1

On the Complexity of Non Universal Polynomial Equation Solving: old and new results

C. Beltrán and L.M. Pardo

Universidad de Cantabria

Abstract

These pages summarize some results on the efficiency of polynomial equation solving. We focus on semantic algorithms, i.e., algorithms whose running time depends on some intrinsic/semantic invariant associated with the input data. Both computer algebra and numerical analysis algorithms are discussed. We show a probabilistic and positive answer to Smale's 17th problem. Estimates of the probability distribution of the condition number of singular complex matrices are also exhibited.

1.1 Introduction

These pages summarize some results on upper and lower complexity bounds in Elimination Theory. They are a revision of the program stated in Pardo (1995).

We focus on *Efficient Polynomial Equation Solving*. This is one of the challenges in the recent history of Computational Mathematics. Two main frameworks in scientific computing deal with this problem. Following different approaches, symbolic/algebraic computing and numerical analysis developed their own techniques for solving polynomial equations. We survey statements of both approaches. New results are contained in Sections 1.4 and 1.5.

Multivariate Polynomial Equation Solving is a central topic both in Computational Mathematics and Computational Algebraic Geometry (Elimination Theory in nineteenth century terminology). Its origin goes back to Sturm, Hermite, Cayley, and Sylvester, among others. Elimination Theory consists of the preparation of input data (polynomial equations and inequalities) to answer questions involving quantifiers. This approach also underlies Kronecker (1882), Hilbert (1890) and further developments in Algebraic Geometry. A central problem in Elimination Theory is the following:

Problem 1 (Hilbert's Nullstellensatz) *Design an efficient algorithm that performs the following task:*
Given a system of multivariate polynomial equations

$$f_1, \ldots, f_s \in \mathbb{C}[X_1, \ldots, X_n],$$

decide whether the following algebraic variety is empty or not:

$$V(f_1, \ldots, f_s) := \{x \in \mathbb{C}^n \ : \ f_i(x) = 0, 1 \leq i \leq s\}.$$

Here the term efficient refers to computational complexity. In the words of Traub &

Werschultz (1998): *"computational complexity is a measure of the intrinsic computational resources required to solve a mathematical problem"*. Computational resources are measured in terms of a computational model or computational device that performs the corresponding algorithm that solves the problem. Intrinsic here means that we measure resources required by the problem and not the concrete algorithm. Hence, computational complexity is the design and analysis of an optimal algorithm (in terms of computational resources) that solves a mathematical problem.

The notion of computational resource requirements is found in the mathematical literature for many years, although not always in a explicit form. For instance, we cite Galois who explicitly described computational requirements in his *Mémoire sur la Résolubilité des Équations par Radicaux*. Galois wrote: *"En un mot, les calculs sont impracticables"*. Galois had developed an algorithm that decides whether a univariate polynomial equation is solvable by radicals, but he realized that the computational complexity required by his procedure is excessive. The phrase thus means that he declined to perform calculations. In fact, he had discovered a central subject in computational complexity: *Intractability*.

In Galois' time, neither the notion of algorithm nor a complexity measure had been established. This relevant step in mathematics history was done *circa* 1933. The works of Gödel, Church and Turing established the notion of algorithm which in later years lead to the existence of computers. We note that Turing's work and his machine concept of algorithm also became the standard pattern for computational complexity. In these pages, we shall measure computational resources in terms of Turing machines as much as is possible.

Computational resources are measured as functions of the input length. The input length is the time we need to write down the data. The (running) time function is the function that relates input length and running time under a concrete computational model.

Intractability is one of the frustrating aspects of computational complexity studies. A mathematical problem is intractable if the computational resources required to solve it are so excessive that there is no hope of solving the problem in practice. Observe that intractability is independent of the algorithm we design. For example, mathematical problems whose running time is at least exponential in the input length are naturally intractable. These are called *exponential problems* and there is no hope of solving them in any real or future computer. The reason is that this exponential time requirement is intrinsic to the problem and not to the concrete algorithm or computer.

Tractable problems are those mathematical problems whose time function is bounded by a polynomial of the input length. Between tractability and intractability there is a wide range of problems for which nobody knows whether they are tractable or not. We call them the Boundary of Intractability (cf. Garey & Johnson (1979)). Hilbert's Nullstellensatz lies in this boundary. This simply means that no one has yet designed a tractable algorithm that solves Hilbert's Nullstellensatz, and it also means that no one has yet proved that this problem is intractable. That is, it is not known whether there is an algorithm that solves **HN** in running time which depends polynomially on the number of variables.

There are several strategies for studying the computational complexity of Hilbert's Nullstellensatz. We classify them in two main groups: *syntactical* and *semantical*.

Although these pages are mainly concerned with semantical strategies, we shall sketch some of the syntactical achievements in the study of **HN**.

Syntactical strategies are characterized by the fact that polynomials are considered as lists of coefficients (dense encoding) in certain vector spaces. They are then treated as vectors and linear algebra methods are applied to answer questions (mainly those involving quantifiers).

Historically, the first syntactical algorithm for **HN** goes back to Hilbert and his student Hermann (cf. Hermann (1926)). Hilbert and Hermann reduced **HN** to the consistency of a system of linear equations.

Hilbert's Nullstellensatz (Hilbert (1890)) states that *given a list of polynomials* $f_1, \ldots, f_s \in \mathbb{C}[X_1, \ldots, X_n]$ *of degree at most* d*, the complex algebraic variety they define* $V(f_1, \ldots, f_s) \subseteq \mathbb{C}^n$ *is empty if an only if there are polynomials* $g_1, \ldots, g_s \in \mathbb{C}[X_1, \ldots, X_n]$ *such that the following equality holds:*

$$1 = g_1 f_1 + \cdots + g_s f_s. \tag{1.1}$$

Identities such as (1.1) are called *Bézout Identities*.

From Hermann's work, we know that there is a function $D(d, n)$ which depends only on the number of variables and the maximum of the degrees, such that the following equivalence holds:

The variety $V(f_1, \ldots, f_s) \subseteq \mathbb{C}^n$ *is empty if and only if there exist polynomials* $g_1, \ldots, g_s$ *in* $\mathbb{C}[X_1, \ldots, X_n]$ *of degree at most* $D(d, n)$ *satisfying identity (1.1).*

Let us observe that Hermann's bound $D(d, n)$ reduces **HN** to the consistency question of a system of linear equations. The unknowns are the coefficients of the (possibly existing) polynomials $g_1, \ldots, g_s$ occurring in (1.1). The linear equations are determined by linear functions in the coefficients of the input polynomials $f_1, \ldots, f_s$. This approach reduces **HN** to the problem of deciding consistency of the linear system given by (1.1) involving

$$s\binom{D(d,n)+n}{n}$$

variables and equations. Its running time is obviously polynomial in this quantity. Hence, sharp upper bounds for the function $D(d, n)$ also imply sharp upper complexity bounds for this approach to solving **HN**. Studies on sharp upper bounds for $D(d, n)$ are called *Effective Nullstellensätze*. We cite Brownawell (1987), Caniglia, Galligo & J. Heintz (1988), Kollár (1988), Berenstein & Yger (1991, 1991a), Krick & Pardo (1996), Hägele, Morais, Pardo & M. Sombra (2000), Krick, Pardo & Sombra (2001) and their references. The known bounds for $D(d, n)$ can be summarized by the following inequalities:

$$d^{n-1} \leq D(d, n) \leq d^n.$$

Thus, this approach is neither efficient not applicable since the time complexity is of order

$$\binom{d^n+n}{n} \approx d^{n^2}.$$

For example, deciding consistency of a system of cubic polynomial equations in 20 variables by this method requires deciding consistency of a system of more than 3^{400} linear equations in a similar number of variables. This is intractable in any actual or future computer.

In Hägele, Morais, Pardo & Sombra (2000) a simply exponential time algorithm

(time of order d^n) to compute Bézout identities was shown, although the technique used in this paper is not syntactical but semantical.

A second syntactical strategy to deal with **HN** is due to rewriting techniques. The most used rewriting method is that of standard/Gröbner basis algorithms. Since the works Hironaka (1964) and Buchberger (1965), a huge list of references has been produced (cf. for instance Becker & Weispfenning (1993), Cox, Little & O'Shea (1997), Mora (2003), Vasconcelos (1998) and references therein). Most of these references discuss algorithms that compute Gröbner bases of an ideal. This strategy has also been fruitful in terms of implementations. Gröbner basis algorithmics is a standard primitive implemented in most computer algebra packages (Maple, Magma or Mathematica, for example). Most efficient implementations are due to Faugère (the *FGb* series). This approach has a serious drawback in terms of computational complexity. Since Mayr & Meyer (1982), we know that computing with Gröbner bases is exponential space complete and this is even worse than the running time of methods based on Effective Nullstellensätze. Computing with Gröbner bases involving more than 15 variables is not yet available. Thus, purely syntactical Gröbner bases techniques do not seem to be the best methods of dealing with **HN**.

A third syntactical strategy uses the underlying concepts of Structural Complexity. Namely, problems are classified into complexity classes and the study of the complexity of a problem consists in locating the appropriate class where this problem is complete. In Blum, Shub & Smale (1989), the authors proved that **HN** is complete in the class $\mathbf{NP}_{\mathbb{C}}$ of non–deterministic polynomial time under the abstract model of complex Turing machines (cf. also Blum, Cucker, Shub & Smale (1998)). Other authors studied the complexity of **HN** within the more realistic Turing machine framework. In Koiran (1996) (see also Rojas (2001, 2003)) the author proved that **HN** belongs to the complexity class **PH** (polynomial hierarchy).

Nevertheless, all these syntactical strategies seem to forget that we are dealing with geometric objects (algebraic varieties) and regular mappings (polynomials viewed as functions and not as mere lists of coefficients). Algebraic varieties and regular mappings are mathematical objects rich in terms of semantic invariants. They have been studied for years as an attempt to describe their topologic, geometric and arithmetic properties. These studies have generated a large number of semantic invariants that must be related to computational complexity. This idea of relating semantical invariants to complexity is not completely new. In fact, semantical invariants of geometric objects have been used to show lower complexity bounds for computational problems (see, for example, Montaña, Morais & Pardo (1996) and references therein). The converse problem was to design an algorithm that solves **HN** in time which depends polynomially on some semantical invariants of the input list of multivariate polynomials. This was achieved by the TERA experience. This TERA experience was more a current of thought than a research project that was active during the nineties. Some of its achievements will be described in Section 1.2.

Somewhere between syntactical and semantical strategies, we may find "sparse" elimination techniques as in Sturmfels (1996) and references therein. However, we do not discuss sparse elimination here.

The rest of the chapter is structured as follows. In Section 1.2 we present an overview of some of the achievements of the TERA experience. In Section 1.3 we discuss an exponential lower time bound for universal algorithms in Elimination Theory. In Section 1.4 we show a positive answer to Smale's 17th Problem. Finally, in Sec-

tion 1.5 we show sharp upper bounds for the probability distribution of the condition number of singular matrices.

1.2 Semantic Algorithms

In the middle nineties, the notion of *semantic algorithms in Elimination Theory* was introduced. Two works initiated this new generation of algorithms. In Pardo (1995), the foundations of a research program were established, whereas Giusti, Heintz, Morais & Pardo (1995) exhibited the first example of a semantical algorithm for Elimination Theory. The program of Pardo (1995) was achieved in the series of papers Giusti, Heintz, Morais, Morgenstern & Pardo (1998), Giusti, Hägele, Heintz, Montaña, Morais & Pardo (1997), Giusti, Heintz, Morais & Pardo (1997). This section is devoted to briefly sketch some of these achievements.

First of all, we reformulate Hilbert's Nullstellensatz in the following form.

Problem 2 *Design an efficient algorithm that performs the following task:*
Given a list of polynomials $f_1, \ldots, f_s, g \in \mathbb{C}[X_1, \ldots, X_n]$ *of degree at most* d*, decide whether the polynomial* g *vanishes at some point of the algebraic variety* $V(f_1, \ldots, f_s) \subseteq \mathbb{C}^n$*.*

This is the usual formulation of elimination polynomials (like resultants and discriminants) in classical Elimination Theory. This is also the usual formulation of **NP**–complete problems (cf. Heintz & Morgenstern (1993) or Pardo (1995) and references therein). Note that all **NP**–complete problems are particular instances of Problem 2 above.

In this formulation, the role played by g and the list of $f_1, \ldots, f_s$ seems to be different.

From a list of polynomials like $f_1, \ldots, f_s$ we want to compute some *information* concerning the variety $V := V(f_1, \ldots, f_s)$ of its common zeros. The information we compute is expected to be used to answer further questions involving the variety V. This information is commonly called a *solution* of the input list of polynomials $f_1, \ldots, f_s$. For instance, in Problem 2, a solution of $f_1, \ldots, f_s$ should be used to decide whether a new polynomial g vanishes at some point in V. The way we choose to represent this information in a computer may be called the *encoding* of the solution variety V.

Obviously, different questions will condition the way we represent the information on the variety V in a computer. Hence, different notions of solution lead to different kinds of algorithms and different encodings of algebraic varieties. In Section 1.4 we recall the Shub–Smale notion of solution (approximate zeros) whose potentiality is still unexplored.

The proposal in Pardo (1995) consists in the design and analysis of a semantic algorithm that performs the following task:

From an input list $f_1, \ldots, f_s$*, the algorithm outputs a description of the solution variety* $V(f_1, \ldots, f_s)$*.*

This algorithm must satisfy two main properties:

- Its running time should be bounded by some intrinsic/semantic quantity that depends on the input list.

- Its output must contain sufficient information to answer any kind of elimination question like the one described in Problem 2.

These two properties lead to a notion of solution that we briefly sketch here. It is called *Kronecker's encoding of an affine algebraic variety* (cf. Kronecker (1882)).

Let $f_1, \ldots, f_i \in \mathbb{C}[X_1, \ldots, X_n]$ be a sequence of polynomials defining a radical ideal $(f_1, \ldots, f_i)$ of codimension i. Let $V := V(f_1, \ldots, f_i) \subseteq \mathbb{C}^n$ be the complex algebraic variety of dimension $n - i$ given by its common zeros. A *Kronecker's encoding* of V is a birational isomorphism of V with some complex algebraic hypersurface in some affine complex space of dimension $n - i + 1$.

Technically, this is is expressed as follows. Firstly, let us assume that the variables $X_1, \ldots, X_n$ are in Noether position with respect to the variety V. Namely, we assume that the following is an integral ring extension:

$$\mathbb{C}[X_1, \ldots, X_{n-i}] \hookrightarrow \mathbb{C}[X_1, \ldots, X_n]/(f_1, \ldots, f_i).$$

Let $u := \lambda_{n-i+1} X_{n-i+1} + \cdots + \lambda_n X_n \in \mathbb{Q}[X_1, \ldots, X_n]$ be a linear form in the dependent variables $\{X_{n-i+1}, \ldots, X_n\}$. A Noether's normalization and the linear mapping u define a linear projection:

$$\mathcal{U} \ : \mathbb{C}^n \longrightarrow \mathbb{C}^{n-i+1} \ : \ (x_1, \ldots, x_n) \longmapsto (x_1, \ldots, x_{n-i}, u(x_1, \ldots, x_n)) \, .$$

Let $\mathcal{U} \mid_V \ : V \longrightarrow \mathbb{C}^{n-i+1}$ be the restriction of the projection $\mathcal{U}$ to the variety V. The image set of the projection $\mathcal{U} \mid_V$ is a complex hypersurface H_u in $\mathbb{C}^{n-i+1}$. Let us denote by $\chi_u \in \mathbb{C}[X_1, \ldots, X_{n-i}, T]$ the minimal equation of H_u. The polynomial χ_u is called the elimination polynomial of u with respect to V.

The linear form u is called a *primitive element* if and only if the projection $\mathcal{U} \mid_V$ defines a birational isomorphism of V with H_u.

A Kronecker solution of the system of polynomial equations $f_1 = 0, \ldots, f_i = 0$ consists of a description of the Noether normalization, the primitive element u, the hypersurface H_u and a description of the inverse of the birational isomorphism, i.e., a description of $(\mathcal{U} \mid_V)^{-1}$. Formally, this list of items can be given as follows:

- The list of variables in Noether position $X_1, \ldots, X_n$ (which includes a description of the dimension of V). It is just a regular matrix that defines a linear change of coordinates that puts the variables in Noether position.
- The primitive element $u := \lambda_{n-i+1} X_{n-i+1} + \cdots + \lambda_n X_n$ given by its coefficients in $\mathbb{Z}$ (or any other computable subfield of $\mathbb{C}$).
- The minimal equation χ_u of the hypersurface H_u.
- A description of $(\mathcal{U} \mid_V)^{-1}$. This description can be given by the following list of polynomials:
  - A non–zero polynomial $\rho \in \mathbb{C}[X_1, \ldots, X_{n-i}]$.
  - A list of polynomials $v_j \in \mathbb{C}[X_1, \ldots, X_{n-i}, T]$, $n - i + 1 \leq j \leq n$.

  These polynomials must satisfy the equality:

  $$(\mathcal{U} \mid_V)^{-1}(x, t) = \left(x_1, \ldots, x_{n-i}, \rho^{-1}(x) v_{n-i+1}(x, t), \ldots, \rho^{-1}(x) v_n(x, t)\right) ,$$

  for all $x := (x_1, \ldots, x_{n-i}) \in \mathbb{C}^{n-i}$, $t \in \mathbb{C}$, such that $(x, t) \in H_u$, $\rho(x) \neq 0$.

In 1882, Kronecker conceived an iterative procedure for solving multivariate systems of equations $F := [f_1, \ldots, f_n]$ defining zero–dimensional complex varieties. Kronecker's idea can be sketched in the following terms:

First, the procedure starts with system $[f_1]$ and it "solves" the equidimensional variety of codimension one $V(f_1) \subseteq \mathbb{C}^n$. Then the procedure runs iteratively: From a Kronecker encoding of $V(f_1, \ldots, f_i)$, the procedure "eliminates" the polynomial f_{i+1} to obtain a Kronecker encoding of the "next" variety $V(f_1, \ldots, f_{i+1})$. Proceed until you reach $i = n$. This iterative procedure has two main drawbacks, which can be explained in the following terms:

- First of all, the space problem arising with the encoding of the intermediate polynomials. The polynomials χ_u, ρ and v_j are polynomials of high degree (eventually of degree d^i) involving $n-i+1$ variables. Thus, to compute with them, the procedure has to handle all their coefficients, which amounts to
$$\binom{d^i + n - i + 1}{n - i + 1}.$$
For example, for $i := n/2$ the procedure must save more than $d^{n^2/4}$ coefficients. Handling such polynomials also requires a time complexity of similar order. This does not seem to be more efficient than the original treatment based on the Effective Nullstellensätze (cf. Section 1.1).
- Secondly, Kronecker's iterative procedure introduces a nesting of interpolation procedures. This nesting is demanded by the iterative process. Every time the procedure computes a new set of variables in Noether position, the procedure makes a recursive call of previously computed objects. This increases the time complexity function to $d^{O(n^2)}$.

The procedure was therefore forgotten by contemporary mathematicians and hardly mentioned in the literature of Algebraic Geometry. Macaulay quotes Kronecker's procedure in Macaulay (1916) and so does König (1903). But both of them thought that this procedure would require excessive running time to be efficient, and so it was progressively forgotten. Traces of this procedure can be found spread over the Algebraic Geometry literature without giving the required reference to it. For example, Kronecker's notion of solution was used in Zariski (1995) to define a notion of dimension for algebraic varieties, claiming that it was also used in the same form by Severi and others.

In Giusti, Heintz, Morais & Pardo (1995) and Pardo (1995), Kronecker's approach for solving was rediscovered without previous knowledge of this ancestor. These two works were able to overcome the first drawback (space problem of representation) of the previous methods. The technical trick was the use of a data structure coming from semi–numerical modeling: straight–line programs. This idea of representing polynomials by programs evaluating them goes back to previous work of the same research group (such as Giusti, Heintz (1991, 1993) or Krick & Pardo (1996), see also the references given in Pardo (1995)).

To overcome the second drawback (nesting), the authors introduced a method based on non–archimedean Newton's operator. The approximate zeros in the corresponding non–archimedean basin of attraction were called *Lifting Fibers* in Giusti, Hägele, Heintz, Morais, Montaña & Pardo (1997) solving the problem of nesting of interpolation procedures by Hensel's Lemma (also called the Implicit Mapping Theorem).

Unfortunately, Giusti, Hägele, Heintz, Morais, Montaña & Pardo (1997) introduced (for the Lifting Fibers) running time requirements which depend on the heights of the intermediate varieties in the sense of Bost, Gillet & Soulé (1994) or Philippon (1991,

1994, 1995). This drawback was finally overcome in Giusti, Heintz, Morais & Pardo (1997), where integer numbers were represented by straight–line programs and the following result was finally established:

Theorem 1 *(Giusti, Heintz, Morais & Pardo (1997)) There exists a bounded error probability Turing machine M which performs the following task: Given a system of multivariate polynomial equations $F := (f_1, \ldots, f_n)$, satisfying the following properties*

- $\deg(f_i) \leq d$ *and* $ht(f_i) \leq h$ *for* $1 \leq i \leq n$ *(h is the bit length of the coefficients),*
- *the ideals $(f_1, \ldots, f_i)$ are radical ideals of codimension i in the ring $\mathbb{Q}[X_1, \ldots, X_n]$ for $1 \leq i \leq n-1$,*
- *the variety $V(f_1, \ldots, f_n) \subseteq \mathbb{C}^n$ is a zero–dimensional complex algebraic variety,*

then the machine M outputs a Kronecker solution of the variety $V(f_1, \ldots, f_n)$. The running time of the machine M is polynomial in the quantities

$$\delta(F),\ n,\ h,\ d,\ L,$$

where $\delta(F)$ is the maximum of the degrees of the intermediate varieties (in the sense of Heintz (1983)), namely

$$\delta(F) := \max\{\deg(V(f_1, \ldots, f_i)) \ : 1 \leq i \leq n-1\},$$

and L is the input length in any natural encoding of multivariate polynomials.

It must be said that the coefficients of the polynomials involved in a Kronecker solution of the variety $V(f_1, \ldots, f_n)$ are given by straight–line programs. However, the complexity estimates for the Turing machine M are independent of the height.

The quantity $\delta(F)$ becomes a kind of condition number for symbolic methods to solve systems of multivariate polynomial equations by Kronecker's deformation technique.

After Giusti, Heintz, Morais, & Pardo (1997), several new authors got into the TERA experience, with several technical improvements, mainly on the exponents occurring in the polynomial upper time bound quoted in Theorem 1. Among them we can cite Giusti & Schost (1999), Lecerf (2001), Heintz, Matera & Weissbein (2001), for instance. This dependence on a semantic invariant was also translated to the problem of computing real solutions of real polynomial equations in the series of papers Bank, Giusti, Heintz & Mbakop (1997,2001), Bank, Giusti, Heintz & Pardo (2004, 2005). The algorithm was successfully implemented by Lecerf and Salvy. This implementation, involving some technical variations, was presented in Giusti, Lecerf & Salvy (2001).

Despite the expected good behavior in practical applications, the package *Kronecker* was not sufficiently efficient to deal with a reasonable number of variables. Hence, a deeper revision of the original goals was needed. Firstly, the reader should observe that the geometric degree of any input system $\delta(F)$ is generically equal to its worst case value (the *Bézout number* $\mathcal{D} := \prod_{i=1}^{n} deg(f_i)$). Secondly, this Bézout number is exponential in the number of variables and so is its running time on average. Thus, Kronecker's solving can only be efficient for a few particular instances (when $\delta(F)$ is "small"). Up to now, we have not found a good class of natural problems with "small" geometric degree $\delta(F)$.

1.3 Universal Solving

TERA experience and the results of the package Kronecker lead to two central questions:

- *Is Bézout's number $\mathcal{D}$ a barrier for the complexity of polynomial equations solvers?.*
- *In case of a positive answer, then explain the meaning of this barrier.*

An attempt to answer these two questions was the notion of *Universal Algorithms* and the results shown in Heintz, Matera, Pardo & Wachenchauzer (1998), Pardo (2000), and Castro, Giusti, Heintz, Matera & Pardo (2003). Roughly speaking, a polynomial equation solver is universal if its output contains sufficient information to answer all elimination questions concerning the solution variety. All known algorithms (either syntactical or semantical) in Elimination Theory are universal. Formalizing this idea requires some additional terminology.

Another feature of semantic algorithms is that they can be adapted to any particular data structure used to represent input polynomials. Data structures of input polynomials are typically defined by a regular morphism from some space of parameters to some space of input data. This can be formalized as follows.

Let $P_d \subseteq \mathbb{C}[X_1, \ldots, X_n]$ be the vector space of all complex polynomial of degree at most d. For a list of degrees $(d) := (d_1, \ldots, d_n)$, let $\mathcal{P}_{(d)}$ be the Cartesian product

$$\mathcal{P}_{(d)} := \prod_{i=1}^{n} P_{d_i}.$$

The vector space of *dense input encoding* $\mathcal{P}_{(d)}$ represents the class of systems of multivariate polynomials $F := [f_1, \ldots, f_n] \in \mathcal{P}_{(d)}$. We denote by $V(F) \subseteq \mathbb{C}^n$ the set of its common zeros, if any. Namely,

$$V(F) := \{x \in \mathbb{C}^n \ : \ f_i(x) = 0, 1 \leq i \leq n\}.$$

For every constructible subset $W \subseteq \mathbb{C}^m$, a *data structure for input systems* and parameters in W is a regular mapping

$$\Phi : W \subseteq \mathbb{C}^m \longrightarrow \mathcal{P}_{(d)}.$$

The mapping Φ associates to every parameter $\alpha \in W$ a system of multivariate polynomial equations $F_\alpha \in \mathcal{P}_{(d)}$. The image set $Im(\Phi)$ is the particular class of systems we want to solve and the varieties $V(F_\alpha) \subseteq \mathbb{C}^n$ are the solution varieties. The constructible set W is called the source space and its dimension $dim(W) \leq m$ is called the *source dimension*. In standard applications, the dimension m of the source space is much smaller than the dimension N of $\mathcal{P}_{(d)}$.

A polynomial equation solver adapted to the unirational family Φ takes as input a system F_α in $Im(\Phi)$ and outputs some encoding of the solution variety $V(F_\alpha) \subseteq \mathbb{C}^n$. The encoding of $V(F_\alpha)$ is written as a point in some affine space $\mathbb{C}^M$. Once again, the dimension M is usually much greater than the source dimension.

For example, the semantic algorithm described in Section 1.2 associates to every $\alpha \in W$ a Kronecker description of $V(F_\alpha)$. Namely, we represent $V(F_\alpha)$ by the list of coefficients of all polynomials occurring in a Kronecker description of $V(F_\alpha)$. This can be done in some affine space $\mathbb{C}^M$, where M is a quantity polynomial in the number of variables and linear in some quantity $\delta(\Phi)$ given as the maximum of the geometric degrees of input systems in $Im(\Phi)$. Namely,

$$\delta(\Phi) := \{deg(V(F_\alpha)) \ : \ \alpha \in W\}.$$

Generically, $\delta(\Phi)$ equals the Bézout number $\mathcal{D} := \prod_{i=1}^{n} d_i$ and, hence, it is exponential in the number of variables n. A similar phenomenon can be observed using either Cayley–Chow encoding or Macaulay's encoding of equidimensional algebraic varieties.

This yields the model described in Castro, Giusti, Heintz, Matera & Pardo (2003) that we briefly sketch. In the sequel, a *unirational family of elimination problems* is a regular morphism

$$\varepsilon \ : W \subseteq \mathbb{C}^m \longrightarrow \mathbb{C}^M.$$

The space $\mathbb{C}^M$ is called the *target space*, a point $y \in Im\ \varepsilon \subseteq \mathbb{C}^M$ is called a *target point* (also a semantical object), and the dimension M of the target space is called the *target dimension*. In the previous notation, for every $\alpha \in W$, $\varepsilon(\alpha)$ is the encoding of the solution variety $V(F_\alpha)$, where $F_\alpha \in Im(\Phi) \subseteq \mathcal{P}_{(d)}$.

Given a unirational family of elimination problems ε, a *mathematical question* concerning target points $y \in Im\ \varepsilon \subseteq \mathbb{C}^M$ is simply a *transformation* of the target space in a neighborhood of $(Im\ \varepsilon, y)$. Namely, a transformation is the germ of a mapping

$$\theta : (\mathbb{C}^M, y) \longrightarrow (\mathbb{C}^\ell, q).$$

The space $\mathbb{C}^\ell$ is called the *space of answers*, and its dimension ℓ is called the dimension of the space of answers. Usual mathematical questions concern spaces of answers of small dimension (with respect to the target dimension). For example, *decisional questions* are transformations of the semantical object into some unidimensional space of answers, i.e., transformations of the form

$$\theta : (\mathbb{C}^M, y) \longrightarrow (\mathbb{C}, q).$$

We claim that the goal of Elimination Theory is the design of algorithms that answer questions concerning target points of unirational families of polynomials.

As the target dimension is usually too big, efficient elimination procedures evaluate an alternative mapping:

$$\mu : W \subseteq \mathbb{C}^m \longrightarrow \mathbb{C}^s.$$

We call μ a *black–box*. It is usually evaluated by an algorithm whose particular form will not be discussed here.

A *versal black–box* associated with a unirational family of elimination problems $\varepsilon : W \longrightarrow \mathbb{C}^M$ is a mapping $\mu : W \longrightarrow \mathbb{C}^s$ such that the following property holds: For every source point $\alpha \in W$ and every question $\theta : (\mathbb{C}^M, \varepsilon(\alpha)) \longrightarrow (\mathbb{C}^\ell, z)$ there is a germ of a mapping $\rho : (\mathbb{C}^s, \mu(\alpha)) \longrightarrow (\mathbb{C}^\ell, z)$ such that the following diagram commutes:

$$\begin{array}{ccccc} (W, \alpha) & \xrightarrow{\varepsilon} & (Im\ \varepsilon, \varepsilon(\alpha)) & \xrightarrow{\theta} & (\mathbb{C}^\ell, z) \\ & \mu \searrow & & \nearrow \rho & \\ & & (Im\ \mu, \mu(\alpha)) & & \end{array}$$

For every source point α, the point $\mu(\alpha) \in \mathbb{C}^s$ is called the *output encoding* of the target $\varepsilon(\alpha)$. The number s of coordinates of $\mu(\alpha)$ is called the *output length*.

Proposition 2 *For every unirational family of elimination problems $\varepsilon \ : W \longrightarrow \mathbb{C}^M$ and a black–box $\mu \ : W \longrightarrow \mathbb{C}^s$, the following properties are equivalent:*

(i) *μ is a versal black–box associated with ε.*
(ii) *For every source point $\alpha \in W$ there is a mapping germ $\rho_\alpha : (\mathbb{C}^s, \mu(\alpha)) \longrightarrow (\mathbb{C}^M, \varepsilon(\alpha))$ such that the following diagram commutes.*

$$\begin{array}{ccc} (W,\alpha) & \xrightarrow{\varepsilon} & (Im\,\varepsilon, \varepsilon(\alpha)) \\ \mu \quad \searrow & & \uparrow \quad \rho_\alpha \\ & & (Im\,\mu, \mu(\alpha)) \end{array}$$

The germ ρ_α is called the interpolation procedure of the versal black–box μ at α.

Let $\varepsilon : W \longrightarrow \mathbb{C}^M$ be a unirational family of elimination problems and let $\mu : W \longrightarrow \mathbb{C}^s$ be some versal black–box associated with ε. We say that μ is *certified* if there is a mapping

$$\varphi : Im\,\varepsilon \longrightarrow Im\,\mu,$$

such that $\mu = \varphi \circ \varepsilon$.

Definition 3 *Let $\varepsilon : W \longrightarrow \mathbb{C}^M$ be a unirational family of elimination problems. A universal black–box associated with ε is a versal and certified black–box $\mu : W \longrightarrow \mathbb{C}^s$ such that the following properties hold :*

(i) *The black–box μ is holomorphic.*
(ii) *For every source point $\alpha \in W$ the interpolation procedure ρ_α of μ is the germ of a holomorphic mapping.*

Definition 4 *A polynomial equation solver is called universal if for every unirational family of elimination problems $\varepsilon : W \longrightarrow \mathbb{C}^M$ the procedure generates a universal black–box $\mu : W \longrightarrow \mathbb{C}^s$ associated with ε.*

Theorem 5 *There is a sequence $\left(\varepsilon_n : W_n \subseteq \mathbb{C}^{m(n)} \longrightarrow \mathbb{C}^{M(n)}\right)_{n\in\mathbb{N}}$ of unirational families of polynomials such that for every $n \in \mathbb{N}$ the following holds:*

(i) *The input length is linear in n. Namely, $m(n) = O(n)$.*
(ii) *The degree of the input space W_n and that of the regular mapping ε_n are also linear in n.*
(iii) *There is an explicit description of the input space W_n of length linear in n.*
(iv) *The target dimension $M(n)$ is exponential in n.*
(v) *For every $n \in \mathbb{N}$ and every universal black–box $\mu_n : W_n \longrightarrow \mathbb{C}^{s_n}$ associated with ε_n the output length s_n is exponential in the source dimension, i.e.,*

$$s_n \geq 2^n.$$

This technical statement can also be stated in the following terms.

Corollary 6 *(Castro, Giusti, Heintz, Matera & Pardo (2003)) Every universal polynomial equation solver requires exponential running time. In particular, the procedure described in Section 1.2 above is essentially optimal as universal polynomial equation solver.*

As shown in Castro, Giusti, Heintz, Matera & Pardo (2003), we can always associate to every unirational family of elimination problems, a certified black–box whose output length is linear in the source dimension. Namely, the smoothness condition on ρ_α is a necessary condition for Theorem 5 to hold. However, non–smooth exact interpolation procedures are difficult to figure out.

It is also possible to define the notion of universal solver in a numerical analysis context. For example, the algorithms implemented by Verschelde and collaborators (cf. Verschelde (2000) and references therein) are universal polynomial solvers. A universal numerical analysis solver takes as input a list of multivariate polynomial equations $F \in \mathcal{P}_{(d)}$ and outputs approximations for all zeros $\zeta \in V(F)$. Since the average number of zeros equals the Bézout number $\mathcal{D}$, it is immediate that universal numerical solvers also require exponential running time.

Corollary 6 must not be understood as a negative result. It is asking for a new generation of algorithms: *Non universal polynomial equations solvers.* The output of a non universal solver will contain only partial information about the variety of solutions. This simple idea also leads to a long series of new problems and questions. The obvious and first is nevertheless the most difficult to answer: *Which questions can be answered with the information provided by a non universal algorithm?* Much more experience with non universal solvers is still required before facing this question.

An example of a symbolic, semantic and non universal polynomial equation solver was given in San Martín & Pardo (2004). However, the worst case complexity of this algorithm is also exponential in the number of variables and, hence, intractable.

The search for non universal solvers naturally leads to numerical analysis polynomial system solvers.

1.4 Shub & Smale Approximate Zero Theory: Bézout $5\frac{1}{2}$

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 solvers in the series of papers Shub & Smale (1993, 1993a, 1993c, 1994, 1996). Other authors also treated this approach as Blum, Cucker, Shub & Smale (1998), Dedieu (2001b, 2003), Kim (1988, 1989), Malajovich (1994), Malajovich & Rojas (2002), Yakoubsohn (1995) and references therein.

Shub and Smale's theory on *approximate zeros* provides an answer to the barrier question stated in Section 1.3. The new results of this section are taken from the still unpublished manuscript Beltrán & Pardo (2005).

As in Shub & Smale (1994), the input space is the space of systems of multivariate homogeneous polynomials with dense encoding and fixed degree list. Namely, for every positive integer $d \in \mathbb{N}$, let $H_d \subseteq \mathbb{C}[X_0, \ldots, X_n]$ be the vector space of all homogeneous polynomials of degree d. 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 $N+1$ 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,

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

Note that with this notation $V(F)$ is always a non–empty projective algebraic variety.

In Beltrán & Pardo (2005), the following statement is proven. It represents a positive answer to Problem 17th of Smale (2000).

Theorem 7 *There is a bounded error probabilistic numerical analysis procedure that solves most systems of multivariate polynomial equations with running time polynomial in*

$$n, N, d.$$

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

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

In this statement the term "solves" means the "algorithm outputs non universal information" about the variety of solutions, whereas the term "most" means "with high probability of success". The precise meaning of this theorem requires some additional technical notions.

A non universal numerical analysis solver takes as input a system $F \in \mathcal{H}_{(d)}$ and outputs local information on some (mostly just one) of the zeros $\zeta \in V(F)$. The local information (close to a zero) 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 Shub & Smale (1993) or Shub (1993)).

For every input system $F \in \mathcal{H}_{(d)}$, let N_F be the projective Newton operator as introduced in Shub (1993). According to Shub & Smale (1993a), 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, the approximate zero z is rich in local information about the zero $\zeta \in V(F)$. In Castro, Hägele, Morais & Pardo (2001), the authors also observed that approximate zeros with rational coordinates contain not only local information about the associated zero, but also algebraic information. But we will not discuss these aspects here.

These basic notions stated, a non universal numerical analysis solver is an algorithm that has 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 kinds of algorithms are not conceived for solving all input systems but a large subclass of them. In principle, singular systems are not intended to be solved by our procedure. This corresponds to a further (and more delicate) analysis.

Let $\Sigma \subseteq \mathcal{H}_{(d)}$ be the class 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 $F \in \mathcal{H}_{(d)} \setminus \Sigma$).

Our main algorithmic scheme is Newton's Homotopic Deformation in the projective space (as described in Shub & Smale (1996)): 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:

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

Then, Newton's operator may be used to follow closely one of these curves $C_i(\Gamma)$ in the incidence variety. 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 based on a Newton Homotopic Deformation Technique.

Definition 8 *A Newton's Homotopic Deformation scheme (NHD 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* $\varphi(F, \varepsilon)$ *"homotopic 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 NHD scheme is an algorithm that constructs a polygonal 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 "homotopic steps" (path following methods) that go from one vertex to the next. Hence, $\varphi(F, \varepsilon)$ is the number of homotopic steps performed by the algorithm. Different subroutines have been designed to perform each one of these "homotopic steps". One of them is projective Newton's operator as described in Shub & Smale (1993), Shub (1993), Malajovich (1994).

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)).

Initial pairs with optimal tradeoff between number of steps and probability of failure are wanted. The following notion is an attempt to fix what this means.

Definition 9 *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 NHD scheme if there is an algorithm based on NHD scheme with initial pair* (G, z_0) *such that the following properties hold:*

(i) *The resource function (i.e., the number of steps)* $\varphi(F, \varepsilon)$ *is bounded by a polynomial in the quantities* $\varepsilon^{-1}, n, N, d$, *where* $d := max\{d_i \ : \ 1 \leq i \leq n\}$.

(ii) *The probability of "failure" (i.e., the probability that a system is not solved) is at most* ε.

Observe that a pair $(G, z_0) \in \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ may be ε–efficient for some positive real number $\varepsilon > 0$ and not ε'–efficient for $\varepsilon' < \varepsilon$.

Moreover, the main outcome in Shub & Smale (1994) proves 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 is an absolute breakthrough regarding the efficiency of numerical analysis polynomial equation solving. It leads to the following procedure based on NHD scheme:

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 Shub & Smale (1994)).*
- *Perform a polynomial (in $\varepsilon^{-1}, n, N, d$) number of homotopic 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.

The procedure seems to give an answer since it may compute approximate zeros for most systems of homogeneous polynomial equations. Here, most means with probability greater than $1 - \varepsilon$.

However, it has three main drawbacks. First of all, Shub & Smale (1994) prove the existence of some $\varepsilon-$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 then you cannot start computing). Shub & Smale (1994) used the term "quasi–algorithm" to explain the result they obtained, whereas Problem 17th in Smale (2000) asks for a "uniform algorithm". In a broad sense, this 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 "Smale's Conjecture" (as in Shub & Smale (1994)) 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 (1994) proved the existence of such $(G_\varepsilon, \zeta_\varepsilon)$ assuming that ζ_ε is a true zero of G_ε (i.e., $\zeta_\varepsilon \in V(G_\varepsilon)$). This means that we not only need to start at some approximate zero of G_ε but it seems that we need to start at a true and exact zero of this initial system.

Thus, any algorithm based on this version of NHD 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 $\varepsilon-$efficient initial pair $(G_\varepsilon, \zeta_\varepsilon)$. Secondly, you need to "solve" the system G_ε in order to compute the "exact" solution ζ_ε.

As "exact" solutions do not seem a good choice, we must proceed in the opposite manner. We must start at some complex point $\zeta_\varepsilon \in \mathbb{P}_n(\mathbb{C})$, given a priori. And

then, we must prove that there is a system G_ε vanishing at ζ_ε such that $(G_\varepsilon, \zeta_\varepsilon)$ is an $\varepsilon-$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 Shub & Smale (1994). But, once again, no hint on how to find G_ε from ζ_ε seems to be known.

In Beltrán & Pardo (2005) we exhibit a solution to these drawbacks. We found 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 notation.

Definition 10 *A class $\mathcal{G} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C})$ is called a correct test class (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*

$$1 - (nNd)^{O(1)}\varepsilon,$$

where $O(1)$ denotes some fixed constant independent of ε, d and n.

Note the analogy between these classes of efficient initial systems and the classes of "correct test sequences" (also "questor sets") for polynomial zero tests (as in Heintz & Schnorr (1982), Krick & Pardo (1996) or Castro, Giusti, Heintz, Matera & Pardo (2003)). The following is shown in Beltrán & Pardo (2005).

Theorem 11 *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 NHD schemes. First of all, note that the class $\mathcal{G}_{(d)}$ does not depend on the positive real number $\varepsilon > 0$ under consideration. Thus, we can define the following NHD 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 homotopic 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 an algorithm. In fact, a simple existential statement as Theorem 11 will not be better than the main outcome in Shub & Smale (1994).

In Beltrán & Pardo (2005), we exhibit an algorithmically tractable subset $\mathcal{G}_{(d)}$ which is proven to be a questor set for efficient initial pairs. This rather technical class can be defined as follows:

Let Δ be the Kostlan matrix as defined in Shub & Smale (1993a). Using this matrix, Shub & Smale (1993a) define 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 $\mathcal{H}_{(d)}$. 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. Moreover, for every subset $A \subseteq \mathbb{P}_n(\mathbb{C})$ the probability $\nu_\Delta[A]$ induced by $d\nu_\Delta$ agrees with the Gaussian measure of its projecting cone $\widetilde{A} \subseteq \mathcal{H}_{(d)}$. 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 projective point $e_0 := (1 : 0 : \cdots : 0) \in \mathbb{P}_n(\mathbb{C})$. Let $L_0 \subseteq \mathcal{H}_{(d)}$ be the class of systems of homogeneous polynomial equations given by the property:
A system $F := [\ell_1, \ldots, \ell_n] \in \mathcal{H}_{(d)}$ belongs to L_0 if and only if for every i, $1 \leq i \leq n$, there is a linear mapping $\lambda_i : \mathbb{C}^n \longrightarrow \mathbb{C}$ such that the following equality holds:

$$\ell_i := X_0^{d_i - 1} \lambda_i(X_1, \ldots, X_n).$$

Let $V_0 \subseteq \mathcal{H}_{(d)}$ be the class of all homogeneous systems $F \in \mathcal{H}_{(d)}$ that vanish at e_0. Namely,

$$V_0 := \{F \in \mathcal{H}_{(d)} \ : \ e_0 \in V(F)\}.$$

Note that L_0 is a vector subspace of V_0.

Next, let $L_0^\perp$ be the orthogonal complement of L_0 in V_0 with respect to Kostlan's metric $\langle \cdot, \cdot \rangle_\Delta$. Note that $L_0^\perp$ is the class of all systems $F \in \mathcal{H}_{(d)}$ that vanishes at e_0 and such that its derivative $DF(e_0)$ also vanishes at e_0. Namely, it is the class of all systems of polynomial equations of order at least 2 at e_0.

Let Y be the following convex affine set, obtained as the product of closed balls:

$$Y := [0, 1] \times B^1(L_0^\perp) \times B^1(\mathcal{M}_{n \times (n+1)}(\mathbb{C})) \subseteq \mathbb{R} \times \mathbb{C}^{N+1},$$

where $B^1(L_0^\perp)$ is the closed ball of radius one in $L_0^\perp$ with respect to the canonical Hermitian metric and $B^1(\mathcal{M}_{n \times (n+1)}(\mathbb{C}))$ 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 of the respective Riemannian structures and the corresponding measures and probabilities.

Let $\tau \in \mathbb{R}$ be the real number given by

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

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

Next, let us denote by $e_0^\perp$ the orthogonal complement of the affine point $(1, 0, \ldots, 0) \in \mathbb{C}^{n+1}$ with respect to the standard Hermitian metric in $\mathbb{C}^{n+1}$. Note that we may identify $e_0^\perp$ with the tangent space $T_{e_0}\mathbb{P}_n(\mathbb{C})$ to the complex manifold $\mathbb{P}_n(\mathbb{C})$ at $e_0 \in \mathbb{P}_n(\mathbb{C})$.

For every matrix $M \in \mathcal{M}_{n\times(n+1)}(\mathbb{C})$ of maximal rank, we may define a linear isomorphism $\ell_M := M\phi(M) : e_0^{\perp} \longrightarrow \mathbb{C}^n$.

Let us define a mapping $\psi_0 : \mathcal{M}_{n\times(n+1)}(\mathbb{C}) \longrightarrow L_0$ in the following terms. For every matrix $M \in \mathcal{M}_{n\times(n+1)}(\mathbb{C})$, we associate the system of homogeneous polynomial equations $\psi_0(M) \in L_0$ given by the equality:

$$\psi_0(M) := [X_0^{d_1-1} d_1^{1/2} \lambda_1(X_1, \ldots, X_n), \ldots, X_0^{d_n-1} d_n^{1/2} \lambda_n(X_1, \ldots, X_n)] \in L_0,$$

where $\ell_M := [\lambda_1, \ldots, \lambda_n] : e_0^{\perp} \longrightarrow \mathbb{C}^n$ is the linear mapping defined by the matrix $M\phi(M)$.

Define a mapping $G_{(d)} : Y \longrightarrow V_0$ in the following terms. For every $(t, h, M) \in Y$, let $G_{(d)}(t, h, M) \in V_0$ be the system of homogeneous polynomial equations given by the identity:

$$G_{(d)}(t, h, M) := \left(1 - \tau^2 t^{\frac{1}{n^2+n}}\right)^{1/2} \frac{\Delta^{-1} h}{||h||_2} + \tau t^{\frac{1}{2n^2+2n}} \psi_0\left(\frac{M}{||M||_F}\right) \in V_0.$$

Finally, let $\mathcal{G}_{(d)}$ be the class defined by the identity:

$$\mathcal{G}_{(d)} := Im(G_{(d)}) \times \{e_0\} \subseteq \mathcal{H}_{(d)} \times \mathbb{P}_n(\mathbb{C}). \tag{1.2}$$

Note that $\mathcal{G}_{(d)}$ is included in the incidence variety and that all systems in $Im(G_{(d)})$ share a common zero e_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 $(G_{(d)}(y), e_0) \in \mathcal{G}_{(d)}$.

The following statement has been shown in Beltrán & Pardo (2005).

Theorem 12 (Main) *With the above notation, the class $\mathcal{G}_{(d)}$ defined by identity (1.2) 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

$$1 - \varepsilon.$$

Additionally, 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 NHD with initial data (G, e_0) performing $O(n^5 N^2 d^4 \varepsilon^{-2})$ steps is at least

$$1 - \varepsilon.$$

As usual, the existence of questor sets immediately yields a probabilistic algorithm. This is Theorem 7 above, which is an immediate consequence of Theorem 12. The following corollary shows how this statement applies.

Corollary 13 *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*

$$O(n^{13} \varepsilon^{-2}),$$

with probability greater than $1 - \varepsilon$.

Taking $\varepsilon = \frac{1}{n^4}$ for instance, this probabilistic algorithm solves a cubic homogeneous system in running time at most $O(n^{21})$ with probability greater than $1 - \frac{1}{n^4}$.

However, randomly choosing a pair $(G, e_0) \in \mathcal{G}_{(d)}$ is not exactly what a computer can perform. We need a discrete class of $\varepsilon-$efficient initial systems. This is achieved by the following argument (that follows those in Castro, San Martín & Pardo (2002,2003)).

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:

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

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

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

Observe that $(4N+6)(\log_2 H+1)$ is a bound for the number of binary digits required to write any point $y \in Y^H$ in a computer.

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

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

We consider $\mathcal{G}^H_{(d)}$ endowed with the pull–back probability distribution obtained from Y^H via $G_{(d)}$. Namely, in order to choose a random point $(g, e_0) \in \mathcal{G}^H_{(d)}$, we choose a random point (uniform distribution) $y \in Y^H$ and we compute the point $(G_{(d)}(y), e_0) \in \mathcal{G}_{(d)}$. Then, the following statement also holds.

Theorem 14 *(Beltrán & Pardo (2005)) There exists a universal constant $C > 0$ such that for every two positive real numbers $\varepsilon > 0, H > 0$ satisfying*

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

the following properties hold.

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

$$1 - 2\varepsilon.$$

- *For $\varepsilon-$efficient initial pairs $(G, e_0) \in \mathcal{G}^H_{(d)}$, the probability that a randomly chosen input $F \in \mathcal{H}_{(d)}$ is solved by NHD with initial data (G, e_0) performing $O(n^5N^2d^4\varepsilon^{-2})$ steps is at least*

$$1 - \varepsilon.$$

Theorem 12 and its consequences represent a small step forward in the theory introduced by Shub and Smale. It simply shows the existence of a true, although probabilistic, algorithm that computes partial information of solution varieties for most homogeneous systems of polynomial equations in time which depends polynomially on the input length.

However, things are not as optimistic as they appear. First of all, the algorithm we propose here is probabilistic and, hence, "uniform" as demanded in Smale (2000). Obviously, a deterministic version is also desirable. Nevertheless, we consider this

a minor drawback that time and some investment of scientific effort will probably overcome.

The second drawback is more important. The algorithm, its efficiency and its probability of success depends upon the generic (dense) encoding of input polynomials. Namely, it is based on the full space $\mathcal{H}_{(d)}$ of input systems. This seems to be an overly mathematical working hypothesis. Note that *the aim of such kind of result is essentially that of explaining why computational numerical analysis methods run efficiently in real life computing, even when there is no well–founded reason proving their efficiency.*

In real life computing, problems modeled by polynomial equations have some structure and are a subset of the generic class of polynomials in dense encoding. Namely, real life problems provide inputs that belong to some particular subclasses of polynomial data (as unirational families of input systems given by regular mappings $\Phi : W \subseteq \mathbb{C}^m \longrightarrow \mathcal{H}_{(d)}$).

Theorem 12 states that $\mathcal{G}_{(d)}$ is a questor set of initial pairs for generic input data. However, it does not mean that F is also a questor set for input systems F in a unirational family of data like $Im(\Phi)$. As we claimed in Section 1.3, source dimension is usually much smaller than the dimension $N+1$ of the space of input systems. Hence, $Im(\Phi)$ is commonly a set of measure zero and it may be unfortunately contained in the class of systems for which $\mathcal{G}_{(d)}$ does not apply. Hence the question is *whether this algorithmic scheme (or anything inspired by these ideas) can be adapted to particular classes of input data.*

In order to deal with this open question, we need to reconsider most of the studies done by Shub and Smale on the generic case $\mathcal{H}_{(d)}$, this time applied to special subsets $Im(\Phi)$ of $\mathcal{H}_{(d)}$.

Theorem 12 owes much of its strength to the good behavior of the probability distribution of a condition number μ_{norm} in the full space of generic inputs $\mathcal{H}_{(d)}$. This is the main semantic invariant involved in the complexity of numerical analysis polynomial equation solvers (as remarked in Shub & Smale (1993)).

Any answer to the above adaptability question requires a preliminary study on the behavior of the probability distribution of the condition number for non–linear systems when restricted to submanifolds, subvarieties and particular subclasses of the generic space $\mathcal{H}_{(d)}$.

In the next section we illustrate some of the main difficulties that may arise in such a study. We deal with the adaptability question of a simpler problem. We study the probability distribution of the condition number associated with the class of singular matrices. It contains some of the main results of Beltrán & Pardo (2004a, 2004b).

1.5 The Distribution of the Condition Number of Singular Complex Matrices

Condition numbers in linear algebra were introduced in Turing (1948). They were also studied in Goldstine & von Neumann (1947) and Wilkinson (1965). Variations of these condition numbers may be found in the literature of numerical linear algebra (cf. Demmel (1988), Golub & van Loan (1996), Higham (2002), Trefethen & Bau (1997) and references therein).

A relevant breakthrough was the study of the probability distribution of these condition numbers. Works as Smale (1985) and, mainly, Edelman (1988, 1992) showed

the exact values of the probability distribution of the condition number of dense complex matrices.

From a computational point of view, these statements can be translated into the following terms. Let $\mathcal{P}$ be a numerical analysis procedure whose space of input data is the space of arbitrary square complex matrices $\mathcal{M}_n(\mathbb{C})$. Then, Edelman's statements mean that *the probability that a randomly chosen dense matrix in* $\mathcal{M}_n(\mathbb{C})$ *is a well–conditioned input for* $\mathcal{P}$ *is high.*

Sometimes, however, we deal with procedures $\mathcal{P}$ whose input space is a proper subset $\mathcal{C} \subseteq \mathcal{M}_n(\mathbb{C})$. Additionally such procedures with particular data lead to particular condition numbers $\kappa_{\mathcal{C}}$ adapted both for the procedure $\mathcal{P}$ and the input space $\mathcal{C}$. Edelman's and Smale's results do not apply with these constraints. In Beltrán & Pardo (2004a, 2004b) we introduced a new technique to study the probability distribution of condition numbers $\kappa_{\mathcal{C}}$. Namely, we introduce a technique to exhibit upper bound estimates for the quantity

$$\frac{vol[\{A \in \mathcal{C} : \kappa_{\mathcal{C}}(A) > \varepsilon^{-1}\}]}{vol[\mathcal{C}]}, \tag{1.3}$$

where $\varepsilon > 0$ is a positive real number, and $vol[\cdot]$ is some suitable measure on the space $\mathcal{C}$ of acceptable inputs for $\mathcal{P}$.

As an example of how our technique applies, let $\mathcal{C} := \Sigma^{n-1} \subseteq \mathcal{M}_n(\mathbb{C})$ be the class of all singular complex matrices. In Kahan (2000) and Stewart & Sun (1990), a condition number for singular matrices $A \in \mathcal{C}$ is considered. This condition number measures the precision required to perform kernel computations. For every singular matrix $A \in \Sigma^{n-1}$ of corank 1, the condition number $\kappa_D^{n-1}(A) \in \mathbb{R}$ is defined by the identity

$$\kappa_D^{n-1}(A) := \|A\|_F \|A^\dagger\|_2,$$

where $\|A\|_F$ is the Frobenius norm of the matrix A, $A^\dagger$ is the Moore–Penrose pseudo–inverse of A and $\|A^\dagger\|_2$ is the norm of $A^\dagger$ as a linear operator.

As Σ^{n-1} is a complex homogeneous hypersurface in $\mathcal{M}_n(\mathbb{C})$ (i.e., a cone of complex codimension 1), it is endowed with a natural volume *vol* induced by the $2(n^2 - 1)-$dimensional Hausdorff measure of its intersection with the unit disk. In Beltrán & Pardo (2004b), we prove the following statement.

Theorem 15 *With the same notation and assumptions as above, the following inequality holds:*

$$\frac{vol[A \in \Sigma^{n-1} : \kappa_D^{n-1}(A) > \varepsilon^{-1}]}{vol[\Sigma^{n-1}]} \leq 18 n^{20} \varepsilon^6,$$

This statement is (almost) immediate consequences of a wider class of results that we state below.

First of all, most condition numbers are by nature projective functions. For example, the classical condition number κ of numerical linear algebra is naturally defined as a function on the complex projective space $\mathbb{P}(\mathcal{M}_n(\mathbb{C}))$ defined by the complex vector space $\mathcal{M}_n(\mathbb{C})$. Namely, we may see κ as a function

$$\kappa : \mathbb{P}(\mathcal{M}_n(\mathbb{C})) \longrightarrow \mathbb{R}_+ \cup \infty.$$

Secondly, statements like the Schmidt–Mirsky–Eckart–Young Theorem (cf. Schmidt (1907), Eckart & Young (1936), Mirsky (1963)) imply that Smale's and Edelman's

estimates are, in fact, estimates of the volume of a tube about a concrete projective algebraic variety in $\mathbb{P}(\mathcal{M}_n(\mathbb{C}))$.

In Beltrán & Pardo (2004b), we prove a general upper bound for the volume of a tube about any (possibly singular) complex projective algebraic variety (see Theorem 16 below).

Estimates on volumes of tubes is a classic topic that began with Weyl's Tube Formula for tubes in the affine space (cf. Weyl (1939)). Formulae for the volumes of some tubes about analytic submanifolds of complex projective spaces are due to Gray (2004) and references therein. However, Gray's results do not apply even in Smale's and Edelman's case. Nor does it to particular classes $\mathcal{C}$, as above. Firstly, Gray's statements are only valid for smooth submanifolds and not for singular varieties (as, for instance, Σ^{n-1}). Secondly, Gray's theorems are only valid for tubes of small enough radius (depending on intrinsic features of the manifold under consideration) which may become dramatically small in the presence of singularities. These two drawbacks motivated us to search for a general statement that may be stated as follows.

Let $d\nu_n$ be the volume form associated with the complex Riemannian structure of $\mathbb{P}_n(\mathbb{C})$. Let $V \subseteq \mathbb{P}_n(\mathbb{C})$ be any subset of the complex projective space and let $\varepsilon > 0$ be a positive real number. We define *the tube of radius ε about V in* $\mathbb{P}_n(\mathbb{C})$ as the subset $V_\varepsilon \subseteq \mathbb{P}_n(\mathbb{C})$ given by the identity:

$$V_\varepsilon := \{x \in \mathbb{P}_n(\mathbb{C}) : d_{\mathbb{P}}(x, V) < \varepsilon\},$$

where $d_{\mathbb{P}}(x, y) := \sin d_R(x, y)$ and $d_R : \mathbb{P}_n(\mathbb{C})^2 \longrightarrow \mathbb{R}$ is the Fubini–Study distance.

Theorem 16 *Let $V \subseteq \mathbb{P}_n(\mathbb{C})$ be a (possibly singular) equidimensional complex algebraic variety of (complex) codimension r in $\mathbb{P}_n(\mathbb{C})$. Let $0 < \varepsilon \leq 1$ be a positive real number. Then, the following inequality holds*

$$\frac{\nu_n[V_\varepsilon]}{\nu_n[\mathbb{P}_n(\mathbb{C})]} \leq 2\deg(V)\left(\frac{e\ n\ \varepsilon}{r}\right)^{2r},$$

where $\deg(V)$ *is the degree of V.*

This theorem can be applied to Edelman's conditions to obtain the estimate:

$$\frac{vol[\{A \in \mathcal{M}_n(\mathbb{C}) : \kappa_D(A) > \varepsilon^{-1}\}]}{vol[\mathcal{M}_n(\mathbb{C})]} \leq 2e^2 n^5 \varepsilon^2,$$

where $\kappa_D(A) := \|A\|_F \|A^{-1}\|_2$, and *vol* is the standard Gaussian measure in $\mathbb{C}^{n^2}$. We also prove that the constants on the right–hand side of the inequality in Theorem 16 are essentially optimal.

The reader will observe that our bound is less sharp than Edelman's or Smale's bounds, although it is a particular instance of a more general statement.

Next, observe that neither Smale's, Edelman's results nor Theorem 16 exhibit upper bounds on the probability distribution described in equation (1.3). In particular, it does not apply to prove Theorem 15. In order to deal with this kind of estimate, *we need an upper bound for the volume of the intersection of an extrinsic tube with a proper subvariety.* This is our main result from Beltrán & Pardo (2004b) and is contained in the following statement.

Theorem 17 *Let $V, V' \subseteq \mathbb{P}_n(\mathbb{C})$ be two projective equidimensional algebraic varieties*

of respective dimensions $m > m' \geq 1$. Let $0 < \varepsilon \leq 1$ be a positive real number. With the same notation as in Theorem 16, the following inequality holds:

$$\frac{\nu_m[V'_\varepsilon \cap V]}{\nu_m[V]} \leq c \deg(V') n \binom{n}{m'}^2 \left[e \frac{n-m'}{m-m'} \varepsilon \right]^{2(m-m')},$$

where $c \leq 4e^{1/3}\pi$, ν_m is the $2m-$dimensional natural measure in the algebraic variety V, and $\deg(V')$ is the degree of V'.

Now, observe that the projective point defined by a corank 1 matrix $A \in \Sigma^{n-1} \subseteq \mathbb{P}(\mathcal{M}_n(\mathbb{C}))$ satisfies:

$$\kappa_d^{n-1}(A) := \frac{1}{d_{\mathbb{P}}(A, \Sigma^{n-2})},$$

where $\Sigma^{n-2} \subseteq \mathbb{P}(\mathcal{M}_n(\mathbb{C}))$ is the projective algebraic variety of all complex matrices of corank at least 2, $d_{\mathbb{P}}(A, \Sigma^{n-2}) := sind_R(A, \Sigma^{n-2})$, and d_R is the Fubini–Study distance in the complex projective space $\mathbb{P}(\mathcal{M}_n(\mathbb{C}))$.

Hence, Theorem 15 becomes an immediate consequence of Theorem 17. In Beltrán & pardo (2004b), other examples of applications of Theorem 17 are shown, including the stratification by corank of the space of complex matrices $\mathcal{M}_n(\mathbb{C})$ and the corresponding condition number.

This is just an example on how research on the adaptability question (discussed in Section 1.4) can be initiated. It is, however, far from being an appropriate treatment for the adaptability question both in linear algebra or in the non–linear case. Future advances in this direction are required.

Acknowledgements

The authors gratefully acknowledge the valuable suggestions for improving the manuscript which were made by Prof. Allan Pinkus. This research was partially supported by the spanish grant MTM2004-01167.

References

B. Bank, M. Giusti, J. Heintz and G. Mbakop (1997), 'Polar varieties, real equation solving and data structures: the hypersurface case', *J. of Complexity* **13**, 5–27.

B. Bank, M. Giusti, J. Heintz and G. Mbakop (2001), 'Polar varieties and efficient real elimination', *Math. Zeits.* **238**, 115–144.

B. Bank, M. Giusti, J. Heintz and L.M. Pardo (2004), 'Generalized polar varieties and an efficient real elimination procedure', *Kybernetika* **40**, (2004) 519–550.

B. Bank, M. Giusti, J. Heintz and L.M. Pardo (2005), 'Generalized polar varieties: Geometry and algorithms', *J. of Complexity* **21**, 377–412.

T. Becker and V. Weispfenning (1993), *Groebner Bases: A Computational Approach to Commutative Algebra,* Volume 141 of *Grad. Texts in Maths*, Springer Verlag.

C. Beltrán and L.M. Pardo (2004a), 'Upper bounds on the distribution of the condition number of singular matrices', *Comptes Rendus Mathématique* **340**, 915–919.

C. Beltrán and Luis M. Pardo (2004b), 'Estimates on the probability distribution of the condition number of singular matrices', submitted to *Found. of Comput. Math.* .

C. Beltrán and L.M. Pardo (2005), 'On Smale's 17th Problem: probabilistic polynomial time', manuscript.

C. Berenstein and A. Yger (1991), 'Effective Bézout identities in $\mathbb{Q}[X_1, \ldots, X_n]$', *Acta Math.* **166**, 69–120.

C. Berenstein and A. Yger (1991a), 'Une formule de Jacobi et ses conséquences', *Ann. Sci. E.N.S.,* 4^{ieme} *série,* **24**, 363–377.

L. Blum, M. Shub and S. Smale (1989), 'On a theory of computation and complexity over the real numbers: **NP**-completeness, recursive functions and universal machines', *Bulletin of the Amer. Math. Soc.* **21**, 1–46.

L. Blum, F. Cucker, M. Shub and S. Smale (1998), *Complexity and Real Computation*, Springer Verlag, New York.

J.-B. Bost, H. Gillet, and C. Soulé (1994), 'Heights of projective varieties and positive green forms', *J. Amer. Math. Soc.* **7**, 903–1027.

W. D. Brownawell (1987), 'Bounds for the degree in the Nullstellensatz', *Annals of Math.* **126**, 577–591.

B. Buchberger (1965), *Ein Algoritmus zum Auffinden der Basiselemente des Restklassenringes nach einen nulldimensionallen Polynomideal*, Leopold–Franzens Universität, Innsbruck, Ph.D. Thesis.

L. Caniglia, A. Galligo and J. Heintz (1988), 'Borne simplement exponentielle pour les degrés dans le théorème des zéros sur un corps de caractéristique quelconque', *C.R. Acad. Sci. Paris, t.* **307**, *Série I*, 255–258.

D. Castro, K. Hägele, J.E. Morais and L.M. Pardo (2001), 'Kronecker's and Newton's approaches to solving: a first comparison', *J. Complexity* **17**, 212–303.

D. Castro, M. Giusti, J. Heintz, G. Matera and L.M. Pardo (2003), 'The hardness of polynomial equation solving', *Found. Comput. Math.* **3**, 347–420.

D. Castro, J.L. Montaña, J. San Martín and L.M. Pardo (2002), 'The distribution of condition numbers of rational data of bounded bit length', *Found. Comput. Math.* **2**, 1–52.

D. Castro, J. San Martín and L.M. Pardo (2003), 'Systems of rational polynomial equations have polynomial size approximate zeros on the average', *J. Complexity* **19**, 161–209.

D. Cox, J. Little, D. O'Shea (1997), *Ideals, Varieties, and Algorithms*,Springer Verlag, New York.

J.P. Dedieu (2001), 'Newton's method and some complexity aspects of the zero–finding problem', in *Foundations of Computational Mathematics (Oxford, 1999)*, Volume 284 of *London Math. Soc. Lecture Note Ser.* , 45–67, Cambridge Univ. Press, Cambridge.

J.P. Dedieu (2003), *Points Fixes, Zéros et la Méthode de Newton,* manuscript, Univ. Paul Sabatier.

J.W. Demmel (1988), 'The probability that a numerical analysis problem is difficult', *Math. Comp.* **50**, 449–480.

C. Eckart, G. Young (1936), 'The approximation of one matrix by another of lower rank', *Psychometrika* **1**, 211–218.

A. Edelman (1988), 'Eigenvalues and condition numbers of random matrices', *SIAM J. Matrix Anal. Appl.* **9**, 543–560.

A. Edelman (1992), 'On the distribution of a scaled condition number', *Math. Comp.* **58**, 185–190.

M.R. Garey and D.S. Johnson (1979), *Computers and Intractability: A Guide to the Theory of NP–Completness,* W.H. Freeman, San Francisco, California.

M. Giusti, K. Hägele, J. Heintz, J.L. Montaña, J.E. Morais and L.M. Pardo (1997), 'Lower bounds for diophantine approximations', *J. Pure Appl. Algebra* **117& 118**, 277–317.

M. Giusti and J. Heintz (1991), 'Algorithmes – disons rapides – pour la décomposition d' une variété algébrique en composantes irréductibles et équidimensionelles', in *Proceedings of MEGA'90,* Volume 94 of Progress in Mathematics, 169–194, ed. T. Mora and C. Traverso, Birkhäuser Verlag, Basel.

M. Giusti and J. Heintz (1993), 'La détermination des points isolés et de la dimension d'une variété algébrique peut se faire en temps polynomial', in *Computational Algebraic Geometry and Commutative Algebra*, volume XXXIV of *Symposia Matematica*,216–256, ed. D. Eisenbud and L. Robbiano, Cambridge University Press, Cambridge.

M. Giusti, J. Heintz, J.E. Morais and L.M. Pardo (1995). 'When polynomial equation systems can be "solved" fast?', in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes (Paris, 1995)*, Volume 948 of *Lecture Notes in Comput. Sci.*, 205–231, ed. A. Cohen, M. Giusti and T. Mora, Springer Verlag, Berlin.

M. Giusti, J. Heintz, J.E. Morais, J. Morgenstern and L.M. Pardo (1998), 'Straight–line programs in geometric elimination theory', *J. Pure Appl. Algebra* **124**, 101–146.

M. Giusti, J. Heintz, J.E. Morais and L.M. Pardo (1997), 'Le rôle des structures de données dans les problèmes d'élimination', *C. R. Acad. Sci. Paris, Sér. I Math.* **325**, 1223–1228.

M. Giusti, G. Lecerf and B. Salvy (2001), 'A Gröbner free alternative for polynomial system solving', *J. of Complexity* **17**, 154–211.

M. Giusti and E. Schost (1999), 'Solving some over–determined systems', in *Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation, ISSAC'99*, 1–8, ed. S. Dooley, ACM Press, New York.

H.H. Goldstine and J. von Neumann (1947), 'Numerical inverting of matrices of high order', *Bull. Amer. Math. Soc.*, **53**, 1021–1099.

G.H. Golub and C.F. Van Loan (1996), *Matrix Computations*, 3rd Edition, *Johns Hopkins Studies in the Mathematical Sciences*, Johns Hopkins University Press, Baltimore, MD.

A. Gray (2004), *Tubes*, 2dn Edition, Volume 221 of *Progress in Mathematics* , Birkhäuser Verlag, Basel.

K. Hägele, J.E. Morais, L.M. Pardo and M. Sombra (2000), 'The intrinsic complexity of the Arithmetic Nullstellensatz', *J. of Pure and App. Algebra* **146**, 103–183.

J. Heintz (1983), 'Definability and fast quantifier elimination in algebraically closed fields'. *Theoret. Comput. Sci.* **24**, 239–27.

J. Heintz and C.P. Schnorr (1980), 'Testing Polynomials which are easy to compute', in *Logic and Algorithmic (an International Symposium in honour of Ernst Specker)*, Monographie n. 30 de *l'Enseignement Mathématique*, 237–254. A preliminary version appeared in *Proc. 12th Ann. ACM Symp. on Computing* (1980) 262–268.

J. Heintz, G. Matera, L.M. Pardo and R. Wachenchauzer (1998), 'The intrinsic complexity of parametric elimination methods', *Electron. J. SADIO* **1**, 37–51.

J. Heintz, G. Matera and A. Weissbein (2001), 'On the time–space complexity of geometric elimination procedures', *App. Algebra Eng. Comm. Comput.* **11**, 239–296.

J. Heintz and J. Morgenstern (1993), 'On the intrinsic complexity of elimination theory', *J. of Complexity* **9**, 471–498.

G. Hermann (1926). 'Die Frage der endlich vielen Schritte in der Theorie der Polynomideale', *Math. Ann.* **95**, 736–788.

N.J. Higham (2002), *Accuracy and Stability of Numerical Algorithms*, 2nd Edition, SIAM, Philadelphia, PA.

D. Hilbert (1890), 'Über theorie der Algebraischen Formen', *Math. Ann.* **36**, 473–534.

H. Hironaka (1964), 'Resolution of singularities of an algebraic variety over a field of characteristic zero', *Annals of Math.* **79**, I : 109–203, II : 205–326.

W. Kahan (2000), *Huge Generalized Inverses of Rank-Deficient Matrices,* unpublished manuscript.

M.H. Kim (1988), 'On approximate zeros and root finding algorithms for a complex polynomial', *Math. Comp.*, **51**, 707–719.

M.H. Kim (1989), 'Topological complexity of a root finding algorithm', *J. Complexity*, **5**, 331–344.

P. Koiran (1996), 'Hilbert's Nullstellensatz is in the Polynomial Hierarchy', *J. of Complexity* **12**, 273–286.

J. Kollár (1988), 'Sharp Effective Nullstellensatz', *J. of Amer. Math. Soc.* **1**, 963–975.

J. König (1903), *Einleitung in die allgemeine Theorie der algebraischen Grözen,* Druck und Verlag von B.G. Teubner, Leipzig.

T. Krick and L. M. Pardo (1996), 'A computational method for diophantine approximation', in *Algorithms in Algebraic Geometry and Applications, Proceedings of MEGA'94*, Volume 143 of *Progress in Mathematics*, 193–254, ed. T. Recio, L. Gonzalez–Vega, Birkhäuser Verlag, Basel.

T. Krick, L.M. Pardo and M. Sombra (2001), 'Sharp estimates for the Arithmetic Nullstellensatz', *Duke Math. Journal* **109**, 521–598.

L. Kronecker (1882), 'Grundzüge einer arithmetischen theorie de algebraischen grössen', *J. reine angew. Math.*, **92**, 1–122.

G. Lecerf (2001), *Une Alternative aux Méthodes de Réécriture pour la Résolution des Systèmes Algébriques*, École Polytechnique, Thèse.

F.S. Macaulay (1916), *The Algebraic Theory of Modular Systems*, Cambridge University Press, Cambridge.

G. Malajovich (1994), 'On generalized Newton algorithms: quadratic convergence, path–following and error analysis', *Theoret. Comput. Sci.* **133**, 65–84.

G. Malajovich and M.J. Rojas (2002), 'Polynomial systems and the momentum map', in *Foundations of Computational Mathematics (Hong Kong, 2000)*,251–266, World Sci. Publishing, River Edge, NJ.

J. San Martín and L.M. Pardo (2004), 'Deformation techniques that solve generalized Pham systems', *Theoret. Comput. Sci.* **315**, 593–625.

E. Mayr and A. Meyer (1982) 'The complexity of the word problem for commutative semigroups', *Advances in Math.* **46**, 305–329.
L. Mirsky (1963), 'Results and problems in the theory of doubly-stochastic matrices', *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **1**, 319–334.
J.L. Montaña, J.E. Morais and L.M. Pardo (1996), 'Lower bounds for arithmetic networks II : Sum of Betti numbers', *Appl. Algebra Engrg. Comm. Comput.* **7**, 41–51.
T. Mora (2003), *Solving Polynomial Equation Systems. I. The Kronecker-Duval Philosophy*, Volume 88 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge.
P. Philippon (1991), 'Sur des hauteurs alternatives, I', *Math. Ann.* **289**, 255–283.
P. Philippon (1994), 'Sur des hauteurs alternatives, II', *Ann. Inst. Fourier, Grenoble* **44**, 1043–1065.
P. Philippon (1995), 'Sur des hauteurs alternatives, III', *J. Math. Pures Appl.* **74**, 345–365.
L.M. Pardo (1995), 'How lower and upper complexity bounds meet in elimination theory', in *Applied Algebra, Algebraic Algorithms and Error–Correcting Codes (Paris, 1995),* Volume 948 of *Lecture Notes in Comput. Sci.*, 33–69, ed. A. Cohen, M. Giusti and T. Mora, Springer Verlag, Berlin.
L.M. Pardo (2000), 'Universal Elimination requires exponential running time', in *Actas EACA'2000*, 25–50, ed. A. Montes, UPC, Barcelona.
M. Rojas (2001), 'Computational arithmetic geometry I: Diophantine sentences nearly within the Polynomial Hierarchy', *J. of Comput and Syst. Sci.* **62**, 216–235.
M.J. Rojas (2003), 'Dedekind Zeta functions and the complexity of Hilbert's Nullstellensatz', Math ArXiV preprint math.NT/0301111.
E. Schmidt (1907), 'Zur Theorie der linearen und nichtlinearen Integralgleichungen. I Tiel. Entwicklung willkÜrlichen Funktionen nach System vorgeschriebener', *Math. Annalen* **63**, 433–476.
M. Shub (1993), 'Some remarks on Bezout's theorem and complexity theory', in *From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990)*, 443–455, Springer Verlag, New York.
M. Shub and S. Smale (1993), 'Complexity of Bézout's theorem I: Geometric aspects', *J. Am. Math. Soc.* **6**, 459–501
M. Shub and S. Smale (1993), 'Complexity of Bezout's theorem. II. Volumes and probabilities', in *Computational algebraic geometry, Proc. MEGA'92,* Volume 109 of *Progress in Mathematics*, 267–285, Birkhäuser Verlag, Basel.
M. Shub and S. Smale (1993), 'Complexity of Bezout's theorem. III. Condition number and packing', *J. Complexity* **9**, 4–14.
M. Shub and S. Smale (1996), 'Complexity of Bezout's theorem. IV. Probability of success; extensions', *SIAM J. Numer. Anal.* **33**, 128–148.
M. Shub and S. Smale (1994), 'Complexity of Bezout's theorem. V. Polynomial time', *Theoret. Comput. Sci.* **133**, 141–164.
S. Smale (1985), 'On the efficiency of algorithms of analysis', *Bull. Amer. Math. Soc. (N.S.)* **13**, 87–121.
S. Smale (2000), 'Mathematical problems for the next century', in *Mathematics: Frontiers and Perspectives*, 271–294, ed. V. Arnold, M. Atiyah, P. Lax and B. Mazur, Amer. Math. Soc., Providence, RI.
G.W. Stewart and J.G. Sun (1990), *Matrix Perturbation Theory,* Computer Science and Scientific Computing, Academic Press, Boston, MA.
B. Sturmfels (1996), *Gröbner Bases and Convex Polytopes* , Volume 8 of *University Lecture Series* Amer. Math. Soc., Providence, RI.
J. Traub and A.G. Werschultz (1998), *Complexity and Information*, Lezioni Lincee, Cambridge University Press, Cambridge.
L.N. Trefethen and D. Bau (1997), *Numerical Linear Algebra*, SIAM, Philadelphia, PA.
A.M. Turing (1948), 'Rounding–off errors in matrix processes', *Quart. J. Mech. Appl. Math.* **1**, 287–308.
W.V. Vasconcelos (1998), *Computational Methods in Commutative Algebra and Algebraic Geometry*, Volume 2 of *Algorithms and Computation in Mathematics*, Springer Verlag, Berlin.
J. Verschelde (2000), 'Toric Newton method for polynomial homotopies', *J. of Symb. Comput.* **29**, 1265–1287.
A. Weyl (1939), 'On the volume of tubes', *Amer. J. Math* **61**, 461–472.
J.H. Wilkinson (1965), *The Algebraic Eigenvalue Problem,* Clarendon Press, Oxford.

J.C. Yakoubsohn (1995), 'A universal constant for the convergence of Newton's method and an application to the classical homotopy method'. *Numer. Algorithms* **9**, 223–244.
O. Zariski (1995).*Algebraic Surfaces*, *Classics in Mathematics*, Springer Verlag, Berlin.