On Smale's 17th Problem: A Probabilistic Positive Solution.

Carlos Beltrán [1,2,3], Luis Miguel Pardo [1,2]

August 17, 2006

Abstract

Smale's 17th Problem asks "Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average [for a suitable probability measure on the space of inputs], in polynomial time with a uniform algorithm?" We present a uniform *probabilistic* algorithm for this problem and prove that its complexity is polynomial. We thus obtain a partial positive solution to Smale 17th Problem.

1 Introduction

In the first half of the nineties, Shub and Smale introduced a seminal conception of the foundations of numerical analysis. They focused on a theory of numerical polynomial equation solving in the series of papers [SS93a, SS93b, SS93c, SS94, SS96]. Other authors also treated this approach in [Ren87, BCSS98, Ded01, Ded06, Kim89, Mal94, MR02, Yak95, CPHM01, CPM03] and references therein.

In these pages we complete part of the program initiated in the series [SS93a] to [SS94]. As in [SS94], the input space is the space of systems of multivariate homogeneous polynomials with dense encoding and fixed degree list. Namely, for every positive integer $l \in \mathbb{N}$, let $H_l \subseteq \mathbb{C}[X_0, \ldots, X_n]$ be the vector space of all homogeneous polynomials of degree l. For a list of degrees $(d) := (d_1, \ldots, d_n) \in \mathbb{N}^n$, let $\mathcal{H}_{(d)}$ be the set of all systems $f := [f_1, \ldots, f_n]$ of homogeneous polynomials of respective degrees $\deg(f_i) = d_i,\ 1 \leq i \leq n$. In other words, $\mathcal{H}_{(d)} := \prod_{i=1}^n H_{d_i}$. We denote by $d := \max\{d_i : 1 \leq i \leq n\}$ the maximum of the degrees. Note that if $(d) = (1) := (1, \ldots, 1)$, the vector

[1] *Dept. de Matemáticas, Estadística y Computación. F. de Ciencias. U. Cantabria. E–39071 SANTANDER, Spain. email: beltranc@unican.es, luis.pardo@unican.es.*

[2] *Research was partially supported by MTM2004-01167*

[3] *Granted by FPU program, Goverment of Spain.*

[4] *AMS Classification: Primary 65H10, 65H20. Secondary 58C35.*

space $\mathcal{H}_{(1)}$ can be identified with the space of all $n \times (n+1)$ complex matrices. Namely,

$$\mathcal{H}_{(1)} = \mathcal{M}_{n\times(n+1)}(\mathbb{C}) = \mathbb{C}^{n\times(n+1)}.$$

We denote by $N+1$ the complex dimension of the vector space $\mathcal{H}_{(d)}$. Note that $N+1$ is the input length for dense encoding of multivariate polynomials. For every system $f \in \mathcal{H}_{(d)}$, we also denote by $V(f)$ the projective algebraic variety of its common zeros. Namely,

$$V(f) := \{x \in \mathbb{P}_n(\mathbb{C}) : f_i(x) = 0, 1 \leq i \leq n\},$$

where $\mathbb{P}_n(\mathbb{C})$ is the n-dimensional complex projective space. Note that with this notation $V(f)$ is always a non-empty projective algebraic variety.

The following is a preliminary version of the main outcome of this manuscript. For a full statement see Theorem 6 of Section 1.2.

Theorem 1 *There is a bounded error probabilistic numerical analysis procedure that solves most systems of multivariate homogeneous polynomial equations with running time polynomial in*

$$n, N, d.$$

The probability that a randomly chosen system $f \in \mathcal{H}_{(d)}$ is solved by this procedure is greater than

$$1 - \frac{1}{N}.$$

This theorem is a positive, although probabilistic, answer to Problem 17 in [Sma00]. Namely, we give a positive answer to the following question.

Problem 1 (Smale, 2000) *"Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average , in polynomial time with a uniform algorithm?"*

We present a uniform *probabilistic* algorithm for this problem and prove that its complexity is polynomial. We thus obtain a partial positive solution to Smale's 17th Problem in [Sma00].

Let us now review some technical notions we'll need to rigorously state our underlying main algorithm.

1.1 Advances and Overcoming Prior Difficulties

This subsection is devoted to introduce the algorithm that satisfies all the claims of Theorem 1. It is an algorithm based on homotopic deformation (cf. [GZ79, HSS01, Mor83, Mor86, MS87b, MS87a]) as established in the series of papers by M. Shub and S. Smale (mainly [SS93a, SS96, SS94]). In terms of algorithmic design, it belongs to the family of "non–universal polynomial equation solvers" as defined in [Par95, BP06a, CGH$^+$03].

We describe our algorithm in four levels. At each level we also introduce the required notions and the main notations to be used in the sequel.

As usual, the first level is devoted to fix the input/output structure of the procedure. At this level we recall the notion of projective Newton's operator (as in [Shu93]) and the notion of approximate zero (as in [SS93a, BCSS98].

At a second level we fix the algorithmic scheme we use: Homotopic deformation with prescribed resource data.

After this second level we are in conditions to fix the main drawback of this kind of algorithmic design: where to begin the homotopy in order to achieve tractable complexity bounds (i.e. a number of steps bounded by a polynomial in the input length). At this level we also discuss the notion of ε-efficient initial pair with prescribed resource function.

Third level is devoted to recall the main algorithmic scheme used to prove the main outcome in [SS94]. At this level we can also explain why the main outcome of [SS94] does not solve Problem 17 of [Sma00]. In [SS94], Shub & Smalke introduced an scheme which is either constructive nor proven to be uniform.

We then arrive to a fourth level of detail and we describe our algorithm satisfying the claims of Theorem 1.

Input/Output Structure. Our algorithm takes as input a system of homogeneous polynomial equations $f \in \mathcal{H}_{(d)}$ and outputs local information close to some (mostly one) of the zeros $\zeta \in V(f)$. The local information we compute is the information provided by an *approximate zero* $z \in \mathbb{P}_n(\mathbb{C})$ of f associated with some zero $\zeta \in V(f)$ (in the sense of [SS93a, BCSS98]).

Projective Newton's operator was introduced in [Shu93]. Let $(f, z) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ be a pair. Let $z^{\perp} := \{w \in \mathbb{C}^{n+1} : \langle w, z \rangle = 0\}$ be the tangent space of $\mathbb{P}_n(\mathbb{C})$ at z. If the restriction of the tangent mapping of f at z, $T_z f := d_z f \mid_{z^{\perp}}$, is nonsingular, we define the Newton iteration of f at z as:

$$N_f(z) := z - (T_z f)^{-1} f(z) \in \mathbb{P}_n(\mathbb{C}).$$

According to [SS93a], an approximate zero $z \in \mathbb{P}_n(\mathbb{C})$ of a system $f \in \mathcal{H}_{(d)}$ with associated zero $\zeta \in V(f) \subseteq \mathbb{P}_n(\mathbb{C})$ is a projective point such that the sequence of iterates $(N_f^k(z))_{k \in \mathbb{N}}$ is well-defined and converges to the actual zero $\zeta \in V(f)$ at a speed which is doubly exponential in the number of iterations. In this sense, an approximate zero is an ideal output for a polynomial system solving algorithm: Approximate zeroes occupy few bits on average (cf. [CPM03]), yet they are close enough to true zeroes for Newton's operator to converge quickly and efficiently yield any desired level of accuracy.

The algorithms we consider will have the following input/output structure:

Input: A system of homogeneous polynomial equations $f \in \mathcal{H}_{(d)}$.

Output: An approximate zero $z \in \mathbb{P}_n(\mathbb{C})$ of f associated with some zero $\zeta \in V(f)$.

Such algorithms are built with the possibility of measure zero "bad" set of inputs (usually called ill-conditioned, singular, or degenerate) on which the algorithm fails. Via certain modifications (such as modifying the definition of approximate zero, restricting the inputs to integer coefficients, or considering the nearest problem with given singularity structure), it is possible to solve polynomial systems in complete generality. We will not pursue these more technical extensions, but let us least point out to the reader that these issues are highly non-trivial already for numerical linear algebra and the setting of one polynomial in one variable (see e.g. [Zen05, LE99, ME98]). Extensions of α-theory and deflation methods for degenerate roots are studied in [Lec01, Lec02, Yak00, Par95, GHMP95, GHMP97, GHH$^+$97, GHM$^+$98, LVZ, GLSY05b, GLSY05a, DS01, BP06b] .

Let $\Sigma \subseteq \mathcal{H}_{(d)}$ be the set of systems f such that $V(f)$ contains a singular zero. We call Σ the discriminant variety. These pages are mainly concerned with procedures that solve systems without singular zeros (i.e., systems $f \in \mathcal{H}_{(d)} \setminus \Sigma$).

Algorithmic Scheme. Our main algorithmic scheme is *Homotopic Deformation* in the projective space (as described in [SS96, SS94]): Given $f, g \in \mathcal{H}_{(d)} \setminus \Sigma$, we consider the "segment" of systems "between" f and g,

$$\Gamma := \{f_t := (1-t)g + tf, t \in [0,1]\}.$$

If $\Gamma \cap \Sigma = \emptyset$, there are non-intersecting and smooth curves of equations-solutions associated with this segment:

$$C_i(\Gamma) := \{(f_t, \zeta_t) : \zeta_t \in V(f_t), t \in [0,1]\}, \qquad 1 \le i \le \mathcal{D} := \prod_{i=1}^{n} d_i.$$

Then, Newton's operator may be used to follow closely one of these curves $C_i(\Gamma)$ in the product space $\mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$. This procedure computes some approximate zero z_1 associated with some zero of f (i.e., $t = 1$) starting at some approximate zero z_0 associated with g (i.e., from $t = 0$). The following definition formalizes this strategy.

Definition 2 *A Homotopic Deformation scheme (HD for short) with initial data* $(g, z_0) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ *and resource function* $\varphi : \mathcal{H}_{(d)} \times \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ *is an algorithmic scheme based on the following strategy:*

Input: $f \in \mathcal{H}_{(d)}$, $\varepsilon \in \mathbb{R}^+$.

Perform $\varphi(f,\varepsilon)$ "homotopy steps" following the segment $(1-t)g+tf$, $t \in [0,1]$, starting at (g,z_0), where z_0 is an approximate zero of g associated with some zero $\zeta_0 \in V(g)$.

Output:

either failure, or
an approximate zero $z_1 \in \mathbb{P}_n(\mathbb{C})$ of f.

An algorithm following the HD scheme is an algorithm that constructs a polygonal line P with $\varphi(f,\varepsilon)$ vertices. The initial vertex of P is the point (g,z_0) and its final vertex is the point (f,z_1) for some $z_1 \in \mathbb{P}_n(\mathbb{C})$. The output of the algorithm is the value $z_1 \in \mathbb{P}_n(\mathbb{C})$. The polygonal is constructed by "homotopy steps" (path following methods) that go from one vertex to the next. Hence, $\varphi(f,\varepsilon)$ is the number of homotopy steps performed by the algorithm. Different subroutines have been designed to perform each one of these "homotopy steps". One of them is projective Newton's operator as described in [Shu93, SS93a, Mal94].

The positive real number ε is currently used both to control the number of steps (through the function $\varphi(f,\varepsilon)$) and the probability of failure (i.e., the probability that a given input $f \in \mathcal{H}_{(d)}$ is not solved in $\varphi(f,\varepsilon)$ steps with initial pair (g,z_0)).

Efficient Initial Pairs. We desire initial pairs with optimal tradeoff between number of steps and probability of failure. We clarify this as follows.

Definition 3 *Let $p \in \mathbb{R}[T_1,T_2,T_3,T_4]$ be some fixed polynomial. Let $\varepsilon > 0$ be a positive real number. We say that an initial pair $(g,z_0) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ is ε-efficient for HD if the HD scheme with initial pair (g,e_0) and resource function*

$$\varphi(f,\varepsilon) := p(n,N,d,\varepsilon^{-1}),\ \ \forall f \in \mathcal{H}_{(d)},\ \varepsilon > 0,$$

satisfies the following property:

" For a randomly chosen system $f \in \mathcal{H}_{(d)}$, the probability that HD outputs an approximate zero of f is at least $1-\varepsilon$".

In order to simplify notations, from now on we consider the polynomial p fixed as follows:

$$p(n,N,d,\varepsilon^{-1}) := 18 \cdot 10^4 n^5 N^2 d^3 \varepsilon^{-2},$$

for every $n,N,d \in \mathbb{N} \setminus \{0\}$, and for every $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$.

The main outcome in [SS94] is that *for every positive real number $\varepsilon > 0$, there is at least one ε-efficient initial pair $(g_\varepsilon,\zeta_\varepsilon) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$.* This statement was a major breakthrough for the efficient numerical resolution of polynomial systems, and its impact is only slowly beginning to be understood. These ε-efficient pairs are used as follows.

Input: $f \in \mathcal{H}_{(d)}$, $\varepsilon \in \mathbb{R}^+$.

- Compute $(g_\varepsilon, \zeta_\varepsilon)$ (the ε-efficient initial pair whose existence is guaranteed by [SS94]).
- Perform $p(n, N, d, \varepsilon^{-1})$ homotopy steps following the segment $(1-t)g + tf$, $t \in [0, 1]$, starting at $(g_\varepsilon, \zeta_\varepsilon)$.

Output:

either failure, or
an approximate zero $z \in \mathbb{P}_n(\mathbb{C})$ of f.

Note that this procedure may output *failure* instead of giving an approximate zero of f. However, the probability that the procedure does the former is bounded above by ε, and we can at least control ε.

However, the procedure has three main drawbacks. First of all, the authors of [SS94] prove the existence of some ε-efficient initial pair, but they give no hint about how to compute such a pair $(g_\varepsilon, \zeta_\varepsilon)$. Note that if there is no method to compute $(g_\varepsilon, \zeta_\varepsilon)$, then the previous scheme is not properly an algorithm (you cannot "write" $(g_\varepsilon, \zeta_\varepsilon)$ and thus you cannot start computing). Shub and Smale used the term "quasi-algorithm" to explain the result they obtained, whereas Problem 17th in [Sma00] asks for a "uniform algorithm". In a broad sense, the last scheme is close to an "oracle machine" where the initial pair $(g_\varepsilon, \zeta_\varepsilon)$ is given by some undefinable oracle. Moreover, the lack of hints on ε-efficient initial pairs leads both to Shub & Smale's Conjecture (as in [SS94]) and to Smale's 17th Problem.

A second drawback is the dependence of $(g_\varepsilon, \zeta_\varepsilon)$ on the value ε.

Thirdly, the reader should observe that the initial pair $(g_\varepsilon, \zeta_\varepsilon)$ must be solved before we can perform any computation. Namely, ζ_ε must be an approximate zero of g_ε. In fact, Shub & Smale in [SS94] proved the existence of such $(g_\varepsilon, \zeta_\varepsilon)$ assuming that ζ_ε is a true zero of g_ε (i.e., $\zeta_\varepsilon \in V(g_\varepsilon)$). However, using [SS93a, Main Thm.] you can relax this condition to assume that ζ_ε is an approximate zero associated with some zero of g_ε. This means that you need to make some precomputation by solving g_ε provided that you know it.

Thus, any algorithm based on this version of HD requires some "a priori" tasks not all of them simple:
First, you have to detect some system of equations g_ε such that some of

its zeros ζ_ε yields an ε-efficient initial pair $(g_\varepsilon, \zeta_\varepsilon)$. Secondly, you need to "solve" the system g_ε in order to compute some approximate zero associated with the exact solution ζ_ε.

As computing either an exact or an approximate zero of some unknown g_ε does not seem a good choice, we should proceed in the opposite manner: Start at some complex point $\zeta_\varepsilon \in \mathbb{P}_n(\mathbb{C})$, given a priori. Then, prove that there is a system g_ε vanishing at ζ_ε such that $(g_\varepsilon, \zeta_\varepsilon)$ is an ε-efficient initial pair. The existence of such a kind of system g_ε for any given $\zeta_\varepsilon \in \mathbb{P}_n(\mathbb{C})$ easily follows from the arguments in [SS94]. But, once again, no hint on how to find g_ε from ζ_ε seems to be known.

Questor Sets. In these pages we exhibit a solution to these drawbacks. We choose a probabilistic approach and, hence, we can give an efficient uniform (i.e. true) algorithm that solves most systems of multivariate polynomial equations. This is achieved using the following notion.

Definition 4 *A set $\mathcal{G} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ is called a correct test set (also questor set) for efficient initial pairs if for every $\varepsilon > 0$ the probability that a randomly chosen pair $(g, \zeta) \in \mathcal{G}$ is ε-efficient is greater than*

$$1 - \varepsilon.$$

Note the analogy between these families of efficient initial systems and the "correct test sequences" (also "questor sets") for polynomial zero tests (as in [HS82, KP96] or [CGH$^+$03]). A similar idea to that used here (i.e. constructing a questor set for deciding where to start an iterative algorithm) has been recently developed in [HSS01]. We prove the following result.

Theorem 5 *For every degree list $(d) = (d_1, \ldots, d_n)$ there is a questor set $\mathcal{G}_{(d)}$ for efficient initial pairs that solves most of the systems in $\mathcal{H}_{(d)}$ in time which depends polynomially on the input length N of the dense encoding of multivariate polynomials.*

The existence of a questor set for initial pairs $\mathcal{G}_{(d)} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ yields another variation (of a probabilistic nature) on the algorithms based on HD schemes. First of all, note that the set $\mathcal{G}_{(d)}$ does not depend on the positive real number $\varepsilon > 0$ under consideration. Thus, we can define the following HD scheme based on some fixed questor set $\mathcal{G}_{(d)}$.

Input: $f \in \mathcal{H}_{(d)}$, $\varepsilon \in \mathbb{R}^+$.

- *Guess at random* $(g, \zeta) \in \mathcal{G}_{(d)}$.

- *Perform a polynomial (in $\varepsilon^{-1}, n, N, d$) number of homotopy steps following the segment $(1-t)g + tf$, $t \in [0,1]$, starting at (g, ζ).*

Output:

either failure, or
an approximate zero $z \in \mathbb{P}_n(\mathbb{C})$ of f.

Observe that the questor set $\mathcal{G}_{(d)}$ is independent of the value ε under consideration. However, the existence of such a questor set does not imply the existence of a uniform algorithm. In fact, a simple existential statement as Theorem 5 will not be better than the main outcome in [SS94]: We need to extract suitable elements of $\mathcal{G}_{(d)}$ *explicitly*. Hence, we exhibit an algorithmically tractable subset $\mathcal{G}_{(d)}$ which is proven to be a questor set for efficient initial pairs. It leads to a "uniform algorithm", although probabilistic. This rather technical set can be defined as follows.

1.2 An answer to Smale's 17th Problem.

Let Δ be the diagonal matrix used in [Kos93, SS93b], (see [BCSS98, pg. 236] for further bibliographical references). With this matrix, Shub & Smale defined a Hermitian product $\langle \cdot, \cdot \rangle_\Delta$ on $\mathcal{H}_{(d)}$ which is invariant under certain natural action of the unitary group $\mathcal{U}_{n+1}$ on $\mathcal{H}_{(d)}$ (see also Section 2 for details). We denote by $|| \cdot ||_\Delta$ the norm on $\mathcal{H}_{(d)}$ defined by $\langle \cdot, \cdot \rangle_\Delta$. This Hermitian product $\langle \cdot, \cdot \rangle_\Delta$ also defines a complex Riemannian structure on the complex projective space $\mathbb{P}(\mathcal{H}_{(d)})$. This complex Riemannian structure on $\mathbb{P}(\mathcal{H}_{(d)})$ induces a volume form $d\nu_\Delta$ on $\mathbb{P}(\mathcal{H}_{(d)})$ and hence a measure on this manifold. The measure on $\mathbb{P}(\mathcal{H}_{(d)})$ also induces a probability on this complex Riemannian manifold (see Section 2). Moreover, for every subset $A \subseteq \mathbb{P}(\mathcal{H}_{(d)})$ the probability $\nu_\Delta[A]$ induced by $d\nu_\Delta$ agrees with the Gaussian measure of the cone $\tilde{A}$ over A in $\mathcal{H}_{(d)}$ (i.e., $\tilde{A}$ modulo scaling yields A). In the sequel, volumes and probabilities in $\mathcal{H}_{(d)}$ and $\mathbb{P}(\mathcal{H}_{(d)})$ always refers to these probabilities and measures defined by $\langle \cdot, \cdot \rangle_\Delta$.

Let us now fix a point $e_0 := (1 : 0 : \cdots : 0) \in \mathbb{P}_n(\mathbb{C})$. Let

$$L_{e_0} := \{f = [f_1, \ldots, f_n] \in \mathcal{H}_{(d)} : f_i = X_0^{d_i - 1} \sum_{j=1}^{n} a_{ij} X_j, \ a_{ij} \in \mathbb{C}\}.$$

Let $\widetilde{V}_{e_0} \subseteq \mathcal{H}_{(d)}$ be the complex vector space of all homogeneous systems in $\mathcal{H}_{(d)}$ that vanish at e_0. Namely,

$$\widetilde{V}_{e_0} := \{f \in \mathcal{H}_{(d)} \ : \ e_0 \in V(f)\}.$$

Note that L_{e_0} is a subspace of $\widetilde{V}_{e_0}$.

Next, let $L_{e_0}^{\perp}$ be the orthogonal complement of L_{e_0} in $\widetilde{V}_{e_0}$ with respect to the Hermitian product $\langle \cdot, \cdot \rangle_{\Delta}$ (see Section 2 for details). Note that $L_{e_0}^{\perp}$ is the family of all homogeneous systems $f \in \mathcal{H}_{(d)}$ that vanish at e_0 and such that its derivative $d_{e_0}f$ also vanishes at e_0.

Let Y be the following convex subset set of the affine space $\mathbb{R} \times \mathbb{C}^{N+1}$:

$$Y := [0,1] \times B^1(L_{e_0}^{\perp}) \times B^1(\mathcal{H}_{(1)}) \subseteq \mathbb{R} \times \mathbb{C}^{N+1},$$

where $B^1(L_{e_0}^{\perp})$ is the closed ball of radius one in $L_{e_0}^{\perp}$ with respect to the canonical Hermitian metric and $B^1(\mathcal{H}_{(1)})$ is the closed ball of radius one in the space of $n \times (n+1)$ complex matrices with respect to the standard Frobenius norm. We assume Y is endowed with the product space probability.

Let

$$\tau := \sqrt{\frac{n^2+n}{N}},$$

and let us fix any mapping $\phi : \mathcal{H}_{(1)} \longrightarrow \mathcal{U}_{n+1}$ such that for every matrix $M \in \mathcal{H}_{(1)}$ of maximal rank, ϕ associates a unitary matrix $\phi(M) \in \mathcal{U}_{n+1}$ satisfying $M\phi(M)e_0 = 0$. In other words, $\phi(M)$ transforms e_0 into a vector in the kernel of M. Our statements below are independent of the chosen mapping ϕ that satisfies this property.

Let us define a mapping $\Theta : \mathcal{H}_{(1)} \longrightarrow L_{e_0}$ as follows. For $M \in \mathcal{H}_{(1)}$, let $(a_{ij} : 1 \leq i,j \leq n)$ be the entries of $M\phi(M)$. Namely,

$$M\phi(M) = \begin{pmatrix} 0 & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

Then, we define $\Theta(M) = [f_1, \ldots, f_n]$, where

$$f_i = d_i^{1/2} X_0^{d_i-1} \sum_{j=1}^{n} a_{ij} X_j.$$

Define a mapping $G_{(d)} : Y \longrightarrow \widetilde{V}_{e_0}$ as follows. For every $(t,h,M) \in Y$, let $G_{(d)}(t,h,M) \in \widetilde{V}_{e_0}$ be the system of homogeneous polynomial equations given by the identity:

$$G_{(d)}(t,h,M) := \left(1 - \tau^2 t^{\frac{1}{n^2+n}}\right)^{1/2} \frac{\Delta^{-1}h}{||h||_2} + \tau t^{\frac{1}{2n^2+2n}} \Theta\left(\frac{M}{||M||_F}\right) \in \widetilde{V}_{e_0},$$

where $\| \cdot \|_F$ is Frobenius norm. Finally, let $\mathcal{G}_{(d)}$ be the set defined by the identity:

$$\mathcal{G}_{(d)} := Im(G_{(d)}) \times \{e_0\} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C}). \tag{1}$$

Note that $\mathcal{G}_{(d)}$ is included in the product $\widetilde{V}_{e_0} \times \mathbb{P}_n(\mathbb{C})$ since all the systems f in $Im(G_{(d)}$ share a common zero e_0 (i.e. $f(e_0) = 0$). Hence initial pairs in $(g, z) \in \mathcal{G}_{(d)}$ always use the same exact zero $z = e_0$. In particular, they are all solved by construction.

We assume that the set $\mathcal{G}_{(d)}$ is endowed with the pull-back probability distribution obtained from Y via $G_{(d)}$. Namely, in order to choose a random point in $\mathcal{G}_{(d)}$, we choose a random point $y \in Y$, and we compute $(G_{(d)}(y), e_0) \in \mathcal{G}_{(d)}$. We present our main result.

Theorem 6 (Main Result) *With the above notation, the set $\mathcal{G}_{(d)}$ defined by identity (1) is a questor set for efficient initial pairs in $\mathcal{H}_{(d)}$.*

More precisely, for every positive real number $\varepsilon > 0$, the probability that a randomly chosen data $(g, e_0) \in \mathcal{G}_{(d)}$ is ε-efficient is greater than

$$1 - \varepsilon.$$

In particular, for these ε-efficient pairs $(g, e_0) \in \mathcal{G}_{(d)}$, the probability that a randomly chosen input $f \in \mathcal{H}_{(d)}$ is solved by HD with initial data (g, e_0) performing $O(n^5 N^2 d^3 \varepsilon^{-2})$ homotopy steps is at least

$$1 - \varepsilon.$$

As usual, the existence of questor sets immediately yields a uniform probabilistic algorithm. This is Theorem 1 above, which is an immediate consequence of Theorem 6. The following corollary shows how this statement applies.

Corollary 7 *There is a bounded error probability algorithm that solves most homogeneous systems of cubic equations (namely inputs are in $\mathcal{H}_{(3)}$) in time of order*

$$O(n^{18}\varepsilon^{-2}),$$

with probability greater than $1 - \varepsilon$.

Taking $\varepsilon = \frac{1}{n^2}$ for instance, this probabilistic algorithm solves a cubic homogeneous system in running time at most $O(n^{22})$ with probability greater than $1 - \frac{1}{n^2}$.

However, randomly choosing a pair $(g, e_0) \in \mathcal{G}_{(d)}$ is not exactly what a computer can perform. Under Church's Thesis, computing is discrete. Thus, we need a discrete set of ε-efficient initial systems. This is achieved by the following argument (that follows similar arguments in [CPM03]).

Observe that $Y \subseteq \mathbb{R} \times \mathbb{C}^{N+1}$ may be seen to be a real semi-algebraic set under the identification $\mathbb{R} \times \mathbb{C}^{N+1} \equiv \mathbb{R}^{2N+3}$. Let $H \geq 0$ be a positive integer number. Let $\mathbb{Z}^{2N+3} \subseteq \mathbb{R}^{2N+3}$ be the lattice consisting of the integer points in $\mathbb{R}^{2N+3}$. Let Y^H be the set of points defined as follows:

$$Y^H := Y \cap \frac{1}{H}\mathbb{Z}^{2N+3},$$

where $\frac{1}{H}\mathbb{Z}^{2N+3}$ is the lattice given by the equality:

$$\frac{1}{H}\mathbb{Z}^{2N+3} := \{\frac{z}{H} : z \in \mathbb{Z}^{2N+3}\}.$$

For any positive real number $H > 0$, we denote by $\mathcal{G}_{(d)}^H \subseteq \mathcal{G}_{(d)}$ the finite set of points given by the equality:

$$\mathcal{G}_{(d)}^H := \{(G_{(d)}(y), e_0) : y \in Y^H\}.$$

Then, the following statement also holds.

Theorem 8 *There exists a universal constant $C > 0$ such that for every two positive real numbers $\varepsilon > 0, H > 0$ satisfying*

$$\log_2 H \geq Cn^2N^3 \log_2 d + 2\log_2 \varepsilon^{-1},$$

the following properties hold.

- *The probability (uniform distribution) that a randomly chosen data $(g, e_0) \in \mathcal{G}_{(d)}^H$ is ε-efficient is greater than*

$$1 - 2\varepsilon.$$

- *In particular, for these ε-efficient initial pairs $(g, e_0) \in \mathcal{G}_{(d)}^H$, the probability that a randomly chosen input $f \in \mathcal{H}_{(d)}$ is solved by HD with initial data (g, e_0) performing $O(n^5N^2d^3\varepsilon^{-2})$ steps is at least*

$$1 - \varepsilon.$$

The lattice estimates in Theorem 8 immediately imply that our probabilistic polynomial algorithm can be implemented on any standard digital computer.

Theorem 6 and its consequences thus represent a small step forward in the theory introduced by Shub and Smale. It simply shows the existence of a uniform, although probabilistic, algorithm that computes approximations of some of the zeros of solution varieties for most homogeneous systems of polynomial equations in time which depends polynomially on the input length. [5]

This paper is structured as follows. In Section 2 we detail the notation we will use, and we continue a series of results appearing in [Shu93, SS93a, SS93b, SS93c, SS94, SS96] that we use to prove our main theorems. We include a brief introduction to the projective Newton operator and the Homotopy Method in Section 3, although we encourage the reader to see this in its original context in [Shu93, SS93a, SS94] or [BCSS98]. Section 5 is devoted to proving Theorem 6 (and hence Theorem 1). Finally, Section 6 contains the proof of Theorem 8.

[5] In the terminology of [Par95, CGH$^+$03, BP06a] this simply means that there is a non universal polynomial equation solver running in probabilistic polynomial time.

2 Basic Notation

2.1 Metrics

For every Hermitian vector space $(F, \langle \cdot, \cdot \rangle)$ of complex dimension m and for every nonsingular matrix $A \in GL(m, \mathbb{C})$, we denote by $\langle \cdot, \cdot \rangle_A : F \times F \longrightarrow \mathbb{C}$ the Hermitian product given by the following identity:

$$\langle x, y \rangle_A := \langle Ax, Ay \rangle,$$

for all $x, y \in F$. Let us denote by $||\cdot||$ and $||\cdot||_A$ the norms on F respectively defined by the Hermitian products $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle_A$.

For every positive real number $t \in \mathbb{R}_+$, we shall use the notations $S^t(F), B^t(F)$, $S^t_A(F), B^t_A(F)$ to denote respectively the spheres and closed balls in F of radius t centered at the origin with respect to the corresponding Hermitian products. For every subspace $L \subseteq F$, we shall denote by $L^\perp$ the orthogonal complement of L in F with respect to some specified Hermitian metric.

As in the Introduction, for every positive integer number $l \in \mathbb{N}$, let $H_l \subseteq \mathbb{C}[X_0, \ldots, X_n]$ be the vector space of all homogenous polynomials of degree l with complex coefficients. The monomial basis of H_l can be identified with the set of multi-indices

$$\mathbb{N}^{n+1}_l := \{\mu = (\mu_0, \mu_1, \ldots, \mu_n) \in \mathbb{N}^{n+1} \ : \ |\mu| := \mu_0 + \cdots + \mu_n = l\}.$$

As in standard elimination theory we can choose a monomial order $\leq_l$ in $\mathbb{N}^{n+1}_l$ (see [CLO97, BW93, GPW03] and references therein for an introduction to monomial orders, Gröbner bases and Computational Commutative Algebra). Any monomial order in $\mathbb{N}^{n+1}_l$ allows us to see the elements of H_l as vectors given by their coordinates (with respect to this monomial order). This is called in standard literature "dense encoding of polynomials". Let N_l be the complex dimension of H_l. For every $\mu \in \mathbb{N}^{n+1}_l$, we define the multinomial coefficient

$$\binom{l}{\mu} := \frac{l!}{\mu_0! \cdots \mu_n!}.$$

We define the matrix $\Delta_l \in \mathcal{M}_{N_l}(\mathbb{C})$ associated with H_l as the diagonal matrix whose μ-th entry (with respect to the monomial order $\leq_l$) at the diagonal is $\binom{l}{\mu}^{-1/2}$. Namely, Δ_l is the diagonal matrix given by the following identity:

$$\Delta_l := \oplus_{\mu \in \mathbb{N}^{n+1}_l} \left(\binom{l}{\mu}^{-1/2} \right).$$

Let $\langle \cdot, \cdot \rangle_l : H_l \times H_l \longrightarrow \mathbb{C}$ be the canonical Hermitian product on H_l.

Let $(d) := (d_1, \ldots, d_n) \in \mathbb{N}^n$ be a list of positive degrees. We also have the canonical Hermitian product on $\mathcal{H}_{(d)}$ given by the following identity:

$$\langle f, g \rangle := \sum_{i=1}^{n} \langle f_i, g_i \rangle_{d_i} \in \mathbb{C},$$

where $f := [f_1, \ldots, f_n], g := [g_1, \ldots, g_n] \in \mathcal{H}_{(d)}$. We finally denote by $\langle \cdot, \cdot \rangle_\Delta$ the Hermitian product over $\mathcal{H}_{(d)}$ defined by the respective matrices Δ_{d_i} and given by the following identity:

$$\langle f, g \rangle_\Delta := \sum_{i=1}^{n} \langle \Delta_{d_i} f_i, \Delta_{d_i} g_i \rangle_{d_i} = \sum_{i=1}^{n} \langle f_i, g_i \rangle_{\Delta_{d_i}},$$

where $f := [f_1, \ldots, f_n], g := [g_1, \ldots, g_n] \in \mathcal{H}_{(d)}$. We denote by Δ the following matrix,

$$\Delta := \bigoplus_{i=1}^{n} \Delta_{d_i} \in \mathcal{M}_{N+1}(\mathbb{C}).$$

In order to simplify the notation, we will denote respectively by $\mathbb{S}$ and $\mathbb{S}_\Delta$ the spheres $S^1(\mathcal{H}_{(d)})$ and $S^1_\Delta(\mathcal{H}_{(d)})$. The volume element in $\mathbb{S}$ will be denoted by $d\nu$.

2.2 Incidence Varieties.

We follow the notation used in the introduction. For every system $f \in \mathcal{H}_{(d)}$, we also denote by f the mapping between complex affine spaces $f : \mathbb{C}^{n+1} \longrightarrow \mathbb{C}^n$.

Let $e_0 := (1 : 0 : \ldots : 0) \in \mathbb{P}_n(\mathbb{C})$ be a point that we may fix as a "north pole". Let $f \in \mathcal{H}_{(d)}$ be a system of homogeneous polynomial equations, and let $\zeta \in V(f)$ be any solution. We denote by $T_\zeta f$ the matrix (in some orthonormal basis) of the restriction of the tangent mapping $d_\zeta f$ to the tangent subspace $T_\zeta \mathbb{P}_n(\mathbb{C}) = \zeta^\perp \subseteq \mathbb{C}^{n+1}$ of all elements of $\mathbb{C}^{n+1}$ which are orthogonal to the complex line $\zeta \in \mathbb{P}_n(\mathbb{C})$. In the case that $\zeta = e_0$ we may identify $T_{e_0} f$ and its matrix in the natural basis $\{e_1, \ldots, e_n\}$. Namely,

$$T_{e_0} f \equiv \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \in \mathcal{M}_n(\mathbb{C}).$$

For every $\zeta \in \mathbb{P}_n(\mathbb{C})$, we shall denote by $\widetilde{V}_\zeta \subseteq \mathcal{H}_{(d)}$ the vector subspace given as the set of systems of homogeneous equations satisfied by ζ. That is,

$$\widetilde{V}_\zeta := \{f \in \mathcal{H}_{(d)} \ : \ f(\zeta) = 0 \in \mathbb{C}^n\}.$$

Note that $\widetilde{V}_\zeta$ is a complex vector subspace of $\mathcal{H}_{(d)}$ of complex codimension n.

We define the incidence variety $V \subseteq \mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C})$ given by the following identity:

$$V := \{(f,\zeta) \in \mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C}) \ : \ \zeta \in V(f)\}.$$

We also consider the two following canonical projections:

$$p_1 : V \longrightarrow \mathbb{S}_\Delta, \qquad p_1(f,\zeta) := f, \forall (f,\zeta) \in V,$$

and

$$p_2 : V \longrightarrow \mathbb{P}_n(\mathbb{C}), \qquad p_2(f,\zeta) := \zeta, \forall (f,\zeta) \in V.$$

We may obviously identify $p_1^{-1}(f) \equiv V(f)$ and $p_2^{-1}(\zeta) \equiv \widetilde{V}_\zeta \cap \mathbb{S}_\Delta = S^1_\Delta(\widetilde{V}_\zeta)$. From now on, we shall denote $V_\zeta := p_2^{-1}(\zeta)$.

The following statement summarizes the basic properties of V, and its proof may be found in [BCSS98].

Proposition 9 *The incidence variety V is a connected submanifold of the product manifold $\mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C})$ of real codimension $2n$. Moreover, the fibers V_ζ are submanifolds of V of real codimension $2n$ in V.*

We shall denote by $\Sigma' \subseteq V$ the critical locus of p_1, i.e. $\Sigma' := \{(f,\zeta) \in V \ : \ T_\zeta f \notin GL(n,\mathbb{C})\}$ (cf. [BCSS98] for details). We also denote by $\Sigma := p_1(\Sigma')$ the critical values of p_1 (also called the discriminant variety). As observed in [SS94], the following proposition follows from the implicit function theorem.

Proposition 10 (Shub & Smale) *Let $g \in \mathbb{S}_\Delta$ be a point, and let $L_\Delta \subseteq \mathbb{S}_\Delta$ be a (real) great circle in $\mathbb{S}_\Delta$ such that $g \in L_\Delta$. Assume that $L_\Delta \cap \Sigma = \emptyset$. Then, $p_1 : p_1^{-1}(L_\Delta) \longrightarrow L_\Delta$ is a $\mathcal{D}$-fold covering map, and $p_1^{-1}(L_\Delta) \setminus p_1^{-1}(\{-g\})$ consists of $\mathcal{D}$ open arcs in V. Let $\zeta \in V(g)$ be a solution of g. We denote by $ARC_{g,\zeta}$ the (unique) open arc of $p_1^{-1}(L_\Delta) \setminus p_1^{-1}(\{-g\})$ that contains the point (g,ζ).*

Note that the vector space L_{e_0} defined at the Introduction is precisely the set of $\ell \in \widetilde{V}_{e_0}$ that satisfy $T_{e_0}\ell \equiv \ell \mid_{e_0^\perp}$ (as linear operators). Then, $L_{e_0}^\perp$ is the subspace of those $f \in \widetilde{V}_{e_0}$ such that $T_{e_0} f \equiv 0$. Namely, it is the family of all homogeneous systems of polynomial equations of order at least 2 at e_0.

Let us denote by $\Delta(d)^{-1/2} \in \mathcal{M}_n(\mathbb{C})$ the diagonal complex matrix given by

$$\Delta(d)^{-1/2} := \begin{pmatrix} d_1^{-1/2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n^{-1/2} \end{pmatrix}.$$

We finally define the mapping

$$\psi_{e_0} : L_{e_0} \longrightarrow \mathcal{M}_n(\mathbb{C})$$

$$\psi_{e_0}(\ell) := \Delta(d)^{-1/2} T_{e_0}\ell,$$

As in [BCSS98, page 235], the following simple fact holds.

Proposition 11 *The mapping ψ_{e_0} defines an isometry between L_{e_0} with the Hermitian product induced by the Hermitian product $\langle \cdot, \cdot \rangle_\Delta$ on $\mathcal{H}_{(d)}$ and $\mathcal{M}_n(\mathbb{C})$ with its canonical (Frobenius) Hermitian product.*

Obviously, ψ_{e_0} also defines an isometry between the spheres $S^t_\Delta(L_{e_0})$ and $S^t(\mathcal{M}_n(\mathbb{C}))$, identifying their respective Riemannian structures.

2.3 Some unitary actions.

Let $\mathcal{U}_{n+1} \subseteq \mathcal{M}_{n+1}(\mathbb{C})$ be the group of unitary matrices. Every $U \in \mathcal{U}_{n+1}$ defines an isometry on the complex projective space: $U : \mathbb{P}_n(\mathbb{C}) \longrightarrow \mathbb{P}_n(\mathbb{C})$. The group $\mathcal{U}_{n+1}$ also defines an action on $\mathcal{H}_{(d)}$ for every $U \in \mathcal{U}_{n+1}$ as follows,

$$f \longrightarrow f \circ U^{-1}.$$

The following statement was proved in [BCSS98].

Proposition 12 *With the notation above, the Hermitian product $\langle \cdot, \cdot \rangle_\Delta$ is invariant under the action of $\mathcal{U}_{n+1}$ over $\mathcal{H}_{(d)}$. Namely, for all $f, g \in \mathcal{H}_{(d)}$ and for all $U \in \mathcal{U}_{n+1}$, the following equality holds:*

$$\langle f, g \rangle_\Delta = \langle f \circ U^{-1}, g \circ U^{-1} \rangle_\Delta.$$

The manifold V is also invariant under the action of $\mathcal{U}_{n+1}$ on the product $\mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C})$. Moreover, every $U \in \mathcal{U}_{n+1}$ defines isometries between the fibers of p_2. In fact, given $\zeta, \zeta' \in \mathbb{P}_n(\mathbb{C})$ two projective points, and given $U \in \mathcal{U}_{n+1}$, such that $U\zeta = \zeta'$, then the mapping

$$\begin{array}{cccc} U^{-1}: & \widetilde{V}_\zeta & \longrightarrow & \widetilde{V}_{\zeta'} \\ & f & \longmapsto & f \circ U^{-1} \end{array}$$

is an isometry. Obviously, the restriction

$$U^{-1} : V_\zeta = S^1_\Delta(\widetilde{V}_\zeta) \longrightarrow V_{\zeta'} = S^1_\Delta(\widetilde{V}_{\zeta'})$$

is also an isometry between spheres. Moreover, the following mapping is also an isometry for any $U \in \mathcal{U}_{n+1}$:

$$\begin{array}{cccc} U: & V & \longrightarrow & V \\ & (f, y) & \mapsto & (f \circ U^{-1}, Uy). \end{array}$$

Observe that the following diagrams commute:

$$\begin{tikzcd} V \arrow[r, "U"] \arrow[d, "p_1"'] & V \arrow[d, "p_1"] \\ \mathbb{S}_\Delta \arrow[r, "U^{-1}"] & \mathbb{S}_\Delta \end{tikzcd} \qquad \begin{tikzcd} V \arrow[r, "U"] \arrow[d, "p_2"'] & V \arrow[d, "p_2"] \\ \mathbb{P}_n(\mathbb{C}) \arrow[r, "U"] & \mathbb{P}_n(\mathbb{C}) \end{tikzcd}$$

Let $f \in V_{e_0}$ be any system. We consider the following number:

$$DET(f, e_0) := det(T_{e_0} f (T_{e_0} f)^*),$$

where the symbol * denotes Hermitian transpose.

The following statement is consequence of the results in [BCSS98].

Proposition 13 *With the notation above, let $(f, \zeta) \in V$ be a regular point of p_1. Then, the following equality holds:*

$$\frac{NJ_{(f,\zeta)}p_1}{NJ_{(f,\zeta)}p_2} = \frac{NJ_{(f\circ U,e_0)}p_1}{NJ_{(f\circ U,e_0)}p_2} = DET(f \circ U, e_0),$$

where $U \in \mathcal{U}_{n+1}$ is any matrix such that $Ue_0 = \zeta$ and $NJ_{(f,\zeta)}p_1$ and $NJ_{(f,\zeta)}p_2$ are respectively the normal jacobians at $(f, \zeta) \in V$ of p_1 and p_2 (as defined for example in [BCSS98, 13.2]).

Proof.– In [BCSS98] this result is proven when considering the projective space $\mathbb{P}(\mathcal{H}_{(d)})$ instead of $\mathbb{S}_\Delta$. Now, we can check that this change does not affect the calculus. The first equality is proved the same way as in [BCSS98]. As for the second one, observe that for any element $(f, e_0) \in V$, the tangent spaces $T_f\mathbb{S}_\Delta$ and $T_f\mathbb{P}(\mathcal{H}_{(d)})$ differ only in the vector $\sqrt{-1}f \in T_f\mathbb{S}_\Delta$. Observe that the vector $(\sqrt{-1}f, 0) \in T_{(f,e_0)}V$ satisfies:

- $\sqrt{-1}f = d_{(f,e_0)}p_1(\sqrt{-1}f, 0)$ is orthonormal to $g = d_{(f,e_0)}p_1(g, x)$ for every $(g, x) \in T_{(f,e_0)}V$ such that $\langle (g, x), (\sqrt{-1}f, 0)\rangle_{T_{(f,e_0)}V} = 0$.
- $(\sqrt{-1}f, 0) \in Ker(d_{(f,e_0)}p_2)$.

Thus, the volume of the images under $d_{(f,e_0)}p_1$ or $d_{(f,e_0)}p_2$ of a unit cube contained in the orthogonal complement of the respective kernel does not vary, and both normal jacobians $NJ_{(f,\zeta)}p_1$ and $NJ_{(f,\zeta)}p_2$ remain the same when considering $\mathbb{S}_\Delta$ or $\mathbb{P}(\mathcal{H}_{(d)})$. ■

2.4 Normalized Condition Numbers.

For every $(f, \zeta) \in V$ we shall denote by $\mu_{\text{norm}}(f, \zeta)$ the normalized condition number introduced in [SS93a] (cf. also [SS93b] or [BCSS98]). Namely,

$$\mu_{\text{norm}}(f, \zeta) := \|(T_\zeta f)^{-1}\Delta(d)^{1/2}\|_2,$$

where the representatives ζ and f are chosen such a way that $\|\zeta\|_2 = \|f\|_\Delta = 1$.

Condition numbers in Linear Algebra were introduced by A. Turing in [Tur48]. They were also studied by J. von Neumann and collaborators (cf.

[NG47]) and by J.H. Wilkinson (cf. also [Wil65]). Variations of these condition numbers may be found in the literature of Numerical Linear Algebra (cf. [Dem88], [GVL96], [Hig02], [TB97] and references therein).

The Condition Number κ_D of Linear Algebra is defined as follows: For any square matrix $A \in \mathcal{M}_k(\mathbb{C})$, $\kappa_D(A) := \|A\|_F \|A^{-1}\|_2$. The following statement immediately follows from the definition of $\mu_{\rm norm}$.

Proposition 14 *With the notation above, the following equality holds for every $(f, \zeta) \in V$:*

$$\mu_{\rm norm}(f, \zeta) = \frac{\kappa_D(\Delta(d)^{-1/2}T_\zeta f)}{||\Delta(d)^{-1/2}T_\zeta f||_F},$$

where the representatives ζ and f are chosen such a way that $\|\zeta\|_2 = \|f\|_\Delta = 1$.

Moreover, the normalized condition number $\mu_{\rm norm}$ is invariant under the action of the unitary group $\mathcal{U}_{n+1}$. Namely, given $(f, \zeta) \in V$ and given $U \in \mathcal{U}_{n+1}$, the following equality holds:

$$\mu_{\rm norm}(f, \zeta) = \mu_{\rm norm}(f \circ U^{-1}, U\zeta).$$

For every positive real number $\varepsilon > 0$, we also introduce the "tube" $\Sigma'_\varepsilon \subseteq V$ given by the following identity:

$$\Sigma'_\varepsilon := \{(f, \zeta) \in V \ : \ \mu_{\rm norm}(f, \zeta) > \varepsilon^{-1}\}.$$

Note that Σ'_ε is invariant under the action of $\mathcal{U}_{n+1}$. We recall the notations of Section 2.2. Let $g \in \mathbb{S}_\Delta$ be a point. For every great circle L_Δ containing g, $L_\Delta \cap \Sigma = \emptyset$, and for every positive number $\varepsilon > 0$, we denote by $\tau_\varepsilon^g(L_\Delta)$ the number of arcs of $p_1^{-1}(L_\Delta) \setminus p_1^{-1}(\{-g\})$ that intersect the set Σ'_ε. In other words,

$$\tau_\varepsilon^g(L_\Delta) := \sharp\{\zeta \in V(g) : ARC_{g,\zeta} \cap \Sigma'_\varepsilon \neq \emptyset\}.$$

This definition makes sense because of Proposition 10. Then, for every positive real number $\varepsilon > 0$, and for every great circle $L_\Delta \subseteq \mathbb{S}_\Delta$ such that $L_\Delta \cap \Sigma = \emptyset$, we define

$$\tau_\varepsilon(L_\Delta) := \sup_{g \in L} \tau_\varepsilon^g(L_\Delta).$$

2.5 A volume estimate for great circles.

Let $\mathbb{S} \times \mathbb{S}$ and $\mathbb{S}_\Delta \times \mathbb{S}_\Delta$ be these Riemannian manifolds, with the product Riemannian structure. For respective measurable subsets $A_1 \subseteq \mathbb{S} \times \mathbb{S}$, $A_2 \subseteq \mathbb{S}_\Delta \times \mathbb{S}_\Delta$, we denote their respective volumes as $\nu[A_1], \nu_\Delta[A_2]$.

Let $\Delta^{-1} \in \mathcal{M}_{N+1}(\mathbb{C})$ be the inverse of the nonsingular matrix Δ. Observe that both

$$\Delta^{-1} : \mathbb{S} \longrightarrow \mathbb{S}_\Delta.$$

and

$$\Delta^{-1} \times \Delta^{-1} : \mathbb{S} \times \mathbb{S} \longrightarrow \mathbb{S}_\Delta \times \mathbb{S}_\Delta.$$

are isometries. Let $\mathcal{L}$ be the Riemannian manifold of great circles (real spherical lines) in $\mathbb{S}$, endowed with the natural orthogonal-invariant Riemannian structure. We denote by $d\mathcal{L}$ the volume form associated with this Riemannian structure. For every measurable subset $A \subseteq \mathcal{L}$, let $\nu_{\mathcal{L}}[A]$ be the volume of A with respect to $d\mathcal{L}$. We may assume that this volume form has been normalized such a way that $\nu_{\mathcal{L}}[\mathcal{L}] = 1$. We recall some basic properties of the Riemannian structures we have introduced. Let $\mathcal{O}_{2N+2}$ be the group of orthogonal square matrices of size $2N + 2$, which acts isometrically over $\mathbb{C}^{N+1} \equiv \mathbb{R}^{2N+2}$. Namely, for every measurable set $A \subseteq \mathbb{S}$, the following holds:

$$\nu[A] = \nu[OA], \qquad \forall O \in \mathcal{O}_{2N+2}.$$

The following mapping is an isometry for every orthogonal matrix $O \in \mathcal{O}_{2N+2}$:

$$\begin{array}{rccl} O : & \mathcal{L} & \longrightarrow & \mathcal{L} \\ & L & \mapsto & OL \ := \{Of \in \mathbb{S} : f \in L\}. \end{array}$$

For every element $L \in \mathcal{L}$, we may consider the great circle $L_\Delta \subseteq \mathbb{S}_\Delta$ defined as $L_\Delta := \Delta^{-1}L = \{\Delta^{-1}f : f \in L\}$. In Subsection 2.4, for every positive number $\varepsilon > 0$ and every great circle $L_\Delta \subseteq \mathbb{S}_\Delta$ not intersecting Σ, we have defined the quantity $\tau_\varepsilon(L_\Delta)$. Thus, for every element $L \in \mathcal{L}$ such that $L_\Delta \cap \Sigma = \emptyset$ we can consider the number $\tau(\varepsilon, L)$ defined as follows.

$$\tau(\varepsilon, L) := \tau_\varepsilon(L_\Delta),$$

For every element $f \in \mathbb{S}_\Delta \setminus \Sigma$, we may consider the positive integer number $\sharp(\varepsilon, f) \in \mathbb{N}$ defined as follows:

$$\sharp(\varepsilon, f) := \sharp\{\zeta \in V(f) : \mu_{norm}(f, \zeta) > \varepsilon^{-1}\}.$$

We also denote by $\mathcal{L}_\Delta$ the set of all the (real) great circles in $\mathbb{S}_\Delta$. We recall a result of M. Shub and S. Smale, which can be found in [SS93b].

Theorem 15 (Shub-Smale) *For every positive number $\varepsilon > 0$, the following inequality holds:*

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f \in \mathbb{S}_\Delta} \sharp(\varepsilon, f) \; d\mathbb{S}_\Delta \leq n^3(n+1)N^2\mathcal{D}\varepsilon^4,$$

where $\mathcal{D} := \prod_{i=1}^n d_i$ is the Bézout number.

The following result is also due to Shub and Smale, as it can be obtained as an immediate consequence of the corollary of Theorem 1 in [SS96].

Lemma 16 (Shub-Smale) *For every positive real number $\varepsilon > 0$ and for every $L \in \mathcal{L}$, the following inequality holds:*

$$\tau(\varepsilon, L) \leq \left(\frac{c\varepsilon^2}{d^{3/2}}\right)^{-1} \int_{f \in L_\Delta} \sharp(2\varepsilon, f) \; dL_\Delta,$$

where $c \geq 0.09$ is a universal constant.

Proof.– From the definition, $\tau(\varepsilon, L) = \tau_\varepsilon(L_\Delta) = \sup_{g \in L_\Delta} \tau_\varepsilon^g(L_\Delta)$. Now, from [SS94, Proof of Cor. 3.5] the quantity $\tau_\varepsilon^g(L_\Delta)$ is bounded for every $g \in L$ by

$$\left(\frac{c\varepsilon^2}{d^{3/2}}\right)^{-1} \int_{f \in L_\Delta} \sharp(2\varepsilon, f) \; dL_\Delta,$$

and the lemma follows. ■

The following result is implicitly stated in [SS94]. We include a short proof for completeness.

Proposition 17 *Let $\varepsilon > 0$ be a positive real number. The following inequality holds.*

$$\int_{L \in \mathcal{L}} \tau(\varepsilon, L) \; d\mathcal{L} \leq \frac{32\pi}{c} d^{3/2} n^3 (n+1) N^2 \mathcal{D} \varepsilon^2,$$

where $c > 0$ is the universal constant of Lemma 16.

Proof.– From Lemma 16,

$$\int_{L \in \mathcal{L}} \tau(\varepsilon, L) \; d\mathcal{L} \leq \left(\frac{c\varepsilon^2}{d^{3/2}}\right)^{-1} \int_{L \in \mathcal{L}} \int_{f \in L_\Delta} \sharp(2\varepsilon, f) \; dL_\Delta \; d\mathcal{L} =$$

$$= \left(\frac{c\varepsilon^2}{d^{3/2}}\right)^{-1} \int_{L \in \mathcal{L}} \int_{f \in L} \sharp(2\varepsilon, \Delta^{-1} f) \; dL \; d\mathcal{L}.$$

The following Santalo-type equality follows from Shub & Smale's arguments in [SS96, Prop. 4b] (cf. [How93, San76] for other similar Integral Geometry formulae):

$$\int_{L \in \mathcal{L}} \int_{f \in L} \sharp(2\varepsilon, \Delta^{-1} f) \; dL \; d\mathcal{L} = 2\pi \frac{\int_{f \in \mathbb{S}} \sharp(2\varepsilon, \Delta^{-1} f) \; d\mathbb{S}}{\nu[\mathbb{S}]}.$$

As Δ^{-1} is an isometry from $\mathbb{S}$ to $\mathbb{S}_\Delta$,

$$\frac{1}{\nu[\mathbb{S}]} \int_{f \in \mathbb{S}} \sharp(2\varepsilon, \Delta^{-1} f) \; d\mathbb{S} = \frac{1}{\nu[\mathbb{S}_\Delta]} \int_{f \in \mathbb{S}_\Delta} \sharp(2\varepsilon, f) \; d\mathbb{S}_\Delta,$$

and Theorem 15 yields:

$$\frac{1}{\nu[\mathbb{S}_\Delta]}\int_{f\in\mathbb{S}_\Delta}\sharp(2\varepsilon,f)\ d\mathbb{S}_\Delta\leq 16n^3(n+1)N^2\mathcal{D}\varepsilon^4.$$

Thus,

$$\int_{L\in\mathcal{L}}\tau(\varepsilon,L)\ d\mathcal{L}\leq\left(\frac{c\varepsilon^2}{d^{3/2}}\right)^{-1}32\pi n^3(n+1)N^2\mathcal{D}\varepsilon^4,$$

and the proposition follows.

■

3 The Homotopy Method.

There is a wide bibliography on Newton-like methods for solving systems of polynomial equations. Some good references are [BCSS98, Ded06, DS00, Mal94]. In [Shu93], the projective Newton operator is introduced, and the series of papers [SS93a, SS93b, SS93c, SS94, SS96] propose a linear homotopy method. We recall now the key ingredients of this method. Most of them are summarized in [BCSS98].

Let $d_T:\mathbb{P}_n(\mathbb{C})\times\mathbb{P}_n(\mathbb{C})\longrightarrow\mathbb{R}$ be the function given by the following equality,

$$d_T(z_1,z_2):=\tan(d_R(z_1,z_2)).$$

Namely, d_T is the tangent of the Riemannian distance. Observe that d_T is not exactly a distance function, but $d_T(z_1,z_2)$ is very similar to $d_R(z_1,z_2)$ for small values of $d_R(z_1,z_2)$. Let $\zeta\in\mathbb{P}_n(\mathbb{C})$ be a zero of $f\in\mathbb{S}_\Delta$.

Definition 18 *We say that $z\in\mathbb{P}_n(\mathbb{C})$ is an approximate zero of f with associated zero ζ if the sequence*

$$z_0:=z,\qquad z_{i+1}:=N_f(z_i)\ \forall i\geq 0$$

is defined, and

$$d_T(\zeta,z_i)\leq\left(\frac{1}{2}\right)^{2^i-1}d_T(\zeta,z),\forall i\geq 0.$$

The following result guarantees the convergence of the Newton sequence under some assumptions:

Theorem 19 (Shub & Smale) *Let $f\in\mathbb{S}_\Delta$, and let $\zeta\in\mathbb{P}_n(\mathbb{C})$ be a zero of f. Let $\gamma_0(f,\zeta)$ be the number defined as follows.*

$$\gamma_0(f,\zeta):=\|\zeta\|\max_{k\geq 1}\left\|(T_\zeta f)^{-1}\frac{D^kf(\zeta)}{k!}\right\|^{\frac{1}{k-1}},$$

where $D^k f(\zeta)$ is the k-th derivative of f, considered as a k-linear map. Let $z \in \mathbb{P}_n(\mathbb{C})$ be such that

$$d_T(z,\zeta)\gamma_0(f,\zeta) \leq \frac{3-\sqrt{7}}{2}.$$

Then, z is an approximate zero of f with associated zero ζ.

3.1 The linear homotopy.

Observe that Theorem 19 does not solve the problem of finding a zero of a given system $f \in \mathbb{S}_\Delta$. In fact, in general it may be hard to find an initial point $z \in \mathbb{P}_n(\mathbb{C})$ satisfying the conditions of Theorem 19. The linear homotopy proposed by Shub and Smale solves this problem considering another system $g \in \mathbb{S}_\Delta$, which has a known zero ζ_0. Then, we consider the segment

$$\Gamma := \{f_t := (1-t)g + tf,\ t \in [0,1]\} \subseteq \mathcal{H}_{(d)}.$$

If $\Gamma \cap \Sigma = \emptyset$, the implicit function theorem defines a path of solutions

$$C(\Gamma) := \{(f_t,\zeta_t) : \zeta_t \in V(f_t),\ t \in [0,1]\}$$

Observe that ζ_1 is a zero of $f_1 = f$. Let $k \geq 1$ be a positive integer, representing the number of homotopy steps to be done. Let $t_i = \frac{i}{k}$, $0 \leq i \leq k$, and consider the following sequence of systems:

$$f_{t_i} = \left(1 - \frac{i}{k}\right) g + \frac{i}{k} f, \qquad 0 \leq i \leq k.$$

Observe that $f_{t_0} = g$, $f_{t_k} = f$. Then, we may consider the sequence of points defined as follows:

$$x_0 := \zeta_0, \qquad x_{i+1} := N_{f_{t_{i+1}}}(x_i), 0 \leq i \leq k-1.$$

The following is the main result of [SS93a] (see also [BCSS98, pg. 271] or [Bel06, Prop. 4.2.6]). It bounds the number of steps k that are necessary to guarantee convergence.

Theorem 20 (Shub & Smale) *With the notations and assumptions above, let $\mu \in \mathbb{R}$ be the number defined as follows:*

$$\mu := \max_{0 \leq t \leq 1} \mu_{norm}(f_t,\zeta_t).$$

Let $k \in \mathbb{N}$ be such that $k \geq 18 d^{3/2}\mu^2$. Then, for every $0 \leq i \leq k$, x_i is an approximate zero of f_{t_i}, with associated zero $\zeta_{i/k}$. In particular, x_k is an approximate zero of f with associated zero ζ_1.

In [SS94], a more intelligent method to construct the homotopy path between two points is proposed. Practical implementations should follow this scheme, instead of the "fixed step size" scheme we propose here. However, the theoretical results we prove are valid for both schemes. Observe that, as shown in the Introduction, the key ingredient for this method is the initial pair (g, ζ_0), satisfying the condition that μ is small for a wide set of input polynomials f.

4 A Series of Reductions.

In this section we will perform a series of geometric reductions from Shub & Smale's statements above. The final expression will be used in the coming sections to prove the main theorems in the Introduction. Every subsection contains one of these reductions.

4.1 From great circles to pairs of systems of equations.

Let $\mathfrak{D} \subseteq \mathbb{S} \times \mathbb{S}$ be the antipodal diagonal in this product space. Namely,

$$\mathfrak{D} := \{(f, g) \in \mathbb{S} \times \mathbb{S} \ : \ f = \pm g\}.$$

We define the mapping

$$\mathfrak{L} : \mathbb{S} \times \mathbb{S} \setminus \mathfrak{D} \longrightarrow \mathcal{L},$$

such that for every $(f, g) \in \mathbb{S} \times \mathbb{S} \setminus \mathfrak{D}$, the line $\mathfrak{L}(f, g) \in \mathcal{L}$ is the unique great circle in $\mathbb{S}$ that contains f and g. We also consider the set

$$\mathfrak{D}_\Delta := \{(f, g) \in \mathbb{S}_\Delta \times \mathbb{S}_\Delta \ : \ f = \pm g\},$$

and the mapping

$$\mathfrak{L}_\Delta : \mathbb{S}_\Delta \times \mathbb{S}_\Delta \setminus \mathfrak{D}_\Delta \longrightarrow \mathcal{L}_\Delta,$$

such that for every $(f, g) \in \mathbb{S}_\Delta \times \mathbb{S}_\Delta \setminus \mathfrak{D}_\Delta$, the line $\mathfrak{L}_\Delta(f, g)$ is the unique great circle in $\mathbb{S}_\Delta$ that contains f and g.

Lemma 21 *Let $F : M \longrightarrow N$ be a map between complex or real Riemannian manifolds M, N. Let $x, y \in M$ be two points in M. Assume that there exist isometries $h : M \longrightarrow M$ and $h_1 : N \longrightarrow N$ such that $h(x) = y$, and the following formula holds:*

$$h_1 \circ F = F \circ h.$$

Then, $NJ_xF = NJ_yF$.

Proof.– As h and h_1 are isometries,

$$(NJ_{F(x)}h_1)(NJ_xF) = NJ_x(h_1 \circ F) = NJ_x(F \circ h) =$$
$$= (NJ_{h(x)}F)(NJ_xh) = (NJ_yF)(NJ_xh).$$

Now, $NJ_{F(x)}h_1 = NJ_xh = 1$ and the lemma follows. $\blacksquare$

We prove the following lemma:

Lemma 22 *Let $\Phi : \mathcal{L} \longrightarrow \mathbb{R}$ be an integrable mapping. Then, the following formula holds:*

$$\frac{\int_{(f,g)\in\mathbb{S}\times\mathbb{S}} \Phi(\mathfrak{L}(f,g))\ d(\mathbb{S}\times\mathbb{S})}{\nu[\mathbb{S}]^2} = \int_{L\in\mathcal{L}} \Phi(L)\ d\mathcal{L}.$$

Proof.– The Coarea formula (see [Fed69, Mor88] or more recently [BCSS98]) applied to the differentiable mapping $\mathfrak{L} : \mathbb{S}\times\mathbb{S}\setminus\mathfrak{D} \longrightarrow \mathcal{L}$, yields:

$$\int_{(f,g)\in\mathbb{S}\times\mathbb{S}} \Phi(\mathfrak{L}(f,g))\ d(\mathbb{S}\times\mathbb{S}) =$$
$$\int_{L\in\mathcal{L}} \Phi(L) \int_{(f,g)\in\mathfrak{L}^{-1}(L)} \frac{1}{NJ_{(f,g)}\mathfrak{L}}\ d\mathfrak{L}^{-1}(L)\ d\mathcal{L}. \tag{2}$$

We check that the inner integral is a constant. In fact, let $L_1, L_2 \in \mathcal{L}$ be two great circles, and let $O \in \mathcal{O}_{2N+2}$ be an orthogonal matrix such that $OL_1 = L_2$. Consider the following isometry:

$$\begin{array}{rccc} O\times O: & \mathbb{S}\times\mathbb{S}\setminus\mathfrak{D} & \longrightarrow & \mathbb{S}\times\mathbb{S}\setminus\mathfrak{D} \\ & (f,g) & \mapsto & (Of,Og). \end{array}$$

Then, $(O\times O)_{|\mathfrak{L}^{-1}(L_1)}$ is an isometry between $\mathfrak{L}^{-1}(L_1)$ and $\mathfrak{L}^{-1}(L_2)$. The Coarea Formula applied to this map yields:

$$\int_{(f_1,g_1)\in\mathfrak{L}^{-1}(L_1)} \frac{1}{NJ_{(f_1,g_1)}\mathfrak{L}}\ d\mathfrak{L}^{-1}(L_1) =$$
$$= \int_{(f_2,g_2)\in\mathfrak{L}^{-1}(L_2)} \frac{1}{NJ_{(O^{-1}f_2,O^{-1}g_2)}\mathfrak{L}}\ d\mathfrak{L}^{-1}(L_2)$$

Now, let $(f_2, g_2) \in \mathfrak{L}^{-1}(L_2)$ be any point. Let $f_2' = O^{-1}f_2$, $g_2' = O^{-1}g_2$ be their respective pre-images by O. Observe that:

$$O\circ\mathfrak{L} = \mathfrak{L}\circ(O\times O), \qquad (O\times O)(f_2', g_2') = (f_2, g_2).$$

Thus, from Lemma 21 the following equality holds:

$$NJ_{(f_2,g_2)}\mathfrak{L} = NJ_{(f_2',g_2')}\mathfrak{L} = NJ_{(O^{-1}f_2,O^{-1}g_2)}\mathfrak{L},$$

and we deduce that the inner integral in equation (2) is a constant. Applying the same equation (2) to the map $\Phi \equiv 1$, we deduce that the value of this constant is $\nu[\mathbb{S}]^2$, and the lemma follows.

$\blacksquare$

Proposition 23 *Let $\Phi : \mathcal{L} \longrightarrow \mathbb{R}$ be an integrable mapping. Then, the following formula holds:*

$$\frac{\int_{(f,g)\in\mathbb{S}_\Delta\times\mathbb{S}_\Delta} \Phi(\mathfrak{L}(\Delta f, \Delta g))\ d(\mathbb{S}_\Delta \times \mathbb{S}_\Delta)}{\nu[\mathbb{S}_\Delta]^2} = \int_{L\in\mathcal{L}} \Phi(L)\ d\mathcal{L}.$$

Proof.– The result is an immediate consequence of Lemma 22, as $\Delta^{-1}\times\Delta^{-1}$ defines an isometry between $\mathbb{S}\times\mathbb{S}$ and $\mathbb{S}_\Delta\times\mathbb{S}_\Delta$. ■

Proposition 24 *With the notation above, the following holds:*

$$\int_{\mathbb{S}_\Delta\times\mathbb{S}_\Delta} \tau_\varepsilon(\mathfrak{L}_\Delta(f,g))d(\mathbb{S}_\Delta\times\mathbb{S}_\Delta) \leq \nu_\Delta[\mathbb{S}_\Delta]^2\frac{32\pi}{c}\varepsilon^2 n^3(n+1)N^2\mathcal{D}d^{3/2},$$

where τ_ε is the mapping introduced at Subsection 2.4 above.

Proof.– Observe that $\tau_\varepsilon(\mathfrak{L}_\Delta(f,g)) = \tau(\varepsilon, \mathfrak{L}(\Delta f, \Delta g))$, as defined in Subsection 2.5. From Proposition 23, the following equality holds:

$$\int_{(f,g)\in\mathbb{S}_\Delta\times\mathbb{S}_\Delta} \tau(\varepsilon, \mathfrak{L}(\Delta f, \Delta g))d(\mathbb{S}_\Delta\times\mathbb{S}_\Delta) = \nu_\Delta[\mathbb{S}_\Delta]^2\int_{L\in\mathcal{L}}\tau(\varepsilon, L)d\mathcal{L}.$$

The inequality follows from Proposition 17. ■

4.2 From pairs of systems to fibers at zeros.

We consider now the product of the incidence variety V with $\mathbb{S}_\Delta$ and define the two following projections:

$$\pi_1 : \mathbb{S}_\Delta\times V \longrightarrow \mathbb{S}_\Delta\times\mathbb{S}_\Delta,$$

given by

$$\pi_1(f,g,\zeta) := (f,g),\ \ \forall f\in\mathbb{S}_\Delta,\ (g,\zeta)\in V,$$

and

$$\pi_2 : \mathbb{S}_\Delta\times V \longrightarrow \mathbb{P}_n(\mathbb{C}),$$

given by

$$\pi_2(f,g,\zeta) := \zeta,\ \ \forall f\in\mathbb{S}_\Delta,\ (g,\zeta)\in V.$$

So, we have the following fibrations:

$$\begin{tikzcd}
\mathbb{S}_\Delta\times V \arrow[r, "i"] \arrow[dr, "\pi_1"'] & \mathbb{S}_\Delta\times\mathbb{S}_\Delta\times\mathbb{P}_n(\mathbb{C}), \arrow[d, "\pi_1"] \\
& \mathbb{S}_\Delta\times\mathbb{S}_\Delta
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbb{S}_\Delta\times V \arrow[r, "i"] \arrow[dr, "\pi_2"'] & \mathbb{S}_\Delta\times\mathbb{S}_\Delta\times\mathbb{P}_n(\mathbb{C}), \arrow[d, "\pi_2"] \\
& \mathbb{P}_n(\mathbb{C})
\end{tikzcd}$$

where $i : \mathbb{S}_\Delta \times V \longrightarrow \mathbb{S}_\Delta \times \mathbb{S}_\Delta \times \mathbb{P}_n(\mathbb{C})$ is the inclusion. Note that $\pi_1 = Id \times p_1$, where Id is the identity on $\mathbb{S}_\Delta$ and p_1 is the projection introduced in Subsection 2.2. On the other hand, $\pi_2 = p_2 \circ \pi$, where p_2 is the projection introduced in 2.2 and π is the projection from $\mathbb{S}_\Delta \times V$ onto V. Hence, the following statement is an immediate consequence of Proposition 13 above.

Proposition 25 *With the notation above, the following equality holds for every $(f, g, \zeta) \in \mathbb{S}_\Delta \times V$:*

$$\frac{NJ_{(f,g,\zeta)}\pi_1}{NJ_{(f,g,\zeta)}\pi_2} = \frac{NJ_{(g,\zeta)}p_1}{NJ_{(g,\zeta)}p_2} = DET(g \circ U, e_0),$$

for any unitary matrix $U \in \mathcal{U}_{n+1}$ such that $Ue_0 = \zeta$.

The following statement follows from Proposition 25 and the Co-area Formula (cf. [Fed69, Mor88]) as used in [SS93b] and [SS96], applied to the previously described fibrations π_1, π_2.

Proposition 26 *Let $\Phi : \mathbb{S}_\Delta \times V \longrightarrow \mathbb{R}_+$ be an integrable function. Assume that Φ is invariant under the action of the unitary group $\mathcal{U}_{n+1}$ on $\mathbb{S}_\Delta \times V$. Namely, for every $(f, g, \zeta) \in \mathbb{S}_\Delta \times V$ and for every $U \in \mathcal{U}_{n+1}$ the following equality holds:*

$$\Phi(f, g, \zeta) = \Phi(f \circ U^{-1}, g \circ U^{-1}, U\zeta).$$

Let I be the quantity given by the following identity:

$$I := \int_{(f,g,\zeta)\in \mathbb{S}_\Delta \times V} \Phi(f, g, \zeta) NJ_{(f,g,\zeta)}\pi_1 d\mathbb{S}_\Delta dV.$$

Then, the two following equalities hold:

$$I = \int_{(f,g)\in \mathbb{S}_\Delta \times \mathbb{S}_\Delta} \sum_{\zeta \in V(g)} \Phi(f, g, \zeta) d(\mathbb{S}_\Delta \times \mathbb{S}_\Delta),$$

and

$$I = \nu_{\mathbb{P}}[\mathbb{P}_n(\mathbb{C})] \int_{f\in\mathbb{S}_\Delta} \int_{g\in V_{e_0}} \Phi(f, g, e_0) DET(g, e_0) dV_{e_0} d\mathbb{S}_\Delta.$$

We apply this proposition as in Section 3 of [SS94]. First of all, the following statement follows from Proposition 10.

Proposition 27 *Let $(f, g, \zeta) \in \mathbb{S}_\Delta \times V$ be a point such that the line $\mathfrak{L}_\Delta(f, g)$ does not intersect Σ. Then, there is one and only one arc of $p_1^{-1}(\mathfrak{L}_\Delta(f, g)) \setminus p_1^{-1}(\{-g\}) \subseteq V$ that contains the point (g, ζ).*

We shall denote by $\mathfrak{L}_\Delta(f, g, \zeta)$ this arc. We shall also denote:

$$\chi_{\mathcal{L}'_\varepsilon}(\mathfrak{L}_\Delta(f, g, \zeta)) := \begin{cases} 1 & \text{if } \mathfrak{L}_\Delta(f, g, \zeta) \cap \Sigma'_\varepsilon \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

Proposition 28 *With the notation above, the following holds:*

$$\nu_{\mathbb{P}}[\mathbb{P}_n(\mathbb{C})]\int_{V_{e_0}} A_\varepsilon(g,e_0)DET(g,e_0)dV_{e_0} \leq \frac{32\pi}{c}\nu_\Delta[\mathbb{S}_\Delta]\varepsilon^2 n^3(n+1)N^2\mathcal{D}d^{3/2},$$

where for every $g\in V_{e_0}$, we define

$$A_\varepsilon(g,e_0) := \frac{1}{\nu_\Delta[\mathbb{S}_\Delta]}\int_{f\in\mathbb{S}_\Delta}\chi_{\mathcal{L}'_\varepsilon}(\mathfrak{L}_\Delta(f,g,e_0))d\mathbb{S}_\Delta.$$

Proof.– First of all, observe that the following inequality holds:

$$\tau_\varepsilon(\mathfrak{L}_\Delta(f,g)) \geq \tau_\varepsilon^g(\mathfrak{L}_\Delta(f,g)) = \sum_{\zeta\in V(g)}\chi_{\mathcal{L}'_\varepsilon}(\mathfrak{L}_\Delta(f,g,\zeta)).$$

Thus, from Proposition 26, the following inequality also holds:

$$\nu_{\mathbb{P}}[\mathbb{P}_n(\mathbb{C})]\int_{V_{e_0}} A_\varepsilon(g,e_0)DET(g,e_0)dV_{e_0} \leq \int_{\mathbb{S}_\Delta\times\mathbb{S}_\Delta}\tau_\varepsilon(\mathfrak{L}_\Delta(f,g))d(\mathbb{S}_\Delta\times\mathbb{S}_\Delta).$$

The statement follows from the inequality of Proposition 24. ∎

4.3 From fibers at zeros to square matrices.

We recover the notations of Subsection 2.2. Let us assume that there exists an index $1\leq i\leq n$ such that $d_i>1$ and let us consider the orthogonal projection

$$\pi_{(d)}:\widetilde{V}_{e_0}\longrightarrow L_{e_0}.$$

This induces an orthogonal projection (that we denote by the same symbol)

$$\pi_{(d)}:V_{e_0}\longrightarrow B^1_\Delta(L_{e_0}).$$

We also consider the mapping ψ_{e_0} defined in Subsection 2.2, and the mapping

$$\Pi_{(d)} := \psi_{e_0}\circ\pi_{(d)}:V_{e_0}\longrightarrow B^1(\mathcal{M}_n(\mathbb{C})),$$

Hence, the situation now is described by the following diagram

$$\begin{tikzcd}
V_{e_0} \arrow[r, "\pi_{(d)}"] \arrow[dr, "\Pi_{(d)}"'] & B^1_\Delta(L_{e_0}) \arrow[d, "\psi_{e_0}"] \\
 & B^1(\mathcal{M}_n(\mathbb{C}))
\end{tikzcd}$$

With the notations introduced in Subsection 2.2, we have

$$\Pi_{(d)}(g) = \Delta(d)^{-1/2} T_{e_0} g,$$

whereas $T_{e_0} g = 0$ for every $g \in L_{e_0}^{\perp}$.

In particular, for every $M \in \mathcal{M}_n(\mathbb{C})$ such that $||M||_F = t \leq 1$, the fiber $\Pi_{(d)}^{-1}(M)$ can be identified with the sphere in $L_{e_0}^{\perp}$ of radius $(1-t^2)^{1/2}$. Namely, we have

$$\Pi_{(d)}^{-1}(M) = S_{\Delta}^{(1-t^2)^{1/2}}(L_{e_0}^{\perp}) := \{(1-t^2)^{1/2} h \ : \ h \in L_{e_0}^{\perp}, ||h||_{\Delta} = 1\}.$$

Moreover, observe that for every $g \in \widetilde{V}_{e_0}$, the following equality holds:

$$DET(g, e_0) := det((T_{e_0} g)^*(T_{e_0} g)) = \mathcal{D} det((\Pi_{(d)}(g))^*(\Pi_{(d)}(g))),$$

as $det(\Delta(d))^{1/2}) = \mathcal{D}^{1/2}$.

As observed in [BCSS98], the normal jacobian of $\Pi_{(d)}$ at a point $g \in V_{e_0}$ satisfies the following chain of equalities:

$$NJ_g \Pi_{(d)} = NJ_g \pi_{(d)} = (1 - ||\pi_{(d)}(g)||_{\Delta}^2)^{1/2} = (1 - ||\Pi_{(d)}(g)||_F^2)^{1/2},$$

provided that $d_i > 1$ for some $i, \ 1 \leq i \leq n$. Then, the proposition below follows from Proposition 28 above and the Coarea Formula applied to $\Pi_{(d)}$.

Proposition 29 *Assume that there exists an index $1 \leq i \leq n$ such that $d_i > 1$. With the notation above, the following inequality holds:*

$$\frac{\nu_{\mathbb{P}}[\mathbb{P}_n(\mathbb{C})] \nu_{\Delta}[S_{\Delta}^1(L_{e_0}^{\perp})]}{\nu_{\Delta}[\mathbb{S}_{\Delta}]} \int_{B^1(\mathcal{M}_n(\mathbb{C}))} det(M^*M)(1 - ||M||_F^2)^{N-n^2-n} B_{\varepsilon}(M, e_0) d\mathcal{M}_n(\mathbb{C})$$
$$\leq \frac{32\pi}{c} \varepsilon^2 n^3 (n+1) N^2 d^{3/2},$$

where $\nu_{\Delta}[S_{\Delta}^1(L_{e_0}^{\perp})]$ is the volume of $S_{\Delta}^1(L_{e_0}^{\perp})$ as a linear subspace of $\mathbb{S}_{\Delta}$, and

$$B_{\varepsilon}(M, e_0) := \frac{1}{\nu_{\Delta}[S_{\Delta}^1(L_{e_0}^{\perp})]} \int_{h \in S_{\Delta}^1(L_{e_0}^{\perp})} A_{\varepsilon}((1 - ||M||_F^2)^{1/2} h + \psi_{e_0}^{-1}(M), e_0) dS_{\Delta}^1(L_{e_0}^{\perp}).$$

Finally, using spherical coordinates we conclude:

Proposition 30 *Assume that there exists an index $1 \leq i \leq n$ such that $d_i > 1$. With the notation above, the following inequality holds:*

$$\int_0^1 (1-t^2)^{N-n^2-n} t^{2n^2+2n-1} K_{\varepsilon}(t, e_0) dt \leq$$
$$\leq \frac{\nu_{\Delta}[\mathbb{S}_{\Delta}]}{\nu_{\Delta}[S_{\Delta}^1(L_{e_0}^{\perp})] \nu[S^1(\mathcal{H}_{(1)})]} \frac{32\pi}{c} \varepsilon^2 n^3 (n+1) N^2 d^{3/2},$$

where

$$K_{\varepsilon}(t, e_0) = \frac{\nu_{\mathbb{P}}[\mathbb{P}_n(\mathbb{C})]}{\nu[S^1(\mathcal{H}_{(1)})]} \int_{S^1(\mathcal{M}_n(\mathbb{C}))} det(M^*M) B_{\varepsilon}(tM, e_0) dS^1(\mathcal{M}_n(\mathbb{C})).$$

4.4 From square matrices to underdetermined linear systems.

This subsection provides an alternative characterization of the quantity $K_\varepsilon(t, e_0)$ (Proposition 32 below). Observe that $\mathcal{H}_{(1)}$ is endowed with the usual Hermitian (Frobenius) product. Let $V_{(1)} := \{(M, x) \in S^1(\mathcal{H}_{(1)}) \times \mathbb{P}_n(\mathbb{C}) : Mx = 0\}$ be the incidence variety. For any matrix $M \in \mathcal{H}_{(1)}$ of rank equal to n, we consider the number:

$$\widetilde{B}_\varepsilon(t, M) := B_\varepsilon(tT_{e_0}(MU), e_0),$$

where $U \in \mathcal{U}_{n+1}$ is any unitary matrix such that $MUe_0 = 0$ and $T_{e_0}(MU)$ is the restriction $MU \mid_{e_0^\perp}$. In other words, $T_{e_0}(MU)$ is the square matrix consisting of the last n columns of MU. The following lemma proves that $\widetilde{B}_\varepsilon$ is well defined.

Lemma 31 *Assume that there exists an index $1 \leq i \leq n$ such that $d_i > 1$. Let $M \in \mathcal{H}_{(1)}$ be a matrix, rank$(M) = n$, and let $0 \leq t \leq 1$ be a real positive number. Let $U_1, U_2 \in \mathcal{U}_{n+1}$ be two unitary matrices such that $MU_1e_0 = MU_2e_0 = 0$. Then, the following equality holds:*

$$B_\varepsilon(tT_{e_0}(MU_1), e_0) = B_\varepsilon(tT_{e_0}(MU_2), e_0).$$

Moreover, for every unitary matrix $U \in \mathcal{U}_{n+1}$ the following equality also holds:

$$\widetilde{B}_\varepsilon(t, M) = \widetilde{B}_\varepsilon(t, MU).$$

Proof.– The second claim is an immediate consequence of the first one. As for the first claim, observe that

$$\frac{B_\varepsilon(tT_{e_0}(MU_1), e_0)}{B_\varepsilon(tT_{e_0}(MU_2), e_0)} = \frac{\int_{h \in \Pi_{(d)}^{-1}(tT_{e_0}(MU_1))} A_\varepsilon(h, e_0)\ d(\Pi_{(d)}^{-1}(tT_{e_0}(MU_1)))}{\int_{h \in \Pi_{(d)}^{-1}(tT_{e_0}(MU_2))} A_\varepsilon(h, e_0)\ d(\Pi_{(d)}^{-1}(tT_{e_0}(MU_2)))}. \quad (3)$$

Let $U := U_1^{-1}U_2 \in \mathcal{U}_{n+1}$ be the unitary matrix such that $U_1U = U_2$. Observe that $Ue_0 = e_0 \in \mathbb{P}_n(\mathbb{C})$. Observe that the following mapping is an isometry:

$$\begin{array}{ccc} \Pi_{(d)}^{-1}(tT_{e_0}(MU_1)) & \longrightarrow & \Pi_{(d)}^{-1}(tT_{e_0}(MU_2)) \\ h & \mapsto & h \circ U. \end{array}$$

Thus, from the Coarea Formula, the expression in equation (3) equals:

$$\frac{\int_{h \in \Pi_{(d)}^{-1}(tT_{e_0}(MU_2))} A_\varepsilon(h \circ U^{-1}, e_0)\ d(\Pi_{(d)}^{-1}(tT_{e_0}(MU_2)))}{\int_{h \in \Pi_{(d)}^{-1}(tT_{e_0}(MU_2))} A_\varepsilon(h, e_0)\ d(\Pi_{(d)}^{-1}(tT_{e_0}(MU_2)))}.$$

Now, $A_\varepsilon(h \circ U^{-1}, e_0) = A_\varepsilon(h \circ U^{-1}, Ue_0) = A_\varepsilon(h, e_0)$ and the lemma follows.

■

The following proposition may be proved in the same manner as Proposition 26 or following the arguments in [BCSS98].

Proposition 32 *Assume that there exists an index $1 \leq i \leq n$ such that $d_i > 1$. Let $\varepsilon > 0$ be a positive real number, and let $t > 0$ be a positive real number, $0 < t \leq 1$. Let $K_\varepsilon(t, e_0)$ be as in Proposition 30. Then, the following equality holds:*

$$K_\varepsilon(t, e_0) = \frac{1}{\nu[S^1(\mathcal{H}_{(1)})]} \int_{M \in S^1(\mathcal{H}_{(1)})} \widetilde{B}_\varepsilon(t, M) \; dS^1(\mathcal{H}_{(1)}).$$

Proof.– For any $M \in \mathcal{M}_n(\mathbb{C})$, let $(0, M) \in \mathcal{H}_{(1)}$ be the matrix obtained by adding to M a first column of zeros. We consider the following fibrations:

$$\begin{tikzcd}
V_{(1)} \arrow[r, "i"] \arrow[dr, "\phi_1"'] & \mathcal{H}_{(1)} \times \mathbb{P}_n(\mathbb{C}), \arrow[d, "\phi_1"] & V_{(1)} \arrow[r, "i"] \arrow[dr, "\phi_2"'] & \mathcal{H}_{(1)} \times \mathbb{P}_n(\mathbb{C}), \arrow[d, "\phi_2"] \\
& \mathcal{H}_{(1)} & & \mathbb{P}_n(\mathbb{C})
\end{tikzcd}$$

where $i : V_{(1)} \longrightarrow \mathcal{H}_{(1)} \times \mathbb{P}_n(\mathbb{C})$ is the inclusion, and $\phi_1 : V_{(1)} \longrightarrow \mathcal{H}_{(1)}$ and $\phi_2 : V_{(1)} \longrightarrow \mathbb{P}_n(\mathbb{C})$ are the canonical projections. Now, observe that the value of the normal jacobians is known from Proposition 13, applied to the particular case that $(d) = (1, \ldots, 1)$:

$$\frac{NJ_{((0,M),e_0)}\phi_1}{NJ_{((0,M),e_0)}\phi_2} = det(MM^*),$$

for every nonsingular matrix $M \in S^1(\mathcal{M}_n(\mathbb{C}))$. The Coarea Formula, as used in Proposition 26, yields:

$$\int_{M \in S^1(\mathcal{H}_{(1)})} \widetilde{B}_\varepsilon(t, M) \; dS^1(\mathcal{H}_{(1)}) =$$

$$\nu_{\mathbb{P}}[\mathbb{P}_n(\mathbb{C})] \int_{S^1(\mathcal{M}_n(\mathbb{C}))} \frac{NJ_{((0,M),e_0)}\phi_1}{NJ_{((0,M),e_0)}\phi_2} \widetilde{B}_\varepsilon(t, (0, M)) dS^1(\mathcal{M}_n(\mathbb{C})),$$

Now, observe that $\widetilde{B}_\varepsilon(t, (0, M)) = B_\varepsilon(tM, e_0)$ and the proposition follows.

■

5 The Last Straight-Line of the argument.

Let $Y \subseteq \mathbb{R} \times \mathbb{C}^{N+1}$ be the set defined in the Introduction. For the sake of readability, let us recall the definition of Y and $G_{(d)}$.

$$Y := [0, 1] \times B^1(L_{e_0}^\perp) \times B^1(\mathcal{H}_{(1)}).$$

We assume Y is endowed with the product space probability. Let $\tau \in \mathbb{R}^+$ be the real number defined as

$$\tau := \sqrt{\frac{n^2+n}{N}}.$$

Let us fix any mapping $\phi : \mathcal{H}_{(1)} \longrightarrow \mathcal{U}_{n+1}$ such that for every matrix $M \in \mathcal{H}_{(1)}$, $\phi(M) \in \mathcal{U}_{n+1}$ is a unitary matrix such that $M\phi(M)e_0 = 0$. Then, the mapping $G_{(d)} : Y \longrightarrow V_{e_0}$ defined in the Introduction may also be defined as follows: For every point $(t, h, M) \in Y$, $G_{(d)}(t, h, M) \in V_{e_0}$ equals

$$\left(1 - \tau^2 t^{\frac{1}{n^2+n}}\right)^{1/2} \frac{\Delta^{-1}h}{\|h\|_2} + \tau t^{\frac{1}{2n^2+2n}} \psi_{e_0}^{-1}\left(T_{e_0}\left(\frac{M}{\|M\|_F}\phi\left(\frac{M}{\|M\|_F}\right)\right)\right).$$

Observe that $G_{(d)}$ is not defined in the case that $M = 0$ or $h = 0$. We are dealing with probabilities, so we can just omit these probability zero cases. This section is devoted to proving Theorem 6 at the Introduction.

In order to prove this theorem, we make use of the following technical statements.

Lemma 33 *Let $f : [0,1] \longrightarrow \mathbb{R}_+$ be a positive real-valued measurable function and assume that for some positive integers M, p, the following inequality holds.*

$$\int_0^1 (1-t^2)^M t^{p-1} f(t)dt \leq H.$$

Then, for every positive real number $t_0 < 1$, the following inequality also holds:

$$(1 - t_0^{2/p})^M \int_0^{t_0} f(t^{1/p})dt \leq pH.$$

Proof.– Observe that

$$p\int_0^1 (1-t^2)^M t^{p-1} f(t)dt = \int_0^1 (1-t^{2/p})^M f(t^{1/p})dt.$$

Thus,

$$(1 - t_0^{2/p})^M \int_0^{t_0} f(t^{1/p})dt \leq \int_0^{t_0} (1-t^{2/p})^M f(t^{1/p})dt \leq pH.$$

■

Corollary 34 *Assume that there exists an index $i \in \{1, \ldots, n\}$ such that $d_i > 1$. Then, following the notation above, the following inequality holds for every positive real number $t_0 \in (0,1)$:*

$$\left(1 - t_0^{\frac{1}{n^2+n}}\right)^{N-n^2-n} \int_0^{t_0} K_\varepsilon\left(t^{\frac{1}{2n^2+2n}}, e_0\right) dt \leq$$

$$\frac{\nu_\Delta[\mathbb{S}_\Delta](2n^2+2n)}{\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]\nu[S^1(\mathcal{H}_{(1)})]}\frac{32\pi}{c}\varepsilon^2 n^3(n+1)N^2 d^{3/2},$$

where K_ε is the function introduced at Proposition 30 above. Moreover, let

$$t_0 := \left(\frac{n^2+n}{N}\right)^{n^2+n}.$$

Then the following inequality holds:

$$\frac{1}{t_0}\int_0^{t_0} K_\varepsilon\left(t^{\frac{1}{2n^2+2n}}, e_0\right)dt \leq 10^4\varepsilon^2 n^5 N^2 d^{3/2}.$$

Proof.– The first inequality is an immediate corollary of Proposition 30 and Lemma 33. As for the second one, observe that

$$t_0\left(1-t_0^{\frac{1}{n^2+n}}\right)^{N-n^2-n} = \frac{(N-n^2-n)^{N-n^2-n}(n^2+n)^{n^2+n}}{N^N}.$$

From sharp Stirling inequalities (cf. for example [Stă01]), this last quantity is greater than

$$\left[\sqrt{2\pi}e^{1/6}\sqrt{n^2+n}\binom{N}{n^2+n}\right]^{-1}.$$

On the other hand,

$$\frac{\nu_\Delta[\mathbb{S}_\Delta]}{\nu_\Delta[S^1_\Delta(L^\perp_{e_0})]\nu[S^1(\mathcal{H}_{(1)})]} = \frac{1}{2n^2+2n}\binom{N}{n^2+n}^{-1},$$

and we obtain that

$$\frac{1}{t_0}\int_0^{t_0} K_\varepsilon\left(t^{\frac{1}{2n^2+2n}}, e_0\right)dt \leq \sqrt{2\pi}e^{1/6}\frac{32\pi}{c}\varepsilon^2 n^{7/2}(n+1)^{3/2}N^2 d^{3/2}.$$

The corollary follows from the inequality

$$\sqrt{2\pi}e^{1/6}32\pi \leq 300,$$

using that $(n+1)^{3/2} \leq 3n^{3/2}$ for every positive integer $n \in \mathbb{N}$. ■

We define now the function $\widetilde{A_\varepsilon} : Y \longrightarrow \mathbb{R}_+$ given by:

$$\widetilde{A_\varepsilon}(t,h,M) := A_\varepsilon(G_{(d)}(t,h,M), e_0).$$

Then, Corollary 34 may be rewritten as follows.

Proposition 35 *With the notation above, the following inequality holds:*

$$E_Y[\widetilde{A_\varepsilon}] \leq 10^4\varepsilon^2 n^5 N^2 d^{3/2}.$$

Proof.– Let X be the following compact affine set:

$$X := [0, t_0] \times S^1_\Delta(L^\perp_{e_0}) \times S^1(\mathcal{H}_{(1)}) \subseteq \mathbb{R} \times \mathbb{C}^{N+1},$$

endowed with the product Riemannian structure. Let $G'_{(d)} : X \longrightarrow V_{e_0}$ be the mapping defined as follows. For a point $(t, h, M) \in X$,

$$G'_{(d)}(t, h, M) := (1 - t^{\frac{1}{n^2+n}})^{1/2} h + t^{\frac{1}{2n^2+2n}} \psi^{-1}_{e_0}(T_{e_0}(M\phi(M))).$$

From the definitions, Proposition 32 and Corollary 34, we obtain that

$$E_X[A_\varepsilon \circ G'_{(d)}] \leq 10^4 \varepsilon^2 n^5 N^2 d^{3/2}.$$

(note the abuse of notation in the expression $A_\varepsilon \circ G'_{(d)}$, in a more correct way we should say $A_\varepsilon \circ (G'_{(d)} \times Id_{e_0})$). Now, let $F : Y \longrightarrow X$ be the mapping defined as follows:

$$F(t, h, M) := \left(t_0 t, \frac{\Delta^{-1} h}{\|h\|_2}, \frac{M}{\|M\|_F} \right).$$

Observe that $G_{(d)} = G'_{(d)} \circ F$ (hence, $\widetilde{A_\varepsilon} = A_\varepsilon \circ G'_{(d)} \circ F$). Thus, the Coarea Formula applied to $F : Y \longrightarrow X$ yields

$$\int_{y \in Y} \widetilde{A_\varepsilon}(y)\, dY = \int_{x \in X} A_\varepsilon \circ G'_{(d)}(x) \int_{y \in F^{-1}(x)} \frac{1}{NJ_y F}\, dF^{-1}(x)\, dX.$$

Following a very similar argument to that in the proof of Lemma 22, we can check that the inner integral is a constant and its value is

$$\frac{\nu_Y[Y]}{\nu_X[X]}.$$

Thus, $E_Y[\widetilde{A_\varepsilon}] = E_X[A_\varepsilon \circ G'_{(d)}]$ and the proposition follows. ■

5.1 Proof of Theorem 6.

Recall the well known Markov's Inequality, which states that for any random variable Z:

$$Probability[Z \geq a] \leq \frac{E[Z]}{a}.$$

From Proposition 35, for a random input $(t, h, M) \in Y$, with probability at least $1 - (10^4 n^5 N^2 d^{3/2})^{1/2} \varepsilon$, the following holds:

$$A_\varepsilon(G_{(d)}(t, h, M), e_0) \leq (10^4 n^5 N^2 d^{3/2})^{1/2} \varepsilon.$$

Now, this result holds for every $\varepsilon > 0$. Thus, we may change each occurrence of ε by $(10^4 n^5 N^2 d^{3/2})^{-1/2}\varepsilon$. Hence, we have that for a random input $(t, h, M) \in Y$, with probability at least $1 - \varepsilon$, the following holds:

$$A_{(10^4 n^5 N^2 d^{3/2})^{-1/2}\varepsilon}(G_{(d)}(t, h, M), e_0) \leq \varepsilon.$$

Now, assume that $g := G_{(d)}(t, h, M)$ satisfies the formula above. We prove that g is a ε-efficient pair. We recover the notations of Section 3. For every $f \in \mathbb{S}_\Delta$, let

$$k_{f,g} := 18 d^{3/2} \max_{(h,z) \in \mathfrak{L}_\Delta(f,g,e_0)} \{\mu_{norm}(h, z)\}^2.$$

Observe that from Theorem 20, the smallest integer greater than $k_{f,g}$ is an upper bound for the number of steps necessary for the linear homotopy with initial pair (g, e_0).

Moreover, observe that the following equality holds:

$$\frac{\nu_\Delta[f \in \mathbb{S}_\Delta : k_{f,g} \geq 18 \cdot 10^4 n^5 N^2 d^3 \varepsilon^{-2}]}{\nu_\Delta[\mathbb{S}_\Delta]} =$$

$$A_{(10^4 n^5 N^2 d^{3/2})^{-1/2}\varepsilon}(g, e_0) \leq \varepsilon.$$

Hence, we have proved that for randomly chosen $f \in \mathbb{S}_\Delta$, with probability at least $1 - \varepsilon$, the HD with initial pair (g, e_0) performing $18 \cdot 10^4 n^5 N^2 d^3 \varepsilon^{-2}$ (the smallest integer greater than this) homotopy steps finds an approximate zero of f. Namely, g is a ε-efficient pair. This finishes the proof of the theorem. ■

6 From continuous to discrete estimations.

In this section we give a discrete version of Theorem 6 using techniques of Geometry of Numbers and Real Semi-algebraic Geometry. As a consequence, we obtain the proof of Theorem 8 at the Introduction. We will follow the ideas of [CPM03, CMPSM02], although for our purposes the affine estimates in these works are enough.

6.1 Technical statements from Semi-algebraic geometry.

Let $M := (M^0, \ldots, M^n) \in \mathcal{H}_{(1)}$ be a maximal rank matrix, with columns $M^0, \ldots, M^n \in \mathbb{C}^n$. We define the vector $v(M) := (v(M)_0, \ldots, v(M)_n) \in \mathbb{C}^{n+1}$ as follows:

$$v(M)_i := \begin{cases} \det(M^1, \ldots, M^n) & \text{if i=0,} \\ (-1)^i \det(M^0, \ldots, M^{i-1}, M^{i+1}, \ldots, M^n) & 1 \leq i \leq n-1, \\ (-1)^n \det(M^0, \ldots, M^{n-1}) & \text{if i=n.} \end{cases}$$

(note the similarity of this definition and that of Plücker coordinates). Let $\phi(M)$ be the matrix defined as follows. We apply the Gram-Schmidt procedure to the set of $n+1$ complex vectors $\{v(M), M_1, \ldots, M_n\}$, where $M_1, \ldots, M_n$ are the rows of M. Let $v_0^M, v_1^M, \ldots, v_n^M$ be the resulting vectors. Then, we define:

$$\phi(M) := transpose \begin{pmatrix} v_0^M \\ \vdots \\ v_n^M \end{pmatrix} \in \mathcal{U}_{n+1}.$$

Observe that for every maximal rank matrix $M \in \mathcal{H}_{(1)}$, $M\phi(M)e_0 = 0$.

Let $W \subseteq \mathbb{R} \times \mathbb{C}^{N+1}$ be a subset definable as semi-algebraic subset under the identification $\mathbb{C} \equiv \mathbb{R}^2$, such that $W \subseteq B_\infty^1(0,1) := \{z \in \mathbb{R} \times \mathbb{C}^{N+1} : \|z\|_\infty \leq 1\}$. We may consider the lattice $\mathbb{Z}^{2N+3} \subseteq \mathbb{R} \times \mathbb{C}^{N+1}$, which is a free module over $\mathbb{Z}$ of dimension $2N+3$. For every positive integer H, we denote by $N_{\mathbb{Z}}(W,H)$ the following number:

$$N_{\mathbb{Z}}(W,H) := \sharp(W \cap \frac{1}{H}\mathbb{Z}^{2N+3}).$$

Let $m \geq 1$ be any positive integer, and let $\{y_i : 1 \leq i \leq m\} \subseteq Y$ be a finite collection of points of Y, where Y is the semi-algebraic set of Section 5. We may consider the discrepancy

$$\mathscr{D}_{\{y_1,\ldots,y_m\}} := \left| \frac{1}{m} \sum_{i=1}^{m} \widetilde{A_\varepsilon}(y_i) - E_Y[\widetilde{A_\varepsilon}] \right|,$$

where $\widetilde{A_\varepsilon}$ is as defined in Section 5. For every positive real number $\varepsilon > 0$ we consider the following subset of $Y \times \mathbb{S}_\Delta$:

$$R_\varepsilon := \{((t,h,M),f) \in Y \times \mathbb{S}_\Delta : f \neq \pm G_{(d)}(t,h,M),$$
$$\mathfrak{L}_\Delta(f, G_{(d)}(t,h,M), e_0) \cap \Sigma'_\varepsilon \neq \emptyset\} \subseteq Y \times \mathbb{S}_\Delta,$$

where we use the notations of Subsection 4.2. For every positive real number $\varepsilon > 0$ and for every $f \in \mathbb{S}_\Delta$, we consider the set $R_{\varepsilon,f}$ defined as follows:

$$R_{\varepsilon,f} := \{(t,h,M) \in Y : ((t,h,M),f) \in R_\varepsilon\} \subseteq Y.$$

The following lemma describes R_ε as a semi-algebraically definable set. The precise definitions required to understand it can be read in detail in [CPM03]. As a brief idea, observe that, for positive integers $m, k, s, d' \in \mathbb{N}$, a set $W \subseteq \mathbb{R}^{m+1}$ is a "k-projection of an (s,d')-definable semi-algebraic set" if there exists a semi-algebraic set definable with at most s equations of degree d', $W' \subseteq \mathbb{R}^{k+m+1}$, such that W is the projection of W' onto $\mathbb{R}^{m+1}$.

Lemma 36 *Assume there exists an $i, 1 \leq i \leq n$, such that $d_i \geq 2$. Then, for every positive real number $\varepsilon > 0$ and for every element $f \in \mathbb{S}_\Delta$, the set $R_{\varepsilon,f}$ is the k-projection of an (s,d')-definable semi-algebraic set, where*

$$\begin{array}{rcl} k & = & 3n+10, \\ s & \leq & d^{O(nN^3)}, \\ d' & \leq & d^{O(nN^3)}. \end{array}$$

Proof.– Observe that a point $((t,h,M),f) \in Y \times \mathbb{S}_\Delta$ is in R_ε if and only if the following property holds:

$$f \neq \pm G_{(d)}(t,h,M) \tag{4}$$

and

$$\exists\ (s_1,s_2) \in S^1(\mathbb{R}), \zeta \in S^1(\mathbb{C}^{n+1}) \text{ such that}$$

$$(s_1 f + s_2 G_{(d)}(t,h,M), \zeta) \in \mathfrak{L}_\Delta(f, G_{(d)}(t,h,M), e_0) \text{ and} \tag{5}$$

$$(s_1 f + s_2 G_{(d)}(t,h,M), \zeta) \in \Sigma'_\varepsilon, \tag{6}$$

where we denote by the same symbol the point $\zeta \in S^1(\mathbb{C}^{n+1})$ and the associated point in $\mathbb{P}_n(\mathbb{C})$.

If we write $g(s_1,s_2,t,h,M) := s_1 f + s_2 G_{(d)}(t,h,M)$, the property (6) is equivalent to

$$\exists \mu \in \mathbb{R} : |\mu| < \varepsilon^2,$$

$$\det(\mu Id_n - \Delta(d)^{-1/2} T_\zeta g(s_1,s_2,t,h,M) T_\zeta g(s_1,s_2,t,h,M)^* \Delta(d)^{-1/2}) = 0,$$

where $T_\zeta g(s_1,s_2,t,h,M)$ is the differential matrix of $g(s_1,s_2,t,h,M)$ restricted to $\zeta^\perp$ for some orthonormal basis. Equivalently, we may write the property (6) as follows.

$$\exists \mu \in \mathbb{R} : |\mu| < \varepsilon^2,$$

$$\det(\mu Id_n - \Delta(d)^{-1/2} d_\zeta g(s_1,s_2,t,h,M) d_\zeta g(s_1,s_2,t,h,M)^* \Delta(d)^{-1/2}) = 0,$$

so that we have eliminated the dependence on the orthogonal space $\zeta^\perp$. This is due to the fact that the singular values of $d_\zeta g(s_1,s_2,t,h,M)$ and $T_\zeta g(s_1,s_2,t,h,M)$ are equal.

Observe that for any sequence of positive real numbers $\lambda_1, \ldots, \lambda_n$, the following equality holds:

$$\phi \begin{pmatrix} \lambda_1 M_1 \\ \vdots \\ \lambda_n M_n \end{pmatrix} = \phi(M). \tag{7}$$

We consider the following sequence of real positive numbers:

$$\begin{array}{rcl} \lambda_0 & := & \frac{1}{\|v(M)\|_2}, \\ \lambda_1 & := & \frac{1}{\|M_1 - \langle M_1, v_0^M \rangle v_0^M \|_2}, \\ \vdots & & \\ \lambda_n & := & \frac{1}{\|M_n - \sum_{i=1}^{n-1} \langle M_n, v_i^M \rangle v_i^M \|_2}. \end{array}$$

Then, from equation (7), the vectors $v_0^M, \ldots, v_n^M$ defining $\phi(M)$ satisfy the following equalities:

$$\begin{array}{rcl} v_0^M & = & \lambda_0 v(M), \\ v_1^M & = & \lambda_1(M_1 - \langle M_1, v_0^M\rangle v_0^M), \\ & \vdots & \\ v_n^M & = & \lambda_n(M_n - \sum_{i=1}^{n-1}\langle M_n, v_i\rangle v_i). \end{array}$$

Hence, we can express every coordinate of v_k^M as a polynomial on these variables, of degree at most $2^{2k}(n+1)$, and every element of $\phi(M)$ may be expressed as a polynomial of degree at most $2^{2n}(n+1)$. Moreover, for every $t \in [0,1]$, we consider points $(t_1, t_2) \in S^1(\mathbb{R})$, and $l_1, l_2 \in \mathbb{R}$ such that:

$$t_2 = \left(\frac{n^2+n}{N}\right)^{1/2} t^{\frac{1}{2n^2+2n}}, \qquad l_1 := \frac{1}{\|h\|_2}, \qquad l_2 := \frac{1}{\|M\|_F}.$$

Then, we can write:

$$g(s_1, s_2, t, h, M) = s_1 f + s_2 \left[t_1 l_1 \Delta^{-1} h + t_2 \psi_{e_0}^{-1}\left(T_{e_0}(l_2 M\phi(M))\right)\right].$$

Thus, every coefficient of $g(s_1, s_2, t, h, M)$ can be expressed as a polynomial of degree at most $2^{2n+1}(n+1)$ on the variables

$$t_1, t_2, s_1, s_2, l_1, l_2, \lambda_0, \ldots, \lambda_n, M, h.$$

Hence, the corresponding expression for the elements of $d_\zeta g(s_1, s_2, t, h, M)$ is a polynomial of degree at most $2^{2n+1}(n+1)d$ on the variables above plus the variables of ζ. We deduce that the equality

$$\det(\mu Id_n - \Delta(d)^{-1/2} d_\zeta g(s_1, s_2, t, h, M) d_\zeta g(s_1, s_2, t, h, M)^* \Delta(d)^{-1/2}) = 0,$$

can be expressed as a polynomial of degree at most $2^{2n+2}(n+1)^2 d$ on the variables $\mu, s_1, s_2, t_1, t_2, l_1, l_2, h, M, \lambda_0, \ldots, \lambda_n, \zeta$. Moreover, for $1 \le i \le n$, λ_i satisfies a polynomial equality in the variables of M of degree at most $2^{2n+1}(n+1)$. We conclude that condition (6) may be written as

$$\exists \mu \in \mathbb{R} : \mu^2 < \varepsilon^4, (P_1 = 0),$$

where P_1 is a polynomial of degree bounded by $2^{2n+2}(n+1)^2$ in the variables above.

With respect to the condition (4), observe that it is equivalent to the fact that the rank of the two row matrix consisting of the coefficients of f and $G_{(d)}(t, h, M)$ is 2. That is an inequality of degree also bounded by $2^{2n+2}(n+1)^2$.

As for condition (5), from [BPR98] the fact that $(g(s_1, s_2, t, h, M), \zeta) \in \mathfrak{L}_\Delta(f, G_{(d)}(t, h, M), e_0)$, may be expressed as a semi-algebraic condition with

$$d^{O(nN^3)}$$

polynomials of degree at most

$$d^{O(nN^3)}.$$

The lemma follows. ■

Lemma 37 *Assume there exists an $i, 1 \le i \le n$, such that $d_i \ge 2$. With the notations as above, for any collection of points $\{y_i : 1 \le i \le m\} \subseteq Y$, the following inequality holds.*

$$\mathscr{D}_{\{y_1,\ldots,y_m\}} \le \frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \left| \frac{1}{m}\sum_{i=1}^{m} \chi_{R_\varepsilon}(y_i, f) - \frac{\nu_Y[R_{\varepsilon,f}]}{\nu_Y[Y]} \right| \, d\mathbb{S}_\Delta.$$

Proof.– First, observe that for every $(t, h, M) \in Y$,

$$\widetilde{A_\varepsilon}(t,h,M) = \frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \chi_{\mathcal{L}'_\varepsilon}(\mathfrak{L}_\Delta(f, G_{(d)}(t,h,M), e_0)) d\mathbb{S}_\Delta =$$

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \int_{f\in\mathbb{S}_\Delta} \chi_{R_\varepsilon}((t,h,M), f) d\mathbb{S}_\Delta.$$

Thus, $\mathscr{D}_{\{y_1,\ldots,y_m\}}$ equals

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \left| \frac{1}{m}\sum_{i=1}^{m} \int_{f\in\mathbb{S}_\Delta} \chi_{R_\varepsilon}(y_i, f) d\mathbb{S}_\Delta - \int_{y\in Y}\int_{f\in\mathbb{S}_\Delta} \chi_{R_\varepsilon}(y, f) d\mathbb{S}_\Delta \; dY \right|.$$

From Fubini Theorem, this last quantity equals

$$\frac{1}{\nu_\Delta[\mathbb{S}_\Delta]} \left| \int_{f\in\mathbb{S}_\Delta} \left(\frac{1}{m}\sum_{i=1}^{m} \chi_{R_\varepsilon}(y_i, f) - \int_{y\in Y} \chi_{R_\varepsilon}(y, f) \; dY \right) d\mathbb{S}_\Delta \right|,$$

and the lemma follows. ■

The following technical lemma follow from Corollary 11 of [CPM03] and Lemma 36.

Lemma 38 *Assume there exists an $i, 1 \le i \le n$, such that $d_i \ge 2$. Let $H \ge (2N+3)^2$ be a positive integer. Let k, s, d' be the numbers of Lemma 36. With the notations above, the following inequality holds for every $f \in \mathbb{S}_\Delta$:*

$$\left| N_{\mathbb{Z}}(R_{\varepsilon,f}, H) - \nu_Y[R_{\varepsilon,f}] H^{2N+3}] \right| \le d^{O(n^2N^3)}.$$

Observe that Corollary 11 of [CPM03] also yields:

$$\left| N_{\mathbb{Z}}(Y, H) - \nu_Y[Y] H^{2N+3} \right| \le 12^{2N+4} H^{2N+2}.$$

Lemma 39 *Let $A, B, C, D, \alpha_1, \alpha_2$ be real positive numbers such that the following inequalities hold.*

$$|A - B| \leq \alpha_1, \qquad |C - D| \leq \alpha_2, \qquad |A| \leq |C|.$$

Then, the following inequality also holds:

$$\left| \frac{A}{C} - \frac{B}{D} \right| \leq \frac{\alpha_1 + \alpha_2}{|D|}.$$

The result below follows from lemmas 38 and 39.

Lemma 40 *Assume there exists an $i, 1 \leq i \leq n$, such that $d_i \geq 2$. Let $H \geq (2N+3)^2$ be a positive integer, and let $f \in \mathbb{S}_\Delta$. With the notations above,*

$$\left| \frac{N_{\mathbb{Z}}(R_{\varepsilon,f}, H)}{N_{\mathbb{Z}}(Y, H)} - \frac{\nu_Y[R_{\varepsilon,f}]}{\nu_Y[Y]} \right| \leq \frac{1}{H} \frac{d^{O(n^2N^3)} + 12^{2N+4}}{\nu_Y[Y]}.$$

In particular,

$$\left| \frac{N_{\mathbb{Z}}(R_{\varepsilon,f}, H)}{N_{\mathbb{Z}}(Y, H)} - \frac{\nu_Y[R_{\varepsilon,f}]}{\nu_Y[Y]} \right| \leq \frac{d^{O(n^2N^3)}}{H}.$$

6.2 Proof of Theorem 8.

Theorem 8 at the Introduction is a consequence of Corollary 42 in this subsection.

Corollary 41 *Assume there exists an $i, 1 \leq i \leq n$, such that $d_i \geq 2$. Let $H \geq (2N+3)^2$ be a positive integer. The following inequality holds:*

$$\left| \frac{1}{N_{\mathbb{Z}}(Y, H)} \sum_{y \in Y \cap \frac{1}{H}\mathbb{Z}^{2N+3}} \widetilde{A_\varepsilon}(y) - E_Y[\widetilde{A_\varepsilon}] \right| \leq \frac{d^{O(n^2N^3)}}{H}.$$

In particular, let $\delta(\varepsilon) := (10^4 n^5 N^2 d^{3/2})^{1/2} \varepsilon$ be this positive number. Then, there exists a universal constant $C > 0$ such that if

$$H \geq d^{Cn^2N^3} H_1,$$

for some positive real number $H_1 \geq 1$, the following inequalities hold:

$$\frac{1}{N_{\mathbb{Z}}(Y, H)} \sum_{y \in Y \cap \frac{1}{H}\mathbb{Z}^{2N+3}} \widetilde{A_\varepsilon}(y) \leq \delta(\varepsilon)^2 + \frac{1}{H_1},$$

and

$$\frac{1}{N_{\mathbb{Z}}(Y, H)} \sharp \left\{ y \in Y \cap \frac{1}{H}\mathbb{Z}^{2N+3} : \widetilde{A_\varepsilon}(y) \geq \delta(\varepsilon) \right\} \leq \delta(\varepsilon) + \frac{1}{\delta(\varepsilon) H_1}.$$

Proof.– Immediate from lemmas 37 and 40. The second inequality follows from the estimation

$$E_Y[\widetilde{A_\varepsilon}] \leq \delta(\varepsilon)^2,$$

obtained in Proposition 35. The last inequality is again a consequence of Markov's Inequality.

■

Let $Y^H := Y \cap \frac{1}{H}\mathbb{Z}^{2N+3}$ be the set defined in the Introduction. Corollary 41 then becomes the following statement:

Corollary 42 *Assume there exists an $i, 1 \leq i \leq n$, such that $d_i \geq 2$. Let $\varepsilon > 0$ be a real positive number. With the notations as above, there exists a universal constant $C > 0$ such that if*

$$\log_2 H \geq CnN^3 log_2 d + h_1,$$

for some positive real number $h_1 > 0$, then the following inequality holds:

$$\frac{1}{\sharp(Y^H)}\sharp\{y \in Y^H : \widetilde{A}_{(10^4 n^5 N^2 d^{3/2})^{-1/2}\varepsilon}(y) \geq \varepsilon\} \leq \varepsilon + \frac{1}{\varepsilon 2^{h_1}}.$$

Proof.– We change each occurrence of ε in Corollary 41 by $(10^4 n^5 N^2 d^{3/2})^{-1/2}\varepsilon$, as in Subsection 5.1. ■

Observe that Theorem 8 is an immediate consequence of this Corollary 42. In fact, it suffices to choose $h_1 = 2\log_2 \varepsilon^{-1}$, and follow the steps of the proof of Theorem 6 (cf. Subsection 5.1).

Aknowledgements The authors gratefully thank Mike Shub for helpful suggestions and discussions. We also thank the two referees for the great work they have done helping us to improve our manuscript.

References

[BCSS98] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and real computation*, Springer-Verlag, New York, 1998. MR MR1479636 (99a:68070)

[Bel06] C. Beltrán, *Sobre el Problema 17 de Smale: Teoría de la Intersección y Geometría Integral*, Ph.D. Thesis., Universidad de Cantabria, 2006.

[BP06a] C. Beltrán and L.M. Pardo, *On the complexity of non–universal polynomial equation solving: old and new results.*, Foundations of Computational Mathematics: Santander 2005. L. Pardo, A. Pinkus, E. Süli, M. Todd editors., Cambridge University Press, 2006, pp. 1–35.

[BP06b] ______, *On the probability distribution of singular varieties of given corank.*, J. Symbolic Comput. **To appear** (2006).

[BPR98] S. Basu, R. Pollack, and M.F. Roy, *Complexity of computing semi-algebraic descriptions of the connected components of a semi-algebraic set*, Proceedings of the 1998 International Symposium on Symbolic and Algebraic Computation (Rostock) (New York), ACM, 1998, pp. 25–29 (electronic). MR MR1805168

[BW93] T. Becker and V. Weispfenning, *Gröbner bases*, Graduate Texts in Mathematics, vol. 141, Springer-Verlag, New York, 1993, A computational approach to commutative algebra, In cooperation with Heinz Kredel. MR MR1213453 (95e:13018)

[CGH$^+$03] D. Castro, M. Giusti, J. Heintz, G. Matera, and L. M. Pardo, *The hardness of polynomial equation solving*, Found. Comput. Math. **3** (2003), no. 4, 347–420. MR MR2009683 (2004k:68056)

[CLO97] D. Cox, J. Little, and D. O'Shea, *Ideals, varieties, and algorithms*, second ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1997. MR MR1417938 (97h:13024)

[CMPSM02] D. Castro, J. L. Montaña, L. M. Pardo, and J. San Martín, *The distribution of condition numbers of rational data of bounded bit length*, Found. Comput. Math. **2** (2002), no. 1, 1–52. MR MR1870855 (2002k:65236)

[CPHM01] D. Castro, L. M. Pardo, K. Hägele, and J. E. Morais, *Kronecker's and Newton's approaches to solving: a first comparison*, J. Complexity **17** (2001), no. 1, 212–303. MR MR1817613 (2002c:68034)

[CPM03] D. Castro, L. M. Pardo, and J. San Martín, *Systems of rational polynomial equations have polynomial size approximate zeros on the average*, J. Complexity **19** (2003), no. 2, 161–209. MR MR1966668 (2004b:11093)

[Ded01] J.P. Dedieu, *Newton's method and some complexity aspects of the zero-finding problem*, Foundations of computational mathematics (Oxford, 1999), London Math. Soc. Lecture Note Ser., vol. 284, Cambridge Univ. Press, Cambridge, 2001, pp. 45–67. MR MR1836614 (2002d:65050)

[Ded06] ______, *Points fixes, zéros et la méthode de newton.*, Collection Mathématiques et Applications, Springer, to appear 2006.

[Dem88] J. W. Demmel, *The probability that a numerical analysis problem is difficult*, Math. Comp. **50** (1988), no. 182, 449–480. MR MR929546 (89g:65062)

[DS00] J.P. Dedieu and M. Shub, *Multihomogeneous Newton methods*, Math. Comp. **69** (2000), no. 231, 1071–1098 (electronic). MR MR1752092 (2000m:65072)

[DS01] ______, *On simple double zeros and badly conditioned zeros of analytic functions of n variables*, Math. Comp. **70** (2001), no. 233, 319–327. MR MR1680867 (2001f:65033)

[Fed69] H. Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969. MR MR0257325 (41 #1976)

[GHH$^+$97] M. Giusti, J. Heintz, K. Hägele, J. L. Montaña, J. E. Morais, and L. M. Pardo, *Lower bounds for Diophantine approximations*, J. Pure Appl. Algebra **117/118** (1997), 277–317. MR MR1457843 (99d:68106)

[GHM$^+$98] M. Giusti, J. Heintz, J. E. Morais, J. Morgenstern, and L. M. Pardo, *Straight-line programs in geometric elimination theory*, J. Pure Appl. Algebra **124** (1998), no. 1-3, 101–146. MR MR1600277 (99d:68128)

[GHMP95] M. Giusti, J. Heintz, J. E. Morais, and L. M. Pardo, *When polynomial equation systems can be "solved" fast?*, Applied algebra, algebraic algorithms and error-correcting codes (Paris, 1995), Lecture Notes in Comput. Sci., vol. 948, Springer, Berlin, 1995, pp. 205–231. MR MR1448166 (98a:68106)

[GHMP97] M. Giusti, J. Heintz, J.E. Morais, and L.M. Pardo, *Le rôle des structures de données dans les problèmes d'élimination*, C. R. Acad. Sci. Paris Sér. I Math. **325** (1997), no. 11, 1223–1228. MR MR1490129 (98j:68068)

[GLSY05a] M. Giusti, G. Lecerf, B. Salvy, and J.C. Yakoubsohn, *On location and approximation of clusters of zeros of analytic functions*, Found. Comput. Math. **5** (2005), no. 3, 257–311. MR MR2168678

[GLSY05b] M. Giusti, G. Lecerf, B. Salvy, and J.P. Yakoubsohn, *On location and approximation of clusters of zeros: case of embedding dimension one*, Found. Comp. Mathematics **to appear** (2005).

[GPW03] M. Giusti, L.M. Pardo, and V. Weispfenning, *Algorithms of commutative algebra and algebraic geometry: Algorithms for polynomial ideals and their varieties*, Handbook of Computer Algebra, Grabmeier, Kaltofen & Weispfenning eds., Springer Verlag, 2003.

[GVL96] Gene H. Golub and Charles F. Van Loan, *Matrix computations*, third ed., Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 1996. MR MR1417720 (97g:65006)

[GZ79] C. B. García and W. I. Zangwill, *Finding all solutions to polynomial systems and other systems of equations*, Math. Programming **16** (1979), no. 2, 159–176. MR MR527572 (80f:65057)

[Hig02] N.J. Higham, *Accuracy and stability of numerical algorithms*, second ed., Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. MR MR1927606 (2003g:65064)

[How93] R. Howard, *The kinematic formula in Riemannian homogeneous spaces*, Mem. Amer. Math. Soc. **106** (1993), no. 509, vi+69. MR MR1169230 (94d:53114)

[HS82] J. Heintz and C.-P. Schnorr, *Testing polynomials which are easy to compute*, Logic and algorithmic (Zurich, 1980), Monograph. Enseign. Math., vol. 30, Univ. Genève, Geneva, 1982, pp. 237–254. MR MR648305 (83g:12003)

[HSS01] John Hubbard, Dierk Schleicher, and Scott Sutherland, *How to find all roots of complex polynomials by Newton's method*, Invent. Math. **146** (2001), no. 1, 1–33. MR MR1859017 (2002i:37059)

[Kim89] M.H. Kim, *Topological complexity of a root finding algorithm*, J. Complexity **5** (1989), no. 3, 331–344. MR MR1018023 (90m:65058)

[Kos93] E. Kostlan, *On the distribution of roots of random polynomials*, From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990) (New York), Springer, 1993, pp. 419–431. MR MR1246137

[KP96] T. Krick and L. M. Pardo, *A computational method for Diophantine approximation*, Algorithms in algebraic geometry and applications (Santander, 1994), Progr. Math., vol.

143, Birkhäuser, Basel, 1996, pp. 193–253. MR MR1414452 (98h:13039)

[LE99] Ross A. Lippert and Alan Edelman, *The computation and sensitivity of double eigenvalues*, Advances in computational mathematics (Guangzhou, 1997), Lecture Notes in Pure and Appl. Math., vol. 202, Dekker, New York, 1999, pp. 353–393. MR MR1661545 (2000e:65043)

[Lec01] G. Lecerf, *Une alternative aux méthodes de réécriture pour la résolution des systemes algébriques*, PhD thesis, École polytechnique, Paris, 2001.

[Lec02] ______, *Quadratic Newton iteration for systems with multiplicity*, Found. Comput. Math. **2** (2002), no. 3, 247–293. MR MR1907381 (2003f:65090)

[LVZ] A. Leykin, J. Verschelde, and A. Zhao, *Higher-order deflation for polynomial systems with isolated singular solutions*, Math ArXiV preprint math.NA/0602031.

[Mal94] G. Malajovich, *On generalized Newton algorithms: quadratic convergence, path-following and error analysis*, Theoret. Comput. Sci. **133** (1994), no. 1, 65–84, Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993). MR MR1294426 (95g:65073)

[ME98] Yanyuan Ma and Alan Edelman, *Nongeneric eigenvalue perturbations of Jordan blocks*, Linear Algebra Appl. **273** (1998), 45–63. MR MR1491598 (99d:15016)

[Mor83] A. Morgan, *A method for computing all solutions to systems of polynomial equations*, ACM Trans. Math. Software **9** (1983), no. 1, 1–17. MR MR715803 (85f:65051)

[Mor86] ______, *A homotopy for solving polynomial systems*, Appl. Math. Comput. **18** (1986), no. 1, 87–92. MR MR815774 (87c:90194)

[Mor88] F. Morgan, *Geometric measure theory: A beginner's guide*, Academic Press Inc., Boston, MA, 1988. MR MR933756 (89f:49036)

[MR02] G. Malajovich and J.M. Rojas, *Polynomial systems and the momentum map*, Foundations of computational mathematics (Hong Kong, 2000), World Sci. Publishing, River Edge, NJ, 2002, pp. 251–266. MR MR2021984 (2004k:65090)

[MS87a] A. Morgan and A. Sommese, *Computing all solutions to polynomial systems using homotopy continuation*, Appl. Math. Comput. **24** (1987), no. 2, 115–138. MR MR914807 (89b:65126)

[MS87b] ______, *A homotopy for solving general polynomial systems that respects m-homogeneous structures*, Appl. Math. Comput. **24** (1987), no. 2, 101–113. MR MR914806 (88j:65110)

[NG47] J.v. Neumann and H. H. Goldstine, *Numerical inverting of matrices of high order*, Bull. Amer. Math. Soc. **53** (1947), 1021–1099. MR MR0024235 (9,471b)

[Par95] L.M. Pardo, *How lower and upper complexity bounds meet in elimination theory*, Applied algebra, algebraic algorithms and error-correcting codes (Paris, 1995), Lecture Notes in Comput. Sci., vol. 948, Springer, Berlin, 1995, pp. 33–69. MR MR1448154 (99a:68097)

[Ren87] J. Renegar, *On the efficiency of Newton's method in approximating all zeros of a system of complex polynomials*, Math. Oper. Res. **12** (1987), no. 1, 121–148. MR MR882846 (88j:65112)

[San76] L.A. Santaló, *Integral geometry and geometric probability*, Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976, Encyclopedia of Mathematics and its Applications, Vol. 1. MR MR0433364 (55 #6340)

[Shu93] M. Shub, *Some remarks on Bezout's theorem and complexity theory*, From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990) (New York), Springer, 1993, pp. 443–455. MR MR1246139 (95a:14002)

[Sma00] S. Smale, *Mathematical problems for the next century*, Mathematics: frontiers and perspectives, Amer. Math. Soc., Providence, RI, 2000, pp. 271–294. MR MR1754783 (2001i:00003)

[SS93a] M. Shub and S. Smale, *Complexity of Bézout's theorem. I. Geometric aspects*, J. Amer. Math. Soc. **6** (1993), no. 2, 459–501. MR MR1175980 (93k:65045)

[SS93b] ______, *Complexity of Bezout's theorem. II. Volumes and probabilities*, Computational algebraic geometry (Nice, 1992), Progr. Math., vol. 109, Birkhäuser Boston, Boston, MA, 1993, pp. 267–285. MR MR1230872 (94m:68086)

[SS93c] ———, *Complexity of Bezout's theorem. III. Condition number and packing*, J. Complexity **9** (1993), no. 1, 4–14, Festschrift for Joseph F. Traub, Part I. MR MR1213484 (94g:65152)

[SS94] ———, *Complexity of Bezout's theorem. V. Polynomial time*, Theoret. Comput. Sci. **133** (1994), no. 1, 141–164, Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993). MR MR1294430 (96d:65091)

[SS96] ———, *Complexity of Bezout's theorem. IV. Probability of success; extensions*, SIAM J. Numer. Anal. **33** (1996), no. 1, 128–148. MR MR1377247 (97k:65310)

[Stă01] P. Stănică, *Good lower and upper bounds on binomial coefficients*, JIPAM. J. Inequal. Pure Appl. Math. **2** (2001), no. 3, Article 30, 5 pp. (electronic). MR MR1876263 (2003g:05018)

[TB97] L.N. Trefethen and D. Bau, III, *Numerical linear algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997. MR MR1444820 (98k:65002)

[Tur48] A. M. Turing, *Rounding-off errors in matrix processes*, Quart. J. Mech. Appl. Math. **1** (1948), 287–308. MR MR0028100 (10,405c)

[Wil65] J. H. Wilkinson, *The algebraic eigenvalue problem*, Clarendon Press, Oxford, 1965. MR MR0184422 (32 #1894)

[Yak95] J.C. Yakoubsohn, *A universal constant for the convergence of Newton's method and an application to the classical homotopy method*, Numer. Algorithms **9** (1995), no. 3-4, 223–244. MR MR1339720 (96d:65092)

[Yak00] ———, *Finding a cluster of zeros of univariate polynomials*, J. Complexity **16** (2000), no. 3, 603–638, Complexity theory, real machines, and homotopy (Oxford, 1999). MR MR1787887 (2001j:65084)

[Zen05] Zhonggang Zeng, *Computing multiple roots of inexact polynomials*, Math. Comp. **74** (2005), no. 250, 869–903 (electronic). MR MR2114653 (2005m:12011)