Chin. Ann. Math. Ser. B
39(2), 2018, 373–396
DOI: 10.1007/s11401-018-1070-8

Chinese Annals of Mathematics, Series B

Grid Methods in Computational Real Algebraic (and Semialgebraic) Geometry*

Felipe CUCKER[1]

(*À Philippe, pour des années d'amitié*)

Abstract In recent years, a family of numerical algorithms to solve problems in real algebraic and semialgebraic geometry has been slowly growing. Unlike their counterparts in symbolic computation they are numerically stable. But their complexity analysis, based on the condition of the data, is radically different from the usual complexity analysis in symbolic computation as these numerical algorithms may run forever on a thin set of ill-posed inputs.

Keywords Numerical algorithms, Complexity, Condition, Semialgebraic geometry
2000 MR Subject Classification 65H10, 14Q20

1 Putting in Context

The primary objects of study of real algebraic geometry are real algebraic sets. These are subsets of $\mathbb{R}^n$ defined as zero sets of finite sets of polynomials. A naturally occurring class of sets, in the context of this study, is that of semialgebraic sets. These are subsets of $\mathbb{R}^n$ defined as Boolean combinations (i.e., unions and intersections) of polynomial equalities and inequalities. Two classic books on real algebraic and semialgebraic geometry are [6, 10].

The fact that polynomials are easy to describe (by, for instance, their degree and the list of their coefficients), together with their ubiquity in computational problems, motivated the blossoming of a subject, within symbolic computation (a.k.a. computer algebra), devoted to problems involving semialgebraic sets. A comprehensive monograph on this subject is [5]; a recent survey is [3].

Central to this subject was the issue of computational complexity. Assume that the problem at hand takes as input a collection of s polynomials in n variables, of degrees $d_1, \cdots, d_s$, respectively. Then the size of this input is the number of coefficients used to describe this collection. That is, this size is

$$N := \sum_{i=1}^{s} \binom{n+d_i}{n}.$$

This quantity is in general exponentially large in n but it does not need to be so. For instance, in the case of systems of quadratic polynomials, $N = \mathcal{O}(sn^2)$.

Manuscript received June 28, 2017. Revised November 3, 2017.
[1]Department of Mathematics, City University of Hong Kong, Hong Kong, China.
E-mail: macucker@cityu.edu.hk
*This work was supported by a GRF grant from the Research Grants Council of the Hong Kong SAR (No. CityU11310716).

Complexity considerations were not always at the forefront in the algorithmics of semialgebraic geometry. One of the first remarkable algorithms in the subject was created by Alfred Tarski in the late 1930's (see [45][1]). This algorithm takes as input an expression of the form

$$Q_1 x_1 Q_2 x_2 \cdots Q_n x_n \, \psi(P_1, \cdots, P_s), \tag{1.1}$$

where $P_1, \cdots, P_s$ are polynomials in $\mathbb{R}[x_1, \cdots, x_n]$, $\psi(P_1, \cdots, P_s)$ is any Boolean combination of these polynomials, and $Q_j x_j$ is a quantification either existential — $\exists x_j$ — or universal —$\forall x_j$. The fact that all variables are quantified implies that such an expression is either true or false. Tarski's algorithm thus returns the truth value of (1.1). The focus at that time was on (theoretical) computability and Tarski's result showed that a number of problems (which ultimately can be reduced to deciding the truth of statements as (1.1)) could be, in theory, solved. In practice the situation was bleaker. The complexity of Tarski's algorithm is of the order of

$$\left. 2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}} \right\} n \text{ times.}$$

In the seventies, Collin [16] and Wüthrich [47] independently devised a method today referred to as Cylindrical Algebraic Decomposition (CAD for short). We will not describe what exactly CAD does; suffice it to say that a CAD of the collection $\{P_1, \cdots, P_s\}$ allows one to decide the truth of any sentence of the form (1.1), no matter the prefix of quantifiers. The complexity of CAD, which is doubly exponential in n,

$$(sD)^{2^{\mathcal{O}(n)}},$$

is far from practical but nonetheless much better than Tarski's. Here $D = \max\{d_1, \cdots, d_s\}$.

A breakthrough was made in the late 1980's with the introduction of the critical points method by Grigoriev and Vorobjov [27–28], and its improvements in [4, 34–36] among others. This method allows one to eliminate quantifiers at a cost which is singly exponential in the number of variables n and doubly exponential only in the number ℓ of quantifier alternations in the prefix of (1.1). That is, it is of the order

$$(sD)^{n^{\mathcal{O}(\ell)}}.$$

The critical points method allowed one to provide single exponential time algorithms for a number of basic problems in semialgebraic geometry. Indeed, let S be a semialgebraic sets defined via a collection $\{P_1, \cdots, P_s\}$ as above. Then the problems (to mention just a few)

Feasibility, i.e., decide whether S is nonempty;

Dimension, i.e., compute the dimension of S;

Counting, i.e., compute the number of points in S (if finite, otherwise return ∞);

Euler, i.e., compute the Euler characteristic of S

could all be solved within time $(sD)^{n^{\mathcal{O}(1)}}$. Among those solvable with CAD but for which no algorithm has been devised that would work within a time bound single exponential in n stands out the following:

[1]The publication of Tarski's results was delayed by the war.

Homology, i.e., compute the homology groups of S.

At this point we note that these bounds may be polynomial in the input size N but may be exponentially large as well, as for the case of families of quadratics polynomials.

In 1989, Blum, Shub and Smale wrote a paper [9] adding a new twist to the complexity considerations above. They developed a theory of complexity over the reals along the lines of the (at that time blooming) discrete theory of complexity. In particular, they defined the class $\mathrm{NP}_{\mathbb{R}}$, the real analog of the acclaimed class NP (see [15]), and they showed that Feasibility is $\mathrm{NP}_{\mathbb{R}}$-complete. Shortly said, this means that Feasibility is universal in the class $\mathrm{NP}_{\mathbb{R}}$ in the sense that whichever algorithmic improvement (viz complexity) can be found for Feasibility the same improvement will hold for all problems in this class. Interestingly, this remains true even when we restrict all polynomials in Feasibility to be quadratic. This foundational paper gave rise to similar completeness results. In particular, Koiran [29] showed that Dimension is $\mathrm{NP}_{\mathbb{R}}$-complete as well, and in [11] the completeness in $\#P_{\mathbb{R}}$ (the real analog of the classical $\#P$) was shown for both Counting and Euler. Other complexity classes over the reals and a list of complete problems for them, all involving real polynomials, was shown in [12].

Towards the end of [9], in a list of open problems, Blum, Shub and Smale mention a theme that was going to occupy a central position in Shub and Smale research (as witnessed by the Bézout series [39–43] and by Part II in [8]). Quoting from Section 5, it would be useful to incorporate round-off error, condition numbers and approximate solutions into our development.

The rationale behind this quote is the assumption of infinite precision in the machine model proposed in [9]: These machines are theoretically capable of store, and operate with, arbitrary real numbers. In the practice of numerical analysis, in contrast, real numbers are replaced by floating-point numbers thus giving rise to round-off errors. These errors accumulate during the computation and the goal of stability analysis is to gauge the quality of an algorithm regarding this accumulation. Blum, Shub and Smale are thus asking to incorporate stability in the complexity theory they developed.

Central to any possible extension of this theory (see [18] for one such extension) there is the notion of condition. A condition number, associated to an input a for a problem $\mathscr{P}$, is, loosely speaking, a measure of how much the outcome of $\mathscr{P}$ is affected by small perturbations of a. It is independent of the choice of an algorithm for solving $\mathscr{P}$ and thus emphasizes the role of a in the stability analysis. Condition numbers usually take values on the interval $[1, \infty]$ and inputs a for which the condition is ∞ are said to be ill-posed. Those are the inputs for which, no matter the algorithm used, no matter how fine its precision, there is no hope of a meaningful, reliable output.

Interestingly, condition numbers play a major role in complexity analysis as well (even assuming infinite precision). Indeed, it is common in numerical analysis to design iterative algorithms. These are procedures that, unlike those mentioned above (CAD, critical points method, $\cdots$), do not have a complexity bound purely in terms of the input size but may take arbitrarily long times even when all input's parameters (e.g., n, s and D above) are fixed. Instead, it is common, at least within certain subjects, that the number of iterations that these algorithms perform is bounded by a function on these parameters and the condition number of the input. In particular, they may loop forever when the input is ill-posed.

Iterative algorithms and their condition-based analysis are the bread and butter of numeri-

cal linear algebra (see [25, 46]). They made a burst in optimization with the pioneering work of Renegar [37–38] and eventually found their way in semialgebraic geometry. An early representative of this work is [23], where an iterative algorithm for Feasibility is proposed and studied.

The motivation behind [23] is the observation that the critical points method performs few operations with large matrices. This is not only costly in time but also unlikely to be numerically stable. In contrast, an algorithm performing a large number of (independent) operations on small matrices, while equally costly in time, should at least be numerically stable.

The ideas in [23] were later extended in [19] to the Counting problem, in [22] to the Homology problem for real projective sets and in [14] to the same problem but for semialgebraic sets. In the next pages we will describe these ideas and the analyses of the resulting algorithms.

2 Setting the Tools

2.1 Spaces of polynomials

Let $\boldsymbol{d} = (d_1, \cdots, d_m) \in \mathbb{N}^m$ be such that $d_i \geq 1$ for all $i \leq m$. To such a tuple we associate the linear space $\mathcal{H}_{\boldsymbol{d}}[m]$ of homogeneous polynomial systems $f = (f_1, \cdots, f_m)$ in $\mathbb{R}[X_0, \cdots, X_n]$ with degree $f_i = d_i$. We make $\mathcal{H}_{\boldsymbol{d}}[m]$ an inner product space by endowing it with the Weyl inner product. For g, h homogeneous of degree d in $\mathbb{R}[X_0, \cdots, X_n]$ we define

$$\langle g, h\rangle = \sum_{|\boldsymbol{a}|=d} g_{\boldsymbol{a}} h_{\boldsymbol{a}} \binom{d}{\boldsymbol{a}}^{-1},$$

where the $\boldsymbol{a} = (\alpha_0, \cdots, \alpha_n)$ are multiindexes satisfying $\alpha_0 + \cdots + \alpha_n = d$, $g = \sum_{|\boldsymbol{a}|=d} g_{\boldsymbol{a}} X^{\boldsymbol{a}}$ (and similarly for h) and

$$\binom{d}{\boldsymbol{a}} = \frac{d!}{\alpha_0! \cdot \cdots \cdot \alpha_n!}$$

is the multinomial coefficient. This definition naturally extends to an inner product in $\mathcal{H}_{\boldsymbol{d}}[m]$ which is no more than the dot product for a weighted monomial basis. The main reason to consider it is that it is invariant under the action of the orthogonal group $\mathcal{O}(n+1)$. That is, for all $f, g \in \mathcal{H}_{\boldsymbol{d}}[m]$ and all $u \in \mathcal{O}(n+1)$, $\langle f, g\rangle = \langle f \circ u, g \circ u\rangle$. This invariance allows for great simplifications in many arguments. See e.g. [13, §16.1] for details. In all what follows, for $f \in \mathcal{H}_{\boldsymbol{d}}[m]$, the expression $\|f\|$ denotes the norm of f induced by the Weyl inner product.

2.2 The condition of a polynomial system at a zero

How much does a zero $\xi \in \mathbb{R}^{n+1}$ of a system $f \in \mathcal{H}_{\boldsymbol{d}}[m]$ move when we slightly perturb the coefficients of f? An answer to this question, the condition $\mu_{\text{norm}}(f, \xi)$ of f at ξ, was provided by Shub and Smale in their "Bézout series" of papers [39–43].

Let

$$\mathrm{D}f(\xi) = \Big(\frac{\partial f_i}{\partial x_j}(\xi)\Big)_{1 \leq i \leq m, 0 \leq j \leq n} : \mathbb{R}^{n+1} \to \mathbb{R}^m$$

be the derivative of f at ξ and

$$\Delta(\xi) := \begin{bmatrix} \|\xi\|^{d_1-1}\sqrt{d_1} & & \\ & \ddots & \\ & & \|\xi\|^{d_m-1}\sqrt{d_m} \end{bmatrix}$$

(we write simply Δ if ξ belongs to the unit sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$). Shub and Smale defined

$$\mu_{\text{norm}}(f,\xi) := \|f\|\|\mathrm{D}f(\xi)^\dagger \Delta(\xi)\|, \tag{2.1}$$

when $\mathrm{D}f(\xi)$ is surjective and $\mu_{\text{norm}}(f,\xi) := \infty$ otherwise. Here $\mathrm{D}f(\xi)^\dagger : \mathbb{R}^m \to \mathbb{R}^{n+1}$ is the Moore-Penrose inverse of the full-rank matrix $\mathrm{D}f(\xi)$, i.e.,

$$\mathrm{D}f(\xi)^\dagger = \mathrm{D}f(\xi)^{\mathrm{t}}(\mathrm{D}f(\xi)\,\mathrm{D}f(\xi)^{\mathrm{t}})^{-1},$$

where $\mathrm{D}f(\xi)^{\mathrm{t}}$ is the transpose of $\mathrm{D}f(\xi)$. This coincides with the inverse of the restricted linear map $\mathrm{D}f(\xi)|_{(\ker \mathrm{D}f(\xi))^\perp}$. Also, the norm in $\|\mathrm{D}f(\xi)^\dagger \Delta(\xi)\|$ is the spectral norm.

We note that the expression on the right of (2.1) is well-defined for arbitrary points $x \in \mathbb{S}^n$, so we can define $\mu_{\text{norm}}(f,x)$ for any such point. This property will allow us to define, for each of the problems in Section 3, a condition number depending only on the considered problem and the data at hand which uses μ_{norm} as a basic ingredient.

Remark 2.1 For any $\lambda \neq 0$ we have $\mu_{\text{norm}}(f,x) = \mu_{\text{norm}}(f,\lambda x)$, since when $\mathrm{D}f(x)$ is surjective, $\mathrm{D}f(\lambda x)^\dagger = (\Lambda \mathrm{D}f(x))^\dagger = \mathrm{D}f(x)^\dagger \Lambda^{-1}$ for

$$\Lambda = \begin{bmatrix} \lambda^{d_1-1} & & \\ & \ddots & \\ & & \lambda^{d_m-1} \end{bmatrix}.$$

Similarly, $\mu_{\text{norm}}(f,\xi) = \mu_{\text{norm}}(\lambda f,\xi)$ for all $\lambda \neq 0$.

Finally, the following result (see [13, Proposition 19.30]) allows one to estimate $\mu_{\text{norm}}(f,y)$ by evaluating $\mu_{\text{norm}}(f,x)$ at a point x sufficiently close to y.

Proposition 2.1 *There exist constants $C, \overline{\varepsilon} > 0$ such that the following is true. For all $\varepsilon \in [0,\overline{\varepsilon}]$, all $f \in \mathcal{H}_{\boldsymbol{d}}[m]$, and all $x, y \in \mathbb{S}^n$, if $D^{\frac{3}{2}} \mu_{\text{norm}}(f,y) d_{\mathbb{S}}(x,y) \leq C\varepsilon$, then*

$$\frac{1}{1+\varepsilon}\mu_{\text{norm}}(f,x) \leq \mu_{\text{norm}}(f,y) \leq (1+\varepsilon)\mu_{\text{norm}}(f,x).$$

2.3 Probabilities

Recall the class of numerical algorithms we mentioned in the introduction. These are iterative algorithms whose number of iterations are analyzed in terms of the condition of the input (as well as its size). Because this condition cannot in general be estimated a priori (it is common that computing the condition of an input is as difficult as solving the problem for which this input is so) a way of gauging the efficiency of algorithms of this kind is via probabilistic estimates. To this end, the space of data is endowed with a probability measure $\mathscr{M}$. Then, $\mathrm{cost}(a)$, the cost of an algorithm at a data a, becomes a random variable and, when possible one attempts to provide bounds for the expectation

$$\mathop{\mathbb{E}}_{x \sim \mathscr{M}} \mathrm{cost}(x).$$

This average case complexity provides, more often than not, a realistic measure of an algorithm's efficiency. But it is sometimes overly pessimistic. Indeed, there are examples of algorithms which are known to be efficient in practice but nonetheless have a large, or even infinite,

average-case complexity. In a recent paper Amelunxen and Lotz [2] identified as a possible cause for this discrepancy the existence of an exceptional set of inputs on which the algorithm runs in superpolynomial time but having a measure that is vanishing exponentially fast when the dimension grows. A prototype of this phenomenon is the behavior of the power method to compute a dominant eigenpair of a symmetric matrix. This algorithm is considered efficient in practice, yet it has been shown that the expectation of the number of iterations performed by the power method, for matrices drawn from the Gaussian Orthogonal Ensemble (see [32]), is infinite [30]. Amelunxen and Lotz [2] showed that, conditioned to exclude a set of measure at most 2^{-n}, this expectation is $\mathcal{O}(n^2)$ for $n \times n$ matrices. The moral of the story is that the power method is efficient in practice because it is so in theory if we disregard a vanishingly small set of outliers. This conditional expectation, in the terminology of [2], shows a weak average polynomial cost for the power method. More generally, we will talk about a complexity bound being weak when this bound holds out of a set of exponentially small measure.

Our probabilistic analyses in Section 3 rely on a fact observed by Demmel [24] namely, that condition numbers are often bounded by the normalized inverse of the distance to the nearest ill-posed problem. For quantities defined in this way, usually referred to as conic condition numbers, the following result (see [13, Theorem 21.1]) is useful. We rephrased the statement in terms of the isotropic Gaussian distribution instead of the uniform distribution on the sphere. The scale invariance of the statement makes both formulations equivalent.

Theorem 2.1 *Let $\Sigma \subseteq \mathbb{R}^{p+1}$ be contained in a real algebraic hypersurface, given as the zero set of a homogeneous polynomial of degree d and let $a \in \mathbb{R}^{p+1}$ be a centered isotropic Gaussian random variable. Then for all $t \geq (2d+1)p$,*

$$\operatorname{Prob}\Big(\frac{\|a\|}{d(a,\Sigma)} \geq t\Big) \leq \frac{11dp}{t}.$$

2.4 Numerical stability for discrete-valued problems

Recall, the machine precision of a finite-precision algorithm is a number $\varepsilon_{\text{mach}} \in (0,1)$ such that the result x of any operation performed by the algorithm is replaced by another number $\widetilde{x}$ satisfying $\widetilde{x} = x(1+\delta)$ for some δ with $|\delta| \leq \varepsilon_{\text{mach}}$. The quantity $\log_2 \varepsilon_{\text{mach}}$, basically, indicates the number of bits in the mantissa of a floating-point representation of real numbers in the execution of $\mathscr{A}$.

Algorithms computing a continuous function of their input have a standard finite-precision analysis estimating how much may the error of the output be for a given $\varepsilon_{\text{mach}}$. But for problems having only a discrete set of values another form of analysis is necessary. The template result below is the most common such form.

Template Result Let $\mathscr{A}$ denote a (fixed-precision) algorithm that, with infinite precision, computes a function $\varphi : \mathcal{D} \to \mathbb{R}^q$. Let $\mathscr{A}^\varphi$ denote the function computed by $\mathscr{A}$ with finite precision. Finally, let $\operatorname{cond}^\varphi(a)$ denote the condition number of an input $a \in \mathcal{D}$. If

$$\varepsilon_{\text{mach}} \leq \frac{1}{g(\operatorname{size}(a), \operatorname{cond}^\varphi(a))},$$

then $\mathscr{A}^\varphi(a) = \varphi(a)$. Here $\operatorname{size}(a)$ denotes the size (given by one or more integers) of data a and g is a function of this size and the condition of a.

2.5 Newton's method and point estimates

Let $f : \mathbb{R}^{n+1} \to \mathbb{R}^m$, $m \leq n+1$, be analytic. The Moore-Penrose Newton operator of f at $x \in \mathbb{R}^{n+1}$ is defined (see [1]) as

$$N_f(x) := x - \mathrm{D}f(x)^\dagger f(x).$$

We say that it is well-defined if $\mathrm{D}f(x)$ is surjective.

Definition 2.1 *Let $x \in \mathbb{R}^{n+1}$. We say that x is an approximate zero of f if the sequence $(x_k)_{k\geq 0}$ defined as $x_0 := x$ and $x_{k+1} := N_f(x_k)$ for $k \geq 0$ is well-defined and there exists a zero ζ of f such that, for all $i \geq 0$,*

$$\|x_i - \zeta\| \leq \left(\frac{1}{2}\right)^{2^i-1} \|x - \zeta\|.$$

The point ζ is said to be the associated zero of x.

Following ideas introduced by Steve Smale [44], the following three quantities are associated to a point $x \in \mathbb{R}^{n+1}$ (see [43]):

$$\begin{aligned}
\beta(f,x) &:= \|\mathrm{D}f(x)^\dagger f(x)\|,\\
\gamma(f,x) &:= \max_{k>1} \left\|\mathrm{D}f(x)^\dagger \frac{\mathrm{D}^k f(x)}{k!}\right\|^{\frac{1}{k-1}},\\
\alpha(f,x) &:= \beta(f,x)\gamma(f,x),
\end{aligned}$$

when $\mathrm{D}f(x)$ is surjective, and $\alpha(f,x) = \beta(f,x) = \gamma(f,x) = \infty$ when $\mathrm{D}f(x)$ is not surjective. The quantity $\beta(f,x) = \|N_f(x) - x\|$ measures the length of the Newton step at x. The value of $\gamma(f,\xi)$, at a zero ξ of f, is related to the radius of the neighborhood of points that converge to the zero ξ of f, and the meaning of $\alpha(f,x)$ is made clear in Theorem 2.2 below.

In this survey we are interested in the theory of point estimates for polynomial maps $f = (f_1, \cdots, f_m)$. When the f_i are homogeneous, the invariants α, β and γ are themselves homogeneous in x. We actually have $\beta(f,\lambda x) = \lambda\beta(f,x)$, $\gamma(f,\lambda x) = \lambda^{-1}\gamma(f,x)$ and $\alpha(f,\lambda x) = \alpha(f,x)$ for all $\lambda \neq 0$. This property motivates the following projective version for them:

$$\begin{aligned}
\beta_{\mathrm{proj}}(f,x) &:= \|x\|^{-1}\|\mathrm{D}f(x)^\dagger f(x)\|,\\
\gamma_{\mathrm{proj}}(f,x) &:= \|x\| \max_{k>1} \left\|\mathrm{D}f(x)^\dagger \frac{\mathrm{D}^k f(x)}{k!}\right\|^{\frac{1}{k-1}},\\
\alpha_{\mathrm{proj}}(f,x) &:= \beta_{\mathrm{proj}}(f,x)\gamma_{\mathrm{proj}}(f,x).
\end{aligned}$$

These projective versions coincide with the previous expressions when $x \in \mathbb{S}^n$ and an α-Theorem for them is easily derived from Theorem 2.2 below. Furthermore, β_{proj} still measures the (scaled) length of the Newton step, and γ_{proj} relates to the condition number via the following bound (known as the Higher Derivative Estimate),

$$\gamma_{\mathrm{proj}}(f,x) \leq \frac{1}{2} D^{\frac{3}{2}} \mu_{\mathrm{norm}}(f,x). \tag{2.2}$$

The proof of this bound is the one of [8, Theorem 2, p. 267] which still holds for $m \leq n$ and $\mathrm{D}f(x)^\dagger$ instead of $\mathrm{D}f(x)|_{T_x}^{-1}$.

The right-hand side in inequality (2.2) is easily computable. We can find similar computable versions for α and β. Indeed, for $x \in \mathbb{S}^n$ we define

$$\begin{aligned}\overline{\beta}(f,x) &:= \mu_{\mathrm{norm}}(f,x)\frac{\|f(x)\|}{\|f\|},\\ \overline{\gamma}(f,x) &:= \frac{1}{2}D^{\frac{3}{2}}\mu_{\mathrm{norm}}(f,x),\\ \overline{\alpha}(f,x) &:= \overline{\beta}(f,x)\overline{\gamma}(f,x) = \frac{1}{2}D^{\frac{3}{2}}\mu^2_{\mathrm{norm}}(f,x)\frac{\|f(x)\|}{\|f\|}.\end{aligned} \tag{2.3}$$

Inequality (2.2) says that $\gamma(f,x) \le \overline{\gamma}(f,x)$. We also observe that $\beta(f,x) \le \overline{\beta}(f,x)$ since

$$\beta(f,x) = \|\mathrm{D}f(x)^\dagger f(x)\| \le \|\mathrm{D}f(x)^\dagger\|\|f(x)\| \le \|f\|\|\mathrm{D}f(x)^\dagger\Delta\|\frac{\|f(x)\|}{\|f\|} = \overline{\beta}(f,x).$$

We finally conclude that $\alpha(f,x) \le \overline{\alpha}(f,x)$.

The main theorem in the theory of point estimates, see [22] and [13, Theorem 19.9] for a proof.

Theorem 2.2 *Let $f : \mathbb{R}^{n+1} \to \mathbb{R}^m$, $m \le n+1$, be analytic. Set $\alpha_0 = 0.125$. Let $x \in \mathbb{R}^{n+1}$ with $\alpha(f,x) < \alpha_0$. Then x is an approximate zero of f and $\|x - \xi\| < 2\beta(f,x)$ where ξ is the associated zero of x. Furthermore, if $n+1 = m$ and $\alpha(f,x) \le 0.02$, then all points in the ball in $\mathbb{S}^n$ with center x and radius $2\overline{\beta}(f,x)$ are approximate zeros of f with the same associated zero.*

2.6 Grids

Our algorithms work on a grid $\mathcal{G}_\eta$ on $\mathbb{S}^n$, which we construct by projecting onto $\mathbb{S}^n$ a grid on the cube. Let $\mathsf{C}^n = \{y \in \mathbb{R}^{n+1} \mid \|y\|_\infty = 1\}$ and $\phi : \mathsf{C}^n \to \mathbb{S}^n$ given by $\phi(y) = \frac{y}{\|y\|}$.

Given $\eta := 2^{-k}$ for some $k \ge 1$, we consider the uniform grid $\mathcal{U}_\eta$ of mesh η on C^n. This is the set of points in C^n whose coordinates are of the form $i2^{-k}$ for $i \in \{-2^k, -2^k+1, \cdots, 2^k\}$ with at least one coordinate equal to 1 or -1. We denote by $\mathcal{G}_\eta$ its image by ϕ in $\mathbb{S}^n$. An argument in elementary geometry shows that for $y_1, y_2 \in \mathsf{C}^n$,

$$\|\phi(y_1) - \phi(y_2)\| \le d_{\mathbb{S}}(\phi(y_1),\phi(y_2)) \le \frac{\pi}{2}\|y_1 - y_2\| \le \frac{\pi}{2}\sqrt{n}\,\|y_1 - y_2\|_\infty, \tag{2.4}$$

where $d_{\mathbb{S}}(x_1,x_2) := \arccos(\langle x_1, x_2\rangle) \in [0,\pi]$ denotes the angular distance, for $x_1, x_2 \in \mathbb{S}^n$.

Given $\varepsilon > 0$ and $x \in \mathbb{S}^n$, we denote by $B(x,\varepsilon) := \{y \in \mathbb{R}^{n+1} \mid \|y - x\| < \varepsilon\}$ the open ball with respect to the Euclidean distance and by $B_{\mathbb{S}}(x,\varepsilon) = \{y \in \mathbb{S}^n \mid d_{\mathbb{S}}(y,x) < \varepsilon\}$ the open ball with respect to the angular distance $d_{\mathbb{S}}$. We also set from now on

$$\mathrm{sep}(\eta) := \eta\sqrt{n}. \tag{2.5}$$

The following lemma is an immediate consequence of (2.4).

Lemma 2.1 *The union $\bigcup_{x\in\mathcal{G}_\eta} B_{\mathbb{S}}(x, \mathrm{sep}(\eta))$ covers the sphere $\mathbb{S}^n$.*

In [19, Lemma 3.1] and [13, Lemma 19.22], the following Exclusion Lemma is proved (the statement there is for $n = m$ but the proof holds for general m).

Lemma 2.2 (Exclusion Lemma) *Let $f \in \mathcal{H}_{\boldsymbol{d}}[m]$ and $x, y \in \mathbb{S}^n$ be such that $0 < d_{\mathbb{S}}(x,y) \leq \sqrt{2}$. Then*

$$\|f(x) - f(y)\| < \|f\|\sqrt{D}\; d_{\mathbb{S}}(x,y).$$

In particular, if $f(x) \neq 0$, there is no zero of f in the ball $B_{\mathbb{S}}\big(x, \frac{\|f(x)\|}{\|f\|\sqrt{D}}\big)$.

3 Solving the Problems

3.1 Feasibility

The first problem we consider is the feasibility problem for systems of homogeneous polynomials. This problem was tackled in [23] and our succinct exposition closely follows [13, §19.6]. We want to algorithmically solve the following problem:

Given a system $f \in \mathcal{H}_{\boldsymbol{d}}[m]$, does there exist $x \in \mathbb{S}^n$ such that $f(x) = 0$?

To analyze our algorithm we will need a notion of condition for the input system. Let $Z_{\mathbb{S}}(f)$ denote the set of zeros of f on the unit sphere $\mathbb{S}^n$. For $f \in \mathcal{H}_{\boldsymbol{d}}[m]$ we define

$$\kappa_{\mathrm{feas}}(f) = \begin{cases} \min\limits_{\zeta \in Z_{\mathbb{S}}(f)} \mu_{\mathrm{norm}}(f,\zeta), & \text{if } Z_{\mathbb{S}}(f) \neq \emptyset, \\ \max\limits_{x \in \mathbb{S}^n} \dfrac{\|f\|}{\|f(x)\|}, & \text{otherwise.} \end{cases}$$

We call f well-posed when $\kappa_{\mathrm{feas}}(f) < \infty$. Note that $\kappa_{\mathrm{feas}}(f) = \infty$ if and only if f is feasible and all its zeros are multiple.

The following is a pseudocode for our algorithm solving the feasibility problem.

Algorithm Feasibility

Input: $f \in \mathcal{H}_{\boldsymbol{d}}[m]$

Preconditions: $f_1, \cdots, f_m \neq 0$

---

let $\eta := \frac{1}{2}$
repeat
  if $\overline{\alpha}(f,x) \leq \alpha_0$ for some $x \in \mathcal{U}_\eta$
    then return "feasible" and halt
  if $\|f(x)\| > \frac{\pi}{2}\eta\sqrt{nD}\|f\|$ for all $x \in \mathcal{U}_\eta$
    then return "infeasible" and halt
  $\eta := \frac{\eta}{2}$

Output: a tag in {feasible, infeasible}

Postconditions: The algorithm halts if $\kappa_{\mathrm{feas}}(f) < \infty$. In this case the tag is feasible iff f has a zero in $\mathbb{S}^n$.

Theorem 3.1 *Algorithm Feasibility works correctly: With input a well-posed system it returns "feasible" (resp. "infeasible") if and only if the system is so. The number of iterations is bounded by $\mathcal{O}(\log_2(Dn\kappa_{\mathrm{feas}}(f)))$ and the total complexity, i.e., the number of arithmetic operations performed by the algorithm, by $(nD\kappa_{\mathrm{feas}}(f))^{\mathcal{O}(n)}$.*

Proof The correctness in the feasible case is a trivial consequence of Theorem 2.2 and the inequality $\alpha(f,x) \leq \overline{\alpha}(f,x)$. The correctness in the infeasible case follows from Lemma 2.2 along with the inequalities (2.4).

To see the complexity bound, assume first that f is feasible and let ζ in the cube C^n, $\zeta \in Z(f)$, be such that $\kappa_{\mathrm{feas}}(f) = \mu_{\mathrm{norm}}(f,\zeta)$. Let k be such that

$$\eta = 2^{-k} \leq \frac{\min\{4\alpha_0, 2C\,\overline{\varepsilon}\}}{\pi D^2 \sqrt{n}\ \kappa^2_{\mathrm{feas}}(f)}.$$

Here C and $\overline{\varepsilon}$ are the constants in Proposition 2.1. Let $x \in \mathcal{U}_\eta$ be such that $\|x-\zeta\|_\infty \leq \eta$. Then, by (2.4),

$$d_{\mathbb{S}}(x,\zeta) \leq \frac{\min\{2\alpha_0, C\,\overline{\varepsilon}\}}{D^2 \kappa^2_{\mathrm{feas}}(f)}.$$

Proposition 2.1 applies, and we have

$$\mu_{\mathrm{norm}}(f,x) \leq (1+\overline{\varepsilon})\mu_{\mathrm{norm}}(f,\zeta) = (1+\overline{\varepsilon})\kappa_{\mathrm{feas}}(f). \tag{3.1}$$

Also, by Lemma 2.2,

$$\|f(x)\| \leq \|f\|\sqrt{D} d_{\mathbb{S}}(x,\zeta) \leq \|f\| \frac{2\alpha_0}{D^{\frac{3}{2}} \kappa^2_{\mathrm{feas}}(f)}.$$

We then have

$$\overline{\alpha}(f,x) \;=\; \frac{D^{\frac{3}{2}}}{2}\mu^2_{\mathrm{norm}}(f,x)\frac{\|f(x)\|}{\|f\|} \;\leq\; \frac{D^{\frac{3}{2}}}{2}\kappa^2_{\mathrm{feas}}(f)\frac{2\alpha_0}{D^{\frac{3}{2}}\kappa^2_{\mathrm{feas}}(f)} \;=\; \alpha_0.$$

It follows that algorithm Feasibility halts at this point, and therefore the number k of iterations performed is at most $\mathcal{O}(\log_2(Dn\kappa_{\mathrm{feas}}(f)))$.

Assume finally that f is infeasible and let k be such that

$$\eta = 2^{-k} < \frac{2}{\pi\sqrt{nD}\ \kappa_{\mathrm{feas}}(f)}.$$

Then, at any point $y \in \mathcal{U}_\eta$ we have

$$\|f(x)\| \geq \frac{\|f\|}{\kappa_{\mathrm{feas}}(f)} > \frac{\pi}{2}\eta\sqrt{nD}\|f\|.$$

Again, algorithm Feasibility halts for this value of η, and the number k of iterations performed is also bounded by $\mathcal{O}(\log_2(Dn\kappa_{\mathrm{feas}}(f)))$.

At each iteration there are $2(n+1)\left(\frac{2}{\eta}\right)^n$ points in the grid. Hence, the number of points in the finest grid (the last run of the iteration) is $\mathcal{O}(n(D^2\sqrt{n}\kappa^2_{\mathrm{feas}}(f))^n)$. For each such point x we evaluate $\mu_{\mathrm{norm}}(f,x)$ and $\|f(x)\|$, both with cost $\mathcal{O}(N)$. The total complexity is therefore bounded by

$$\mathcal{O}((\log_2(Dn\kappa_{\mathrm{feas}}(f)))nN((D^2\sqrt{n}\kappa^2_{\mathrm{feas}}(f))^n) = (nD\kappa_{\mathrm{feas}}(f))^{\mathcal{O}(n)},$$

where we used that $N \leq m(n+D)^n$.

Remark 3.1 A finite-precision version of algorithm Feasibility can be implemented as well following the template result in Subsection 2.4 (see [23] for details). The running time remains of the same order and the returned tag is correct. The finest precision required by the algorithm satisfies

$$\varepsilon_{\mathrm{mach}} = \frac{1}{(Dn\kappa_{\mathrm{feas}}(f))^{\mathcal{O}(1)}}.$$

This shows that the required number of mantissa's bits to correctly decide feasibility is logarithmic in D, n and the condition $\kappa_{\mathrm{feas}}(f)$. The reason behind this reasonable bounds for an exponential time algorithm is that the algorithm performs an exponentially large of independent computations, each of them requiring only polynomial time and having a very simple nature.

3.2 Counting

The second problem we consider is the counting problem for square systems of homogeneous polynomials. That is, we consider systems in $\mathcal{H}_{\boldsymbol{d}} := \mathcal{H}_{\boldsymbol{d}}[n]$ with n polynomials in $n+1$ homogeneous variables. The zero set of one such system f is a finite set of, say ℓ, lines passing through the origin in $\mathbb{R}^{n+1}$. We can see this set as the set of zeros of f in projective space $\mathbb{P}(\mathbb{R}^{n+1})$ and we note that this set arises from the identification of antipodal points of the zero set $Z_{\mathbb{S}}(f)$, which consists of 2ℓ points. We thus want to algorithmically solve the following problem.

Given a system $f \in \mathcal{H}_{\boldsymbol{d}}$, count the number of points $x \in \mathbb{P}(\mathbb{R}^{n+1})$ such that $f(x) = 0$?

This problem was studied in [19–21] and, again, our exposition closely follows [13, §19.4]. As with the feasibility problem, our first step is to define an appropriate condition number. To do so now for a system f it is not enough to just consider the condition at its zeros. For points $x \in \mathbb{R}^{n+1}$ where $\|f(x)\|$ is non-zero but small, small perturbations of f can turn x into a new zero (and thus change its number of zeros). The following definition goes back to [17]:

$$\kappa(f, x) := \frac{\|f\|}{\{\|f\|^2 \mu_{\mathrm{norm}}^{-2}(f, x) + \|f(x)\|^2\}^{\frac{1}{2}}}, \tag{3.2}$$

where $\mu_{\mathrm{norm}}(f, x)$ is defined as in (2.1) for $x \in \mathbb{S}^n$, with the convention that $\infty^{-1} = 0$ and $0^{-1} = \infty$, and

$$\kappa(f) := \max_{x \in \mathbb{S}^n} \kappa(f, x). \tag{3.3}$$

We observe that, because of Remark 2.1, $\kappa(\lambda f) = \kappa(f)$. We also note that $\kappa(f) = \infty$ if and only if there exists $\xi \in \mathbb{S}^n$ such that $f(\xi) = 0$ (i.e., $\xi \in \mathcal{M}_{\mathbb{S}}$) and $\mathrm{D}f(\xi)$ is not surjective, i.e., f belongs to the set

$$\Sigma_{\mathrm{count}} := \{f \in \mathcal{H}_{\boldsymbol{d}} \mid \exists\, \xi \in \mathbb{S}^n \text{ such that } f(\xi) = 0 \text{ and } \mathrm{rank}(\mathrm{D}f(\xi)) < n\}. \tag{3.4}$$

The following result (see [20] or [13, Theorem 19.3]) relates condition to the distance to ill-posedness. It is a version, for polynomial systems, of the classical Eckard-Young theorem (see [26]).

Proposition 3.1 *For all $f \in \mathcal{H}_{\boldsymbol{d}}$,*

$$\kappa(f) = \frac{\|f\|}{\mathrm{dist}(f, \Sigma_{\mathrm{count}})}.$$

Our algorithm to solve the counting problem (described below) relies on some graph theoretical ideas. We consider, as in the previous section, a number $\eta \in (0,1)$ and the grid $\mathcal{G}_\eta$. We associate to this grid the graph G_η defined as follows. We set

$$A(f) := \{x \in \mathbb{S}^n \mid \overline{\alpha}(f,x) < \alpha_0\}.$$

The vertices of the graph are the points in $\mathcal{G}_\eta \cap A(f)$. Two vertices $x, y \in G_\eta$ are joined by an edge if and only if $\overline{B}(x) \cap \overline{B}(y) \neq \emptyset$, where

$$\overline{B}(x) = \{z \in \mathbb{S}^n \mid d_{\mathbb{S}}(x,z) \leq 2\overline{\beta}(f,x)\}.$$

Note that as a simple consequence of Theorem 2.2, we obtain the following lemma.

Lemma 3.1 (a) *For each $x \in A(f)$ there exists $\zeta_x \in Z_{\mathbb{S}}(f)$ such that $\zeta_x \in \overline{B}(x)$. Moreover, for each point z in $\overline{B}(x)$, the Newton sequence starting at z converges to ζ_x.*

(b) *Let $x, y \in A(f)$. Then $\zeta_x = \zeta_y \Leftrightarrow \overline{B}(x) \cap \overline{B}(y) \neq \emptyset$.*

We define $W(G_\eta) := \bigcup_{x \in G_\eta} \overline{B}(x) \subset \mathbb{S}^n$, where $x \in G_\eta$ has to be understood as x running over all the vertices of G_η. Similarly, for a connected component U of G_η, we define

$$W(U) := \bigcup_{x \in U} \overline{B}(x).$$

The following lemma implies that the connected components of the graph G_η are of a very special nature: they are cliques. It also implies that

$$|Z_{\mathbb{S}}(f)| \geq \#\text{ connected components of } G_\eta. \tag{3.5}$$

Lemma 3.2 (a) *For each component U of G_η, there is a unique zero $\zeta_U \in Z_{\mathbb{S}}(f)$ such that $\zeta_U \in W(U)$. Moreover, $\zeta_U \in \bigcap_{x \in U} \overline{B}(x)$.*

(b) *If U and V are different components of G_η, then $\zeta_U \neq \zeta_V$.*

Proof (a) Let $x \in U$. Since $x \in A(f)$, by Lemma 3.1(a) there exists a zero ζ_x of f in $\overline{B}(x) \subseteq W(U)$. This shows the existence. For the uniqueness and the second assertion, assume that there exist zeros ζ and ξ of f in $W(U)$. Let $x, y \in U$ be such that $\zeta \in \overline{B}(x)$, and $\xi \in \overline{B}(y)$. Since U is connected, there exist $x_0 = x, x_1, \cdots, x_{k-1}, x_k := y$ in $A(f)$ such that (x_i, x_{i+1}) is an edge of G_η for $i = 0, \cdots, k-1$, that is, $\overline{B}(x_i) \cap \overline{B}(x_{i+1}) \neq \emptyset$. If ζ_i and ζ_{i+1} are the associated zeros of x_i and x_{i+1} in $Z_{\mathbb{S}}(f)$ respectively, then by Lemma 3.1(b) we have $\zeta_i = \zeta_{i+1}$, and thus $\zeta = \xi \in \overline{B}(x) \cap \overline{B}(y)$.

(b) Let $\zeta_U \in \overline{B}(x)$ and $\zeta_V \in \overline{B}(y)$ for $x \in U$ and $y \in V$. If $\zeta_U = \zeta_V$, then $\overline{B}(x) \cap \overline{B}(y) \neq \emptyset$ and x and y are joined by an edge; hence $U = V$.

If equality holds in (3.5), we can compute $|Z_{\mathbb{S}}(f)|$ by computing the number of connected components of G_η. The reverse inequality in (3.5) amounts to the fact that there are no zeros of f in $\mathbb{S}^n$ that are not in $W(G_\eta)$. To verify that this is the case, we want to find, for each point $x \in \mathcal{G}_n \setminus A(f)$, a ball centered at x such that $f \neq 0$ on this ball. In addition, we want the union of these balls to cover $\mathbb{S}^n \setminus W(G_\eta)$.

These considerations lead to the following algorithm:

Algorithm Zero_Counting

Input: $f \in \mathcal{H}_{\boldsymbol{d}}$

Preconditions: $f \neq 0$

---

let $\eta := \frac{1}{2}$
repeat
  let $U_1, \cdots, U_r$ be the connected components of G_η
  if
    (a) for $1 \leq i < j \leq r$
      for all $x_i \in U_i$ and all $x_j \in U_j$
        $d_{\mathbb{S}}(x_i, x_j) > \pi\eta\sqrt{n}$
    and
    (b) for all $x \in \mathcal{G}_\eta \setminus A(f)$
      $\|f(x)\| > \frac{\pi}{4}\eta\sqrt{nD}\|f\|$
  then return $r/2$ and halt
  else $\eta := \frac{\eta}{2}$

Output: $\ell \in \mathbb{N}$

Postconditions: The algorithm halts if $f \notin \Sigma_{\text{count}}$. In this case f has exactly ℓ zeros in $\mathbb{P}(\mathbb{R}^{n+1})$.

Theorem 3.2 *Given an input $f \in \mathcal{H}_{\boldsymbol{d}} \setminus \Sigma_{\text{count}}$, Algorithm Zero_Counting*:

(a) *Returns the number of zeros of f in $\mathbb{P}(\mathbb{R}^{n+1})$.*

(b) *Performs $\mathcal{O}(\log_2(nD\kappa(f)))$ iterations and has a total cost (number of arithmetic operations) satisfying*

$$\text{cost}(f) \leq (nD\kappa(f))^{\mathcal{O}(n)}.$$

(c) *It can be modified to return, in addition, at the same cost, and for each real zero $\zeta \in \mathbb{P}(\mathbb{R}^{n+1})$ of f, an approximate zero x of f with associated zero ζ.*

(d) *Assume $\mathcal{H}_{\boldsymbol{d}}$ is endowed with the standard Gaussian. Then, with probability at least $1 - (nD)^{-n}$ we have $\text{cost}(f) \leq (nD)^{\mathcal{O}(n^2)}$.*

(e) *Similarly, with probability at least $1 - 2^{-N}$ we have $\text{cost}(f) \leq 2^{\mathcal{O}(N^{\frac{3}{2}})}$.*

Sketch of Proof (a) This part claims the correctness of algorithm Zero_Counting. To prove it, we use some notions of spherical convexity.

Let H^n be an open hemisphere in $\mathbb{S}^n$ and $x_1, \cdots, x_q \in H^n$. Recall that the spherical convex hull of $\{x_1, \cdots, x_q\}$ is defined by

$$\text{sconv}(x_1, \cdots, x_q) := \text{cone}(x_1, \cdots, x_q) \cap \mathbb{S}^n,$$

where $\text{cone}(x_1, \cdots, x_q)$ is the smallest convex cone with vertex at the origin and containing the points $x_1, \cdots, x_q$.

Lemma 3.3 *Let $x_1, \cdots, x_q \in H^n \subset \mathbb{R}^{n+1}$. If $\bigcap_{i=1}^{q} B_{\mathbb{S}}(x_i, r_i) \neq \emptyset$, then $\text{sconv}(x_1, \cdots, x_q) \subseteq \bigcup_{i=1}^{q} B_{\mathbb{S}}(x_i, r_i)$.*

Proof Let $x \in \operatorname{sconv}(x_1, \cdots, x_q)$ and $y \in \bigcap_{i=1}^{q} B_{\mathbb{S}}(x_i, r_i)$. We will prove that $x \in B_{\mathbb{S}}(x_i, r_i)$ for some i. Without loss of generality we assume $x \neq y$. Let H be the open half-space

$$H := \{z \in \mathbb{R}^{n+1} : \langle z, y - x\rangle < 0\}.$$

We have

$$\begin{aligned} z \in H &\Leftrightarrow \langle z, y - x\rangle < 0 \Leftrightarrow -\langle z, x\rangle < -\langle z, y\rangle \\ &\Leftrightarrow \|z\|^2 + \|x\|^2 - 2\langle z, x\rangle \ < \ \|z\|^2 + \|y\|^2 - 2\langle z, y\rangle \\ &\Leftrightarrow \|z - x\|^2 < \|z - y\|^2, \end{aligned}$$

the second line following from $\|x\| = \|y\| = 1$. Therefore the half-space H is the set of points z in $\mathbb{R}^{n+1}$ such that the Euclidean distance $\|z - x\|$ is less than $\|z - y\|$.

On the other hand, H must contain at least one point of the set $\{x_1, \cdots, x_q\}$, since if this were not the case, the convex set $\operatorname{cone}(x_1, \cdots, x_q)$ would be contained in $\{z : \langle z, y - x\rangle \geq 0\}$, contradicting $x \in \operatorname{sconv}(x_1, \cdots, x_q)$. Therefore, there exists i such that $x_i \in H$. It follows that

$$\|x - x_i\| < \|y - x_i\|.$$

Since the function $z \mapsto 2 \arcsin\left(\frac{x}{2}\right)$ giving the length of an arc as a function of its chord is nondecreasing, we obtain

$$d_{\mathbb{S}}(x, x_i) < d_{\mathbb{S}}(y, x_i) \leq r_i.$$

We can now proceed. Assume that Algorithm Zero_Countinghalts. We want to show that if r equals the number of connected components of G_η, then $\#_{\mathbb{R}}(f) = \#Z_{\mathbb{S}}(f)/2 = \frac{r}{2}$. We already know by Lemma 3.2 that each connected component U of G_η determines uniquely a zero $\zeta_U \in Z_{\mathbb{S}}(f)$. Thus it is enough to prove that $Z_{\mathbb{S}}(f) \subseteq W(G_\eta)$. This would prove the reverse inequality in (3.5).

Assume, by way of contradiction, that there is a zero ζ of f in $\mathbb{S}^n$ such that ζ is not in $W(G_\eta)$. Let $B_\infty(\phi^{-1}(\zeta), \eta) := \{y \in \mathcal{U}_\eta \mid \|y - \phi^{-1}(\zeta)\|_\infty \leq \eta\} = \{y_1, \cdots, y_q\}$, the set of all neighbors of $\phi^{-1}(\zeta)$ in $\mathcal{U}_\eta$, and let $x_i = \phi(y_i)$, $i = 1, \cdots, q$. Clearly, $\phi^{-1}(\zeta)$ is in the cone spanned by $\{y_1, \cdots, y_q\}$, and hence $\zeta \in \operatorname{sconv}(x_1, \cdots, x_q)$.

We claim that there exists $j \leq q$ such that $x_j \notin A(f)$. Indeed, assume that this is not the case. We consider two cases.

(i) All the x_i belong to the same connected component U of G_η. In this case Lemma 3.2 ensures that there exists a unique zero $\zeta_U \in \mathbb{S}^n$ of f in $W(U)$ and $\zeta_U \in \bigcap_i \overline{B}(x_i)$. Since $x_1, \cdots, x_q$ lie in an open half-space of $\mathbb{R}^{n+1}$, we may apply Lemma 3.3 to deduce that

$$\operatorname{sconv}(x_1, \cdots, x_q) \subseteq \cup \overline{B}(x_i).$$

It follows that for some $i \in \{1, \cdots, q\}$, $\zeta \in \overline{B}(x_i) \subseteq W(U)$, contradicting that $\zeta \notin W(G_\eta)$.

(ii) There exist $\ell \neq s$ and $1 \leq j < k \leq r$ such that $x_\ell \in U_j$ and $x_s \in U_k$. Since condition (a) in the algorithm is satisfied, $d_{\mathbb{S}}(x_\ell, x_s) > \pi\eta\sqrt{n}$. But by the bounds (2.4),

$$d_{\mathbb{S}}(x_\ell, x_s) \leq \frac{\pi}{2}\sqrt{n}\|y_\ell - y_s\|_\infty$$

$$\leq \frac{\pi}{2}\sqrt{n}(\|y_\ell - \phi^{-1}(\zeta)\|_\infty + \|\phi^{-1}(\zeta) - y_s\|_\infty) \;\leq\; \pi\eta\sqrt{n},$$

a contradiction.

We have thus proved the claim. Let then $1 \leq j \leq q$ be such that $x_j \notin A(f)$. Then, using Lemma 2.2,

$$\|f(x_j)\| = \|f(x_j) - f(\zeta)\| \leq \|f\|\sqrt{D}\, d_{\mathbb{S}}(x_j, \zeta) \leq \frac{\pi}{2}\eta\sqrt{nD}\|f\|.$$

This is in contradiction with condition (b) in the algorithm being satisfied.

(b) We next prove the claimed bound for the cost. The idea is to show that when η becomes small enough, as a function of $\kappa(f), n, N$ and D, then conditions (a) and (b) in algorithm Zero_Counting are satisfied. To do so, we rely on the following result (see [13, Lemma 19.26]).

Lemma 3.4 *Let $x_1, x_2 \in G_\eta$ with associated zeros $\zeta_1 \neq \zeta_2$. If*

$$\eta \leq \frac{0.08}{D^{\frac{3}{2}}\pi\kappa(f)\sqrt{n}},$$

then $d_{\mathbb{S}}(x_1, x_2) > \pi\eta\sqrt{n}$.

Lemma 3.5 *Let $x \in \mathbb{S}^n$ be such that $x \notin A(f)$. Suppose $\eta \leq \frac{\alpha_0}{nD^2\kappa(f)^2}$. Then $\|f(x)\| > \frac{\pi}{4}\eta\sqrt{nD}\|f\|$.*

Proof Since $x \notin A(f)$, we have $\overline{\alpha}(f, x) \geq \alpha_0$. Also, $\kappa(f) \geq \kappa(f, x)$ which implies, by (3.2),

$$\kappa(f)^{-2} \leq 2\max\Big\{\mu_{\mathrm{norm}}(f, x)^{-2}, \frac{\|f(x)\|^2}{\|f\|^2}\Big\}.$$

We accordingly divide the proof into two cases.

Assume firstly that $\max\big\{\mu_{\mathrm{norm}}(f, x)^{-2}, \frac{\|f(x)\|^2}{\|f\|^2}\big\} = \frac{\|f(x)\|^2}{\|f\|^2}$.
In this case

$$\eta \leq \frac{\alpha_0}{nD^2\kappa(f)^2} \leq \frac{2\alpha_0\|f(x)\|^2}{nD^2\|f\|^2},$$

which implies

$$\|f(x)\| \geq \frac{\sqrt{\eta}\sqrt{n}D\|f\|}{\sqrt{2\alpha_0}} > \frac{\pi}{4}\eta\sqrt{nD}\|f\|,$$

the second inequality since $\eta \leq \frac{1}{2} < \frac{8D}{\pi^2\alpha_0}$.

Now assume instead that $\max\big\{\mu_{\mathrm{norm}}(f, x)^{-2}, \frac{\|f(x)\|^2}{\|f\|^2}\big\} = \mu_{\mathrm{norm}}(f, x)^{-2}$.
In this case

$$\eta \leq \frac{\alpha_0}{nD^2\kappa(f)^2} \leq \frac{2\alpha_0}{nD^2\mu_{\mathrm{norm}}(f, x)^2},$$

which implies $\alpha_0 \geq \frac{1}{2}\eta n D^2 \mu_{\mathrm{norm}}(f, x)^2$. Also

$$\alpha_0 \leq \overline{\alpha}(f, x) = \frac{1}{2}\beta(f, x)\mu_{\mathrm{norm}}(f, x)D^{\frac{3}{2}} \leq \frac{1}{2\|f\|}\mu_{\mathrm{norm}}(f, x)^2 D^{\frac{3}{2}}\|f(x)\|.$$

Putting both inequalities together, we obtain

$$\frac{1}{2}\eta n D^2 \mu_{\mathrm{norm}}(f, x)^2 \leq \frac{1}{2\|f\|}\mu_{\mathrm{norm}}(f, x)^2 D^{\frac{3}{2}}\|f(x)\|,$$

which implies

$$\|f(x)\| \geq \eta n D^{\frac{1}{2}} \|f\| > \frac{\pi}{4} \eta \sqrt{nD} \|f\|.$$

We can now conclude the proof of part (b). Assume

$$\eta \leq \eta_0 := \min \Big\{ \frac{0.08}{\pi D^{\frac{3}{2}} \sqrt{n}\, \kappa(f)}, \frac{\alpha_0}{nD^2 \kappa(f)^2} \Big\}.$$

Then the hypotheses of Lemmas 3.4–3.5 hold. The first of these lemmas ensures that condition (a) in algorithm Zero_Counting is satisfied, the second, that condition (b) is satisfied as well. Therefore, the algorithm halts as soon as $\eta \leq \eta_0$. This gives a bound of $\mathcal{O}(\log_2(nD\kappa(f)))$ for the number of iterations.

At each iteration there are $K := 2(n+1)\big(\frac{2}{\eta}\big)^n$ points in the grid. For each such point x we evaluate $\mu_{\mathrm{norm}}(f,x)$ and $\|f(x)\|$, both with cost $\mathcal{O}(N)$. We can therefore decide with cost $\mathcal{O}(KN)$ which of these points are vertices of G_η and for those points x compute the radius $2\overline{\beta}(f,x)$ of the ball $\overline{B}_f(x)$. Therefore, with cost $\mathcal{O}(K^2N)$ we can compute the edges of G_η. The number of connected components of G_η is then computed with $\mathcal{O}(K^2N)$ operations as well by standard algorithms in graph theory (see the Notes for references).

Since $d_{\mathbb{S}}$ is computed with $\mathcal{O}(n)$ operations, the total cost of verifying condition (a) is at most $\mathcal{O}(K^2 n)$, and the additional cost of verifying (b) is $\mathcal{O}(K)$. It follows that the cost of each iteration is $\mathcal{O}(K^2N)$. Furthermore, since at these iterations $\eta \geq \eta_0$, we have $K \leq (\mathbf{C}(n+1)D^2\kappa(f)^2)^{n+1}$. Using this estimate in the $\mathcal{O}(K^2N)$ cost of each iteration and multiplying by the bound $\mathcal{O}(\log_2(nD\kappa(f)))$ for the number of iterations, the claimed bound for the total cost follows.

(c) To prove this part just note that for $i = 1, \cdots, r$, any vertex x_i of U_i is an approximate zero of the only zero of f in $W(U_i)$.

(d) The probabilistic analysis relies, among other results, on understanding the nature of a particular complex algebraic set. For $q \leq n+1$ let

$$\Sigma^{\mathbb{C}}_{\boldsymbol{d}}[q] := \{F \in \mathcal{H}^{\mathbb{C}}_{\boldsymbol{d}}[q] \mid \exists x \in \mathbb{P}(\mathbb{C}^{n+1})\ F(x) = 0,\ \operatorname{rank} DF(x)_{|T_x} < q\}.$$

Here $\mathcal{H}^{\mathbb{C}}_{\boldsymbol{d}}[q]$ denotes the linear space of q-tuples of complex homogeneous polynomials in $n+1$ variables with degree pattern $\boldsymbol{d}$ and $\mathbb{P}(\mathbb{C}^{n+1})$ the complex projective space of dimension n. The following result (see [14, Proposition 4.20]) tells us some basic features of $\Sigma^{\mathbb{C}}_{\boldsymbol{d}}[q]$.

Proposition 3.2 *For any $q \leq n+1$, the set $\Sigma^{\mathbb{C}}_{\boldsymbol{d}}[q] \subseteq \mathcal{H}^{\mathbb{C}}_{\boldsymbol{d}}[q]$ is an algebraic hypersurface defined by an irreducible polynomial with integer coefficients of degree at most $n2^nD^n$.*

We now note that $\Sigma_{\mathrm{count}} \subset \Sigma^{\mathbb{C}}_{\boldsymbol{d}}[n] \cap \mathbb{R}^N$ where $N = \dim_{\mathbb{R}} \mathcal{H}_{\boldsymbol{d}} = \dim_{\mathbb{C}} \mathcal{H}^{\mathbb{C}}_{\boldsymbol{d}}$. Therefore, by the proposition above, Σ_{count} is included in a real algebraic surface given as the zero set of a polynomial of degree at most $n2^nD^n$. An immediate application of Proposition 3.1 and Theorem 2.1 yields

$$\operatorname*{Prob}_{f\sim N(0,\mathrm{Id})}(\kappa(f) \geq t) = \operatorname*{Prob}_{f\sim N(0,\mathrm{Id})}\Big(\frac{\|f\|}{d(f,\Sigma_{\mathrm{count}})} \geq t\Big) \leq \frac{11n2^nD^nN}{t}.$$

By taking $t = (nD)^{cn}$ for a constant c large enough we have

$$\operatorname*{Prob}_{f\sim N(0,\mathrm{Id})}\big(\kappa(f) \geq (nD)^{cn}\big) \leq \frac{11n2^nD^nN}{(nD)^{cn}} \leq (nD)^{-n}.$$

Hence, using part (b), for c large enough, we have

$$\operatorname*{Prob}_{f\sim N(0,\mathrm{Id})}\left(\mathrm{cost}(f)\geq (nD)^{\mathcal{O}(n^2)}\right)\leq (nD)^{-n}.$$

(e) We now take $t=2^{cN}$ with a constant c large enough and use that $N\geq \frac{n^2}{2}$ (as we may assume that there is at least one polynomial of degree 2).

Remark 3.2 As with algorithm Feasibility, a finite-precision version of algorithm Zero_Counting, has been analyzed in [19]. Again, the running time remains of the same order and the returned value is the number of zeros of the input f in $\mathbb{P}(\mathbb{R}^{n+1})$ as long as the round-off unit satisfies

$$\varepsilon_{\mathrm{mach}}\leq \frac{1}{\mathcal{O}(D^2n^{\frac{5}{2}}\kappa(f)^3(\log_2 N+n^{\frac{3}{2}}D^2\kappa(f)^2))}.$$

We note now that the probability tail for $\kappa(f)$ implies that, with probability at least $1-(nD)^{-n}$ the number of mantissa's bits necessary to ensure correctness of the computed result is $\mathcal{O}(\log(nD))$.

3.3 Homology

The last problem we consider is the computation of the homology of basic semialgebraic sets. These are subsets of $\mathbb{R}^n$ defined by a system of equalities and inequalities

$$f_1(x)=\cdots=f_q(x)=0\quad\text{and}\quad g_1(x)\succ 0,\cdots,g_s(x)\succ 0, \tag{3.6}$$

where $F=(f_1,\cdots,f_q)$ and $G=(g_1,\cdots,g_s)$ are tuples of polynomials with real coefficients and the expression $g(x)\succ 0$ stands for either $g(x)\geq 0$ or $g(x)>0$ (we use this notation to emphasize the fact that our main result does not depend on whether the inequalities in (3.6) are strict). Let $W(F,G)$ denote the solution set of the semialgebraic system (3.6), for a vector $\mathbf{d}=(d_1,\cdots,d_{q+s})$ of $q+s$ positive integers, let $\mathcal{P}_{\boldsymbol{d}}$ (or $\mathcal{P}_{\boldsymbol{d}}[q;s]$ to emphasize the number of components) denote the linear space of the $(q+s)$-tuples of real polynomials in n variables of degree $d_1,\cdots,d_{q+s}$, respectively. Our (clearly more ambitious) goal now is the following:

Given $(F,G)\in\mathcal{P}_{\boldsymbol{d}}[q;s]$, compute the homology groups of $W(F,G)$.

A numerical algorithm solving this problem was recently given in [14]. Our exposition follows this paper.

A first remark is that the set $W(F,G)$ may be unbounded and hence, not suitable for our grid methods. The obvious solution is to "fit it" into a sphere via homogeneization. Given a degree pattern $\boldsymbol{d}=(d_1,\cdots,d_{q+s})$, the homogeneization of polynomials (with respect to that pattern) yields an isomorphism of linear spaces

$$\mathcal{P}_{\boldsymbol{d}}[q;s]\to\mathcal{H}_{\boldsymbol{d}}[q;s],\quad \psi=(F,G)\mapsto\psi^{\mathrm{hm}}=(F^{\mathrm{hm}},G^{\mathrm{hm}}),$$

where F^{h} denotes the homogeneization of F with homogeneizing variable X_0. The Weyl inner product on $\mathcal{H}_{\boldsymbol{d}}[q;s]$ induces an inner product on $\mathcal{P}_{\boldsymbol{d}}[q;s]$ such that the above map is isometric. The set $W(F,G)\subseteq\mathbb{R}^n$ is diffeomorphic to the subset of $\mathbb{S}^n$ defined by $F^{\mathrm{hm}}=0$, $G^{\mathrm{hm}}\succ 0$ and $X_0>0$ and we can therefore focus on computing the homology of the latter. As it happens, we can further relax the inequalities in this system. But to justify this we need to understand condition.

A (closed) homogeneous semialgebraic system has the form

$$f_1(x) = 0, \cdots, f_q(x) = 0 \quad \text{and} \quad g_1(x) \geq 0, \cdots, g_s(x) \geq 0, \tag{3.7}$$

where f_i and g_j are homogeneous polynomials in $\mathbb{R}[X_0, X_1, \cdots, X_n]$. The system is therefore an element $(F, G) \in \mathcal{H}_{\boldsymbol{d}}[q; s]$. The set of solutions $x \in \mathbb{S}^n$ of system (3.7), which we will denote by $S(F, G)$, is a spherical basic semialgebraic set. To such a system (F, G) we associate the condition number $\kappa_*(F, G)$ as follows. For a subtuple $L = (g_{j_1}, \cdots, g_{j_\ell})$ of G, we denote by F^L the system obtained from F by appending the polynomials from L, that is,

$$F^L := (f_1, \cdots, f_q, g_{j_1}, \cdots, g_{j_\ell}) \in \mathcal{H}_{\boldsymbol{d}}[q + \ell],$$

where now $\boldsymbol{d}$ denotes the appropriate degree pattern in $\mathbb{N}^{q+\ell}$. Abusing notation, we will frequently use set notations $L \subseteq G$ or $g \in G$ to denote subtuples or coefficients of G.

Definition 3.1 *Let $q \leq n+1$, $(F, G) \in \mathcal{H}_{\boldsymbol{d}}[q; s]$. The condition number of the homogeneous semialgebraic system (F, G) is defined as*

$$\kappa_*(F, G) := \max_{\substack{L \subseteq G \\ q + |L| \leq n+1}} \kappa(F^L).$$

We define Σ_* as the set of all $(F, G) \in \mathcal{H}_{\boldsymbol{d}}[q; s]$ such that $\kappa_*(F, G) = \infty$.

Clearly, Σ_* is semialgebraic and invariant under scaling of the $q + s$ components.

The next result, Proposition 4.14 in [14], shows that we can relax inequalities.

Proposition 3.3 *Let $(F, G) \in \mathcal{H}_{\boldsymbol{d}}[q; s]$ be such that $\kappa_*(F, G) < \infty$. Put $S := S(F, G)$, let $r \leq s$, and let $S' \subseteq S$ be the solution set in $\mathbb{S}^n$ of the semialgebraic system*

$$f_1 = \cdots = f_q = 0, \quad g_1 \geq 0, \cdots, g_r \geq 0, \quad g_{r+1} > 0, \cdots, g_s > 0.$$

Moreover, let ∂S denote the boundary of S in $S(F, \varnothing)$. Then $S \backslash \partial S \subseteq S'$ and S' is homotopically equivalent to S.

We have thus reduced our problem to the computation of the homology of the set $S(H(F, G))$ where $H(F, G)$ denotes the homogeneous semialgebraic system

$$F^{\mathrm{hm}} = 0, \quad G^{\mathrm{hm}} \geq 0, \quad \|(F, G)\| X_0 \geq 0.$$

Note the coefficient $\|(F, G)\|$ in the last polynomial. This coefficient does not change the solution set in $\mathbb{S}^n$ but allows us to control the condition (by avoiding artificial differences in the scaling of the polynomials of $H(F, G)$). Indeed, for $\psi = (F, G) \in \mathcal{P}_{\boldsymbol{d}}[q; s]$, we define

$$\kappa_*^{\mathrm{aff}}(\psi) := \kappa_*(H(\psi))$$

and call $\Sigma_*^{\mathrm{aff}} := H^{-1}(\Sigma_*)$ the set of ill-posed affine semialgebraic systems. In summary, if $\kappa_*^{\mathrm{aff}}(F, G) < \infty$, then the spherical set $S(H(F, G))$ is homotopically equivalent to $W(F, G)$.

Furthermore, we have the following result [14, Proposition 4.16], extending Proposition 3.1.

Proposition 3.4 *For any nonzero $\psi \in \mathcal{P}_{\boldsymbol{d}}[q; s]$,*

$$\kappa_*^{\mathrm{aff}}(\psi) \leq \frac{4D\|\psi\|}{d(\psi, \Sigma_*^{\mathrm{aff}})}.$$

3.3.1 The reach of a closed set

Let E be a real Euclidean space of finite dimension. For a nonempty subset $W \subseteq E$ and $x \in E$, we denote by $d_W(x) := \inf_{p \in W} \|x - p\|$ the distance of x to W.

Definition 3.2 *Let $W \subseteq E$ be a nonempty closed subset. The medial axis of W is defined as the closure of the set*

$$\Delta_W := \{x \in E \mid \exists p, q \in W,\ p \neq q \text{ and } \|x - p\| = \|x - q\| = d_W(x)\}.$$

The reach (or local feature size) of W at a point $p \in W$ is defined as $\tau(W, p) := d_{\Delta_W}(p)$. The (global) reach of W is defined as $\tau(W) := \inf_{p \in W} \tau(W, p)$. We also set $\tau(\varnothing) := +\infty$.

By the (open) neighborhood of radius $r \geq 0$ around a nonempty set $S \subseteq E$ we understand the set

$$\mathcal{U}(S, r) := \{p \in E \mid d_S(p) < r\}.$$

In [33, Proposition 7.1], Niyogi, Smale and Weinberger gave an answer to the following question: Given a compact submanifold $S \subseteq E$, a finite set $\mathcal{X} \subset E$ and $\varepsilon > 0$, which conditions do we need to ensure that S is a deformation retract of $\mathcal{U}(\mathcal{X}, \varepsilon)$?

Theorem 3.3 below (see [14, §2.2] for a proof) gives an extension to their result to any compact subsets $S, \mathcal{X}$ provided S has positive reach $\tau(S)$.

Recall that the Hausdorff distance between two nonempty closed subsets $A, B \subseteq E$ is defined as

$$d_H(A, B) := \max\Big(\sup_{a \in A} d_B(a), \sup_{b \in B} d_A(b)\Big).$$

Theorem 3.3 *Let S and $\mathcal{X}$ be nonempty compact subsets of E. The set S is a deformation retract of $\mathcal{U}(\mathcal{X}, \epsilon)$ for any ϵ such that $3\, d_H(S, \mathcal{X}) < \epsilon < \frac{1}{2}\, \tau(S)$.*

Theorem 3.3 is the main stepping stone towards computing the homology groups of $S(H(F, G))$. The reach $\tau(S(H(F, G)))$ of this set, however, is not easily computable. The following result, Theorem 4.12 in [14], shows a simple bound for it in terms of κ_*.

Theorem 3.4 *For any homogeneous semialgebraic system (F, G) defining a semialgebraic set $S := S(F, G) \subseteq \mathbb{S}^n$, if $\kappa_*(F, G) < \infty$, then*

$$D^{\frac{3}{2}}\, \tau(S)\, \kappa_*(F, G) \geq \tfrac{1}{7}.$$

3.3.2 Tubes and relaxations

For a subset $A \subseteq \mathbb{S}^n$ we denote by

$$\mathcal{U}_{\mathbb{S}}(A, r) := \{x \in \mathbb{S} \mid d_{\mathbb{S}}(x, A) < r\}$$

the open r-neighborhood of A with respect to the geodesic distance $d_{\mathbb{S}}$ on the sphere $\mathbb{S}^n$. Also, for a homogeneous system $(F, G) \in \mathcal{H}_{\boldsymbol{d}}[q; s]$ and $r > 0$, we define the r-relaxation of $S(F, G)$:

$$\mathrm{Approx}(F, G, r) := \{x \in \mathbb{S}^n \mid \forall f \in F,\ |f(x)| < \|f\| r \text{ and } \forall g \in G,\ g(x) > -\|g\| r\}.$$

It is clear that $S(F,G) \subseteq \mathrm{Approx}(F,G,r)$ for any $r > 0$. Also, it is easy to see that $\mathrm{Approx}(F,G,r)$ converges to S with respect to the Hausdorff distance, when $r \to 0$. The next two results (see [14, §4.2]) quantify more precisely this behaviour in terms of the condition number $\kappa_*(F,G)$. Recall that D denotes the maximum degree of the components of F and G.

Proposition 3.5 *For any $r > 0$,*

$$\mathcal{U}_{\mathbb{S}}(S(F,G), D^{-\frac{1}{2}}r) \subseteq \mathrm{Approx}(F,G,r).$$

Theorem 3.5 *Let $q \leq n+1$. For any positive number $r < (13D^{\frac{3}{2}}\kappa_*^2)^{-1}$ we have*

$$\mathrm{Approx}(F,G,r) \ \subseteq \ \mathcal{U}_{\mathbb{S}}(S(F,G), 3\kappa_* r).$$

3.3.3 Algorithms and analyses

We can now describe an algorithm computing a covering for $S(F,G)$, with $(F,G) \in \mathcal{H}_{\boldsymbol{d}}[q;s]$, as in Theorem 3.3.

Algorithm Covering

Input: $(F,G) \in \mathcal{H}_{\boldsymbol{d}}[q;s]$

Preconditions: $q \leq n$

$\eta := 1$
repeat
  $\eta := \frac{\eta}{2}$
  $r = \mathrm{sep}(\eta)$
  $K_* := \max\{\kappa(F^L, x) \mid x \in \mathcal{G}_\eta$ and $L \subseteq G$ such that $|L| \leq n - q + 1\}$
until $71D^{\frac{5}{2}}K_*^2 r < 1$
return the set $\mathcal{X} := \mathcal{G}_\eta \cap \mathrm{Approx}(F,G,D^{\frac{1}{2}}r)$ and $\varepsilon := 5DK_*r$

Output: A finite subset $\mathcal{X}$ of $\mathbb{S}^n$ and an $\epsilon > 0$.

Postconditions: The algorithm halts if $(F,G) \notin \Sigma_*$. In this case $\mathcal{U}(\mathcal{X},\epsilon)$ is homotopically equivalent to $S(F,G)$.

Proposition 3.6 *On input F and G, algorithm covering outputs a finite set $\mathcal{X}$ and an $\varepsilon > 0$ such that $\mathcal{U}(\mathcal{X},\varepsilon)$ is homotopically equivalent to $S(F,G)$. Moreover, the computation performs $((s+n)D\kappa_*)^{\mathcal{O}(n)}$ arithmetic operations, where $s = |G|$ and $\kappa_* = \kappa_*(F,G)$, and the number $|\mathcal{X}|$ of points in $\mathcal{X}$ is $(nD\kappa_*)^{\mathcal{O}(n)}$.*

Proof Let $\kappa_* := \kappa_*(F,G)$, $S := S(F,G)$ and let η, r, and K_* be the values of the corresponding variables after the repeat loop terminates in algorithm covering. By design,

$$\frac{1}{2} < 71\, D^{\frac{5}{2}}K_*^2 r < 1. \tag{3.8}$$

We will first show that

$$\kappa_* \leq \Big(1 + \frac{1}{100}\Big)K_*. \tag{3.9}$$

Let $L \subseteq G$ and $y \in \mathbb{S}^n$ be such that $\kappa_* = \kappa(F^L) = \kappa(F^L, y)$. Because of Lemma 2.1 there is some $x \in \mathcal{G}_\eta$ such that $d_{\mathbb{S}}(x,y) < r$, and $\kappa(F^L, x) \leq K_*$ by the definition of K_*. Since the map

$x \mapsto \frac{1}{\kappa(F^L,x)}$ is D-Lipschitz continuous (see [14, Proposition 4.7]), we have

$$\kappa_* = \kappa(F^L, y) \leq \frac{\kappa(F^L, x)}{1 - D\kappa(F^L, x)r} \leq \frac{K_*}{1 - DK_*r}.$$

Inequality (3.8) shows that

$$DK_*r < \frac{1}{71\, D^{\frac{3}{2}} K_*} \leq \frac{1}{101},$$

the last as $D \geq 2$ and $K_* \geq 1$, and Inequality (3.9) follows.

Let $\mathcal{X} := \mathcal{G}_\eta \cap \mathrm{Approx}(F, G, D^{\frac{1}{2}}r)$ and $\epsilon := 5DK_*r$, that is, the finite set and the real number output by the algorithm. We will now prove that $\mathcal{U}(\mathcal{X}, \epsilon)$ is homotopically equivalent to S. By Theorem 3.3, it is enough to prove the inequalities

$$3d_H(\mathcal{X}, S) < \epsilon < \frac{1}{2}\tau(S). \tag{3.10}$$

The second inequality follows from Inequalities (3.8)–(3.9) and Theorem 3.4:

$$\varepsilon = 5DK_*r < \frac{5}{71}\,\frac{1}{D^{\frac{3}{2}}K_*} \leq \frac{505}{7100}\,\frac{1}{D^{\frac{3}{2}}\kappa_*} \;\leq\; \frac{3535}{7100}\tau(S) \leq \frac{1}{2}\tau(S).$$

Concerning the inequality $3d_H(\mathcal{X}, S) < \epsilon$, let $x \in S$. Because of Lemma 2.1, there is some $y \in \mathcal{G}_\eta$ with $d_{\mathbb{S}}(x, y) < r$. Hence y lies in $\mathrm{Approx}(F, G, D^{\frac{1}{2}}r)$, by Proposition 3.5. Thus $y \in \mathcal{X}$ and $d(x, \mathcal{X}) < d_{\mathbb{S}}(x, y) < r < \frac{1}{3}\varepsilon$.

Next, let $x \in \mathcal{X}$. Then, $x \in \mathrm{Approx}(F, G, D^{\frac{1}{2}}r)$ and

$$13\, D^{\frac{3}{2}}\kappa_*^2 r < 13 \cdot 4\, D^{\frac{5}{2}}\kappa_*^2 r < 1,$$

the last by Inequality (3.8). Hence, Theorem 3.5 applies and shows that

$$d(x, S) \leq 3\kappa_* r \leq \Big(3 + \frac{3}{100}\Big)K_* r < \frac{1}{3}\epsilon,$$

where we used $D \geq 2$ for the last inequality. Thus we have shown that $d_H(\mathcal{X}, S) < \frac{1}{3}\epsilon$. This concludes the proof of (3.10) and of the homotopy equivalence.

Lastly, we deal with the complexity analysis. We can approximate $\kappa(F^L, x)$ within a factor of 2 in $\mathcal{O}(N + n^3)$ operations (see [31, §2.5]) and this is enough for our needs. For simplicity, we will do as if we could compute κ exactly.

The repeat loop performs $\mathcal{O}(\log(D\kappa_*))$ iterations. Each iteration can be done in $\mathcal{O}(|\mathcal{G}_\eta| M(N + n^3))$ operations, where $M = \sum_{i=0}^{n+1-q} \binom{s}{i} \leq (s+1)^{n+1-q}$. Moreover, $|\mathcal{X}| \leq |\mathcal{G}_\eta| = (nD\kappa_*)^{\mathcal{O}(n)}$ and $N + n^3 = (nD)^{\mathcal{O}(n)}$. Therefore, the total number of operations is bounded by $((s+n)D\kappa_*)^{\mathcal{O}(n)}$.

Once in the possession of a pair $(\mathcal{X}, \varepsilon)$ such that S is a deformation retract of $\mathcal{U}(\mathcal{X}, \varepsilon)$, the computation of the homology groups of S is a known process. One computes the nerve $\mathcal{N}$ of the covering $\{B(x, \varepsilon) \mid x \in \mathcal{X}\}$ (this is a simplicial complex whose elements are the subsets N of $\mathcal{X}$ such that $\bigcap_{x \in N} B(x, \varepsilon)$ is not empty) and from it, its homology groups $H_k(\mathcal{N})$. Since the intersections of any collection of balls is convex, the Nerve Theorem (see, e.g., [7, Theorem 10.7]) ensures that

$$H_k(\mathcal{N}) \simeq H_k(\mathcal{U}(\mathcal{X}, \varepsilon)) \simeq H_k(S),$$

the last because S is a deformation retract of $\mathcal{U}(\mathcal{X}, \varepsilon)$. The process is described in detail in [22, §4] where the proof for the following result can be found.

Proposition 3.7 *Given a finite set $\mathcal{X} \subseteq \mathbb{R}^{n+1}$ and a positive real number ϵ, one can compute the homology of $\bigcup_{x\in\mathcal{X}} B(x,\epsilon)$ with $|\mathcal{X}|^{\mathcal{O}(n)}$ operations.*

We finally can put all the ingredients together in our main result.

Theorem 3.6 *There is an algorithm Homology, working over the reals and numerically stable that, given a system $(F,G) \in \mathcal{P}_{\boldsymbol{d}}$ with $q \le n$ equalities and s inequalities, computes the homology groups of $W(F,G)$. Moreover, the number of arithmetic operations in $\mathbb{R}$ performed by Homology on input (F,G), denoted $\operatorname{cost}(F,G)$, satisfies*

(a) $\operatorname{cost}(F,G) = ((s+n)D\delta^{-1})^{\mathcal{O}(n^2)}$ *where δ is the distance of $\frac{1}{\|(F,G)\|}(F,G)$ to Σ_*^{aff}.*

Furthermore, if (F,G) is drawn from the Gaussian measure on $\mathcal{P}_{\boldsymbol{d}}$, then

(b) $\operatorname{cost}(F,G) \le ((s+n)D)^{\mathcal{O}(n^3)}$ *with probability at least $1-((s+n)D)^{-n}$,*

(c) $\operatorname{cost}(F,G) \le 2^{\mathcal{O}(N^2)}$ *with probability at least $1-2^{-N}$.*

Sketch of Proof Following our previous stream of thoughts we consider the following algorithm:

Algorithm Homology

Input: $(F,G) \in \mathcal{P}_{\boldsymbol{d}}[q;s]$

Preconditions: $q \le n$, $\kappa_*^{\mathrm{aff}}(f) < \infty$

$(\mathcal{X},\varepsilon) :=$ Covering$(H(F,G))$
compute the nerve $\mathcal{N}$ of $\mathcal{U}(\mathcal{X},\varepsilon)$
return the homology groups of $\mathcal{N}$

Output: A description of a finite set of groups.

Postconditions: The algorithm halts if $(F,G) \notin \Sigma_*^{\mathrm{aff}}$. In this case the groups returned are the homology groups of $W(F,G)$.

Proposition 3.6 shows that $(\mathcal{X},\varepsilon)$ is computed with cost $((s+n)D\kappa_*(F,G)))^{\mathcal{O}(n)}$ and that $|\mathcal{X}| = (nD\kappa_*(F,G))^{\mathcal{O}(n)}$. Proposition 3.7 shows that we can further compute the homology of $\bigcup_{x\in\mathcal{X}} B(x,\varepsilon)$ with cost $|\mathcal{X}|^{\mathcal{O}(n)} = (nD\kappa_*(F,G))^{\mathcal{O}(n^2)}$. This homology coincides with that of $S(H(F,G))$ which in turn coincides, we have already argued, with that of $W(F,G)$. This, together with Proposition 3.4, shows part (a).

Parts (b) and (c) are proved as parts (d) and (e) in Theorem 3.2 using that Σ_*^{aff} is included in a union of at most $(s+1)^{n+1-q}$ sets, each of them having the form $\Sigma_{\boldsymbol{d}}^{\mathbb{C}}[m] \cap \mathbb{R}^M$ for appropriate values of $m \le n+1$ and $M \le N$.

Remark 3.3 (i) As with algorithms Feasibility and Zero_Counting, a finite-precision version of Homology can be implemented. This has not been done but would follow the same steps as in the other two cases and ensure that the computed homology groups (i.e., the Betti numbers and torsion coefficients corresponding to these groups) are correct as long as

$$\varepsilon_{\mathrm{mach}} \le \frac{1}{(Dn\kappa_*(F,G))^{\mathcal{O}(1)}}.$$

A discussion of this theme is in [22, §7].

(ii) It is immediate to check that $\kappa_{\mathrm{feas}}(f) \le \kappa_*(f)$. Hence, the probabilistic analysis for the latter gives bounds for the former as well. But the fact that the codimension of $\Sigma_{\mathrm{feas}} := \{f \mid$

$\kappa_{\text{feas}}(f) = \infty\}$ is much greater than that of Σ_{count} (which is equal to one) suggests that one might obtain substantially better bounds. There are no results as of today, however, supporting this suggestion.

References

[1] Allgower, E. L. and Georg, K., Numerical Continuation Methods, Springer-Verlag, Berlin, 1990.

[2] Amelunxen, D. and Lotz, M., Average-case complexity without the black swans, *J. Complexity*, **41**, 2017, 82–101.

[3] Basu, S., Algorithmic semi-algebraic geometry and topology—recent progress and open problems, Surveys on discrete and computational geometry, Contemp. Math., **453**, Amer. Math. Soc., Providence, RI, 2008, 139–212.

[4] Basu, S., Pollack, R. and Roy, M.-F., On the combinatorial and algebraic complexity of quantifier elimination, 35th Annual IEEE Symp. on Foundations of Computer Science, 1994, 632–641.

[5] Basu, S., Pollack, R. and Roy, M.-F., Algorithms in Real Algebraic Geometry, 2nd ed., Algorithms and Computation in Mathematics, **10**, Springer-Verlag, Berlin, 2006.

[6] Benedetti, R. and Risler, J.-J., Real algebraic and semi-algebraic sets, Hermann, Paris, 1990.

[7] Björner, A., Topological methods, Handbook of Combinatorics, R. Graham, M. Grotschel and L. Lovasz (eds.), North-Holland, Amsterdam, 1995, 1819–1872.

[8] Blum, L., Cucker, F., Shub, M. and Smale, S., Complexity and Real Computation, Springer-Verlag, New York, 1998.

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

[10] Bochnak, J., Coste, M. and Roy, M.-F., Real Algebraic Geometry, Ergebnisse der Mathematik und Ihrer Grenzgebiete (3), **36**, Springer-Verlag, Berlin, 1998.

[11] Bürgisser, P. and Cucker, F., Counting complexity classes for numeric computations II: Algebraic and semialgebraic sets, *J. Complexity*, **22**, 2006, 147–191.

[12] Bürgisser, P. and Cucker, F., Exotic quantifiers, complexity classes, and complete problems, *Found. Comput. Math.*, **9**, 2009, 135–170.

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

[14] Bürgisser, P., Cucker, F. and Lairez, P., Computing the homology of basic semialgebraic sets in weakly exponential time, 2017, arXiv: 1706.07473.

[15] Carlson, J., Jaffe, A., Wiles, A., et al., The Millenium Prize Problems, Clay Mathematics Institute, Cambridge, MA, American Mathematical Society, Providence, RI, 2006.

[16] Collins, G. E., Quantifier elimination for real closed fields by cylindrical algebraic deccomposition, Lect. Notes in Comp. Sci., **33**, Springer-Verlag, Berlin, 1975, 134–183.

[17] Cucker, F., Approximate zeros and condition numbers, *J. Complexity*, **15**, 1999, 214–226.

[18] Cucker, F., A theory of complexity, condition and roundoff, *Forum of Mathematics, Sigma*, e4, 2015.

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

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

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

[22] Cucker, F., Krick, T. and Shub, M., Computing the homology of real projective sets, *Found. Comp. Math.,* 2016, arXiv:1602.02094.

[23] Cucker, F. and Smale, S., Complexity estimates depending on condition and round-off error, *Journal of the ACM*, **46**, 1999, 113–184.

[24] Demmel, J., On condition numbers and the distance to the nearest ill-posed problem, *Numer. Math.*, **51**, 1987, 251–289.

[25] Demmel, J. W., Applied Numerical Linear Algebra, SIAM, Philadelphia, 1997.

[26] Eckart, C. and Young, G., The approximation of one matrix by another of lower rank, *Psychometrika*, **1**, 1936, 211–218.

[27] Grigoriev, D. Yu., Complexity of deciding Tarski algebra, *Journal of Symbolic Computation*, **5**, 1988, 65–108.

[28] Grigoriev, D. Yu. and Vorobjov, N. N., Solving systems of polynomial inequalities in subexponential time, *Journal of Symbolic Computation*, **5**, 1988, 37–64.

[29] Koiran, P., The real dimension problem is $\mathrm{NP}_{\mathbb{R}}$-complete, *Journal Complexity*, **15**, 1999, 227–238.

[30] Kostlan, E., Complexity theory of numerical linear algebra, *Journal of Computational and Applied Mathematics*, **22**, 1988, 219–230.

[31] Lairez, P., A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time, *Found. Comput. Math.*, 2017, DOI: 10.1007/s10208-016-9319-7.

[32] Mehta, M. L., Random matrices, 3rd ed., Pure and Applied Mathematics (Amsterdam), **142**, Elsevier/Academic Press, Amsterdam, 2004.

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

[34] Renegar, J., On the computational complexity and geometry of the first-order theory of the reals, Part I, *Journal of Symbolic Computation*, **13**, 1992, 255–299.

[35] Renegar, J., On the computational complexity and geometry of the first-order theory of the reals, Part II, *Journal of Symbolic Computation*, **13**, 1992, 301–327.

[36] Renegar, J., On the computational complexity and geometry of the first-order theory of the reals, Part III, *Journal of Symbolic Computation*, **13**, 1992, 329–352.

[37] Renegar, J., Incorporating condition measures into the complexity theory of linear programming, *SIAM J. Optim.*, **5**, 1995, 506–524.

[38] Renegar, J., Linear programming, complexity theory and elementary functional analysis, *Math. Program.*, **70**, 1995, 279–351.

[39] Shub, M. and Smale, S., Complexity of Bézout's theorem I: Geometric aspects, *Journal of the Amer. Math. Soc.*, **6**, 1993, 459–501.

[40] Shub, M. and Smale, S., Complexity of Bézout's theorem II: Volumes and probabilities, Computational Algebraic Geometry, F. Eyssette and A. Galligo (eds.), Progress in Mathematics, **109**, Birkhäuser, Boston, 1993, 267–285.

[41] Shub, M. and Smale, S., Complexity of Bézout's theorem III: Condition number and packing, *Journal of Complexity*, **9**, 1993, 4–14.

[42] Shub, M. and Smale, S., Complexity of Bézout's theorem V: Polynomial time, *Theoret. Comp. Sci.*, **133**, 1994, 141–164.

[43] Shub, M. and Smale, S., Complexity of Bézout's Theorem IV: Probability of success, extensions, *SIAM J. of Numer. Anal.*, **33**, 1996, 128–148.

[44] Smale, S., Newton's method estimates from data at one point, The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics, R. Ewing, K. Gross and C. Martin (eds.), Springer-Verlag, New York, 1986.

[45] Tarski, A., A Decision Method for Elementary Algebra and Geometry, University of California Press, 1951.

[46] Trefethen, L. N. and Bau III, D., Numerical Linear Algebra, SIAM, Philadelphia, 1997.

[47] Wüthrich, H. R., Ein Entscheidungsverfahren für die Theorie der reell-abgeschlossenen Körper, Komplexität von Entscheidungsproblemen, E. Specker and V. Strassen (eds.), Lect. Notes in Comp. Sci., **43**, Springer-Verlag, Berlin, 1976, 138–162.