Functional norms, condition numbers and numerical algorithms in algebraic geometry*

Felipe Cucker†
Dept. of Mathematics
City University of Hong Kong
HONG KONG
macucker@cityu.edu.hk

Alperen A. Ergür
Dept. of Mathematics
University of Texas at San Antonio
San Antonio, Texas 78249, USA
alperen.ergur@utsa.edu

Josué Tonelli-Cueto‡
Inria Paris & IMJ-PRG
Sorbonne Université
Paris, FRANCE
josue.tonelli.cueto@bizkaia.eu

Abstract

In numerical linear algebra, a well-established principle is to choose a norm that exploits the structure of the problem in hand in order to optimize accuracy-computational complexity trade-off. In numerical polynomial algebra, a single norm (attributed to Bombieri and Weyl) dominates the entire literature. This article initiates the use of L_p norms for numerical algebraic geometry, with an emphasis on L_∞. This classical idea yields strong improvements in the analysis of the number of steps performed by numerous iterative algorithms. In particular, we exhibit three algorithms where, despite the complexity of computing L_∞-norm, the usage of L_p-norms substantially reduce computational complexity: a subdivision-based algorithms in real algebraic geometry for computing the homology of semialgebraic sets, a well-known meshing algorithm in computational geometry, and the computation of zeros of systems of complex quadratic polynomials (a particular case of Smale's 17th problem).

1 Introduction

It is common in numerical linear algebra to chose a matrix norm to exploit the structure of the problem at hand and optimize computational efficiency. In numerical polynomial algebra, however, a norm introduced by Weyl [54] dominates the entire literature. Two features of this norm explain this omnipresence: Firstly, the cost of computing it (assuming dense or sparse representation of the involved polynomials) is optimal. Secondly, it is unitarily (respectively orthogonally if we consider real polynomials) invariant.

In this paper we initiate the use of L_p norms in numerical algebraic geometry with a focus on an L_∞-norm. We should say from the outset that, currently, we don't have an

*This work was supported by the Einstein Foundation Berlin.

†Partially supported by GRF grant CityU 11300220.

‡Supported by a postdoctoral fellowship of the 2020 "Interaction" program of the Fondation Sciences Mathématiques de Paris. Partially supported by the ANR JCJC GALOP (ANR-17-CE40-0009), the PGMO grant ALMA, and the PHC GRAPE.

efficient way to approximate $\| \|_\infty$. For polynomials in $n+1$ homogeneous variables whose degrees are bounded by $\mathbf{D}$ our current fastest algorithm doing so takes time polynomial in $\mathbf{D}$ and exponential in n. The computation of $\| \|_\infty$ amounts, however, to a polynomial optimization problem and efficient algorithms exist for particular classes of polynomials. This is the case e.g., with sums of squares [38, 9], sparse polynomials [30, 20], and other structures [5]. Unrestricted efficient algorithms cannot be expected to be designed because it is well-known that polynomial optimization reduces to the feasibility problem over the reals and the latter is $\mathrm{NP}_{\mathbb{R}}$-complete. The fact that for most applications we only need a coarse approximation of $\| \|_\infty$, however, allows for some optimism.

The major claim of our paper is that, notwithstanding its generally high cost, the use of the ∞-norm reduces the number of iterations in various numerical algorithms and, in some cases, achieves a reduction in total complexity. To show this, in a nutshell, we define a version of the relevant condition number that scales with $\| \|_\infty$ instead of with the Weyl norm and show that the condition-based complexity estimates obtained with these new condition numbers are essentially the same as those obtained with the original ones.

Then, we eliminate the occurrences of condition numbers in the cost bounds via considering random data: here is where the complexity reductions take place. The reason behind these reductions lies on the value of the norm itself for random systems f. In this context, and whereas for the Weyl norm we have $\|f\|_W \sim \binom{n+\mathbf{D}}{n}^{\frac{1}{2}}$, for the ∞-norm we have $\|f\|_\infty \sim \sqrt{n}\log\mathbf{D}$. This drop passes to the condition numbers defined via $\| \|_\infty$ (instead of $\| \|_W$) and, eventually, to the complexity estimates. We show this for a few algorithms in three different settings.

Firstly, in § 4.1, we consider a family of algorithms, we refer to them as grid-based, that solve various problems in real algebraic and semialgebraic geometry. The best numerical algorithms for these problems have exponential complexity. In §4.1 we replace the Weyl norm by $\| \|_\infty$ in the design of one such algorithm (to compute Betti numbers) and in §4.3 we show a decrease in its complexity. This has no extra cost because there is only one norm computation and it is done, so to speak, along the way. The gain in the reduction of the number of iterations thus directly yields a reduction in total complexity (see Corollary 4.29).

Secondly, in §4.2, we consider the Plantinga-Vegter algorithm as it is described and analyzed in [22]. Again, replacing the Weyl norm by $\| \|_\infty$ in the algorithm's design results in a gain in efficiency.And again, the computation of $\| \|_\infty$ is not a burden as it done only once and its cost is dominated by that of the rest of the algorithm. The Plantinga-Vegter algorithm is usually considered with $n = 2$ or $n = 3$. Remark 4.33 exhibits the improvement achieved on average complexity for these two cases. For larger values of n the improvement is more substantial. (see Remark 4.33).

Thirdly, in Section 5, we consider the problem of computing a zero of a system of complex quadratic equations. For this question, a particular case of Smale's 17th problem, we consider the algorithms proposed in [8, 12] and, again, we design versions of them where the Weyl norm is replaced by $\| \|_\infty$. Again, this results in a small, but measurable, gain in efficiency (from n^7 to $n^{6.875}$). A crucial fact to achieve this is that, even though n is general, we can find an efficient way to compute $\| \|_\infty$ using the fact that $\mathbf{D} = 2$.

In all three cases, we are able to show a clear reduction on the expected number of iterations of some iterative algorithm. We believe that this is a common pattern; Theorem 5.5 is a case in point. But, in general, this reduction does not immediately translate into a reduction in total complexity. This puts the focus on the search for efficient algorithms to approximate $\| \|_\infty$ -note that our setting allows rough but fast approximations-, and on the complexity and accuracy trade-off by using numerous L_p-norms.

Before showing the results above, in Section 2, we define the norms we will consider in the paper, work out various examples with the goal of providing a clear understanding of the differences between these norms and the Weyl norm, and prove several properties for them. Then, in Section 3, we define L_p-norm versions M and K of the well-known condition numbers μ_{norm} and κ (for complex and real problems, respectively) and prove that the main properties of μ_{norm} and κ —those allowing them to feature in condition-based cost estimates— hold for M and K as well.

We conclude the paper, in Section 6, with a minor digression. Because a natural habitat for functional norms are spaces of continuous functions we consider extensions of the real condition number κ to the space $C^1[q] := C^1(\mathbb{S}^n, \mathbb{R}^q)$ and we prove (somehow unexpectedly) Condition Number Theorems for these extensions. We do not analyze algorithms here. We nonetheless point out that substantial literature on algorithms on spaces of continuous functions exists [52, 45, 43] where these theorems might be useful.

Contents

Acknowledgments. The second author is grateful to Hakan and Bahadır Ergür for their cheerful response to his sudden all-day availability throughout the pandemic times. The third author is grateful to Evgenia Lagoda for moral support and Gato Suchen for useful suggestions for this paper.

2 Norms for polynomials

Let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$. Let also $n, d \in \mathbb{N}$, $n, d \geq 1$. We denote by $\mathcal{H}_d^{\mathbb{F}}[1]$ the linear space of homogeneous polynomials of degree d in the $n+1$ variables $X_0, X_1, \ldots, X_n$ with coefficients in $\mathbb{F}$. Let $\boldsymbol{d} = (d_1, \ldots, d_q) \in \mathbb{N}^q$ and $n \in \mathbb{N}$ as above. We denote by $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{F}}[q]$ the space $\mathcal{H}_{d_1}^{\mathbb{F}}[1] \times \cdots \times \mathcal{H}_{d_q}^{\mathbb{F}}[1]$. If $\mathbb{F}$ is clear from the context, or if it is not relevant to the argument, we will omit the superscript. We will use the following conventions for dimension counting:

$$N_i := \binom{n+d_i}{d_i} = \dim_{\mathbb{F}} \mathcal{H}_{d_i}^{\mathbb{F}}[1] \quad \text{and} \quad N := \sum_{i=1}^{q} \binom{n+d_i}{d_i} = \dim_{\mathbb{F}} \mathcal{H}_{\boldsymbol{d}}^{\mathbb{F}}[q].$$

We also use $\mathbf{D} := \max\{d_1, \ldots, d_q\}$ and denote by Δ the $q \times q$ diagonal matrix with d_i in its ith diagonal entry.

In all what follows, $\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} \mid \|x\| = 1\}$ will be the (real) n-sphere and $\mathbb{P}^n := \mathbb{C}^{n+1}/\mathbb{C}^*$ the complex projective space of dimension n. We note that there will be no ambiguity, as the sphere is the usual space to work with real polynomials and the projective space the usual one for complex polynomials.

Remark 2.1. In what follows, we will write $z \in \mathbb{P}^n$ instead of $[z] \in \mathbb{P}^n$ and we will assume that the representative $z \in \mathbb{C}^{n+1}$ always satisfies $\|z\| = 1$. This simplifies the form of many of our definitions. This convention can be made w.l.o.g. as every point in $\mathbb{P}^n$ has a representative of norm 1.

2.1 Euclidean norms

The simplest norm considered on $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ is the one induced by the standard Euclidean inner product in monomial basis. Every $f \in \mathcal{H}_d^{\mathbb{F}}[1]$ can be uniquely represented as

$$f = \sum_{|\alpha|=d} f_\alpha X^\alpha \tag{2.1}$$

where $\alpha = (\alpha_0, \ldots, \alpha_n) \in \mathbb{N}^{n+1}$ and $|\alpha| = \alpha_0 + \cdots + \alpha_n$. The norm induced by the standard Euclidean inner product is therefore

$$\|f\|_{\mathrm{std}} := \sqrt{\sum_{|\alpha|=d} |f_\alpha|^2}.$$

For $f = (f_1, \ldots, f_q) \in \mathcal{H}_{\boldsymbol{d}}[q]$ the norm extends as $\|f\|_{\mathrm{std}}^2 := \|f_1\|_{\mathrm{std}}^2 + \cdots + \|f_q\|_{\mathrm{std}}^2$.

The most commonly used norm on $\mathcal{H}_{\boldsymbol{d}}[q]$ is the *Weyl norm*. For a polynomial as in (2.1), this is given by

$$\|f\|_W := \sqrt{\sum_{|\alpha|=d} \binom{d}{\alpha}^{-1} |f_\alpha|^2}$$

where $\binom{d}{\alpha}$ is the multinomial coefficient $\frac{d!}{\alpha_0!\dots\alpha_n!}$. Again, for $f \in \mathcal{H}_{\boldsymbol{d}}[q]$ this extends by $\|f\|_W^2 := \|f_1\|_W^2 + \cdots + \|f_q\|_W^2$. The Weyl norm is also induced by an inner product, and this inner product is invariant under the action of the unitary group (respectively the orthogonal group when the underlying field is $\mathbb{R}$). It is straightforward to check that, for $f \in \mathcal{H}_{\boldsymbol{d}}[q]$,

$$\|f\|_W \leq \|f\|_{\mathrm{std}} \leq \max_{i\leq q} \max_{|\alpha|=d_i} \binom{d_i}{\alpha} \|f\|_W.$$

Here, and in all what follows, for any $x \in \mathbb{S}^n$ and $f \in \mathcal{H}_{\boldsymbol{d}}[q]$, $\mathrm{D}_x f : \mathrm{T}_x\mathbb{S}^n \to \mathbb{R}^q$ is the derivative of f at x restricted to the tangent space $\mathrm{T}_x\mathbb{S}^n$ of $\mathbb{S}^n$ at x. A similar convention applies in the complex case replacing $\mathbb{S}^n$ and $\mathrm{T}_x\mathbb{S}^n$ by $\mathbb{P}^n$ and $\mathrm{T}_z\mathbb{P}^n$. The following property (see [13, Prop. 16.16]) is one of the most important properties of the Weyl norm from the viewpoint of the complexity of numerical algorithms.

Proposition 2.2. *For all $x \in \mathbb{S}^n$ the map*

$$\mathcal{H}_{\boldsymbol{d}}[q] \ni f \mapsto \mathrm{ev}_x f := \left(f(x), \Delta^{-\frac{1}{2}} \mathrm{D}_x f\right)$$

is an orthogonal projection from $\mathcal{H}_{\boldsymbol{d}}[q]$ endowed with the Weyl norm onto $\mathbb{R}^q \times \mathrm{T}_x\mathbb{S}^n \simeq \mathbb{R}^{q+n}$ equipped with the standard Euclidean norm. An analogous statement holds in the complex case. □

2.2 Functional norms

We will consider functional norms that arise from evaluating polynomials at points on the sphere. One might consider other norms (as we do in Section 6), but L_p-norms suffice for obtaining the computational improvements we aim for.

We will consider the two following classes of L-norms on $\mathcal{H}_{\boldsymbol{d}}[q]$:

($\mathbb{R}$) *Real L_p-norm*: For $p \in [1, \infty]$,

$$\|f\|_p^{\mathbb{R}} := \begin{cases} \max_{x\in\mathbb{S}^n} \|f(x)\|_\infty = \max_{x\in\mathbb{S}^n} \max_i |f_i(x)| & \text{if } p = \infty \\ \left(\mathop{\mathbb{E}}_{x\in\mathbb{S}^n} \|f(x)\|_p^p\right)^{1/p} = \left(\mathop{\mathbb{E}}_{x\in\mathbb{S}^n} \left(\sum_{i=1}^q |f_i(x)|^p\right)\right)^{1/p} & \text{otherwise} \end{cases}$$

where the expectations are taken over the uniform distribution of the n-dimensional sphere $\mathbb{S}^n \subseteq \mathbb{R}^{n+1}$.

($\mathbb{C}$) *Complex L_p-norm*: For $p \in [1, \infty]$,

$$\|f\|_p^{\mathbb{C}} := \begin{cases} \max_{z\in\mathbb{P}^n} \|f(z)\|_\infty = \max_{z\in\mathbb{P}^n} \max_i |f_i(z)| & \text{if } p = \infty \\ \left(\mathop{\mathbb{E}}_{z\in\mathbb{P}^n} \|f(z)\|_p^p\right)^{1/p} = \left(\mathop{\mathbb{E}}_{z\in\mathbb{P}^n} \left(\sum_{i=1}^q |f_i(z)|^p\right)\right)^{1/p} & \text{otherwise} \end{cases}$$

where the expectations are taken over the uniform distribution of the the complex n-dimensional projective space $\mathbb{P}^n := \mathbb{P}^n_{\mathbb{C}}$.

In general, we will omit the superscript when the context is clear. It will be common for us to work with the norms $\| \, \|_p^{\mathbb{R}}$ in $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and the norms $\| \, \|_p^{\mathbb{C}}$ in $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$.[1]

Our definition has some arbitrary choices. These are motivated by the following two properties:

(D) For $p \in [1,\infty]$ and $f \in \mathcal{H}_{\boldsymbol{d}}[q]$,

$$\|f\|_p^{\mathbb{R}} = \left\| \left(\|f_1\|_p^{\mathbb{R}}, \dots, \|f_q\|_p^{\mathbb{R}} \right) \right\|_p \quad \text{and} \quad \|f\|_p^{\mathbb{C}} = \left\| \left(\|f_1\|_p^{\mathbb{C}}, \dots, \|f_q\|_p^{\mathbb{C}} \right) \right\|_p .$$

This commutativity is why we take the p-average of the p-norm of $f(x)$ instead of taking the p-average of a fixed norm.

(I) We have actions of the qth power of the (real) orthogonal group, $\mathscr{O}(n+1)^q$, on $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$, given by $(A,f) \mapsto (f_i^{A_i}) := (f_i(A_iX))$. Similarly, we have an action of the qth power of the unitary group, $\mathscr{U}(n+1)^q$, on $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$. The norms $\| \, \|_p^{\mathbb{R}}$ and $\| \, \|_p^{\mathbb{C}}$ are invariant under these actions.

We perform some simple computations to have a better grasp on the introduced norms.

Example 2.3 (Monomials). We consider the value of the norms for a monomial $X^\alpha \in \mathcal{H}_d[1]$ of degree d. In this case we have that for $p \in [1,\infty)$,

$$\|X^\alpha\|_p^{\mathbb{R}} = \left(\frac{\Gamma\left(\frac{n+1}{2}\right) \prod_{i=0}^n \Gamma\left(\frac{p\alpha_i+1}{2}\right)}{\pi^{\frac{n+1}{2}} \Gamma\left(\frac{pd+n+1}{2}\right)} \right)^{\frac{1}{p}} \quad \text{and} \quad \|X^\alpha\|_p^{\mathbb{C}} = \left(n! \frac{\prod_{i=0}^n \Gamma\left(\frac{p\alpha_i}{2}+1\right)}{\Gamma\left(\frac{pd}{2}+n+1\right)} \right)^{\frac{1}{p}}$$

where Γ is Euler's Gamma function, and that

$$\|X^\alpha\|_\infty^{\mathbb{R}} = \|X^\alpha\|_\infty^{\mathbb{C}} = \prod_{i=0}^n \left(\frac{\alpha_i}{d}\right)^{\frac{\alpha_i}{2}} = \sqrt{\frac{1}{d^d} \prod_{i=0}^n \alpha_i^{\alpha_i}}.$$

For the calculations of L_p-norms of monomials we refer the reader to [33]. Although the calculation is only illustrated over the reals in the reference, the complex case is similar. For the second one, note that for monomials real and complex ∞-norms are equivalent. Once this is clear, we are just using the method of Lagrange multipliers to compute the maximum over the sphere.

Example 2.4 (Linear functions). Let $\mathbb{1} = (1,1,\dots,1) \in \mathbb{N}^q$ and $f \in \mathcal{H}_{\mathbb{1}}[q]$. Then f can be identified with a matrix A of size $q \times (n+1)$. We can see that

$$\|f\|_\infty = \|A\|_{2,\infty} := \sup_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_2}$$

where $\| \, \|_{2,\infty}$ is the operator norm where the domain vector space has the usual Euclidean norm $\| \, \|_2$ and the codomain the ∞-norm $\| \, \|_\infty$.

For $p \in [1,\infty)$,

$$\|f\|_p^{\mathbb{R}} = \|X_0\|_p^{\mathbb{R}} \left\| \left(\|A^1\|_2, \dots, \|A^q\|_2 \right) \right\|_p \quad \text{and} \quad \|f\|_p^{\mathbb{C}} = \|X_0\|_p^{\mathbb{C}} \left\| \left(\|A^1\|_2, \dots, \|A^q\|_2 \right) \right\|_p$$

where A^i is the ith row of A and X_0 is a variable (and hence $\|X_0\|_p^{\mathbb{F}}$ is given by the expressions in Example 2.3). Note that $\left\| \left(\|A^1\|_2, \dots, \|A^q\|_2 \right) \right\|_p$ is just the p-norm of the vector of 2-norms of the rows of A.

[1] Observe, however, that the $\| \, \|_p^{\mathbb{R}}$ are also norms for $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$ since a complex homogeneous polynomial cannot vanish on the real sphere without being zero.

Example 2.5 (Sum of squares). Let $f := \sum_{i=0}^n X_i^2 \in \mathcal{H}_2[1]$. As this function is constant on the real sphere, we have that for all $p \in [1,\infty]$,

$$\|f\|_p^{\mathbb{R}} = 1.$$

However, on $\mathbb{P}^n$, f does not behave as a constant function as it has a positive dimensional zero set. Again, arguing as in [33], we can conclude that

$$\|f\|_p^{\mathbb{C}} = \left(\frac{1}{\pi^{n+1}}\frac{n!}{(n+p)!}\int_{z\in\mathbb{C}^{n+1}} |f(z)|^p e^{-|z|^2}\right)^{\frac{1}{p}}$$

for $p \in [1,\infty)$. Now, if p is even, we can obtain the expression

$$\|f\|_p^{\mathbb{C}} = \left(\binom{n+2}{2}^{-1} \sum_{\substack{\alpha\in\mathbb{N}^{n+1}\\ |\alpha|=p/2}} \binom{p/2}{\alpha}^2 \binom{p}{2\alpha}^{-1}\right)^{\frac{1}{p}},$$

after writing $|f(z)|^p = f(z)^{\frac{p}{2}}\overline{f(z)}^{\frac{p}{2}}$, expanding and using separation of variables. In particular, for $p=2$, we obtain that

$$\|f\|_2^{\mathbb{C}} = \sqrt{\frac{2}{n+2}} \neq 1.$$

This shows how the norms $\|\ \|_p^{\mathbb{C}}$ may be smaller than their corresponding norm $\|\ \|_p^{\mathbb{R}}$ for $p \in [1,\infty)$.

Example 2.6 (Cosine polynomials). Let $d \geq 2$ and consider the family of homogenous polynomials

$$c_d := \sum_{k=0}^{\lfloor d/2\rfloor} \binom{d}{2k}(-1)^k X^{d-2k}Y^{2k} = \frac{1}{2}(X+iY)^d + \frac{1}{2}(X-iY)^d \in \mathcal{H}_d[1].$$

Since $c_d(\cos\theta,\sin\theta) = \cos d\theta$, we have that

$$\|c_d\|_\infty^{\mathbb{R}} = 1.$$

Also, c_d is unitarily equivalent to $2^{\frac{d}{2}-1}(X^d+Y^d)$. Hence

$$\|c_d\|_\infty^{\mathbb{C}} = 2^{\frac{d}{2}-1},$$

since $\|X^d+Y^d\|_\infty^{\mathbb{C}} = 1$ for $d \geq 2$. This shows that for degrees $d \geq 3$, the norms $\|\ \|_\infty^{\mathbb{R}}$ and $\|\ \|_\infty^{\mathbb{C}}$ disagree on real polynomials.

The following proposition lists simple inequalities between the functional norms.

Proposition 2.7. *Let $1 \leq p < p' < \infty$ and $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$. Then for all $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{F}}[q]$, the following inequalities hold:*

$$\frac{1}{q^{\frac{1}{p}}}\|f\|_p^{\mathbb{F}} \leq \frac{1}{q^{\frac{1}{p'}}}\|f\|_{p'}^{\mathbb{F}} \leq \|f\|_\infty^{\mathbb{F}} \leq \|f\|_\infty^{\mathbb{C}}.$$

Sketch of proof. It is a direct consequence of the inequalities between p-means. $\square$

The Bombieri-Weyl norm is essentially a scaled version of the complex L_2 norm.

Proposition 2.8. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$, then*

$$\|f\|_W = \sqrt{\sum_{i=1}^{q} N_i \left(\|f_i\|_2^{\mathbb{C}}\right)^2}.$$

In particular, for $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[1]$,

$$\|f\|_W^{\mathbb{C}} = \sqrt{N}\|f\|_2^{\mathbb{C}}.$$ $\square$

Sketch of proof. We only need to show this in the case $q = 1$. Now, both the Bombieri-Weyl norm and the complex L_2-norm are unitarily invariant Hermitian norms of $\mathcal{H}_d^{\mathbb{C}}$. For the Bombieri-Weyl norm, see [13, Theorem 16.3]; for the complex L_2-norm, this is property (I). Since $\mathcal{H}_d^{\mathbb{C}}$ is an irreducible representation of $\mathscr{U}(n+1)$, this means that the two norms are equal up to a constant. Using Example 2.3 with $f = X_0^d$, one can check that this constant is $\sqrt{N}$. $\square$

From Proposition 2.2 we get the following result.

Proposition 2.9. *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $f \in \mathcal{H}_{\boldsymbol{d}}[q]$. Then for all $p \geq 2$,*

$$\|f\|_p^{\mathbb{F}} \leq \|f\|_W.$$

Sketch of proof. For each $x \in \mathbb{S}^n$,

$$\|f(x)\|_p \leq \|f(x)\| \leq \|f\|_W$$

where the the first inequality is known and the second follows from Proposition 2.2. $\square$

We finish this subsection by noting how the L_∞-norms relate to the Weyl norm. We note that this is very related to the so-called best rank-one approximation of a symmetric tensor [1, 55], and the inequality for the real case below was already present in [55, Theorem 2.4].

Proposition 2.10. *Let $f \in \mathcal{H}_{\boldsymbol{d}}[q]$. Then*

$$\|f\|_\infty^{\mathbb{C}} \leq \|f\|_W \leq \sqrt{N}\|f\|_\infty^{\mathbb{C}}.$$

If $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$. Then

$$\|f\|_\infty^{\mathbb{R}} \leq \|f\|_W \leq (n+1)^{\frac{\mathbf{D}}{2}}\|f\|_\infty^{\mathbb{R}}.$$

Proof. The first part follows from Proposition 2.8 and 2.9. The left-hand side of the second part uses Proposition 2.9.

Now, for $f \in \mathcal{H}_d[1]$, Corollary 2.20 implies that for each α, $|f_\alpha| = \left\|\frac{1}{\alpha!}\overline{\mathrm{D}}_x f\right\| \leq \binom{d}{\alpha}$. The right-hand inequality follows from here. $\square$

Example 2.11. Proposition 2.10 is almost optimal for $n = 1$. In [1], it was shown that for the cosine polynomials c_d of Example 2.6 we have

$$\|c_d\|_W = 2^{\frac{d-1}{2}}$$

and that c_d is the real polynomial of real L_∞ norm one with largest Weyl norm. Curiously, in this case, the Weyl norm and the complex L_∞ are almost equal, the former being the latter times $\sqrt{2}$.

2.3 Kellogg's Theorem

We will denote by $\overline{\mathrm{D}}$ the operation of taking all partial derivatives with respect to all variables, i.e., $f \mapsto \overline{\mathrm{D}}f$ is a linear map $\mathcal{H}_{\boldsymbol{d}}[q] \to \mathcal{H}_{\boldsymbol{d}-\mathbb{1}}[(n+1)q]$ and, for $x \in \mathbb{F}^{n+1}$, $\overline{\mathrm{D}}_x f : \mathbb{F}^{n+1} \to \mathbb{F}^q$ is a linear map. We will write $\overline{\mathrm{D}}_X f$, with capital X, to emphasize that we view $\overline{\mathrm{D}}_X f$ as a polynomial tuple in $\mathcal{H}_{\boldsymbol{d}-\mathbb{1}}[(n+1)q]$, and $\overline{\mathrm{D}}_x f$, with lowercase x, to emphasize that we view $\mathrm{D}_x f$ as the linear map $\mathbb{F}^{n+1} \to \mathbb{F}^q$ defined at the point x. We also recall that $\mathrm{D}_x f$ is the tangent map $\mathrm{T}_x\mathbb{S}^n \to \mathbb{R}^q$ in the real case, and the tangent map $\mathrm{T}_x\mathbb{P}^n \to \mathbb{C}^q$ in the complex case.

The following result plays the role of Proposition 2.2 for the infinity norm instead of the Weyl one. It is a reformulation of a well-known inequality proved in [36].

Theorem 2.12 (Kellogg's Inequality). *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, $f \in \mathcal{H}^{\mathbb{F}}_{\boldsymbol{d}}[q]$ and $v \in \mathbb{F}^{n+1}$, then*

$$\left\|\Delta^{-1}\overline{\mathrm{D}}_X f v\right\|_\infty^{\mathbb{F}} \leq \|f\|_\infty^{\mathbb{F}}\|v\|.$$

Corollary 2.13. *Let $f \in \mathcal{H}^{\mathbb{R}}_{\boldsymbol{d}}[q]$ and $x \in \mathbb{S}^n$. Then*

$$\max\left\{\|f(x)\|_\infty, \|\Delta^{-1}\mathrm{D}_x f\|_{2,\infty}\right\} \leq \|f\|_\infty^{\mathbb{R}}.$$

Corollary 2.14. *Let $f \in \mathcal{H}^{\mathbb{C}}_{\boldsymbol{d}}[q]$ and $z \in \mathbb{P}^n$. Then*

$$\max\left\{\|f(z)\|_\infty, \left\|\Delta^{-1}\mathrm{D}_z f\right\|_{2,\infty}\right\} \leq \|f\|_\infty^{\mathbb{C}}.$$

Remark 2.15. We note that the left-hand sides in Corollaries 2.13 and 2.14 are not optimal. In general, we have that

$$\|\Delta_{-1}\overline{\mathrm{D}}_x f\|_{2,\infty} = \max_i \sqrt{|f_i(x)|^2 + \frac{1}{d_i^2}\|\mathrm{D}_x f_i\|_{2,\infty}^2}.$$

The following examples show how the bound of Theorem 2.12 looks like in a few particular cases.

Example 2.16. Consider the cosine polinomials c_d of Example 2.6. A direct computation shows that

$$\frac{1}{d}\mathrm{D}_X c_d v = v_X c_{d-1} - v_Y s_{d-1}$$

where $s_{d-1} := -\frac{i}{2}(X+iY)^{d-1} + \frac{i}{2}(X-iY)$ is the *sine polynomial* for which $s_d(\cos\theta, \sin\theta) = \sin d\theta$.

In the real case, this gives

$$\left\|\frac{1}{d}\mathrm{D}_X c_d v\right\|_\infty^{\mathbb{R}} = \|v\| = \|c_d\|_\infty^{\mathbb{R}}\|v\|,$$

using the Cauchy-Schwarz inequality. In the complex case, $\frac{1}{d}\mathrm{D}_X c_d v = v_X c_{d-1} - v_Y s_{d-1}$ is unitarily equivalent to

$$\frac{2^{\frac{d-1}{2}}}{d}\left[(v_x - iv_Y)X^{d-1} + (v_x + iv_Y)Y^{d-1}\right].$$

Now, $\left\|(v_x - iv_Y)x^{d-1} + (v_x + iv_Y)y^{d-1}\right\| \leq \sqrt{2}\|v\|(|x|^{d-1} + |y|^{d-1}) \leq \|v\|$ for $d \leq 3$ and v real, when $|x|^2 + |y|^2 \leq 1$. Thus

$$\left\|\frac{1}{d}\mathrm{D}_X c_d v\right\|_\infty^{\mathbb{C}} = \frac{2^d}{d}\|v\| = \frac{\sqrt{2}}{d}\|c_d\|_\infty^{\mathbb{C}}\|v\|.$$

This shows that the real version of Kellogg's theorem is tight for c_d, but the complex version is not.

Example 2.17. The reverse situation is true for the polynomial X_0^d. One can see that

$$\left\|\frac{1}{d}\mathrm{D}_X X_0^d e_0\right\|_\infty^{\mathbb{C}} = \|X_0^d\|_\infty^{\mathbb{C}}.$$

Now it is the complex Kellogg's theorem the one which is tight. We note, however, that one might still improve Corollary 2.14. For example, is it possible to substitute Δ by $\Delta^{\frac{1}{2}}$ in this corollary?

Remark 2.18. The examples here motivates development of a randomized Kellog's theorem that holds with high probability for random polynomials and has a tighter right hand side.

Proof of Theorem 2.12. We only prove the real case. The complex case is proven in an analogous way (see [36, §8] for the complex version of the results we use in the real case).

By [36, Theorem IV], we have that for all i and all $x \in \mathbb{S}^n$,

$$\left|\overline{\mathrm{D}}_x f_i v\right| \leq d_i \|f_i\|_\infty^{\mathbb{R}}\|v\|,$$

since $\overline{\mathrm{D}}_x f_i v$ is the directional derivative of f at x in the direction of v scaled by its norm, $\|v\|$. Therefore for all $x \in \mathbb{S}^n$,

$$\left\|\Delta^{-1}\overline{\mathrm{D}}_x f v\right\|_\infty = \max_i \frac{1}{d_i}\left|\overline{\mathrm{D}}_x f v\right| \leq \max_i \|f_i\|_\infty^{\mathbb{R}}\|v\| = \|f\|_\infty^{\mathbb{R}}\|v\|.$$

Now, $\left\|\Delta^{-1}\overline{\mathrm{D}}_X f v\right\|_\infty^{\mathbb{R}} = \max_{x\in\mathbb{S}^n} \|\Delta^{-1}\overline{\mathrm{D}}_x f v\|_\infty$ by definition of $\| \ \|_\infty^{\mathbb{R}}$, so we are done. □

Remark 2.19. We note that the application of [36, Theorem IV] using the scaling with the diagonal matrix was not used in [31, Theorem 2.4]. This can be used to improve by a factor of the degree some of the bounds there.

Proof of Corollaries 2.13 and 2.14. We only prove Corollary 2.13, the proof of Corollary 2.14 is essentially the same.

Recall that, by Euler's formula for homogeneous functions,

$$\Delta^{-1}\overline{\mathrm{D}}_x f x = f(x). \tag{2.2}$$

In this way, for $x \in \mathbb{S}^n$, $\lambda \in \mathbb{R}$ and $w \in \mathrm{T}_x\mathbb{S}^n = x^\perp$,

$$\Delta^{-1}\overline{\mathrm{D}}_x f(\lambda x + w) = \lambda f(x) + \Delta^{-1}\mathrm{D}_x f w.$$

When $\lambda x + w = x$, this expression yields $f(x)$; and when $\lambda x + w = w$, it yields $\Delta^{-1}\mathrm{D}_x f w$. In this way,

$$\max_{\lambda x + w \neq 0} \frac{\|\Delta^{-1}\overline{\mathrm{D}}_x f(\lambda x + w)\|_\infty}{\sqrt{|\lambda|^2 + \|w\|^2}} \geq \max\left\{\|f(x)\|_\infty, \max_{v\in\mathrm{T}_x\mathbb{S}^n\setminus 0} \frac{\|\Delta^{-1}\mathrm{D}_x v\|_\infty}{\|v\|}\right\}.$$

The left-hand side is bounded by $\|f\|_\infty^{\mathbb{R}}$ by Theorem 2.12, and the right-hand side equals $\max\{\|f(x)\|_\infty, \|\Delta^{-1}\mathrm{D}_x f\|_{2,\infty}\}$. Thus the desired inequality follows. □

The following corollaries (which are closely related to [55, Theorem 2.1]) will be useful later. For a real k-multilinear map $A : (\mathbb{R}^n)^k \to \mathbb{R}^q$, we define

$$\|A\|_{2,\infty}^{\mathbb{R}} := \sup_{v_1,\dots,v_k\neq 0} \frac{\|A(v_1,\dots,v_k)\|_\infty}{\|v_1\|\cdots\|v_k\|}. \tag{2.3}$$

We define $\|A\|_{2,\infty}^{\mathbb{C}}$ for a complex k-multilinear map $A : (\mathbb{C}^n)^k \to \mathbb{C}^q$ in a similar manner. Note that, by the following corollaries and Example 2.6,

$$\left\|\frac{1}{k!}\overline{\mathrm{D}}_X^k c_k\right\|_{2,\infty}^{\mathbb{C}} = \|c_k\|_\infty^{\mathbb{C}} = 2^{\frac{k}{2}-1} > 1 = \|c_k\|_\infty^{\mathbb{R}} = \left\|\frac{1}{k!}\overline{\mathrm{D}}_X^k c_k\right\|_{2,\infty}^{\mathbb{R}},$$

so for real A, $\|A\|_{2,\infty}^{\mathbb{R}}$ and $\|A\|_{2,\infty}^{\mathbb{C}}$ are not necessarily equal and can differ by a factor exponential in k.

Corollary 2.20. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $x \in \mathbb{S}^n$. Then, for all $k \geq 1$ and $v_1,\dots,v_k \in \mathbb{R}^{n+1}$,*

$$\left\|\frac{1}{k!}\Delta^{-1}\overline{\mathrm{D}}_x^k f(v_1,\dots,v_k)\right\|_\infty \leq \frac{1}{k}\binom{\mathbf{D}-1}{k-1}\|f\|_\infty^{\mathbb{R}}\|v_1\|\cdots\|v_k\|.$$

In particular, $\left\|\frac{1}{k!}\Delta^{-1}\overline{\mathrm{D}}_x^k f\right\|_{2,\infty} \leq \frac{1}{k}\binom{\mathbf{D}-1}{k-1}\|f\|_\infty^{\mathbb{R}}$.

Corollary 2.21. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$ and $z \in \mathbb{P}^n$. Then, for all $k \geq 1$ and $v_1,\dots,v_k \in \mathbb{C}^{n+1}$,*

$$\left\|\frac{1}{k!}\Delta^{-1}\overline{\mathrm{D}}_z^k f(v_1,\dots,v_k)\right\|_\infty \leq \frac{1}{k}\binom{\mathbf{D}-1}{k-1}\|f\|_\infty^{\mathbb{C}}\|v_1\|\cdots\|v_k\|.$$

In particular, $\left\|\frac{1}{k!}\Delta^{-1}\overline{\mathrm{D}}_z^k f\right\|_{2,\infty} \leq \frac{1}{k}\binom{\mathbf{D}-1}{k-1}\|f\|_\infty^{\mathbb{C}}$.

Proof of Corollaries 2.20 and 2.21. Both corollaries follow from Theorem 2.12 by induction, followed by an application of Corollaries 2.13 and 2.14. □

Remark 2.22. Although the results in this section were proved only for $\|\ \|_\infty^{\mathbb{F}}$, some of them can be generalized to other norms. For example, similar results can be obtained for $\|\ \|_2^{\mathbb{R}}$ (see [47]) and certainly for other norms. We defer to future work the application of these extensions to the analysis of numerical algorithms in algebraic geometry. We also note that Kellogg's theorem (in the form of Corollary 2.13) can be generalized to smooth real algebraic varieties other than the sphere (see [10]).

3 Condition numbers for the L_∞-norm

In this section, we will consider condition numbers that capture "how singular" an element in $\mathcal{H}_{\boldsymbol{d}}[q]$ is at a point $x \in \mathbb{S}^n$. We will define condition numbers and develop a geometric understanding of them for the L_∞-norms defined in the preceding section.

We define the *real local condition number* —of $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ at $x \in \mathbb{S}^n$— as

$$\mathsf{K}(f,x) := \frac{\sqrt{q}\|f\|_\infty^{\mathbb{R}}}{\max\left\{\|f(x)\|, \|\mathrm{D}_x f^\dagger \Delta\|_{2,2}^{-1}\right\}} \tag{3.1}$$

and the *real global condition number* —of $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$— as

$$\mathsf{K}(f) := \sup_{y\in\mathbb{S}^n} \mathsf{K}(f,y). \tag{3.2}$$

And we define the *complex local condition number* —of $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$ at $\zeta \in \mathbb{P}^n$— as

$$\mathsf{M}(f,\zeta) = \sqrt{q}\|f\|_\infty^{\mathbb{C}} \left\|\mathrm{D}_\zeta f^\dagger \Delta\right\|_{2,2} \tag{3.3}$$

and the *complex global condition number* —of $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$ (with $q \leq n$)— as

$$\mathsf{M}(f) := \sup\{\mathsf{M}(f,\zeta) \mid \zeta \in \mathbb{P}^n,\ f(\zeta) = 0\}. \tag{3.4}$$

Remark 3.1. By convention, we assume that $\|A^\dagger\|_{2,2} = \infty$ when A is not surjective. We do this, as for $A \in \mathbb{C}^{q\times n}$ surjective,

$$\left\|A^\dagger\right\|_{2,2}^{-1} = \sigma_q(A)$$

where σ_q is the qth singular valuer. As the latter is continuous, this choice guarantees that $A \mapsto \|A^\dagger\|_{2,2}^{-1}$ is continuous.

We can see that K is a variant of the real local condition number κ used in [24, 25, 26, 27], which is given by

$$\kappa(f,x) := \frac{\|f\|_W}{\sqrt{\|f(x)\|_2^2 + \left\|\mathrm{D}_x f^\dagger \Delta^{1/2}\right\|_{2,2}^2}} \quad \text{and} \quad \kappa(f) := \sup_{y\in\mathbb{S}^n} \kappa(f,y) \tag{3.5}$$

for $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $x \in \mathbb{S}^n$. Also, we can see that M is a variant of the μ-condition number introduced by Shub and Smale [48], and given by

$$\mu(f,\zeta) := \|f\|_W \left\|\mathrm{D}_x f^\dagger \Delta^{1/2}\right\|_{2,2} \tag{3.6}$$

for $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$ and $\zeta \in \mathbb{P}^n$. We note that the main difference lie at the fact that we are substituting the norms that we are using in the numerator for the polynomial tuples. The fact that we use a different scaling factor ($\Delta^{1/2}$ instead of Δ) or different norms in the denominators ($\|\ \|_\infty$ instead of $\|\ \|_2$ and so on) does only have an effect up to a $\sqrt{2q\mathbf{D}}$ factor, which does not affect complexity so dramatically. This will be made more explicit in Proposition 4.26. Note that despite these changes, we still have that the local local condition numbers, K and M, become ∞ at a singular zero and that they are finite otherwise.

The remainder of this section is devoted to prove the main properties of K and M, which are the reason we have defined these numbers the way we did. The properties we will show are those needed for a condition-based complexity analysis of the forthcoming algorithms following the lines of [24, 27, 14, 15, 16] (as shown in [50]) and of [13, Ch. 17]. Because of this, the definition of our condition numbers does not raise from a geometric approach as the one in [13, Ch. 14], but rather from a complexity-centered approach in line with the philosophies of [21] and [40].

3.1 Properties of the real condition number K

Recall (see, e.g., [13, Def. 16.35]) that for $f \in \mathcal{H}_{\boldsymbol{d}}[q]$ and $x \in \mathbb{S}^n$, the *Smale's projective gamma* is given by

$$\gamma(f,x) := \sup_{k\geq 2} \left\| \frac{1}{k!} \mathrm{D}_x f^\dagger \overline{\mathrm{D}}_x^k f \right\|^{\frac{1}{k-1}}$$

where $\| \; \| = \| \; \|_{2,2}$ is the operator norm (with respect to the Euclidean norms) of a multilinear map.

Theorem 3.2. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $x \in \mathbb{S}^n$. The following holds:*

- ***Regularity Inequality:*** *Either*

$$\frac{\|f(x)\|}{\sqrt{q}\|f\|_\infty^{\mathbb{R}}} \geq \frac{1}{\mathsf{K}(f,x)} \quad \text{or} \quad \sqrt{q}\|f\|_\infty^{\mathbb{R}} \left\| \mathrm{D}_x f^\dagger \Delta \right\|_{2,2} \leq \mathsf{K}(f,x).$$

  In particular, if $\mathsf{K}(f,x)\frac{\|f(x)\|}{\sqrt{q}\|f\|_\infty^{\mathbb{R}}} < 1$, then $\mathrm{D}_x f : \mathrm{T}_x\mathbb{S}^n \to \mathbb{R}^q$ is surjective and its pseudoinverse $(\mathrm{D}_x f)^\dagger$ exists.

- ***1st Lipschitz property:*** *The maps*

$$\begin{aligned} \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q] &\to [0,\infty) \\ g &\mapsto \frac{\|g\|_\infty^{\mathbb{R}}}{\mathsf{K}(g,x)} \end{aligned} \qquad \text{and} \qquad \begin{aligned} \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q] &\to [0,\infty) \\ g &\mapsto \frac{\|g\|_\infty^{\mathbb{R}}}{\mathsf{K}(g)} \end{aligned}$$

  are 1-Lipschitz with respect the real L_∞-norm. In particular,

$$\mathsf{K}(f,x) \geq 1 \quad \text{and} \quad \mathsf{K}(f) \geq 1.$$

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

$$\begin{aligned} \mathbb{S}^n &\to [0,1] \\ y &\mapsto \frac{1}{\mathsf{K}(f,y)} \end{aligned}$$

  is $\mathbf{D}$-Lipschitz with respect the geodesic distance on $\mathbb{S}^n$.

- ***Higher Derivative Estimate:*** *If $\mathsf{K}(f,x)\frac{|f(x)|}{\|f\|_\infty^{\mathbb{R}}} < 1$, then*

$$\gamma(f,x) \leq \frac{1}{2}(\mathbf{D}-1)\mathsf{K}(f,x).$$

We now discuss the role of the above properties.

Regularity Inequality. The regularity inequality guarantees that, when $\mathsf{K}(f,x) < \infty$, either x is far away from the zero set of f or $\mathrm{D}_x f^\dagger$ exist and is well-defined. The latter is important, because it allows us to do various geometric arguments that rely on this pseudoinverse being defined or, equivalently, on $\mathrm{D}_x f$ being surjective.

1st Lipschitz Property. The main use of the 1st Lipschitz inequality is to control the variation of K with respect to f. This property implies that

$$\frac{1 - \frac{\|f-\tilde{f}\|_\infty^{\mathbb{R}}}{\|f\|_\infty^{\mathbb{R}}}}{1 + \mathsf{K}(f,x)\frac{\|f-\tilde{f}\|_\infty^{\mathbb{R}}}{\|f\|_\infty^{\mathbb{R}}}} \mathsf{K}(f,x) \leq \mathsf{K}\left(\tilde{f},x\right) \leq \frac{1 + \frac{\|f-\tilde{f}\|_\infty^{\mathbb{R}}}{\|f\|_\infty^{\mathbb{R}}}}{1 - \mathsf{K}(f,x)\frac{\|f-\tilde{f}\|_\infty^{\mathbb{R}}}{\|f\|_\infty^{\mathbb{R}}}} \mathsf{K}(f,x) \tag{3.7}$$

whenever $\mathsf{K}(f,x)\frac{\|f-\tilde{f}\|_\infty^{\mathbb{R}}}{\|f\|_\infty^{\mathbb{R}}} < 1$. The latter gauges how K of an approximation of f relates to K of f. We note that the error for f entering this bound is relative.

2nd Lipschitz Property. The 2nd Lipschitz property allows us to gauge the variation of K with respect to x. In this sense, it is very similar to the first Lipschitz property and it implies that

$$\frac{1}{1+\mathsf{K}(f,x)\mathrm{dist}_{\mathbb{S}}(x,\tilde{x})}\mathsf{K}(f,x) \leq \mathsf{K}\left(f,\tilde{x}\right) \leq \frac{1}{1-\mathsf{K}(f,x)\mathrm{dist}_{\mathbb{S}}(x,\tilde{x})}\mathsf{K}(f,x) \tag{3.8}$$

whenever $\mathsf{K}(f,x)\mathrm{dist}_{\mathbb{S}}(x,\tilde{x}) < 1$. Note, however, that in this case, the relative error for x does not play a role as the points we consider lie on the sphere.

Higher Derivative Estimate. Smale's projective gamma, $\gamma(f,\zeta)$, controls many aspects of the local geometry around a zero ζ of the function f. Notably, in the case $q=n$, the radius of the basin of attraction at ζ of Newton's operator N_f associated to f. Recall (see [13, Def. 16.34]) that we say that $x \in \mathbb{S}^n$ is an *approximate zero* of $f \in \mathcal{H}_{\boldsymbol{d}}[n]$ with associated zero $\zeta \in \mathbb{S}^n$ when for all $k \geq 1$,

$$d_{\mathbb{S}}(N_f^k,\zeta) \leq \left(\frac{1}{2}\right)^{2^k-1} d_{\mathbb{S}}(x,\zeta).$$

We have the following result (see [13, Thm. 16.38 and Table 16.1]).

Theorem 3.3. *Let $f \in \mathcal{H}_{\boldsymbol{d}}[n]$ and $\zeta \in \mathbb{S}^n$ such that $f(\zeta)=0$. Let $z \in \mathbb{S}^n$ be such that $d_{\mathbb{S}}(z,\zeta) \leq \frac{1}{45}$ and $d_{\mathbb{S}}(z,\zeta)\gamma(f,\zeta) \leq 0.17708$. Then, z is an approximate zero of f with associated zero ζ.* □

The computation of $\gamma(f,x)$ appears to require all the derivatives of f. The Higher Derivative Estimate allows one to estimate $\gamma(f,x)$ in terms of the first derivative only.

Proof of Theorem 3.2. **Regularity Inequality.** By definition,

$$\frac{1}{\mathsf{K}(f,x)} = \max\left\{\frac{\|f(x)\|}{\sqrt{q}\|f\|_\infty^{\mathbb{R}}}, \frac{1}{\sqrt{q}\|f\|_\infty^{\mathbb{R}}\,\|\mathrm{D}_x f^\dagger \Delta\|_{2,2}}\right\}.$$

Hence either $\frac{1}{\mathsf{K}(f,x)} = \frac{\|f(x)\|}{\sqrt{q}\|f\|_\infty^{\mathbb{R}}}$ or $\mathsf{K}(f,x) = \sqrt{q}\|f\|_\infty^{\mathbb{R}}\left\|\mathrm{D}_x f^\dagger \Delta\right\|_{2,2}$, which finishes the proof.

1st Lipschitz property. We have that

$$\frac{\|g\|_\infty^{\mathbb{R}}}{\mathsf{K}(g,x)} = \max\left\{\frac{\|g(x)\|}{\sqrt{q}}, \frac{\sigma_q\left(\Delta^{-1}\mathrm{D}_x g\right)}{\sqrt{q}}\right\}.$$

Hence, we only need to show that $g \mapsto \|g(x)\|/\sqrt{q}$ and $g \mapsto \sigma_q\left(\Delta^{-1}\mathrm{D}_x g\right)/\sqrt{q}$ are 1-Lipschitz. Now,

$$\left|\frac{\|g(x)\|}{\sqrt{q}} - \frac{\|\tilde{g}(x)\|}{\sqrt{q}}\right| \leq \frac{\|\left(g-\tilde{g}\right)(x)\|}{\sqrt{q}} \leq \|\left(g-\tilde{g}\right)(x)\|_\infty \leq \|g-\tilde{g}\|_\infty^{\mathbb{R}},$$

by the reverse triangle inequality, $\| \; \| \leq \sqrt{q}\| \; \|_\infty$ and the definition of the real L_∞-norm; and

$$\left|\frac{\sigma_q\left(\Delta^{-1}\mathrm{D}_x g\right)}{\sqrt{q}} - \frac{\sigma_q\left(\Delta^{-1}\mathrm{D}_x \tilde{g}\right)}{\sqrt{q}}\right| \leq \frac{\left\|\Delta^{-1}\mathrm{D}_x\left(g-\tilde{g}\right)\right\|_{2,2}}{\sqrt{q}} \leq \left\|\Delta^{-1}\mathrm{D}_x\left(g-\tilde{g}\right)\right\|_{\infty,2}\|g-\tilde{g}\|_\infty^{\mathbb{R}},$$

by the fact that σ_q is 1-Lipschitz with respect $\| \|_{2,2}$, $\| \| \leq \sqrt{q}\| \|_\infty$ and Kellogg's Inequality (Theorem 2.12). Thus our claims follow.

The claim for $g \mapsto \|g\|_\infty^{\mathbb{R}}/\mathsf{K}(g)$ follows from the fact that the supremum of a family of 1-Lipschitz functions is 1-Lipschitz and from

$$\frac{\|g\|_\infty^{\mathbb{R}}}{\mathsf{K}(g)} = \max_{y\in\mathbb{S}^n} \frac{\|g\|_\infty^{\mathbb{R}}}{\mathsf{K}(g,x)}.$$

For the lower bound, just note that

$$\frac{\|f\|_\infty^{\mathbb{R}}}{\mathsf{K}(f,x)} = \left|\frac{\|f\|_\infty^{\mathbb{R}}}{\mathsf{K}(f,x)} - \frac{\|0\|_\infty^{\mathbb{R}}}{\mathsf{K}(0,x)}\right| \leq \|f-0\|_\infty^{\mathbb{R}} = \|f\|_\infty^{\mathbb{R}}$$

by the proven Lipschitz property, and so $\mathsf{K}(f,x) \geq 1$. Similarly with $\mathsf{K}(f)$.

2nd Lipschitz property. Without loss of generality, assume that $\|f\|_\infty^{\mathbb{R}} = 1$, after scaling f by an appropriate constant —note that this does not change the value of K—. Let $y, \tilde{y} \in \mathbb{S}^n$ and $u \in \mathscr{O}(n+1)$ be the planar rotation taking y into $\tilde{y}$. Then

$$\left|\frac{1}{\mathsf{K}\left(f,y\right)} - \frac{1}{\mathsf{K}\left(f,\tilde{y}\right)}\right| = \left|\frac{1}{\mathsf{K}\left(f,y\right)} - \frac{1}{\mathsf{K}\left(f^u,y\right)}\right| \leq \|f - f^u\|_\infty^{\mathbb{R}},$$

where $f^u := f(uX)$ and where the equality follows from the fact that the L_∞-norm is orthogonally invariant along with the inequality from the 1st Lipschitz property.

Now, arguing as when proving the 1st Lipschitz property, we have that for all $z \in \mathbb{S}^n$,

$$|f(z) - f(uz)| \leq \mathbf{D}\,\mathrm{dist}_{\mathbb{S}}(z, uz).$$

By the choice of u, we have that $\mathrm{dist}_{\mathbb{S}}(z,uz) \leq \mathrm{dist}_{\mathbb{S}}(y,\tilde{y})$. Therefore $\|f - f^u\|_\infty^{\mathbb{R}} \leq \mathbf{D}\,\mathrm{dist}_{\mathbb{S}}(y,\tilde{y})$ and we are done.

We note that a variational argument showing that both $y \mapsto \|g(y)\|/\sqrt{q}$ and $y \mapsto \sigma_q(\Delta^{-1}\mathrm{D}_y f))/\sqrt{q}$ are Lipschitz is possible. This argument would be almost identical to the one used for proving the 1st Lipschitz property, but varying the point in the sphere instead of the polynomial. We use the above argument since it is simpler and it gives a slightly better bound.

Higher Derivative Estimate. Again, without loss of generality, we assume that $\|f\|_\infty^{\mathbb{R}} = 1$, since multiplying f by a scalar affects neither the value of K nor Smale's projective gamma. Then

$$\begin{aligned}
\left\|\frac{1}{k!}\mathrm{D}_x f^\dagger \overline{\mathrm{D}}_x^k f\right\| &\leq \left\|\mathrm{D}_x f^\dagger \Delta\right\|_{2,2} \left\|\frac{\Delta^{-1}}{k!}\overline{\mathrm{D}}_x^k f\right\|_{2,2} && \text{(Inequalities for operator norms)}\\
&\leq \sqrt{q}\left\|\mathrm{D}_x f^\dagger \Delta\right\|_{2,2} \left\|\frac{\Delta^{-1}}{k!}\overline{\mathrm{D}}_x^k f\right\|_{2,\infty} && \| \|/\sqrt{q} \leq \| \|_\infty\\
&\leq \mathsf{K}(f,x) \left\|\frac{\Delta^{-1}}{k!}\overline{\mathrm{D}}_x^k f\right\|_{2,\infty} && \text{(Assumption + Regularity Inequality)}\\
&\leq \frac{1}{k}\binom{\mathbf{D}-1}{k-1}\mathsf{K}(f,x). && \text{(Corollary 2.20)}
\end{aligned}$$

Taking $(k-1)$th roots, we have that $\mathsf{K}(f,x)^{\frac{1}{k-1}} \leq \mathsf{K}(f,x)$, since $\mathsf{K}(f,x) \geq 1$ by Corollary 2.13; and that

$$\left(\frac{1}{k}\binom{\mathbf{D}-1}{k-1}\right)^{\frac{1}{k-1}} \leq \frac{\mathbf{D}-1}{2},$$

using that $\frac{1}{k}\binom{\mathbf{D}-1}{k-1} \leq (\mathbf{D}-1)^{k-1}/2^{k-1}$. Putting this together, we obtain the desired bound for Smale's projective gamma. $\square$

We state without proof the following proposition for the sake of completeness. In Section 6, we will discuss other condition number theorems for variants of K and prove the following proposition.

Proposition 3.4. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $x \in \mathbb{S}^n$. Then*

$$\frac{\|f\|_\infty^{\mathbb{R}}}{\operatorname{dist}_\infty^{\mathbb{R}}(f, \Sigma_{\boldsymbol{d},x}^{\mathbb{R}}[q])} \leq \mathsf{K}(f,x) \leq 2\sqrt{\sum_{i=1}^q d_i^2}\,\frac{\|f\|_\infty^{\mathbb{R}}}{\operatorname{dist}_\infty^{\mathbb{R}}(f, \Sigma_{\boldsymbol{d},x}^{\mathbb{R}}[q])}$$

and

$$\frac{\|f\|_\infty^{\mathbb{R}}}{\operatorname{dist}_\infty^{\mathbb{R}}(f, \Sigma_{\boldsymbol{d}}^{\mathbb{R}}[q])} \leq \mathsf{K}(f) \leq 2\sqrt{\sum_{i=1}^q d_i^2}\,\frac{\|f\|_\infty^{\mathbb{R}}}{\operatorname{dist}_\infty^{\mathbb{R}}(f, \Sigma_{\boldsymbol{d}}^{\mathbb{R}}[q])}$$

where $\operatorname{dist}_\infty^{\mathbb{R}}$ is the distance induced by $\|\ \|_\infty^{\mathbb{R}}$,

$$\Sigma_{\boldsymbol{d},x}^{\mathbb{R}}[q] := \left\{g \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q] \mid g(x) = 0,\, \operatorname{rank} \mathrm{D}_x g < q\right\}, \quad \textit{and} \quad \Sigma_{\boldsymbol{d}}^{\mathbb{R}}[q] := \bigcup_{x \in \mathbb{S}^n} \Sigma_{\boldsymbol{d},x}^{\mathbb{R}}[q].$$

3.2 Properties of the complex condition number M

In the complex case, Theorem 3.2 takes the form of the following result, whose proof we omit as it is identical to that of Theorem 3.2. We do not consider a regularity inequality for M since over complex numbers one usually has the luxury to only consider $\mathsf{M}(f,\zeta)$ for a zero ζ of f.

Theorem 3.5. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$ and $\zeta \in \mathbb{P}^n$. The following holds:*

- ***1st Lipschitz property:*** *The maps*

$$\begin{aligned} \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q] &\to [0,\infty) \\ g &\mapsto \frac{\|g\|_\infty^{\mathbb{C}}}{\mathsf{M}(g,\zeta)} \end{aligned} \qquad \textit{and} \qquad \begin{aligned} \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q] &\to [0,\infty) \\ g &\mapsto \frac{\|g\|_\infty^{\mathbb{C}}}{\mathsf{M}(g)} \end{aligned}$$

  are 1-Lipschitz with respect the complex L_∞-norm. In particular,

$$\mathsf{M}(f,\zeta) \geq 1 \quad \textit{and} \quad \mathsf{M}(f) \geq 1.$$

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

$$\begin{aligned} \mathbb{P}^n &\to [0,1] \\ [\eta] &\mapsto \frac{1}{\mathsf{K}(f,[\eta])} \end{aligned}$$

  is $\mathbf{D}$-Lipschitz with respect the geodesic distance on $\mathbb{P}^n$.

- ***Higher Derivative Estimate:*** *We have*

$$\gamma(f,\zeta) \leq \frac{1}{2}(\mathbf{D}-1)\mathsf{M}(f,\zeta). \qquad \square$$

We finish with the following proposition, which combines the 1st and 2nd Lipschitz properties of M, as it will play a fundamental role in our analysis of linear homotopy in Section 5. We note that this proposition is to M what [13, Proposition 16.55] is for μ.

Proposition 3.6. *Let $f, \tilde{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$, $\zeta, \tilde{\zeta} \in \mathbb{P}^n$ and $\varepsilon \in (0,1)$. If*

$$\mathsf{M}(f,\zeta) \max\left\{\frac{2\|\tilde{f}-f\|_\infty^{\mathbb{R}}}{\|f\|_\infty^{\mathbb{C}}}, \mathbf{D}\,\mathrm{dist}_{\mathbb{P}}(\zeta,\tilde{\zeta})\right\} \leq \frac{\varepsilon}{4},$$

then

$$\frac{1}{1+\varepsilon}\mathsf{M}\left(f,\zeta\right) \leq \mathsf{M}\left(\tilde{f},\tilde{\zeta}\right) \leq (1+\varepsilon)\mathsf{M}\left(f,\zeta\right).$$

Proof. Note that

$$\left|\frac{1}{\mathsf{M}(f,\zeta)} - \frac{1}{\mathsf{M}\left(\tilde{f},\tilde{\zeta}\right)}\right| \leq \left|\frac{1}{\mathsf{M}(f,\zeta)} - \frac{1}{\mathsf{M}\left(\tilde{f},[\zeta]\right)}\right| + \left|\frac{1}{\mathsf{M}\left(\tilde{f},\zeta\right)} - \frac{1}{\mathsf{M}\left(\tilde{f},\tilde{\zeta}\right)}\right|.$$

For the first summand, we have

$$\left|\frac{1}{\mathsf{M}\left(f,\zeta\right)} - \frac{1}{\mathsf{M}\left(\tilde{f},[\zeta]\right)}\right| = \left|\frac{1}{\mathsf{M}\left(\frac{f}{\|f\|_\infty^{\mathbb{C}}},\zeta\right)} - \frac{1}{\mathsf{M}\left(\frac{\tilde{f}}{\|\tilde{f}\|_\infty^{\mathbb{C}}},[\zeta]\right)}\right| \leq \left\|\frac{f}{\|f\|_\infty^{\mathbb{C}}} - \frac{\tilde{f}}{\|\tilde{f}\|_\infty^{\mathbb{C}}}\right\|_\infty^{\mathbb{C}}$$

by the 1st Lipschitz property of M (Theorem 3.5). Now,

$$\left\|\frac{f}{\|f\|_\infty^{\mathbb{C}}} - \frac{\tilde{f}}{\|\tilde{f}\|_\infty^{\mathbb{C}}}\right\|_\infty^{\mathbb{C}} \leq \left\|\frac{f}{\|f\|_\infty^{\mathbb{C}}} - \frac{\tilde{f}}{\|f\|_\infty^{\mathbb{C}}}\right\|_\infty^{\mathbb{C}} + \left\|\frac{\tilde{f}}{\|f\|_\infty^{\mathbb{C}}} - \frac{\tilde{f}}{\|\tilde{f}\|_\infty^{\mathbb{C}}}\right\|_\infty^{\mathbb{C}} \leq \frac{2\|\tilde{f}-f\|_\infty^{\mathbb{C}}}{\|f\|_\infty^{\mathbb{C}}}.$$

For the second summand, we have

$$\left|\frac{1}{\mathsf{M}\left(\tilde{f},[\zeta]\right)}\right| + \left|\frac{1}{\mathsf{M}\left(\tilde{f},\zeta\right)} - \frac{1}{\mathsf{M}\left(\tilde{f},\tilde{\zeta}\right)}\right| \leq \mathbf{D}\mathrm{dist}_{\mathbb{P}}(\zeta,\tilde{\zeta})$$

by the 2nd Lipschitz property of M (Theorem 3.5).

Hence, we have

$$\left|\frac{1}{\mathsf{M}(f,\zeta)} - \frac{1}{\mathsf{M}\left(\tilde{f},\tilde{\zeta}\right)}\right| \leq \frac{2\|\tilde{f}-f\|_\infty^{\mathbb{C}}}{\|f\|_\infty^{\mathbb{C}}} + \mathbf{D}\mathrm{dist}_{\mathbb{P}}(\zeta,\tilde{\zeta}).$$

By assumption, after multiplying by $\mathsf{M}(f,\zeta)$, we have

$$\left|1 - \frac{\mathsf{M}(f,\zeta)}{\mathsf{M}\left(\tilde{f},\tilde{\zeta}\right)}\right| \leq \frac{\varepsilon}{2}$$

and so, from

$$1 - \frac{\mathsf{M}(f,\zeta)}{\mathsf{M}\left(\tilde{f},\tilde{\zeta}\right)} \leq \frac{\varepsilon}{2} \text{ and } \frac{\mathsf{M}(f,\zeta)}{\mathsf{M}\left(\tilde{f},\tilde{\zeta}\right)} - 1 \leq \frac{\varepsilon}{2},$$

we get

$$\frac{1}{1+\frac{\varepsilon}{2}}\mathsf{M}(f,\zeta) \leq \mathsf{M}\left(\tilde{f},\tilde{\zeta}\right) \leq \frac{1}{1-\frac{\varepsilon}{2}}\mathsf{M}(f,\zeta).$$

Since $\varepsilon < 1$, the desired inequalities follow. $\square$

4 Numerical Algorithms in Real Algebraic Geometry

There is a growing literature on numerical algorithms that address basic computational tasks in real algebraic geometry such as counting real zeros [24, 25, 26], computing homology of algebraic [27] and semialgebraic sets [14, 15, 16], and meshing real curves and surfaces [44, 22]. These works rely on condition numbers to control precision, and to estimate computational complexity of the algorithms.

In this section we show how the complexity estimates in these works are improved by using the real L_∞-norm in the algorithm's design. These improvements rely on three observations:

1. The only properties of the real condition number κ that are used in the complexity analyses are those stated in Theorem 3.2: the regularity inequality, the 1st and 2nd Lipschitz properties and the Higher Derivative Estimate. As these properties hold as well for K, an almost identical condition-based cost analysis can be derived when we pass from the Weyl norm to the real L_∞-norm and from κ to K. We showcase this in §4.1 and §4.2.

2. When we consider random or semi-random input models, the gains in the complexity estimates become more evident. In §4.3, we show that the ratio of the new K to κ is roughly $\sqrt{n}/\sqrt{N}$ for a random or semi-random polynomial systems. Since $N \sim n^d$ for $n > d$ and $N \sim d^n$ for $d > n$, this yields a significant reduction in the complexity estimates .

3. Computing the Weyl norm is cheaper than computing the real L_∞-norm, but this does not affect the overall complexity: We only compute the L_∞-norm once, and the cost of this computation is dominated by that of the remaining steps in real algebraic geometry algorithms.

In what follows, we will focus on algorithms dealing with real algebraic sets. The algorithms we have in mind are the ones in [24, 25, 26, 27] and the Plantinga-Vegter algorithm [44] as described and analyzed in [23] (cf. [22]). Our condition number K as defined in preceding section will improve the overall computational complexity of these algorithms. Similar results can be obtained for the algorithms dealing with semialgebraic sets in [14, 15, 16] using following variants of K,

$$\overline{\mathsf{K}}(f) := \max\left\{\mathsf{K}\left(f^L\right) \mid L \subseteq [q],\, |L| \leq n+1\right\}$$

where $f^L := (f_i)_{i\in L}$ and

$$\mathsf{K}_*(f,E) := \max\left\{\mathsf{K}\left(f^L\right) \mid E \subseteq L \subseteq [q],\, |L| \leq n+1\right\}$$

for $E \subseteq [q]$ —E indexes the polynomial equalities in the basic semialgebraic formula— which are the natural extensions of

$$\overline{\kappa}(f) := \max\left\{\kappa\left(f^L\right) \mid L \subseteq [q],\, |L| \leq n+1\right\}$$

and

$$\kappa_*(f,E) := \max\left\{\kappa\left(f^L\right) \mid E \subseteq L \subseteq [q],\, |L| \leq n+1\right\}$$

used in [14, 15, 16]. We don't delve into further details here as this would only add length and technicalities without adding substance.

4.1 A grid-based algorithm and its condition-based complexity

A *grid-based algorithm* is a subdivision-based method which constructs a grid to discretize the original problem and solves the latter by working on the grid points only (selecting and finding proximity relations between its points). The algorithms in [24, 25, 26], [27], and [14, 15, 16] (cf. [50]) are grid-based. Their basic structure is (simplifying to the extreme) the following:

1. Estimate the condition number of the problem (with a sequence of grids of increasing fineness).

2. Create an extra grid (if necessary), whose mesh is determined by the condition number.

3. Select points in the grid and use them to obtain a solution to the problem.

In general, grid-based algorithms have complexity $\Omega(\mathbf{D}^n)$. This fact allows us to estimate the norm $\|f\|_\infty^{\mathbb{R}}$ of the data f without affecting the overall complexity of the algorithms. Moreover, the fact that K is smaller than κ results in a cost reduction.

In this subsection, we focus on an algorithm for the computation of the Betti numbers of a spherical algebraic set. This covers the case of counting zeros of a square polynomial system treated in [24, 25, 26] and the computation of the Betti numbers of a projective real variety [27]. For simplicity of expositions we omit some computational aspects: 1) The presentation we do of the algorithms follows the paradigm of construction-selection of [14, 15, 16] instead of the one of inclusion-exclusion of [24, 25, 26, 27]. This makes easier the exposition of the algorithms without compromising their complexity. 2) Our focus on Betti numbers to avoid describing the more involved computation of torsion coefficients in the homology groups. 3) We deal with neither parallelization nor finite precision. The interested reader can find the details about this is the cited references.

The backbone of existing grid-based algorithms in numerical real algebraic geometry [24, 25, 26, 27, 14, 15, 16] is an effective construction of spherical nets. The basic construction was done originally in [24] and it is based on projecting the uniform grid in the boundary of a unit cube onto the unit sphere.

Recall that a *(spherical) δ-net* is a finite subset $\mathcal{G} \subset \mathbb{S}^n$ such that for all $x \in \mathbb{S}^n$, $\mathrm{dist}_{\mathbb{S}}(x, \mathcal{G}) < \delta$. We will omit the term 'spherical' as all the considered nets are spherical.

Proposition 4.1. *There is an algorithm* GRID *that on input* $(n, k) \in \mathbb{N} \times \mathbb{N}$ *outputs a* 2^{-k}*-net* $\mathcal{G}_k \subset \mathbb{S}^n$ *with*

$$|\mathcal{G}_k| = \mathcal{O}\left(2^{n \log n + nk}\right).$$

The cost of this algorithm is $\mathcal{O}\left(2^{n \log n + nk}\right)$. □

Although the grid construction in Proposition 4.1 has been used in the cited references, it's not optimal. The latter is due to the $2^{n \log n}$ factor in the estimates. An algorithm that can construct such a spherical 2^{-k}-net of size $2^{\mathcal{O}(n)} 2^{k(n+1)}$ in $2^{\mathcal{O}(n)} 2^{k(n+1)}$-time was given in [2, Theorem 1.9 (1)]. The main reason to use the sub-optimal result of Proposition 4.1 is to focus on the effect of just changing the norm and doing fair comparison between the old and new versions of the algorithms. The reader who prefers to have an optimal result in memory can neglect the $\log(n)$ factors in the exponents and assume the mentioned near optimal grid construction is being used.

4.1.1 Computation of $\| \ \|_\infty^{\mathbb{R}}$

The following is an easy consequence of Kellogg's theorem.

Proposition 4.2. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $\mathcal{G} \subset \mathbb{S}^n$ a δ-net. If $\mathbf{D}\delta \leq 1$, then*

$$\max_{x\in\mathcal{G}} \|f(x)\|_\infty \leq \|f\|_\infty^{\mathbb{R}} \leq \frac{1}{1-\frac{\mathbf{D}^2}{2}\delta^2} \max_{x\in\mathcal{G}} \|f(x)\|_\infty.$$

Proof. We only need to show the right-hand inequality, the other being trivial. Without loss of generality, assume that $q = 1$, i.e., f is a homogeneous polynomial of degree $\mathbf{D}$.

Let x_* be the maximum of $|f|$ on $\mathbb{S}^n$, $x \in \mathcal{G}$ such that $\mathrm{dist}_{\mathbb{S}}(x_*, x) \leq \delta$ and $[0,1] \ni t \mapsto x_t$ the geodesic on $\mathbb{S}^n$ going from x_* to x with constant speed. Then, for the function $t \mapsto M(t) := f(x_t)$, we have that $|M(1)| \leq |M(0)| + |M'(0)| + \max_{s\in[0,1]} \frac{M''(s)}{2}$ by Taylor's theorem. Furthermore, $|M(0)| = |f(x_*)| = \|f\|_\infty^{\mathbb{R}}$, $|M(1)| = |f(x)|$ and $M'(0) = 0$. The latter is due to the fact that x_* is an extremal point of f and so of M. Now,

$$M''(t) = \overline{\mathrm{D}}_{x_t}^2 f(\dot{x}_t, \dot{x}_t) - \mathbf{D} f(x_t)\mathrm{dist}_{\mathbb{S}}(x_*, x)^2,$$

since $\ddot{x}_t = -\mathrm{dist}_{\mathbb{S}}(x_*, x)^2 x_t$, as x_t is a geodesic on $\mathbb{S}^n$ of constant speed $\mathrm{dist}_{\mathbb{S}}(x_*, x)$, and $\overline{\mathrm{D}}_{x_t} f(x_t) = \mathbf{D} f(x_t)$ by Euler's formula (2.2). Then, by Corollary 2.20,

$$\max_{s\in[0,1]} \frac{|M''(s)|}{2} \leq \binom{\mathbf{D}}{2}\|f\|_\infty^{\mathbb{R}} + \frac{\mathbf{D}}{2}\|f\|_\infty^{\mathbb{R}} = \frac{\mathbf{D}^2}{2}\|f\|_\infty^{\mathbb{R}}.$$

Thus $\|f\|_\infty^{\mathbb{R}} \leq |f(x)| + \frac{\mathbf{D}^2}{2}\|f\|_\infty^{\mathbb{R}}\delta^2$, and the desired inequality follows. $\square$

Remark 4.3. Note that this is an improvement of [31, Lemma 2.5] in which we improve the relation between the fineness of the grid and the precision of the computed estimation of the maximum.

Proposition 4.2 suggests the following algorithm, whose complexity is given in Proposition 4.4 below.

Algorithm 4.1: NormApprox$\mathbb{R}$

Input : $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$, $k \in \mathbb{N}$

$\mathcal{G} \leftarrow$ Grid $(n, \lceil (k-1)/2 + \log \mathbf{D} \rceil)$
$t \leftarrow \left(1 - 2^{-k}\right)^{-1} \max\{\|f(x)\|_\infty \mid x \in \mathcal{G}\}$

Output : $t \in [0, \infty)$
Postcondition: $\left(1 - 2^{-k}\right) t \leq \|f\|_\infty^{\mathbb{R}} \leq t$

Proposition 4.4. *Algorithm* NormApprox$\mathbb{R}$ *is correct. On input $(f, k) \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q] \times \mathbb{N}$, its cost is bounded by*

$$\mathcal{O}\left(2^{n\log n}\mathbf{D}^n 2^{\frac{(k+1)n}{2}} N\right).$$

Proof. This is a direct consequence of Propositions 4.1 and 4.2 and the fact that f can be evaluated at $x \in \mathbb{S}^n$ with $\mathcal{O}(N)$ arithmetic operations (see [13, Lemma 16.31]). $\square$

Remark 4.5. The ideas in this subsubsection can be applied to compute $\|f\|_\infty^{\mathbb{C}}$.

4.1.2 Estimation of K

In many grid-based algorithms, the estimation of condition numbers is done implicitly along the way; this does not affect the overall computational cost and it makes for an easier understanding of these algorithms. The next proposition is the core of the estimation of K. Note that the mesh of the needed grid to estimate K depends on K itself.

Proposition 4.6. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $\mathcal{G} \subset \mathbb{S}^n$ be a δ-net. If*

$$\delta\, \mathbf{D} \max_{x\in\mathcal{G}} \mathsf{K}(f,x) < 1,$$

then

$$\max_{x\in\mathcal{G}} \mathsf{K}(f,x) \leq \mathsf{K}(f) \leq \frac{1}{1-\delta\,\mathbf{D}\ \max_{x\in\mathcal{G}} \mathsf{K}(f,x)} \max_{x\in\mathcal{G}} \mathsf{K}(f,x).$$

Proof. We only have to prove the right-hand side inequality, since the other one is obvious. Let $x_* \in \mathbb{S}^n$ such that $\mathsf{K}(f) = \mathsf{K}(f,x_*)$ and $x \in \mathcal{G}$ such that $\operatorname{dist}_{\mathbb{S}}(f,x) \leq \delta$. Then, by the 2nd Lipschitz property (Theorem 3.2), we have

$$\frac{1}{\mathsf{K}(f,x)} - \frac{1}{\mathsf{K}(f,x_*)} \leq \mathbf{D}\operatorname{dist}_{\mathbb{S}}(x_*,x) \leq \mathbf{D}\,\delta.$$

Hence $1/\mathsf{K}(f,x_*) \leq (1-\delta\,\mathbf{D}\,\mathsf{K}(f,x))/\mathsf{K}(f,x)$ and the desired inequality follows under the given assumptions. ☐

Proposition 4.6 suggest the following algorithm which computes only one L_∞-norm.

Algorithm 4.2: K-ESTIMATE

Input : $f \in \mathcal{H}_{\boldsymbol{d}}[q]$, $k \in \mathbb{N}$, $b \in \mathbb{N} \cup \{\infty\}$

$t \leftarrow \text{NORMAPPROX}\mathbb{R}(f,k+1)$
$\ell \leftarrow 0$
repeat
  $\ell \leftarrow \ell + 1$
  $K \leftarrow \max\{\sqrt{q}t/\max\{\|f(x)\|, \|\mathrm{D}_x f^\dagger \Delta\|^{-1}\} \mid x \in \text{GRID}(n,\ell)\}$
until $\mathbf{D}\,K2^{-\ell} \leq 2^{-(k+1)}$ *or* $2^b \leq K$
if $B \leq K$ **then**
  return `fail`
else
  $\mathcal{K} \leftarrow (1-2^{-k})^{-1}K$
  return $\mathcal{K}$

Output : `fail` or $\mathcal{K} \in (0,\infty)$
Postcondition: $2^b \leq \mathsf{K}(f)$, if `fail`;
$(1-2^{-k})\mathsf{K}(f) \leq \mathcal{K} \leq \mathsf{K}(f)$, otherwise

Proposition 4.7. *Algorithm* K-ESTIMATE *is correct. On input $(f,k,b) \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q] \times \mathbb{N} \times (\mathbb{N} \cup \{\infty\})$, its cost is bounded by*

$$2^{\mathcal{O}(n(k+\log n))} D^n N \min\{\mathsf{K}(f)^n, 2^{nb}\}.$$

Proof. The correctness follows from Propositions 4.4 and 4.6, and $(1-2^{-(k+1)})^2 > 1-2^{-k}$.

The cost of the first line of the algorithm is bounded by Proposition 4.4. The number of evaluations of

$$\sqrt{q}t/\max\{\|f(x)\|, \|\mathrm{D}_x f^\dagger \Delta\|^{-1}\}$$

in the ℓth iteration of the loop is given by Proposition 4.1. We need $\mathcal{O}(N+n^3)$ operations for each such evaluation, by [13, Proposition 16.32].

In this way, if the loop runs ℓ_0 iterations, it performs a total of

$$\mathcal{O}(2^{n\log n}(D^n 2^{\frac{(k+2)n}{2}} N + 2^{n(\ell_0+1)}(N+n^3)))$$

operations.

If the algorithm outputs $\mathcal{K}$, then $\ell_0 = \lceil k + \log \mathbf{D} + \log \mathcal{K} - \log(1-2^{-k})\rceil$. Moreover, from the correctness, $\log \mathcal{K} - \log(1-2^{-k}) \leq \log \mathsf{K}(f)$, and so $\ell_0 \leq k+1+\log \mathbf{D} + \log \mathsf{K}(f)$.

If the algorithm outputs `fail`, then the first criterion had to fail and so, as long as the second criterion fails too, we have

$$\ell < k + \log \mathbf{D} + b.$$

And so, in this case, $\ell_0 \leq k + 1 + \log \mathbf{D} + \log b$.

We conclude from the bounds above and some straightforward computations. □

By setting k to 7 and b large enough, we have the following important corollary.

Corollary 4.8. *There is an algorithm,* K-ESTIMATE*, *that on input* $(f) \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ *computes* $\mathcal{K} \in [1,\infty)$ *such that*

$$0.99\mathcal{K} \leq \mathsf{K}(f) \leq \mathcal{K}.$$

This algorithm halts if and only if $\mathsf{K}(f) < \infty$ *and its cost is bounded by*

$$2^{\mathcal{O}(n\log n)} D^n N \mathsf{K}(f)^n.$$

□

4.1.3 Complexity analysis of grid-based algorithms using K

To get the grid-method to work, we need two ingredients: a method for selecting the points in the grid near the geometric object of interest and a way of controlling distances between these two sets.

Theorem 4.9 (Construction-Selection). *Let* $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ *and* $\mathcal{G} \subseteq \mathbb{S}^n$ *be a* δ*-net. If*

$$4\mathbf{D}^2 \mathsf{K}(f)^2 \delta < 1,$$

and $Q \in \mathbb{R}$ *is such that* $0.99Q \leq \|f\|_\infty^{\mathbb{R}} \leq Q$, *then*

$$\mathrm{dist}_H\left(\left\{x \in \mathcal{G} \mid \frac{\|f(x)\|}{\sqrt{q}Q} < \mathbf{D}\,\delta\right\}, \mathcal{Z}_{\mathbb{S}}(f)\right) < 2\mathbf{D}\mathsf{K}(f)\delta,$$

where $\mathrm{dist}_H(A,B) := \max\{\sup\{\mathrm{dist}(a,B) \mid a \in A\}, \sup\{\mathrm{dist}(b,A) \mid b \in B\}\}$ *is the Hausdorff distance.*

Following [32], recall that the *medial axis* Δ_X of a closed set $X \subset \mathbb{R}^n$ is the set

$$\Delta_X := \{p \in \mathbb{R}^n \mid \#\{x \in X \mid \mathrm{dist}(p,x) = \mathrm{dist}(p,X)\} \geq 2\},$$

consisting of those points for which there is more than one nearest point in X, and that the *reach* $\tau(X)$ of X is the quantity

$$\tau(X) := \mathrm{dist}(X, \Delta_X),$$

measuring the size of the neighborhood of X at which the nearest point projection is well-defined. If X is finite, then Δ_X is the union of the boundaries of the cells of the Voronoi diagram of X, and $\tau(X)$ is half the minimum distance between two distinct points of X. Thus, for $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[n]$, $2\tau(\mathcal{Z}_{\mathbb{S}}(f))$ is the separation of the zeros of f in the sphere.

Theorem 4.10. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$. Then*

$$\tau(\mathcal{Z}_{\mathbb{S}}(f)) \geq \frac{1}{7\mathbf{D}\mathsf{K}(f)}.$$

Proof of Theorem 4.9. Let $x_0 \in \mathcal{Z}_{\mathbb{S}}(f)$, then there is some $x_1 \in \mathcal{G}$ such that $\mathrm{dist}_{\mathbb{S}}(x_0, x_1) \leq \delta$. Let $[0,1] \ni t \mapsto x_t$ be the geodesic joining them. By Taylor's theorem,

$$\|f(x_1)\| \leq \|f(x_0)\| + \delta \sup_{s \in [0,1]} \|\mathrm{D}_x f\|,$$

and so, by Kellogg's theorem (Corollary 2.13) and $f(x_0) = 0$, we have that $\frac{\|f(x_1)\|}{\sqrt{q}\|f\|_\infty^{\mathbb{R}}} \leq \mathbf{D}\,\delta$. Hence $\frac{\|f(x_1)\|}{\sqrt{q}Q} \leq \mathbf{D}\,\delta$ and

$$\mathrm{dist}\left(x_0, \left\{x \in \mathcal{G} \mid \frac{\|f(x)\|}{\sqrt{q}Q} \leq \mathbf{D}\,\delta\right\}\right) \leq \mathrm{dist}(x_0, x_1) \leq \mathrm{dist}_{\mathbb{S}}(x_0, x_1) \leq \delta.$$

Let now $x_2 \in \mathcal{G}$ be such that $\frac{\|f(x_2)\|}{\sqrt{q}Q} < \mathbf{D}\,\delta$. Then

$$\frac{\|f(x_2)\|}{\sqrt{q}\|f\|_\infty^{\mathbb{R}}} < 1.02\mathbf{D}\,\delta \leq \frac{1}{4\mathbf{D}\mathsf{K}(f)^2} < \frac{1}{\mathsf{K}(f, x_2)} \tag{4.1}$$

the second inequality by our hypothesis. Because of the Regularity Inequality (Theorem 3.2) we must then have $\sqrt{q}\|f\|_\infty^{\mathbb{R}}\|\mathrm{D}_{x_2} f^\dagger \Delta^{1/2}\| \leq \mathsf{K}(f, x_2)$. It follows that

$$\begin{aligned} \|\mathrm{D}_{x_2} f^\dagger f(x_2)\|\gamma(f, x_2) &< \frac{\mathsf{K}(f, x_2)}{\sqrt{q}\|f\|_\infty^{\mathbb{R}}}\gamma(f, x_2) \leq \frac{1}{2}\mathbf{D}\mathsf{K}(f)^2 \frac{\|f(x_2)\|}{\sqrt{q}\|f\|_\infty^{\mathbb{R}}} \\ &< \frac{1.02}{2}\mathbf{D}^2\mathsf{K}(f)^2\delta < \frac{1.02}{8} < 0.13071\ldots \end{aligned}$$

where we used the Higher Derivative Estimate (Theorem 3.2) in the first line, and (4.1) and the hypothesis in the second. This means that Smale's α-criterion holds for x_2 and $f_{|\mathrm{T}_{x_0}\mathbb{S}^n}$ by [28, Théorème 128]. Hence there is $x_3 \in \mathrm{T}_{x_2}\mathbb{S}^n$ such that $f(x_3) = 0$ and that

$$\mathrm{dist}(x_2, x_3) \leq 1.64\|\mathrm{D}_{x_2} f^\dagger f(x_2)\| \leq 1.64 \cdot 1.02\,\mathbf{D}\mathsf{K}(f)\delta < 2\mathbf{D}\mathsf{K}(f)\delta.$$

Since $\mathrm{dist}(x_2, x_3/\|x_3\|) = \arctan \mathrm{dist}(x_2, x_3) \leq \mathrm{dist}(x_2, x_3)$, we are done. □

Remark 4.11. The proof also shows that we have convergence of the Newton method associated to $f_{|\mathrm{T}_x\mathbb{S}^n}$ for every $x \in \mathcal{G}$ such that $\frac{\|f(x)\|}{\sqrt{q}\|f\|_\infty^{\mathbb{R}}} \leq \mathbf{D}\,\delta$. Hence, we can refine our approximations if wanted.

Sketch of proof of Theorem 4.10. The proof is very similar to the one of [14, Theorem 4.12]. By [14, Lemma 2.7] and [14, Theorem 3.3], we have that

$$\tau(\mathcal{Z}_{\mathbb{S}}(f)) \geq \min\left\{1, \frac{1}{14\max\{\gamma(f,x) \mid x \in \mathcal{Z}_{\mathbb{S}}(f)\}}\right\}.$$

Hence, by the Higher Derivative Estimate (Theorem 3.2), the desired bound follows. □

The following theorem is a variant of the so-called Niyogi-Smale-Weinberger theorem [42, Propoposition 7.1].

Theorem 4.12. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$, $\mathcal{G} \subset \mathbb{S}^n$ be a δ-net, and $t \in \mathbb{R}$ be such that $0.99Q \leq \|f\|_\infty^{\mathbb{R}} \leq Q$. If $90\mathbf{D}^2\mathsf{K}(f)^2\delta < 1$, then for every*

$$\varepsilon \in \left(6\mathbf{D}\mathsf{K}(f)\delta, \frac{1}{14\mathbf{D}\mathsf{K}(f)}\right),$$

the sets $\mathcal{Z}_{\mathbb{S}}(f)$ and

$$\bigcup\left\{B(x,\varepsilon) \mid x \in \mathcal{G}, \frac{\|f(x)\|}{\sqrt{q}Q} < \mathbf{D}\delta\right\}$$

are homotopically equivalent. In particular, they have the same Betti numbers.

Proof. This is just [14, Theorem 2.8] combined with Theorems 4.9 and 4.10. □

We can now describe the algorithm. In doing so, we will call a black box Betti for computing the Betti numbers of a union of balls. This is a standard procedure in topological data analysis.

Algorithm 4.3: PolyBetti$_\infty$

Input : $f \in \mathcal{H}_{\boldsymbol{d}}[q]$
Precondition : $q \leq n$, f has no singular zeros (i.e. $\mathsf{K}(f) < \infty$)

$Q \leftarrow$ NormApprox$\mathbb{R}(f, 7)$
$k \leftarrow$ K-Estimate$^*(f)$
$\ell \leftarrow 7 + \lceil 2\log\mathbf{D} + 2\log k\rceil$
$\mathcal{G} \leftarrow$ Grid(n, ℓ)
$\mathcal{X} \leftarrow \{x \in \mathcal{G} \mid \|f(x)\| < \sqrt{q}\,\mathbf{D}Q2^{-\ell}\}$
$\varepsilon \leftarrow 3/(50\mathbf{D}\mathsf{K}(f))$
$(\beta_0, \ldots, \beta_n) \leftarrow$ Betti$(\mathcal{X}, \varepsilon)$
return $\beta_0, \ldots, \beta_n$

Output : $\beta_0, \ldots, \beta_n \in \mathbb{N}$
Postcondition: $\beta_0, \ldots, \beta_n$ are the Betti numbers of $\mathcal{Z}_{\mathbb{S}}(f)$

Proposition 4.13. *Algorithm* PolyBetti$_\infty$ *is correct and its cost is bounded by*

$$2^{\mathcal{O}(n^2\log n)}D^{10n^2}\mathsf{K}(f)^{10n^2}.$$

Proof. Correctness is a consequence of Theorem 4.12 and tyhe fact that the computed Q satisfies $0.99Q \le \|f\|_\infty^{\mathbb{R}} \le Q$ by Proposition 4.4.

For the complexity, we apply Proposition 4.2 for the first line, Corollary 4.8 for the second line, and Proposition 4.1 for the fourth and fifth line. We know that BETTI has cost $\mathcal{O}\left(2^{\mathcal{O}(n\log n)}|\mathcal{X}|^{5n}\right)$ (see [27, §5] for example) and that $|\mathcal{X}| = \mathcal{O}(2^{n\log n}\mathbf{D}^{2n}\mathsf{K}(f)^{2n})$, by Proposition 4.1.

We note that our bound uses $k \le 1.02\mathsf{K}(f)$ in order to get the cost dependent on $\mathsf{K}(f)$ instead of on the computed estimate k. □

The complexity estimate in Proposition 4.13 does not differ too much in form from those in other grid-based algorithms. We will see in §4.3, however, that the occurrence of K in the place of κ leads to substantial improvements when one goes beyond the-worst case framework and considers random or semi-random input models.

4.2 Complexity of the Plantinga-Vegter algorithm

The ideas above can also be applied to the Plantinga-Vegter algorithm [44]. Since in a recent work [23] (cf. [22]) we performed an extensive analysis of this algorithm including details for finite precision arithmetic. So we will be here scarce on details, referring the reader to [23] for details, and only focus on the (exact) interval version of the algorithm.

4.2.1 The Plantinga-Vegter Subdivision Algorithm

Let $\mathcal{P}_d$ be the space of polynomials in $X_1,\dots,X_n$ of degree at most d. The Plantinga-Vegter algorithm [44] is a subdivision-based algorithm for obtaining a piecewise linear approximation[2] of the zero set of $f \in \mathcal{P}_d$ inside $[-a,a]^n$. As customary, we will focus on the complexity analysis of the subdivision routine only. The idea is to subdivide $[-a,a]^n$ into succesive boxes, i.e., sets of the form $B = m(B) + [-w(B)/2, w(B)/2]^n$ (here $m(B) \in \mathbb{R}^n$ is the center of B and $w(B) > 0$ is its width), until every box B satisfies the following condition:

$$C_f(B) \,:\, \text{either } 0 \notin f(B) \text{ or } 0 \notin \langle \partial f(B), \partial f(B)\rangle$$

where $\langle\ ,\ \rangle$ is the standard inner product and ∂f is the gradient vector of f. Once this approximation criterion is satisfied by all cell is the sub-division the Plantinga-Vegter algorithms halts, and returns a topologically accurate apprimation of the zero set of f in the region $[a,-a]^n$. We will use to check if the termination criterion is satisfied, the use of the condition numbers for precision control and for the complexity analysis of finite-precision implementation can be read off from our earlier work [23].

For $f \in \mathcal{P}_d$, we define

$$\|f\|_\infty := \max\{|f^{\mathsf{h}}(x)| \mid x \in \mathbb{S}^n\} = \|f^{\mathsf{h}}\|_\infty^{\mathbb{R}}$$

where $f^{\mathsf{h}} \in \mathcal{H}_d[1]$ is the homogenization of f. Taking the maps (2.3), (2.4), (2.5) in [23] and substituting the Weyl norm by the real L_∞-norm on them we get

$$h(x) = \frac{1}{\|f\|_\infty (1+\|x\|^2)^{(d-1)/2}} \quad \text{and} \quad h'(x) = \frac{1}{d\|f\|_\infty (1+\|x\|^2)^{d/2-1}} \tag{4.2}$$

[2] The original algorithm [44] only dealt with dimensions two and three. For the extension to dimensions four or higher see [34].

together with

$$\widehat{f} : x \mapsto h(x)f(x) = \frac{f(x)}{\|f\|_\infty(1+\|x\|^2)^{(d-1)/2}} \tag{4.3}$$

and

$$\widehat{\partial f} : x \mapsto h'(x)\mathrm{D}f(x) = \frac{\partial f(x)}{d\|f\|_\infty(1+\|x\|^2)^{d/2-1}}. \tag{4.4}$$

One can use this maps to produce interval approximations as we do in [23]. Recall that an *interval approximation* of $f : \mathbb{R}^n \to \mathbb{R}^q$ is a function $\square f : \square\mathbb{R}^n \to \mathbb{R}^q$ that maps boxes in $\mathbb{R}^n$ to boxes in $\mathbb{R}^q$ in such a way that $f(B) \subseteq \square f(B)$.

Proposition 4.14. *Let $f \in \mathcal{P}_d$. Then*

$$\square[hf] : B \mapsto \widehat{f}(m(B)) + (1+d)\sqrt{n}\left[-\frac{w(B)}{2}, \frac{w(B)}{2}\right]$$

is an interval approximation of hf, and

$$\square[\|h'\mathrm{D}f\|] : B \mapsto \|\widehat{\mathrm{D}f}(m(B))\| + d\sqrt{n}\left[-\frac{w(B)}{2}, \frac{w(B)}{2}\right]$$

is an interval approximation of $\|h'\mathrm{D}f\|$.

Sketch of proof. Using the bounds from Kellogg's theorem (Theorem 2.12) and its corollaries, we can deduce easily (as it is done in the proof of Theorem 3.2) that the maps

$$g/\|g\|_\infty^{\mathbb{R}} : \mathbb{S}^n \to [-1,1] \text{ and } \overline{\mathrm{D}}g(v)/(d\|g\|_\infty^{\mathbb{R}}\|v\|) : \mathbb{S}^n \to [-1,1]$$

are d- and $(d-1)$-Lipschitz (with respect to the geodesic distance) for $g \in \mathcal{H}_d^{\mathbb{R}}[1]$.

We now argue as in [23, §4], but using these Lipschitz properties, to prove that $\hat{f}$ and $\widehat{\partial f}$ are $(1+d)$- and d-Lipschitz. For the latter, we use the fact that for $v \in \mathbb{R}^n$, $\overline{\mathrm{D}}_X f^{\mathsf{h}}\begin{pmatrix}0\\v\end{pmatrix} = (\langle \partial f, v\rangle)^{\mathsf{h}}$ and that $\|\widehat{\partial f}\|$ is d-Lipschitz if $\langle\widehat{\partial f}, v\rangle$ is so for every $v \in \mathbb{S}^{n-1}$. □

Using the interval approximations and their Lipschitz properties in Proposition 4.14 we can rewrite the condition $C_f(B)$. We only need to use [23, Lemma 4.2] for the second clause of the condition.

Theorem 4.15. *Let $B \in \square\mathbb{R}^n$. If the condition*

$$C_f^{\square}(B) \,:\, \left|\widehat{f}(m(B))\right| > 2d\sqrt{n}w(B) \text{ or } \left\|\widehat{\mathrm{D}f}(m(B))\right\| > 2\sqrt{2}d\sqrt{n}w(B).$$

is satisfied, then $C_f(B)$ is true. □

The subdivision procedure of the Plantinga-Vegter algorithm thus takes the following form (where StandardSubdivision is a procedure that given a box divides it into 2^n equal boxes and $\square[-a,a]^n$ is the set of boxes within $[-a,a]^n$).

Algorithm 4.4: PV-INTERVAL$_\infty$

Input : $f \in \mathcal{P}_d$
$a \in (0, \infty)$
Precondition : $\mathcal{Z}(f)$ is smooth inside $[-a, a]^n$

$Q \leftarrow$ NORMAPPROX$\mathbb{R}(f, 7)$
$\tilde{\mathcal{S}} \leftarrow \{[-a, a]^n\}$
$\mathcal{S} \leftarrow \varnothing$
repeat
  Take B in $\tilde{\mathcal{S}}$
  $\tilde{\mathcal{S}} \leftarrow \tilde{\mathcal{S}} \setminus \{B\}$
  if $|f(m(B))| > 2d\sqrt{n}w(B)Q(1 + \|m(B)\|^2)^{\frac{d-1}{2}}$ **then**
    $\mathcal{S} \leftarrow \mathcal{S} \cup \{B\}$
  else if $\|\partial f(m(B))\| > 2\sqrt{2}d\sqrt{n}w(B)Q(1 + \|m(B)\|^2)^{\frac{d}{2}-1}$ **then**
    $\mathcal{S} \leftarrow \mathcal{S} \cup \{B\}$
  else
    $\tilde{\mathcal{S}} \leftarrow \tilde{\mathcal{S}} \cup$ STANDARDSUBDIVISION(B)
until $\tilde{\mathcal{S}} = \varnothing$
return $\mathcal{S}$

Output : Subdivision $\mathcal{S} \subseteq \square[-a, a]^n$ of $[-a, a]^n$
Postcondition: For all $B \in \mathcal{S}$, $C_f(B)$ is true

4.2.2 Complexity of PV-INTERVAL$_\infty$

Without effort, [23, Proposition 5.1] transforms into the following proposition. The essential step is to multiply the inequalities in that proposition by $\|f^{\mathsf{h}}\|_W/\|f\|_\infty$.

Proposition 4.16. *Let $f \in \mathcal{P}_d$ and $x \in \mathbb{R}^n$. Then either*

$$\left|\widehat{f}(x)\right| > \frac{1}{2\sqrt{2d}\,\mathsf{K}(f^{\mathsf{h}}, \text{IO}(x))} \text{ or } \left\|\widehat{\partial f}(x)\right\| > \frac{1}{2\sqrt{2d}\,\mathsf{K}(f^{\mathsf{h}}, \text{IO}(x))},$$

where $\text{IO}(x) = \frac{1}{\sqrt{1+\|x\|^2}}\begin{pmatrix}1\\ x\end{pmatrix} \in \mathbb{S}^n$. ☐

With Proposition 4.16 and the Lipschitz properties shown for $\hat{f}$ and $\widehat{\partial f}$, one can produce a so-called *local size bound* for $C'_f(B)$. This is a function that evaluated at a point x gives a lower bound on the volume of all possible boxes not satisfying $C'_f(B)$ and containing x.

Theorem 4.17. *The map*

$$x \mapsto 1/\left(2^{3/2}d^{\frac{3}{2}}\sqrt{n}\mathsf{K}(f^{\mathsf{h}}, \text{IO}(x))\right)^n$$

is a local size bound for C'_f (of Theorem 4.15). ☐

Then using the continuous amortization of [17, 18, 19] (see [23, Theorem 6.1]), we conclude the following. Note that we need some care with the call to NormApprox$\mathbb{R}$, but this is taken care of by Proposition 4.2.

Theorem 4.18. *The number of boxes in the final subdivision $\mathcal{S}$ of* PV-Interval$_\infty$ *on input (f,a) is at most*

$$d^{\frac{3}{2}n}\max\{1,a^n\}2^{\frac{1}{2}n\log n+11n}\mathop{\mathbb{E}}_{\mathfrak{x}\in[-a,a]^n}\left(\mathsf{K}(f^{\mathsf{h}},\mathrm{IO}(\mathfrak{x}))^n\right).$$

The number of arithmetic operations performed by PV-Interval$_\infty$ *on input (f,a) is at most*

$$\mathcal{O}\left(d^{\frac{3}{2}n+1}\max\{1,a^n\}2^{\frac{1}{2}n\log n+11n}N\mathop{\mathbb{E}}_{\mathfrak{x}\in[-a,a]^n}\left(\mathsf{K}(f^{\mathsf{h}},\mathrm{IO}(\mathfrak{x}))^n\right)\right). \qquad \square$$

The condition-based estimates in Theorem 4.18 are very similar to those of [23, Theorem 6.3]. It is important to observe that only one norm computation is performed by PV-Interval$_\infty$ (in its very first step) and that the cost of this computation is already included in the cost bound in Theorem 4.18. We will see in §4.3.3 that the occurrence of K in the place of κ results in significant improvements in overall complexity when we consider average or smoothed analysis.

4.3 Probabilistic Analysis of Algorithms

In the preceding sections, we have shown that existing grid-based and subdivision-based algorithms based on κ can be modified to use K instead. Moreover, we have shown that the condition-based complexity estimates in terms of K are similar to those in terms of κ. In this section we will show that when we consider random inputs, in contrast, the cost (expected or in probability) substantially decreases.

We first introduce the randomness model with some useful probabilistic results. Then we prove a general comparison result which shows that when substituting κ by K one can expect to reduce the size of the condition number by a factor of $\sqrt{N}$. Finally, we apply these probabilistic estimates for the Plantinga-Vegter algorithm where the improvement can be easily appreciated.

For most algorithms in real algebraic geometry, condition-based estimates show a dependence on either κ^n or on K^n. When this occurs the complexity estimates improve by a factor of the form $N^{\frac{n}{2}}$ when we pass from κ to K. The final complexity changes from having an exponent quadratic in n to an exponent quasilinear in n.

4.3.1 The Randomness Model: Dobro Random Polynomials

Given a random variable $\mathfrak{x}\in\mathbb{R}$ we say that:

(i) $\mathfrak{x}$ is *centered* if $\mathbb{E}\,\mathfrak{x}=0$.

(ii) A random variable $\mathfrak{x}\in\mathbb{R}$ is *subgaussian* if there is a constant K such that for all $p\geq 1$,

$$(\mathbb{E}\,|\mathfrak{x}|^p)^{\frac{1}{p}}\leq K\sqrt{p}.$$

The smallest K satisfying this condition is called the ψ_2-*norm* of $\mathfrak{x}$, and is denoted $\|\mathfrak{x}\|_{\psi_2}$.

(iii) $\mathfrak{x}$ has the *anti-concentration property with constant* ρ if for all $u \in \mathbb{R}$ and $\varepsilon > 0$,

$$\mathbb{P}(|\mathfrak{x} - u| < \varepsilon) \leq 2\rho\varepsilon.$$

Note that this is equivalent to $\mathfrak{x}$ having a density (with respect to the Lebesgue measure) bounded by ρ.

We now extend to tuples, the class of real random polynomials, introduced in [22]. These random tuples generalized the so-called KSS random tuples.

Definition 4.19. A *dobro random polynomial tuple* $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ with parameters K and ρ is a tuple of random polynomials

$$\left(\sum_{|\alpha|=d_1} \binom{d_1}{\alpha}^{\frac{1}{2}} \mathfrak{c}_{1,\alpha} X^\alpha, \dots, \sum_{|\alpha|=d_q} \binom{d_q}{\alpha}^{\frac{1}{2}} \mathfrak{c}_{q,\alpha} X^\alpha \right)$$

such that the $\mathfrak{c}_{i,\alpha}$ are independent centered subgaussian random variables with ψ_2-norm at most K and anti-concentration property with constant ρ.

Remark 4.20. Probabilistic estimates for a dobro polynomial $\mathfrak{f}$ will depend on $K\rho$. This product is invariant under scalar multiplication of $\mathfrak{f}$ since $\lambda\mathfrak{f}$ is dobro with parameters $|\lambda|K$ and $\rho/|\lambda|$. Moreover, note that[3] $6K\rho \geq 1$. In the particular case of KSS polynomials, we can take $K\rho = 2/\sqrt{\pi}$.

Example 4.21. A dobro random polynomial tuple $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ such that the $\mathfrak{c}_\alpha$ are are i.d.d. normal random variables of mean zero and variance one is called a *KSS (real) polynomial tuple*[4]. In this case, we can take $K\rho = 2/\sqrt{\pi}$.

Example 4.22. A dobro random polynomial tuple $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ such that the $\mathfrak{c}_\alpha$ are are i.d.d. uniform random variables in $[-1, 1]$ is a *Weyl uniform (real) polynomial tuple*. In this case, we can take $K\rho = 1/2$.

We now state and prove several probabilistic results that will be used later on.

Proposition 4.23 (Subgaussian tail bounds). *Let $\mathfrak{x} \in \mathbb{R}$ be a random variable.*

1. *If $\mathfrak{x}$ is subgaussian with ψ_2-norm at most K, then for all $t > 0$, $\mathbb{P}(|\mathfrak{x}| \geq t) \leq 2e^{-\frac{t^2}{2K^2}}$.*
2. *If there are $C > 1$ and $K > 0$ such that for all $t > 0$, $\mathbb{P}(|\mathfrak{x}| \geq t) \leq Ce^{-\frac{t^2}{K^2}}$, then $\mathfrak{x}$ is subgaussian with ψ_2-norm at most $K\left(1 + \sqrt{2 \ln C}\right)$.*

Proposition 4.24 (Hoeffding inequality). *Let $\mathfrak{x} \in \mathbb{R}^N$ be a random vector such that its components $\mathfrak{x}_i$ are centered subgaussian random varianbles with ψ_2-norm at most K and $a \in \mathbb{S}^{N-1}$. Then, $a^*\mathfrak{x}$ is a subgaussian random variable with ψ_2-norm at most $\frac{5}{4}K$. In particular, for all $t \geq 0$,*

$$\mathbb{P}_{\mathfrak{x}}\left(|a^*\mathfrak{x}| \geq t\right) \leq 2e^{-\frac{8t^2}{25K^2}}.$$

[3]This follows from $2tK\rho \geq \mathbb{P}_{\mathfrak{x}}(|\mathfrak{x}| \leq Kt) \geq 1 - \mathbb{P}_{\mathfrak{x}}(|\mathfrak{x}| > Kt) \geq 1 - 2e^{-t^2/2}$ and optimizing, where $\mathfrak{x}$ is subgaussian with ψ_2-norm K and the anti-concentration property with constant ρ.

[4]In this definition, KSS refers to Kostlan-Shub-Smale. An alternative term is 'Shub-Smale random polynomial tuple', following [4], but we use 'KSS' instead, as this is consistent with the use we have made of the term in the case of a single polynomial.

Proposition 4.25 (Anti-concentration bound). *Let $\mathfrak{x} \in \mathbb{R}^N$ be a random vector such that its components $\mathfrak{x}_i$ are random variables with anti-concentration property with constant ρ. Then, for every $A \in \mathbb{R}^{k\times N}$ with rank k and measurable $U \subseteq \mathbb{R}^k$,*

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

Proof of Proposition 4.23. This is just [53, Proposition 2.5.2] with a twist. For the first part, we only have to follow the constants in the proof. For the second one, note that

$$\mathbb{E}\,|\mathfrak{x}|^p = K^p\left(2\ln C\right)^{\frac{p}{2}} + \int_0^\infty u^{p-1}e^{-\frac{u^2}{2K}}\,\mathrm{d}u,$$

which follows from

$$\mathbb{P}(|\mathfrak{x}| > u) \leq \begin{cases} 1 & \text{if } u \leq K\sqrt{2\ln C} \\ e^{-\frac{u^2}{2K^2}} & \text{if } u \geq K\sqrt{2\ln C}, \end{cases}$$

dividing the integration domain into $[0, K\sqrt{2\ln C}]$ and $[K\sqrt{2\ln C}, \infty]$, and applying some straightforward calculations and bounds.

Now, applying the change of variables $t = \frac{u^2}{2K}$ and Stirling's inequality, we obtain

$$\int_0^\infty u^{p-1}e^{-\frac{u^2}{2K}}\,\mathrm{d}u = K^p 2^{\frac{p}{2}-1}\Gamma\left(\frac{p}{2}\right) \leq 2K^p e^{-\frac{p}{2}}p^{\frac{p}{2}}.$$

Hence

$$\mathbb{E}\,|\mathfrak{x}|^p \leq K^p\left(\left(2\ln C\right)^{\frac{p}{2}} + 2p^{\frac{p}{2}}\right),$$

from where the second part follows. □

Proof of Proposition 4.24. This is an application of [53, Proposition 2.6.1], where we only have to find the explicit constants hidden in the proofs of [53, (2.5) and (2.6)] —the constants are given as absolute constants in the statement, but one can find the precise constants in the proofs—. Note however that for us the ψ_2-norm is the K_1 in [53, Proposition 2.5.2], while in [53] it is the constant K_4 of that proposition.

The last claim is just applying Proposition 4.23. □

Proof of Proposition 4.25. This is just a rewriting of [46, Theorem 1.1]. To get explicit constants, we use [39]. This rewriting was first given in [51, Proposition 2.5]. We provide the argument for the sake of completeness.

By the SVD, we have $A = P\Sigma Q$ where P is an isometry, $\Sigma \in \mathbb{R}^{k\times k}$ a positive diagonal matrix and Q an orthogonal projection. Hence

$$\mathbb{P}_{\mathfrak{x}}\left(A\mathfrak{x} \in U\right) = \mathbb{P}_{\mathfrak{x}}\left(Q\mathfrak{x} \in \Sigma^{-1}P^*U\right)$$

and, since $\mathrm{vol}(\Sigma^{-1}P^*U) = \mathrm{vol}(U)/\det\Sigma = \mathrm{vol}(U)/\sqrt{\det(AA^*)}$, we only have to prove the claim for the case in which A is an orthogonal projection.

Now, by [46, Theorem 1.1] (see [39, Theorem 1.1] for getting the constant), we have that $A\mathfrak{x}$ has density bounded by $\sqrt{2}\rho$. Thus $\mathbb{P}(A\mathfrak{x} \in U) \leq \mathrm{vol}(U)(\sqrt{2}\rho)^k$, as we wanted to show. □

4.3.2 K vs. κ: Meassuring the effect of the L_∞-norm on the Grid Method

The condition-based complexity estimates we obtained in this section essentially substitute the κ in the cost estimates of the original algorithm by K. In this way, the comparison between the two algorithm reduces to estimate K/κ. The following proposition shows that, in turn, this amounts to look at the quotient $\|f\|_\infty^{\mathbb{R}}/\|f\|_W$.

Proposition 4.26. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $x \in \mathbb{S}^n$. Then*

$$\frac{\|f\|_\infty^{\mathbb{R}}}{\|f\|_W} \leq \frac{\mathsf{K}(f,x)}{\kappa(f,x)} \leq \sqrt{2q\mathbf{D}}\frac{\|f\|_\infty^{\mathbb{R}}}{\|f\|_W}$$

and

$$\frac{\|f\|_\infty^{\mathbb{R}}}{\|f\|_W} \leq \frac{\mathsf{K}(f)}{\kappa(f)} \leq \sqrt{2q\mathbf{D}}\frac{\|f\|_\infty^{\mathbb{R}}}{\|f\|_W}.$$

Proof. It follows from

$$\frac{\mathsf{K}(f,x)}{\kappa(f,x)} = \sqrt{q}\frac{\|f\|_\infty^{\mathbb{R}}}{\|f\|_W}\frac{\sqrt{\|f(x)\|^2+\sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_xf\right)^2}}{\max\left\{\|f(x)\|,\sigma_q\left(\Delta^{-1}\mathrm{D}_xf\right)\right\}}$$

and

$$\frac{1}{\sqrt{\mathbf{D}}}\,\sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_xf\right) \leq \sigma_q\left(\Delta^{-1}\mathrm{D}_xf\right) \leq \sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_xf\right).$$

□

In general, we have that $\frac{\|f\|_\infty^{\mathbb{R}}}{\|f\|_W} \leq 1$ so the corresponding quotient of condition numbers worsens by a factor of at most $\sqrt{2q\mathbf{D}}$. Our main result derives from the fact that $\frac{\|f\|_\infty^{\mathbb{R}}}{\|f\|_W}$ is, for a substantial number of f's, much smaller than one: we can expect it to be smaller than $\sqrt{\mathbf{D}\ln(e\mathbf{D})/N}$ with very high probability.

Theorem 4.27. *Let $q \leq n+1$, $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ be dobro with parameters K and rho and $\ell \in \mathbb{N}$. Then for any power ℓ with $1 \leq \ell \leq \frac{\sqrt{N}}{2}$ we have*

$$\mathop{\mathbb{E}}_{\mathfrak{f}}\left(\frac{\|\mathfrak{f}\|_\infty^{\mathbb{R}}}{\|\mathfrak{f}\|_W}\right)^\ell \leq \left(\frac{140\sqrt{2}K\rho\sqrt{n\ln(eD)\ell}}{\sqrt{N-2\ell}}\right)^\ell.$$

Corollary 4.28. *Let $q \leq n+1$ and $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ be dobro with parameters K and ρ. Then for $1 \leq \ell \leq \frac{\sqrt{N}}{2}$ we have,*

$$\mathop{\mathbb{E}}_{\mathfrak{f}}\left(\frac{\mathsf{K}(\mathfrak{f})}{\kappa(\mathfrak{f})}\right)^\ell \leq \left(\frac{280K\rho\sqrt{qn\mathbf{D}\ln(eD)\ell}}{\sqrt{N-2\ell}}\right)^\ell.$$

□

Let PolyBetti$_W$ be the version of PolyBetti$_\infty$ using the Weyl norm. An analysis along the lines of [27] (or [14]) shows that the run-time of PolyBetti$_W$ is

$$2^{\mathcal{O}(n^2\log n)}\mathbf{D}^{10n}\kappa(f)^{10n}$$

which is very similar to the cost bound for $\textsc{PolyBetti}_\infty$ in Proposition 4.13. It follows from both bounds that

$$\frac{\text{run-time}(\textsc{PolyBetti}_\infty, f)}{\text{run-time}(\textsc{PolyBetti}_W, f)} \leq \left(\frac{\mathsf{K}(f)}{\kappa(f)}\right)^{10n}.$$

Using Corollary 4.28 and Markov's inequality, it is easy to prove the following estimate.

Corollary 4.29. *Let $q \leq n+1$, $N > 20n$ and let $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ be a dobro random polynomial tuple of parameters K and ρ,*

$$\frac{\textit{run-time}(\textsc{PolyBetti}_\infty, f)}{\textit{run-time}(\textsc{PolyBetti}_W, f)} \leq \left(\frac{896K\rho n\sqrt{q\mathbf{D}\ln(e\mathbf{D})}}{\sqrt{N-20n}}\right)^{10n}$$

with probability at least $1-1/N$. Note that for fixed n and large $\mathbf{D}$, the ratio in the right-hand side is of order

$$\left(\frac{K\rho\sqrt{\ln(e\mathbf{D})}}{D^{\frac{n-1}{2}}}\right)^{10n}. \qquad \square$$

We proceed to prove Theorem 4.27.

Proposition 4.30. *Let $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ be dobro with parameters K and ρ. Then, for all $t > 0$,*

$$\mathbb{P}\left(\|\mathfrak{f}\|_\infty^{\mathbb{R}} \geq t\right) \leq \frac{q3^{n-1}\mathbf{D}^n}{\sqrt{n+1}} e^{-t^2/(3K)^2}.$$

In particular, if $q \leq n+1$, for all $\ell \geq 1$, $\left(\mathbb{E}_{\mathfrak{f}}\left(\|\mathfrak{f}\|_\infty^{\mathbb{R}}\right)^\ell\right)^{\frac{1}{\ell}} \leq 10K\sqrt{n\ln(e\mathbf{D})\ell}$.

Proof of Theorem 4.27. By the Cauchy-Schwarz inequality,

$$\mathop{\mathbb{E}}_{\mathfrak{f}}\left(\frac{\|\mathfrak{f}\|_\infty^{\mathbb{R}}}{\|\mathfrak{f}\|_W}\right)^\ell \leq \sqrt{\mathop{\mathbb{E}}_{\mathfrak{f}}(\|\mathfrak{f}\|_\infty^{\mathbb{R}})^{2\ell}}\sqrt{\mathop{\mathbb{E}}_{\mathfrak{f}}\frac{1}{\|\mathfrak{f}\|_W^{2\ell}}}.$$

The first term in the right is bounded by Propositions 4.30.

For the second term, we will use [41, Theorem 1.11]. By Proposition 4.25 (applied with the identity and using orthogonal coordinates with respect the Weyl norm) and Striling's approximation, we have that $\mathfrak{f}$ has the SBA with constant $2\sqrt{\pi e}\rho$. Thus, by [41, Theorem 1.11],

$$\mathop{\mathbb{E}}_{\mathfrak{f}}\frac{1}{\|\mathfrak{f}\|_W^{2\ell}} \leq (14\rho)^{2\ell}\mathop{\mathbb{E}}_{\mathfrak{f}}\frac{1}{\|\mathfrak{g}\|_W^{2\ell}}$$

where $\mathfrak{g} \in \mathcal{H}_{\boldsymbol{d}}[q]$ is KSS. Now, for KSS $\mathfrak{g}$, $\|\mathfrak{g}\|_W^2$ is distributed according to a χ^2-distribution with N degrees of freedom. Therefore

$$\mathop{\mathbb{E}}_{\mathfrak{g}}\frac{1}{\|\mathfrak{g}\|_W^{2\ell}} = \int_0^\infty t^{-\ell}\frac{1}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)}t^{\frac{N}{2}-1}e^{-\frac{t}{2}}\,\mathrm{d}t = \frac{\Gamma\left(\frac{N}{2}-\ell\right)}{2^\ell\Gamma\left(\frac{N}{2}\right)} = \frac{1}{(N-2)(N-4)\cdots(N-2\ell)}.$$

The desired claim follows now. $\square$

Proof of Proposition 4.30. Fix $\delta \in [0, 1/\mathbf{D}]$. By the proof of Proposition 4.2, we have that $\|\mathfrak{f}\|_\infty^{\mathbb{R}} > t$ implies $\operatorname{vol}\left\{x \in \mathbb{S}^n \mid \|\mathfrak{f}(x)\|_\infty \geq \left(1 - \frac{\mathbf{D}^2}{2}\delta^2\right) t\right\} \geq \operatorname{vol} B_{\mathbb{S}}(x_*, \delta)$, where $x_* \in \mathbb{S}^n$ maximizes $\|f(x)\|_\infty$. Therefore

$$\mathbb{P}\left(\|\mathfrak{f}\|_\infty^{\mathbb{R}} \geq t\right) \leq \mathbb{P}_{\mathfrak{f}}\left(\mathbb{P}_{\mathfrak{x}\in\mathbb{S}^n}\left(\|\mathfrak{f}(\mathfrak{x})\|_\infty \geq \left(1 - \frac{\mathbf{D}^2}{2}\delta^2\right) t\right) \geq \operatorname{vol} B_{\mathbb{S}}(x_*, \delta)/\operatorname{vol}\mathbb{S}^n\right).$$

Now, by Stirling's approximation [13, Eq. (2.14)] and [13, Lemma 2.31] (plus some estimations of $\int_0^\delta \sin^{n-1}\theta \,\mathrm{d}\theta$), we have that

$$\operatorname{vol} B_{\mathbb{S}}(x_*, \delta)/\operatorname{vol}\mathbb{S}^n \geq 3\sqrt{n+1}\left(\frac{2}{5}\delta\right)^n.$$

In this way,

$$\begin{aligned}
&\mathbb{P}\left(\|\mathfrak{f}\|_\infty^{\mathbb{R}} \geq t\right) \\
&\leq \mathbb{P}_{\mathfrak{f}}\left(\mathbb{P}_{\mathfrak{x}\in\mathbb{S}^n}\left(\|\mathfrak{f}(\mathfrak{x})\|_\infty \geq \left(1 - \frac{\mathbf{D}^2}{2}\delta^2\right) t\right) \geq 3\sqrt{n+1}\left(\frac{2}{5}\delta\right)^n\right) \\
&\leq \frac{1}{3\sqrt{n+1}}\left(\frac{5}{2\delta}\right)^n \mathop{\mathbb{E}}_{\mathfrak{f}} \mathbb{P}_{\mathfrak{x}\in\mathbb{S}^n}\left(\|\mathfrak{f}(\mathfrak{x})\|_\infty \geq \left(1 - \frac{\mathbf{D}^2}{2}\delta^2\right) t\right) && \text{(Markov's inequality)} \\
&\leq \frac{1}{3\sqrt{n+1}}\left(\frac{5}{2\delta}\right)^n \mathop{\mathbb{E}}_{\mathfrak{x}\in\mathbb{S}^n} \mathbb{P}_{\mathfrak{f}}\left(\|\mathfrak{f}(\mathfrak{x})\|_\infty \geq \left(1 - \frac{\mathbf{D}^2}{2}\delta^2\right) t\right) && \text{(Tonelli's theorem)} \\
&\leq \frac{1}{3\sqrt{n+1}}\left(\frac{5}{2\delta}\right)^n \max_{x\in\mathbb{S}^n} \mathbb{P}_{\mathfrak{f}}\left(\|\mathfrak{f}(x)\|_\infty \geq \left(1 - \frac{\mathbf{D}^2}{2}\delta^2\right) t\right) \\
&\leq \frac{q}{3\sqrt{n+1}}\left(\frac{5}{2\delta}\right)^n \max_{i,x\in\mathbb{S}^n} \mathbb{P}_{\mathfrak{f}}\left(|\mathfrak{f}_i(x)|_\infty \geq \left(1 - \frac{\mathbf{D}^2}{2}\delta^2\right) t\right) && \text{(Union bound).}
\end{aligned}$$

Since $\mathfrak{f}$ is dobro, for all i and $x \in \mathbb{S}^n$, $\mathfrak{f}_i(x)$ is a subgaussian random variable with ψ_2-norm at most $\frac{5}{4}K$. Note that we are using that in the coordinates of an orthogonal monomial basis for the Weyl norm, the following holds: 1) a dobro random polynomial looks like random vector whose components are independent and subgaussian of ψ_2-norm at most K, and 2) evaluation at a point of the sphere becomes inner product with a vector of norm one.

Hence

$$\mathbb{P}\left(\|\mathfrak{f}\|_\infty^{\mathbb{R}} \geq t\right) \leq \frac{q}{3\sqrt{n+1}}\left(\frac{5}{2\delta}\right)^n \exp\left(-\left(1 - \frac{\mathbf{D}^2}{2}\delta^2\right)^2 \frac{8t^2}{25K^2}\right).$$

The claim follows taking $\delta = 5/(6\mathbf{D})$. For the other inequalities on the moments use Proposition 4.23. □

4.3.3 Complexity of Plantinga-Vegter Algorithm

In [23] (cf. [22]), we proved the following result (which we are just adapting to the notation[5] of this paper).

[5] Note that there is a slight difference in how we define the anti-concentration constant in [23] and how we define it here.

Theorem 4.31. *[23, Theorem 8.4 and Theorem 7.3] Let $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[1]$ be dobro with parameters K and ρ. For all $x \in \mathbb{S}^n$ and $t \geq e$,*

$$\mathbb{P}\left(\kappa(\mathfrak{f}, x) \geq t\right) \leq 2\left(\frac{N}{n+1}\right)^{\frac{n+1}{2}} (30K\rho)^{n+1} \frac{\ln^{\frac{n+1}{2}} t}{t^{n+1}}.$$

In particular, for Plantinga-Vegter algorithm implemented using the domain $[-a, a]^n$ the number of hypercubes in the subdivision is at most

$$a^n \mathbf{D}^n N^{\frac{n+1}{2}} 2^{n \log n + 13n + \frac{3}{2} \log n + \frac{17}{2}} (K\rho)^{n+1}. \qquad \square$$

Our objective is the following theorem, which shows how the $N^{\frac{n+1}{2}}$ factor vanishes from these estimates when we pass from κ to K. This shows that the version of Plantinga-Vegter using K is faster than the one using κ, i.e., the one in [23].

Theorem 4.32. *Let $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[1]$ be dobro with parameters K and ρ. For all $x \in \mathbb{S}^n$ and $t \geq e$,*

$$\mathbb{P}\left(\mathsf{K}(\mathfrak{f}, x) \geq t\right) \leq \mathbf{D}^{\frac{n}{2}} (\ln e\mathbf{D})^{\frac{n+1}{2}} (40K\rho)^{n+1} \frac{\ln^{\frac{n+1}{2}} t}{t^{n+1}}.$$

It follows that for every compact $\Omega \subseteq \mathbb{S}^n$,

$$\mathop{\mathbb{E}}_{\mathfrak{f}} \mathop{\mathbb{E}}_{\mathfrak{x} \in \Omega} \left(\mathsf{K}(\mathfrak{f}, \mathfrak{x})^n\right) \leq \mathbf{D}^{\frac{n}{2}} (\ln e\mathbf{D})^{\frac{n+1}{2}} 2^{\frac{1}{2} n \log n + 5n + 2\log(n) + \frac{19}{2}} (K\rho)^{n+1}.$$

In particular, for Plantinga-Vegter algorithm implemented using the domain $[-a, a]^n$ the number of hypercubes in the subdivision is at most

$$a^n \mathbf{D}^{\frac{3n}{2}} (\ln e\mathbf{D})^{\frac{n+1}{2}} 2^{\frac{3}{2} n \log n + 13n + 2\log(n) + \frac{19}{2}}$$

Remark 4.33. Theorem 4.32 allows us to compare the efficiency of Plantinga-Vegter for the versions based on the Weyl-norm and the ∞-norm. One can observe that (in the region of interest $\mathbf{D} > n$) the term $N^{\frac{n}{2}} \sim \mathbf{D}^{\frac{n^2}{2}}$ in the estimate for the Weyl-norm version is replaced with $(\mathbf{D} \log \mathbf{D})^{\frac{n}{2}}$ in the ∞-norm. Basically the exponent of $\mathbf{D}$ goes from $\mathcal{O}(n^2)$ to $\mathcal{O}(n)$. If we focus on the original cases of interest (cf. [44]), that is $n = 2$ and $n = 3$, with the average complexity analysis from [23], it is shown in Theorem 3.1 there that PV-Interval$_W$ has an average complexity of

$$\begin{aligned} &\mathcal{O}\left(d^8 \max\{1, a^2\} (K\rho)^3\right) && \text{for } n = 2, \text{ and} \\ &\mathcal{O}\left(d^{13} \max\{1, a^3\} (K\rho)^4\right) && \text{for } n = 3. \end{aligned}$$

It follows from Theorems 4.18 and 4.32 that the average complexity of PV-Interval$_\infty$ is

$$\begin{aligned} &\mathcal{O}\left(d^7 \log^{1.5}(d) \max\{1, a^2\} (K\rho)^3\right) && \text{for } n = 2, \text{ and} \\ &\mathcal{O}\left(d^{10} \log^2(d) \max\{1, a^3\} (K\rho)^4\right) && \text{for } n = 3. \end{aligned}$$

We next proceed to prove Theorem 4.32.

Proof of Theorem 4.32. Let $u, t \geq 0$, then

$$\begin{aligned} &\mathbb{P}_{\mathfrak{f}}\left(\mathsf{K}(\mathfrak{f}, x) \geq t\right) \\ &\quad \leq \mathbb{P}_{\mathfrak{f}}\left(\|f\|_\infty^{\mathbb{R}} \geq u \text{ or } \max\left\{|\mathfrak{f}(x)|, \frac{\|\mathrm{D}_x \mathfrak{f}\|}{\mathbf{D}}\right\} \leq \frac{u}{t}\right) && \text{(Implication bound)} \\ &\quad \leq \mathbb{P}_{\mathfrak{f}}\left(\|f\|_\infty^{\mathbb{R}} \geq u\right) + \mathbb{P}_{\mathfrak{f}}\left(\max\left\{|\mathfrak{f}(x)|, \frac{\|\mathrm{D}_x \mathfrak{f}\|}{\mathbf{D}}\right\} \leq \frac{u}{t}\right), && \text{(Union bound)} \end{aligned}$$

where we used the fact that for $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[1]$, $\mathsf{K}(f,x) = \|f\|_\infty^{\mathbb{R}} / \max\{|f(x)|, \|\mathrm{D}_x f\|/\mathbf{D}\}$.

On the one hand, $\mathbb{P}_{\mathfrak{f}}\left(\|f\|_\infty^{\mathbb{R}} \geq u\right)$ is bounded by Proposition 4.30. On the other hand, the map

$$f \mapsto \begin{pmatrix} f(x) & \frac{\mathrm{D}_x f}{\mathbf{D}} \end{pmatrix}$$

has singular values $1, 1/\sqrt{\mathbf{D}}, \ldots, 1/\sqrt{\mathbf{D}}$ in the coordinates of a monomial basis orthogonal with respect to the Weyl inner product. And since in such a basis a dobro polynomial is a vector whose coefficients are independent and have the anti-concentration property with constant ρ, we deduce that

$$\begin{aligned}\mathbb{P}_{\mathfrak{f}}\left(\max\left\{|\mathfrak{f}(x)|, \frac{\|\mathrm{D}_x\mathfrak{f}\|}{\mathbf{D}}\right\} \leq \frac{u}{t}\right) &\leq \mathbf{D}^{\frac{n}{2}} \operatorname{vol}\left\{(x_0,x) \in \mathbb{R}^{n+1} \mid |x_0|, \|x\| \leq u/t\right\} \left(\sqrt{2}\rho\right)^{n+1} \\ &\leq \omega_n \mathbf{D}^{\frac{n}{2}} \left(\frac{\sqrt{2}\rho u}{t}\right)^{n+1} \leq 9^n \mathbf{D}^{\frac{n}{2}} \left(\frac{u\rho}{\sqrt{n}}\right)^{n+1} \frac{1}{t^{n+1}},\end{aligned}$$

where ω_n is the volume of the n-ball, using Proposition 4.25 and Stirling's estimation [13, Eq. (2.14)].

Hence, combining the inequalities above,

$$\mathbb{P}_{\mathfrak{f}}\left(\mathsf{K}(\mathfrak{f},x) \geq t\right) \leq \frac{3^{n-1}\mathbf{D}^n}{\sqrt{n+1}} e^{-u^2/(3K)^2} + 9^n \mathbf{D}^{\frac{n}{2}} \left(\frac{u\rho}{\sqrt{n}}\right)^{n+1} \frac{1}{t^{n+1}}.$$

Taking $t \geq e$ and $u = 3K\sqrt{n\ln(3e^2\mathbf{D})\ln t} \geq 3K\sqrt{n\ln(3\mathbf{D}) + (n+1)\ln t}$, we get

$$\mathbb{P}_{\mathfrak{f}}\left(\mathsf{K}(\mathfrak{f},x) \geq t\right) \leq \frac{1}{3\sqrt{n+1}} \frac{1}{t^{n+1}} + 9^n \mathbf{D}^{\frac{n}{2}} \left(3K\rho \ln^{\frac{1}{2}}(3e^2\mathbf{D})\right)^{n+1} \frac{\ln^{\frac{n+1}{2}} t}{t^{n+1}}.$$

This proves the first statement.

By Tonelli's theorem, in order to prove the second statement it is enough to bound $\mathbb{E}_{\mathfrak{f}} \mathsf{K}(\mathfrak{f},x)^n$ for a fixed $x \in \mathbb{S}^n$. Now,

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathfrak{f}} \mathsf{K}(\mathfrak{f},x)^n &= \int_0^\infty \mathbb{P}_{\mathfrak{f}}\left(\mathsf{K}(\mathfrak{f},x) \geq t^{\frac{1}{n}}\right) \\ &\leq e^n + \int_{e^n}^\infty \mathbb{P}_{\mathfrak{f}}\left(\mathsf{K}(\mathfrak{f},x) \geq t^{\frac{1}{n}}\right) \leq e^n + \int_{e^n}^\infty \mathbf{D}^{\frac{n}{2}} \ln^{\frac{n+1}{2}}(e\mathbf{D})(40K\rho)^{n+1} \frac{\ln(t^{\frac{1}{n}})^{\frac{n+1}{2}}}{t^{1+\frac{1}{n}}}\,\mathrm{d}t \\ &\leq e^n + \mathbf{D}^{\frac{n}{2}} \ln^{\frac{n+1}{2}}(e\mathbf{D})(40K\rho)^{n+1} \int_1^\infty \frac{\ln(t^{\frac{1}{n}})^{\frac{n+1}{2}}}{t^{1+\frac{1}{n}}}\,\mathrm{d}t.\end{aligned}$$

By changing variables, $t = e^{sn}$, we can see that

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

where the inequality comes from Stirling's approximation [13, Eq. (2.14)]. Hence we get

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathfrak{f}} \mathsf{K}(\mathfrak{f},x)^n &\leq e^n + \sqrt{2\pi e} n\sqrt{n+1}\mathbf{D}^{\frac{n}{2}} \ln^{\frac{n+1}{2}}(e\mathbf{D})(18\sqrt{n+1}K\rho)^{n+1} \\ &\leq 8n\sqrt{n+1}\mathbf{D}^{\frac{n}{2}} \ln^{\frac{n+1}{2}}(e\mathbf{D})(18\sqrt{n+1}K\rho)^{n+1}.\end{aligned}$$

The second bound now follows after some easy bounds. □

5 Linear Homotopy for Computing Complex Zeros

Smale's 17th problem asks if a complex zero of n complex polynomial equations in n unknowns be found on average polynomial time [49]. A probabilistic solution to Smale's 17th problem was given by Beltrán and Pardo in 2009 [6, 7]. The construction of Beltrán and Pardo was probabilistic in the sense that they exhibited a *randomized* algorithm.

The distribution underlying the average-case analysis for Beltrán-Pardo algorithm is a Gaussian vector in $\mathbb{C}^N$ where the coordinates have prescribed variances coming from multinomial coefficients, or equivalently the distribution is $\mathbb{N}(0, \mathbb{I})$ when the vector space $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$ is endowed with the Weyl inner product; this particular distribution is called KSS complex random polynomial (see Subsection 5.3) . Finally, the algorithm of Beltrán-Pardo is polynomial time with respect to $N = \dim_{\mathbb{C}} \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$.

When we have n equations and n unknowns, i.e. $q = n$, we call it a square system. A generic square system of equations with degrees $d_1, d_2, \ldots, d_n$ has $\mathcal{D} := d_1 \cdot \cdots \cdot d_n$ many zeros, and Smale's 17th problems asks to compute one of these zeros. Following the initial work by Shub and Smale [48], the hearth of Beltrán-Pardo solution is a *linear homotopy*, let's call it ALH. It takes as input the system f for which a zero is sought, along with an initial pair $(g, \zeta) \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n] \times \mathbb{P}^n$ satisfying $g(\zeta) = 0$. If we define $q_t := tf + (1-t)g$, for $t \in [0,1]$, then generically, the segment $[g, f]$ in $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$ lifts to a curve $\{(q_t, \zeta_t) \mid t \in [0,1]\}$ in the *solution variety*

$$\mathcal{V} := \{(f, \zeta) \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n] \times \mathbb{P}^n \mid f(\zeta) = 0\}.$$

The idea of ALH, in a nutshell, is to "follow" this curve (for which we know its origin (g, ζ)) close enough so that we end up with an approximation to the zero ζ_1 of $f = q_1$.

The breakthrough in [6, 7] was to come up with a randomized algorithm to produce the (long sought) initial pair (g, ζ). To state this result, we endow $\mathcal{V}$ with the *standard distribution* ρ_{std} defined via the following procedure:

- draw a KSS complex random polynomial $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$.
- draw ζ from the $\mathcal{D}$ zeros of f with the uniform distribution.

For details on ρ_{std} see [13, §17.5]. The following version of Beltrán and Pardo randomization allows one to efficiently sample from ρ_{std}.

Proposition 5.1. *([13, Proposition 17.21]) There is a randomized algorithm which, with input n and $\boldsymbol{d}$, returns a pair $(g, \zeta) \in \mathcal{V}$ drawn from ρ_{std}. The algorithm performs $2(N + n^2 + n + 1)$ draws of random real numbers from the standard Gaussian distribution and $O(\mathbf{D}nN + n^3)$ arithmetic operations.* □

With the randomization procedure at hand, the structure of the algorithm to compute approximate zeros is simple.

Algorithm 5.1: SOLVE

Input : $f \in \mathcal{H}_{\boldsymbol{d}}[n]$
Precondition : $f \neq 0$

draw $(g, \zeta) \in \mathcal{V}$ from ρ_{std}
run ALH on input (f, g, ζ)

Output : $z \in \mathbb{C}_*^{n+1}$
Postcondition: z is an approximate zero of f
Halting cond.: The lifting of $[g, f]$ at ζ does not cut $\tilde{\Sigma} \subseteq \mathcal{V}$

Here $\tilde{\Sigma} := \{(f, \zeta) \in \mathcal{V} \mid \det \mathrm{D}_\zeta f = 0\}$. This set has complex codimension 1 in $\mathcal{V}$. Hence, because the lifting of the segment $[g, f]$ corresponding to ζ has real dimension 1, generically, it does not cut $\tilde{\Sigma}$. That is, algorithm SOLVE almost surely terminates for almost all inputs $f \in \mathcal{H}_{\boldsymbol{d}}[n]$.

Regarding complexity, the total cost of SOLVE is dominated by that of running ALH, which is given by the number of steps K performed by the homotopy times the cost of each step. In previous work ([48, 6, 7, 12, 3] among others) the latter is essentially optimal as it is $O(N + n^3)$ (which is $O(N)$ if $d_i \geq 2$ for $i = 1, \dots, n$). The former depends on the input at hand and it is there where average considerations play a role. In [8, 12] ALH∞ was implemented using the Weyl norm to compute step-lengths. Its average complexity is $O(n\mathbf{D}^{3/2}N)$. The average complexity of the resulting algorithm, let us call it SOLVE$_W$, is then $O(n\mathbf{D}^{3/2}N^2)$.

The goal of this section is to analyze a version ALH$_\infty$ of ALH with step-lengths based on $\| \ \|_\infty$. We show that this can be done in a straightforward manner and that, maybe surprisingly, the average complexity of ALH with step lengths based on our new condition number is $\mathcal{O}(n^3\mathbf{D}^2 \ln(n\mathbf{D}))$; the cost is independent of N. Unfortunately, this computational gain is not decisive for a general input model due to high cost of computing $\| \ \|_\infty$ norms.

Nonetheless, for the particular —but highly relevant— case of quadratic polynomials, we can efficiently compute the ∞-norm and this results in SOLVE$_\infty$ being faster than SOLVE$_W$.

5.1 Description of the linear homotopy

The algorithm below is, essentially, the one in [12] and [13, Ch. 17]. The only change is in the computation of the step-length Δ_t where we replace the original (here $\mathrm{dist}_\mathbb{S}$ denotes angle)

$$\frac{0.008535284}{\mathrm{dist}_\mathbb{S}(f, g)\mathbf{D}^{3/2}\mu^2(q, z)}$$

by

$$\frac{0.03\, \|q\|_\infty^\mathbb{C}}{\|f - g\|_\infty^\mathbb{C} \mathbf{D}\mathsf{M}^2(q, z)}. \tag{5.1}$$

This change amounts —leaving aside difference in constants and a smaller exponent in $\mathbf{D}$— to the use of the ∞-norm instead of the Weyl one, and consequently, the use of M instead of μ.

Algorithm 5.2: ALH∞

Input : $f, g \in \mathcal{H}_{\boldsymbol{d}}[n]$ and $\zeta \in \mathbb{P}^n$
Precondition : $g(\zeta) = 0$

$t \leftarrow 0$, $q \leftarrow g$, $z \leftarrow \zeta$
repeat
  $\Delta t \leftarrow \frac{0.03\,\|q\|_\infty}{\|f-g\|_\infty \mathbf{D}\mathsf{M}^2(q,z)}$
  $t \leftarrow \min\{t + \Delta t, 1\}$
  $q \leftarrow tf + (1-t)g$
  $z \leftarrow N_q(z)$
until $t = 1$
return z and halt

Output : $z \in \mathbb{C}^{n+1}_*$
Postcondition: The algorithm halts if $q_t \notin \Sigma_{\zeta_t}$ for all $t \in [0, 1]$. In this case, z is an approximate zero of f

5.2 A bound on the number of iterations

The analysis of ALH∞ closely follows the steps in [13]. It requires, however, to prove for $\| \; \|_\infty$ a number of results we know for the Weyl norm. The first one shows that M is a condition number in the standard sense of this expression (it measures how solutions change when data is perturbed). To simplify the notation, in the rest of this section, we will often omit the reference to the base field $\mathbb{C}$.

Theorem 5.2. *Suppose that the lifting of the segment $[g, f]$ in $\mathcal{V}$ corresponding to ζ does not cut Σ'. Then the algorithm* ALH∞ *stops after at most K steps with*

$$K \leq 45\,\mathbf{D}\,\|f - g\|_\infty \int_0^1 \frac{\mathsf{M}^2(q_t, \zeta_t)}{\|q_t\|_\infty}\mathrm{d}t.$$

The returned point z is an approximate zero of f with associated zero ζ_1.

Corollary 5.3. *The bound K in Theorem 5.2 satisfies*

$$K \leq 45\,n\,\mathbf{D} \int_0^1 (\|f\|_\infty + \|g\|_\infty)^2 \|\mathrm{D}_{\zeta_t} q_t^{-1} \Delta\|^2 \mathrm{d}t.$$

Proposition 5.4. *Let $t \mapsto (f_t, \zeta_t) \in V$ be a smooth path. Then, for all t,*

$$\|\dot{\zeta}_t\| \leq \mathsf{M}(f_t, \zeta_t) \frac{\|\dot{f}_t\|_\infty}{\|f_t\|_\infty}.$$

Proof in Theorem 5.2. The proof follows the lines of [13, Theorem 17.3]. We will therefore only offer a brief sketch. Set $\varepsilon := \frac{1}{4}$ and $C = \frac{\varepsilon}{4} = \frac{1}{16}$. Let $q_t := tf + (1-t)g$. Also, let $0 < t_1 < \ldots < t_K = 1$ and $\zeta_0 = z_0, \ldots, z_K$ be the sequence of t-values and points in $\mathbb{P}^n$, respectively, generated by the algorithm in its K first iterations. To simplify notation we write q_i and ζ_i instead of q_{t_i} and ζ_{t_i}.

As in [13, Theorem 17.3], but using Proposition 3.6 in the place of [13, Proposition 16.2] and Theorem 3.5 in the place of [13, Theorem 16.1], one proves by induction the following statements for $i = 0, \dots, K-1$:

(a,i) $\mathrm{dist}_{\mathbb{P}}(z_i, \zeta_i) \le \frac{C}{\mathbf{D}\mathsf{M}(q_i,\zeta_i)}$
(b,i) $\frac{\mathsf{M}(q_i,z_i)}{1+\varepsilon} \le \mathsf{M}(q_i,\zeta_i) \le (1+\varepsilon)\mathsf{M}(q_i,z_i)$
(c,i) $\|q_i - q_{i+1}\|_\infty \le \frac{C\|q_i\|_\infty}{\mathbf{D}\mathsf{M}(q_i,\zeta_i)}$
(d,i) $\mathrm{dist}_{\mathbb{P}}(\zeta_i, \zeta_{i+1}) \le \frac{C}{\mathbf{D}\mathsf{M}(q_i,\zeta_i)}\frac{1-\varepsilon}{1+\varepsilon}$
(e,i) $\mathrm{dist}_{\mathbb{P}}(z_i, \zeta_{i+1}) \le \frac{2C}{(1+\varepsilon)\mathbf{D}\mathsf{M}(q_i,\zeta_i)}$
(f,i) z_i is an approximate zero of q_{i+1} with associated zero ζ_{i+1}

We proceed by induction showing that:

- $(\mathrm{a}, i) \Rightarrow (\mathrm{b}, i) \Rightarrow ((\mathrm{c}, i)\ \&\ (\mathrm{d}, i))$
- $((\mathrm{a}, i)\ \&\ (\mathrm{d}, i)) \Rightarrow (\mathrm{e}, i)$
- $((\mathrm{c}, i), (\mathrm{d}, i)\ \&\ (\mathrm{e}, i)) \Rightarrow ((\mathrm{f}, i)\ \&\ (\mathrm{a}, i+1))$.

The base case, $(\mathrm{a}, 0)$, is trivial.
$(\mathrm{a}, i) \Rightarrow (\mathrm{b}, i)$: By assumption,

$$\mathbf{D}\,\mathsf{M}(q_i,\zeta_i)\mathrm{dist}_{\mathbb{P}}(z_i,\zeta_i) \le C \le \frac{\varepsilon}{4}$$

and so, by Proposition 3.6,

$$\frac{\mathsf{M}(q_i,z_i)}{1+\varepsilon} \le \mathsf{M}(q_i,\zeta_i) \le (1+\varepsilon)\mathsf{M}(q_i,z_i).$$

Thus (b, i) holds.
$(\mathrm{b}, i) \Rightarrow ((\mathrm{c}, i)\ \&\ (\mathrm{d}, i))$:
By definition of q_t, we have that for $t \in [t_i, t_{i+1}]$,

$$\|q_t - q_i\|_\infty = \|(t-t_i)(f-g)\|_\infty = |t-t_i|\|f-g\|_\infty \le \Delta t_i\|f-g\|_\infty,$$

and so

$$\|q_t - q_i\|_\infty \le \frac{0.03\|q_i\|_\infty}{\mathbf{D}\,\mathsf{M}(q_i,z_i)^2} \le \frac{C\|q_i\|_\infty}{\mathbf{D}\,\mathsf{M}(q_i,\zeta_i)}, \tag{5.2}$$

where we use that $\mathsf{M}(f, z_i) \ge 1$, by the 1st Lipschitz property (Theorem 3.5), (b, i) and our choice of C. This shows (c, i).

Let $t \in [t_i, t_{i+1}]$. Because of the continuity of $t \mapsto \zeta_t$, we can assume t sufficiently small, so that

$$\mathbf{D}\,\mathsf{M}(f_i,\zeta_i)\mathrm{dist}_{\mathbb{P}}(\zeta_i,\zeta_t) \le C = \frac{\varepsilon}{4}, \tag{5.3}$$

holds. It follows from (5.2) and (5.3) that the hypothesis of Proposition 3.6 hold. Then

$$\begin{aligned}
\mathrm{dist}_{\mathbb{P}}(\zeta_i,\zeta_t) &= \int_{t_i}^{t} \|\dot{\zeta}_s\|\,\mathrm{d}s \\
&\leq \int_{t_i}^{t} \mathsf{M}(q_s,\zeta_s)\frac{\|\dot{q}_s\|_\infty}{\|q_s\|_\infty}\,\mathrm{d}s && \text{(Proposition 5.4)}\\
&\leq \int_{t_i}^{t} (1+\varepsilon)\mathsf{M}(q_i,\zeta_i)\frac{\|\dot{q}_s\|_\infty}{\|q_s\|_\infty}\,\mathrm{d}s && \text{(Proposition 3.6)}\\
&= \int_{t_i}^{t} \frac{(1+\varepsilon)\|f-g\|_\infty\mathsf{M}(q_i,\zeta_i)\mathrm{d}s}{\|q_s\|_\infty} && (\|\dot{q}_s\|_\infty=\|f-g\|_\infty)\\
&\leq \int_{t_i}^{t} \frac{(1+\varepsilon)\|f-g\|_\infty\mathsf{M}(q_i,\zeta_i)\mathrm{d}s}{\|q_i\|_\infty-(s-t_i)\|f-g\|_\infty} && \text{(Triangle inequality)}\\
&\leq \frac{(1+\varepsilon)\|f-g\|_\infty\mathsf{M}(q_i,\zeta_i)(t-t_i)}{\|q_i\|_\infty\left(1-\frac{\|f-g\|_\infty}{\|q_i\|_\infty}(t-t_i)\right)}\\
&\leq \frac{(1+\varepsilon)\|f-g\|_\infty\mathsf{M}(q_i,\zeta_i)\Delta t_i}{\|q_i\|_\infty\left(1-\frac{\|f-g\|_\infty}{\|q_i\|_\infty}\Delta t_i\right)} && (t-t_i\leq t_{i+1}-t_i=\Delta t_i)\\
&\leq \frac{C}{\mathbf{D}\,\mathsf{M}(q_i,z_i)}\frac{1-\varepsilon}{(1+\varepsilon)^2} && \text{(See (5.4) below)}\\
&\leq \frac{C}{\mathbf{D}\,\mathsf{M}(q_i,\zeta_i)}\frac{1-\varepsilon}{1+\varepsilon} && (\mathrm{b},i)
\end{aligned}$$

where we note that the denominator in the integrand doesn't vanish since for $t\in[t_i,t_{i+1}]$, $t-t_i\leq t_{i+1}-t_i<\|q_i\|_\infty/\|f-g\|_\infty$ by construction (as $C<1$ and $\mathsf{M}(q_i,z_i)\geq 1$). In the inequality before the last one, we have used that

$$\Delta t_i \leq \frac{C\|q_i\|_\infty}{\|f-g\|_\infty\mathbf{D}\,\mathsf{M}(q_i,\zeta_i)^2}\frac{(1-C)(1-\varepsilon)}{(1+\varepsilon)^2}\quad\text{and}\quad \Delta t_i\leq\frac{\|q_i\|_\infty}{\|f-g\|_\infty}C, \tag{5.4}$$

which hold due to the choice of the constant in the step-size Δt_i and $\mathsf{M}(q_i,z_i)\geq 1$.

The upper bound obtained implies that (5.3) holds for all $t\in[t_i,t_{i+1}]$. If (5.3) holds for t_{i+1}, this is obvious. If it does not hold, then we can take $t_*=\inf\{t\in[t_i,t_{i+1}]\mid \mathbf{D}\,\mathsf{M}(f_i,\zeta_i)\mathrm{dist}_{\mathbb{P}}(\zeta_i,\zeta_t)>C\}\in[t_i,t_{i+1}]$. By continuity, it satisfies (5.3), and so, by our deduction above,

$$\mathrm{dist}_{\mathbb{P}}(\zeta_i,\zeta_t)\leq\frac{C}{\mathbf{D}\,\mathsf{M}(q_i,\zeta_i)}\frac{1-\varepsilon}{1+\varepsilon}<\frac{C}{\mathbf{D}\,\mathsf{M}(q_i,\zeta_i)},$$

which gives a contradiction by continuity with t_* being the infimum.
$((\mathrm{a},i)\ \&\ (\mathrm{d},i))\Rightarrow(\mathrm{e},i)$. By the triangle inequality,

$$d_{\mathbb{S}}(z_i,\zeta_{i+1})\leq d_{\mathbb{S}}(z_i,\zeta_i)+d_{\mathbb{S}}(\zeta_i,\zeta_{i+1}).$$

This inequality, together with (a,i) and (d,i), proves (e,i).
$((\mathrm{c},i),,(\mathrm{d},i)\ \&\ (\mathrm{e},i))\Rightarrow((\mathrm{f},i)\ \&\ (\mathrm{a},i+1))$. By the Higher Derivative Estimate (Theorem 3.5), we have

$$\gamma(q_{i+1},\zeta_{i+1})d_{\mathbb{S}}(z_i,\zeta_{i+1})\leq\frac{1}{2}\mathbf{D}\mathsf{M}(q_{i+1},\zeta_{i+1})d_{\mathbb{S}}(z_i,\zeta_{i+1})<0.17708,$$

where the last inequality follows from (d, i) and our choices for C and ε. Thus, by Theorem 3.3, z_i is an approximate zero of q_{i+1} with associated zero ζ_{i+1}. This proves (f, i). Moreover, we have that $z_{i+1} := N_{q_{i+1}}(z+i)$ satisfies

$$d_{\mathbb{P}}(z_{i+1}, \zeta_{i+1}) \leq \frac{1}{2} d_{\mathbb{P}}(z_i, \zeta_{i+1}) \leq \frac{C}{(1+\varepsilon)\mathbf{D}\mathsf{M}(q_i, \zeta_i)} \leq \frac{C}{\mathbf{D}\mathsf{M}(q_{i+1}, \zeta_{i+1})},$$

where the first inequality follows from (e, i) and the second one from Proposition 3.6, (c, i) and (d, i). This proves $(\mathrm{a}, i+1)$.

By Proposition 3.6, (c, i), (d, i) and our choice of C and ε, we have that for all $t \in [t_i, t_{i+1}]$,

$$\frac{4}{5}\mathsf{M}(q_i, \zeta_i) \leq \mathsf{M}(q_t, \zeta_t) \leq \frac{5}{4}\mathsf{M}(q_i, \zeta_i). \tag{5.5}$$

And, by the triangle inequality and (b, i), for $t \in [t_i, t_{i+1}]$,

$$\frac{\|q_t\|_\infty}{\|q_i\|_\infty} \leq 1 + C = \frac{17}{16}. \tag{5.6}$$

The statement now easily follows. Consider any $i \in \{0, 1, \ldots, K-1\}$. Then

$$\begin{aligned} \int_{t_i}^{t_{i+1}} \frac{\mathsf{M}^2(q_t, \zeta_t)}{\|q_t\|_\infty} \mathrm{d}t &\geq \frac{64}{85} \int_{t_i}^{t_{i+1}} \frac{\mathsf{M}^2(q_i, z_i)}{\|q_i\|_\infty} \mathrm{d}t = \frac{64}{85} \frac{\mathsf{M}^2(q_i, z_i)}{\|q_i\|_\infty} |t_{i+1} - t_i| && ((5.5), (5.6)) \\ &= \frac{64}{85} \frac{0.03}{\|f-g\|_\infty \mathbf{D}} && (\mathrm{ALH}\infty) \end{aligned}$$

Hence

$$\int_0^1 \frac{\mathsf{M}^2(q_t, z_t)}{\|q_t\|_\infty} \mathrm{d}t \geq \frac{192}{8500} \frac{K}{\|f-g\|_\infty \mathbf{D}} \geq \frac{K}{45\|f-g\|_\infty \mathbf{D}}$$

and the result follows. □

Proof of Corollary 5.3. It immediately follows from the definition of $\mathsf{M}(q_t, \zeta_t)$ and the inequality $\|q_t\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. □

Proof of Proposition 5.4. Recall from [13, § 14.3] that the zero ζ_t is given by $\zeta_t = G(f_t)$ where $G : U \subset \mathcal{H}_{\boldsymbol{d}}[n] \to \mathbb{P}^n$ is a local inverse of the projection $\pi_1 : \mathcal{V} \to \mathcal{H}_{\boldsymbol{d}}[n]$. Hence, for all $\dot{f}_t \in \mathcal{H}_{\boldsymbol{d}}[n]$ we have

$$\dot{\zeta}_t = \mathrm{D}_{f_t} G(\dot{f}_t) = -(\mathrm{D}_{\zeta_t} f_t)^{-1}(\dot{f}_t(\zeta_t)) \tag{5.7}$$

where the second equality is shown in the course of the proof of [13, Prop. 16.10]. Using this equality along with the fact that $(\mathrm{D}_{\zeta_t} f_t)^{-1} = (\mathrm{D}_{\zeta_t} f_t)^\dagger$ (as $q = n$) we deduce that

$$\begin{aligned} \|\dot{\zeta}_t\| &= \max_{\|\dot{f}_t\|_\infty = 1} \|(\mathrm{D}_{\zeta_t} f_t)^{-1}(\dot{f}_t(\zeta_t))\| && (\text{By } (5.7)) \\ &\leq \left(\max_{\|\dot{f}_t\|_\infty = 1} \|\dot{f}_t(\zeta_t)\|\right) \|(\mathrm{D}_{\zeta_t} f_t)^{-1}\| && (\text{Operator norm inequality}) \\ &\leq \sqrt{n} \left(\max_{\|\dot{f}_t\|_\infty = 1} \|\dot{f}_t(\zeta_t)\|_\infty\right) \|(\mathrm{D}_{\zeta_t} f_t)^{-1}\| && \|\ \| \leq \sqrt{n}\|\ \|_\infty \\ &\leq \sqrt{n}\|(\mathrm{D}_{\zeta_t} f_t)^{-1}\| && (\text{Definition of } \|\ \|_\infty) \\ &= \frac{\sqrt{n}\|f_t\|_\infty \|(\Delta^{-1}\mathrm{D}_{\zeta_t} f_t)^{-1}\|}{\|f_t\|_\infty} = \frac{\mathsf{M}(f_t, \zeta_t)}{\|f_t\|_\infty}. && (\text{Definition of } \mathsf{M}) \end{aligned}$$

We recall that the norms where we have omitted the subscript correspond to the usual norm in the case of vector and the usual operator norm in the case of linear maps. □

5.3 Average complexity analysis of $\textsc{Solve}_\infty$

The execution of $\textsc{Solve}_\infty$ on an input $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$ amounts to calling ALH∞ on input $(f, \mathfrak{g}, \mathfrak{z})$ where $(\mathfrak{g}, \mathfrak{z}) \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n] \times \mathbb{P}^n$ is a standard random pair. Consequently, the number of iterations of $\textsc{Solve}_\infty$ amounts to the number of iterations done by ALH∞. The latter is a random variable as $(\mathfrak{g}, \mathfrak{z})$ is random. We will further consider f random and bound the average complexity of $\textsc{Solve}$ by taking the expectation over both $(\mathfrak{g}, \mathfrak{z})$ and f. Recall that a *KSS complex random polynomial system* $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$ is a tuple of random polynomials

$$\left(\sum_{|\alpha|=d_1} \binom{d_1}{\alpha}^{\frac{1}{2}} \mathfrak{c}_{1,\alpha} X^\alpha, \ldots, \sum_{|\alpha|=d_n} \binom{d_n}{\alpha}^{\frac{1}{2}} \mathfrak{c}_{n,\alpha} X^\alpha\right)$$

such that the $\mathfrak{c}_{i,\alpha}$ are i.d.d. complex normal random variables of mean zero and variance one.

Our main result is the following one.

Theorem 5.5. *Let $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$. On input $\mathfrak{f}$, Algorithm $\textsc{Solve}_\infty$ halts with probability 1 and it performs*

$$\mathcal{O}(n^3 \mathbf{D}^2 \ln(e\mathbf{D}))$$

iteration steps on average.

Remark 5.6. The bound in Theorem 5.5 is independent on N: it is a polynomial in n and $\mathbf{D}$. The possibility of such a bound for the number of iterations of a linear homotopy was explored in [3], where the dependence on N was reduced from linear to $\mathcal{O}(\sqrt{N})$. Pierre Lairez subsequently exhibited one such bound but for a *rigid* homotopy [37]. To the best of our knowledge, Theorem 5.5 is the first result doing so for a linear homotopy.

We will use the following two results. The first is the complex version of Proposition 4.30 and has an almost identical proof. The main difference lies in the needed volume computations as the geometry of the complex projective space $\mathbb{P}^n$ is somewhat different from that of the real sphere $\mathbb{S}^n$. The second is a known result on random complex Gaussian matrices.

Proposition 5.7. *Let $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$ be a KSS complex random polynomial tuple. Then, for all $t > 0$,*

$$\mathbb{P}\left(\|\mathfrak{f}\|_\infty^{\mathbb{C}} \geq t\right) \leq 2n \left(\frac{3\mathbf{D}}{2}\right)^{2n} e^{-(t/3)^2}.$$

In particular, for all $\ell \geq 1$, $\left(\mathbb{E}_{\mathfrak{f}} \left(\|\mathfrak{f}\|_\infty^{\mathbb{C}}\right)^\ell\right)^{\frac{1}{\ell}} \leq 12\sqrt{\ell n \ln(eD)}$. □

Proposition 5.8. *[13, Proposition 4.27]. Let $\mathfrak{A} \in \mathbb{C}^{n\times(n+1)}$ be a random complex matrix whose entries are i.i.d. complex normal Gaussian variables. Then for all $t \geq 0$,*

$$\operatorname{Prob}\left\{\|\mathfrak{A}^\dagger\| \geq t\right\} \leq \frac{1}{16} \frac{n^2}{t^4}.$$

In particular, for $\ell \in [1, 4)$, $\left(\mathbb{E}_{\mathfrak{A}} \|\mathfrak{A}^\dagger\|^\ell\right)^{\frac{1}{\ell}} \leq \frac{\sqrt{n}}{2} \left(\frac{4}{4-\ell}\right)^{\frac{1}{\ell}}$ □

Proof of Theorem 5.5. We are calling Algorithm ALH∞ with input $(\mathfrak{f},\mathfrak{g},\mathfrak{z})$ where $\mathfrak{f}\in\mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$ is a KSS complex polynomial system and $(\mathfrak{g},\mathfrak{z})\in\mathcal{H}_{\boldsymbol{d}}[n]$ is an standard pair.

Let $\Sigma:=\{h\in\mathcal{H}_{\boldsymbol{d}}[n]\mid \exists\zeta\in\mathbb{P}^n \text{ such that } (h,\zeta)\in\tilde{\Sigma}\}$. By classic results in algebraic geometry, this set is a complex algebraic hypersurface and so it has real codimension 2. Hence, with probability one, the segment $[\mathfrak{g},\mathfrak{f}]$ does not intersect it and, for each zero $\zeta^{(i)}$ of $\mathfrak{g}$, we obtain a unique lifted path

$$t\mapsto(\mathfrak{q}_t,\zeta_t^{(i)})\in\mathcal{V}.$$

Here, for each t, the $\zeta_t^{(i)}$ cover all the $d_1\cdots d_n$ different zeros of $\mathfrak{q}_t:=t\mathfrak{f}+(1-t)\mathfrak{g}$. Recall that behind this lifting lies the fact that the map $\mathcal{V}\setminus\tilde{\Sigma}\mapsto\mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]\setminus\Sigma$, $(f,\eta)\mapsto f$, is a regular covering map of degree $d_1\cdots d_n$.

In this way, the random zero $\mathfrak{z}$ of $\mathfrak{g}$ defines, following its lifted path, a zero $\mathfrak{z}_t$ of $\mathfrak{q}_t$. Moreover, since the original $\mathfrak{z}$ is chosen uniformly, the $\mathfrak{z}_t$ is a uniformly chosen zero of $\mathfrak{q}_t$. Hence

$$\left(\frac{\mathfrak{q}_t}{\sqrt{t^2+(1-t)^2}},\mathfrak{z}_t\right)\in\mathcal{V}$$

is a standard random pair, since $\frac{\mathfrak{q}_t}{\sqrt{t^2+(1-t)^2}}$ is a KSS complex random polynomial and $\mathfrak{z}_t$ is a uniformly drawn zero of this system.

By Corollary 5.3, the number of iterations of SOLVE$_\infty$ with input $\mathfrak{f}$ is given by

$$45n\,\mathbf{D}\int_0^1\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}\left((\|\mathfrak{f}\|_\infty^2+\|\mathfrak{g}\|_\infty^2)^2\|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{-1}\Delta\|^2\right)\,\mathrm{d}t,\tag{5.8}$$

where we have moved the expectation inside the integral using Tonelli's theorem. Now, by Hölder's inequality,

$$\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}\left((\|\mathfrak{f}\|_\infty^2+\|\mathfrak{g}\|_\infty^2)^2\|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{-1}\Delta\|^2\right)\leq\left(\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}(\|\mathfrak{f}\|_\infty^2+\|\mathfrak{g}\|_\infty^2)^6\right)^{\frac{1}{3}}\left(\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}\|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{-1}\Delta\|^3\right)^{\frac{2}{3}}.\tag{5.9}$$

By Proposition 5.7, we have that

$$\left(\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}(\|\mathfrak{f}\|_\infty^2+\|\mathfrak{g}\|_\infty^2)^6\right)^{\frac{1}{3}}=\mathcal{O}(n\ln(e\mathbf{D})).$$

To apply the proposition we expanded the binomial and used the fact that $\mathfrak{f}$ and $\mathfrak{g}$ are independent.

Because $(\mathfrak{q}_t/\sqrt{t^2+(1-t)^2},\mathfrak{z}_t)$ is a random standard pair, we have that

$$\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}\|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{-1}\Delta\|^3=\left(t^2+(1-t)^2\right)^{\frac{3}{2}}\mathop{\mathbb{E}}_{(\mathfrak{h},\mathfrak{y})\sim\rho_{\mathsf{std}}}\|\mathrm{D}_{\mathfrak{y}}\mathfrak{h}^{-1}\Delta\|^3.\tag{5.10}$$

Now, by the Beltrán-Pardo trick [13, Algorithm 17.6], we have that the matrix

$$\Delta^{-1/2}\overline{\mathrm{D}}_{\mathfrak{y}}\mathfrak{h}\in\mathbb{C}^{n\times(n+1)}$$

is a random complex Gaussian matrix. Moreover, $\|\mathrm{D}_{\mathfrak{y}}\mathfrak{h}^{-1}\Delta^{\frac{1}{2}}\|=\|\overline{\mathrm{D}}_{\mathfrak{y}}\mathfrak{h}^{\dagger}\Delta^{\frac{1}{2}}\|$, since $\mathfrak{y}$ is a zero of $\mathfrak{h}$ and $\mathrm{D}_{\mathfrak{y}}\mathfrak{h}$ is just $\overline{\mathrm{D}}_{\mathfrak{y}}\mathfrak{h}$ restricted to the orthogonal complement of $\mathfrak{y}$, which we can view as $\mathrm{T}_{\mathfrak{y}}\mathbb{P}^n$. Because of this, by Proposition 5.8,

$$\mathop{\mathbb{E}}_{(\mathfrak{h},\mathfrak{y})\sim\rho_{\mathsf{std}}}\|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{1-}\Delta\|^3\leq\mathbf{D}^{\frac{3}{2}}\mathop{\mathbb{E}}_{(\mathfrak{h},\mathfrak{y})\sim\rho_{\mathsf{std}}}\left\|\left(\Delta^{-\frac{1}{2}}\overline{\mathrm{D}}_{\mathfrak{z}_t}\mathfrak{q}_t\right)^{\dagger}\right\|^3\leq\frac{1}{2}\mathbf{D}^{\frac{3}{2}}n^{\frac{3}{2}}.$$

Hence, integrating (5.10),

$$\left(\int_0^1 \mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})} \|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{-1}\Delta\|^3\,\mathrm{d}t\right)^{\frac{2}{3}} = \mathcal{O}(n\,\mathbf{D}). \tag{5.11}$$

Putting together (5.8), (5.9) and (5.11) the desired result follows. □

5.4 Systems of quadratic equations

Theorem 5.5 is an improvement over the average number of iterations of $\textsc{Solve}_W$ — which is $\mathcal{O}(nDN)$. Furthermore, in the case of quadratic systems, we can compute each iteration with low cost, ensuring that the average total complexity keeps smaller than the one for $\textsc{Solve}_W$ —which is $\mathcal{O}(n^7)$. The major task, unsurprisingly, is to compute $\|q\|_\infty^{\mathbb{C}}$ in (5.1). But we can use that, for a quadratic polynomial q_i, we can write $q_i(X)$ as X^TA_iX with A_i complex symmetric and that $\|q_i\|_\infty = \|A_i\|$. We can then compute, for a quadratic system $q \in \mathcal{H}_{\mathbf{2}}[n]$ the norm $\|q\|_\infty = \max\|q_i\|_\infty$. A naive approach to compute each $\|q_i\|_\infty$ leads to an $\mathcal{O}(n^4)$ cost for the computation of $\|q\|_\infty$ as it uses $\mathcal{O}(n^3)$ operations to compute each $\|q_i\|_\infty$. Proposition 5.10 below shows we can do better. All in all, we obtain the following result.

Theorem 5.9 (Solving Systems of Quadratic Equations). *Algorithm* $\textsc{Solve}_\infty$ *finds a common complex zero of a system of quadratic equations* $f \in \mathcal{H}_{\mathbf{2}}[n]$ *within* $\mathcal{O}(n^{4.5+\omega})$ *time on average, where* $\omega < 2.375$ *is the exponent of matrix multiplication.*

Proposition 5.10. *Let* $q \in \mathcal{H}_{\mathbf{2}}[n]$ *be a quadratic system such that for each* i, $q_i = X^TA_iX$. *Then*

$$\|q\|_\infty^{\mathbb{C}} \le \sqrt{\left\|\sum_{i=1}^n A_i^*A_i\right\|} \le \sqrt{n}\|q\|_\infty^{\mathbb{C}}$$

where the norm $\|\ \|$ *in the middle formula is the usual operator norm. Moreover, the number* $\sqrt{\|\sum_{i=1}^n A_i^*A_i\|}$ *can be computed with* $\mathcal{O}(n^{1+\omega})$ *operations, where* ω *is the exponent of matrix multiplication.*

Proof of Theorem 5.9. By Proposition 5.10, we can estimate the step-length of our homotopy

$$\frac{0.015\,\|q\|_\infty^{\mathbb{C}}}{\|f-g\|_\infty^{\mathbb{C}}\mathsf{M}^2(q,z)} = \frac{0.06}{\|f-g\|_\infty^{\mathbb{C}}\mathbf{D}\|q\|_\infty^{\mathbb{C}}\|\mathrm{D}_zq^{-1}\|^2}$$

by the smaller

$$\frac{0.06}{\|f-g\|_\infty^{\mathbb{C}}\sqrt{\|\sum_{i=1}^n A_i^*A_i\|}\,\|\mathrm{D}_zq^{-1}\|^2}$$

where $q = (X^TA_iX)_i$. In doing so, the algorithm still terminates, but gets an extra factor of $\sqrt{n}$.

Now, $\|f-g\|_\infty$ can be computed in $\mathcal{O}(n^4)$ operations at the beginning of the algorithm a single time, so we don't need to compute it in each iteration. By Proposition 5.10, we can compute $\sqrt{\|\sum_{i=1}^n A_i^*A_i\|}$ in $\mathcal{O}(n^{1+\omega})$ operations, and, by [13, Proposition 16.32], the remaining arithmetic operations can be done in $\mathcal{O}(n^3)$ operations. Combining this, with the bound of Theorem 5.5 and adding the extra factor $\sqrt{n}$ gives the desired estimate. □

Proof of Proposition 5.10. By the so-called Autonne–Takagi factorization [35, Problem 33], we have that

$$A_i = U_i^T D_i U_i$$

for some real diagonal matrix D_i with non-negative entries and some unitary matrix U_i. Now, it is easy to check that

$$\|q_i\|_\infty^{\mathbb{C}} = \|D_i\| = \sqrt{\|D_i^* D_i\|} = \sqrt{\|A_i^* A_i\|} \leq \sqrt{\left\|\sum_{i=1}^n A_i^* A_i\right\|},$$

where the last inequality follows from the fact that the operator norm is non-decreasing with respect to the order of psd matrices. So $\|q\|_\infty^{\mathbb{C}} \leq \sqrt{\|\sum_{i=1}^n A_i^* A_i\|}$, as we wanted to show.

For the other inequality, observe that

$$\sqrt{\left\|\sum_{i=1}^n A_i^* A_i\right\|} \leq \sqrt{\sum_{i=1}^n \|A_i^* A_i\|} = \sqrt{\sum_{i=1}^n (\|q_i\|_\infty^{\mathbb{C}})^2} \leq \sqrt{n}\|q\|_\infty^{\mathbb{C}},$$

where the equality follows from reversing the equalities in the previously displayed formula. This finish the proof of the inequalities.

For the computation, note that computing $A_i^* A_i$ takes $\mathcal{O}(n^\omega)$ operations, so computing all the $A_i^* A_i$ requires $\mathcal{O}(n^{1+\omega})$ operations. Then adding the $A_i^* A_i$ requires $\mathcal{O}(n^3)$ operations and computing $\|\sum_{i=1}^n A_i^* A_i\|$ another $\mathcal{O}(n^3)$ operations. We thus get $\mathcal{O}(n^{1+\omega})$ operations in total, as we wanted to show. the multiplications $A_i^* A_i$. □

Remark 5.11. Observe that if we are allowed to perform the multiplication of the $A_i^* A_i$ in parallel, which requires only $\mathcal{O}(n)$ processors, we can bring the complexity of our homotopy down to $\mathcal{O}(n^{6.5})$ unconditionally on the value of ω.

6 Extension to spaces of C^1-maps

We now prove some condition number theorems for the space of C^1-functions over $\mathbb{S}^n$, $C^1[q] := C^1(\mathbb{S}^n, \mathbb{R}^q)$. Note that $C^1[q]$ is not complete with respect $\| \|_\infty$. Consider instead, for $f \in C^1[q]$,

$$\|f\|_{\overline{\infty}} := \max_{x\in\mathbb{S}^n} \sqrt{\|f(x)\|_2^2 + \|\mathrm{D}_x f\|_{2,2}^2} = \max_{\substack{x\in\mathbb{S}^n \\ v\in \mathrm{T}_x\mathbb{S}^n}} \sqrt{\|f(x)\|^2 + \frac{\|\mathrm{D}_x f v\|_2^2}{\|v\|_2^2}}.$$

This is a variant of the C^1-norm and so one can show that $C^1[q]$ is complete with respect to $\| \|_{\overline{\infty}}$. Let's see how this norm looks like on an easy kind of C^1-maps.

Example 6.1 (Linear functions). Let $A \in q \times (n+1)$ be a linear matrix and consider the map $\mathcal{A} \in C^1[q]$ given by $x \mapsto Ax$. We can show that

$$\|\mathcal{A}\|_{\overline{\infty}} = \sqrt{\sigma_1(A)^2 + \sigma_2(A)^2}$$

where σ_1 and σ_2 are, respectively the first and second singular values. Recall that σ_1 is also the operator norm.

To see the above equality, note that

$$\|\mathcal{A}\|_{\infty} = \max_{\substack{v,w\in\mathbb{S}^n\\ v\perp w}} \sqrt{\|Av\|_2^2+\|Aw\|_2^2}.$$

Since $\begin{pmatrix}Av & Aw\end{pmatrix}$ has rank at most 2,

$$\sqrt{\|Av\|_2^2+\|Aw\|_2^2} = \left\|\begin{pmatrix}Av & Aw\end{pmatrix}\right\|_F = \sqrt{\sigma_1\left(\begin{pmatrix}Av & Aw\end{pmatrix}\right)^2+\sigma_2\left(\begin{pmatrix}Av & Aw\end{pmatrix}\right)^2};$$

and, since $\begin{pmatrix}Av & Aw\end{pmatrix}$ is an orthogonal projection, by the Interlacing Theorem for Singular Values (c.f. [35, 3.1.3],

$$\sigma_1\left(\begin{pmatrix}Av & Aw\end{pmatrix}\right) \leq \sigma_1(A) \text{ and } \sigma_2\left(\begin{pmatrix}Av & Aw\end{pmatrix}\right) \leq \sigma_2(A).$$

Hence $\|\mathcal{A}\|_\infty \leq \sqrt{\sigma_1(A)^2+\sigma_2(A)^2}$. And we actually have equality as we can take v and w to be, respectively, the 1st and 2nd (right) singular vectors of A.

6.1 Condition Number Theorems for $C^1[q]$

Given $x \in \mathbb{S}^n$, we can consider the set of C^1-maps whose zero set in $\mathbb{S}^n$ have a singularity at x,

$$\Sigma_x^1[q] := \left\{g \in C^1[q] \mid g(x)=0,\, \mathrm{rank}\mathrm{D}_x g < q\right\}.$$

Similarly, we can consider the set of C^1-maps having a singular zero,

$$\Sigma^1[q] := \bigcup_{x\in\mathbb{S}^n} \Sigma_x^1[q].$$

The following result shows a way to compute the distance of a C^1-map to these sets.

Theorem 6.2 (Condition Number Theorem). *Let $f \in C^1[q]$ and $x \in \mathbb{S}^n$, then*

$$\mathrm{dist}_{\infty}(f,\Sigma_x^1[q]) = \sqrt{\|f(x)\|^2+\sigma_q(\mathrm{D}_x f)^2}$$

and

$$\mathrm{dist}_{\infty}(f,\Sigma^1[q]) = \min_{x\in\mathbb{S}^n} \sqrt{\|f(x)\|^2+\sigma_q(\mathrm{D}_x f)^2}$$

where dist_∞ is the distance induced by $\|\ \|_\infty$ and σ_q is the qth singular value.

We call this result "Condition Number Theorem" as it is so for the following two condition numbers for C^1-maps:

$$\mathsf{K}_{\infty}(f,x) := \frac{\|f\|_{\infty}}{\sqrt{\|f(x)\|^2+\sigma_q\left(\mathrm{D}_x f\right)^2}}$$

and

$$\mathsf{K}_{\infty}(f) := \sup_{x\in\mathbb{S}^n} \mathsf{K}_{\infty}(f,x).$$

These condition numbers are very similar to K and one might try (but we won't here) to prove an analogous of Theorem 3.2 for them when restricted to polynomial maps. For C^1-maps, instead, such a theorem would require dealing with multiple technical problems.

For $\mathsf{K}_{\overline{\infty}}(f)$, one has the following,

$$\mathsf{K}_{\overline{\infty}}(f) := \frac{\max\left\{\sqrt{\|f(x)\|_2^2 + \|a^*\mathrm{D}_x f\|_2^2} \mid x \in \mathbb{S}^n,\, a \in \mathbb{S}^{q-1}\right\}}{\min\left\{\sqrt{\|f(x)\|_2^2 + \|a^*\mathrm{D}_x f\|_2^2} \mid x \in \mathbb{S}^n,\, a \in \mathbb{S}^{q-1}\right\}}.$$

This formula shows that $\mathsf{K}_{\overline{\infty}}(f)$ is similar to the condition number associated to an operator norm of a linear map.

Proof of Theorem 6.2. Using the triangular inequality and that σ_q is Lipschitz with respect to the operator norm, we can see that, for $f, g \in C^1[q]$,

$$\left|\sqrt{\|f(x)\|^2 + \sigma_q(\mathrm{D}_x f)^2} - \sqrt{\|g(x)\|^2 + \sigma_q(\mathrm{D}_x g)^2}\right| \leq \|f - g\|_{\overline{\infty}}.$$

From here, we deduce that

$$\sqrt{\|f(x)\|^2 + \sigma_q(\mathrm{D}_x f)^2} \leq \mathrm{dist}_{\overline{\infty}}(f, \Sigma_x^1[q])$$

by taking $g \in \Sigma_x^1[q]$ and minimizing over the right-hand side. For the reversed inequality, let

$$\mathrm{D}_x f = U \begin{pmatrix} s_1 & & & \\ & \ddots & & \mathbf{0} \\ & & s_q & \end{pmatrix} V$$

be the SVD of $D_x f$, where U and V are orthogonal and $\mathbf{0}$ is the zero matrix.

Since orthogonal transformations leave invariant $\| \; \|_{\overline{\infty}}$, we can assume, without loss of generality, that $x = e_0$ and that V is the identity matrix. Consider now

$$g_i := f_i - f_i(e_0)X_0 - u_{i,q}s_q X_q.$$

We have then that $g \in \Sigma_{e_0}^1[q]$, since $g(e_0) = 0$ and $\sigma_q(\mathrm{D}_{e_0} g) = 0$, and that

$$f - g = f(e_0)X_0 + s_q u_q X_q.$$

By arguing as in Example 2.4 and noting that $f(e_0)X_0 + s_q u_q X_q$ has rank at most 2, we have that

$$\begin{aligned}\|f(e_0)X_0 + s_q u_q X_q\|_{\overline{\infty}} &= \left\|\begin{pmatrix} f(e_0) & s_q u_q \end{pmatrix}\right\|_F \\ &= \sqrt{\|f(e_0)\|_2^2 + \|s_q u_q\|_2^2} = \sqrt{\|f(e_0)\|^2 + \sigma_q(\mathrm{D}_{e_0} f)^2}/\end{aligned}$$

Hence

$$\mathrm{dist}_{\overline{\infty}}(f, \Sigma_{e_0}^1[q]) \geq \|f - q\|_{\overline{\infty}} = \sqrt{\|f(e_0)\|^2 + \sigma_q(\mathrm{D}_{e_0} f)^2}$$

finishing the proof of the first equality.

The second equality follows immediate from the first one. □

6.2 Structured Condition Number Theorem for $C^1[q]$

Recall that, for $\boldsymbol{d} \in \mathbb{N}^q$, Δ is the diagonal $q \times q$ matrix whose diagonal is $\boldsymbol{d}$. We consider the following variant of $\| \ \|_{\overline{\infty}}$,

$$\|f\|_{\overline{\infty},\boldsymbol{d}} := \max_{x\in\mathbb{S}^n} \sqrt{\|f(x)\|_2^2 + \|\Delta^{-\frac{1}{2}}\mathrm{D}_x f\|_{2,2}^2} = \max_{\substack{x\in\mathbb{S}^n \\ v\in\mathrm{T}_x\mathbb{S}^n}} \sqrt{\|f(x)\|^2 + \frac{\|\Delta^{-\frac{1}{2}}\mathrm{D}_x f v\|_2^2}{\|v\|_2^2}}$$

for $f \in C^1[q]$. The following example shows a class of functions for which this norm can be computed exactly.

Example 6.3. Let

$$M_{a,b} := \left(aX_0^{d_i} + \Delta^{\frac{1}{2}} bX_0 d_i - 1X_1\right)_i \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q].$$

Then, we can see that

$$\|M_{a,b}\|_{\overline{\infty},\boldsymbol{d}} = \|M_{a,b}\|_W = \sqrt{\|a\|^2 + \|b\|^2}.$$

Indeed, by Proposition 2.2, we have that for all $x \in \mathbb{S}^n$,

$$\sqrt{\|M_{a,b}(x)\|_2^2 + \left\|\Delta^{-\frac{1}{2}}\mathrm{D}_x M_{a,b}\right\|_2^2} \leq \|M_{a,b}\|_W.$$

Thus $\|M_{a,b}\|_{\overline{\infty},\boldsymbol{d}} \leq \|M_{a,b}\|_W$, where we have equality for $x = e_0$.

We can also associate to $\| \ \|_{\overline{\infty},\boldsymbol{d}}$, for $f \in C^1[q]$ and $x \in \mathbb{S}^n$, the quantities

$$\mathsf{K}_{\overline{\infty},\boldsymbol{d}}(f,x) := \frac{\|f\|_{\overline{\infty}}}{\sqrt{\|f(x)\|^2 + \sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right)^2}}$$

and

$$\mathsf{K}_{\overline{\infty},\boldsymbol{d}}(f) := \sup_{x\in\mathbb{S}^n} \mathsf{K}_{\overline{\infty},\boldsymbol{d}}(f,x).$$

For these variants of $\mathsf{K}_{\overline{\infty}}$, we have the following structured condition number theorem for perturbations by homogeneous polynomials.

Theorem 6.4 (Structured Condition Number Theorem). *Let $f \in C^1[q]$, $x \in \mathbb{S}^n$ and $\boldsymbol{d} \in \mathbb{N}^q$, then*

$$\mathrm{dist}_{\overline{\infty},\boldsymbol{d}}\left(f, \Sigma_x^1[q] \cap (f + \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q])\right) = \sqrt{\|f(x)\|^2 + \sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right)^2}$$

and

$$\mathrm{dist}_{\overline{\infty},\boldsymbol{d}}\left(f, \Sigma^1[q] \cap (f + \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q])\right) = \min_{x\in\mathbb{S}^n} \sqrt{\|f(x)\|^2 + \sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right)^2}$$

where $\mathrm{dist}_{\overline{\infty},\boldsymbol{d}}$ is the distance induced by $\| \ \|_{\overline{\infty},\boldsymbol{d}}$ and σ_q is the qth singular value.

Corollary 6.5. *Let for $\boldsymbol{d} \in \mathbb{N}^d$, $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$, and $x \in \mathbb{S}^n$. Then*

$$\mathrm{dist}_{\overline{\infty},\boldsymbol{d}}(f, \Sigma_{\boldsymbol{d},x}[q]) = \sqrt{\|f(x)\|^2 + \sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right)^2} = \mathrm{dist}_W(f, \Sigma_{\boldsymbol{d},x}[q])$$

and

$$\mathrm{dist}_{\overline{\infty},\boldsymbol{d}}(f, \Sigma_{\boldsymbol{d}}[q]) = \min_{x\in\mathbb{S}^n} \sqrt{\|f(x)\|^2 + \sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right)^2} = \mathrm{dist}_W(f, \Sigma_{\boldsymbol{d}}[q])$$

where $\mathrm{dist}_{\overline{\infty},\boldsymbol{d}}$ is the distance induced by $\| \ \|_{\overline{\infty},\boldsymbol{d}}$ and σ_q is the qth singular value.

Note that the adjective 'structured' refers to the fact that we only allow perturbations of f by C^1-maps in $\mathcal{H}^{\mathbb{R}}_{\boldsymbol{d}}[q]$. However, we might still be interested in general perturbations. If this is the case, we can get them using the relationship between $\| \ \|_{\infty,\boldsymbol{d}}$ and $\| \ \|_{\overline{\infty}}$. We will explore this more detail in the next subsection.

Proof of Theorem 6.4. This proof is almost the same than the one of Theorem 6.2. We only have to modify the part were we find an explicit minimizer for the distance. Again, we write

$$\Delta^{-\frac{1}{2}}\mathrm{D}_x f = U \begin{pmatrix} s_1 & & & \\ & \ddots & & \mathbf{0} \\ & & s_q & \end{pmatrix} V$$

where $s_1, \ldots, s_q > 0$, U and V are orthogonal and $\mathbf{0}$ is the zero matrix. Again, without loss of generality, we assume that $x = e_0$ and that V is the identity. We consider

$$g_i := f_i - x_0^{d-1}(f_i(e_0)X_0 - \sqrt{d_i}u_{i,q}s_qX_q)$$

so that $g \in \Sigma^1_{e_0}[q]$, as $g(e_0) = 0$ and $\sigma_q(\mathrm{D}_{e_0}g) = 0$, and

$$f - g = \left(f_i(e_0)X_0^{d_i} + \sqrt{d_i}s_qu_qX_q\right)_i.$$

Because of Example 6.3, for

$$h = \left(a_iX_0^{d_i} + \sqrt{d_i}bX_0^{d_i-1}X_1\right)_i \in \mathcal{H}^{\mathbb{R}}_{\boldsymbol{d}}[q],$$

we have that $\|h\|_{\overline{\infty},\boldsymbol{d}} = \sqrt{\|a\|_2^2 + \|b\|_2^2}$. Hence,

$$\mathrm{dist}_{\overline{\infty},\boldsymbol{d}}(f, \Sigma^1_{e_0}[q]) \geq \|f - g\|_{\overline{\infty}} = \sqrt{\|f(e_0)\|^2 + \sigma_q(\Delta^{-\frac{1}{2}}\mathrm{D}_{e_0}f)^2}$$

and the first equality follows. The second equality immediately follows from the first one. □

Proof of Corollary 6.5. This is Theorem 6.5 together with [14, Theorem 4.4]. □

6.3 Relationship between norms

As it happens with K and κ (see Subsection 4.3), the relations between the condition numbers K, κ, $\mathsf{K}_{\overline{\infty}}$ and $\mathsf{K}_{\overline{\infty},\boldsymbol{d}}$ reduces to the relations between the corresponding norms.

We therefore prove the following propositions relating these norms. Note that for $C^1[q]$, we compare $\| \ \|_{\overline{\infty}}$ with $\| \ \|_{\overline{\infty},\boldsymbol{d}}$, and for $\mathcal{H}^{\mathbb{R}}_{\boldsymbol{d}}[q]$, we compare $\| \ \|^{\mathbb{R}}_{\infty}$, $\| \ \|_W$, $\| \ \|_{\overline{\infty}}$ and $\| \ \|_{\overline{\infty},\boldsymbol{d}}$.

Proposition 6.6. *Let $f \in C^1[q]$. Then for all $\boldsymbol{d}, \widetilde{\boldsymbol{d}} \in \mathbb{N}^q$,*

$$\frac{1}{\max_i \sqrt{d_i}}\|f\|_{\overline{\infty}} \leq \|f\|_{\overline{\infty},\boldsymbol{d}} \leq \|f\|_{\overline{\infty}}$$

and

$$\min\left\{1, \min_i \sqrt{\frac{\tilde{d}_i}{d_i}}\right\}\|f\|_{\overline{\infty},\widetilde{\boldsymbol{d}}} \leq \|f\|_{\overline{\infty},\boldsymbol{d}} \leq \max\left\{1, \max_i \sqrt{\frac{\tilde{d}_i}{d_i}}\right\}\|f\|_{\overline{\infty},\widetilde{\boldsymbol{d}}}.$$

Proposition 6.7. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$. Then the following inequalities hold:*

$$\frac{1}{\sqrt{2q}\,\mathbf{D}}\|f\|_{\overline{\infty}} \leq \|f\|_{\infty}^{\mathbb{R}} \leq \|f\|_{\overline{\infty},\boldsymbol{d}} \leq \|f\|_{\overline{\infty}} \tag{6.1}$$

$$\frac{1}{\sqrt{2q}\mathbf{D}}\|f\|_{\overline{\infty},\boldsymbol{d}} \leq \|f\|_{\infty}^{\mathbb{R}} \leq \|f\|_{\overline{\infty},\boldsymbol{d}} \tag{6.2}$$

$$\|f\|_{\infty}^{\mathbb{R}} \leq \|f\|_{\overline{\infty},\boldsymbol{d}} \leq \|f\|_{W} \tag{6.3}$$

Proof of Proposition 6.6. It is enough to show that

$$\|f\|_{\overline{\infty},\boldsymbol{d}} \leq \max\left\{1, \max_i \sqrt{\frac{\tilde{d}_i}{d_i}}\right\} \|f\|_{\overline{\infty},\tilde{\boldsymbol{d}}},$$

since the rest of inequalities are derived from this claim in an straightforward way. For the latter, note that $\|\ \|_{\overline{\infty}} = \|\ \|_{\overline{\infty},\mathbb{1}}$ where $\mathbb{1} = (1, \ldots, 1)$.

Now, one can easily check that for $A \in \mathbb{R}^{q\times n}$,

$$\left\|\Delta^{-\frac{1}{2}}A\right\|_{2,2} = \left\|\Delta^{-\frac{1}{2}}\tilde{\Delta}^{\frac{1}{2}}\,\tilde{\Delta}^{-\frac{1}{2}}A\right\|_{2,2} \leq \left\|\Delta^{-\frac{1}{2}}\tilde{\Delta}^{\frac{1}{2}}\right\|_{2,2}\left\|\tilde{\Delta}^{-\frac{1}{2}}A\right\|_{2,2} = \max_i\sqrt{\frac{\tilde{d}_i}{d_i}}\left\|\tilde{\Delta}^{-\frac{1}{2}}\right\|_{2,2},$$

and that for $a, b, t \in \mathbb{R}^2$ that

$$\sqrt{a^2 + (tb)^2} \leq \max\{1, |t|\}\sqrt{a^2 + b^2}.$$

Combining these bounds together, we get

$$\sqrt{\|f(x)\|^2 + \left\|\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right\|_{2,2}^2} \leq \max\left\{1, \max_i \sqrt{\frac{\tilde{d}_i}{d_i}}\right\}\sqrt{\|f(x)\|^2 + \left\|\tilde{\Delta}^{-\frac{1}{2}}\mathrm{D}_x f\right\|_{2,2}^2}$$

and so the desired claim. □

Proof of Proposition 6.7. Arguing as in Proposition 6.6, we can prove that, for all $x \in \mathbb{S}^n$,

$$\frac{1}{\sqrt{2q}\,\mathbf{D}}\sqrt{\|f(x)\|^2 + \|\mathrm{D}_x f\|_{2,2}^2} \leq \max\left\{\|f(x)\|_\infty, \left\|\tilde{\Delta}^{-1}\mathrm{D}_x f\right\|_{\infty,2}\right\} \leq \sqrt{\|f(x)\|^2 + \|\mathrm{D}_x f\|_{2,2}^2}$$

and

$$\begin{aligned}\frac{1}{\sqrt{2q}\mathbf{D}}\sqrt{\|f(x)\|^2 + \left\|\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right\|_{2,2}^2} &\leq \max\left\{\|f(x)\|_\infty, \left\|\tilde{\Delta}^{-1}\mathrm{D}_x f\right\|_{\infty,2}\right\}\\ &\leq \sqrt{\|f(x)\|^2 + \left\|\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right\|_{2,2}^2}.\end{aligned}$$

Maximizing over $z \in \mathbb{S}^n$ gives the inequalities in (6.1) and (6.2).

It only remains to prove $\|f\|_{\infty,\boldsymbol{d}} \leq \|f\|_W$ in (6.3). To do this, note that by Proposition 2.2, for all $x \in \mathbb{S}^n$,

$$\sqrt{\|f(x)\|^2 + \left\|\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right\|_{2,2}^2} \leq \|f\|_W.$$

The result follows from maximizing over $x \in \mathbb{S}^n$. □

We finish with the following theorem, similar in flavour to [29, Proposition 3] and [11, Theorem 7], where it was shown that the distance of a polynomial tuple to polynomial tuples with singularities bounds the distance of this polynomial to C^1-functions with singularities.

Theorem 6.8. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $x \in \mathbb{S}^n$. Then*

$$\frac{1}{\sqrt{\mathbf{D}}}\operatorname{dist}_{\overline{\infty}}(f, \Sigma_x^1[q]) \leq \operatorname{dist}_{\overline{\infty},\boldsymbol{d}}(f, \Sigma_{\boldsymbol{d},x}[q]) = \operatorname{dist}_W(f, \Sigma_{\boldsymbol{d},x}[q]) \leq \operatorname{dist}_{\overline{\infty}}(f, \Sigma_x^1[q]),$$

and

$$\frac{1}{\sqrt{\mathbf{D}}}\operatorname{dist}_{\overline{\infty}}(f, \Sigma^1[q]) \leq \operatorname{dist}_{\overline{\infty},\boldsymbol{d}}(f, \Sigma_{\boldsymbol{d}}[q]) = \operatorname{dist}_W(f, \Sigma_{\boldsymbol{d}}[q]) \leq \operatorname{dist}_{\overline{\infty}}(f, \Sigma^1[q]),$$

where $\operatorname{dist}_{\overline{\infty}}$ and $\operatorname{dist}_{\overline{\infty},\boldsymbol{d}}$ are, respectively, the distances induced by $\| \ \|_{\overline{\infty}}$ and $\| \ \|_{\overline{\infty},\boldsymbol{d}}$.

Sketch of proof. The proof is similar to that of Proposition 6.6. Arguing as there, we can prove that for all $x \in \mathbb{S}^n$,

$$\frac{1}{\sqrt{\mathbf{D}}}\sqrt{\|f(x)\|_2^2 + \sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right)^2} \leq \sqrt{\|f(x)\|_2^2 + \sigma_q\left(\mathrm{D}_x f\right)^2} \leq \sqrt{\|f(x)\|_2^2 + \sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right)^2}.$$

Minimizing over $x \in \mathbb{S}^N$ and applying Theorems 6.2 and Corollary 6.5, we conclude. □

References

[1] A. Agrachev, K. Kozhasov, and A. Uschmajew. Chebyshev polynomials and best rank-one approximation ratio. *SIAM J. Matrix Anal. Appl.*, 41(1):308–331, 2020.

[2] N. Alon, T. Lee, A. Shraibman, and S. Vempala. The approximate rank of a matrix and its algorithmic applications. In *STOC'13—Proceedings of the 2013 ACM Symposium on Theory of Computing*, pages 675–684. ACM, New York, 2013.

[3] D. Armentano, C. Beltrán, P. Bürgisser, F. Cucker, and M. Shub. Condition length and complexity for the solution of polynomial systems. *Found. Comput. Math.*, 16:1401–1422, 2016.

[4] J.-M. Azaïs and M. Wschebor. On the roots of a random system of equations. The theorem on Shub and Smale and some extensions. *Found. Comput. Math.*, 5(2):125–144, 2005.

[5] Alexander Barvinok. Integration and optimization of multivariate polynomials by restriction onto a random subspace. *Foundations of Computational Mathematics*, 7(2):229–244, 2007.

[6] C. Beltrán and L.M. Pardo. On Smale's 17 problem: a probabilistic positive solution. *Found. Comput. Math.*, 8:1–43, 2008.

[7] C. Beltrán and L.M. Pardo. Smale's 17th problem: average polynomial time to compute affine and projective solutions. *J. Amer. Math. Soc.*, 22(2):363–385, 2009.

[8] C. Beltrán and L.M. Pardo. Fast linear homotopy to find approximate zeros of polynomial systems. *Found. Comput. Math.*, 11(1):95–129, 2011.

[9] V. Bhattiprolu, V. Guruswami, and E. Lee. Sum-of-squares certificates for maxima of random tensors on the sphere. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[10] L. P. Bos, N. Levenberg, P. D. Milman, and B. A. Taylor. Tangential Markov inequalities on real algebraic varieties. *Indiana Univ. Math. J.*, 47(4):1257–1272, 1998.

[11] P. Breiding, H. Keneshlou, and A. Lerario. Quantitative Singularity Theory for Random Polynomials. *International Mathematics Research Notices*, 10 2020.

[12] P. Bürgisser and F. Cucker. On a problem posed by Steve Smale. *Annals of Mathematics*, 174:1785–1836, 2011.

[13] P. Bürgisser and F. Cucker. *Condition*, volume 349 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2013.

[14] P. Bürgisser, F. Cucker, and P. Lairez. Computing the homology of basic semialgebraic sets in weak exponential time. *J. ACM*, 66(1):5:1–5:30, 2018.

[15] P. Bürgisser, F. Cucker, and J. Tonelli-Cueto. Computing the homology of semialgebraic sets. i: Lax formulas. *Foundations of Computational Mathematics*, May 2019. On-line from May of 2019. arXiv:1807.06435.

[16] P. Bürgisser, F. Cucker, and J. Tonelli-Cueto. Computing the Homology of Semialgebraic Sets. II: General formulas. *Foundations of Computational Mathematics*, 1 2021. On-line from January of 2021. arXiv: 1903.10710.

[17] M. Burr, F. Krahmer, and C. Yap. Continuous amortization: A non-probabilistic adaptive analysis technique. *Electronic Colloquium on Computational Complexity*, Report. No. 136, 2009.

[18] M. A. Burr. Continuous amortization and extensions: with applications to bisection-based root isolation. *J. Symbolic Comput.*, 77:78–126, 2016.

[19] M. A. Burr, S. Gao, and E. P. Tsigaridas. 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*, pages 61–68. ACM, New York, 2017.

[20] Venkat Chandrasekaran and Parikshit Shah. Relative entropy relaxations for signomial optimization. *SIAM Journal on Optimization*, 26(2):1147–1173, 2016.

[21] F. Cucker. A theory of complexity, condition, and roundoff. *Forum Math. Sigma*, 3:Paper No. e4, 50, 2015.

[22] F. Cucker, A. A. Ergür, and J. Tonelli-Cueto. Plantinga-Vegter Algorithm Takes Average Polynomial Time. In *Proceedings of the 2019 on International Symposium on Symbolic and Algebraic Computation*, pages 114–121. ACM, 7 2019.

[23] F. Cucker, A. A. Ergür, and J. Tonelli-Cueto. On the Complexity of the Plantinga-Vegter Algorithm, 4 2020. Journal version of [22]. arXiv:2004.06879.

[24] F. Cucker, T. Krick, G. Malajovich, and M. Wschebor. A numerical algorithm for zero counting. I: Complexity and accuracy. *J. Complexity*, 24:582–605, 2008.

[25] F. Cucker, T. Krick, G. Malajovich, and M. Wschebor. A numerical algorithm for zero counting. II: Distance to ill-posedness and smoothed analysis. *J. Fixed Point Theory Appl.*, 6:285–294, 2009.

[26] F. Cucker, T. Krick, G. Malajovich, and M. Wschebor. A numerical algorithm for zero counting. III: Randomization and condition. *Adv. Applied Math.*, 48:215–248, 2012.

[27] F. Cucker, T. Krick, and M. Shub. Computing the Homology of Real Projective Sets. *Found. Comput. Math.*, 18:929–970, 2018.

[28] J.-P. Dedieu. *Points fixes, zéros et la méthode de Newton*, volume 54 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Berlin, 2006. With a preface by Steve Smale.

[29] D. N. Diatta and A. Lerario. Low degree approximation of random polynomials, 12 2018. arXiv:1812.10137.

[30] M. Dressler, S. Iliman, and T. De Wolff. A positivstellensatz for sums of nonnegative circuit polynomials. *SIAM Journal on Applied Algebra and Geometry*, 1(1):536–555, 2017.

[31] Alperen A. Ergür, Grigoris Paouris, and J. Maurice Rojas. Probabilistic condition number estimates for real polynomial systems I: A broader family of distributions. *Found. Comput. Math.*, 19:131–157, 2019.

[32] H. Federer. Curvature measures. *Trans. Amer. Math. Soc.*, 93:418–491, 1959.

[33] G. B. Folland. How to integrate a polynomial over a sphere. *Amer. Math. Monthly*, 108(5):446–448, 2001.

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

[35] R. A. Horn and C. R. Johnson. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[36] O. D. Kellogg. On bounded polynomials in several variables. *Math. Z.*, 27(1):55–64, 1928.

[37] P. Lairez. Rigid continuation paths I. Quasilinear average complexity for solving polynomial systems. *J. Amer. Math. Soc.*, 33(2):487–526, 2020.

[38] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging applications of algebraic geometry*, pages 157–270. Springer, 2009.

[39] G. Livshyts, G. Paouris, and P. Pivovarov. On sharp bounds for marginal densities of product measures. *Israel Journal of Mathematics*, 216(2):877–889, 2016.

[40] G. Malajovich and M. Shub. A theory of NP-completeness and ill-conditioning for approximate real computations. *J. ACM*, 66(4):Art. 27, 38, 2019.

[41] S. Mendelson and G. Paouris. Stable recovery and the coordinate small-ball behaviour of random vectors, 2019. arXiv:1904.08532.

[42] P. Niyogi, S. Smale, and S. Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.*, 39(1-3):419–441, 2008.

[43] E. Novak, I.H. Sloan, J.F. Traub, and H. Woźniakowski. *Essays on the complexity of continuous problems*. European Mathematical Society (EMS), Zürich, 2009.

[44] S. Plantinga and G. Vegter. Isotopic approximation of implicit curves and surfaces. In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 245–254. ACM, 2004.

[45] L. Plaskota. *Noisy information and computational complexity*. Cambridge University Press, Cambridge, 1996.

[46] M. Rudelson and R. Vershynin. Small ball probabilities for linear images of high-dimensional distributions. *Int. Math. Res. Not. IMRN*, 19:9594–9617, 2015.

[47] R. T. Seeley. Spherical harmonics. *Amer. Math. Monthly*, 73(4, part II):115–121, 1966.

[48] M. Shub and S. Smale. Complexity of Bézout's Theorem I: geometric aspects. *Journal of the Amer. Math. Soc.*, 6:459–501, 1993.

[49] S. Smale. Mathematical problems for the next century. In V. Arnold, M. Atiyah, P. Lax, and B. Mazur, editors, *Mathematics: Frontiers and Perspectives*, pages 271–294. AMS, 2000.

[50] J. Tonelli-Cueto. *Condition and Homology in Semialgebraic Geometry*. Doctoral thesis, Technische Universität Berlin, 12 2019.

[51] J. Tonelli-Cueto and E. Tsigaridas. Condition numbers for the cube. i: Univariate polynomials and hypersurfaces. In *Proceedings of the 45th International Symposium on Symbolic and Algebraic Computation*, ISSAC '20, page 434–441, New York, NY, USA, 2020. Association for Computing Machinery.

[52] J.F. Traub, G.W. Wasilkowski, and H. Woźniakowski. *Information-based complexity.* Computer Science and Scientific Computing. Academic Press, Inc., Boston, MA, 1988. With contributions by A. G. Werschulz and T. Boult.

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

[54] H. Weyl. *The Theory of Groups and Quantum Mechanics.* Dover, 1932.

[55] X. Zhang, C. Ling, and L. Qi. The best rank-1 approximation of a symmetric tensor and related spherical optimization problems. *SIAM J. Matrix Anal. Appl.*, 33(3):806–821, 2012.