Condition Numbers for the Cube. I: Univariate Polynomials and Hypersurfaces

Josué Tonelli-Cueto, Elias Tsigaridas

Condition Numbers for the Cube.
I: Univariate Polynomials and Hypersurfaces

Josué Tonelli-Cueto

Inria Paris & IMJ-PRG, Sorbonne Université
4 place Jussieu, F-75005, Paris, France

Elias Tsigaridas

Inria Paris & IMJ-PRG, Sorbonne Université
4 place Jussieu, F-75005, Paris, France

Abstract

The condition-based complexity analysis framework is one of the gems of modern numerical algebraic geometry and theoretical computer science. One of the challenges that it poses is to expand the currently limited range of random polynomials that we can handle. Despite important recent progress, the available tools cannot handle random sparse polynomials and Gaussian polynomials, that is polynomials whose coefficients are i.i.d. Gaussian random variables.

We initiate a condition-based complexity framework based on the norm of the cube that is a step in this direction. We present this framework for real hypersurfaces and univariate polynomials. We demonstrate its capabilities in two problems, under very mild probabilistic assumptions. On the one hand, we show that the average run-time of the Plantinga-Vegter algorithm is polynomial in the degree for random sparse (alas a restricted sparseness structure) polynomials and random Gaussian polynomials. On the other hand, we study the size of the subdivision tree for Descartes' solver and run-time of the solver by Jindal and Sagraloff (2017). In both cases, we provide a bound that is polynomial in the size of the input (size of the support plus logarithm of the degree) for not only on the average, but all higher moments.

Keywords: condition number; random polynomial; subdivision algorithm; univariate solver;

1. Introduction

The complexity of numerical algorithms is not uniform. It depends on a measure of the numerical sensitivity of the output with respect to perturbations of the input, called *condition number* and introduced originally by Turing (1948) and von Neumann and Goldstine (1947). When the condition number of an input is large, then this means that small numerical perturbations of

Email addresses: josue.tonelli.cueto@bizkaia.eu (Josué Tonelli-Cueto), elias.tsigaridas@inria.fr (Elias Tsigaridas)
URL: https://tonellicueto.xyz (Josué Tonelli-Cueto),
https://who.paris.inria.fr/Elias.Tsigaridas/ (Elias Tsigaridas)

the input can significantly change the solution of the problem at hand. Consequently, numerical algorithms need to use more computational resources to guarantee a correct computation.

The above phenomenon motivates the condition-based complexity analysis of numerical algorithms. Although a condition-based complexity analysis can explain the success of an algorithm for a given input, it cannot explain, at least on its own, why a numerical algorithm is efficient. The reason is that condition-based complexity analyses are not input-independent. Thus a common technique that goes back Goldstine and von Neumann (1951), Demmel (1987, 1988) and Smale (1997) is to randomize the input. In this way, we obtain a probabilistic complexity analysis that can explain the successful behaviour of an algorithm. Moreover, the framework of smoothed analysis Spielman and Teng (2002) fully explains the practical behaviour of an algorithm. We refer the reader to Bürgisser and Cucker (2013) and references therein for more details about this complexity paradigm for numerical algorithms.

After the complete solution[1] of Smale's 17th problem by Lairez (2017), following the steps of Beltrán and Pardo (2008) and Bürgisser and Cucker (2011), the main challenge in numerical algebraic geometry is to extend (and analyze) the current algorithms for solving polynomial systems to handle more general inputs; for example sparse and structured polynomials.

In the complex setting, Malajovich (2019, 2020) and Malajovich and Rojas (2002, 2004) did groundbreaking work in the development for numerical algorithms for finding a solution of sparse polynomial systems, and recently Bürgisser et al. (2020) (following the ideas of Lairez (2020)) introduces efficient numerical algorithms for finding a solution of determinant polynomial systems. Additionally, Armentano and Beltrán (2019) provided a probabilistic estimate of the condition number for the Polynomial Eigenvalue Problem, although they did not provide an algorithm.

In the real setting, the situation is far more difficult. For example, as of today, the real version of Smale's 17th problem (asking to decide numerically the feasibility of a real polynomial system) remains open as no algorithm running in finite expected time is known (see (Bürgisser and Cucker, 2013, Ch. 17 and P.18) for more details). Of course, ideally, we want to solve the real sparse Smale's 17th problem (which was proposed by Rojas and Ye (2005)):

> Find an algorithm that finds all the real roots of a real fewnomial system with average run-time bounded by a polynomial in the size of the system (which is the size of the support and the logarithm of the degree).

However, such a result (although motivated by Khovanskiĭ (1991) and Kushnirenko's hypothesis) seems to be out of reach today. Nevertheless, some progress has been made by Rojas (2020) (although not in the numerical setting). Regarding structured systems, there are some results by Beltrán and Kozhasov (2019) and Ergür, Paouris and Rojas (2019); Ergür, Paouris and Rojas (2018).

A common problem with many of the current techniques is that they rely on unitary/orthogonal invariance. Therefore, it is central for an effective algorithmic framework to develop techniques that do not rely on this invariance to compute with sparse/structured polynomials and more general probability distributions. We make one step in this research direction by developing a condition-based complexity framework that relies on the ∞-norm of the cube, and which consequently does not rely on unitary invariance.

[1] One should notice that the solution of Lairez (2017) does not construct a good starting system, but exploits the randomness of the input as a source of randomness for a deterministic algorithm. Constructing such initial systems is hard, although there was a breakthrough construction for the univariate case by Etayo et al. (2020).

We develop the above framework for univariate polynomials and hypersurfaces. We hope to extend it for polynomial systems in a future work. To illustrate its advantages we apply it to two problems. First, to study the complexity of the Plantinga-Vegter algorithm (Plantinga and Vegter, 2004; Burr et al., 2017). Then, to study the separation bounds of the roots of real univariate polynomials. Using the latter bounds we deduce a bound on the average number of subdivisions that Descartes' solver performs to isolate the real roots of univariate polynomials and we estimate the average bit complexity of the algorithm by Jindal and Sagraloff (2017) for solving sparse univariate polynomials. The latter bound is polynomial in the input size, providing a first approximation to the real sparse version of Smale's 17th problem by Rojas and Ye (2005).

In the case of the Plantinga-Vegter algorithm, we demonstrate its efficiency by showing that its complexity is polynomial on the average, for a wide class of random sparse polynomials (Theorem 7.19). This significantly extends the results by Cucker et al. (2019), (cf. Cucker et al., 2020b). Additionally, our approach applies to Gaussian random polynomials, when all coefficients have the same variance.

We note that our aim is not to show that the Plantinga-Vegter is the most efficient algorithm for random sparse polynomials, but that it remains efficient when we restrict it to a wide class of random sparse polynomials. We note that our bounds depend polynomially on the degree and not logarithmically. A similar approach was employed by Ergür et al. (2018) for the algorithm for finding real zeros of real polynomial systems from Cucker et al. (2008). However, unlike Ergür et al. (2018), our analysis applies to structured polynomials that are sparse, but with a combinatorial restriction on the support. We note that our sparseness condition is similar to that of Renegar (1987) and so is the bound we obtain; the latter is polynomial in the degree and the size of the support and exponential in the number of variables. Many computational problems in real algebraic geometry lack algorithms that are polynomial in the degree, so such bounds push the limits of the state-of-the-art.

In the case of univariate polynomials, our results imply that the complex roots of a random real univariate sparse polynomial around the unit interval are well separated with high probability. The logarithm of the separation bound is an important parameter that controls the complexity of many, if not all, univariate solvers. We exploit its relation with the condition number to obtain bounds on the size of the subdivision tree of Descartes' solver and on the average run-time of the sparse univariate solver of Jindal and Sagraloff (2017).

In both cases, that for both Descartes' solver and the solver of Jindal and Sagraloff (2017), the bounds that we obtain are: 1) polynomial in the size of the sparse polynomial (meaning polynomial in the size of the support and the logarithm of the degree, and 2) extend to all higher moments. The importance of these bounds is that they are the first step towards solving the real sparse version of Smale's 17th problem, as stated by Rojas and Ye (2005).

Our framework is based on the one hand on variational properties of the polynomials and the corresponding condition numbers and on the other on probabilistic techniques from geometric functional analysis. The former follows the variational approach to condition numbers of (Tonelli-Cueto, 2019, $2^{\S 2}$) and extends Cucker et al. (2020a) to new norms. The latter has been already applied by Ergür et al. (2019); Ergür et al. (2018) and by Cucker et al. (2019), but the way that we apply these methods takes them to the maximum development.

The 1-norm on the space of polynomials behaves as the "dual"norm to the ∞-norm on the cube. This norm is naturally suited for subdivision methods on the cube. The analysis of the Plantinga-Vegter subdivision process using our framework serves the purpose to convince the reader of the advantages of the new framework for the analysis of algorithms. It also has the ambition to bring new insights in the study of algorithms in numerical algebraic geometry. Our

approach continues the trend started by Cucker et al. (2019) of bringing further interactions between the communities of numerical algebraic geometry and symbolic computation.

A preliminary version of the paper appeared in the proceedings of ISSAC 2020 (Tonelli-Cueto and Tsigaridas, 2020). Compared with the conference paper, the current paper extends significantly the probabilistic model, incorporating random polynomials whose coefficient are subexponential, and extends significantly the treatment of the univariate case, adding several new results. Among these results, one can find polynomial (in the size of the support and logarithm of the degree) bounds for all the moments of the size of the subdivision tree of Descartes' solver and for the bit complexity of the algorithm by Jindal and Sagraloff (2017) for solving a random sparse polynomial. The latter, to our knowledge, is the first such bound.

Notation. We denote by $\mathcal{P}_{n,d}$ the space of polynomials in n variables of total degree at most d. Then a polynomial is $f = \sum_{|\alpha|\leq d} f_\alpha X^\alpha \in \mathcal{P}_{n,d}$, where $\alpha \in \mathbb{N}^n$; nevertheless, we commonly omit the summation index. By $\mathcal{H}_{n,d}$ we denote the space of homogeneous polynomials of degree d in $n+1$ variables.

The unit cube is $I^n := [-1,1]^n \subset \mathbb{R}^n$ and $B_{\mathbb{C}}(x,r)$ is the complex disk centered at x with radius r. The polydisc is $D^n := \overline{B}_{\mathbb{C}}(x,1)^n \subseteq \mathbb{C}^n$.

For $A \subseteq \mathbb{R}^n$, we denote by $\mathcal{B}(A)$ the set of boxes (i.e., cubes) contained in A. For any $B \in \mathcal{B}(\mathbb{R}^n)$, we denote by $m(B)$ its *midpoint* and by $w(B)$ its *width*, so that $B = m(B) + w(B)/2[-1,1]^n$.

Outline of the paper. In the next section, we outline and discuss the main results of the paper: the average run-time of the Plantinga-Vegter algorithm, and the average size of the subdivsion tree of Descartes' solver and the average run-time of the algorithm by Jindal and Sagraloff (2017). In Section 3, we introduce the norms with which we will be working and their main properties. In Section 4, we introduce a new condition number adapted to the introduced norms and we prove its main properties. In Section 5, we perform the condition-based complexity analysis of the subdivision routine of the Plantinga-Vegter algorithm; and in Section 6, we introduce the separation bound, give condition-based bounds for it and apply them to the Descartes' solver and the solver by Jindal and Sagraloff (2017). Finally, in Section 7, we introduce the randomness model that we will consider, zintzo random polynomials and p-zintzo random polynomials and provide the relevant probabilistic bounds to prove our results.

2. Overview

We present a condition-based framework that allows us to control the probability of numerical algorithms with respect to random polynomials that are sparse and do not have any scaling in their coefficients, as it has been usual with the so-called KSS or dobro random polynomials introduced in Cucker et al. (2019). We illustrate our techniques by analyzing the expected complexity of the Plantinga-Vegter algorithm and the univariate solvers Descartes and JindalSagraloff for a class of random sparse polynomials.

We will consider a very general class of random polynomials: the class of zintzo random polynomials (see Definition 7.8). Moreover, our probabilistic estimates are both in the average and smoothed paradigm (see Proposition 7.12). However, for the sake of concreteness, we will expose our results here only for Gaussian and uniform random polynomials.

Definition 2.1. Let $M \subseteq \mathbb{N}^n$ be such that it contains $0, e_1, \ldots, e_n$.

(G) A *Gaussian polynomial supported on M* is a random polynomial $\mathfrak{f} = \sum_{\alpha\in M} \mathfrak{f}_\alpha X^\alpha$ supported on M whose coefficients $\mathfrak{f}_\alpha$ are i.i.d. Gaussian random variables of mean 0 and variance 1.

(U) A *uniform random polynomial supported on M* is a random polynomial $\mathfrak{f} = \sum_{\alpha\in M} \mathfrak{f}_\alpha X^\alpha$ supported on M whose coefficients $\mathfrak{f}_\alpha$ are i.i.d. uniform random variables on $[-1, 1]$.

The condition $0, e_1, \ldots, e_n \in M$ is a technicality that we need for the proofs. In layman's terms, this technical condition states that all the terms of the first order approximation of $\mathfrak{f}$ at 0, $\mathfrak{f}_0 + \mathfrak{f}_{e_1} X_1 + \cdots + \mathfrak{f}_{e_n} X_n$, appear with probability one. When we translate this condition to an homogeneous setting, this condition would be translated into

$$M \subseteq \{(d-1)e_i + e_j \mid i, j \in \{0, \ldots, n\}\},$$

which means that the support would contain not only the vertices of the standard simplex, but also the adjacent lattice points to those vertices. We observe this sparseness condition, considered already by Renegar (1987), is a kind of pseudo-sparseness condition. Nevertheless, we note that this is an improvement over the restrictions of other existing analysis such as the one by Ergür et al. (2018).

2.1. Expected complexity of the Plantinga-Vegter algorithm

The probabilistic complexity bound for the subdivision routine of the Plantinga-Vegter algorithm, PV-Subdivsion, is the following one. We refer to Section 5 for more details on the Plantinga-Vegter algorithm.

Theorem 2.2. *Let $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a random polynomial supported on M. The average number of boxes of the final subdivision of* PV-Subdivsion *using the interval approximations* (3) *and* (4) *on input $\mathfrak{f}$ is at most*

$$2n^{\frac{3}{2}} (10(n+1))^{n+1} d^{2n} |M|^{n+2}$$

if $\mathfrak{f}$ is Gaussian, and

$$2n\, 32^{n+1} d^{2n} |M|^{n+2}$$

if $\mathfrak{f}$ is uniform.

We notice that the above theorem is a particular case of Theorem 7.19, which gives the claim for the more general class of zintzo random polynomials.

We notice that the bounds on the number of boxes are polynomial in the degree, as in (Cucker et al., 2019). This is an additional theoretical justification of the practical success of the Plantinga-Vegter algorithm. However, unlike the estimates in (Cucker et al., 2019), the bounds we present justify the success of the Plantinga-Vegter algorithm even for sparse random polynomials. This is one of the first such probabilistic complexity estimates in numerical algebraic geometry.

2.2. Complexity results on univariate solvers

In the setting of univariate solvers, we present two results. First for the Descartes solver in Sagraloff and Mehlhorn (2016) we bound the average size of the subdivision tree. Second, we bound the average bit complexity of algorithm by Jindal and Sagraloff (2017) for isolating the roots of sparse univariate polynomial. In both cases, we don't only bound the average (1st moment), but all the higher moments.

Theorem 2.3. *Let $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a random polynomial supported on M that is either Gaussian or uniform. The average size of the subdivision tree of* Descartes *on input $\mathfrak{f}$ is at most*

$$O(|M| \log d).$$

Moreover, the kth moment of the size is bounded by

$$O(k|M| \log d)^k.$$

The above result shows that Descartes' univariate solver can perform well in practice for sparse polynomials in $[-1, 1]$. It shows that in average the size of the subdivision tree will be polynomial in the size of the sparse polynomial. This provides an insight on the special character of Mignotte-like 4-mials. The above theorem is a particular case of Theorem 7.23, which states the result for zintzo random polynomials.

Theorem 2.4. *Let $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a random polynomial supported on M that is either Gaussian or uniform. The average bit-complexity of* JindalSagraloff *on input $(\mathfrak{f}, I)$ is at most*

$$O\left(|M|^{12} \log^7 d\right).$$

Moreover, the kth moment of the bit-run-time is bounded by

$$O\left(k|M|^{12} \log^7 d\right)^k.$$

This theorem provides a first step into the solution of the Rojas-Ye version of Smale's 17th problem for sparse systems (Rojas and Ye, 2005). The later theorem is a particular case of Theorem 7.24, where the claim is shown for a restricted class of zintzo random polynomials. In future work, we hope to be able to extend this analysis to polynomials distributed not only with continuous probability distributions, but with discrete probability distributions.

3. Norms for the cube and polynomials over the cube

In the traditional setting, for homogeneous polynomial $F = \sum_{|\alpha|=d} F_\alpha X^\alpha \in \mathcal{H}_{n,d}$ of degree d in $n+1$ variables, we consider the *Weyl norm*,

$$\|F\|_{\mathrm{W}} := \sqrt{\sum_\alpha \binom{d}{\alpha}^{-1} |F_\alpha|^2},$$

to control the evaluations of the polynomial, $F(p)$, and its gradient, $\nabla_p F$, at points $p \in \mathbb{S}^n$. Unfortunately, the scaling introduced by the norm in the coefficients affects the probabilistic model and forces us to consider random polynomials with a particular variance structure that excludes Gaussian polynomials.

To avoid the scaling of the coefficients, we work in the cube and we will use the ∞-norm,

$$\|x\|_\infty := \max_i |x_i|.$$

One of the main disadvantages of this norm is that it does not come from an inner product. However, we can overcome this problem as shown by Cucker et al. (2020a).

For $\mathcal{P}_{n,d}$, the space of affine polynomials of degree at most d in n variables. For a polynomial $f := \sum_{|\alpha|\leq d} f_\alpha X^\alpha \in \mathcal{P}_{n,d}$, motivated by duality, we consider the following norm

$$\|f\|_1 := \sum_\alpha |f_\alpha|. \tag{1}$$

To demonstrate that all the results generalize to the complex case we will prove the various bounds for polydics, $z \in D^n := \overline{B}_{\mathbb{C}}(0,1)^n$, which is the complex analogous of the cube.

The motivation to choose the 1-norm emanates from the following proposition which shows that we can control the evaluation of f at $x \in I^n := [-1,1]^n$, that is $f(x)$, using the 1-norm for f.

Proposition 3.1. *Let $f \in \mathcal{P}_{n,d}$ and $z \in D^n$. Then $|f(z)| \leq \|f\|_1$.*

Proof. It holds $|f(z)| = \left|\sum_\alpha f_\alpha z^\alpha\right| \leq \sum_\alpha |f_\alpha| |z|_\infty^{|\alpha|} \leq \|f\|_1$; as $z \in D^n$ implies that $\|z\|_\infty \leq 1$. □

Remark 3.2. A reader might wonder why we do not choose another norm. For example, if we choose $\|f\|_2 := \sqrt{\sum_\alpha |f_\alpha|^2}$, then we can prove that for all $z \in D^n$, it holds $|f(z)| \leq \sqrt{N}\|f\|_2$, where N is the number of terms in f. This gives worse bounds than using $\|f\|_1$ since

$$\|f\|_2 \leq \|f\|_1 \leq \sqrt{N}\|f\|_2,$$

which makes us prefer $\|f\|_1$ to $\sqrt{N}\|f\|_2$ as a bounding quantity.

Notation 3.3. Before continuing, let us clarify notations so that the statements that we consider are clear. By convention,

$$\mathrm{D}\, f := \begin{pmatrix} \frac{\partial f}{\partial X_1} & \cdots & \frac{\partial f}{\partial X_n} \end{pmatrix}$$

represents the formal tangent covector of f, whose entries are the formal partial derivatives of f. Similarly, the formal tensor of kth derivatives of f is

$$\mathrm{D}^k f := \left(\frac{\partial^k f}{\partial X_{i_1} \cdot \partial X_{i_k}} \right)_{i_1, \cdots, i_k}.$$

When we want to refer to the gradient vector, the tangent covector or the tensor of kth derivatives of f (evaluated) at a point $z \in \mathbb{R}^n$, we will use respectively $\mathrm{D}_z f$ and $\mathrm{D}_z^k f$. Thus we have that $\mathrm{D}_z f$ is a covector $\mathrm{T}_z(\mathbb{R}^n) \cong \mathbb{R}^n \to \mathbb{R}$ and that $\mathrm{D}_z^k f$ is a multilinear map $\mathrm{T}_z(\mathbb{R}^n)^k \cong (\mathbb{R}^n)^k \to \mathbb{R}$.

An important feature of the 1-norm is that for polynomials we can use it to control the 1-norm of their derivatives. In our notation, if $v_1, \ldots, v_k \in \mathbb{R}^n$ and $f \in \mathcal{P}_{n,d}$, then $\mathrm{D}^k f(v_1, \ldots, v_k)$ is a polynomial of degree $\leq d - k$ and so it makes sense to compute its 1-norm.

Proposition 3.4. *Let $f \in \mathcal{P}_{n,d}$ and $v \in \mathbb{R}^n$. Then*

$$\|\mathrm{D}\, f(v)\|_1 \leq d\|f\|_1\|v\|_\infty.$$

In particular, for all $z \in D^n$, $|\mathrm{D}_z f(v)| \leq d\|f\|_1\|v\|_\infty$.

Corollary 3.5. *Let $f \in \mathcal{P}_{n,d}$, $d \geq k \geq 0$ and $v_1, \ldots, v_k \in \mathbb{R}^n$. Then*

$$\left\| \frac{1}{k!} \mathrm{D}^k f(v_1, \ldots, v_k) \right\|_1 \leq \binom{d}{k} \|f\|_1 \|v_1\|_\infty \cdots \|v_k\|_\infty.$$

In particular, for all $z \in D^n$, $\left|\frac{1}{k!} \mathrm{D}_z^k f(v_1, \ldots, v_k)\right| \leq \binom{d}{k} \|f\|_1 \|v_1\|_\infty \cdots \|v_k\|_\infty$.

Proof of Proposition 3.4. We have

$$d\|f\|_1\|v\|_\infty = \sum_\alpha d|f_\alpha|\|v\|_\infty \text{ and } \|\mathrm{D}\, f(v)\|_1 \le \sum_\alpha |f_\alpha|\|\mathrm{D}\, X^\alpha(v)\|_1.$$

Therefore, it is enough to prove the claim for X^α. By a simple computation, we get

$$\|\mathrm{D}\, X^\alpha(v)\|_1 = \left\| \sum_{i=1}^n \alpha_i v_i X^\alpha / X_i \right\|_1 \le \sum_{i=1}^n \alpha_i |v_i| \le \|\alpha\|_1 \|v\|_\infty \le d\|v\|_\infty,$$

which is the desired inequality for the 1-norm. The last claim is Proposition 3.1. □

Proof of Corollary 3.5. By induction hypothesis, we have that

$$\begin{aligned}\left\| \frac{1}{k!} \mathrm{D}^k f(v_1, \dots, v_k) \right\|_1 &= \frac{1}{k} \left\| \frac{1}{(k-1)!} \mathrm{D}^{k-1} \left(\mathrm{D}\, f(v_k)\right)(v_1, \dots, v_{k-1}) \right\|_1 \\ &\le \frac{1}{k}\binom{d-1}{k-1} \|\mathrm{D}\, f(v_k)\|_1 \|v_1\|_\infty \cdots \|v_{k-1}\|_\infty.\end{aligned}$$

Proposition 3.4 finishes the induction step and provides the base case for induction. The last claim is again Proposition 3.1. □

The bounds on the derivatives allows us to bound the (Lipschitz constants) of the variations of a polynomial f and all its derivatives inside D^n.

Proposition 3.6. *Let $f \in \mathcal{P}_{n,d}$, $d \ge k \ge 0$ and $v_1, \dots, v_k \in \mathbb{R}^n$ such that $\|v_i\|_\infty = 1$. Then the map*

$$\begin{aligned} D^n &\to [-1, 1]^n \\ z &\mapsto \frac{1}{d^k \|f\|_1} \mathrm{D}^k_z f(v_1, \dots, v_k) \end{aligned}$$

is well-defined and $(d-k)$-Lipschitz with respect the ∞-norm.

Proof. Without loss of generality, assume that $\|f\|_1 = 1$. Let $x, y \in D^n$. By the fundamental theorem of calculus,

$$\mathrm{D}^k_y f(v_1, \dots, v_k) - \mathrm{D}^k_x f(v_1, \dots, v_k) = \int_0^1 \mathrm{D}^k_{ty+(1-t)x} f(v_1, \dots, v_k, y - x)\, dt.$$

Hence, taking absolute values and using Corollary 3.5, we get that

$$\left| \frac{1}{d^k} \mathrm{D}^k_y f(v_1, \dots, v_k) - \frac{1}{d^k} \mathrm{D}^k_x f(v_1, \dots, v_k) \right| \le \frac{d!}{d^k (d-k-1)!} \|y - x\|_\infty \le (d-k)\|x - y\|_\infty,$$

which gives the Lipschitz property. The choice of the co-domain follows from Corollary 3.5. □

Recall that for a multilinear map $A : (\mathbb{R}^n)^k \to \mathbb{R}$ we can consider the induced norm

$$\|A\|_{\infty,\infty} := \sup_{v_1,\dots,v_k \neq 0} \frac{|A(v_1, \dots, v_k)|}{\|v_1\|_\infty \cdots \|v_k\|_\infty},$$

instead of the 1-norm

$$\|A\|_1 := \sum_{i_1,\dots,i_k} |A_{i_1,\dots,i_k}|.$$

Although $\|A\|_{\infty,\infty} \le \|A\|_1$ is not an equality in general, it is so in the case that A is a linear map, which allows us to deduce the following.

Proposition 3.7. *Let $f \in \mathcal{P}_{n,d}$ and $d \geq k \geq 0$. Then the map*

$$\begin{aligned} D^n &\to [0,1]^n \\ z &\mapsto \frac{1}{d^k \|f\|_1} \|\mathrm{D}_z^k f\|_{\infty,\infty} \end{aligned}$$

is well-defined and $(d-k)$-Lipschitz with respect the ∞-norm.

Proof. Without loss of genreality, assume that $\|f\|_1 = 1$. By Proposition 3.6, we have that for all $x, y \in D^n$ and all $v_1, \ldots, v_k \in \mathbb{R}^n$ such taht $\|v_i\|_\infty = 1$,

$$\left| \left(\frac{1}{d^k} \mathrm{D}_y^k f - \frac{1}{d^k} \mathrm{D}_x^k f \right)(v_1, \ldots, v_k) \right| \leq (d-k)\|y - x\|_\infty.$$

By maximixing the left-hand side with respect $v_1, \ldots, v_k$, we get that for all $x, y \in D^n$,

$$\left\| \frac{1}{d^k} \mathrm{D}_y^k f - \frac{1}{d^k} \mathrm{D}_x^k f \right\|_{\infty,\infty} \leq (d-k)\|y - x\|_\infty.$$

This gives the Lipzchitz property. The choice of codomain is justified in a similar way. □

Corollary 3.8. *Let $f \in \mathcal{P}_{n,d}$. Then the maps*

$$\begin{aligned} D^n &\to [0,1] \\ z &\mapsto \frac{1}{\|f\|_1}|f(z)| \end{aligned} \qquad \textit{and} \qquad \begin{aligned} D^n &\to [0,1] \\ z &\mapsto \frac{1}{d\|f\|_1}\|\mathrm{D}_z f\|_1 \end{aligned}$$

are well-defined and d-Lipschitz with respect the ∞-norm.

Proof. Just note that $\|\mathrm{D}_z f\|_{\infty,\infty} = \|\mathrm{D}_z f\|_1$. The rest is straightforward from Proposition 3.7. □

We finish with a slightly stronger version of some of the above results that will be useful later. When we are outside the polydisk D^n, then the non-linear factors of the polynomials dominate. However, we can retain control around D^n if we are not too far.

Proposition 3.9. *Let $f \in \mathcal{P}_{n,d}$, $\varepsilon > 0$ and $D_\varepsilon^n := \overline{B}_{\mathbb{C}}(0, 1+\varepsilon)^n$. If $\varepsilon \leq \frac{1}{d}$, then:*

1. *For all $z \in D_\varepsilon^n$, $k \geq 0$, and all $v_1, \ldots, v_k \in \mathbb{R}^n$,*

$$\left| \frac{1}{k!} \mathrm{D}_z^k f(v_1, \ldots, v_k) \right| \leq \mathrm{e}\binom{d}{k} \|f\|_1 \|v_1\|_\infty \cdots \|v_k\|_\infty.$$

2. *The maps*

$$\begin{aligned} D_\varepsilon^n &\to [0, \mathrm{e}] \\ z &\mapsto \frac{1}{\|f\|_1}|f(z)| \end{aligned} \qquad \textit{and} \qquad \begin{aligned} D_\varepsilon^n &\to [0, \mathrm{e}] \\ z &\mapsto \frac{1}{d\|f\|_1}\|\mathrm{D}_z f\|_1 \end{aligned}$$

are well-defined and $\mathrm{e}d$-Lipschitz with respect the ∞-norm.

Proof. We consider the polynomial

$$g := f\left((1+\varepsilon)X\right).$$

We can see that

$$\|g\|_1 \leq e\|f\|_1,$$

since for $k \in \{0, 1, \ldots, d\}$,

$$(1+\varepsilon)^k \leq \left(1+\frac{1}{d}\right)^d \leq \mathrm{e}.$$

Moreover, for $z \in D_\varepsilon^n$,

$$\mathrm{D}_z^k f = \frac{1}{(1+\varepsilon)^k}\,\mathrm{D}_{z/(1+\varepsilon)}^k g,$$

and so, by Corollary 3.5,

$$\left|\frac{1}{k!}\,\mathrm{D}_z^k f\right| \leq \frac{1}{(1+\varepsilon)^k}\binom{d}{k}\|g\|_1 \leq \mathrm{e}\binom{d}{k}\|f\|_1.$$

This proves 1.

Point 2 follows from point 1 in the same way we prove Corollary 3.8 from Corollary 3.5 (after passing through Propositions 3.6 and 3.7). □

4. Condition and its properties

In this section, we define and study the properties of the condition number. The following definition adapts the real local condition number (Bürgisser and Cucker, 2013, Chapter 19) (cf. Cucker et al. (2018)) to our setting.

Definition 4.1. Let $f \in \mathcal{P}_{n,d}$ and $x \in I^n$, the *local condition number of f at x* is the quantity

$$\mathtt{C}_1(f,x) := \frac{\|f\|_1}{\max\left\{|f(x)|, \frac{1}{d}\|\,\mathrm{D}_x f\|_1\right\}}.$$

The intuition behind this condition number is simple. It holds $\mathtt{C}_1(f,x) = \infty$ if and only if x is a singular zero of f. Thus $\mathtt{C}_1(f,x)$ measures how close is f to have a singularity at x. An important point to note is that, since we will be dealing with problems from real algebraic geometry, we do not only have to guarantee that zeros are smooth, but also that they exist. The latter is the reason why the term $|f(x)|$ appears in the denominator and it is fundamental in numerical real algebraic geometry, where a perturbation of a polynomial not only perturbs the zeros, but also can make them disappear.

In (Tonelli-Cueto, 2019, $2^{\S 2}$) a series of explicit properties of the condition number are underlined as the important properties to carry out a condition-based complexity analysis. These properties are: the regularity inequality, the 1st and the 2nd Lipschitz property, and the higher derivative estimate. The following theorem shows that the condition number in (4.1) has these properties. We recall that Smale's gamma, γ, is the invariant[2] given by

$$\gamma(f,z) := \sup_{k\geq 2}\left(\frac{1}{\|\,\mathrm{D}_z f\|_2}\left\|\frac{1}{k!}\,\mathrm{D}_z^k f\right\|_{2,2}\right)^{\frac{1}{k-1}}, \tag{2}$$

[2]The formula looks different from the usual one because it is simplified for the case of one multivariate polynomial.

for a polynomial $f \in \mathcal{P}_{n,d}$ and $z \in \mathbb{C}^n$, where $\|\cdot\|_{2,2}$ is the induced norm for multilinear maps for the usual Euclidean norm.

Theorem 4.2. *Let $f \in \mathcal{P}_{n,d}$ and $x \in I^n$. Then:*

- ***Regularity inequality**: Either*

$$\frac{|f(x)|}{\|f\|_1} \geq \frac{1}{\mathcal{C}_1(f,x)} \quad \text{or} \quad \frac{\|\mathrm{D}_x f\|_1}{d\|f\|_1} \geq \frac{1}{\mathcal{C}_1(f,x)}.$$

  In particular, if $\mathcal{C}_1(f,x)\frac{|f(x)|}{\|f\|_1} < 1$, then $\mathrm{D}_x f \neq 0$.

- ***1st Lipschitz inequality**: The map*

$$\begin{aligned} \mathcal{P}_{n,d} &\to [0,\infty) \\ g &\mapsto \frac{\|g\|_1}{\mathcal{C}_1(g,x)} \end{aligned}$$

  is 1-Lipschitz with respect the 1-norm. In particular, $\mathcal{C}_1(f,x) \geq 1$.

- ***2nd Lipschitz inequality**: The map*

$$\begin{aligned} I^n &\to [0,1] \\ y &\mapsto \frac{1}{\mathcal{C}_1(f,y)} \end{aligned}$$

  is d-Lipschitz with respect the ∞-norm.

- ***Higher derivative estimate**: If $\mathcal{C}_1(f,x)\frac{|f(x)|}{\|f\|_1} < 1$, then*

$$\gamma(f,x) \leq \frac{\sqrt{n}(d-1)}{2}\,\mathcal{C}_1(f,x).$$

Remark 4.3. We note that the theorem still holds if we replace I^n by D^n.

Before continuing with the proof of Theorem 4.2, let's discuss why these properties are important for us.

- The regularity inequality tells us (in a quantitative way depending on the condition number) that either the value of f at a point x is big or that the gradient of f at that point x is big. In this way, the regularity inequality guarantees us that the covector field $x \mapsto \mathrm{D}_x f$ does not vanish near the zero set of f. The latter allows us to guarantee that the Newton operator is well-defined or that geometric arguments based on following the gradient flow work near the zero set.

- The 1st and 2nd Lipschitz properties allow us to guarantee that $\mathtt{C}_1(f,x)$ can be numerically evaluated at (f,x), since it guarantees us that $\mathtt{C}_1(f,x)$ can be bounded by $\mathtt{C}_1(\tilde{f},\tilde{x})$ for sufficiently good approximations $\tilde{f}$ of f and $\tilde{x}$ of x. Making this concrete, we have that

$$\mathtt{C}_1(f,x) \leq \frac{\mathtt{C}_1\left(\tilde{f},\tilde{x}\right)}{1 - \mathtt{C}_1\left(\tilde{f},\tilde{x}\right)\left(2\frac{\|\tilde{f}-f\|_1}{\|\tilde{f}\|_1} + \|\tilde{x}-x\|_\infty\right)} \leq \mathtt{C}_1\left(\tilde{f},\tilde{x}\right)(1+\delta),$$

  whenever $2\frac{\|\tilde{f}-f\|_1}{\|\tilde{f}\|_1} + \|\tilde{x}-x\|_\infty < \frac{1}{\mathtt{C}_1(\tilde{f},\tilde{x})}\delta$ for some $\delta \in (0,1)$.

- The higher derivative estimate allows us to control how the Newton method converges near the zero set. This is based on Smale's α-theory, which can be found in (Bürgisser and Cucker, 2013, 15.2) and (Dedieu, 2006, Chapter 4) among many other references.

Proof of Theorem 4.2. For the regularity inequality, note that $\frac{1}{\mathsf{C}_1(f,x)}$ is the maximum of $\frac{|f(x)|}{\|f\|_1}$ and $\frac{\|\mathrm{D}_x f\|_1}{d\|f\|_1}$. Note that we don't only get an inequality, but an equality, but this is not important for later arguments.

For the 1st Lipschitz property, let $g_0, g_1 \in \mathcal{P}_{n,d}$. Then:

$$\begin{aligned}
&\left|\frac{\|g_0\|_1}{\mathsf{C}_1(g_0,x)} - \frac{\|g_1\|_1}{\mathsf{C}_1(g_1,x)}\right| \\
&\quad= \left|\max\left\{|g_0(x)|, \frac{1}{d}\|\mathrm{D}_x g_0\|_1\right\} - \max\left\{|g_1(x)|, \frac{1}{d}\|\mathrm{D}_x g_1\|_1\right\}\right| && \text{(Definition 4.1} \\
&\quad\le \max\left\{|g_0(x) - g_1(x)|, \frac{1}{d}\|\mathrm{D}_x g_0 - \mathrm{D}_x g_1\|_1\right\} && \text{(Triangle inequality)} \\
&\quad\le \max\left\{|(g_0 - g_1)(x)|, \frac{1}{d}\|\mathrm{D}_x(g_0 - g_1)\|_1\right\} \\
&\quad\le \|g_0 - g_1\|_1 && \text{(Corollary 3.8)}
\end{aligned}$$

Note that we use that the codomain is $[0,1]$ in Corollary 3.8. For the later inequality, note that $\frac{\|0\|_1}{\mathsf{C}_1(0,x)} = 1$ (or simply use Corollary 3.8 again).

For the 2nd Lipschitz property, the argument is similar to the one above. Without loss of generality, assume that $\|f\|_1 = 1$. Let $y_0, y_1 \in I^n$, then:

$$\begin{aligned}
&\left|\frac{1}{\mathsf{C}_1(f,y_0)} - \frac{1}{\mathsf{C}_1(f,y_1)}\right| \\
&\quad= \left|\max\left\{\frac{|f(y_0)|}{\|f\|_1}, \frac{\|\mathrm{D}_{y_0} f\|_1}{d\|f\|_1}\right\} - \max\left\{\frac{|f(y_1)|}{\|f\|_1}, \frac{\|\mathrm{D}_{y_1} f\|_1}{d\|f\|_1}\right\}\right| && \text{(Definition 4.1} \\
&\quad\le \max\left\{\left|\frac{|f(y_0)|}{\|f\|_1} - \frac{|f(y_1)|}{\|f\|_1}\right|, \left|\frac{\|\mathrm{D}_{y_0} f\|_1}{d\|f\|_1} - \frac{\|\mathrm{D}_{y_1} f\|_1}{d\|f\|_1}\right|\right\} && \text{(Triangle inequality)} \\
&\quad\le \|y_0 - y_1\|_\infty && \text{(Corollary 3.8)}
\end{aligned}$$

For the higher derivative estimate, note that for a multilinear map $A : (\mathbb{R}^n)^k \to \mathbb{R}$,

$$\|A\|_{2,2} = \sup_{x_1,\dots,x_k \neq 0} \frac{|A(x_1,\dots,x_k)|}{\|x_1\|_2 \cdots \|x_k\|_2} \le \sup_{x_1,\dots,x_k \neq 0} \frac{|A(x_1,\dots,x_k)|}{\|x_1\|_\infty \cdots \|x_k\|_\infty} = \|A\|_{\infty,\infty},$$

since $\|z\|_2 \ge \|z\|_\infty$ for all z, and that for a linear map $a : \mathbb{R}^n \to \mathbb{R}$,

$$\|a\|_2 \ge \frac{\|a\|_1}{\sqrt{n}}.$$

In this way, for $k \ge 2$,

$$\frac{1}{\|\mathrm{D}_x f\|_2}\left\|\frac{1}{k!}\mathrm{D}_x^k f\right\|_{2,2} \le \frac{\sqrt{n}}{\|\mathrm{D}_x f\|_1}\left\|\frac{1}{k!}\mathrm{D}_x^k f\right\|_{\infty,\infty} = \frac{\sqrt{n}\|f\|_1}{\|\mathrm{D}_x f\|_1}\frac{\left\|\frac{1}{k!}\mathrm{D}_x^k f\right\|_{\infty,\infty}}{\|f\|_1}.$$

By the above, the regularity inequality and Corollary 3.5, the above becomes

$$\frac{1}{\|\mathrm{D}_x f\|_2}\left\|\frac{1}{k!}\mathrm{D}^k_x f\right\|_{2,2} \leq \frac{\sqrt{n}}{d\,\mathtt{C}_1(f,x)}\binom{d}{k} = \frac{\sqrt{n}}{d}\binom{d}{k}\frac{1}{\mathtt{C}_1(f,x)}.$$

Now, for $k \geq 2$, $\left|\frac{1}{\mathtt{C}_1(f,x)}\right|^{\frac{1}{k-1}} \leq \frac{1}{\mathtt{C}_1(f,x)}$, since $\mathtt{C}_1(f,x) \geq 1$, and

$$\left|\frac{\sqrt{n}}{d}\binom{d}{k}\right|^{\frac{1}{k-1}} \leq \sqrt{n}\frac{((d-1)\cdots(d-k+1))^{\frac{1}{k-1}}}{(k!)^{\frac{1}{k-1}}} \leq \frac{\sqrt{n}(d-1)}{2},$$

since $(k!)^{\frac{1}{k-1}} \geq 2$. This finishes the proof. □

Using the above theorem, we can provide a sort of geometric interpretation of the condition number above. Fix $x \in I^n$ and consider

$$\Sigma_x := \{g \in \mathcal{P}_{n,d} \mid g(x) = 0,\ \nabla_x g = 0\} \subset \mathcal{P}_{n,d},$$

which is the subset of polynomials that are singular at 0. The following proposition, usually referred as 'Condition Number Theorem' relates the distance of a polynomial to Σ_x to the defined condition number. Although the version we provide does not give an equality, it provides a bound in both relations.

Proposition 4.4 (Condition Number Theorem). *For all $f \in \mathcal{P}_{n,d}$ and $x \in I^n$,*

$$\frac{\|f\|_1}{\mathrm{dist}_1(f,\Sigma_x)} \leq \mathtt{C}_1(f,x) \leq (1+d)\frac{\|f\|_1}{\mathrm{dist}_1(f,\Sigma_x)},$$

where dist_1 *is the distance induced by the* 1*-norm.*

Proof. The left hand side follows from the 1st Lipschitz property (Theorem 4.2), since for $g \in \Sigma_x$,

$$\frac{\|f\|_1}{\mathtt{C}_1(f,x)} = \left|\frac{\|f\|_1}{\mathtt{C}_1(f,x)} - \frac{\|g\|_1}{\mathtt{C}_1(g,x)}\right| \leq \|f-g\|_1.$$

Thus, minimizing for $g \in \Sigma_x$,

$$\frac{\|f\|_1}{\mathtt{C}_1(f,x)} \leq \mathrm{dist}_1(f,\Sigma_x).$$

For the right hand side, consider the polynomial

$$g := f - f(x) - \sum_{i=1}^{n} \partial_i f(x) X_i.$$

Then $g \in \Sigma_x$ and $\|f-g\|_1 \leq |f(x)| + \|\mathrm{D}_x f\|_1$. Hence

$$\mathrm{dist}_1(f,\Sigma_x) \leq \|f-g\|_1 \leq (1+d)\max\left\{|f(x)|, \frac{1}{d}\|\mathrm{D}_x f\|_1\right\} = (1+d)\frac{\|f\|_1}{\mathtt{C}_1(f,x)},$$

as desired. □

We conclude this section, introducing the global condition number and stating its properties.

Definition 4.5. Let $f \in \mathcal{P}_{n,d}$, the *global condition number of* f is the quantity

$$\mathtt{C}_1(f) := \max\{\mathtt{C}_1(f, x) \mid x \in I^n\}.$$

Notice that $\mathtt{C}_1(f)$ is infinity if and only if the zero set of f has a singularity in I^n. The following proposition quantifies this fact, which interprets geometrically f. We denote by

$$\Sigma_{n,d} := \{g \in \mathcal{P}_{n,d} \mid \text{for some } x \in I^n,\ g(x) = 0,\ \mathrm{D}_x\, g = 0\} = \bigcup_{x \in I^n} \Sigma_x \subset \mathcal{P}_{n,d}$$

the set of polynomials whose zero sets have a singularity in I^n.

Proposition 4.6. *The map*

$$f \to \frac{\|f\|_1}{C_1(f)}$$

is 1*-Lipschitz. Moreover, for every* $f \in \mathcal{P}_{n,d}$,

$$\frac{\|f\|_1}{\mathrm{dist}_1(f, \Sigma_{n,d})} \leq C_1(f) \leq (1 + d)\frac{\|f\|_1}{\mathrm{dist}_1(f, \Sigma_{n,d})},$$

where dist_1 *is the distance induced by the* 1*-norm.*

Proof. This follows immediately from the 1st Lipschitz property, Theorem 4.2 and the Condition Number Theorem (Proposition 4.4). □

5. Plantinga-Vegter Algorithm and its complexity

The Plantinga-Vegter algorithm (Plantinga and Vegter, 2004) is a subdivision-based algorithm that computes an isotopically correct approximation of the zeros of a univariate polynomial in an interval, of a curve in the plane, or of a surface in the 3-dimensional space. This algorithm can be generalized to hypersurfaces of arbitrary dimension as shown by Galehouse (2009) and to singular curves as shown by Burr et al. (2012).

Following Burr et al. (2017) and Cucker et al. (2020b) (cf. Cucker et al. (2019)), we focus on the subdivision procedure. To analyze the complexity, we notice that three levels that we focus on (following Xu and Yap (2019) (cf. Yap (2019))):

A) Abstract level: Evaluations are modelled with exact arithmetic.

I) Interval level: Evaluations are modelled with exact interval arithmetic.

E) Effective level: Evaluations are modelled with finite precision interval arithmetic.

Cucker et al. (2020b) did a complexity analysis of the Plantinga-Vegter algorithm for all three levels. However, our objective is not to reproduce the analysis in (Cucker et al., 2020b) to the last detail, but to show that changing from the Weyl norm to the 1-norm improves the complexity. For this reason, we only focus on the interval level. We do this, because if we reproduce all the arguments for the effective level, then the analysis would not only be technically tedious, but will also distract us from highlighting the complexity improvement.

While analyzing the interval level, we estimate the number of boxes of the final subdivision. This is our measure of complexity. We refer to Burr et al. (2017), Cucker et al. (2020b) and (Tonelli-Cueto, 2019, $5^{\S 2}$) for further justifications of this approach.

Algorithm 1: PV-Subdivsion

Input : $f \in \mathcal{P}_{n,d}$ which is non-singular in I^n
Output : A subdivision $\mathcal{S}$ of I^n into boxes such that for all $B \in \mathcal{S}$, $C_f(B)$ holds

$\mathcal{S}_0 \leftarrow \{I^n\}$, $\mathcal{S} \leftarrow \varnothing$;
while $\mathcal{S}_0 \neq \varnothing$ **do**
  Take $B \in \mathcal{S}_0$;
  if $C_f(B)$ *holds* **then**
    $\mathcal{S} \leftarrow \mathcal{S} \cup \{B\}$, $\mathcal{S}_0 \leftarrow \mathcal{S}_0 \setminus \{B\}$;
  else
    $\mathcal{S}_0 \leftarrow \mathcal{S}_0 \setminus \{B\} \cup$ StandardSubdivsion(B);
RETURN $\mathcal{S}$;

5.1. Interval version of the PV Algorithm

The subdivision routine of the PV algorithm, PV-Subdivsion, relies on subdividing the unit cube I^n until each box B in the subdivision satisfies the condition

$$C_f(B): \text{ either } 0 \notin f(B) \text{ or } 0 \notin \left\{ \mathrm{D}_x f \, \mathrm{D}_y f^{\mathrm{T}} \mid x, y \in B \right\}.$$

Note that $\mathrm{D}_x f \, \mathrm{D}_y f^{\mathrm{T}} = \sum_{i=1}^n \partial_i f(x) \partial_i f(y)$, as $\mathrm{D}_z f$ is a covector (and so a row-vector).

To implement this algorithm one uses interval arithmetic. Recall that an *interval approximation* of a map $g : I^n \to \mathbb{R}^q$ is a map $\square[g] : \square[I^n] \to \square[\mathbb{R}^q]$, where $\square[X]$ is the set of (coordinate) boxes contained in X, such that for all $B \in \square[I^n]$, we have

$$g(B) \subseteq \square[g](B).$$

The following proposition provides interval approximations for f and $\|\mathrm{D} f\|_1$.

Proposition 5.1. *Let $f \in \mathcal{P}_{n,d}$. Then*

$$\square[f](B) := f(m(B)) + d\|f\|_1 \frac{w(B)}{2}[-1, 1] \tag{3}$$

and

$$\square[\|\mathrm{D} f\|_1](B) := \|\mathrm{D}_{m(B)} f\|_1 + \sqrt{2n} d^2 \|f\|_1 \frac{w(B)}{2}[-1, 1]. \tag{4}$$

Proof. We only need to show, respectively, that $f(B) \subseteq f(m(B)) + d\|f\|_1 \frac{w(B)}{2}[-1, 1]$ and that $\|\mathrm{D} f\|_1(B) \subseteq \|\nabla_{m(B)} f\|_1 + d^2 \|f\|_1 \frac{w(B)}{2}[-1, 1]$. However, this follows from Corollary 3.8. □

We now show how to test $C_f(B)$ using the above interval approximations. The reason we have the factor $\sqrt{n}$ in (4) is so that the next proposition provides a nicer statement.

Proposition 5.2. *The condition $C_f(B)$ is implied by the condition*

$$C'_f(B): \ |f(m(B))| > d\|f\|_1 \frac{w(B)}{2} \ \text{ or } \ \|\mathrm{D}_{m(B)} f\|_1 > d^2 \sqrt{2n} \|f\|_1 \frac{w(B)}{2}$$

Hence, PV-Subdivsion *with the interval approximations given in* (3) *and* (4) *is correct if we substitute the condition* $C_f(B)$ *by*

$$C_f^{\square}(B): \text{ either } 0 \notin \square[f](B) \text{ or } 0 \notin \square[\|\nabla f\|_1](B).$$

Proof. On the one hand, by Corollary 3.8, the map $|f|$ is $d\|f\|_1$-Lipschitz. Thus $|f(m(B))| > d\|f\|_1 \frac{w(B)}{2}$ implies that for all $x \in B$, $|f(x)| \geq |f(m(B))| - d\|f\|_1 \frac{w(B)}{2}$. This is the first clause of $C_f(B)$.

On the other hand, by (Cucker et al., 2019, Lemma 4.4), we have that if for all $x \in B$, $\|\mathrm{D}_x f - \mathrm{D}_{m(B)} f\|_2 \leq \frac{1}{\sqrt{2}}\|\mathrm{D}_{m(B)} f\|$, then for all $x, y \in B$, $\mathrm{D}_x f \, \mathrm{D}_y f^{\mathrm{T}} \neq 0$, which is the second clause of $C_f(B)$. For $x \in B$,

$$\|\mathrm{D}_x f - \mathrm{D}_{m(B)} f\|_2 \leq \|\mathrm{D}_x f - \mathrm{D}_{m(B)} f\|_1 \leq d^2\|f\|_1 \frac{w(B)}{2},$$

due to Corollary 3.8. Hence $d^2\|f\|_1 \frac{w(B)}{2} \leq \frac{1}{\sqrt{2}}\|\mathrm{D}_{m(B)} f\|_2$ implies the second clause of $C_f(B)$, and so does $\|\mathrm{D}_{m(B)} f\|_1 \geq d^2\sqrt{2n}\|f\|_1 w(B)$, since $\|y\|_1 \leq \sqrt{n}\|y\|_2$.

The two paragraphs above together give the desired claim. □

In what follows the interval version of PV-Subdivsion will be a variant that exploits the interval approximations in (3) and (4).

5.2. Complexity analysis of the interval version

As in Burr et al. (2017) and Cucker et al. (2019), our complexity analysis relies on the construction of a local size bound for PV-Subdivsion and the application of the continuous amortization developed by Burr et al. (2009); Burr (2016).

We recall the definition of the local size bound and the result that we will exploit in our complexity analysis.

Definition 5.3. A *local size bound* for the interval version of PV-Subdivsion on input f is a function $b_f : I^n \to [0, 1]$ such that for all $x \in \mathbb{R}^n$,

$$b_f(x) \leq \inf\{\mathrm{vol}(B) \mid x \in B \in \mathcal{B}(I^n) \text{ and } C_f^{\square}(B) \text{ false}\}.$$

Theorem 5.4. *(Burr et al., 2009; Burr, 2016; Burr et al., 2017) The number of boxes of the final subdivision of the interval version of* PV-Subdivsion *on input* f *is at most*

$$4^n \mathbb{E}_{\mathfrak{x} \in I^n} \frac{1}{b_f(\mathfrak{x})}.$$

In addition, the bound is finite if and only if PV-Subdivsion *terminates.* □

Theorem 5.5. *The function*

$$x \mapsto \left(d\sqrt{2n}\, C_1(f, x)\right)^{-n}$$

is a local size bound for PV-Subdivsion *on input* f.

Proof. Without loss of generality, assume that $\|f\|_1 = 1$. Let $x \in B \in \mathcal{B}(I^n)$. Then $\|m(B) - x\|_\infty \leq w(B)/2$ and so, by Corollary 3.8 and the regularity inequality (Theorem 4.2), we have that either

$$|f(m(B))| > \frac{1}{\mathtt{C}_1(f,x)} - d\frac{w(B)}{2} \tag{5}$$

or

$$\|\,\mathrm{D}_{m(B)}\, f\|_1 > d\frac{1}{\mathtt{C}_1(f,x)} - d^2\frac{w(B)}{2}. \tag{6}$$

Hence, $C_f^{\square}(B)$ is true as long as either $\mathtt{C}_1(f,x)^{-1} \geq dw(B)$, or $\mathtt{C}_1(f,x)^{-1} > d\sqrt{2n}w(B)$. The result follows, since $\mathrm{vol}(B) = w(B)^n$. □

Theorem 5.4 and Theorem 5.5 result the following corollary, which is the preamble of one of our results.

Corollary 5.6. *The number of boxes of the final subdivision of the interval version of* PV-Subdivsion *on input f is at most*

$$2^{\frac{5}{2}n} n^{\frac{n}{2}} d^n \mathbb{E}_{x\in I^n}\, \mathtt{C}_1(f,\mathfrak{x})^n.$$

Theorem 7.19 follows now from the corollary above and the following proposition.

Remark 5.7. A similar argument as in the proof of (Cucker et al., 2019, Theorem 6.4) shows that we can bound the local size bound of Burr et al. (2017) in terms of $1/\,\mathtt{C}_1(f,x)^n$. Since the interval approximation of the analyzed version is simpler, requiring a single evaluation, we only analyze the complexity of this.

Remark 5.8. We note that previous bound can be generalized for a cube $[-a,a]^n$ bigger than the unit cube. A straightforward argument might try to consider the polynomial $f(aX)$ which considered on I^n behaves like f inside $[-a,a]^n$. Unfortunately,

$$\|f(aX)\|_1 \leq a^d \|f\|_1,$$

which will complicated things as the bounds would become exponential in the degree. To avoid this, one should reprove Corollary 3.8. The trick is to consider the maps

$$x \mapsto \frac{|f(x)|}{d\|f\|_1 \max\left\{1, \|x\|_\infty^d\right\}} \text{ and } x \mapsto \frac{|f(x)|}{d^2\|f\|_1 \max\left\{1, \|x\|_\infty^{d-1}\right\}}$$

and prove that they are Lipschitz. Let us demonstrate this approach for the first map. We only have to consider the map as the composition of

$$\partial I^n \ni \begin{pmatrix} x_0 \\ x \end{pmatrix} \mapsto \frac{x_0^d}{d\|f\|_1} f\left(\frac{x}{x_0}\right)$$

together with

$$\mathbb{R}^{1+n} \ni x \mapsto \frac{1}{\max\{1, \|x\|_\infty\}} \begin{pmatrix} 1 \\ x \end{pmatrix} \in \partial I^n,$$

and since each map is Lipschitz, this is also the case for their composition.

6. Condition, separation bounds and univariate solvers

In this section we turn our attention to the separation of the roots of a real univariate polynomial. In general, we are interested in the separation between the real roots and the separation between the conjugate complex roots, as these affect the complexity of some univariate solvers. As we will focus on searching roots in I, we will consider the following separation quantities.

Definition 6.1. Let $f \in \mathcal{P}_{1,d}$. Then we define:

(R) The *real separation of* f, $\Delta(f)$, is given by

$$\Delta(f) := \min\left\{\left|\zeta - \tilde{\zeta}\right| \mid \zeta, \tilde{\zeta} \in I,\ f(\zeta) = f\left(\tilde{\zeta}\right) = 0\right\},$$

if f has no double roots in I, and it is zero otherwise.

(C) Let $\varepsilon \in \left(0, \frac{1}{d}\right)$. The *$\varepsilon$-real separation of* f, $\Delta_\varepsilon(f)$, is

$$\Delta_\varepsilon(f) := \min\left\{\left|\zeta - \tilde{\zeta}\right| \mid \zeta, \tilde{\zeta} \in I_\varepsilon := \{z \in \mathbb{Z} \mid \operatorname{dist}(z, I) \leq \varepsilon\},\ f(\zeta) = f\left(\tilde{\zeta}\right) = 0\right\},$$

if f has no double roots in I and it is zero otherwise.

We observe that $\Delta(f)$ gives the minimum distance between two roots of f in I, while $\Delta_\varepsilon(f)$ takes into account how near to the real line complex roots near I are. The quantity $\Delta(f)$ plays a role controlling the complexity of univariate solvers that do not depend on the complex roots, such as the Sturm's solver, while $\Delta_\varepsilon(f)$ plays a role controlling the complexity of of univariate solvers that depend on the complex roots near the real line, such as the Descartes' solver.

Our objective is to give lower bounds on these two quantities in terms of the condition number and use them for analysing two univariate solvers: Descartes and JindalSagraloff.

6.1. Condition-based bounds for separation

For bounding the real separation, we follow the ideas of Raffalli (2014) which allow us to obtain a bound depending on the square root of the global condition number. The main idea is to exploit that between two consecutive roots there is a point where the derivative vanishes and so a point where the Taylor expansion becomes quadratic.

Theorem 6.2. *Let $f \in \mathcal{P}_{1,d}$. Then*

$$\Delta(f) \geq \frac{2\sqrt{2}}{d\sqrt{\mathcal{C}_1(f)}}.$$

Proof. Let $\zeta, \tilde{\zeta} \in I$ be the pair of real roots of f in I with the minimum distance. By Rolle's theorem, there is $x_0 \in I$ between these roots such that

$$f'(x_0) = 0.$$

Without loss of generality assume that ζ is the roots closest to x_0, so that $\left|\zeta - \tilde{\zeta}\right|/2 \geq |x_0 - \zeta|$. Then by Taylor's theorem,

$$0 = f(\zeta) = f(x_0) + \frac{1}{2} f''(x)(\zeta - x_0)^2,$$

for some $x \in I$ between ζ and x_0. Hence

$$|f(x_0)| = \frac{1}{2}|f''(x)||\zeta - x_0|^2 \leq \frac{\left|\zeta - \tilde{\zeta}\right|^2}{8}|f''(x)| = \frac{(\Delta(f))^2}{8}|f''(x)|.$$

Since $f'(x_0) = 0$, we obtain the desired result dividing by $\|f\|_1$ and applying Corollary 3.5. □

For the ε-real separation bound, our results are based on (Dedieu, 1997, Theorem 3.2 and Theorem 5.1). The main idea is to use the higher derivative estimate and the fact that the inverse of the Smale's γ is Lipschitz.

Theorem 6.3. *Let $f \in \mathcal{P}_{1,d}$. Then for all $\varepsilon \in \left(0, \frac{1}{\mathrm{e}d\,\mathtt{C}_1(f)}\right)$,*

$$\Delta_\varepsilon(f) \geq \frac{1}{12d\,\mathtt{C}_1(f)}.$$

Proof. Let $\zeta \in I_\varepsilon$ be a complex root. By (Dedieu, 2006, Théorème 91), the Newton method converges for any point in $B_{\mathbb{C}}(\zeta, 1/(6\gamma(f,\zeta))$, where γ is Smale's gamma. Hence, taking $\zeta \in I_\varepsilon$ maximizing $\gamma(f,\zeta)$, we have that

$$\frac{1}{3\gamma(f,\zeta)} \leq \Delta_\varepsilon(f),$$

since $\frac{1}{3\gamma(f,\zeta)}$ is the distance from ζ to the rest of the roots of f. By (Dedieu, 2006, Lemme 98),

$$\gamma(f,\zeta) \leq \frac{\gamma(f,x)}{(1-\gamma(f,x)\varepsilon)(1-4\gamma(f,x)\varepsilon + 2\gamma(f,x)^2\varepsilon^2)}$$

for some $x \in I$ such that $|\zeta - x| \leq \varepsilon$.

By Proposition 3.9, we have that

$$\frac{|f(x)|}{\|f\|_1} \leq \mathrm{e}d\varepsilon \leq \frac{1}{\mathtt{C}_1(f,x)},$$

where the last inequality is by assumption. Hence, by the above inequality and the higher derivative estimate (Theorem 4.2),

$$\gamma(f,\zeta) \leq 4(d-1)\,\mathtt{C}_1(f,x),$$

since $\gamma(f,x)\varepsilon \leq \frac{1}{2\mathrm{e}}$. Hence

$$\frac{1}{3\gamma(f,\zeta)} \geq \frac{1}{24(d-1)\,\mathtt{C}_1(f)},$$

which concludes the proof. □

6.2. Complexity of univariate solvers

By "univariate solver", we refer to an algorithm that given a univariate polynomial $f \in \mathcal{P}_{1,d}$ and an interval J where f has only simple roots, it outputs a set of isolating intervals for the roots of f in J. The latter means that we are focusing on finding real roots. In our case, we will focus in the case where the interval is I.

In general, an univariate solver will be of the form

$$\mathcal{P}_{1,d} \times \mathbb{N} \ni (f, L) \mapsto \{(x_i, r_i, n_i)\}_{i=0}^{t} \subseteq I \times (0, 2^{-L}) \times \mathbb{N}$$

where the input is a polynomial f and the natural number L and the output, $\{B_{\mathbb{C}}(x_i, r_i)\}_{i=0}^{t}$, is a disjoint family of complex disks with centers at I and radius at 2^{-L}, each containing at most n_i roots; that is for all i it holds $\left|f^{-1}(0) \cap B_{\mathbb{C}}(x_i, r_i)\right| \leq n_i$. In the particular case, where it also holds that for all i,

$$1 \leq \left|f^{-1}(0) \cap B_{\mathbb{C}}(x_i, r_i)\right| = n_i,$$

we say that $\{(x_i, r_i, n_i)\}_{i=0}^{t}$ is an (I, L)-*covering of* f (Jindal and Sagraloff, 2017, Definition 1).

In our computational model, the input polynomials will be *bitstream polynomials* (Eigenwillig, 2010, Definition 3.35), i.e., we can access the coefficients of the polynomial at any desired precision and we will consider the bit complexity. We will focus on two solvers: Descartes and JindalSagraloff.

6.2.1. Descartes *Solver*

Descartes is a prominent representative of the subdivision-based algorithms for isolating the real roots of polynomials, usually with integer coefficients. The algorithm is extremely efficient in practice (Rouillier and Zimmermann, 2004; Hemmer et al., 2009). We refer to Eigenwillig (2010) for general exposition about Descartes.

For simplicity, we would focus on the size of the subdivision tree of Descartes. Recall that the *subdivision tree* of a subdivision-based algorithm is the tree of intervals, ordered by containment, that the algorithm processes during its execution.

Proposition 6.4. *Given $f \in \mathcal{P}_{1,d}$ with support $|M|$, the algorithm* Descartes *computes isolating intervals for all the roots of f in I in at most $O(\log \mathtt{C}_1(f) + \log d)$ iterations (this is the depth of the subdivision tree). In particular, the subdivision tree of* Descartes *at input (f, I) has size*

$$O\left(|M|\left(\log \mathtt{C}_1(f) + \log d\right)\right).$$

Proof. We note that an interval $J \subseteq I$ is a terminal interval for Descartes as long as $B_{\mathbb{C}}(m(J), w(J)/2)$ does not contain roots or $B_{\mathbb{C}}(m(J) + i\sqrt{3}w(J)/6,)$ contains exactly one real root, by the one-circle and two-circles theorems (Mehlhorn and Sagraloff, 2011, Theorem 2). Hence the depth of the subdivision tree is at most $O\left(\log \max\{\varepsilon^{-1}, \Delta_\varepsilon(f)^{-}1\}\right)$ for some $\varepsilon > 0$, and so, by Theorem 6.3, at most $O\left((\log \mathtt{C}_1(f) + \log d)\right)$.

Now, since f has support $|M|$, the number of sign variations in $(0, \infty)$ and $(-\infty, 0)$ is at most $O(|M|)$. Hence, no level of the subdivision tree can have more than $O(|M|)$ nodes by the Schoenberg's theorem (Mehlhorn and Sagraloff, 2011, Theorem 3). This finishes the proof. □

6.2.2. JindalSagraloff *Solver*

Jindal and Sagraloff (2017) propose an algorithm, JindalSagraloff, to solve sparse polynomials. In this setting , the representation of the polynomial $f \in \mathcal{P}_{1,d}$ consists of its support $M \subset \mathbb{N}$, which has size $|M| \log d$, and a sequence of $|M|$ coefficients.

Theorem 6.5. *The algorithm* JindalSagraloff *outputs an (L, I)-covering of f on input $(f, L) \in \mathcal{P}_{1,d} \times \mathbb{N}$ with f supported on $M \subseteq \mathbb{N}$. The algorithm* JindalSagraloff *runs with at most*

$$O\left(|M|^{12} \log^3 d \max\{\log^2 \|f\|_1, L^2\}\right)$$

bit operations.

Proof. This is essentially (Jindal and Sagraloff, 2017, Lemma 8) rewritten. For the rest of the claims, one has just to read the assumptions in the introduction. □

Remark 6.6. We note that for dense polynomials the estimate of JindalSagraloff is far from optimal, because of that we will state many of the later results only as $\mathrm{poly}(|M|, \log d)$, since determining the exact form of the polynomial does not lead to optimal bounds.

The following proposition it is based on the fact that in order to find the roots of a polynomial with an (I, L)-cover, we only need to compute (I, k)-covers until $k > \max\{\log \Delta_\varepsilon(f), \log \varepsilon^{-1}\}$ for some $\varepsilon > 0$.

Proposition 6.7. *Given $f \in \mathcal{P}_{1,d}$ with support $|M|$, the algorithm* JindalSagraloff *computes isolating intervals for all the roots of f in I with a run-time of*

$$O\left(|M|^{12} \log^3 d \max\{\log^2 \|f\|_1, \log^3 C_1(f)\}\right)$$

bit operations.

Proof. We only need to apply the bounds in Theorem 6.3. □

7. Probability estimates of the condition number and complexity

We refine the techniques of Cucker et al. (2019) to obtain explicit constants in the bounds and to deal with a restricted classes of sparse polynomials. We also add certain variations of the randomness models we consider. Of very special interest are the relaxation of the hypotheses

7.1. Probabilistic concepts and toolbox

We introduce now the relevant probabilistic definitions and results. There will be two kind of probabilistic results: tail bounds, mainly subgaussian and subexponential; and anti-concentration bounds.

7.1.1. Subgaussian and other tail bounds

The more usual tail bound is some form of Markov's inequality (Vershynin, 2018, Proposition 1.2.4) where for a random variable $\mathfrak{x} \in \mathbb{R}$,

$$\mathbb{P}(\|\mathfrak{x}\| \geq t) \leq \frac{\mathbb{E}|\mathfrak{x}|^k}{t^k}.$$

The variables in which we will focus will satisfy such a tail bound, but in an stronger sense.

Definition 7.1. (see Vershynin, 2018, 2.5 and 2.7) Let $\mathfrak{x} \in \mathbb{R}$ be a random variable.

(SE) We call $\mathfrak{x}$ *subexponential*, if there exist a $E > 0$ such that for all $t \geq E$,

$$\mathbb{P}(|\mathfrak{x}| > t) \leq 2\exp(-t^2/E^2).$$

The smallest such E is the *subexponential constant* of $\mathfrak{x}$.

(SG) We call $\mathfrak{x}$ *subgaussian*, if there exist a $K > 0$ such that for all $t \geq K$,

$$\mathbb{P}(|\mathfrak{x}| > t) \leq 2\exp(-t^2/K^2).$$

The smallest such K is the *subgaussian constant* of $\mathfrak{x}$.

(pSE) Let $p \geq 1$. We call $\mathfrak{x}$ *p-subexponential*, if there exist a $L > 0$ such that for all $t \geq L$,

$$\mathbb{P}(|\mathfrak{x}| > t) \leq 2\exp(-t^p/L^p).$$

The smallest such L is the *p-subexponential* of $\mathfrak{x}$.

Remark 7.2. We note that the technical term for the subexponetial, subgaussian and p-subexponential constants are, respectively, the ψ_1-, ψ_2- and ψ_p-norms (see Vershynin, 2018, 2.5 and 2.7). However, note that these norms can be defined in many ways, although all of them are equivalent up to an absolute constant.

Remark 7.3. Note that we are not requiring the random variables to be centered, i.e., to have zero expectation. This plays a role into having an uniform approach for the average and smoothed analyses.

The main inequality is a variant of Hoeffding inequality which we will use in our bounds of the condition number.

Proposition 7.4. *Let $\mathfrak{x} \in \mathbb{R}^M$ be a random vector.*

(SE) If for each $\alpha \in M$, $\mathfrak{x}_\alpha$ is subexponential with subexponential constant E_α. Then for all $t \geq \sum_\alpha E_\alpha$, we have

$$\mathbb{P}(\|\mathfrak{x}\|_1 \geq t) \leq 2|M| \exp\left(-t/\left(\sum_{\alpha\in M} K_\alpha\right)\right).$$

(SG) If for each $\alpha \in M$, $\mathfrak{x}_\alpha$ is subgaussian with subgaussian constant K_α. Then for all $t \geq \sum_\alpha K_\alpha$, we have

$$\mathbb{P}(\|\mathfrak{x}\|_1 \geq t) \leq 2|M| \exp\left(-t^2/\left(\sum_{\alpha\in M} K_\alpha\right)^2\right).$$

(pSE) Let $p \geq 1$. If for each $\alpha \in M$, $\mathfrak{x}_\alpha$ is p-subexponential with subexponential constant L_α. Then for all $t \geq \sum_\alpha L_\alpha$, we have

$$\mathbb{P}(\|\mathfrak{x}\|_1 \geq t) \leq 2|M| \exp\left(-t^p/\left(\sum_{\alpha\in M} K_\alpha\right)^p\right).$$

Proof. We prove the last claim, as the other two are only particular cases of this. We have that for $t \geq \sum_{\alpha\in M} L_\alpha$,

$$\begin{aligned}
\mathbb{P}\left(\textstyle\sum_{\alpha\in M} |\mathfrak{x}_\alpha| \geq t\right) &= \mathbb{P}\left(\textstyle\sum_{\alpha\in M} |\mathfrak{x}_\alpha| \geq \sum_{\alpha\in M} L_\alpha t/\left(\sum_{\alpha\in M} L_\alpha\right)\right) && \left(\sum_{\alpha\in M} L_\alpha t/\left(\textstyle\sum_{\alpha\in M} L_\alpha\right) = t\right) \\
&\leq \mathbb{P}\left(\exists \alpha \in M \mid |\mathfrak{x}_\alpha| \geq L_\alpha t/\left(\textstyle\sum_{\alpha\in M} L_\alpha\right)\right) && \text{(Implication bound)} \\
&\leq |M| \max_{\alpha\in M} \mathbb{P}\left(|\mathfrak{x}_\alpha| \geq L_\alpha t/\left(\textstyle\sum_{\alpha\in M} L_\alpha\right)\right) && \text{(Union bound)} \\
&\leq 2|M| \exp\left(-t^2/\left(\textstyle\sum_{\alpha\in M} L_\alpha\right)^2\right), && \text{(Hypothesis)}
\end{aligned}$$

where the implication bound is based on the fact that $\sum_{i=1}^n x_i \geq \sum_{i=1}^n y_i$ implies that for some i, $x_i \geq y_i$ and the application of the hypothesis follows from the fact that $L_\alpha t/\left(\sum_{\alpha\in M} L_\alpha\right) \geq L_\alpha$ for $t \geq \sum_{\alpha\in M} L_\alpha$. □

7.1.2. Anti-concentration bounds

A random variable $\mathfrak{x} \in \mathbb{R}$ such that for some $x \in \mathbb{R}$ has $\mathbb{P}(\mathfrak{x} = x) > 0$ is said to be concentrated at x. If this phenomenon happens with some of the coefficients of our random polynomial, we cannot guarantee that our random polynomial has finite condition almost-surely,it might happen that it equals a particular ill-posed polynomial (one with infinite condition) with non-zero probability. Because of this, we introduce the following.

Definition 7.5. A random variable $\mathfrak{x}$ has the *anti-concentration property*, if there exists a $\rho > 0$, such that for all $\varepsilon > 0$,

$$\max\{\mathbb{P}(|\mathfrak{x} - u| \leq \varepsilon) \mid u \in \mathbb{R}\} \leq 2\rho\varepsilon.$$

The smallest such ρ is the anti-concentration constant of $\mathfrak{x}$.

The above definition is good at giving the property in a clear way. The following proposition characterizes anti-concentration in terms of having a bounded continuous density. We will use this equivalence without mentioning it.

Proposition 7.6. *Let $\mathfrak{x} \in \mathbb{R}$ be a random variable. Then $\mathfrak{x}$ has the anti-concentration property with anti-concentration constant $\leq \rho$ if and only if $\mathfrak{x}$ is absolutely continuous with respect the Lebesgue measure with density $\delta_{\mathfrak{x}}$ bounded by ρ.*

Proof. This is (Rudelson and Vershynin, 2015, Proposition 2.2). The precise bounds follow immediately once we have shown the equivalence. □

We conclude with the following proposition which is a generalization of (Rudelson and Vershynin, 2015, Theorem 1.1) for more general linear maps. The explicit constants are thanks to the work of Livshyts et al. (2016).

Proposition 7.7. *Let $A \in \mathbb{R}^{k\times N}$ be a surjective linear map and $\mathfrak{x} \in \mathbb{R}^N$ be a random vector such that the $\mathfrak{x}_i$'s are independent random variables with densities (with respect the Lebesgue measure) bounded almost everywhere by ρ. Then, for all measurable $U \subseteq \mathbb{R}^k$,*

$$\mathbb{P}(A\mathfrak{x} \in U) \leq \operatorname{vol}(U)\left(\sqrt{2}\rho\right)^k / \sqrt{\det AA^*}.$$

Proof. Using SVD, write $A = QSP$ where, $P \in \mathbb{R}^{k\times N}$ is an orthogonal projection, S a diagonal matrix containing the singular values of A, and Q an orthogonal matrix.

By (Rudelson and Vershynin, 2015, Theorem 1.1), see also (Livshyts et al., 2016, Theorem 1.1) for the explicit constant, we have that $P\mathfrak{x} \in \mathbb{R}^k$ is a random vector with density bounded, almost everywhere, by $(\sqrt{2}\rho)^k$. Hence

$$\mathbb{P}(A\mathfrak{x} \in U) = \mathbb{P}(P\mathfrak{x} \in (QS)^{-1}U) \leq \operatorname{vol}\left((QS)^{-1}U\right)(\sqrt{2}\rho)^k.$$

This suffices to conclude the proof, since we have $\operatorname{vol}\left((QS)^{-1}U\right) = \operatorname{vol}(U)/\det(QS)$ and $\det(QS) = \sqrt{\det AA^*}$. □

7.2. Zintzo random polynomials

We introduce a new class of random polynomials; we call them *zintzo*[3] polynomials. The zintzo polynomials have many simililarities with the dobro random polynomials, introduced by Cucker et al. (2019). The main difference between the two is in the variance structure of the coefficients. While dobro polynomials scale the coefficients according to the weights induced by the Weyl norm, zintzo polynomials do not. This property makes zintzo random polynomials a more natural model of random polynomials. Moreover, it allows us to explicitly include sparseness in the model of randomness.

Definition 7.8. Let $M \subseteq \mathbb{N}^n$ be a finite set such that $0, e_1, \dots, e_n \in M$. A *zintzo random polynomial supported on M* is a random polynomial $\mathfrak{f} = \sum_{\alpha \in M} \mathfrak{f}_\alpha X^\alpha \in \mathcal{P}_{n,d}$ such that the coefficients $\mathfrak{f}_\alpha$ are independent subgaussian random variables with the anti-concentration property.

Remark 7.9. We recall that the condition on the support is important to guarantee the smoothness of zero and necessary for our technical assumptions. We stress this, as we will be assuming this in all this section.

The bounds and the complexity estimates involving zintzo random polynomials that we present in the sequel depend on (the product of) the following two quantities:

1. The *subgaussian constant* of $\mathfrak{f}$ which is

$$K_{\mathfrak{f}} := \sum\nolimits_{\alpha \in M} K_\alpha, \tag{7}$$

   where K_α is the subgaussian constant of $\mathfrak{f}_\alpha$.

2. The *anti-concentration constant* of $\mathfrak{f}$ which is

$$\rho_{\mathfrak{f}} := \sqrt[n+1]{\rho_0 \rho_{e_1} \cdots \rho_{e_n}}, \tag{8}$$

   where ρ_0 is the anti-concentration constant of $\mathfrak{f}_0$ and ρ_{e_i} is the anti-concentration constant of $\mathfrak{f}_{e_i}$, for $i \in [n]$.

Example 7.10. Both Gaussian and uniform random polynomials are zintzo random polynomials with

$$K_{\mathfrak{f}} \rho_{\mathfrak{f}} \leq \frac{|M|}{2}. \tag{9}$$

In general, we should think of zintzo random polynomials as a robust version of Gaussian polynomials.

The product $K_{\mathfrak{f}} \rho_{\mathfrak{f}}$ is invariant under multiplication of $\mathfrak{f}$ by a non-zero scalar and it is bounded from below.

Proposition 7.11. *Let $\mathfrak{f}$ be a zintzo random polynomial supported on M. Then $K_{\mathfrak{f}} \rho_{\mathfrak{f}} > (n+1)/4 \geq 1/2$.*

[3]The word "zintzo" is a Basque word that means honest, upright, righteous. We note that this is the same meaning as that of the word "dobro" in a certain sense. This coincidence is not by chance, since it show that dobro and zintzo random polynomials are polynomials of the same kind, but different.

Proof. Using the positivity of the subgaussian constants, K_α, of the coefficients of the zintzo polynomial $\mathfrak{f}$ and the arithmetic-geometric inequality,

$$K_{\mathfrak{f}}\rho_{\mathfrak{f}} \geq (n+1)\sqrt[n+1]{(K_0\rho_0)(K_{e_1}\rho_{e_1})\cdots(K_{e_n}\rho_{e_n})}.$$

Hence, it suffices to show that for a random variable with X with subgaussian constant K and anti-concentration constant ρ, $K\rho \geq 1/4$. Now, by definition,

$$3K\rho \geq \mathbb{P}\left(|X| \leq 1.5K\right) = 1 - \mathbb{P}\left(|X| > 1.5K\right) \geq 1 - 2\exp\left(-2.25\right).$$

Calculating we get $K\rho \geq 1/4$, as desired. □

7.2.1. Smoothed analysis

Recall that in the context of smoothed analysis, as introduced by Spielman and Teng (2002), we study the complexity algorithms when the input polynomial is a fixed polynomial with a random perturbation. The importance of smoothed analysis lies in that it explains the behaviour of an algorithm in practice better than the average case, since in practice we tend to have a fixed input with a perturbation produced by errors. In our setting, it is natural to consider this perturbation proportional to the norm of the polynomial. The following proposition shows that for zintzo random polynomials the average complexity already includes the smoothed case as a particular case. Because of that, there is no need to give a smoothed analysis of any of our results.

Proposition 7.12. *Let $\mathfrak{f}$ be a zintzo random polynomial supported on M, $f \in \mathcal{P}_{n,d}$ a polynomial supported on M, and $\sigma > 0$. Then, $\mathfrak{f}_\sigma := f + \sigma\|f\|_1\mathfrak{f}$ is a zintzo random polynomial supported on M such that $K_{\mathfrak{f}_\sigma} \leq \|f\|_1(1+\sigma K_{\mathfrak{f}})$ and $\rho_{\mathfrak{f}_\sigma} \leq \rho_{\mathfrak{f}}/(\sigma\|f\|_1)$. In particular,*

$$K_{\mathfrak{f}_\sigma}\rho_{\mathfrak{f}_\sigma} = (K_{\mathfrak{f}} + 1/\sigma)\rho_{\mathfrak{f}}.$$

Note that both the worst case and the average case are limit cases of the smoothed case. When the perturbation, σ, becomes zero, the smoothed case becomes the worst case, and when the perturbation becomes of infinite magnitude, the smoothed case becomes the average case. Because of this, it is not surprising that

$$\lim_{\sigma\to 0} K_{\mathfrak{f}_\sigma}\rho_{\mathfrak{f}_\sigma} = \infty \text{ and } \lim_{\sigma\to\infty} K_{\mathfrak{f}_\sigma}\rho_{\mathfrak{f}_\sigma} = K_{\mathfrak{f}}\rho_{\mathfrak{f}},$$

which shows that this is the case.

Proof of Proposition 7.12. It is enough to show that for $x, s \in \mathbb{R}$ and a random variable $\mathfrak{x}$ with subgaussian constant K and anti-concentration constant ρ, $x + s\mathfrak{x}$ is a random variable with subgaussian constant $\leq |x| + sK$ and anti-concentration constant $\leq \rho/s$. We note that the latter follows directly from the definition, so we only prove the former.

Now, for all $t \geq |x| + sK$,

$$\mathbb{P}(|x + s\mathfrak{x}| \geq t) \leq \mathbb{P}(|\mathfrak{x}| \geq (t - |x|)/s) \leq 2\exp(-(t-|x|)^2/(sK)^2).$$

We can easily check that $t \geq |x| + sK$ implies $(t-|x|)/(sK) \geq t/(|x| + sK)$. Hence, the claim follows. □

7.2.2. p-zintzo random polynomials

Zintzo random polynomials have subgaussian coefficients, we can relax or tighten up this condition which leads to p-zintzo random polynomials. We note that 1-zintzo polynomials will have subexponential coefficients, extending the results that concern zintzo polynomials further.

Definition 7.13. Let $p \geq 1$ and $M \subseteq \mathbb{N}^n$ be a finite set such that $0, e_1, \ldots, e_n \in M$. A *$p$-zintzo random polynomial supported on M* is a random polynomial $\mathfrak{f} = \sum_{\alpha \in M} \mathfrak{f}_\alpha X^\alpha \in \mathcal{P}_{n,d}$ such that the coefficients $\mathfrak{f}_\alpha$ are independent p-subexpoential random variables with the anti-concentration property.

Given a p-zintzo random polynomial $\mathfrak{f} \in \mathcal{P}_{n,d}$, we define the following quantities:

1. The *tail constant* of $\mathfrak{f}$ which is
$$L_{\mathfrak{f}} := \sum\nolimits_{\alpha \in M} L_\alpha, \tag{10}$$
where L_α is the p-subgaussian constant of $\mathfrak{f}_\alpha$.

2. The *anti-concentration constant* of $\mathfrak{f}$ which is
$$\rho_{\mathfrak{f}} := \sqrt[n+1]{\rho_0 \rho_{e_1} \cdots \rho_{e_n}}, \tag{11}$$
where ρ_0 is the anti-concentration constant of $\mathfrak{f}_0$ and ρ_{e_i} is the anti-concentration constant of $\mathfrak{f}_{e_i}$, for $i \in [n]$.

Example 7.14. We note that Gaussian polynomials are p-zintzo for $p \in [1, 2]$, while uniform random polynomials are p-zintzo random polynomials with
$$L_{\mathfrak{f}} \rho_{\mathfrak{f}} \leq \frac{|M|}{2} m$$
for $p \geq 1$. This is the reason we have slightly better results for uniform random polynomials, as we can take the bound for p-zintzo random polynomials when $p \to \infty$.

We note that we only vary the constants that control the how the tail goes to zero as we goes to infinity. Again, our estimates will depend on $L_{\mathfrak{f}} \rho_{\mathfrak{f}}$ which has a universal lower bound. The proof is analogous to that of Proposition 7.11.

Proposition 7.15. *Let $p \geq 1$ and $\mathfrak{f}$ a p-zintzo random polynomial supported on M. Then $L_{\mathfrak{f}} \rho_{\mathfrak{f}} > 9(n+1)/50 \geq 9/25$.* □

In the same way, a version of Proposition 7.12 holds also for p-zintzo random polynomials.

7.3. Condition of zintzo random polynomials

The following two theorems are our main probabilistic results.

Theorem 7.16. *Let $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a zintzo random polynomial supported on M. Then for all $t \geq e$,*
$$\mathbb{P}(\mathcal{C}_1(\mathfrak{f}, x) \geq t) \leq \sqrt{n} d^n |M| \left(\frac{8 K_{\mathfrak{f}} \rho_{\mathfrak{f}}}{\sqrt{n+1}} \right)^{n+1} \frac{\ln^{\frac{n+1}{2}} t}{t^{n+1}}.$$

Theorem 7.17. *Let $p \geq 1$ and $\mathfrak{f} \in \mathcal{P}_{n,d}$ a p-zintzo random polynomial supported on M. Then for all $t \geq e$,*

$$\mathbb{P}(\mathtt{C}_1(\mathfrak{f}, x) \geq t) \leq \sqrt{n} d^n |M| \left(\frac{8 L_{\mathfrak{f}} \rho_{\mathfrak{f}}}{(n+1)^{1-\frac{1}{p}}} \right)^{n+1} \frac{\ln^{\frac{n+1}{p}} t}{t^{n+1}}.$$

The main idea of the proof is to use a union bound to control $\mathbb{P}(\mathtt{C}_1(\mathfrak{f}, x) \geq t)$ by the sum of $\mathbb{P}(\|\mathfrak{f}\|_1 \geq u)$ and $\mathbb{P}(\|R_x \mathfrak{f}\| \leq u/t)$ which are controlled, respectively, by Propositions 7.4 and 7.7. The reason that this works is because $\mathbb{P}(\|\mathfrak{f}\|_1 \geq t)$ decreases exponentially fast. To apply Proposition 7.7, we need the following lemma, which we proof at the end of this subsection.

Lemma 7.18. *Let $M \subseteq \mathbb{N}^n$ as in Definition 7.8 and $\mathcal{P}_{n,d}(M)$ the set of polynomials in $\mathcal{P}_{n,d}$ supported on M. Let $R_x : \mathcal{P}_{n,d}(M) \to \mathbb{R}^{n+1}$ be the linear map given by*

$$R_x : f \mapsto \begin{pmatrix} f(x) & \frac{1}{d}\partial_1 f(x) & \cdots & \frac{1}{d}\partial_n f(x) \end{pmatrix}^*,$$

and $S : \mathcal{P}_{n,d}(M) \to \mathcal{P}_{n,d}(M)$ be the linear map given by

$$S : f = \sum_{\alpha \in M} f_\alpha X^\alpha \mapsto \sum_{\alpha \in M} \rho_\alpha f_\alpha X^\alpha,$$

where $\rho \in \mathbb{R}_+^M$. Then for $\tilde{R}_x := R_x S^{-1}$ we have that

$$\sqrt{\det \tilde{R}_x \tilde{R}_x^*} \geq \frac{1}{d^n \rho_0 \rho_{e_1} \cdots \rho_{e_n}},$$

with respect to coordinates induced by the standard monomial basis.

Proof of Theorem 7.16. We write $\mathtt{C}_1(\mathfrak{f}, x) = \|f\|_1 / \|R_x \mathfrak{f}\|$, where R_x is as in Lemma 7.18 and the norm $\|\cdot\|$ in the denominator is given by $\|y\| = \max\{|y_1|, |y_2| + \cdots + |y_{n+1}|\}$. By the union bound, we have that for $u, s > 0$, it holds

$$\mathbb{P}(\mathtt{C}_1(\mathfrak{f}, x) \geq t) \leq \mathbb{P}(\|\mathfrak{f}\|_1 \geq u) + \mathbb{P}(\|R_x \mathfrak{f}\| \leq u/t). \tag{12}$$

By Propositions 7.4, we have that for $u \geq K_{\mathfrak{f}}$,

$$\mathbb{P}(\|\mathfrak{f}\| \geq u) \leq 2|M| \exp\left(-u^2 / K_{\mathfrak{f}}^2\right). \tag{13}$$

Let $S : \mathcal{P}_{n,d}(M) \to \mathcal{P}_{n,d}(M)$ be as in Lemma 7.18 with ρ_α the anti-concentration constant of $\mathfrak{f}_\alpha$. Then, we have that $S\mathfrak{f}$ has independent random coefficients with densities bounded (almost everywhere) by 1 and so we can apply to it to the Proposition 7.7. We do so with the help of Lemma 7.18, so that we obtain

$$\mathbb{P}(\|R_x \mathfrak{f}\| \leq u/t) = \mathbb{P}(\|\tilde{R}_x (S\mathfrak{f})\| \leq u/t) \leq \frac{2^{n+1}}{n!} d^n (\sqrt{2} \rho_{\mathfrak{f}} u/t)^{n+1}, \tag{14}$$

where $\tilde{R}_x$ is as in Lemma 7.18.

Combining (12), (13), and (14) with $u = K_{\mathfrak{f}} \sqrt{(n+1)\ln t}$, we get

$$\mathbb{P}(\mathtt{C}_1(\mathfrak{f}, x) \geq t) \leq \frac{2|M|}{t^{n+1}} + \frac{2^{n+1}}{n!} d^n \left(\sqrt{2} K_{\mathfrak{f}} \rho_{\mathfrak{f}} \sqrt{n+1} \right)^{n+1} \frac{\ln^{\frac{n+1}{2}} t}{t^{n+1}}.$$

By Stirling's formula, $(n+1)^{n+1}/n! \leq \sqrt{n} e^{n+1}$, and so the desired claim follows for $t \geq e$, by Proposition 7.11. □

Proof of Theorem 7.17. The proof is as that of Theorem 7.16, but with p instead of 2 in the exponents of (13), $u = L_{\mathfrak{f}}((n+1)\ln t)^{\frac{1}{p}}$ instead of $u = K_{\mathfrak{f}}\sqrt{(n+1)\ln t}$, and Proposition 7.15 instead of Proposition 7.11. □

Proof of Lemma 7.18. The maximal minor of A_x is given by

$$R_X = \begin{pmatrix} 1 & x^* \\ 0 & \frac{1}{d}\mathbb{I} \end{pmatrix}.$$

This is precisely the minor associated to the subset $\{1, X_1, \ldots, X_n\}$ of the standard monomial basis of $\mathcal{P}_{n,d}(M)$. Note that at this point we require the assumption that $0, e_1, \ldots, e_n \in M$.

By the Cauchy-Binet identity, $\sqrt{\det A_x A_x^*}$ is lower-bounded by the absolute value of the determinant of the given maximal minor. Hence the lemma follows. □

7.4. Probabilistic complexity analysis for the Plantinga-Vegter algorithm

We now proceed to the complexity analysis of the condition-based quantities in the complexity analysis of the Plantinga-Vegter algorithm. The theorems of interest are the following two.

Theorem 7.19. *Let $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a zintzo random polynomial supported on M. The average number of boxes of the final subdivision of* PV-Subdivsion *using the interval approximations* (3) *and* (4) *on input $\mathfrak{f}$ is at most*

$$2n^{\frac{3}{2}} d^{2n} |M| \left(20(n+1)K_{\mathfrak{f}}\rho_{\mathfrak{f}}\right)^{n+1}.$$

Theorem 7.20. *Let $p \geq 1$ and $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a p-zintzo random polynomial supported on M. The average number of boxes of the final subdivision of* PV-Subdivsion *using the interval approximations* (3) *and* (4) *on input $\mathfrak{f}$ is at most*

$$2n^{\frac{3}{2}}(n+1)^{\frac{1}{p}-\frac{1}{2}} d^{2n} |M| \left(64 n^{\frac{1}{p}} (n+1)^{\frac{1}{p}} L_{\mathfrak{f}}\rho_{\mathfrak{f}}\right)^{n+1}.$$

The above two results follow from Corollary 5.6 together with the following two results.

Proposition 7.21. *Let $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a zintzo random polynomial supported on M. Then,*

$$\mathbb{E}_{\mathfrak{f}} \mathbb{E}_{\mathfrak{f} \in I^n} \mathtt{C}_1(f,x)^n \leq 2n^2 d^n |M| \left(7\sqrt{n+1} K_{\mathfrak{f}}\rho_{\mathfrak{f}}\right)^{n+1}.$$

Proposition 7.22. *Let $p \geq 1$ and $\mathfrak{f} \in \mathcal{P}_{n,d}$ a p-zintzo random polynomial supported on M. Then,*

$$\mathbb{E}_{\mathfrak{f}} \mathbb{E}_{\mathfrak{f} \in I^n} \mathtt{C}_1(f,x)^n \leq 2n^2 (n+1)^{\frac{1}{p}-\frac{1}{2}} d^n |M| \left(\frac{e^{1-\frac{1}{p}} 8}{p^{\frac{1}{p}}} n^{\frac{1}{p}-\frac{1}{2}} (n+1)^{\frac{1}{p}} L_{\mathfrak{f}}\rho_{\mathfrak{f}}\right)^{n+1}.$$

Proof of Proposition 7.21. By the Fubini-Tonelli theorem, we have

$$\mathbb{E}_{\mathfrak{f}} \mathbb{E}_{\mathfrak{f} \in I^n} \mathtt{C}_1(f,x)^n = \mathbb{E}_{\mathfrak{f} \in I^n} \mathbb{E}_{\mathfrak{f}} \mathtt{C}_1(f,x)^n,$$

so it is enough to compute $\mathbb{E}_{\mathfrak{f}} \mathtt{C}_1(f,x)^n = \int_1^\infty \mathbb{P}(\mathtt{C}_1(\mathfrak{f},x)^n \geq t)$. The latter, by Theorem 7.16, is bounded from above by

$$e^n \sqrt{n} d^n |M| \left(\frac{8K_{\mathfrak{f}}\rho_{\mathfrak{f}}}{\sqrt{n}\sqrt{n+1}}\right)^{n+1} \int_1^\infty \frac{\ln^{\frac{n+1}{2}} t}{t^{1+\frac{1}{n}}} \, \mathrm{d}t.$$

After straightforward calculations, we obtain

$$\int_1^\infty \frac{\ln^{\frac{n+1}{2}} t}{t^{1+\frac{1}{n}}}\,\mathrm{d}t = n^{\frac{n+3}{2}}\Gamma\left(\frac{n+3}{2}\right) \leq e\sqrt{\pi}n^{\frac{n+4}{2}}\left(\frac{n+1}{2e}\right)^{\frac{n+1}{2}},$$

where Γ is Euler's Gamma function and the second inequality follows from Striling's approximation. Hence, the bound follows. □

Proof of Proposition 7.22. We do as in the proof of Proposition 7.21, but applying Theorem 7.17 instead of Theorem 7.16. We obtain this way that the desired quantity is bounded by

$$e^n\sqrt{n}d^n|M|\left(\frac{8L_{\mathfrak{f}}\rho_{\mathfrak{f}}}{\sqrt{n}(n+1)^{1-\frac{1}{p}}}\right)^{n+1}\int_1^\infty \frac{\ln^{\frac{n+1}{p}} t}{t^{1+\frac{1}{n}}}\,\mathrm{d}t.$$

Computing the integral turns into

$$\int_1^\infty \frac{\ln^{\frac{n+1}{p}} t}{t^{1+\frac{1}{n}}}\,\mathrm{d}t = n^{1+\frac{n+1}{p}}\Gamma\left(\frac{n+1}{p}+1\right) \leq e\sqrt{\pi}n^{2+\frac{n+1}{p}}\left(\frac{n+1}{pe}\right)^{\frac{n+1}{p}},$$

after we apply Stirling's approximation in the last inequality. The result now follows. □

7.5. Probabilistic complexity analysis for univariate solvers

For Descartes and JindalSagraloff, we prove the two following general result.

Theorem 7.23. *Let $p \geq 1$ and $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a zintzo random polynomial supported on M. The average size of the subdivision tree of* Descartes *on input $\mathfrak{f}$ is at most*

$$O\left(|M|\left|\log(dL_{\mathfrak{f}}\rho_{\mathfrak{f}})\right|\right).$$

Moreover, the kth moment of the size is bounded by

$$O\left(k|M|\left|\log(dL_{\mathfrak{f}}\rho_{\mathfrak{f}})\right|\right)^k.$$

Theorem 7.24. *Let $p \geq 1$ and $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a p-zintzo random polynomial supported on M. The average bit-complexity of* JindalSagraloff *on input $(\mathfrak{f}, I)$ is at most*

$$O\left(|M|^{12}\log^6 d\left|\log(L_{\mathfrak{f}}\rho_{\mathfrak{f}})\right|\right).$$

Moreover, the kth moment of the bit-run-time is bounded by

$$O\left(k|M|^{12}\log^6 d\left|\log(L_{\mathfrak{f}}\rho_{\mathfrak{f}})\right|\right)^k.$$

Remark 7.25. The global assumption on the anti-concentrarion constants is to control $\log\|\mathfrak{f}\|_1$ when $\|\mathfrak{f}\|_1$ is small.

These results follow from Propositions 6.4 and 6.7 and the following two propositions. We give the computations below.

Proposition 7.26. *Let $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a zintzo random polynomial supported on M. Then, for all $t > 2e$,*

$$\mathbb{P}(\mathtt{C}_1(\mathfrak{f}) \geq t) \leq 2\sqrt{n}d^{2n}|M|\left(\frac{16K_{\mathfrak{f}}\rho_{\mathfrak{f}}}{\sqrt{n+1}}\right)^{n+1}\frac{\ln^{\frac{n+1}{2}} t}{t} \leq 2\sqrt{n}d^{2n}|M|\left(10K_{\mathfrak{f}}\rho_{\mathfrak{f}}\right)^{n+1}\frac{1}{\sqrt{t}}.$$

Proposition 7.27. *Let $p \geq 1$ and $\mathfrak{f} \in \mathcal{P}_{n,d}$ be a p-zintzo random polynomial supported on M. Then, for all $t > 2e$,*

$$\mathbb{P}(\mathtt{C}_1(\mathfrak{f}) \geq t) \leq 2\sqrt{n}d^{2n}|M|\left(\frac{16L_{\mathfrak{f}}\rho_{\mathfrak{f}}}{(n+1)^{1-\frac{1}{p}}}\right)^{n+1}\frac{\ln^{\frac{n+1}{p}} t}{t} \leq 2\sqrt{n}d^{2n}|M|\left(\frac{16L_{\mathfrak{f}}\rho_{\mathfrak{f}}}{(n+1)^{1-\frac{2}{p}}}\right)^{n+1}\frac{1}{t^{1-\frac{1}{p}}}.$$

Proof of Proposition 7.26. The idea is to use an efficient ε-net of I^n and the 2nd Lipschitz property to turn our local estimates into global ones, as is done in (Tonelli-Cueto, 2019, Theorem 1$^{\S 2}$19). Recall, that an ε-net of I^n (with respect to the ∞-norm) is a finite subset $\mathcal{G} \subseteq I^n$ such that, for all $y \in I^n$, $\operatorname{dist}_\infty(y, \mathcal{G}) \leq \varepsilon$.

Note that for every $\varepsilon \in (0,1)$, we have an ε-net $\mathcal{G}_\varepsilon \subseteq I^n$ of size $\leq 2\varepsilon^{-n}$. To construct it, we take the uniform grid in the cube.

If $\mathtt{C}_1(\mathfrak{f}) \geq t$, then

$$\max\left\{\mathtt{C}_1(\mathfrak{f}, x) \mid x \in \mathcal{G}_{\frac{1}{dt}}\right\} \geq \frac{t}{2}$$

by the 2nd Lipschitz property (Theorem 4.2). Effectively, let $x_* \in I^n$ such that $\mathtt{C}_1(\mathfrak{f}) = \mathtt{C}_1(\mathfrak{f}, x_*)$, then there is $x \in \mathcal{G}_{\frac{1}{dt}}$ such that $\operatorname{dist}_\infty(x, x_*) \leq \frac{1}{dt}$ and therefore

$$\max\{\mathtt{C}_1(\mathfrak{f}, x) \mid x \in \mathcal{G}_{\frac{1}{2dt}}\} \geq \mathtt{C}_1(\mathfrak{f}, x) = \left(\frac{1}{\mathtt{C}_1(\mathfrak{f}, x)}\right)^{-1} \geq \left(\frac{1}{\mathtt{C}_1(\mathfrak{f}, x_*)} + \frac{1}{t}\right)^{-1} = \left(\frac{1}{\mathtt{C}_1(\mathfrak{f})} + \frac{1}{t}\right)^{-1} \geq \frac{t}{2}.$$

To obtain the first inequality, we argue as follows:

$$\begin{aligned}
\mathbb{P}(\mathtt{C}_1(\mathfrak{f}) \geq t) &\leq \mathbb{P}\left(\exists x \in \mathcal{G}_{\frac{1}{dt}} \mid \mathtt{C}_1(\mathfrak{f}, x) \geq \frac{t}{2}\right) && \text{(Implication bound)}\\
&\leq \left|\mathcal{G}_{\frac{1}{dt}}\right| \max\left\{\mathbb{P}\left(\mathtt{C}_1(\mathfrak{f}, x) \geq \frac{t}{2}\right) \mid x \in \mathcal{G}_{\frac{1}{dt}}\right\} && \text{(Union bound)}\\
&\leq 2d^n t^n \max\left\{\mathbb{P}\left(\mathtt{C}_1(\mathfrak{f}, x) \geq \frac{t}{2}\right) \mid x \in \mathcal{G}_{\frac{1}{dt}}\right\} && \left(\left|\mathcal{G}_{\frac{1}{dt}}\right| \leq 2d^n t^n\right)\\
&\leq 2\sqrt{n}d^{2n}|M|\left(\frac{16K_{\mathfrak{f}}\rho_{\mathfrak{f}}}{\sqrt{n+1}}\right)^{n+1}\frac{\ln^{\frac{n+1}{2}} t}{t} && \text{(Theorem 7.16)}
\end{aligned}$$

For the second one, note that $\ln^{n+1} t \leq \frac{(n+1)^{(n+1)}}{e^{n+1}} t$ □

Proof of Proposition 7.26. Like the proof of Proposition 7.26, but using Theorem 7.16 instead of Theorem 7.17. □

We restrict ourselves now to the technical results needed in the proofs. The following propistion proves Theorem 7.23 and so Theorem 2.3.

Proposition 7.28. *Let $p \geq 1$ and $\mathfrak{f} \in \mathcal{P}_{1,d}$ be a p-zintzo random polynomial supported on M. Then for $k \geq 1$,*

$$\mathbb{E}_{\mathfrak{f}} \log^k \mathtt{C}_1(\mathfrak{f}) \leq O\left(k(\log(dL_{\mathfrak{f}}\rho_{\mathfrak{f}}))\right)^k.$$

Proof. Let $\mathfrak{x} \in [1, \infty)$ be a random variable such that

$$\mathbb{P}(\mathfrak{x} \geq t) \leq Ct^{-\alpha}$$

for all $t \geq t_0$, where $C, \alpha > 0$ and $t_0 \geq 2$. Then we have that

$$\mathbb{E}_{\mathfrak{x}} \log^k \mathfrak{x} \leq \log^k t_0 + \frac{Ck^k}{\alpha^k t_0^{\frac{\alpha}{2}}}.$$

To see this note that

$$\int_0^\infty \mathbb{P}\left(\log^k \mathfrak{x} \geq s\right) \mathrm{d}s \leq \log^k t_0 + \int_{\log^k t_0}^\infty \mathbb{P}\left(\log^k \mathfrak{x} \geq t\right) \mathrm{d}s \leq \log^k t_0 + C \int_{\log^k t_0}^\infty 2^{-\alpha s^{\frac{1}{k}}} \mathrm{d}s.$$

Then, we have that

$$\begin{aligned}
\int_{\log^k t_0}^\infty 2^{-\alpha s^{\frac{1}{k}}} \mathrm{d}s &= \frac{k}{\alpha^k} \int_{\alpha \ln 2 \log t_0}^\infty u^{k-1} e^{-u} \mathrm{d}u && \left(s = \alpha \ln 2\, t^{\frac{1}{k}}\right) \\
&\leq \frac{k}{\alpha^k} \left(\frac{2(k-1)}{e}\right)^{k-1} \int_{\alpha \ln 2 \log t_0}^\infty e^{-\frac{u}{2}} \mathrm{d}u && \left(u^{k-1} e^{-\frac{u}{2}} \leq \left(\frac{2(k-1)}{e}\right)^{k-1}\right) \\
&= \frac{k}{\alpha^k} \left(\frac{2(k-1)}{e}\right)^{k-1} t_0^{-\frac{\alpha}{2}} \\
&\leq \frac{k^k}{\alpha^k t_0^{\frac{\alpha}{2}}}.
\end{aligned}$$

Now, we take $\log t_0 = \Omega\left(\log d + \log(L_{\mathfrak{f}} \rho_{\mathfrak{f}})\right)$. □

The following two propositions gives the proof of Theorem 7.24 and so of Theorem 2.4. For applying it, just notice that

$$\mathbb{E} \max\{|\mathfrak{x}|, |\mathfrak{y}|\} \leq \mathbb{E}|\mathfrak{x}| + \mathbb{E}|\mathfrak{y}|.$$

Proposition 7.29. *Let $p \geq 1$ and $\mathfrak{f} \in \mathcal{P}_{1,d}$ be a p-zintzo random polynomial supported on M. Assume that for all $\alpha \in M$, the anti-concentration constant of $\mathfrak{f}_\alpha$ is ≤ 1. Then for $k \geq 1$,*

$$\mathbb{E}_{\mathfrak{f}} |\log^k \|f\|_1| \leq O\left(k^k + \left(1 + \frac{k^k}{p^k}\right) \log^k L_{\mathfrak{f}}\right).$$

Proof. We have that

$$\mathbb{E}_{\mathfrak{f}} |\log^k \|f\|_1| = \int_0^1 |\log^k t| \delta_{\|f\|_1}(t) \mathrm{d}t + \int_1^\infty |\log^k t| \delta_{\|f\|_1}(t) \mathrm{d}t,$$

where $\delta_{\|\mathfrak{f}\|_1}$ is the density of $\mathfrak{f}$.

By our assumption, $\mathfrak{f}$ has density (with respect the Lebesgue measure) and this density is bounded by 1. Hence we have that for $J \subseteq [0, 1]$,

$$\mathbb{P}(\|\mathfrak{f}\|_1 \in J) = \mathrm{vol}\{x \in \mathbb{R}^M \mid \|x\|_1 \in J\} \leq 2w(J);$$

and so $\|\mathfrak{f}\|_1$ has density bounded by 2. To finish the estimation of the first summand, note now that

$$\int_0^1 |\log^k t| \, \mathrm{d}t = \in_0^\infty \; e^{-s^{\frac{1}{k}}} \, \mathrm{d}s = k!.$$

If we define the random variable $\mathfrak{x}$ to be $\|f\|_1$ if $\|f\|_1 \geq 1$ and zero otherwise, then

$$\int_1^\infty |\log^k t| \delta_{\|f\|_1}(t) \, \mathrm{d}t = \mathbb{E}_{\mathfrak{x}} \log^k \mathfrak{x}$$

and $\mathfrak{x}$ satisfies

$$\mathbb{P}(\mathfrak{x} \geq t) \leq e^{-\frac{t^p}{L_{\mathfrak{f}}^p}} \leq \frac{L_{\mathfrak{f}}^p}{t^p}$$

for $t \geq L_{\mathfrak{f}}$, by Proposition 7.4. Arguing as in the proof of Proposition 7.28, we obtain then that

$$\mathbb{E}_{\mathfrak{x}} \log^k \mathfrak{x} \leq \left(1 + \frac{k^k}{p^k}\right) \log^k L_{\mathfrak{f}}.$$

□

Acknowledgments

The first author is supported by a postdoctoral fellowship of the 2020 "Interaction" program of the Fondation Sciences Mathématiques de Paris. The authors are partially supported by ANR JCJC GALOP (ANR-17-CE40-0009), the PGMO grant ALMA and the PHC GRAPE.

Both authors are grateful to Alperen Ergür for various discussions and suggestions. The first author is grateful to Evgenia Lagoda for moral support.

References

Armentano, D., Beltrán, C., 2019. The polynomial eigenvalue problem is well conditioned for random inputs. SIAM J. Matrix Anal. Appl. 40, 175–193. doi:10.1137/17M1139941.

Beltrán, C., Kozhasov, K., 2019. The real polynomial eigenvalue problem is well conditioned on the average. Foundations of Computational Mathematics On-line First. doi:10.1007/s10208-019-09414-2.

Beltrán, C., Pardo, L.M., 2008. On Smale's 17th problem: a probabilistic positive solution. Found. Comput. Math. 8, 1–43. doi:10.1007/s10208-005-0211-0.

Bürgisser, P., Cucker, F., 2011. On a problem posed by Steve Smale. Ann. of Math. (2) 174, 1785–1836. doi:10.4007/annals.2011.174.3.8.

Bürgisser, P., Cucker, F., 2013. Condition. volume 349 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin. doi:10.1007/978-3-642-38896-5.

Bürgisser, P., Cucker, F., Lairez, P., 2020. Rigid continuation paths II. Structured polynomial systems. arXiv:2010.10997.

Burr, M., Choi, S.W., Galehouse, B., Yap, C.K., 2012. Complete subdivision algorithms, II: Isotopic meshing of singular algebraic curves. J. Symbolic Comput. 47, 131–152. doi:10.1016/j.jsc.2011.08.021.

Burr, M., Krahmer, F., Yap, C., 2009. Continuous amortization: A non-probabilistic adaptive analysis technique. Electronic Colloquium on Computational Complexity 16.

Burr, M.A., 2016. Continuous amortization and extensions: with applications to bisection-based root isolation. J. Symbolic Comput. 77, 78–126. doi:10.1016/j.jsc.2016.01.007.

Burr, M.A., Gao, S., Tsigaridas, E.P., 2017. The complexity of an adaptive subdivision method for approximating real curves, in: ISSAC'17—Proceedings of the 2017 ACM International Symposium on Symbolic and Algebraic Computation. ACM, New York, pp. 61–68. doi:10.1145/3087604.3087654.

Cucker, F., Ergür, A.A., Tonelli-Cueto, J., 2019. Plantinga-Vegter Algorithm Takes Average Polynomial Time, in: Proceedings of the 2019 on International Symposium on Symbolic and Algebraic Computation, ACM, New York, Beijing, China. pp. 114–121. doi:10.1145/3326229.3326252.

Cucker, F., Ergür, A.A., Tonelli-Cueto, J., 2020a. Functional norms, condition numbers and numerical algorithms in algebraic geometry. Manuscript.

Cucker, F., Ergür, A.A., Tonelli-Cueto, J., 2020b. On the Complexity of the Plantinga-Vegter Algorithm. arXiv:2004.06879.

Cucker, F., Krick, T., Malajovich, G., Wschebor, M., 2008. A numerical algorithm for zero counting. I: Complexity and accuracy. J. Complexity 24, 582–605. doi:10.1016/j.jco.2008.03.001.

Cucker, F., Krick, T., Shub, M., 2018. Computing the Homology of Real Projective Sets. Found. Comput. Math. 18, 929–970. doi:10.1007/s10208-017-9358-8.

Dedieu, J.P., 1997. Estimations for the Separation Number of a Polynomial System. Journal of Symbolic Computation 24, 683–693.

Dedieu, J.P., 2006. Points fixes, zéros et la méthode de Newton. volume 54 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Berlin.

Demmel, J.W., 1987. On condition numbers and the distance to the nearest ill-posed problem. Numer. Math. 51, 251–289. doi:10.1007/BF01400115.

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

Eigenwillig, A., 2010. Real root isolation for exact and approximate polynomials using Descartes' rule of signs. Ph.D. thesis. Universität des Saarlandes.

Ergür, A.A., Paouris, G., Rojas, J.M., 2018. Probabilistic Condition Number Estimates for Real Polynomial Systems II: Structure and Smoothed Analysis. arXiv:1809.03626.

Ergür, A.A., Paouris, G., Rojas, J.M., 2019. Probabilistic condition number estimates for real polynomial systems I: A broader family of distributions. Found. Comput. Math. 19, 131–157. doi:10.1007/s10208-018-9380-5.

Etayo, U., Beltrán, C., Marzo, J., Ortega-Cerdà, J., 2020. A sequence of polynomials with optimal condition number. Journal of the American Mathematical Society doi:10.1090/jams/956.

Galehouse, B.T., 2009. Topologically accurate meshing using domain subdivision techniques. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)–New York University.

Goldstine, H.H., von Neumann, J., 1951. Numerical inverting of matrices of high order. II. Proc. Amer. Math. Soc. 2, 188–202. doi:10.2307/2032484.

Hemmer, M., Tsigaridas, E.P., Zafeirakopoulos, Z., Emiris, I.Z., Karavelas, M.I., Mourrain, B., 2009. Experimental evaluation and cross-benchmarking of univariate real solvers, in: Proc. of the ACM Int'l Conference on Symbolic Numeric Computation (SNC), pp. 45–54.

Jindal, G., Sagraloff, M., 2017. Efficiently computing real roots of sparse polynomials, in: ISSAC'17—Proceedings of the 2017 ACM International Symposium on Symbolic and Algebraic Computation, ACM, New York. pp. 229–236. doi:10.1145/3087604.3087652.

Khovanskiĭ, A.G., 1991. Fewnomials. volume 88 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI. Trans. from the Russian by S. Zdravkovska.

Lairez, P., 2017. A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time. Found. Comput. Math. 17, 1265–1292. doi:10.1007/s10208-016-9319-7.

Lairez, P., 2020. Rigid continuation paths I. Quasilinear average complexity for solving polynomial systems. J. Amer. Math. Soc. 33, 487–526. doi:10.1090/jams/938.

Livshyts, G., Paouris, G., Pivovarov, P., 2016. On sharp bounds for marginal densities of product measures. Israel Journal of Mathematics 216, 877–889. doi:10.1007/s11856-016-1431-5.

Malajovich, G., 2019. Complexity of sparse polynomial solving: homotopy on toric varieties and the condition metric. Found. Comput. Math. 19, 1–53. doi:10.1007/s10208-018-9375-2.

Malajovich, G., 2020. Complexity of Sparse Polynomial Solving 2: Renormalization. arXiv:2005.01223.

Malajovich, G., Rojas, J.M., 2002. Polynomial systems and the momentum map, in: Foundations of computational mathematics (Hong Kong, 2000). World Sci. Publ., River Edge, NJ, pp. 251–266.

Malajovich, G., Rojas, J.M., 2004. High probability analysis of the condition number of sparse polynomial systems. Theoret. Comput. Sci. 315, 524–555. doi:10.1016/j.tcs.2004.01.006.

Mehlhorn, K., Sagraloff, M., 2011. A deterministic algorithm for isolating real roots of a real polynomial. J. Symbolic Comput. 46, 70–90. doi:10.1016/j.jsc.2010.09.004.

von Neumann, J., Goldstine, H.H., 1947. Numerical inverting of matrices of high order. Bull. Amer. Math. Soc. 53, 1021–1099. doi:10.1090/S0002-9904-1947-08909-6.

Plantinga, S., Vegter, G., 2004. Isotopic approximation of implicit curves and surfaces, in: Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, ACM, New York, NY, USA. pp. 245–254. doi:10.1145/1057432.1057465.

Raffalli, C., 2014. Distance to the discriminant. arXiv:1404.7253.

Renegar, J., 1987. On the efficiency of Newton's method in approximating all zeros of a system of complex polynomials. Math. Oper. Res. 12, 121–148. doi:10.1287/moor.12.1.121.

Rojas, J.M., 2020. Counting positive roots in polynomial-time for systems supported on circuits. arXiv:2012.04868.

Rojas, J.M., Ye, Y., 2005. On solving univariate sparse polynomials in logarithmic time. J. Complexity 21, 87–110. doi:10.1016/j.jco.2004.03.004.

Rouillier, F., Zimmermann, Z., 2004. Efficient isolation of polynomial's real roots. J. Comput. & Applied Math. 162, 33–50.

Rudelson, M., Vershynin, R., 2015. Small ball probabilities for linear images of high-dimensional distributions. Int. Math. Res. Not. IMRN 19, 9594–9617. doi:10.1093/imrn/rnu243.

Sagraloff, M., Mehlhorn, K., 2016. Computing real roots of real polynomials. J. Symbolic Comput. 73, 46–86. doi:10.1016/j.jsc.2015.03.004.

Smale, S., 1997. Complexity theory and numerical analysis, in: Iserles, A. (Ed.), Acta Numerica. Cambridge University Press, Cambridge, UK, pp. 523–551. doi:10.1017/S0962492900002774.

Spielman, D.A., Teng, S.H., 2002. Smoothed analysis of algorithms, in: Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002), Higher Ed. Press, Beijing. pp. 597–606.

Tonelli-Cueto, J., 2019. Condition and Homology in Semialgebraic Geometry. Doctoral thesis. Technische Universität Berlin. DepositOnce Repository. doi:10.14279/depositonce-9453.

Tonelli-Cueto, J., Tsigaridas, E.P., 2020. Condition numbers for the cube. I: Univariate polynomials and hypersurfaces, in: Emiris, I.Z., Zhi, L. (Eds.), ISSAC '20: Int'l Symposium on Symbolic and Algebraic Computation, Kalamata, Greece, July 20-23, 2020, ACM. pp. 434–441. doi:10.1145/3373207.3404054.

Turing, A.M., 1948. Rounding-off errors in matrix processes. Quart. J. Mech. Appl. Math. 1, 287–308. doi:10.1093/qjmam/1.1.287.

Vershynin, R., 2018. High-dimensional probability: An introduction with applications in data science. volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge. doi:10.1017/9781108231596.

Xu, J., Yap, C., 2019. Effective subdivision algorithm for isolating zeros of real systems of equations, with complexity analysis, in: ISSAC'19—Proceedings of the 2019 ACM International Symposium on Symbolic and Algebraic Computation, ACM, New York. pp. 355–362.

Yap, C., 2019. Towards soft exact computation (invited talk), in: International Workshop on Computer Algebra in Scientific Computing, Springer. pp. 12–36.