HAL open science

Functional norms, condition numbers and numerical algorithms in algebraic geometry

Felipe Cucker, Alperen A. Ergür, Josué Tonelli-Cueto

To cite this version:

Felipe Cucker, Alperen A. Ergür, Josué Tonelli-Cueto. Functional norms, condition numbers and numerical algorithms in algebraic geometry. Forum of Mathematics, Sigma, 2022, 10, pp.e103. 10.1017/fms.2022.89 . hal-03151436v2

HAL Id: hal-03151436

https://inria.hal.science/hal-03151436v2

Submitted on 22 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Forum of Mathematics, Sigma (2022), **Vol. 10**:e103 1–49
doi:10.1017/fms.2022.89

RESEARCH ARTICLE

Functional norms, condition numbers and numerical algorithms in algebraic geometry

Felipe Cucker [1], Alperen A. Ergür [2] and Josué Tonelli-Cueto [3]

[1]Department of Mathematics, City University of Hong Kong, Hong Kong; E-mail: macucker@cityu.edu.hk.
[2]Department of Mathematics, The University of Texas at San Antonio, San Antonio, TX, United States;
E-mail: alperen.ergur@utsa.edu.
[3]OURAGAN team, Inria Paris & Institut de mathématiques de Jussieu-Paris Rive Gauche, Paris, France;
E-mail: josue.tonelli.cueto@bizkaia.eu.

Received: 24 February 2021; **Revised:** 5 April 2022; **Accepted:** 12 October 2022

2020 Mathematics Subject Classification: *Primary* – 14Q20, 65Y20; *Secondary* – 68Q25, 68U05

Abstract

In numerical linear algebra, a well-established practice is to choose a norm that exploits the structure of the problem at hand to optimise accuracy or computational complexity. In numerical polynomial algebra, a single norm (attributed to Weyl) dominates the 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 use of L_p-norms substantially reduces computational complexity: a subdivision-based algorithm 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).

Contents

1 Introduction **2**
2 Norms for polynomials **4**
2.1 Euclidean norms . . . 4
2.2 Functional norms . . . 5
2.3 Kellogg's theorem . . . 9
3 Condition numbers for the L_∞-norm **11**
3.1 Properties of the real condition number K . . . 12
3.2 Properties of the complex condition number M . . . 16
4 Numerical algorithms in real algebraic geometry **17**
4.1 A grid-based algorithm and its condition-based complexity . . . 18
4.1.1 Computation of $\| \|_\infty^{\mathbb{R}}$ . . . 19
4.1.2 Estimation of K . . . 20
4.1.3 Complexity analysis of grid-based algorithms using K . . . 21
4.2 Complexity of the Plantinga-Vegter algorithm . . . 24
4.2.1 The Plantinga-Vegter subdivision algorithm . . . 24
4.2.2 Complexity of PV-Interval$_\infty$ . . . 25

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is included and the original work is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use.

4.3 Probabilistic analysis of algorithms . . . 26
4.3.1 The randomness model: dobro random polynomials . . . 27
4.3.2 K vs. κ: Measuring the effect of the L_∞-norm on the grid method . . . 29
4.3.3 Complexity of the Plantinga-Vegter algorithm . . . 32
5 Linear homotopy for computing complex zeros **34**
5.1 Description of the linear homotopy . . . 36
5.2 A bound on the number of iterations . . . 36
5.3 Average complexity analysis of Solve_∞ . . . 38
5.4 Systems of quadratic equations . . . 40
A Extension to spaces of C^1-maps **42**
A.1 Condition number theorems for $C^1[q]$ . . . 42
A.2 Structured condition number theorem for $C^1[q]$ . . . 44
A.3 Relationship between norms . . . 46

1. Introduction

In numerical analysis, it matters how we measure errors. Change the metric we measure the perturbations with, and a well-conditioned input may turn badly conditioned (a remarkable example is in [22]). Because of this, a careful choice of how we measure errors is a fundamental step in the design and analysis of algorithms. A main example is numerical linear algebra, where it is commonplace to carefully choose a matrix norm depending on the problem at hand: the goal is to exploit the structure of the problem and optimise computational efficiency.

Unlike numerical linear algebra, a single norm – the Weyl norm – prevails in numerical algebraic geometry. The nice properties of the Weyl norm, ease of computing and unitary invariance, explain this prevalence. Nevertheless, the absence of complexity analyses using other norms in numerical algebraic geometry reflects badly on the theoretical strength of our analyses, which appear to rely on a specific choice of metric.

In this paper, we aim to show that using other norms is possible in numerical algebraic geometry. To do so, we consider an L_∞-norm in the space of polynomial systems and show how this leads to numerical algorithms and a complexity framework analogous to the one we have with the Weyl norm. Furthermore, we show that the change of norms leads to significant improvements in complexity bounds thanks to the better probabilistic behaviour of this L_∞ norm with respect to the Weyl norm. We show this in three relevant cases: 1) computation of the homology of algebraic sets, 2) the Plantinga-Vegter algorithm and 3) the homotopy continuation method for quadratic polynomial systems.

We now discuss in more detail the aspects we have mentioned in passing to put our results in context within the wider setting of complexity theory for numerical algorithms and numerical algebraic geometry.

Complexity paradigm. The behaviour of numerical algorithms varies from input to input. This phenomenon is due not necessarily to the algorithms themselves but rather to the numerical sensitivity – how much the output varies with respect to a perturbation of the input – of the input we are processing. The numerical sensitivity of an input is captured by the so-called *condition number*. Then, in turn, condition numbers allow one to analyse numerical algorithms and explain why numerical algorithms handle some inputs faster than others.

Central to our paper is the fact that the choice of the metric under which we measure perturbations determines the condition number of the data. An example of this is given by the polynomial $X^d - 1$, which is well-conditioned (for the zero finding problem) with respect to the standard norm in equation (2.2) but badly conditioned with respect to the Weyl norm in equation (2.3) [14, Example 14.3].

A drawback of condition-based complexity analyses is that, as we don't know a priori the condition of the input at hand, we cannot foresee the running time for this input. We can nonetheless get an idea of how the algorithm behaves in general by randomising the input. This allows one to obtain probabilistic estimates for the practical performance of the numerical algorithm.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Again, we note that the metric we choose to measure perturbations affects the probabilistic models we consider. This is so because probabilistic parameters such as the variance are always given with respect to some metric, so when we change the metric, we change the values of these parameters.

We refer to [14] for a more detailed overview of this complexity paradigm based on condition numbers. In the rest of the paper, we will show how this complexity framework works for each of the three cases mentioned above.

Choice of the norm. Arguably, one disadvantage of the L_∞-norm is that we don't have an efficient way to approximate $\| \|_\infty$. For polynomials in $n+1$ homogeneous variables whose degrees are bounded by $\mathbf{D}$, our current fastest algorithm takes time polynomial in $\mathbf{D}$ and exponential in n. However, the computation of $\| \|_\infty$ amounts to a polynomial optimisation problem, and efficient algorithms exist for particular classes of polynomials. This is the case, for example, with sums of squares [43, 10], sparse polynomials [31, 21] and other structures [5]. Unrestricted efficient algorithms are not expected to be designed because it is well-known that polynomial optimisation reduces to the feasibility problem over the reals, and the latter is $\mathrm{NP}_{\mathbb{R}}$-complete. Nonetheless, for most applications we only need a coarse approximation of $\| \|_\infty$, which allows for some optimism.

Our choice of the L_∞-norm is due to the inequalities shown in Kellogg's theorem (Theorem 2.13), which we haven't found for other L_p-norms. A way around Kellogg's theorem for general L_p-norms would certainly lead to new results regarding the use of these norms in algorithm analysis.

Despite the high cost of computing the L_∞ norm, its use may yield substantially better cost bounds for some algorithms. This improvement rests on two facts:

1. For a homogeneous polynomial f with $n+1$ variables and degree $\mathbf{D}$, we always have $\|f\|_\infty \leq \|f\|_W$, and for a random homogeneous polynomial $\mathfrak{f}$, we have $\|\mathfrak{f}\|_\infty \lesssim \sqrt{n \log \mathbf{D}}$, whereas $\|\mathfrak{f}\|_W \sim \binom{n+\mathbf{D}}{n}^{\frac{1}{2}}$. An analogous situation holds for polynomial systems (see Theorem 4.28 and Proposition 4.32).
2. Condition numbers with respect to the L_∞-norm yield condition-based complexity estimates (i.e., cost bounds in terms of both n, $\mathbf{D}$ and a condition number) almost identical to those obtained using the condition numbers with respect to the Weyl norm (see Section 3).

In this way, the reduction in the probabilistic estimates in passing to $\| \|_\infty$ from $\| \|_W$ immediately translates to reductions in the magnitude of the corresponding condition numbers and, in turn, reductions in the complexity estimates.

Considered algorithms. We showcase three algorithms where despite the high cost of computing the L_∞-norm, the reductions in the total cost bounds remain significant.

Firstly, in Section 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 Section 4.1, we replace the Weyl norm by $\| \|_\infty$ in the design of one such algorithm (to compute Betti numbers); and in Section 4.3, we show a decrease in its cost bounds. We take advantage of the fact that there is only one norm computation, and it is done, so to speak, along the way. The gain in the reduction of the estimate for the number of iterations directly yields a reduction in the total cost bound (see Corollary 4.31).

Secondly, in Section 4.2, we consider the Plantinga-Vegter algorithm as it is described and analysed in [23]. Again, we replace the Weyl norm by $\| \|_\infty$ in the algorithm's design results in improved cost bounds. And again, the computation of $\| \|_\infty$ is not a burden as it is 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 4.35 exhibits the improvement achieved on average complexity bounds for these two cases. For larger values of n, the improvement is more substantial.

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 [9, 13] and, again, design versions of them where the Weyl norm is replaced by $\| \|_\infty$. Again, this results in a small but measurable reduction in the cost bounds (from n^7 to $n^{6.875}$). A crucial fact in

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

achieving 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 that the use of L_∞-norm yields a clear reduction in the estimates for the expected number of iterations. We believe this is a common pattern. But in general, the reduction in the number of steps does not immediately translate into a reduction in total computational cost. This motivates the search for efficient algorithms that (roughly) approximate $\| \|_\infty$ and for a better understanding of the complexity and accuracy of computing with L_p-norms in polynomial spaces.

Organisation of the paper. In Section 2, we define the norms that will be considered in this paper and work out several examples. We also recall basic properties of these norms and highlight their differences from the Weyl norm. Then, in Section 3, we define condition numbers M and K that scale with the L_∞-norm. These condition numbers are similar to their widely used Weyl versions μ_{norm} and κ (for complex and real problems, respectively). We also prove in Section 3 that the main properties of μ_{norm} and κ – those allowing them to feature in condition-based cost estimates – hold for M and K. Section 4.1, Section 4.2 and Section 5 are the home of three algorithms that are designed using L_∞-scaled condition numbers. We compare the cost bounds of these algorithms to those of their Weyl counterparts and highlight computational gains.

We conclude in Appendix A with a minor digression. Because a natural habitat for functional norms is 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 analyse algorithms here. We nonetheless point out that substantial literature on algorithms on spaces of continuous functions exists [57, 50, 48], where these theorems might be useful.

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 that follows, $\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} \mid \|x\|_2 = 1\}$ 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 is 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 $\|z\|_2 = 1$. This simplifies the form of many of our definitions. This convention can be made without loss of generality 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 a monomial basis. Every $f \in \mathcal{H}_d^{\mathbb{F}}[1]$ can be uniquely represented as

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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

$$\|f\|_{\mathrm{std}} := \sqrt{\sum_{|\alpha|=d} |f_\alpha|^2}. \tag{2.2}$$

For $f = (f_1, \dots, f_q) \in \mathcal{H}_{\boldsymbol{d}}[q]$, 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 equation (2.1), this is given by

$$\|f\|_W := \sqrt{\sum_{|\alpha|=d} \binom{d}{\alpha}^{-1} |f_\alpha|^2}, \tag{2.3}$$

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 $f \in \mathcal{H}_{\boldsymbol{d}}[q]$,

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

Here, and in all that 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 [14, Proposition 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 A), but L_p-norms suffice for obtaining the computational improvements we aim for. Although in the sequel we will only use the L_∞-norm, we present the full family of L_p-norms since we consider that these norms will be useful in the future. Moreover, presenting the full family of L_p-norms allows us to appreciate how the L_∞ differs from and relates to these other norms.

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

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

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

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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

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

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

Remark 2.3. In the case of a single polynomial, the definitions above become simpler. For $f \in \mathcal{H}_{\boldsymbol{d}}[1]$,

$$\|f\|_p^{\mathbb{R}} := \begin{cases} \max_{x \in \mathbb{S}^n} |f(x)| & \text{if } p = \infty \\ \left(\mathop{\mathbb{E}}_{\mathfrak{x} \in \mathbb{S}^n} |f(\mathfrak{x})|^p \right)^{1/p} & \text{otherwise} \end{cases} \quad \text{and} \quad \|f\|_p^{\mathbb{C}} := \begin{cases} \max_{z \in \mathbb{P}^n} |f(z)| & \text{if } p = \infty \\ \left(\mathop{\mathbb{E}}_{\mathfrak{z} \in \mathbb{P}^n} |f(\mathfrak{z})|_p^p \right)^{1/p} & \text{otherwise,} \end{cases}$$

which amount to taking the p-mean of $|f|$ over, respectively, $\mathbb{S}^n$ and $\mathbb{P}^n$.

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 identity is why we take the p-mean of the p-norm of $f(x)$ instead of taking the p-mean of a fixed norm.

(I) We have actions of the qth power of the (real) orthogonal group, $\mathcal{O}(n+1)^q$, on $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$, given by $(A, f) \mapsto (f_i^{A_i}) := (f_i(A_i X))$. Similarly, we have an action of the qth power of the unitary group, $\mathcal{U}(n+1)^q$, 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 of the introduced norms.

Example 2.4 (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 [36]. 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.

[1]Observe, however, that $\| \ \|_p^{\mathbb{R}}$ 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.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Example 2.5 (Linear functions). Let $\mathbb{1} = (1, 1, \ldots, 1) \in \mathbb{N}^q$ 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 $p \in [1, \infty)$,

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

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 2.4). Note that $\Big\| \big(\|A^1\|_2, \ldots, \|A^q\|_2 \big) \Big\|_p$ is just the p-norm of the vector of 2-norms of the rows of A.

Example 2.6 (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 [36], we can conclude that

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

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 2.7 (Cosine polynomials). Let $d \geq 2$, and consider the family of homogeneous 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.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

The following proposition lists simple inequalities between the functional norms. For a converse of some of the inequalities below, where the L_∞ norm is bounded in terms of L_p norms, see [6].

Proposition 2.8. *Let $1 \le p < p' < \infty$ and $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Then for all $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{F}}[q]$, the following inequalities hold:*

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

Sketch of proof. It is a direct consequence of the inequalities between p-means. □

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

Proposition 2.9. *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 $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[1]$,

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

Sketch of proof. We only need to show this in the case $q = 1$. Now both the Weyl norm and the complex L_2-norm are unitarily invariant Hermitian norms of $\mathcal{H}_d^{\mathbb{C}}$. For the Weyl norm, see [14, Theorem 16.3]; for the complex L_2-norm, this is property (I). Since $\mathcal{H}_d^{\mathbb{C}}$ is an irreducible representation of $\mathcal{U}(n+1)$, this means the two norms are equal up to a constant. Using Example 2.4 with $f = X_0^d$, one can check that this constant is $\sqrt{N}$. □

From Proposition 2.2, we get the following result.

Proposition 2.10. *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $f \in \mathcal{H}_{\boldsymbol{d}}[q]$. Then for all $p \ge 2$,*

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

Sketch of proof. By Proposition 2.2, $f \mapsto f(x)$ is an orthogonal projection with respect to the Weyl norm, so $\|f(x)\|_2 \le \|f\|_W$. Hence, for every $x \in S^{n-1}$, $\|f(x)\|_p \le \|f(x)\|_2 \le \|f\|_W$, where the first inequality follows from Minkowski's inequality. □

We finish this subsection by noting how the L_∞-norms relate to the Weyl norm. We note that this is related to the so-called best rank-one approximation of a symmetric tensor [1, 59]; the inequality for the real case below was already present in [59, Theorem 2.4].

Proposition 2.11. *Let $f \in \mathcal{H}_{\boldsymbol{d}}[q]$. Then*

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

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

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

Proof. The first part follows from Proposition 2.9 and 2.10. The left-hand side of the second part uses Proposition 2.10.

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Example 2.12. Proposition 2.11 is almost optimal for $n = 1$. In [1], it was shown that for the cosine polynomials c_d of Example 2.7, we have

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

and that c_d is the real polynomial of real L_∞ norm 1 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: that is, $f \mapsto \overline{\mathrm{D}} f$ 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 a capital X, to emphasise that we view $\overline{\mathrm{D}}_X f$ as a polynomial tuple in $\mathcal{H}_{\boldsymbol{d}-\mathbb{1}}[(n+1)q]$; and we will write $\overline{\mathrm{D}}_x f$, with a lowercase x, to emphasise that we view $\overline{\mathrm{D}}_x f$ 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 [40].

Theorem 2.13 (Kellogg's inequality). *Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{F}}[q]$ and $v \in \mathbb{F}^{n+1}$; then*

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

Corollary 2.14. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{F}}[q]$ and $z \in \mathbb{S}^n$ (if $\mathbb{F} = \mathbb{R}$) or $z \in \mathbb{P}^n$ (if $\mathbb{F} = \mathbb{C}$). Then*

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

Before proving Theorem 2.13 and Corollary 2.14, we discuss some features of these results.

Remark 2.15. We note that the left-hand side in Corollary 2.14 is 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 2.13 looks in a few particular cases.

Example 2.16. Consider the cosine polynomials c_d of Example 2.7. A direct computation shows that

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

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

$$\left\|\frac{1}{d}\overline{\mathrm{D}}_X c_d v\right\|_\infty^{\mathbb{R}} = \|v\|_2 = \|c_d\|_\infty^{\mathbb{R}}\|v\|_2,$$

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, $\left\|(v_X - iv_Y)x^{d-1} + (v_X + iv_Y)y^{d-1}\right\| \le \sqrt{2}\|v\|_2(|x|^{d-1} + |y|^{d-1}) \le \|v\|_2$ for $d \le 3$, and v real when $|x|^2 + |y|^2 \le 1$. Thus

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

$$\left\|\frac{1}{d}\overline{\mathrm{D}}_X c_d v\right\|_\infty^{\mathbb{C}} = \frac{2^d}{d}\|v\|_2 = \frac{\sqrt{2}}{d}\|c_d\|_\infty^{\mathbb{C}}\|v\|_2.$$

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

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

Now it is the complex Kellogg's theorem that 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. Examples 2.16 and 2.17 motivate the search of a randomised Kellogg's theorem that holds with high probability for random polynomials and has a tighter right-hand side.

Proof of Theorem 2.13. We only prove the real case. The complex case is proven analogously (see [40, Section 8] for the complex version of the results we use in the real case).

By [40, 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. 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 [40, Theorem IV] using the scaling with the diagonal matrix was not used in [33, Theorem 2.4] and [34]. This can be used to improve by a factor of the degree some of the bounds there.

Proof of Corollary 2.14. We only prove the real case, the proof for the complex case being essentially the same. Recall that by Euler's formula for homogeneous functions,

$$\Delta^{-1}\overline{\mathrm{D}}_x f x = f(x). \tag{2.4}$$

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 2.13, and the right-hand side equals $\max\{\|f(x)\|_\infty, \|\Delta^{-1}\mathrm{D}_x f\|_{2,\infty}\}$. Thus the desired inequality follows. □

Following the notations introduced above, we will write $\overline{\mathrm{D}}_x^k f$ to denote the kth derivative map of $f \in \mathcal{H}_{\boldsymbol{d}}[q]$ at $x \in \mathbb{F}^{n+1}$. This is the k-multilinear map $(\mathbb{F}^{n+1})^k \to \mathbb{F}^q$ given by the kth derivatives of f

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

at x. Also, $\overline{\mathrm{D}}_X^k f(v_1,\dots,v_k)$, where $v_1,\dots,v_k \in \mathbb{F}^{n+1}$ will denote the corresponding polynomial tuple in $\mathcal{H}_{\boldsymbol{d}-k\mathbb{1}}[q]$. For a real k-multilinear map $A : (\mathbb{R}^n)^k \to \mathbb{R}^q$, we define

$$\|A\|_{2,\infty}^{\mathbb{R}} := \sup_{v_1,\dots,v_k\neq 0} \frac{\|A(v_1,\dots,v_k)\|_\infty}{\|v_1\|\cdots\|v_k\|}. \tag{2.5}$$

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 for $k > 2$, by the following corollary and Example 2.7,

$$\left\|\frac{1}{k!}\overline{\mathrm{D}}_0^k c_k\right\|_{2,\infty}^{\mathbb{C}} = \|c_k\|_\infty^{\mathbb{C}} = 2^{\frac{k}{2}-1} > 1 = \|c_k\|_\infty^{\mathbb{R}} = \left\|\frac{1}{k!}\overline{\mathrm{D}}_0^k c_k\right\|_{2,\infty}^{\mathbb{R}},$$

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. The following corollary (which is closely related to [59, Theorem 2.1]) will be useful later.

Corollary 2.20. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{F}}[q]$ and $z \in \mathbb{S}^n$ (if $\mathbb{F} = \mathbb{R}$) or $z \in \mathbb{P}^n$ (if $\mathbb{F} = \mathbb{C}$). Then for all $k \geq 1$ and $v_1,\dots,v_k \in \mathbb{F}^{n+1}$,*

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

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

Proof. It follows from Theorem 2.13 by induction, followed by an application of Corollary 2.14. □

Remark 2.21. Although the results in this section were proved only for $\|\ \|_\infty^{\mathbb{F}}$, some of them can be generalised to other norms. For example, similar results can be obtained for $\|\ \|_2^{\mathbb{R}}$ (see [52]) 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 Corollary 2.14 for $\mathbb{F} = \mathbb{R}$ can be generalised to smooth real algebraic varieties other than the sphere (see [11]).

3. Condition numbers for the L_∞-norm

In this section, we will consider condition numbers that capture 'how near to being singular' a system $f \in \mathcal{H}_{\boldsymbol{d}}[q]$ 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.

Recall the local and global versions of the real condition number κ used in [25, 26, 27, 28]. For $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $x \in \mathbb{S}^n$, they are defined by

$$\kappa(f,x) := \frac{\|f\|_W}{\sqrt{\|f(x)\|_2^2 + \left\|\mathrm{D}_x f^\dagger \Delta^{1/2}\right\|_{2,2}^{-2}}} \quad \text{and} \quad \kappa(f) := \sup_{y\in\mathbb{S}^n} \kappa(f,y). \tag{3.1}$$

Here, for a surjective linear map A, $A^\dagger := A^*(AA^*)^{-1}$ denotes its Moore-Penrose inverse [14, Section 1.6]. Also recall the μ-condition number introduced by Shub and Smale [53]: for $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[q]$ and $\zeta \in \mathbb{P}^n$, $\mu(f,\zeta)$ is defined by

$$\mu_{\mathrm{norm}}(f,\zeta) := \|f\|_W \left\|\mathrm{D}_\zeta f^\dagger \Delta^{1/2}\right\|_{2,2}. \tag{3.2}$$

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

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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

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

$$\mathsf{K}(f,x) := \frac{\sqrt{q}\|f\|_\infty^{\mathbb{R}}}{\max\left\{\|f(x)\|, \left\|\mathrm{D}_x f^\dagger \Delta\right\|_{2,2}^{-1}\right\}} \tag{3.3}$$

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

$$\mathsf{K}(f) := \sup_{y \in \mathbb{S}^n} \mathsf{K}(f,y). \tag{3.4}$$

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

$$\mathsf{M}(f,\zeta) = \sqrt{q}\|f\|_\infty^{\mathbb{C}}\left\|\mathrm{D}_\zeta f^\dagger \Delta\right\|_{2,2} \tag{3.5}$$

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

$$\mathsf{M}(f) := \sup\{\mathsf{M}(f,\zeta) \mid \zeta \in \mathbb{P}^n,\ f(\zeta) = 0\}. \tag{3.6}$$

We can see that K is a variant of κ and M is a variant of μ_{norm}. We note that the main difference lies in the fact that we are substituting all occurrences of $\| \ \|_W$ with occurrences of $\| \ \|_\infty$. The fact that we use a different scaling factor ($\Delta^{1/2}$ instead of Δ) or different norms for vectors ($\| \ \|_\infty$ instead of $\| \ \|_2$ and so on) only affects these quantities up to a $\sqrt{2q\mathbf{D}}$ factor. This has little consequence for complexity. We will be more explicit in Proposition 4.27. Note that despite these changes, we still have that the 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 proving the main properties of K and M, which are the reason we defined these numbers the way we did. The properties we will show are those needed for a condition-based complexity analyses of the algorithms in Sections 4 and 5 following the lines of the analyses in [25, 28, 15, 16, 17] (see also [55]) and [14, Chapter 17].

3.1. *Properties of the real condition number* K

Recall (see, e.g., [14, Definition 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

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

where $\| \ \| = \| \ \|_{2,2}$ is the operator norm (with respect to 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*

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

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.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

- *1st Lipschitz property: The maps*

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

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

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

- *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 to 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 $\mathrm{D}_x f^\dagger$ exists 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. In the particular case of K, we could state it with equalities (see its proof below), but we leave the statement with inequalities as this is the one holding for κ as well and it is enough for our purposes.

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

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

whenever $\mathsf{K}(f,x)\frac{\|f-\tilde{f}\|_\infty^{\mathbb{R}}}{\|f\|_\infty^{\mathbb{R}}} < 1$. This formula shows how the condition number of an approximation of f relates to that of f.

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

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

whenever $\mathsf{K}(f,x)\mathrm{dist}_{\mathbb{S}}(x,\tilde{x}) < 1$. Here $\mathrm{dist}_{\mathbb{S}}$ denotes the geodesic distance in $\mathbb{S}^n$.

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 with f. Recall (see [14, Definition 16.34]) that we say $x \in \mathbb{S}^n$ is an *approximate zero* of $f \in \mathcal{H}_{\boldsymbol{d}}[n]$ with associated zero $\zeta \in \mathbb{S}^n$ when for all $k \geq 1$, the kth iteration

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

N_f^k of N_f satisfies

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

We have the following result (see [14, Theorem 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 $\operatorname{dist}_{\mathbb{S}}(z, \zeta) \leq \frac{1}{45}$ and $\operatorname{dist}_{\mathbb{S}}(z, \zeta)\gamma(f, \zeta) \leq 0.17708$. 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,

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

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,

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

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

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

because σ_q is 1-Lipschitz with respect to $\|\ \|_{2,2}$, $\|\ \| \leq \sqrt{q}\|\ \|_\infty$ and Kellogg's inequality (Theorem 2.13). Thus our claims follow.

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

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

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, 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 $u \in \mathcal{O}(n+1)$ be

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

the planar rotation taking y into $\tilde{y}$. Then

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

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 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 2.14, 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. □

The following proposition, which we state here for the sake of completeness, will be proved in Section A.

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

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

and

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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

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

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 is identical, so we omit it. We do not consider a regularity inequality for M since over complex numbers one usually considers $\mathsf{M}(f,\zeta)$ for a zero ζ of f (or a point nearby).

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*

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

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

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

- *2nd Lipschitz property: The map*

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

  is $\mathbf{D}$*-Lipschitz with respect to the geodesic distance* $\mathrm{dist}_{\mathbb{P}}$ *on* $\mathbb{P}^n$.
- *Higher derivative estimate: We have*

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

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 [14, Proposition 16.55] is to μ_{norm}.

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

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

Proof. Note that

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

For the first term in the sum, we have

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

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 term, we have

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

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

Hence, we have

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

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

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

so, from

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

we get

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

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

4. Numerical algorithms in real algebraic geometry

There is a growing literature on numerical algorithms that addresses basic computational tasks in real algebraic geometry, such as counting real zeros [25, 26, 27], computing homology of algebraic [28] and semialgebraic sets [15, 16, 17], and meshing real curves and surfaces [49, 23]. These works rely on condition numbers to control precision and estimate computational complexity.

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 Section 4.1 and Section 4.2.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

2. When we consider random input models, the gains in the complexity estimates become more evident. In Section 4.3, we show that the ratio of the new K to κ is typically of the order of $\sqrt{n}/\sqrt{N}$ for a random polynomial system. 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 what follows, we will focus on algorithms dealing with real algebraic sets. The algorithms we have in mind are the ones in [25, 26, 27, 28] and the Plantinga-Vegter algorithm [49] as described and analysed in [24] (compare to [23]). Our condition number K as defined in the preceding section will improve the overall computational complexity of these algorithms. Similar results can be obtained for the algorithms dealing with semialgebraic sets in [15, 16, 17] (compare to [55]) using natural extensions $\overline{\mathsf{K}}$ and K_* of the condition numbers $\overline{\kappa}$ and κ_* used in these papers.

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

A *grid-based algorithm* is a subdivision-based method that constructs a grid to discretise 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 [25, 26, 27], [28] and [15, 16, 17] (compare to [55]) 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 [25, 26, 27] and the computation of the Betti numbers of a projective real variety [28]. For simplicity of exposition, we omit some computational aspects: 1) our presentation of the algorithms follows the construction-selection paradigm of [15, 16, 17] instead of the inclusion-exclusion paradigm of [25, 26, 27, 28]. This makes the exposition of the algorithms easier without compromising their computational complexity. 2) We focus on Betti numbers to avoid describing the more involved computation of torsion coefficients in the homology groups. 3) We deal with neither parallelisation nor finite precision. The interested reader can find details about these in the cited references.

The backbone of existing grid-based algorithms in numerical real algebraic geometry [25, 26, 27, 28, 15, 16, 17] is an effective construction of spherical nets. The basic construction was done originally in [25] and 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 $x \in \mathbb{S}^n$, $\operatorname{dist}_{\mathbb{S}}(x, \mathcal{G}) < \delta$. We will omit the term 'spherical' as all nets we consider are so.

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*

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

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

Remark 4.2. The grid construction in Proposition 4.1, which occurs in [25, 26, 27, 28, 15, 16, 17], is not optimal. This is due to the $2^{n \log n}$ factor in the estimates, which can be decreased to $2^{\mathcal{O}(n)}$. An

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

algorithm doing this – that is, constructing 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 – is given in [2, Theorem 1.9(1)]. We use the suboptimal result of Proposition 4.1 to focus on the effect of just changing the norm when comparing the old and new versions of the algorithms. But we observe here that by using the nets in [2], one can remove the $\log(n)$ factors in the exponents.

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

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

Proposition 4.3. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $\mathcal{G} \subset \mathbb{S}^n$ be a δ-net. If $\mathbf{D}\delta < \sqrt{2}$, then*

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

Proof. We only need to show the right-hand inequality, the other being trivial. Without loss of generality, assume that $q = 1$: that is, 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 $\mathrm{dist}_{\mathbb{S}}(x_*, x) \le \delta$ 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 $|M(1)| \le |M(0)| + |M'(0)| + \max_{s\in[0,1]} \frac{M''(s)}{2}$ by Taylor's theorem. Furthermore, $|M(0)| = |f(x_*)| = \|f\|_\infty^{\mathbb{R}}$, $|M(1)| = |f(x)|$ and $M'(0) = 0$. The latter is because 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 in equation (2.4). Then by Corollary 2.20,

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

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

Remark 4.4. Proposition 4.3 is a slight improvement of [33, Lemma 2.5].

Proposition 4.3 suggests the following algorithm.

Algorithm 4.1: NormApproxℝ

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

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

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

Proposition 4.5. *Algorithm NormApproxℝ 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 4.3 and the fact that f can be evaluated at $x \in \mathbb{S}^n$ with $\mathcal{O}(N)$ arithmetic operations (see [14, Lemma 16.31]). □

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Remark 4.6. The ideas here can also 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 grid needed to estimate K depends on K itself.

Proposition 4.7. *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

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

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 $\operatorname{dist}_{\mathbb{S}}(f,x) \le \delta$. Then by the 2nd Lipschitz property (Theorem 3.2), we have

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

Hence $1/\mathsf{K}(f,x_*) \le (1-\delta\,\mathbf{D}\,\mathsf{K}(f,x))/\mathsf{K}(f,x)$, and the desired inequality follows from the hypothesis. □

Proposition 4.7 suggests the following algorithm, which involves only one L_∞-norm computation.

Algorithm 4.2: K-Estimate

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

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

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

Proposition 4.8. *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*

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Proof. The correctness follows from Propositions 4.5 and 4.7 and $(1-2^{-(k+1)})^2 > 1-2^{-k}$.

The cost of the first line of the algorithm is bounded by Proposition 4.5. 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 [14, 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, $\log \mathcal{K} - \log(1-2^{-k}) \le \log \mathsf{K}(f)$, so $\ell_0 \le k+1+\log \mathbf{D} + \log \mathsf{K}(f)$.

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

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

So, in this case, $\ell_0 \le k+1+\log \mathbf{D} + \log b$.

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

By setting k to 7 and $b=\infty$, we have the following important corollary.

Corollary 4.9. *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*

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

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 4.10 (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* $0.99Q \le \|f\|_\infty^{\mathbb{R}} \le Q$*, 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 [35], recall that the *medial axis* Δ_X of a closed set $X \subset \mathbb{R}^n$ is the set

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

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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

measuring the size of the neighbourhood of X within 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, when $\mathcal{Z}_{\mathbb{S}}(f)$ is zero-dimensional, $2\tau(\mathcal{Z}_{\mathbb{S}}(f))$ is the separation of the zeros of f in the sphere.

Theorem 4.11. *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 4.10. Let $x_0 \in \mathcal{Z}_{\mathbb{S}}(f)$. Then there is some $x_1 \in \mathcal{G}$ such that $\operatorname{dist}_{\mathbb{S}}(x_0, x_1) \leq \delta$. Let $[0,1] \ni t \mapsto x_t$ be the geodesic joining them. By Taylor's theorem,

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

so, by Kellogg's theorem (Corollary 2.14) 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

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

Now let $x_2 \in \mathcal{G}$ 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 equation (4.1) and the hypothesis in the second. This means Smale's α-criterion holds for x_2 and $f_{|\mathrm{T}_{x_0}\mathbb{S}^n}$ by [29, Théorème 128]. Hence there is $x_3 \in \mathrm{T}_{x_2}\mathbb{S}^n$ such that $f(x_3) = 0$ and

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

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

Remark 4.12. The proof also shows the convergence of Newton's method associated with $f_{|\mathrm{T}_x\mathbb{S}^n}$ 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 needed.

Sketch of proof of Theorem 4.11. The proof is very similar to the one of [15, Theorem 4.12]. By [15, Lemma 2.7] and [15, 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. □

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

The following theorem is a variant of the so-called Niyogi-Smale-Weinberger theorem [47, Proposition 7.1].

Theorem 4.13. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$, $\mathcal{G} \subset \mathbb{S}^n$ be a δ-net and $Q \in \mathbb{R}$ 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 [15, Theorem 2.8] combined with Theorems 4.10 and 4.11. □

We can now describe the algorithm. 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 [32].

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$\mathbb{R}(f, 7)$
$\mathcal{K} \leftarrow$ K-Estimate* (f)
$\ell \leftarrow 7 + \lceil 2\log\mathbf{D} + 2\log\mathcal{K}\rceil$
$\mathcal{G} \leftarrow$ Grid(n, ℓ)
$\mathcal{X} \leftarrow \{x \in \mathcal{G} \mid \|f(x)\| < \sqrt{q}\,\mathbf{D}Q2^{-\ell}\}$
$\varepsilon \leftarrow 3/(50\mathbf{D}\mathcal{K})$
$(\beta_0, \ldots, \beta_n) \leftarrow$ 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 4.14. *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 4.13 and the fact that the computed Q satisfies $0.99Q \leq \|f\|_\infty^{\mathbb{R}} \leq Q$ by Proposition 4.5.

For the complexity, we apply Proposition 4.3 for the first line, Corollary 4.9 for the second line and Proposition 4.1 for the fourth and fifth lines. We know that Betti has cost $\mathcal{O}\big(2^{\mathcal{O}(n\log n)}|\mathcal{X}|^{5n}\big)$ (see [28, Section 5] for example) and that $|\mathcal{X}| = \mathcal{O}(2^{n\log n}\mathbf{D}^{2n}\mathsf{K}(f)^{2n})$, by Proposition 4.1. Note that we have eliminated N from the bounds. We have done so using the fact that as $q \leq n$ (by the precondition of the input), $N \leq 2^{n\log n}\mathbf{D}^n$.

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

The complexity estimate in Proposition 4.14 does not differ much from those in other grid-based algorithms. We will see in Section 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 input models.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

4.2. Complexity of the Plantinga-Vegter algorithm

The ideas above can also be applied to the Plantinga-Vegter algorithm [49]. In a recent work [24] (compare to [23]), we performed an extensive analysis of this algorithm, including details for finite precision arithmetic. So we will be brief here, referring the reader to [24] for details, and will 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 [49][2] is a subdivision-based algorithm for obtaining a piecewise linear approximation 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 iteratively subdivide some boxes – that is, sets of the form $B = m(B) + [-w(B)/2, w(B)/2]^n$ (here $m(B) \in \mathbb{R}^n$ is the centre of B and $w(B) > 0$ is its width) – in $[-a, a]^n$ until every box B in the subdivision satisfies the following condition:

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

where $\langle\ ,\ \rangle$ is the standard inner product and ∇f is the gradient vector of f. Once this criterion is satisfied by all boxes in the subdivision, the Plantinga-Vegter algorithm returns a topologically accurate approximation of the zero set of f in the region $[a, -a]^n$ and halts (see [49] ($n \leq 3$) and [37] (arbitrary n) for details on how this is done).

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 homogenisation of f. Taking the maps (2.3), (2.4), (2.5) in [24] and substituting on them the Weyl norm by the real L_∞-norm, we get

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

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

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

One can use these maps to produce interval approximations as we do in [24]. For $X \subseteq \mathbb{R}^m$, we denote by $\square X$ the set of boxes contained in X. Recall that an *interval approximation* of $f : \mathbb{R}^n \to \mathbb{R}^q$ is a function $\square f : \square\mathbb{R}^n \to \square\mathbb{R}^q$ 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 4.15. *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

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

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

[2]The original algorithm [49] only dealt with dimensions two and three. For the extension to dimensions four or higher, see [37].

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Sketch of proof. Using the bounds from Kellogg's theorem (Theorem 2.13) and its corollaries, we can easily deduce (as 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 [24, Section 4], but using these Lipschitz properties, to prove that $\hat{f}$ and $\widehat{\nabla f}$ are $(1+d)$- and d-Lipschitz, respectively. For the latter, we use the fact that for $v \in \mathbb{R}^n$, $\overline{\mathrm{D}}_X f^{\mathsf{h}}\begin{pmatrix}0\\ v\end{pmatrix} = (\langle \nabla f, v\rangle)^{\mathsf{h}}$and that $\|\widehat{\nabla f}\|$ is d-Lipschitz if $\langle \widehat{\nabla f}, v\rangle$ is so for every $v \in \mathbb{S}^{n-1}$.

Using the interval approximations and their Lipschitz properties in Proposition 4.15, we can rewrite the condition $C_f(B)$. We only need to use [24, Lemma 4.2] for the second clause of the condition.

Theorem 4.16. *Let $B \in \square\mathbb{R}^n$. If the condition*

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

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. Recall that $\square[-a,a]^n$ 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 $\|\nabla f(m(B))\| > 2\sqrt{2}d\sqrt{n}w(B)Q(1+\|m(B)\|^2)^{\frac{d}{2}-1}$ **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 much effort, [24, Proposition 5.1] transforms into the following proposition. The essential step is multiplying the inequalities in that proposition by $\|f^{\mathsf{h}}\|_W/\|f\|_\infty$.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Proposition 4.17. *Let $f \in \mathcal{P}_d$ and $x \in \mathbb{R}^n$. Then either*

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

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

With Proposition 4.17 and the Lipschitz properties shown for $\hat{f}$ and $\widehat{\nabla f}$, one can produce a *local size bound* for $C_f^{\square}(B)$. This is a function that, evaluated at a point x, gives a lower bound on the volume of any possible box containing x and not satisfying the predicate $C'_f(B)$.

Theorem 4.18. *The map*

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

is a local size bound for $C_f^{\square}$ (of Theorem 4.16).

Then using the continuous amortisation of [20, 18, 19] (see [24, Theorem 6.1]), we conclude the following, which takes into account the cost of calling NormApprox$\mathbb{R}$ (Proposition 4.3).

Theorem 4.19. *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

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

The condition-based estimates in Theorem 4.19 are very similar to those of [24, 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 4.19. We will see in Section 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 that use (in their design and/or analysis) 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 along with some useful probabilistic results. Then we prove a general comparison result showing 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 estimates to both PolyBetti and the Plantinga-Vegter algorithm and highlight the complexity improvements.

For most algorithms in real algebraic geometry, condition-based estimates show a dependence on either κ^n or 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 estimates thus change from having an exponent quadratic in n to an exponent quasilinear in n.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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 *centred* if $\mathbb{E}\,\mathfrak{x} = 0$.
(ii) $\mathfrak{x}$ is *subgaussian* if there is a constant $K > 0$ 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 [23].

Definition 4.20. 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

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

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

Remark 4.21. 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] $6K\rho \geq 1$.

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 independent and identically distributed 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 4.23. A dobro random polynomial tuple $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ such that the $\mathfrak{c}_\alpha$ are are independent and identically distributed 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.

Proposition 4.24 (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 $t > 0$, $\mathbb{P}(|\mathfrak{x}| \geq t) \leq \mathrm{e}^{1-\frac{t^2}{6K^2}}$.*
2. *If there are $C \geq \mathrm{e}$ and $K > 0$ such that for all $t > 0$, $\mathbb{P}(|\mathfrak{x}| \geq t) \leq C\mathrm{e}^{-\frac{t^2}{K^2}}$, then $\mathfrak{x}$ is subgaussian with ψ_2-norm at most $K\left(\sqrt{\pi/2} + \sqrt{2\ln C}\right)$.*

Proposition 4.25 (Hoeffding inequality). *Let $\mathfrak{x} \in \mathbb{R}^N$ be a random vector such that its components $\mathfrak{x}_i$ are centred subgaussian random variables with ψ_2-norm at most K and $a \in \mathbb{S}^{N-1}$. Then for all $t \geq 0$,*

$$\mathbb{P}_{\mathfrak{x}}(|a^*\mathfrak{x}| \geq t) \leq 2\mathrm{e}^{-\frac{t^2}{11K^2}}.$$

In particular, $a^*\mathfrak{x}$ is a subgaussian random variable with ψ_2-norm at most $5K$.

[3]This follows from $2tK\rho \geq \mathbb{P}_{\mathfrak{x}}(|\mathfrak{x}| \leq Kt) \geq 1 - \mathbb{P}_{\mathfrak{x}}(|\mathfrak{x}| > Kt) \geq 1 - 2\mathrm{e}^{-t^2/2}$ and optimising, 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.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Proposition 4.26 (Anti-concentration bound). *Let $\mathfrak{x} \in \mathbb{R}^N$ be a random vector such that its components $\mathfrak{x}_i$ are independent random variables with anti-concentration property with constant ρ. Then for every $A \in \mathbb{R}^{k\times N}$ with rank k and measurable $U \subseteq \mathbb{R}^k$,*

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

Proof of Proposition 4.24. This is just [58, Proposition 2.5.2] with improved constants.

For the first part, we give a proof since we don't explicitly use the constants in the proof of [58, Proposition 2.5.2]. Fix $\lambda > 0$. Then by Markov's inequality and expanding the exponential as a power series,

$$\mathbb{P}(|\mathfrak{x}| \geq t) = \mathbb{P}\Big(\mathrm{e}^{\lambda^2\mathfrak{x}^2} \geq \mathrm{e}^{\lambda^2 t^2}\Big) \leq \mathrm{e}^{-\lambda^2 t^2} \sum_{p=0}^{\infty} \frac{\lambda^{2p}\,\mathbb{E}\,\mathfrak{x}^{2p}}{p!} \leq \mathrm{e}^{-\lambda^2 t^2} \sum_{p=0}^{\infty} \frac{(\lambda^2 2pK^2)^p}{p!}.$$

Now, by setting the value of λ to $\frac{1}{\sqrt{6}K}$, $\mathbb{P}(|\mathfrak{x}| \geq t) \leq \mathrm{e}^{-\frac{t^2}{8K^2}} \sum_{p=0}^{\infty} \frac{(p/3)^p}{p!}$. The right-hand series is convergent, and after adding the series numerically, we can see that $\sum_{p=0}^{\infty} \frac{(p/3)^p}{p!} = 2.625\ldots \leq \mathrm{e}$, which finishes the proof of the first part. Following the constants in the proof of [58, Proposition 2.5.2] directly seems to give $4\mathrm{e} \simeq 10.8$ in the denominator of the exponent instead of 6.

For the second, note that

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

which follows from

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

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 $t = \frac{u^2}{2K}$, we obtain

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

Hence

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

from which the second part follows. □

Proof of Proposition 4.25. This is a version of [58, Proposition 2.6.1]. Let us sketch a proof to see the values of the chosen constants.

Let $\mathfrak{y} \in \mathbb{R}$ be a centred random variable with ψ_2-norm at most K. Arguing as in part '**ii** $\Rightarrow$ **iii**' of the proof of [58, Proposition 2.5.2], we have that for all $\lambda \in [-1/\sqrt{2\mathrm{e}}, 1/\sqrt{2\mathrm{e}}]$,

$$\mathbb{E}\,\mathrm{e}^{\lambda^2\mathfrak{y}^2} \leq \mathrm{e}^{\mathrm{e}K^2\lambda^2},$$

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

using $n! \geq \sqrt{2\pi}(n/e)^n$ and that for $x \in [-1/2, 1/2]$, we have $1 + \frac{1}{\sqrt{2\pi}} \frac{x^2}{1-x^2} \leq \mathrm{e}^{x^2/2}$. Then arguing as in part '**iii** $\Rightarrow$**v**' of the proof of [58, Proposition 2.5.2], we get that for all $\lambda \in \mathbb{R}$,

$$\mathbb{E}\, \mathrm{e}^{\lambda \mathfrak{y}} \leq \mathrm{e}^{\mathrm{e}K^2\lambda^2}. \tag{4.5}$$

In this way, we have that

$$\begin{aligned}
\mathbb{P}(|a^*\mathfrak{x}| \geq t) &\leq 2\mathbb{P}(a^*\mathfrak{x} \geq t) && \text{(symmetry)}\\
&= 2\mathbb{P}(\mathrm{e}^{a^*\mathfrak{x}} \geq \mathrm{e}t)\\
&\leq 2\mathrm{e}^{-\lambda t}\, \mathbb{E}\, \mathrm{e}^{\lambda a^*\mathfrak{x}} && \text{(Markov's inequality)}\\
&= 2\mathrm{e}^{-\lambda t} \prod_{i=1}^{N} \mathbb{E}\, \mathrm{e}^{\lambda a_i \mathfrak{x}_i} && (a_1\mathfrak{x}_1, \ldots, a_N\mathfrak{x}_N \text{ independent})\\
&\leq 2\mathrm{e}^{-\lambda t} \prod_{i=1}^{N} \mathrm{e}^{\mathrm{e}a_i^2K^2\lambda^2} && (4.5)\\
&= 2\mathrm{e}^{-\lambda t + \mathrm{e}K^2\lambda^2} && (\|a\|_2 = 1).
\end{aligned}$$

Taking $\lambda = \frac{t}{2\mathrm{e}K^2}$, we get the desired tail bound. The last claim immediately follows from Proposition 4.24. □

Proof of Proposition 4.26. This is a rewriting of [51, Theorem 1.1] using [44] to get explicit constants. This rewriting was first given in [56, 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

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

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 [51, Theorem 1.1] (see [44, 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: Measuring 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. This way, the comparison between the two algorithms reduces to estimate K/κ. The following proposition shows that, in turn, this amounts to looking at the quotient $\|f\|_\infty^{\mathbb{R}}/\|f\|_W$.

Proposition 4.27. *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

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Proof. It follows from

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

and

$$\frac{1}{\sqrt{\mathbf{D}}}\,\sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right) \le \sigma_q\left(\Delta^{-1}\mathrm{D}_x f\right) \le \sigma_q\left(\Delta^{-\frac{1}{2}}\mathrm{D}_x f\right). \qquad \square$$

In general, we have that $\frac{\|f\|_\infty^{\mathbb{R}}}{\|f\|_W} \le 1$, 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 fs, much smaller than 1: we can expect it to be smaller than $\sqrt{n\ln(\mathrm{e}\mathbf{D})/N}$ with very high probability. Recall that $K\rho$ is a constant from the randomness model.

Theorem 4.28. *Let $q \le n+1$, $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ be dobro with parameters K and ρ and $\ell \in \mathbb{N}$. For any power ℓ with $1 \le \ell < \frac{N}{2}$, we have*

$$\mathop{\mathbb{E}}_{\mathfrak{f}}\left(\frac{\|\mathfrak{f}\|_\infty^{\mathbb{R}}}{\|\mathfrak{f}\|_W}\right)^\ell \le \left(\frac{890\sqrt{2}K\rho\sqrt{n\ln(eD)\ell}}{\sqrt{N-2\ell}}\right)^\ell.$$

In particular,

$$\mathop{\mathbb{E}}_{\mathfrak{f}}\frac{\|\mathfrak{f}\|_\infty^{\mathbb{R}}}{\|\mathfrak{f}\|_W} \le \mathcal{O}\left(K\rho\sqrt{\frac{n\ln(e\mathbf{D})}{N}}\right).$$

Remark 4.29. In the study of tensors, the quotients $\|\mathfrak{f}\|_\infty^{\mathbb{R}}/\|\mathfrak{f}\|_W$ and their nonsymmetric analogue play an important role. Because of this, we can consider Theorem 4.28 a symmetric analogue of the results shown in [38] and [46]. In a paper under preparation by Kozhasov and the third author [41], the probabilistic techniques introduced in this paper are developed further to study $\|\mathfrak{f}\|_\infty^{\mathbb{R}}/\|\mathfrak{f}\|_W$ in several settings.

Corollary 4.30. *Let $q \le n+1$ and $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ be dobro with parameters K and ρ. Then for $1 \le \ell < \frac{N}{2}$, we have*

$$\mathop{\mathbb{E}}_{\mathfrak{f}}\left(\frac{\mathsf{K}(\mathfrak{f})}{\kappa(\mathfrak{f})}\right)^\ell \le \left(\frac{1780}{K}\rho\sqrt{qn\mathbf{D}\ln(eD)\ell}\sqrt{N-2\ell}\right)^\ell.$$

Let $\textsc{PolyBetti}_W$ be the version of $\textsc{PolyBetti}_\infty$ using the Weyl norm and κ. An analysis along the lines of [28] (or [15]) shows that the cost of $\textsc{PolyBetti}_W$ 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 4.14. Let us denote by $\text{est-run-time}(\textsc{PolyBetti}_\infty, f)$ and $\text{est-run-time}(\textsc{PolyBetti}_W, f)$ these cost bounds. It follows that

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

Using Corollary 4.30 and Markov's inequality, it is easy to prove the following estimate.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Corollary 4.31. *Let $q \le n+1$, $N > 20n$ and $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ be dobro with parameters K and ρ,*

$$\frac{\text{est-run-time}(\textsc{PolyBetti}_\infty, f)}{\text{est-run-time}(\textsc{PolyBetti}_W, f)} \le \left(\frac{5700}{K}\rho n\sqrt{q\mathbf{D}\ln(\mathrm{e}\mathbf{D})}\sqrt{N-20n}\right)^{10n}$$

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 the order of

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

We proceed to prove Theorem 4.28.

Proposition 4.32. *Let $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ be dobro with parameters K and ρ. Then for all $t > 0$,*

$$\mathbb{P}\Big(\|\mathfrak{f}\|_\infty^{\mathbb{R}} \ge t\Big) \le q\sqrt{2\pi}\sqrt{n+1}\left(\frac{\mathrm{e}\mathbf{D}}{2}\right)^n \mathrm{e}^{-\frac{t^2}{17K^2}}.$$

In particular, if $q \le n+1$, for all $\ell \ge 1$, $\left(\mathbb{E}_{\mathfrak{f}}\left(\|\mathfrak{f}\|_\infty^{\mathbb{R}}\right)^\ell\right)^{\frac{1}{\ell}} \le 63K\sqrt{n\ln(\mathrm{e}\mathbf{D})}\ell$.

Proof of Theorem 4.28. By the Cauchy-Schwarz inequality,

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

The first term on the right is bounded by Proposition 4.32.

For the second term, we will use [45, Theorem 1.11]. We note that $\mathfrak{x} \in \mathbb{R}^N$ satisfies the *small ball assumption (SBA) with constant $\mathcal{L}$* [45, Assumption 1.1.] if for every $k \in \{1,\dots,N-1\}$, every orthogonal projection $P \in \mathbb{R}^{k\times N}$, every $y \in \mathbb{R}^k$ and every $\varepsilon > 0$,

$$\mathbb{P}\Big(\|P\mathfrak{x} - y\|_2 \le \sqrt{k}\varepsilon\Big) \le (\mathcal{L}\varepsilon)^k.$$

By Proposition 4.26 (applied with coordinates orthogonal with respect to the Weyl inner product) and Stirling's approximation, we have that $\mathfrak{f}$ has the SBA with constant $2\sqrt{\pi e}\rho$. Thus, by [45, Theorem 1.11],

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

where $\mathfrak{g} \in \mathcal{H}_{\boldsymbol{d}}[q]$ is KSS. Since $\mathfrak{g}$ is a Gaussian vector for all coordinate systems orthogonal with respect to the Weyl inner product, $\|\mathfrak{g}\|_W^2$ is distributed according to a χ^2-distribution with N degrees of freedom. Therefore,

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

The desired claim now follows. □

Proof of Proposition 4.32. Fix $\delta \in [0, 1/\mathbf{D}]$. By the proof of Proposition 4.3, we have that $\|\mathfrak{f}\|_\infty^{\mathbb{R}} > t$ implies $\operatorname{vol}\left\{x \in \mathbb{S}^n \mid \|\mathfrak{f}(x)\|_\infty \ge \left(1 - \frac{\mathbf{D}^2}{2}\delta^2\right)t\right\} \ge \operatorname{vol} B_{\mathbb{S}}(x_*, \delta)$, where $x_* \in \mathbb{S}^n$ maximises $\|f(x)\|_\infty$. Therefore,

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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

By [14, Lemma 2.25], [14, Lemma 2.31] and $\int_0^\delta n\sin^{n-1}\theta\,\mathrm{d}\theta \geq (1-\delta^2/6)^n\delta^n$, we have that

$$\operatorname{vol}_n B_{\mathbb{S}}(x_*,\delta)/\operatorname{vol}_n \mathbb{S}^n \geq \frac{\left(1-\delta^2/6\right)^n}{\sqrt{2\pi}\sqrt{n+1}}\delta^n.$$

In this way,

$$\begin{aligned}
&\mathbb{P}\Big(\|\mathfrak{f}\|_\infty^{\mathbb{R}} \geq t\Big)\\
&\leq \mathbb{P}_{\mathfrak{f}}\left(\mathbb{P}_{\mathfrak{x}\in\mathbb{S}^n}\left(\|\mathfrak{f}(\mathfrak{x})\|_\infty \geq \left(1-\frac{\mathbf{D}^2}{2}\delta^2\right)t\right) \geq \frac{\left(1-\delta^2/6\right)^n}{\sqrt{2\pi}\sqrt{n+1}}\delta^n\right)\\
&\leq \frac{\sqrt{2\pi}\sqrt{n+1}}{\left(1-\delta^2/6\right)^n\delta^n}\,\mathop{\mathbb{E}}_{\mathfrak{f}}\mathbb{P}_{\mathfrak{x}\in\mathbb{S}^n}\left(\|\mathfrak{f}(\mathfrak{x})\|_\infty \geq \left(1-\frac{\mathbf{D}^2}{2}\delta^2\right)t\right) && \text{(Markov's inequality)}\\
&\leq \frac{\sqrt{2\pi}\sqrt{n+1}}{\left(1-\delta^2/6\right)^n\delta^n}\,\mathop{\mathbb{E}}_{\mathfrak{x}\in\mathbb{S}^n}\mathbb{P}_{\mathfrak{f}}\left(\|\mathfrak{f}(\mathfrak{x})\|_\infty \geq \left(1-\frac{\mathbf{D}^2}{2}\delta^2\right)t\right) && \text{(Tonelli's theorem)}\\
&\leq \frac{\sqrt{2\pi}\sqrt{n+1}}{\left(1-\delta^2/6\right)^n\delta^n}\,\max_{x\in\mathbb{S}^n}\mathbb{P}_{\mathfrak{f}}\left(\|\mathfrak{f}(x)\|_\infty \geq \left(1-\frac{\mathbf{D}^2}{2}\delta^2\right)t\right)\\
&\leq \frac{q\sqrt{2\pi}\sqrt{n+1}}{\left(1-\delta^2/6\right)^n\delta^n}\,\max_{i,x\in\mathbb{S}^n}\mathbb{P}_{\mathfrak{f}}\left(|\mathfrak{f}_i(x)|_\infty \geq \left(1-\frac{\mathbf{D}^2}{2}\delta^2\right)t\right) && \text{(Union bound).}
\end{aligned}$$

In the coordinates of a monomial basis orthogonal for the Weyl inner product, the following holds: (1) a dobro random polynomial $\mathfrak{f}$ looks like a random vector whose components are independent and subgaussian of ψ_2-norm at most K, and (2) evaluation at a point of the sphere, $\mathfrak{f}(x)$, becomes the inner product with a vector of norm 1 (by Proposition 2.2). Hence, by Proposition 4.24,

$$\mathbb{P}\Big(\|\mathfrak{f}\|_\infty^{\mathbb{R}} \geq t\Big) \leq \frac{q\sqrt{2\pi}\sqrt{n+1}}{\left(1-\delta^2/6\right)^n\delta^n}\exp\left(-\left(1-\frac{\mathbf{D}^2}{2}\delta^2\right)^2\frac{t^2}{11K^2}\right).$$

The claim follows taking $\delta = 5/(6\mathbf{D})$ and $\left(1-\frac{1}{2}\left(\frac{5}{6}\right)^2\right)\frac{1}{11} \geq \frac{1}{17}$. For the other inequalities on the moments, use Proposition 4.24. □

4.3.3. Complexity of the Plantinga-Vegter algorithm

In [24] (compare to [23]), we proved the following result (which we are just adapting to the notation[5] of this paper).

Theorem 4.33 [24, 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 $t \geq e$,*

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

In particular, for the Plantinga-Vegter algorithm with input $\mathfrak{f}$ over the domain $[-a,a]^n$, the expected number of hypercubes in the final subdivision is at most

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

[5]There is a slight difference in the way the anti-concentration constant is defined in [24] and here.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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 yields better cost bounds than the one using κ: that is, the one in [24].

Theorem 4.34. *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 $t \geq e$,*

$$\mathbb{P}(\mathsf{K}(\mathfrak{f}, x) \geq t) \leq \mathbf{D}^{\frac{n}{2}} (\ln \mathrm{e}\mathbf{D})^{\frac{n+1}{2}} 2^{6n+4} (K\rho)^{n+1} \frac{\ln^{\frac{n+1}{2}} t}{t^{n+1}}.$$

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

$$\mathop{\mathbb{E}}_{\mathfrak{f}} \mathop{\mathbb{E}}_{\mathfrak{x} \in \Omega} (\mathsf{K}(\mathfrak{f}, \mathfrak{x})^n) \leq \mathbf{D}^{\frac{n}{2}} (\ln \mathrm{e}\mathbf{D})^{\frac{n+1}{2}} 2^{\frac{1}{2} n \log n + 5n + 2\log(n) + 7} (K\rho)^{n+1}.$$

In particular, for the Plantinga-Vegter algorithm with input $\mathfrak{f}$ over the domain $[-a, a]^n$, the expected number of hypercubes in the final subdivision is at most

$$a^n \mathbf{D}^{\frac{3n}{2}} (\ln \mathrm{e}\mathbf{D})^{\frac{n+1}{2}} 2^{\frac{3}{2} n \log n + 13n + 2\log(n) + 7}.$$

Remark 4.35. Theorem 4.34 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 (compare to [49]) – that is, $n = 2$ and $n = 3$, with the average complexity analysis from [24] – it is shown in Theorem 3.1 there that PV-Interval$_W$ has an average complexity of

$$\mathcal{O}\Big(d^8 \max\{1, a^2\} (K\rho)^3\Big) \qquad \text{for } n = 2, \text{ and}$$
$$\mathcal{O}\Big(d^{13} \max\{1, a^3\} (K\rho)^4\Big) \qquad \text{for } n = 3.$$

It follows from Theorems 4.19 and 4.34 that the average complexity of PV-Interval$_\infty$ is

$$\mathcal{O}\Big(d^7 \log^{1.5}(d) \max\{1, a^2\} (K\rho)^3\Big) \qquad \text{for } n = 2, \text{ and}$$
$$\mathcal{O}\Big(d^{10} \log^2(d) \max\{1, a^3\} (K\rho)^4\Big) \qquad \text{for } n = 3.$$

We next proceed to prove Theorem 4.34.

Proof of Theorem 4.34. Let $u, t \geq 0$, then

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

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, $\mathbb{P}_{\mathfrak{f}}(\|f\|_\infty^{\mathbb{R}} \geq u)$ is bounded by Proposition 4.32. On the other hand, the map

$$f \mapsto \left(f(x) \;\; \frac{\mathrm{D}_x f}{\mathbf{D}} \right)$$

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

has singular values $1, 1/\sqrt{\mathbf{D}}, \dots, 1/\sqrt{\mathbf{D}}$ 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

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

where ω_n is the volume of the unit n-ball, and we used Proposition 4.26 and Stirling's estimation [14, Equation (2.14)].

Hence, combining the inequalities above,

$$\mathbb{P}_{\mathfrak{f}}(\mathsf{K}(\mathfrak{f}, x) \ge t) \le \sqrt{2\pi(n+1)}\left(\frac{\mathrm{e}\mathbf{D}}{2}\right)^n \mathrm{e}^{-\frac{u^2}{17K^2}} + 9^n \mathbf{D}^{\frac{n}{2}}\left(\frac{u\rho}{\sqrt{n}}\right)^{n+1}\frac{1}{t^{n+1}}.$$

Taking $t \ge e$ and $u = \sqrt{17}K\sqrt{n\ln(\mathrm{e}^2\mathbf{D})\ln t} \ge \sqrt{17}K\sqrt{n\ln\mathbf{D} + (n+1)\ln t}$, we get

$$\mathbb{P}_{\mathfrak{f}}(\mathsf{K}(\mathfrak{f}, x) \ge t) \le \frac{\sqrt{2\pi(n+1)}}{(2\mathrm{e})^n}\frac{1}{t^{n+1}} + 9^n \mathbf{D}^{\frac{n}{2}}\left(\sqrt{17}K\rho\ln^{\frac{1}{2}}(\mathrm{e}^2\mathbf{D})\right)^{n+1}\frac{\ln^{\frac{n+1}{2}} t}{t^{n+1}}.$$

This proves the first statement.

By Tonelli's theorem, 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,

$$\mathop{\mathbb{E}}_{\mathfrak{f}}\mathsf{K}(\mathfrak{f}, x)^n = \int_0^\infty \mathbb{P}_{\mathfrak{f}}\left(\mathsf{K}(\mathfrak{f}, x) \ge t^{\frac{1}{n}}\right)$$
$$\le \mathrm{e}^n + \int_{\mathrm{e}^n}^\infty \mathbb{P}_{\mathfrak{f}}\left(\mathsf{K}(\mathfrak{f}, x) \ge t^{\frac{1}{n}}\right) \le \mathrm{e}^n + \int_{\mathrm{e}^n}^\infty \mathbf{D}^{\frac{n}{2}}\ln^{\frac{n+1}{2}}(\mathrm{e}\mathbf{D})2^{6n+4}(K\rho)^{n+1}\frac{\ln(t^{\frac{1}{n}})^{\frac{n+1}{2}}}{t^{1+\frac{1}{n}}}\,\mathrm{d}t$$
$$\le \mathrm{e}^n + \mathbf{D}^{\frac{n}{2}}\ln^{\frac{n+1}{2}}(\mathrm{e}\mathbf{D})2^{6n+4}(K\rho)^{n+1}\int_1^\infty \frac{\ln(t^{\frac{1}{n}})^{\frac{n+1}{2}}}{t^{1+\frac{1}{n}}}\,\mathrm{d}t.$$

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

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

where the inequality comes from Stirling's approximation [14, Equation (2.14)]. Hence, we get

$$\mathop{\mathbb{E}}_{\mathfrak{f}}\mathsf{K}(\mathfrak{f}, x)^n \le \mathrm{e}^n + \sqrt{2\pi\mathrm{e}}n\sqrt{n+1}\mathbf{D}^{\frac{n}{2}}\ln^{\frac{n+1}{2}}(\mathrm{e}\mathbf{D})\frac{2^{6n+4}}{(2\mathrm{e})^{\frac{n+1}{2}}}(\sqrt{n+1}K\rho)^{n+1}$$
$$\le 8n\sqrt{n+1}\mathbf{D}^{\frac{n}{2}}\ln^{\frac{n+1}{2}}(\mathrm{e}\mathbf{D})\frac{2^{6n+4}}{(2\mathrm{e})^{\frac{n+1}{2}}}(\sqrt{n+1}K\rho)^{n+1}.$$

The second statement 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 $n+1$ homogeneous unknowns can be found on average polynomial time [54]. A probabilistic solution to Smale's

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

17th problem was given by Beltrán and Pardo in 2009 [7, 8]. The construction of Beltrán and Pardo was probabilistic in the sense that they exhibited a *randomised* algorithm.

The distribution underlying the average-case analysis for the Beltrán-Pardo algorithm is the complex version of the KSS distribution (see Example 4.22). Finally, the expected running time of Beltrán-Pardo's algorithm is polynomial in $N = \dim_{\mathbb{C}} \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$.

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 [53], 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 that we end up with an approximation to the zero ζ_1 of $f = q_1$.

The breakthrough in [7, 8] was to come up with a randomised 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 complex KSS system $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$.
- Draw ζ from the $\mathcal{D}$ zeros of $\mathfrak{f}$ with the uniform distribution.

For details on ρ_{std}, see [14, Section 17.5]. The description of ρ_{std} above is not constructive: it merely describes the distribution. It is remarkable, however, that it is possible to efficiently sample from ρ_{std}.

Proposition 5.1. ([14, Proposition 17.21]). *There is a randomised algorithm that, 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 this randomisation 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 ([53, 7, 8, 13, 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 $i = 1, \ldots, n$). The former depends on the input at hand, and that is where average considerations play a role. In [9, 13], ALH was implemented using the Weyl norm to compute step lengths. Its average

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

number of iterations is $O(n\mathbf{D}^{3/2}N)$. The average total 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 analyse a version ALH_∞ 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 number of iterations of ALH with step lengths based on our new condition number is $\mathcal{O}(n^3\mathbf{D}^2 \ln(n\mathbf{D}))$: a bound independent of N. Unfortunately, this gain is not decisive for a general input model due to the high cost of computing $\| \|_\infty$ norms.

Nonetheless, for the particular – but highly relevant – case of quadratic polynomials, we can efficiently compute the ∞-norm. As a result, we derive bounds that show the expected complexity of Solve_∞ is smaller than the expected complexity of Solve_W.

5.1. Description of the linear homotopy

The algorithm below is, essentially, the one in [13] and [14, Chapter 17]. The only change is in the computation of the step-length Δ_t, where we replace the original (here $\text{dist}_{\mathbb{S}}$ denotes angle)

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

by

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

This change amounts – leaving aside the difference in the 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_{norm}. Recall that N_q is the Newton operator associated to $q \in \mathcal{H}_{\boldsymbol{d}}[n]$.

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 [14]. It uses the properties of M shown in Theorem 3.5 and one more result (we know for μ_{norm}): namely, that M is a condition number in the standard sense of this expression – it measures how solutions change when data is perturbed (see Proposition 5.4 below). To simplify the notation, in the rest of this section, we will often omit the reference to the base field $\mathbb{C}$.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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*

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

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*

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

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 [14, 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 first K iterations. To simplify notation, we write q_i and ζ_i instead of q_{t_i} and ζ_{t_i}.

As in [14, Theorem 17.3], but using Proposition 3.6 in the place of [14, Proposition 16.2] and Theorem 3.5 in the place of [14, Theorem 16.1], one proves by induction the following statements for $i = 0, \ldots, K-1$:

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

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

$$\frac{4}{5}\mathsf{M}(q_i, \zeta_i) \leq \mathsf{M}(q_t, \zeta_t) \leq \frac{5}{4}\mathsf{M}(q_i, \zeta_i). \tag{5.2}$$

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

$$\frac{\|q_t\|_\infty}{\|q_i\|_\infty} \leq 1 + C = \frac{17}{16}. \tag{5.3}$$

The statement now easily follows. Consider any $i \in \{0, 1, \ldots, K-2\}$. Then

$$\begin{aligned}
\int_{t_i}^{t_{i+1}} \frac{\mathsf{M}^2(q_t, \zeta_t)}{\|q_t\|_\infty} \mathrm{d}t &\geq \frac{64}{85} \int_{t_i}^{t_{i+1}} \frac{\mathsf{M}^2(q_i, z_i)}{\|q_i\|_\infty} \mathrm{d}t = \frac{64}{85} \frac{\mathsf{M}^2(q_i, z_i)}{\|q_i\|_\infty} |t_{i+1} - t_i| && ((5.2) \text{ and } (5.3)) \\
&= \frac{64}{85} \frac{0.03}{\|f - g\|_\infty \mathbf{D}}. && (\text{choice of } \Delta t)
\end{aligned}$$

Hence

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

and the result follows. □

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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 [14, Section 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

$$\dot{\zeta}_t = \mathrm{D}_{f_t} G(\dot{f}_t) = -(\mathrm{D}_{\zeta_t} f_t)^{-1}(\dot{f}_t(\zeta_t)), \tag{5.4}$$

where the second equality is shown in the course of the proof of [14, Proposition 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

$$\begin{aligned}
\|\dot{\zeta}_t\| &= \max_{\|\dot{f}_t\|_\infty = 1} \|(\mathrm{D}_{\zeta_t} f_t)^{-1}(\dot{f}_t(\zeta_t))\| && \text{(By (5.4))}\\
&\leq \left(\max_{\|\dot{f}_t\|_\infty = 1} \|\dot{f}_t(\zeta_t)\|\right) \|(\mathrm{D}_{\zeta_t} f_t)^{-1}\| && \text{(operator norm inequality)}\\
&\leq \sqrt{n}\left(\max_{\|\dot{f}_t\|_\infty = 1} \|\dot{f}_t(\zeta_t)\|_\infty\right) \|(\mathrm{D}_{\zeta_t} f_t)^{-1}\| && \|\ \| \leq \sqrt{n}\|\ \|_\infty\\
&\leq \sqrt{n}\|(\mathrm{D}_{\zeta_t} f_t)^{-1}\| && \text{(definition of } \|\ \|_\infty)\\
&= \frac{\sqrt{n}\|f_t\|_\infty \|(\Delta^{-1}\mathrm{D}_{\zeta_t} f_t)^{-1}\|}{\|f_t\|_\infty} = \frac{\mathsf{M}(f_t, \zeta_t)}{\|f_t\|_\infty}. && \text{(definition of } \mathsf{M})
\end{aligned}$$

We recall that the norms where we have omitted subscripts are the usual norm in the case of vectors and the usual operator norm in the case of linear maps. □

5.3. *Average complexity analysis of Solve$_\infty$*

The execution of Solve$_\infty$ on an input $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$ amounts to calling ALH$_\infty$ 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 Solve$_\infty$ amounts to the number of iterations done by ALH$_\infty$. 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 Solve$_\infty$ 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 independent and identically distributed complex normal random variables of mean 0 and variance 1.

Our main result is the following.

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

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

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 [42]. To the best of our knowledge, Theorem 5.5 is the first such bound for a linear homotopy.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

We will use the following two results. The first is the complex version of Proposition 4.32 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 $\mathfrak{f} \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{C}}[n]$ be a KSS complex random polynomial tuple. Then for all $t > 0$,*

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

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

Proposition 5.8 [14, Proposition 4.27]. *Let $\mathfrak{A} \in \mathbb{C}^{n\times(n+1)}$ be a random complex matrix whose entries are independent and identically distributed complex normal Gaussian variables. Then for all $t \geq 0$,*

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

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 a 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, 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 $\mathcal{D} = d_1 \cdots d_n$.

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 from the $\mathcal{D}$ zeros of $\mathfrak{g}$, 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 expected number of iterations of $\textsc{Solve}_\infty$ with input $\mathfrak{f}$ is bounded by

$$45n\,\mathbf{D}\int_0^1 \mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}\left((\|\mathfrak{f}\|_\infty^2 + \|\mathfrak{g}\|_\infty^2)^2 \|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{-1}\Delta\|^2\right) \mathrm{d}t, \tag{5.5}$$

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

$$\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}\left((\|\mathfrak{f}\|_\infty^2 + \|\mathfrak{g}\|_\infty^2)^2 \|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{-1}\Delta\|^2\right) \leq \left(\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}(\|\mathfrak{f}\|_\infty^2 + \|\mathfrak{g}\|_\infty^2)^6\right)^{\frac{1}{3}} \left(\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})} \|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{-1}\Delta\|^3\right)^{\frac{2}{3}}. \tag{5.6}$$

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

By Proposition 5.7, we have that

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

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

$$\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}\|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{-1}\Delta\|^3=\left(t^2+(1-t)^2\right)^{\frac{3}{2}}\mathop{\mathbb{E}}_{(\mathfrak{h},\mathfrak{y})\sim\rho_{\mathrm{std}}}\|\mathrm{D}_{\mathfrak{y}}\mathfrak{h}^{-1}\Delta\|^3. \tag{5.7}$$

Now, since $(\mathfrak{h},\mathfrak{y})$ is a random standard pair, 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. This is the so-called Beltrán-Pardo trick [14, Proposition 17.21(a)]. 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,

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

Hence, integrating equation (5.7),

$$\left(\int_0^1\mathop{\mathbb{E}}_{(\mathfrak{f},\mathfrak{g},\mathfrak{z})}\|\mathrm{D}_{\mathfrak{z}_t}\mathfrak{q}_t^{-1}\Delta\|^3\,\mathrm{d}t\right)^{\frac{2}{3}}=\mathcal{O}(n\,\mathbf{D}). \tag{5.8}$$

Putting together equations (5.5), (5.6) and (5.8), 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 remains smaller than the one for $\textsc{Solve}_W$, which is $\mathcal{O}(n^7)$. The major task left, unsurprisingly, is to compute $\|q\|_\infty^{\mathbb{C}}$ in equation (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 $q\in\mathcal{H}_2[n]$ 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 $f\in\mathcal{H}_2[n]$ within $\mathcal{O}(n^{4.5+\omega})$ time on average, where ω is the exponent for the cost of matrix multiplication. We currently have $\omega<2.375$.*

Proposition 5.10. *Let $q\in\mathcal{H}_2[n]$ be a quadratic system such that for each i, $q_i=X^TA_iX$. Then*

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

where the norm $\| \ \|$ *in the middle formula is the usual operator norm. Moreover, the number* $\sqrt{\left\|\sum_{i=1}^n A_i^* A_i\right\|}$ *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

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

by the smaller

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

where $q = (X^T A_i X)_i$. 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 $\sqrt{\left\|\sum_{i=1}^n A_i^* A_i\right\|}$ in $\mathcal{O}(n^{1+\omega})$ operations, and by [14, 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 factorisation [39, Problem 33], we have that

$$A_i = U_i^T D_i U_i$$

for some real diagonal matrix D_i with nonnegative entries and some unitary matrix U_i. Now it is easy to check that

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

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

For the other inequality, observe that

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

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

Regarding cost, 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 $\left\|\sum_{i=1}^n A_i^* A_i\right\|$ another $\mathcal{O}(n^3)$ operations. We thus get $\mathcal{O}(n^{1+\omega})$ operations in total, as we wanted to show. □

Acknowledgements. 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. We are thankful to the reviewers of this paper for useful suggestions that helped improve the presentation and to Khazhgali Kozhasov for pointing out an error in a constant used in Proposition 4.24.

Conflict of Interest. The authors have no conflict of interest to declare.

Financial support. This work was supported by the Einstein Foundation Berlin. The first author was partially supported by GRF grant CityU 11300220. The second author was partially supported by NSF CCF 211 00 75. The last author was 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.

A. Extension to spaces of C^1-maps

In this appendix, we 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 to $\| \|_\infty$. 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, 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-map.

Example A.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

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

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 (compare to [39, 3.1.3],

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

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

A.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,

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

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.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Theorem A.2 (Condition number theorem). *Let $f \in C^1[q]$ and $x \in \mathbb{S}^n$. Then*

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

and

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

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

We call this result the 'condition number theorem' as it is so for the following condition numbers for C^1-maps:

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

and

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

These condition numbers are very similar to K, and one might try (but we won't here) to prove an analogue 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

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

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

Proof of Theorem A.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.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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

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

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

$$f - g = f(e_0)X_0 + s_qu_qX_q.$$

By arguing as in Example 2.5 and noting that $f(e_0)X_0 + s_qu_qX_q$ has rank at most 2, we have that

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

Hence

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

finishing the proof of the first equality.

The second equality follows immediately from the first one. □

A.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 $\| \ \|_{\overline{\infty}}$

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

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

Example A.3. Let

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

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 $x \in \mathbb{S}^n$,

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

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

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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 A.4 (Structured condition number theorem). *Let $f \in C^1[q]$, $x \in \mathbb{S}^n$ and $\boldsymbol{d} \in \mathbb{N}^q$. Then*

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

and

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

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

Corollary A.5. *Let $\boldsymbol{d} \in \mathbb{N}^d$, $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $x \in \mathbb{S}^n$. Then*

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

and

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

where $\operatorname{dist}_{\overline{\infty},\boldsymbol{d}}$ *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 $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$. 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 in more detail in the next subsection.

Proof of Theorem A.4. This proof is almost the same as the one of Theorem A.2. We only have to modify the part where we find an explicit minimiser 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 $s_1, \dots, s_q > 0$, 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

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

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

$$f - g = \Big(f_i(e_0)X_0^{d_i} + \sqrt{d_i}s_q u_q X_q\Big)_i.$$

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

Because of Example A.3, for

$$h = \left(a_i X_0^{d_i} + \sqrt{d_i} b X_0^{d_i - 1} X_1\right)_i \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q],$$

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

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

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

Proof of Corollary A.5. This is Theorem A.4 together with [15, Theorem 4.4]. □

A.3. Relationship between norms

As it happens with K and κ (see Section 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 $\mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$, we compare $\| \|_{\infty}^{\mathbb{R}}$, $\| \|_W$, $\| \|_{\overline{\infty}}$ and $\| \|_{\overline{\infty},\boldsymbol{d}}$.

Proposition A.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 A.7. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$. Then the following inequalities hold:*

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

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

$$\|f\|_{\infty}^{\mathbb{R}} \leq \|f\|_{\overline{\infty},\boldsymbol{d}} \leq \|f\|_W \tag{A.3}$$

Proof of Proposition A.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 the inequalities are derived from this claim in a 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 $A \in \mathbb{R}^{q \times n}$,

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

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

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

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

Combining these bounds, 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 A.7. Arguing as in Proposition A.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 equations (A.1) and (A.2).

It only remains to prove $\|f\|_{\infty,\boldsymbol{d}} \leq \|f\|_W$ in equation (A.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 [30, Proposition 3] and [12, 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 A.8. *Let $f \in \mathcal{H}_{\boldsymbol{d}}^{\mathbb{R}}[q]$ and $x \in \mathbb{S}^n$. Then*

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

and

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

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

Sketch of proof. The proof is similar to that of Proposition A.6. Arguing as there, we can prove that for all $x \in \mathbb{S}^n$,

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

Minimizing over $x \in \mathbb{S}^n$ and applying Theorems A.2 and Corollary A.5, we conclude. □

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

References

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

[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*, (2013), 675–684. ACM, New York.

[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** (2016), 1401–1422.

[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) (2005), 125–144.

[5] A. Barvinok. 'Integration and optimization of multivariate polynomials by restriction onto a random subspace', *Found. Comput. Math.*, **7**(2) (2007), 229–244.

[6] Alexander Barvinok. 'Estimating L_∞norms by L_{2k}norms for functions on orbits', *Foundations of Computational Mathematics*, **2**(4) (2002), 393–412.

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

[8] 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) (2009), 363–385.

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

[10] 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*, volume 81 of *LIPIcs. Leibniz Int. Proc. Inform.* Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2017.

[11] 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) (1998), 1257–1272.

[12] P. Breiding, H. Keneshlou and A. Lerario. 'Quantitative singularity theory for random polynomials', *International Mathematics Research Notices*, **10** (2020).

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

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

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

[16] P. Bürgisser, F. Cucker and J. Tonelli-Cueto. 'Computing the homology of semialgebraic sets. I: Lax formulas', *Found. Comput. Math.*, **20** (2020), 71–118. doi:10.1007/s10208-019-09418-y.

[17] P. Bürgisser, F. Cucker and J. Tonelli-Cueto. 'Computing the homology of semialgebraic sets. II: General formulas', *Found. Comput. Math.*, **21**(5) (2021), 1279–1316. doi:10.1007/s10208-020-09483-8.

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

[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*, (2017), 61–68. ACM, New York.

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

[21] V. Chandrasekaran and P. Shah. 'Relative entropy relaxations for signomial optimization', *SIAM J. Optim.*, **26**(2) (2016), 1147–1173.

[22] D. Cheung and F. Cucker. 'Componentwise condition numbers of random sparse matrices', *SIAM J. Matrix Anal. Appl.*, **31** (2009), 721–731.

[23] F. Cucker, A. A. Ergür and J. Tonelli-Cueto. 'Plantinga-Vegter algorithm takes average polynomial time', In *Proceedings of the 2019 International Symposium on Symbolic and Algebraic Computation*, (2019), 114–121. ACM.

[24] F. Cucker, A. A. Ergür and J. Tonelli-Cueto. 'On the complexity of the Plantinga-Vegter algorithm', *Discrete Comput. Geom.*, **68** (2022), 664–708. doi:10.1007/s00454-022-00403-x.

[25] F. Cucker, T. Krick, G. Malajovich and M. Wschebor. 'A numerical algorithm for zero counting. I: Complexity and accuracy', *J. Complexity*, **24**(5–6) (2008), 582–605.

[26] 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** (2009), 285–294.

[27] F. Cucker, T. Krick, G. Malajovich and M. Wschebor. 'A numerical algorithm for zero counting. III: Randomization and condition', *Adv. Applied Math.*, **48** (2012), 215–248.

[28] F. Cucker, T. Krick and M. Shub. 'Computing the homology of real projective sets', *Found. Comput. Math.*, **18** (2018), 929–970.

[29] 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.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press

[30] D. N. Diatta and A. Lerario. 'Low degree approximation of random polynomials', *Found. Comput. Math.*, **22** (2022), 77–97. doi:10.1007/s10208-021-09506-y.

[31] M. Dressler, S. Iliman and T. de Wolff. 'A Positivstellensatz for sums of nonnegative circuit polynomials', *SIAM J. Appl. Algebra Geom.*, **1**(1) (2017), 536–555.

[32] H. Edelsbrunner. 'The union of balls and its dual shape', *Discrete Comput. Geom.*, **13**(3–4) (1995), 415–440.

[33] A. A. Ergür, G. Paouris and J. M. Rojas. 'Probabilistic condition number estimates for real polynomial systems I: A broader family of distributions', *Found. Comput. Math.*, **19** (2019), 131–157.

[34] A. A. Ergür, G. Paouris and J. M. Rojas. 'Smoothed analysis for the condition number of structured real polynomial systems', *Math. Comp.*, **90**(331) (2021), 2161–2184.

[35] H. Federer. 'Curvature measures', *Trans. Amer. Math. Soc.*, **93** (1959), 418–491.

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

[37] B. T. Galehouse. 'Topologically Accurate Meshing Using Domain Subdivision Techniques', ProQuest LLC, Ann Arbor, MI, 2009. Ph.D. thesis, New York University.

[38] D. Gross, S. T. Flammia and J. Eisert. 'Most quantum states are too entangled to be useful as computational resources', *Phys. Rev. Lett.*, **102**(19) (2009), 190501.

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

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

[41] Kh. Kozhasov and J. Tonelli-Cueto. 'Bounds for the average rank-one approximation ratio', 2021. In progress.

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

[43] M. Laurent. 'Sums of squares, moment matrices and optimization over polynomials', In *Emerging Applications of Algebraic Geometry*, volume **149** of *IMA Vol. Math. Appl.*, (2009), 157–270. Springer, New York.

[44] G. Livshyts, G. Paouris and P. Pivovarov. 'On sharp bounds for marginal densities of product measures', *Israel Journal of Mathematics*, **216**(2) (2016), 877–889.

[45] S. Mendelson and G. Paouris. Stable recovery and the coordinate small-ball behaviour of random vectors, Preprint, 2019. arXiv:1904.08532.

[46] N. H. Nguyen, P. Drineas and T. D. Tran. 'Tensor sparsification via a bound on the spectral norm of random tensors', *Inf. Inference*, **4**(3) (2015), 195–229.

[47] P. Niyogi, S. Smale and S. Weinberger. 'Finding the homology of submanifolds with high confidence from random samples', *Discrete Comput. Geom.*, **39**(1–3) (2008), 419–441.

[48] E. Novak, I. H. Sloan, J. F. Traub and H. Woźniakowski. *Essays on the Complexity of Continuous Problems*. European Mathematical Society (EMS), Zürich, 2009.

[49] S. Plantinga and G. Vegter. 'Isotopic approximation of implicit curves and surfaces', In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*, (2004), 245–254. ACM.

[50] L. Plaskota. *Noisy Information and Computational Complexity*. Cambridge University Press, Cambridge, 1996.

[51] M. Rudelson and R. Vershynin. 'Small ball probabilities for linear images of high-dimensional distributions', *Int. Math. Res. Not. IMRN*, **19** (2015), 9594–9617.

[52] R. T. Seeley. 'Spherical harmonics', *Amer. Math. Monthly*, **73**(4, part II) (1966), 115–121.

[53] M. Shub and S. Smale. 'Complexity of Bézout's theorem. I. Geometric aspects', *J. Amer. Math. Soc.*, **6**(2) (1993), 459–501.

[54] S. Smale. 'Mathematical problems for the next century', In V. Arnold, M. Atiyah, P. Lax and B. Mazur, editors, *Mathematics: Frontiers and Perspectives*, (2000), 271–294. AMS.

[55] J. Tonelli-Cueto. *Condition and Homology in Semialgebraic Geometry*. Doctoral thesis, Technische Universität Berlin, 2019.

[56] J. Tonelli-Cueto and E. P. Tsigaridas. 'Condition numbers for the cube. i: Univariate polynomials and hypersurfaces', In *Proceedings of the 45th International Symposium on Symbolic and Algebraic Computation*, (2020), 434–441. ACM.

[57] 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.

[58] 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.

[59] 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) (2012), 806–821.

https://doi.org/10.1017/fms.2022.89 Published online by Cambridge University Press