norm of the deterministic vector a but also on the arithmetic structure of a. It is possible that our methods combined with the work of Rudelson and Veshynin in the Littlewood-Offord problem [38] can extend our main results to discrete distributions such as Bernoulli. In this work, we will content ourselves with continuous distributions.

The examples of random vectors from the preceding paragraph do not necessarily have independent coordinates and this provides important extra flexibility. There are also interesting examples of random vectors with independent coordinates. In particular, if $X_1, \ldots, X_m$

are independent centered random variables that all satisfy the Sub-Gaussian inequality with constant K and the Small Ball condition with c_o, then the random vector $X = (X_1, \dots, X_m)$ also satisfies the Sub-Gaussian and Small Ball inequalities with constants C_1K and C_2c_0, where C_1 and C_2 are universal constants. This is a relatively new result of Rudelson and Vershynin [40]. The best possible universal constant C_2 is discussed in [33, 37]. To create a random variable satisfying the Small Ball and Sub-Gaussian properties one can, for instance, start by fixing any $p \geq 2$ and then considering a random variable with density function $f(t) := c_p e^{-|t|^p}$ for suitably chosen positive c_p.

1.4. **Main Results.** We consider the linear structure E to be given and assume that we can find a basis for E orthonormal with respect to the Weyl-Bombieri inner product (defined in Section 2 below).

Theorem 1.5. *Let $D = (d_1, \dots, d_{n-1})$ be a vector with positive integer coordinates, let $E_i \subseteq H_{d_i}$ be full linear subspaces, and let $E = (E_1, \dots, E_{n-1})$. Assume $\dim(E) \geq n\log(ed)$ and $n \geq 3$. Let $p_i \in E_i$ be independent random elements of E_i that satisfy the Centering property, the Sub-Gaussian property with constant K, and the Small Ball property with constant c_0 all with respect to the Bombieri-Weyl inner product. We set $d := \max_i d_i$, and $M := nK\sqrt{\dim(E)}(Cc_0Kd^2\log(ed)\sigma(E))^{2n-2}$, where $C \geq 4$ is a universal constant. Then for the random polynomial system $P = (p_1, \dots, p_{n-1})$, we have*

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

Moreover, for $0 < q < \frac{1}{2} - \frac{1}{2\log(ed)}$, we have $\mathbb{E}(\tilde{\kappa}(P)^q) \leq M^q(1 + 4q\log(ed))$. In particular, $\mathbb{E}\log(\tilde{\kappa}(P)) \leq 1 + \log M$.

Following this condition number estimate, our next goal is a smoothed-analysis type result for $\tilde{\kappa}$. For this we will need a slightly stronger assumption on the random input. This slightly stronger property is called the *Anti-Concentration Property* and it replaces the Small Ball assumption in our model of randomness. We will need a bit of terminology to define anti-concentration.

Definition 1.6 (Concentration Function)**.** *For any real-valued random variable Z and $t \geq 0$, the* concentration function*, $F(Z,t)$, is defined as $F(Z,t) := \max_{u\in\mathbb{R}} \mathrm{Prob}\{|Z-u| \leq t\}$. Let $\langle\cdot,\cdot\rangle$ denote the standard inner product on $\mathbb{R}^n$. We then say a random vector $X \in \mathbb{R}^n$ satisfies the* Anti-Concentration Property with constant c_0 *if we have $F(\langle X, \theta\rangle, \varepsilon) \leq c_0\varepsilon$ for all $\theta \in S^{n-1}$.*

It is easy to check that if the random variable Z has bounded density f then $F(Z,t) \leq \|f\|_\infty t$. Moreover, the Lebesque Differentiation theorem states that upper bounds for the function $t^{-1}F(Z,t)$ for all t imply upper bounds for $\|f\|_\infty$. See [39] for the details.

Theorem 1.7. *Let $E \subset H_D$ be a full linear subspace, and $Q \in E$ a fixed (deterministic) polynomial system. Assume $\dim(E) \geq n\log(ed)^2$, and $n \geq 3$. Now let $G \in E$ be a random polynomial system given by the same model of randomness as in Theorem 1.5, but with the Small Ball Property replaced by the Anti-Concentration Property. Set $d := \max_i d_i$, and $M := nK\sqrt{\dim(E)}\,(Cc_oKd^2\log(ed)\sigma(E))^{2n-2}\left(1 + \frac{\|Q\|_W}{\sqrt{n}K\log(ed)}\right)^{2n-1}$, where $C \geq 4$ is a universal constant. Then for the randomly perturbed polynomial system $P = Q + G$, we have*

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

Moreover, for $0 < q < \frac{1}{2} - \frac{1}{2\log(ed)}$*, we have* $\mathbb{E}(\tilde{\kappa}(P)^q) \leq M^q(1 + 4q\log(ed))$*. In particular,* $\mathbb{E}\log(\tilde{\kappa}(P)) \leq 1 + \log M$*.*

We prove a stronger version of Theorem 1.7 in Section 3.4: See Theorem 3.12 and Remark 3.14 there. As a corollary of this smoothed-analysis theorem we derive the following structural result.

Theorem 1.8. *Let* $E_i \subseteq H_{d_i}$ *be full linear subspaces, let* $E = (E_1, \ldots, E_{n-1})$*, and let* $Q \in E$*. Then, for every* $0 < \varepsilon < 1$*, there is a polynomial system* $P_\varepsilon \in E$ *with the following properties:*

$$\|P_\varepsilon - Q\|_W \leq \varepsilon \|Q\|_W \left(\frac{\sqrt{\dim(E)}}{\log(ed)\sqrt{n}} \right)$$

and

$$\tilde{\kappa}(P_\varepsilon) \leq \sqrt{n}\sqrt{\dim(E)} \left(\frac{Cd^2 \log(ed)\sigma(E)}{\varepsilon} \right)^{2n-2}$$

for a universal constant C*.*

One can view this result as a metric entropy type statement as follows: Suppose we are given a bounded set $\mathbb{T} \subset E$ with $\sup_{P \in T} \|P\|_W \leq 1$, and we would like to cover $\mathbb{T}$ with balls of radius δ, i.e., $\mathbb{T} = \bigcup_i B(p_i, \delta)$. Moreover, suppose we want the ball-centers p_i to have a controlled condition number. We can start with an arbitrary $\frac{\delta}{2}$ covering of $\mathbb{T} = \bigcup_i B(q_i, \frac{\delta}{2})$, and use Theorem 1.8 with $\varepsilon = \frac{\delta\sqrt{n}}{2\sqrt{\dim(E)}}$ to find a p_i with controlled condition number in each one of the balls $B(q_i, \frac{\delta}{2})$. Then $\mathbb{T} = \bigcup_i B(p_i, \delta)$ gives a δ-covering of $\mathbb{T}$ where p_i has controlled condition number.

Remark 1.9. *In the literature on random polynomials, it is customary to consider a model of randomness expressed in terms of a fixed basis and random coefficients. To create such a model of randomness one can consider the following: Let* $D = (d_1, \ldots, d_{n-1})$ *be a vector with positive integer coordinates, assume that* $E_i \subseteq H_{d_i}$ *are full linear subspaces with* $\dim(E_i) = t_i$*, and let* $E = (E_1, \ldots, E_{n-1})$*. Suppose for each* i *that* $\{u_{ij}\}_{j=1}^{t_i}$ *is an orthonormal basis for the linear space* E_i *with respect to the Weyl-Bombieri inner product. Let* $C_i \in \mathbb{R}^{t_i}$ *be independent random vectors satisfying the Centering, Sub-Gaussian, and Small Ball properties with constants* K *and* c_0*. Consider the random polynomials*

$$p_i = \sum_{j=1}^{t_i} c_{ij} u_{ij}(x)$$

where c_{ij} *is the* j*th coordinate of* C_i*. Then, the random polynomial system* $P := (p_1, \ldots, p_{n-1})$ *satisfies the three assumptions in Theorem 1.5. Similarly, if we replace the Small Ball assumption with the Anti-Concentration Property then the assumptions of Theorem 1.7 are satisified by* P*.* $\diamond$

2. Preliminaries

We start by introducing the main inner product (on vector spaces of polynomials) that will be used throughout our paper. For n-variate degree d homogenous polynomials $f(x) :=$

$\sum_{|\alpha|=d} b_\alpha x^\alpha$ and $g(x) := \sum_{|\alpha|=d} c_\alpha x^\alpha$ 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., [28, Thm. 4.1]) that for any $U \in O(n)$ we have

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

We equip the vector space of n-variate degree d homogenous polynomials H_d with the Weyl-Bombieri inner product and we consider pointwise evaluations H_d. More precisely, for any $v \in S^{n-1}$, consider the $l_v : H_d \to \mathbb{R}$, $l_v(f) = f(v)$. For the natural resulting Hilbert space structure on H_d, the Riesz Representation Theorem tells us that for any $v \in S^{n-1}$ there exists a corresponding unique $q_v \in H_d$ such that, for all $f \in H_d$, we have

$$l_v(f) = f(v) = \langle f, q_v \rangle_W.$$

It is easy to show that for the norm $\|\cdot\|_W$ on H_d induced by $\langle \cdot, \cdot \rangle_W$, we have $\|q_v\|_W = 1$ for all $v \in S^{n-1}$. In particular, for the Weyl-Bombieri inner product it is a simple exercise to verify that $q_v(x) = \langle v, x \rangle^d$.

Now, let $E \subseteq H_d$ be a linear subspace, and let $\Pi_E : H_d \longrightarrow E$ denote orthogonal projection onto E. Then, for all $f \in E$, we have $f(v) = \langle f, q_v \rangle_W = \langle f, \Pi_E(q_v) \rangle_W$. Even though $\|q_v\|_W$ is fixed for all $v \in S^{n-1}$, $\|\Pi_E(q_v)\|_W$ can vary arbitrarily between 0 to 1. This is due to the fact that H_d is closed under $O(n)$-action, and all directions $v \in S^{n-1}$ are "equal". In particular, E need not be closed under this action and every $v \in S^{n-1}$ has weight $\|\Pi_E(q_v)\|_W$.

We now state some basic probabilistic estimates for our structured setup.

Lemma 2.1. *Let $E \subseteq H_d$ be a full linear subspace, and let $p \in E$ be a random element of E that satisfy the Centering Property, the Sub-Gaussian Property with constant K, and the Small Ball Property with constant c_0. Then, for all $v \in S^{n-1}$, the following estimates hold:*

$$\text{Prob}\left(|p(v)| \geq t \|\Pi_E(q_v)\|_W\right) \leq \exp\left(1 - \frac{t^2}{K^2}\right)$$

$$\text{Prob}\left(|p(v)| \leq \varepsilon \|\Pi_E(q_v)\|_W\right) \leq c_o \varepsilon$$

Proof. By definition, we have $p(v) = \langle p, \Pi_E(q_v) \rangle_W$. So by the Cauchy-Schwartz Inequality and the 3 assumed properties on the distribution for p, we are done. ∎

For notational convenience, we define two extremal quantities for a linear subspace $E' \subseteq H_d$:

$$\sigma_{\max}(E') := \max_{v \in S^{n-1}} \|\Pi_{E'}(q_v)\|_W$$

$$\sigma_{\min}(E') := \min_{v \in S^{n-1}} \|\Pi_{E'}(q_v)\|_W .$$

The following is an immediate corollary of Lemma 2.1.

Corollary 2.2. *Assume the hypothesis of Lemma 2.1, then for all $w \in S^{n-1}$ the following estimates hold:*

$$\text{Prob}\left(|p(w)| \geq t\sigma_{\max}(E')\right) \leq \exp\left(1 - \frac{t^2}{K^2}\right)$$

$$\text{Prob}\left(|p(w)| \leq \varepsilon\sigma_{\min}(E')\right) \leq c_o \varepsilon.$$

We now define the Weyl-Bombieri inner product for polynomial systems and derive some basic probabilistic estimates. Let $D := (d_1, \ldots, d_{n-1})$ and let H_D denote the space of (real) $(n-1)\times n$ systems of homogenous n-variate polynomials with respective degrees $d_1, \ldots, d_{n-1}$. Then for $F := (f_1, \ldots, f_{n-1}) \in H_D$ and $G := (g_1, \ldots, g_{n-1}) \in H_D$ we define their Weyl-Bombieri inner product to be $\langle F, G\rangle_W := \sum_{i=1}^{n-1} \langle f_i, g_i\rangle_W$. We also let $\|F\|_W := \sqrt{\langle F, F\rangle}$.

Let $E_i \subseteq H_{d_i}$ be linear subspaces, and let $E = (E_1, \ldots, E_{n-1})$ denote a linear space of polynomial systems $P = (p_1, \ldots, p_{n-1})$ with $p_i \in E_i$ for all i. We define the following quantities for notational convenience: $\sigma_{\max}(E) := \max_i \sigma_{\max}(E_i)$ and $\sigma_{\min}(E) := \min_i \sigma_{\min}(E_i)$. We are now ready to present our basic tool.

Lemma 2.3. *Let $D = (d_1, \ldots, d_{n-1})$ be a vector with positive integer coordinates, let $E_i \subseteq H_{d_i}$ be full linear subspaces, and let $E := (E_1, \ldots, E_{n-1})$. Let $p_i \in E_i$ be random elements of E_i that satisfy the Centering Property, the Sub-Gaussian Property with constant K, and the Small Ball Property with constant c_0, all with respect to Bombieri-Weyl inner product. Then, for the random polynomial system $P = (p_1, \ldots, p_{n-1})$ and all $w \in S^{n-1}$, the following estimates hold:*

$$\mathrm{Prob}\left(\|P(w)\|_2 \geq t\sigma_{\max}(E)\sqrt{n-1}\right) \leq \exp\left(1 - \frac{a_1 t^2 (n-1)}{K^2}\right)$$

$$\mathrm{Prob}\left(\|P(w)\|_2 \leq \varepsilon\sigma_{\min}(E)\sqrt{n-1}\right) \leq (a_2 c_o \varepsilon)^{n-1}$$

where a_1 and a_2 are absolute constants.

The reader has perhaps already observed the contrast between the two inequalities in Lemma 2.3. The first inequality controls how large $\|P(w)\|_2$ is, while the second inequality controls how small $\|P(w)\|_2$ is. This contrast between these inequalities will be our main tool throughout this paper. So we will need to somehow use the ratio of $\sigma_{\max}(E)$ and $\sigma_{\min}(E)$.

Definition 2.4. *Let $E' \subseteq H_d$ be a linear subspace, and let $\sigma_{\min}(E')$ and $\sigma_{\max}(E')$ be as defined above. Then we call $\sigma(E') := \frac{\sigma_{\max}(E')}{\sigma_{\min}(E')}$ the* dispersion constant *of E'. Now, for $E_i \subseteq H_{d_i}$ and $E := (E_1, \ldots, E_m)$, we then define the* dispersion constant (for an n-tuple of linear spaces) *to be $\sigma(E) = \frac{\sigma_{\max}(E)}{\sigma_{\min}(E)}$.* $\diamond$

For now, the reader should be aware of two things: $\sigma(E)$ will appear in main theorems in the following section, and the last section of this article is completely devoted to understanding $\sigma(E)$. So after one finishes reading preliminaries, the remaining two sections can be read independently.

Now we prove Lemma 2.3. For the proof we need to recall some theorems from probability theory and some basic tools developed in Part I [20] of our present work. These basic lemmata will also be used later in other proofs. We start with a theorem which is reminiscent of Hoeffding's classical inequality [21].

Theorem 2.5. [49, Prop. 5.10] *There is an absolute constant $\tilde{c}_1 > 0$ with the following property: If $X_1, \ldots, X_n$ are centered, sub-Gaussian random variables with constant K, and $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ and $t \geq 0$, then*

$$\mathrm{Prob}\left(\left|\sum_i a_i X_i\right| \geq t\right) \leq 2\exp\left(\frac{-\tilde{c}_1 t^2}{K^2 \|a\|_2^2}\right). \qquad \blacksquare$$

We will also need the following standard lemma (see, e.g., [38, Lemma 2.2]).

Lemma 2.6. *Assume $Z_1, \dots, Z_n$ are independent random variables that have the property that $F(Z_i, t)) \leq c_0 t$ for all $t > 0$. Then $F(W, t\sqrt{n}) \leq (cc_0 t)^n$, $t > 0$, where $W := \|(Z_1, \dots, Z_n)\|_2$. Moreover, if $\xi_1, \dots, \xi_k$ are independent random variables such that, for every $\varepsilon > 0$, we have* $\mathrm{Prob}\,(|\xi_i| \leq \varepsilon) \leq c_0 \varepsilon$, *then there is a universal constant $\tilde{c} > 0$ such that for every $\varepsilon > 0$ we have* $\mathrm{Prob}\left(\sqrt{\xi_1^2 + \cdots + \xi_k^2} \leq \varepsilon\sqrt{k}\right) \leq (\tilde{c} c_0 \varepsilon)^k$. ■

Now we have the basic tools to "tensorize" sub-Gaussian tail bounds and small ball inequalities. We proceed with deterministic inequalities for polynomial systems.

The following lemma was proved in our earlier paper [20], generalizing a classical Theorem of Kellog [24]. To state the lemma we need a bit of terminology. For any system of homogeneous polynomials $P := (p_1, \dots, p_{n-1}) \in (\mathbb{R}[x_1, \dots, x_n])^{n-1}$ define $\|P\|_\infty := \sup_{x \in S^{n-1}} \sqrt{\sum_{i=1}^{n-1} p_i(x)^2}$. Let $DP(x)$ denote the Jacobian matrix of the polynomial system at point x, let $DP(x)(u)$ denote the image of the vector u under the linear operator $DP(x)$, and set $\left\|D^{(1)}P\right\|_\infty := \sup_{x,u \in S^{n-1}} \|DP(x)(u)\|_2 = \sup_{x,u \in S^{n-1}} \sqrt{\sum_{i=1}^{n-1} \langle \nabla p_i(x), u \rangle^2}$.

Lemma 2.7. *Let $P := (p_1, \dots, p_{n-1}) \in (\mathbb{R}[x_1, \dots, x_n])^{n-1}$ 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 $\left\|D^{(1)}P\right\|_\infty \leq d^2 \|P\|_\infty$ 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 $i \in \{1, \dots, m\}$ then we also have $\left\|D^{(1)}P\right\|_\infty \leq d \|P\|_\infty$.* ■

The last lemma we need is a discretization tool for homogenous polynomial systems that was developed in [20] based on Lemma 2.7.

Lemma 2.8. *Let $P = (p_1, \dots, p_{n-1})$ be a system of homogenous polynomials p_i with n variables and $\deg(p_i) = d_i$. Let $\mathcal{N}$ be a δ-net on S^{n-1}. Let $\max_{\mathcal{N}}(P) = \sup_{y \in \mathcal{N}} \|P(y)\|_2$ and $\|P\|_\infty = \sup_{x \in S^{n-1}} \|P(x)\|_2$. Similarly let us define,*

$$\max_{\mathcal{N}^{k+1}}(D^{(k)}P) = \sup_{x, u_1, \dots, u_k \in \mathcal{N}} \left\|D^{(k)}P(x)(u_1, \dots, u_k)\right\|_2$$

$$\left\|D^{(k)}P\right\|_\infty = \sup_{x, u_1, \dots, u_k \in S^{n-1}} \left\|D^{(k)}P(x)(u_1, \dots, u_k)\right\|_2$$

1. *When $\deg(p_i) = d$ for all $i \in \{1, 2, \dots, m\}$ we have $\|P\|_\infty \leq \frac{\max_{\mathcal{N}}(P)}{1 - d\delta}$ and $\left\|D^{(k)}P\right\|_\infty \leq \frac{\max_{\mathcal{N}^{k+1}}(D^{(k)}P)}{1 - \delta d\sqrt{k+1}}$.*
2. *When $\max_i\{deg(p_i)\} \leq d$ we have $\|P\|_\infty \leq \frac{\max_{\mathcal{N}}(P)}{1 - d^2\delta}$ and $\left\|D^{(k)}P\right\|_\infty \leq \frac{\max_{\mathcal{N}^{k+1}}(D^{(k)}P)}{1 - \delta d^2\sqrt{k+1}}$.* ■

Proof of Lemma 2.3: We begin with the first claim. Using Lemma 2.1 and the fact that $\sigma_{\max} \geq \sigma_{\max}(E_i)$ yield the following estimate for any $p_i \in E_i$ and for all $w \in S^{n-1}$:

$$\mathrm{Prob}\,(|p_i(w)| \geq t\sigma_{\max}(E)) \leq \exp\left(1 - \frac{t^2}{K^2}\right)$$

Now let $a = (a_1, \dots, a_{n-1}) \in \mathbb{R}^{n-1}$ with $\|a\|_2 = 1$, and apply Lemma 2.5 to the sub-Gaussian random variables $\frac{p_i(w)}{\sigma_{\max}(E)}$ and the vector a:

$$\mathrm{Prob}\left(\left|\sum_i a_i p_i(w)\right| \geq t\sigma_{\max}(E)\right) \leq \exp\left(1 - \frac{\tilde{c}_1 t^2}{K^2}\right)$$

Observe that $\|P(w)\|_2 = \max_{a\in S^{n-2}} |\langle a, P(w)\rangle|$. For any fixed point $w \in S^{n-1}$ and a free variable $a \in \mathbb{R}^n$, we have that $\langle a, P(w)\rangle$ is a linear polynomial on a. We then use Lemma 2.8 for this linear polynomial, which gives us the following estimate:

$$\text{Prob}\left(\|P(w)\|_2 \geq \frac{t\sigma_{\max}(E)}{1-\delta}\right) \leq |N| \exp\left(1 - \frac{\tilde{c}_1 t^2}{K^2}\right)$$

We need to control $|N| = \left(\frac{3}{\delta}\right)^{n-1} = e^{(n-1)\tilde{c}\log(\frac{1}{\delta})}$. So we set $t = 2t\sqrt{n-1}$ (i.e assuming $t \geq 2\sqrt{n-1}$), $\delta = \frac{1}{2}$, and we have the following estimate with some universal constant a_1.

$$\text{Prob}\left(\|P(w)\|_2 \geq t\sigma_{\max}(E)\sqrt{n-1}\right) \leq \exp\left(1 - \frac{a_1 t^2(n-1)}{K^2}\right)$$

We continue with proof of the second claim. Using Lemma 2.1 and the fact that $\sigma_{\min}(E) \leq \sigma_{\min}(E_i)$ for all i, we deduce the following estimate hold for all p_i and for any $\varepsilon > 0$:

$$\text{Prob}\left(\left|\frac{p_i(w)}{\sigma_{\min}(E)}\right| \leq \varepsilon\right) \leq c_0\varepsilon$$

Using Lemma 2.6 on random variables $\left|\frac{p_i(w)}{\sigma_{\min}(E)}\right|$ gives the following estimate:

$$\text{Prob}\left(\|P(w)\|_2 \leq \varepsilon\sigma_{\min}(E)\sqrt{n-1}\right) \leq (\tilde{c}_2 c_0 \varepsilon)^{n-1}. \blacksquare$$

3. Probabilistic Analysis of Condition Number for Structured Polynomial Systems

In this section we will prove our main theorem on probabilistic analysis of the condition number $\tilde{\kappa}(P)$ for structured polynomial systems. Recall that for a given $D = (d_1, \ldots, d_k)$, we defined H_D to be the vector space of homogenous polynomial systems $P = (p_1, \ldots, p_k)$ with $\deg(p_i) = d_i$. We call a system of homogenous polynomials overdetermined if $k > n-1$, square if $k = n-1$, and underdetermined if $k < n-1$. Our techniques in this section work for probabilistic analysis of $\tilde{\kappa}(P)$ for square and overdetermined polynomial systems. However, for the sake of simplicity in presentation we will state and prove our theorems only for the square case. Deriving estimates for the overdetermined case based on the arguments presented in this article is routine. For technical details of the overdetermined case, the interested reader is invited to consult our earlier paper [20].

The definitions of local condition number $\tilde{\kappa}(P, x)$ and global condition number $\tilde{\kappa}(P)$ were given in the introduction as follows: $\tilde{\kappa}(P, x) := \frac{\|P\|_W}{\sqrt{\|P\|_W^2 \tilde{\mu}_{\text{norm}}(P,x)^{-2} + \|P(x)\|_2^2}}$ and $\tilde{\kappa}(P) := \sup_{x\in S^{n-1}} \tilde{\kappa}(P, x)$. The idea behind this definition can be summarized in a theorem as follows: First, for $x \in S^{n-1}$, 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 then define $\Sigma_{\mathbb{R}}$ (the *real part of the disciminant variety for* H_D) to be:

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

Using the Weyl-Bombieri norm to measure the distance in space H_D, we point out the following important geometric characterization of $\tilde{\kappa}$.

Theorem 3.1. [14, Prop. 3.1] *For all* $P \in H_D$, *we have* $\tilde{\kappa}(P) = \frac{\|P\|_W}{\text{Dist}(P,\Sigma_{\mathbb{R}})}$. $\blacksquare$

After this theorem, it should be clear that there are two main factors determining the behavior of $\tilde{\kappa}(P)$: the normalization factor $\|P\|_W$ and the denominator $\text{Dist}(P, \Sigma_{\mathbb{R}})$. $\|P\|_W$

has strong concentration properties for a broad family of random polynomials and is relatively easy to handle. So we will be mainly interested in probabilistic analysis of the denominator in the definition of $\tilde{\kappa}$. We define the following quantity for later convenience.

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

It follows directly that $\frac{\|P\|_W}{\tilde{\kappa}(P,x)} = \sqrt{\|P\|_W^2 \tilde{\mu}_{\text{norm}}(P,x)^{-2} + \|P(x)\|_2^2} = \inf\limits_{\substack{y \perp x \\ y \in S^{n-1}}} \mathcal{L}(x,y)$. So we set $L(P,x) = \frac{\|P\|_W}{\tilde{\kappa}(P,x)}$ and $L(P) = \min_{x \in S^{n-1}} L(P,x)$. We then have the following equalities: $L(P,x) = \inf\limits_{\substack{y \perp x \\ y \in S^{n-1}}} \mathcal{L}(x,y)$, $\tilde{\kappa}(P,x) = \frac{\|P\|_W}{L(P,x)}$, and $\tilde{\kappa}(P) = \frac{\|P\|_W}{L(P)}$.

3.1. **Operator Norm Type Estimates.** In this section we will estimate the absolute maximum of a random polynomial system on the sphere. Recall that for a homogenous polynomial system $P = (p_1, \ldots, p_{n-1})$ the sup-norm is defined as $\|P\|_\infty = \sup_{x \in S^{n-1}} \|P(x)\|_2$. The following lemma is our sup-norm estimate for random polynomial system P.

Lemma 3.2. *Let $D = (d_1, \ldots, d_{n-1})$ be a vector with positive integer coordinates, let $E_i \subseteq H_{d_i}$ be full linear subspaces, and let $E = (E_1, \ldots, E_{n-1})$. Let $p_i \in E_i$ be independent random elements of E_i that satisfy the Centering Property, the Sub-Gaussian Property with constant K, and the Small Ball Property with constant c_0, each with respect to Bombieri-Weyl inner product. Let $\mathcal{N}$ be a δ-net on S^{n-1}. Then for $P = (p_1, \ldots, p_{n-1})$ we have*

$$\text{Prob}\left(\max_{x \in \mathcal{N}} \|P(x)\|_2 \geq t\sigma_{\max}(E)\sqrt{n}\right) \leq |\mathcal{N}| \exp\left(1 - \frac{a_1 t^2 n}{K^2}\right).$$

In particular, for $d = \max_i \deg(p_i)$, $\delta = \frac{1}{3d^2}$, and $t = s\log(ed)$ with $s \geq 1$ this gives us the following estimate with an absolute constant a_3.

$$\text{Prob}\left(\|P\|_\infty \geq s\sigma_{\max}(E)\sqrt{n}\log(ed)\right) \leq \exp\left(1 - \frac{a_3 s^2 n \log(ed)}{K^2}\right)$$

Proof. The first statement is proven by just taking a union bound over $\mathcal{N}$ and using Lemma 2.3. The second part of the statement immediately follows by using the first part and Lemma 2.8. ■

3.2. **Small Ball Type Estimates.** In this section, we prove a small ball type estimate to control behavior of the denominator $L(p)$. We first need to recall a technical lemma from our earlier paper [20], which builds on an idea of Nguyen [36].

Lemma 3.3. *Let $n \geq 2$, let $P := (p_1, \ldots, p_{n-1})$ 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 $w = x + \beta r y + \beta^2 z$ 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 $\|P(w)\|_2^2 \leq 8(\alpha^2 + (2+e^4)\beta^4 d^4 \gamma^2)$.*
2. *If $\deg(p_i) = d$ for all $i \in [n-1]$, and $0 < \beta \leq d^{-2}$ then $\|P(w)\|_2^2 \leq 8(\alpha^2 + (2+e^4)\beta^4 d^4 \gamma^2)$.*

Theorem 3.4. *Let $D = (d_1, \ldots, d_{n-1})$ be a vector with positive integer coordinates, let $E_i \subseteq H_{d_i}$ be full linear subspaces, and let $E = (E_1, \ldots, E_{n-1})$. Let $p_i \in E_i$ be independent random elements of E_i that satisfy the Centering Property, the Sub-Gaussian Property with* *constant K, and the Small Ball Property with constant c_0, each with respect to Bombieri-Weyl inner product. Let $\gamma \geq 1$, $d := \max_i d_i$, and assume $\alpha \leq \min\{d^{-8}, n^{-1}\}$. Then for $P = (p_1, \ldots, p_{n-1})$ we have*

$$\mathrm{Prob}(L(P) \leq \alpha) \leq \mathrm{Prob}\left(\|P\|_\infty \geq \gamma\right) + c\alpha^{\frac{1}{2}}\sqrt{n}\left(\frac{Cc_0 d^2\gamma}{\sigma_{\min}(E)\sqrt{n}}\right)^{n-1}$$

where C is a universal constant.

The proof of Theorem 3.4 is similar to a proof in our earlier paper [20]. We reproduce the proof due to the importance of Theorem 3.4 in the flow of the current paper.

Proof. We assume the hypotheses of Assertion (1) in Lemma 3.3: Let $\alpha, \gamma > 0$ and $\beta \leq d^{-4}$. Let $\mathbf{B} := \{P \mid \|P\|_\infty \leq \gamma\}$ and let

$\mathbf{L} := \{P \mid L(P) \leq \alpha\} = \{P \mid$ There exist $x, y \in S^{n-1}$ with $x \perp y$ 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 3.3 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 $V := \mathrm{Vol}(V_{x,y})$. 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

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

Using Markov's Inequality, Fubini's Theorem, and Lemma 2.3, we can estimate the probability of this event. Indeed,

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

Now recall that $\mathrm{Vol}(B_2^n) = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}$. Then $\frac{\mathrm{Vol}(B_2^n)}{\mathrm{Vol}(B_2^{n-1})} \leq \frac{c'}{\sqrt{n}}$ 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

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

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 2.3, that for every $w \in (1 + 2\beta^2)B_2^n \setminus B_2^n$ we have

$$\mathrm{Prob}\left(\|P(w)\|_2^2 \leq \Gamma\right) \leq \mathrm{Prob}\left(\|P(\tilde{w})\|_2^2 \leq \Gamma\right) \leq \left(cc_0\sqrt{\frac{\Gamma}{n\sigma_{\min}(E)^2}}\right)^{n-1}.$$

So we conclude that $\text{Prob}\left(\mathbf{B} \cap \mathbf{L}\right) \leq c\sqrt{n}\beta\beta^{2-2n}\left(cc_0\sqrt{\frac{\Gamma}{n\sigma_{\min}(E)^2}}\right)^{n-1}$. Since $\text{Prob}\left(L(P) \leq \alpha\right) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right) + \text{Prob}(\mathbf{B} \cap \mathbf{L})$ we then have

$$\text{Prob}\left(L(P) \leq \alpha\right) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right) + c\sqrt{n}\beta\beta^{2-2n}\left(cc_0\sqrt{\frac{\Gamma}{n\sigma_{\min}(E)^2}}\right)^{n-1}$$

Recall that $\Gamma = 8(\alpha^2 + (5+e^4)\beta^4 d^4\gamma^2)$. We set $\beta^2 := \alpha$. Our choice of β and the assumption that $\gamma \geq 1$ then imply that $\Gamma \leq C\alpha^2\gamma^2 d^4$ for some constant C. So we obtain

$$\text{Prob}(L(P) \leq \alpha) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right) + c\sqrt{n}(\alpha)^{\frac{3}{2}-n}\left(\frac{Cc_0\alpha d^2\gamma}{\sigma_{\min}(E)\sqrt{n}}\right)^{n-1}$$

$$\text{Prob}(L(P) \leq \alpha) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right) + c\sqrt{n}(\alpha)^{\frac{1}{2}}\left(\frac{Cc_0 d^2\gamma}{\sigma_{\min}(E)\sqrt{n}}\right)^{n-1}$$

and our proof is complete. ■

3.3. **Condition Number Theorem.** In this section we will prove our main result on probabilistic analysis of condition numbers for structured real polynomial systems. We first need to estimate $\|P\|_W$ for such systems. The following lemma is more or less standard, and it follows from Lemma 2.5.

Lemma 3.5. *Let $D = (d_1, \ldots, d_{n-1})$ be a vector with positive integer coordinates, let $E_i \subseteq H_{d_i}$ be full linear subspaces, and let $E = (E_1, \ldots, E_{n-1})$. Let $p_i \in E_i$ be random elements of E_i that satisfy the Centering Property and the Sub-Gaussian Property with constant K, each with respect to Bombieri-Weyl inner product. Then for all $t \geq 1$, we have*

$$\text{Prob}\left(\|p_i\|_W \geq t\sqrt{\dim(E_i)}\right) \leq \exp\left(1 - \frac{t^2\dim(E_i)}{K^2}\right)$$

and

$$\text{Prob}\left(\|P\|_W \geq t\sqrt{\dim(E)}\right) \leq \exp\left(1 - \frac{t^2\dim(E)}{K^2}\right).$$

Now we have all the necessary tools to prove our probabilistic condition number theorem.

Theorem 3.6. *Let $D = (d_1, \ldots, d_{n-1})$ be a vector with positive integer coordinates, let $E_i \subseteq H_{d_i}$ be full linear subspaces, and let $E = (E_1, \ldots, E_{n-1})$. We assume that $\dim(E) \geq n\log(ed)$ and $n \geq 3$. Let $p_i \in E_i$ be independent random elements of E_i that satisfy the Centering Property, the Sub-Gaussian Property with constant K, and the Small Ball Property with constant c_0, each with respect to the Bombieri-Weyl inner product. We set $d := \max_i d_i$, and*

$$M := nK\sqrt{\dim(E)}(Cc_0Kd^2\log(ed)\sigma(E))^{2n-2}$$

where $C \geq 4$ is a universal constant. Then for $P = (p_1, \ldots, p_{n-1})$, we have

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

Proof. Let $m = \dim(E)$. A first observation to start our proof is the following:

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

The first probability on the right hand side will be controlled by Lemma 3.5, and the latter will be controlled by Theorem 3.4. Theorem 3.4 states that for any $\gamma \geq 1$ and for $\frac{sK\sqrt{m}}{tM} \leq \min\{d^{-8}, n^{-1}\}$, we have

$$\text{Prob}\left(L(P) \leq \frac{sK\sqrt{m}}{tM}\right) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right) + \left(\frac{sK\sqrt{m}}{tM}\right)^{\frac{1}{2}} \sqrt{n}\left(\frac{Cc_0\gamma d^2}{\sigma_{\min}(E)\sqrt{n}}\right)^{n-1}$$

To have $\frac{sK\sqrt{m}}{tM} \leq \min\{d^{-8}, n^{-1}\}$ is equivalent to $tM\min\{d^{-8}, n^{-1}\} \geq sK\sqrt{m}$. We will check this condition at the end of the proof. Now, for $\gamma = u\sigma_{\max}(E)\sqrt{n}\log(ed)K$ with $u \geq 1$, from Lemma 3.2 we have $\text{Prob}\left(\|P\|_\infty \geq u\sigma_{\max}(E)\sqrt{n}\log(ed)K\right) \leq \exp(1 - a_3u^2n\log(ed))$. That is, for $\gamma = u\sigma_{\max}(E)\sqrt{n}\log(ed)K$, we have the following estimate:

$$\text{Prob}\left(L(P) \leq \frac{s\sqrt{m}}{tM}\right) \leq \exp(1 - a_3u^2n\log(ed)) + \left(\frac{sK\sqrt{m}}{tM}\right)^{\frac{1}{2}} \sqrt{n}\left(\frac{Cc_0u\sigma_{\max}(E)\log(ed)d^2K}{\sigma_{\min}(E)}\right)^{n-1}$$

Since $\sigma(E) = \frac{\sigma_{\max}(E)}{\sigma_{\min}(E)}$ and $M = n\sqrt{m}K(Cc_0\log(ed)d^2K\sigma(E))^{2n-2}$, we have

$$\text{Prob}\left(L(P) \leq \frac{s\sqrt{m}}{tM}\right) \leq \exp(1 - a_3u^2n\log(ed)) + \left(\frac{s}{t}\right)^{\frac{1}{2}} u^{n-1}.$$

Using Lemma 3.5 and the assumption that $m \geq n\log(ed)$ we then have

$$\text{Prob}\left(\tilde{\kappa}(P) \geq tM\right) \leq \exp(1 - s^2n\log(ed)) + \exp(1 - a_3u^2n\log(ed)) + \left(\frac{s}{t}\right)^{\frac{1}{2}} u^{n-1}.$$

If $t \leq e^{2n\log(ed)}$ then setting $s = u = 1$ gives the desired inequality. If $t \geq e^{2n\log(ed)}$ then we set $s = u^2 = \frac{\log(t)}{2\tilde{a}n\log(ed)}$, where $\tilde{a} > a_3 > 0$ is a constant greater than 1. Then we have

$$\text{Prob}\left(\tilde{\kappa}(P) \geq tM\right) \leq \exp\left(2 - \frac{1}{2}\log(t)\right) + \left(\frac{\log(t)}{2n\log(ed)}\right)^{\frac{n}{2}} \frac{1}{\sqrt{t}}.$$

Observe that $\exp\left(2 - \frac{1}{2}\log t\right) = \frac{e^2}{\sqrt{t}}$. So we have $\text{Prob}\left(\tilde{\kappa}(P) \geq tM\right) \leq \left(\frac{\log(t)}{2n\log(ed)}\right)^{\frac{n}{2}} \frac{e^2+1}{\sqrt{t}}$. To finalize our proof we need to check if $tM\min\{d^{-8}, n^{-1}\} \geq sK\sqrt{m}$. So we check the following:

$$tKn\sqrt{m}(Cc_0\log(ed)d^2K\sigma(E))^{2n-2}\min\{d^{-8}, n^{-1}\} \overset{?}{\geq} \frac{\log(t)}{2n\log(ed)}K\sqrt{m}.$$

For $n \geq 3$, $(d^2\log(ed))^{2n-2} > d^8$. Since $Kc_0 \geq \frac{1}{4}$, $C \geq 4$, and $\sigma(E) \geq 1$, we have

$$(Cc_0\log(ed)d^2K\sigma(E))^{2n-2} > d^8.$$

Hence, it suffices to check if $t \geq \frac{\log(t)}{2n\log(ed)}$, which is clear. ■

We would like to complete the proof of Theorem 1.5 as presented in the introduction, for which the following easy observation suffices.

Lemma 3.7. *For $t \geq e^{2n\log(ed)}$, we have* $\left(\frac{\log(t)}{2n\log(ed)}\right)^{\frac{n}{2}} \leq t^{\frac{1}{4\log(ed)}}$.

Proof. Let $t = xe^{2n\log(ed)}$ where $x \geq 1$. Then

$$\left(\frac{\log(t)}{2n\log(ed)}\right)^{\frac{n}{2}} = \left(1 + \frac{\log(x)}{2n\log(ed)}\right)^{\frac{n}{2}} \leq e^{\frac{\log(x)}{4\log(ed)}} = x^{\frac{1}{4\log(ed)}}$$

Since $x \leq t$, we are done. ■

We now state the resulting bounds on the expectation of the condition number.

Corollary 3.8. *Under the assumptions of Theorem 3.6,* $0 < q < \frac{1}{2} - \frac{1}{2\log(ed)}$ *implies that* $\mathbb{E}(\tilde{\kappa}(P)^q) \leq M^q(1 + 4q\log(ed))$. *In particular,* $\mathbb{E}\log(\tilde{\kappa}(P)) \leq 1 + \log M$.

3.4. **Smoothed Analysis.** We had 3 initial randomness assumptions in the introduction of our paper: Centering, sub-Gaussian tails, and Small Ball. The sub-Gaussian tails are a large deviation inequality for magnitude of a random variable. Such tails can be seen in some sense as concentration around the expectation. Small Ball inequalities are in the opposite spirit: They ask for anti-concentration in a small ball around 0. For smoothed analysis type estimates we will need to assume a stronger anti-concentration property, which is nevertheless satisfied by all the examples presented in the introduction. We recall the definition of concentration function of a random variable $X \in \mathbb{R}$ for $t \geq 0$ is

$$F(Z,t) := \max_{u\in\mathbb{R}} \text{Prob}\{|Z - u| \leq t\}$$

In this section, we assume that the random vectors C_i satisfy the following anti-concentration inequality for all $\theta \in S^{d_i-1}$:

$$F(\langle C_i, \theta\rangle, \varepsilon) \leq c_0\varepsilon.$$

Using this assumption, we will be able to analyze the condition number $\tilde{\kappa}(P)$ for a random polynomial system P that is created as a random perturbation around a fixed polynomial system Q. More precisely, let $D = (d_1, \ldots, d_{n-1})$, let $E_i \subseteq H_{d_i}$ be linear spaces, let $E = (E_1, \ldots, E_{n-1})$, and let $Q \in E$ be a polynomial system. Let $g_i \in E_i$ be independent random elements of E_i that satisfy the Centering Property, the Sub-Gaussian Property with constant K, and the Anti-Concentration Property with constant c_0, each with respect to the Bombieri-Weyl inner product. We define the random polynomial system $G := (g_1, g_2, \ldots, g_{n-1})$ and the perturbed polynomial system is defined by $P := Q + G$. We will this notation for P, Q and G for the rest of this section.

Lemma 3.9. *Assume the introduced hypothesis on the smoothed system* $P = Q + G$. *Then we have*

$$\text{Prob}\left(\|P\|_\infty \geq s\sigma_{\max}(E)\sqrt{n}\log(ed) + \|Q\|_\infty\right) \leq \exp\left(1 - \frac{a_3 s^2 n\log(ed)}{K^2}\right)$$

where a_3 *is an absolute constant.*

Proof. The triangle inequality implies $\|P\|_\infty \leq \|Q\|_\infty + \|G\|_\infty$. We complete the proof by using Lemma 3.2 for the random system G. ■

Lemma 3.10. *Assume the introduced hypothesis on the smoothed system* $P = Q+G$. *Then, for all* $\varepsilon > 0$, *and for any* $w \in S^{n-1}$ *we have*

$$\text{Prob}\left(\|P(w)\|_2 \leq \varepsilon\sigma_{\min}(E)\sqrt{n-1}\right) \leq (a_2 c_o\varepsilon)^{n-1}$$

a_2 *is an absolute constant.*

Proof. By the Anti-Concentration Property, for all $1 \leq i \leq n-1$, we have

$$\text{Prob}\{|g_i(w) + q_i(w)| \leq c_o\varepsilon\sigma_{\min}(E_i)\} \leq c_0\varepsilon$$

We use Lemma 2.6 to tensorize this inequality with random variables being $g_i(w) + q_i(w)$. ■

Lemma 3.11. *Assume the introduced hypothesis on the smoothed system $P = Q + G$. Then for all $t \geq 1$, we have*

$$\text{Prob}\left(\|P\|_W \geq tK\sqrt{\dim(E)} + \|Q\|_W\right) \leq \exp(1 - t^2 m)$$

Proof. For all $1 \leq i \leq n-1$, by triangle inequality $\|p_i\|_W \leq \|q_i\|_W + \|g_{C_i}\|_W$. So using the first claim of Lemma 3.5 gives

$$\text{Prob}\left(\|p_i\|_W \geq t\sqrt{\dim(E_i)} + \|q_i\|_W\right) \leq \exp\left(1 - \frac{t^2 \dim(E_i)}{K^2}\right)$$

Tensorizing this inequality completes the proof. ■

Theorem 3.12. *Assume the introduced hypothesis on the smoothed system $P = Q + G$, let $\gamma \geq 1$, $d := \max_i d_i$, and assume $\alpha \leq \min\{d^{-8}, n^{-1}\}$. Then*

$$\text{Prob}(L(P) \leq \alpha) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right) + c\alpha^{\frac{1}{2}}\sqrt{n}\left(\frac{Cc_0 d^2 \gamma}{\sigma_{\min}(E)\sqrt{n}}\right)^{n-1}$$

where C is a universal constant.

Proof of Theorem 3.12 is identical to Theorem 3.4, so we skip. Now, we are ready to state main theorem of this section.

Theorem 3.13. *Assume the introduced hypothesis on smoothed system $P = Q + G$, let $d := \max_i d_i$, and set*

$$M = nK\sqrt{\dim(E)}\left(Cc_o K d^2 \log(ed)\sigma(E)\right)^{2n-2}(1 + \frac{\|Q\|_W}{\sqrt{n}K\log(ed)})^{2n-1}$$

where $C \geq 4$ is a universal constant. Also assume that $\dim(E) \geq n\log(ed)^2$, *and $n \geq 3$. Then*

$$\text{Prob}(\tilde{\kappa}(P) \geq tM) \leq \begin{cases} \frac{3}{\sqrt{t}} & \textit{if } 1 \leq t \leq e^{2n\log(ed)} \\ \frac{e^2+1}{\sqrt{t}}\left(\frac{\log t}{2n\log(ed)}\right)^{\frac{n}{2}} & \textit{if } e^{2n\log(ed)} \leq t \end{cases}$$

Proof. We need a quick observation before we start our proof: for any $Q \in E$ and $w \in S^{n-1}$, we have $\|Q(w)\|_2^2 \leq \sum_{i=1}^{n-1} \|q_i\|_W^2 \, \sigma_{\max}(E_i) \leq \|Q\|_W^2 \, \sigma_{\max}(E)$. So we have

$$\|Q\|_\infty \leq \|Q\|_W \, \sigma_{\max}(E).$$

Using this upper bound on $\|Q\|_\infty$ and the assumption that $\dim(E) \geq n\log(ed)^2$, we deduce

$$M \geq nK\sqrt{\dim(E)}\left(Cc_o K d^2 \log(ed)\sigma(E)\right)^{2n-2}(1+\frac{\|Q\|_W}{nK\sqrt{\dim(E)}})(1+\frac{\|Q\|_\infty}{\sqrt{n}\log(ed)K\sigma_{\max}(E)})^{2n-2}$$

We will use this lower bound on M later in our proof. Now, let $m = \dim(E)$. We start our proof with the following.

$$\text{Prob}\left(\tilde{\kappa}(P) \geq tM\right) \leq \text{Prob}\left(\|P\|_W \geq sK\sqrt{m} + \|Q\|_W\right) + \text{Prob}\left(L(P) \leq \frac{sK\sqrt{m} + \|Q\|_W}{tM}\right)$$

Firstly, Lemma 3.11 states that

$$\text{Prob}\left(\|P\|_W \geq sK\sqrt{m} + \|Q\|_W\right) \leq \exp(1 - s^2 m).$$

Secondly, Theorem 3.12 states that for $\frac{sK\sqrt{m}+\|Q\|_W}{tM} \leq \min\{d^{-8}, n^{-1}\}$ we have

$$\text{Prob}\left(L(P) \leq \frac{sK\sqrt{m}+\|Q\|_W}{tM}\right) \leq \text{Prob}\left(\|P\|_\infty \geq \gamma\right)+c(\frac{sK\sqrt{m}+\|Q\|_W}{tM})^{\frac{1}{2}}\sqrt{n}\left(\frac{Cc_0 d^2\gamma}{\sigma_{\min}(E)\sqrt{n}}\right)^{n-1}$$

We set $\gamma = u\sigma_{\max}(E)\sqrt{n}\log(ed)K + \|Q\|_\infty$. From Lemma 3.9, we have

$$\text{Prob}\left(\|P\|_\infty \geq u\sigma_{\max}(E)\sqrt{n}\log(ed)K + \|Q\|_\infty\right) \leq \exp(1 - a_3u^2 n\log(ed))$$

We also have

$$\left(\frac{Cc_0 d^2\gamma}{\sigma_{\min}(E)\sqrt{n}}\right)^{n-1} = \left(Cc_oKud^2\log(ed)\sigma(E)\right)^{n-1}(1+\frac{\|Q\|_\infty}{u\sqrt{n}\log(ed)K\sigma_{\max}(E)})^{n-1}$$

Using $u \geq 1$, $s \geq 1$, $m \geq n\log(ed)^2$, and the lower bound $(*)$, we then have

$$\text{Prob}\left(\tilde{\kappa}(P) \geq tM\right) \leq \exp(1 - s^2 n\log(ed)) + \exp(1 - a_3u^2 n\log(ed)) + \left(\frac{s}{t}\right)^{\frac{1}{2}} u^{n-1}$$

The rest of the proof is identical to the proof of Theorem 3.6. ■

Repeating the same as in Lemma 3.7, and Theorem 3.8 we complete the proof of Theorem 1.7. As an application of our smoothed analysis, we at last derive our existence result (Theorem 1.8 from our introduction) for well-conditioned structured systems near a given structured system.

Proof of Theorem 1.8: Define a random polynomial system $F_\varepsilon = Q + G$ where G is Gaussian random polynomial system with $K = \frac{\varepsilon\|Q\|_W}{\sqrt{n}\log(ed)}$ and $c_oK = \frac{1}{\sqrt{2\pi}}$. Using Lemma 3.5 with $t = 1$, we have with probability at least $1 - \exp(1 - \dim(E))$ that

$$\|F_\varepsilon - Q\|_W = \|G\|_W \leq \frac{\varepsilon\,\|Q\|_W\sqrt{\dim(E)}}{\sqrt{n}\log(ed)}$$

For the condition estimate we will use Theorem 3.13. Firs note that with $K = \frac{\varepsilon\|Q\|_W}{\sqrt{n}\log(ed)}$ and $c_oK = \frac{1}{\sqrt{2\pi}}$, M in Theorem 3.13 is the following

$$M = \frac{\varepsilon\sqrt{n}\sqrt{\dim(E)}}{\log(ed)}\left(\frac{Cd^2\log(ed)\sigma(E)}{\sqrt{2\pi}}\right)^{2n-2}(1+\frac{1}{\varepsilon})^{2n-1}$$

So we have $M \leq 2\sqrt{n}\sqrt{\dim(E)}\left(\frac{2}{\varepsilon}\right)^{2n-2}\left(\frac{Cd^2\log(ed)\sigma(E)}{\sqrt{2\pi}}\right)^{2n-2}$. We use Theorem 3.13 with $t = 36$ and we deduce that with probability greater than $\frac{1}{2}$, we have

$$\tilde{\kappa}(F_\varepsilon) \leq 2\sqrt{n}\sqrt{\dim(E)}\left(\frac{Cd^2\log(ed)\sigma(E)}{\varepsilon}\right)^{2n-2}$$

Since the union of the complement of these two events has measure less than $\frac{1}{2} + \exp(1 - \dim(E))$, their intersection has positive measure, and the proof is completed. ■

Remark 3.14. *We remark that the proof of Theorem 3.12 actually works for*

$$M = nK\sqrt{\dim(E)}\left(Cc_oKd^2\log(ed)\sigma(E)\right)^{2n-2}(1+\|Q\|_W)(1+\frac{\|Q\|_\infty}{\sqrt{n}\log(ed)K\sigma_{\max}(E)})^{2n-2}$$

which is often much more smaller than the M used in the theorem statement.

4. Dispersion Constant of a Linear Structure in Polynomial Systems

Consider a linear space $E \subseteq H_D$, where $D = (d_1, \dots, d_{n-1})$, let G be a Gaussian element of E, and let $v \in S^{n-1}$ be a point on the sphere. How large is $\|G(v)\|_2$? In the case $E = H_D$, the answer does not depend on v;

$$\mathbb{E}\,\|G(v)\|_2 \sim \sqrt{n}$$

for any $v \in S^{n-1}$. However, for most of the full subspaces E, the answer does depend on the point v. In contrast to this, we analyze a global quantity $\tilde{\kappa}(P)$ for a random element $P \in E$ which only depends on the polynomial system P but not on any point v. So we need to measure how much fluctuation is caused by the structure of E on the pointwise estimates $\|P(v)\|_2$. In the preliminaries section, we defined the notion of dispersion constant $\sigma(E)$ to measure this fluctuation. Let us recall the definition.

Definition 4.1 (Dispersion Constant)**.** *Let $E_1 \subseteq H_{d_1}$ be a linear subspace, and let $\sigma_{\min}(E_1)$ and $\sigma_{\max}(E_1)$ be as defined in the preliminaries section. Then*

$$\sigma(E_1) := \frac{\sigma_{\max}(E_1)}{\sigma_{\min}(E_1)}$$

is the dispersion constant of E_1.
For $E_i \subseteq H_{d_i}$, and $E = (E_1, \dots, E_{n-1})$, recall that $\sigma_{\max}(E) := \max_i \sigma_{\max}(E_i)$ and $\sigma_{\min}(E) := \min_i \sigma_{\min}(E_i)$. Then

$$\sigma(E) := \frac{\sigma_{\max}(E)}{\sigma_{\min}(E)}$$

is the dispersion constant of E.

From the definition, we always have $\sigma(E_i) \geq 1$, and we clearly have $\sigma(H_{d_i}) = 1$. Moreover, $\sigma(E_i)$ is not determined by the dimension of E_i. For instance, consider a point $v \in S^{n-1}$ and polynomials vanishing on v: $F_v := \{p \in H_{d_i} : p(v) = 0\}$. Then $\dim(F_v) = \dim(H_{d_i}) - 1$ and $\sigma_{\min}(F_v) = 0$. On the other hand, one can construct subspaces E_i of small dimension where $\sigma(E_i)$ is close to 1. So rather than considering $\sigma(E)$ as linear algebraic notion we need pursue a more geometric understanding of the quantity.

Let H_{d_1} be space of homogenous degree d_1 polynomials and let $q_v = \langle v, x\rangle^{d_1}$ be the pointwise evaluation polynomials as introduced in the preliminaries section. We define the following set, which can also be interpreted as the image of the Veronese map.

$$B_{d_1} := \{q_v : v \in S^{n-1}\}$$

For a linear subspace $E_1 \subseteq H_{d_1}$, and orthogonal projection map Π_{E_1} on E_1, we can interpret $\sigma(E_1)$ as follows:

$$\sigma(E_1) = \frac{\max_{x \in B_{d_1}} \|\Pi_{E_1}(x)\|_W}{\min_{x \in B_{d_1}} \|\Pi_{E_1}(x)\|_W}$$

It is easy to prove that $T \circ B_{d_1} = B_{d_1}$ for all $T \in O(n)$. Invariance of B_{d_1} under $O(n)$ action directly implies that $\sigma(E_1) = \sigma(T \circ E_1)$ for all $T \in O(n)$. Hence, we consider $\sigma(E_1)$ as a quantity of the $O(n)$-orbit of E_1. Similarly, we can consider the obvious action of $O(n)^{n-1}$ on $(n-1)$-tuples of linear structures $E = (E_1, \dots, E_{n-1})$. Following the definitions, it again follows that for any $T \in O(n)^{n-1}$ we have $\sigma(T \circ E) = \sigma(E)$. So dispersion constant of linear structure E is invariant under any isometric change of variables.

To make this discussion a bit more concrete, we would like to write down $\sigma(E)$ quantity in terms of a basis.

Lemma 4.2. *Let $D = (d_1, \ldots, d_{n-1})$ be a degree vector, and suppose $E \subseteq H_D$ with $E = (E_1, \ldots, E_{n-1})$ where $E_i \subseteq H_{d_i}$ are linear spaces with $\dim(E_i) = t_i$. Let $\{u_{ij}\}_{j=1}^{t_1}$ be orthonormal basis for E_i with respect to Bombieri-Weyl inner product, and set*

$$\bar{u}_i(x) := \left(\sum_{j=1}^{t_i} u_{ij}(x)^2 \right)^{\frac{1}{2}}$$

Then, for all $1 \leq i \leq n-1$, we have

$$\sigma_{\min}(E_i) = \min_{x \in S^{n-1}} \bar{u}_i(x) \text{ and } \sigma_{\max}(E_i) = \max_{x \in S^{n-1}} \bar{u}_i(x)$$

Proof. We will write the proof for E_1. We define an auxilliary polynomial:

$$g_v(x) := \sum_{j=1}^{t_1} u_{1j}(v) u_{1j}(x)$$

Note that for any $f \in E_1$, we have $f(x) = \sum_{j=1}^{t_1} \langle f, u_{1j} \rangle_W u_{1j}(x)$. This implies the following:

$$\langle f, g_v \rangle_W = \sum_{j=1}^{t_1} \langle f, u_j \rangle_W u_{1j}(v) = f(v)$$

Now we compare g_v with $\Pi_E(q_v)$:

$$\langle f, g_v \rangle_W = f(v) = \langle f, \Pi_E(q_v) \rangle_W$$

So we have $g_v = \Pi_{E_1}(q_v)$ and $\|\Pi_{E_1}(q_v)\|_W = \|g_v\|_W = (\sum_{j=1}^{t_1} u_j(v)^2)^{\frac{1}{2}}$ which completes the proof. ■

This lemma shows that $\sigma(E)$ can be computed by running an optimization procedure over sums of squares of basis polynomials. So in principle, for a given linear structure E, we can compute $\sigma(E)$ to quantify conditioning aspects of E. One can also wonder if for a given linear structure E, do we always gain speed by running numerical algorithms in a structure preserving way? Our condition number results is meant to answer this question. Results in previous section involve $\dim(E)$ and $\sigma(E)$. For structured spaces E, one has in mind that $\dim(E)$ is perhaps exponentially smaller than $N = \sum_{i=1}^{n-1} \binom{n+d_i-1}{d_i}$. On the other hand, $\sigma(E)$ can become arbitrarily large. So how large is $\sigma(E)$ typically for a low-dimensional linear structure? Do we typically get a speed-up or not? Our next section answers this question in a quite positive way: For a typical low-dimensional structure, we gain exponentially in terms of condition number analysis.

4.1. **Dispersion Constant of a Typical Structure.** We are interested in analyzing $\sigma(E)$ for a random linear m dimensional subspace E of H_d drawn from Haar measure on the $Gr(m, \dim(H_d))$. We will use the following standard notion from high dimensional probability.

Definition 4.3 (Gaussian Complexity)**.** *Let $X \subseteq \mathbb{R}^n$ be a set, then the Gaussian complexity of X denoted by $\gamma(X)$ is defined as follows:*

$$\gamma(X) := \mathbb{E} \sup_{x \in X} |\langle G, x \rangle|$$

where G is distributed according to standard normal distribution $\mathcal{N}(0,\mathbb{I})$ on $\mathbb{R}^n$.

A corollary of Lemma 2.1 and Lemma 2.8 is the following.

Corollary 4.4 (Gaussian Complexity of the Veronese Embedding)**.** *Let H_d be the vector space of degree d homogenous polynomials in n variables. Let $B_d := \{q_v : v \in S^{n-1}\}$ as defined before. Then*

$$\gamma(B_d) \le c\sqrt{n}\log(ed)$$

for a universal constant c.

Proof. We need to consider a Gaussian element G in the vector space H_d. Note that for $G \sim \mathcal{N}(0,\mathbb{I})$ in H_d, $\langle G, \sqrt{\binom{d}{\alpha}}x^\alpha\rangle_W \sim \mathcal{N}(0,1)$ since $\sqrt{\binom{d}{\alpha}}x^\alpha$ is an orthonormal basis w.r.t to the Bombieri-Weyl inner product. This means Gaussian elements of H_d are included in our model of randomness for the special case $K = 1$. Since $\sigma_{\max}(H_d) = 1$, Lemma 2.1 gives us the following estimate for pointwise evaluations of the Gaussian element $G \sim \mathcal{N}(0,\mathbb{I})$ in H_d.

$$\text{Prob}\{|G(v)| \ge t\} \le \exp\left(1 - \frac{t^2}{2}\right)$$

Note that $\|G\|_\infty = \max_{v\in S^{n-1}} |G(v)| = \max_{q_v\in B_d} |\langle G, q_v\rangle|$. So to estimate Gaussian complexity of the Veronese embedding B_d, we need to estimate $\mathbb{E}\,\|G\|_\infty$. Let $\mathcal{N}$ be a δ-net on the sphere S^{n-1}. We then have

$$\text{Prob}\{\max_{v\in\mathcal{N}} |G(v)| \ge t\} \le |N| \exp\left(1 - \frac{t^2}{2}\right)$$

Setting $\delta = \frac{1}{d}$, and using Lemma 2.8 for $t \ge a_1\sqrt{n}\log(ed)$ gives the following.

$$\text{Prob}\{\|G\|_\infty \ge a_1 t\sqrt{n}\log(ed)\} \le |\mathcal{N}| \exp\left(1 - \frac{a_1^2 t^2 n\log(ed)}{2}\right)$$

It is known that $|\mathcal{N}| \le \exp(a_0 n\log d)$. So we have

$$|\mathcal{N}| \exp\left(1 - \frac{a_1^2 t^2 n\log(ed)}{2}\right) \le \exp(1 - a_2 t^2 n\log(ed))$$

for some constant a_2. Hence,

$$\text{Prob}\{\|G\|_\infty \ge a_1 t\sqrt{n}\log(ed)\} \le \exp(1 - a_2 t^2 n\log(ed))$$

Using this inequality one can routinely derive the estimate for $\mathbb{E}\,\|G\|_\infty$. ■

Controlling Gaussian complexity of a geometric set provides control on the behavior of random linear maps on the set. This is a general principle behind most of the dimensionality reduction tools in data science and engineering. For the particular case of random projections, we recall a well-known estimate below. This estimate can be found in many references, we refer the reader to [32] for a nice exposition. The above statement has first appeared in [26].

Theorem 4.5. *Let F be a random m dimensional subspace of R^n drawn from Haar measure on $Gr(m,n)$, and let P_F be orthogonal projection map on F. Let $X \subseteq \mathbb{R}^n$ be a set. Then*

$$\sup_{x\in X} \left|\sqrt{n}\,\|P_F(x)\| - \sqrt{m}\,\|x\|\right| \le Ct\gamma(X)$$

with probability greater than $1 - e^{-t^2}$.

Corollary 4.6. *Let F be a random m dimensional subspace of H_d drawn from the Haar measure on $Gr(m, \dim(H_d))$, where $m \geq 16Cn\log(ed)^2$. Then*

$$\sigma(F) \leq \frac{\sqrt{m} + Ct\sqrt{n}\log(ed)}{\sqrt{m} - Ct\sqrt{n}\log(ed)}$$

with probability greater than $1 - e^{-t^2}$, where C is an absolute constant.

Proof. Since $\|q_v\|_W = 1$ for all $v \in S^{n-1}$, applying Theorem 4.5 to the set B_d we have

$$\sup_{x \in B_d} \left| \binom{n+d-1}{d}^{\frac{1}{2}} \|\Pi_F(x)\| - \sqrt{m} \right| \leq Ct\sqrt{n}\log(ed)$$

with probability greater than $1 - e^{-t^2}$ for all $t \geq 1$. Equivalently,

$$\frac{\sqrt{m} - Ct\sqrt{n}\log(ed)}{\binom{n+d-1}{d}^{\frac{1}{2}}} \leq \sigma_{\min}(F) \leq \sigma_{\max}(F) \leq \frac{\sqrt{m} + Ct\sqrt{n}\log(ed)}{\binom{n+d-1}{d}^{\frac{1}{2}}}$$

with probability greater than $1 - e^{-t^2}$. ■

We can now prove Theorem 1.4 as stated in the introduction.
Proof of Theorem 1.4: We use Corollary 4.6 with $t' = t\log(n)$ for all E_i in the theorem statement of 1.4, then we do a union bound. ■

References

[1] P. Bagos, I. Valavanis, and I. Emiris, *"Algebraic methods for Protein structure prediction from RDC data,"* in: Intern. Workshop on Algebraic Biology 2007, Austria, Springer, LNCS 4545.

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

[3] 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.

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

[5] 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.

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

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

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

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

[10] 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.

[11] Tianran Chen; Tsung-Lin Lee; and Tien-Yien Li, *"Mixed cell computation in Hom4PS-3,"* Journal of Symbolic Computation (2017), pp. 516–534.

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

[13] 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

[14] 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.

[15] 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.

[16] 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`.)

[17] 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.

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

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

[20] Alperen A. Ergür, Grigoris Paouris, J. Maurice Rojas, *"Probabilistic Condition Number Estimates For Real Polynomial Systems I: A Broader Family Of Distributions"*, Foundations of Computational Mathematics, to appear. DOI: 10.1007/s10208-018-9380-5

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

[22] Birkett Huber, Bernd Sturmfels, *"A Polyhedral Method for Solving Sparse Polynomial Systems, "* Mathematics of Computation Vol. 64, No. 212 (Oct., 1995), pp. 1541–1555.

[23] Badal Joshi and Anne Shiu, *"Which small reaction networks are multistationary?,"* SIAM Journal on Applied Dynamical Systems, vol. 16, pp. 802–833 (2017).

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

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

[26] Bo'az Klartag and Shahar Mendelson, *Empirical processes and random projections* , J. Functional Analysis, Vol. 225, no. 1 (2005) 229–245.

[27] 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.

[28] 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.

[29] 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.

[30] Pierre Lairez, *"Rigid continuation paths I. Quasilinear average complexity for solving polynomial systems,"* https://arxiv.org/abs/1711.03420 .

[31] Michel Ledoux, *"The Concentration of Measure Phenomemon,"* Mathematical Surveys & Monographs, Book 89, AMS Press, 2005.

[32] C. Liaw, A. Mehrabian, Y. Plan, and R. Vershynin, *"A simple tool for bounding the deviation of random matrices on geometric sets,"* Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 2014–2016, Lecture Notes in Mathematics 2169, Springer, 2017, pp. 277–299.

[33] 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.

[34] Gregorio Malajovich, *"Complexity of Sparse Polynomial Solving: Homotopy on Toric Varieties and the Condition Metric,"* Foundations of Computational Mathematics, https://www.springerprofessional.de/en/complexity-of-sparse-polynomial-solving-homotopy-on-toric-variet/15421270 .

[McL05] Andrew McLennan, *"The Expected Number of Nash Equilibria of a Normal Form Game,"* Econometrica **73** (2005), pp. 141–174.

[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.

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

[37] G. Paouris and P. Pivovarov, *"Randomized Isoperimetric Inequalities"*, IMA Volume "Discrete Structures: Analysis and Applications" (Springer)
[38] Mark Rudelson and Roman Vershynin, *"The Littlewood-Offord Problem and Invertibility of Random Matrices,"* Adv. Math. 218 (2008), no. 2, pp. 600–633.
[39] Mark Rudelson and Roman Vershynin, *"The Smallest Singular Value of Rectangular Matrix,"* Communications on Pure and Applied Mathematics 62 (2009), pp. 1707–1739.
[40] 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.
[41] Igor R. Shafarevich, *Basic Algebraic Geometry 1: Varieties in Projective Space*, 3rd edition, Springer-Verlag (2013).
[42] Mike Shub and Steve Smale, *"Complexity of Bezout's Theorem I. Geometric Aspects,"* J. Amer. Math. Soc. 6 (1993), no. 2, pp. 459–501.
[43] Steve Smale, *"Mathematical Problems for the next Century,"* The Mathematical Intelligencer, 1998, 20.2, pp. 7-15.
[44] D. A. Spielman and S.-H. Teng, *"Smoothed analysis of algorithms,"* Proc. Int. Congress Math. (Beijing, 2002), Volume I, 2002, pp. 597–606.
[45] Jan Verschelde, *"Algorithm 795: PHCpack: a general-purpose solver for polynomial systems by homotopy continuation"*, ACM Transactions on Mathematical Software (TOMS) Volume 25 Issue 2, June (1999), pp. 251–276
[46] Jan Verschelde, Pierre Verlinden, Ronald Cools, *"Homotopies Exploiting Newton Polytopes for Solving Sparse Polynomial Systems, "* SIAM J. Numer. Anal. 31(3), (1994), 915–930
[47] 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
[48] Roman Vershynin , *"Four Lectures on Probabilistic Methods for Data Science"*, available at https://www.math.uci.edu/ rvershyn/papers/four-lectures-probability-data.pdf
[49] Roman Vershynin, *"Introduction to the Non-Asymptotic Analysis of Random Matrices,"* Compressed sensing, pp. 210–268, Cambridge Univ. Press, Cambridge, 2012.

Technische Universität Berlin, Institut für Mathematik, Sekretariat MA 3-2, Strasse 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