Bounds on the infimum of polynomials over a generic semi-algebraic set using asymptotic critical values

Boulos El Hilany, Elias Tsigaridas

BOUNDS ON THE INFIMUM OF POLYNOMIALS OVER A GENERIC SEMI-ALGEBRAIC SET USING ASYMPTOTIC CRITICAL VALUES

Boulos El Hilany * **Elias Tsigaridas** †

July 25, 2024

Abstract

We present precise bit and degree estimates for the optimal value of the polynomial optimization problem $f^* := \inf_{x \in \mathcal{X}} f(x)$, where $\mathcal{X}$ is a semi-algebraic set satisfying some non-degeneracy conditions. Our bounds depend on the degree, the bitsize of f, and the polynomials defining $\mathcal{X}$, and are single exponential with respect to the number of variables. They generalize the single exponential bounds from Jeronimo, Perrucci, and Tsigaridas (SIAM Journal on Optimization, 23(1):241–255, 2013) for the minimum of a polynomial function on a compact connected component of a basic closed semi-algebraic set.

The tools that we use allow us to obtain specialized bounds and dedicated algorithms for two large families of polynomial optimization problems in which the optimum value might not be attained. The first family forms a dense set of real polynomial functions with a fixed collection of Newton polytopes; we provide the best approximation yet for the bifurcation set, which contains the optimal value, and we deduce an effective method for computations. As for the second family, we consider any unconstrained polynomial optimization problem; we present more precise bounds, together with a better bit complexity estimate of an algorithm to compute the optimal value.

***K*eywords** asymptotic critical value · bifurcation set · polynomial optimization · unconstrained polynomial optimization · Newton polytope · bit complexity

*Institut für Analysis und Algebra, TU Braunschweig, Germany, boulos.hilani@gmail.com, b.el-hilany@tu-braunschweig.de

†Inria Paris and Institut de Mathématiques de Jussieu–Paris Rive Gauche, Sorbonne Université and Paris Université, elias.tsigaridas@inrial.fr

Contents

1 Introduction

Let $f : \mathcal{X} \longrightarrow \mathbb{R}$ be a semi-algebraic function over a basic semi-algebraic set $\mathcal{X} \subset \mathbb{R}^n$. We consider the problem of computing effective bounds on the infimum $f^* := \inf_{x\in\mathcal{X}} f(x)$, assuming it lies in $\mathbb{R}$, as a function of the number of variables, the number of polynomials defining the semialgebraic set, their degrees and bitsize.

Problem 1. *The optimal value of the polynomial optimization problem $f^* = \inf_{x\in\mathcal{X}} f(x)$ is a real algebraic number, when the defining polynomials have rational coefficients. Provide an effective lower bound on its value and an upper bound on degree of the defining polynomial that depend on the number of variables n, the degrees, and the coefficient (bit)size of f and the polynomials defining $\mathcal{X}$.*

The computation of effective, explicit, and if possible (asymptotically) optimal bounds on the size and degree of the optimal value is an important problem on its own, as these are intrinsic quantities of the polynomial optimization problem that characterize and measure its difficulty. Furthermore, they are fundamental quantities for the analysis of various symbolic and numerical algorithms, with a wide range of applications, e.g., [44, 40, 5, 28]. As the optimal value is an algebraic number, the degree of the defining polynomial is essential in complexity statements, as it measures the algebraic complexity of the problem, e.g., [10, 1].

If the semi-algebraic subset $\mathcal{X}$ is compact, then the infimum of f is attained at $\mathcal{X}$. In this case, explicit bounds for Problem 1 and an effective method for computing the infimum already exist. Namely, if $\mathcal{X}$ is defined by r equalities, s inequalities with polynomials in $\mathbb{Z}[x_1, \ldots, x_n]$, having degrees at most d, and coefficients of absolute values at most H, then either f^* is not reached over $\mathcal{X}$, or is an algebraic number of degree at most

$$\max_{0\leq i\leq \min\{r+s,n\}} \binom{n}{i} d^i (d-1)^{n-i} \leq 2^{n-1} d^n, \tag{1}$$

and if it is not zero, then

$$|f^*| \geq (2^{5-\frac{n}{2}} \tilde{H} d^n)^{-n 2^n d^n}, \tag{2}$$

where $\tilde{H} := \max(H,\ n + r + s)$ [24, Theorem 1.1]. If we let $H \leq 2^\tau$, then $|f^*| \geq 2^{-\widetilde{\mathcal{O}}(n d^n \tau)}$. We notice that the bound is single exponential with respect to the number of variables; hence polynomial when the number of variables is fixed. Even more, it is (asymptotically) optimal. Jeronimo et al [24] proved the inequality in (2) by bounding the isolated roots of carefully deformed polynomial systems; this was the first precise bound for the optimum value of the polynomial optimization problem. While the initial motivation in [24] was to approximate the minimum distance between two semi-algebraic sets, the bounds has already several other applications, e.g., [44, 40]. Let us also mention that there is also an efficient (symbolic) algorithm to compute it due to Jeronimo and Perrucci [23].

While the bound of (2) is quite important, the assumption that the optimum value is attained is rather restrictive in several problems and applications. Not all semi-algebraic functions attain their infima; as an example, consider the infimum of the function $x \mapsto 1/x$ over $\{x > 0\}$ is 0. We make a step towards closing this gap in the literature by considering the Problem 1 where f admits a non-trivial infimum $f^* \in \mathbb{R}$, but it is not attained on $\mathcal{X}$; under some assumptions on $\mathcal{X}$ that we detail in the sequel. The computation of the optimum value, both in the *"attained"* and in the *"unattained"* case, appears frequently in science and engineering. Namely, systems and problems in applications, e.g., computer aided design [16], robotics [43], control systems [26], to mention a few, are modeled as semi-algebraic subsets, and so a recurring task is to decide whether two semi-algebraic subsets are disjoint, e.g., [41]. One approach to solve this problem is to figure out whether the images (under the same function) of the two semi-algebraic sets are disjoint or not. Accordingly, if the infimum of one function is larger than the maximum of the other, the two semi-algebraic sets are disjoint.

A straightforward algorithm to compute the optimum value of Problem 1 is to use (the general purpose approach of) cylindrical algebraic decomposition (CAD) for stratifying semi-algebraic projections, e.g. [7]. However, this approach requires us to compute many algebraic objects that are possibly unnecessary, as it relies on repeated projections. Even more, it has a worst-case complexity that is double exponential in the number of variables [3, 4]. It will also result double exponential bounds for the optimal value(s), f^*. Another approach consists in describing the problem as a general decision problem for the theory of reals and then perform quantifier elimination [2]. However, it is not clear, at least to us, what is the optimal formulation in this setting and how to obtain precise bit and degree estimates. We opt for a different approach.

We provide effective bounds on the infimum f^* of a semi-algebraic function $\mathcal{X} \longrightarrow \mathbb{R}$, when the polynomials involved in $\mathcal{X}$ satisfy a non-degeneracy condition. The bound is single exponential with respect to the number of variables (Theorem 1). Up to the non-degeneracy assumptions, our contribution fills the gap in the literature for Problem 1 and to the best of our knowledge is the first bound for when the optimal value is not attained.

Our approach builds on the works of Jelonek and Kurdyka [21, 20] and Jelonek and Tibăr [22] for approximating the *bifurcation set at infinity*, $\mathcal{B}_f^\infty$ of a polynomial function f [45, 47, 51]. The set $\mathcal{B}_f^\infty$ is the smallest set of values, outside of which f is a locally trivial $\mathcal{C}^\infty$-fibration "at infinity" (i.e., outside a large ball), and has been a subject of intensive study in the last fifty years, e.g., [47]. Clearly, the set $\mathcal{B}_f^\infty$ contains the infimum f^* of a real polynomial function. Consequently, and for a plethora of other reasons, it remains an important open problem to derive an efficient algorithm to compute $\mathcal{B}_f^\infty$ for any polynomial function f [47, 51]. The results of Jelonek, Kurdyka, and Tibăr lead to an efficient computation of the set of points, $\mathcal{K}_f^\infty$, called *Rabier set* of f, or the *asymptotic critical values* of f, as

$$\mathcal{K}_f^\infty := \big\{z \in \mathbb{R} \mid \exists \{\boldsymbol{x}_\ell\}_{\ell\in\mathbb{N}} \subset X,\ \|\boldsymbol{x}_\ell\| \underset{\ell\to\infty}{\longrightarrow} \infty,\ \|\boldsymbol{x}_\ell\| \cdot \|\operatorname{grad} f(\boldsymbol{x}_\ell)\| \longrightarrow 0,\ f(\boldsymbol{x}_\ell) \longrightarrow z\big\}. \tag{3}$$

This is a finite that contains $\mathcal{B}_f^\infty$, that is $\mathcal{B}_f^\infty \subseteq \mathcal{K}_F^\infty$ [35]. Therefore, these important results provide new computational tools for f^*. There are several related methods dedicated in computing $\mathcal{B}_f^\infty$ for large families of polynomial functions (e.g., [31, 54, 34]). We build on methods that require the fewer possible assumptions on the input (polynomial) functions.

One of our contributions is generalize the previous techniques from polynomial functions to semi-algebraic functions. In this way, we derive the best known estimates for bounds on f^*, and develop algorithms to compute f^* in two important special cases of Problem 1.

The first (special) case considers the semi-algebraic set $\mathcal{X}$ of Problem 1 to be a real affine variety that satisfies a mild non-degeneracy condition. This condition (Theorem 18) applies for a dense family of functions sharing the same collection of Newton polytopes and gives rise to functions having a non-trivial bifurcation set. We present single exponential bit and degree bounds for f^* and we introduce an algorithm to compute it with precise (single exponential) bit complexity estimates. This algorithm computes directly the bifurcation set and exploits the collection of Newton polytopes of the input polynomials. In this way, we avoid to compute the larger set of critical values at infinity and we also exploit the sparsity of the input polynomials. To the best of our knowledge, prior to our work, there was no dedicated algorithm to compute the infimum of real polynomial functions above.

The second (special) case is the unconstrained polynomial optimization problem, that is when $\mathcal{X} = \mathbb{R}^n$. An effective algorithm for approximating $\mathcal{K}_f^\infty$ (together with its arithmetic complexity analysis, and a method for computing f^*) was developed by Safey El Din [38]. As was pointed out by Jelonek and Tibăr [22], this would only provide a subset of $\mathcal{K}_f^\infty$. We present single exponential (bit and degree) bounds for f^* and a probabilistic algorithm to compute it, that is based on [20] and makes no assumptions on the input (Theorem 25) along with precise (single exponential) bit complexity estimates. Even though a method for computing f^* exists [38], to the best of our knowledge, neither bit/degree bounds were known before, nor precise bit complexity estimates for an algorithm to compute it; without any assumptions on the input.

In the rest of this section we present in detail the theorems that support our bounds for f^* and the corresponding algorithms, along with a bird's eye view of the proof techniques that we employ.

1.1 Presentation of the main results

We denote by $\mathcal{O}$, resp. $\mathcal{O}_B$, the arithmetic, resp. bit, complexity and we use the soft-O notations, $\widetilde{\mathcal{O}}$, respectively $\widetilde{\mathcal{O}}_B$, to ignore (poly-)logarithmic factors. We refer to § 1.3 for further details on the notation that we use.

Let $\mathcal{X} \subset \mathbb{R}^n$ be a semi-algebraic defined by a set of polynomials $\boldsymbol{f} \subset \mathbb{R}[x_1, \ldots, x_n]$. We call $\mathcal{X}$ *complete* if it is closed, connected, and for each subset $\boldsymbol{g} \subseteq \boldsymbol{f}$, the variety $\mathbb{V}(\boldsymbol{g}) = \{\boldsymbol{x} \in \mathbb{C}^n \mid g(\boldsymbol{x}) = 0, \text{ for all } g \in \boldsymbol{g}\} \subset \mathbb{C}^n$ is a complete intersection and smooth; the latter conditions means that the Jacobian of $\boldsymbol{g}$ evaluated at the points of $\mathbb{V}(\boldsymbol{g})$ has full rank.

The following theorem provides bit and degree bounds for the infimum of Problem 1 when $\mathcal{X}$ is complete.

Theorem 1. *Let $\mathcal{X}$ be a complete semi-algebraic set given by polynomial (in)equalities*

$$\mathcal{X} := \{g_1 = \cdots = g_r = 0,\ g_{r+1}, \ldots, g_s \geq 0,\ g_1, \ldots g_s \in \mathbb{R}[x_1, \ldots, x_n]\},$$

let $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a polynomial function of degree d_1, and assume that the infimum f^* of the function $f := F|_{\mathcal{X}} : \mathcal{X} \longrightarrow \mathbb{R}$ is not attained. Then f^* is an algebraic number of degree $\mathcal{O}((nrd_1)^{n^2})$, such that $2^{-\eta} \leq |f^*| \leq 2^\eta$, where $d = \max\{\deg g_1, \ldots, \deg g_s\}$ and

$$\eta := \widetilde{\mathcal{O}}(r(d+n+\tau)+d_1)(nrd_1)^{n^2}).$$

If we compare the bounds of the minimum induced by Eq. (1) and (2) with the bounds on the infimum of the previous theorem, then we notice that both bounds are single exponential with respect to the number of variables. However, the

bound of Theorem 1 admits a worse exponent, n^2 instead of n. It is worth noting that the dependence, in both cases, in the number of polynomials and the degrees is polynomial.

To give an overview of the approach that we follow to prove Theorem 1, let $X \subset \mathbb{R}^n$ be a real affine variety. Then, for every $z \in \mathbb{R}$ close enough to f^*, the preimage $f^{-1}(z)$ is empty if $z < f^*$. Hence, the infimum f^* is a point in $\mathcal{B}_f^\infty$. Since we have the inclusion

$$\mathcal{B}_f^\infty \subseteq \mathcal{K}_f^\infty, \tag{4}$$

we can compute $\mathcal{K}_f^\infty$ instead. Notice that the previous inclusion can be sharp for some functions [51, Example 3.8].

A straightforward analogue of the bifurcation set at infinity does not exist for arbitrary semi-algebraic sets $\mathcal{X}$. Because of this, we decompose $\mathcal{X}$ from Theorem 1 into finitely-many semi-algebraic subsets and consider their real Zariski closures, which we assume to be smooth and irreducible. For each such algebraic $X \subset \mathbb{R}^n$, we show that the infimum f^* belongs to the bifurcation set of the restricted polynomial function $F|_X$, where F is the canonical extension of f to the space $\mathbb{R}^n$ (Theorem 8). Then, we can approximate the bifurcation set at infinity by the Rabier set of $F|_X$. This way, we can then use effective methods developed in [21], to compute a superset of the Rabier set.

Using the above description, one can extract lower bounds on the infimum; roughly speaking, following [21], we express the Rabier set as the intersection of (the closure of) the image of a polynomial map with a linear subspace. This results in a univariate polynomial whose real roots include the infimum. Since the number of polynomials in the process is much larger than the number of variables, we should avoid using Gröbner basis computations, because it is difficult to bound their (bit) complexity and, even worse, they might induce double exponential bounds on the degree and bitsize of the resulting polynomial. Instead, we use resultant systems, e.g. [48, 53] which can express the image of the polynomial map using minors of a Macaulay-type matrix whose size exploits the single exponential bounds of the effective Nullstellensatz. In this way, we bound the degree and bitsize of the univariate polynomial that has the infimum among its real roots. A straightforward algorithm based on this approach still has double exponential complexity. However, our goal is not to compute the infimum, but to bound it.

The algorithms tailored for $\mathcal{K}_f^\infty$ [21] have prohibitive complexity, mainly because the require the computation of exponentially many minors of a Jacobian matrix (see §3 for details). This will result in double exponential complexity bound for an algorithm to compute f^*. However, if we impose some conditions on the semi-algebraic function, then might be able to overcome this computational obstacle. Indeed we present two such important special cases where this is possible and where not only we present improved (compared to Theorem 1) lower bounds, but also efficient algorithms for f^*, supported by precise bit complexity estimates.

Unconstrained polynomial optimization

Consider $\mathcal{X} = \mathbb{R}^n$. Then, we can approximate the Rabier set by considering the *complexification* $\mathbb{C}f : \mathbb{C}^n \longrightarrow \mathbb{C}$ of the polynomial function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, which simply extends the domain of f. Jelonek and Kurdyka in [20] presented one of the first methods for computing $\mathcal{K}_f^\infty$, while there are also other similar approaches and implementations [21, 37, 22]. Based on [20], we present an algorithm (Alg. 1) in §5, where we exploit resultants instead of Gröbner basis to perform algebraic elimination. Note that, once we compute $\mathcal{K}_f^\infty$, then we can identify the infimum by testing the non-emptiness of fibers in generic points of the intervals $\mathbb{R} \setminus \mathcal{K}_f^\infty$ [38]. A precise analysis of the degenerate conditions and bound on the bitsize of the algebraic objects involved in the computation of Alg. 1 lead to the following theorem for bounding and computing $f^* \in \mathcal{K}_f^\infty$.

Theorem 2 (Theorem 25)**.** *Let $f \in \mathbb{Z}[x]$ be of degree d and bitsize τ. Then, $f^* = \min_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x})$ is an algebraic number of degree $\mathcal{O}(d^{n-1})$, such that*

$$2^{-\eta} \leq |f^*| \leq 2^{\eta}, \quad \textit{where } \eta := \widetilde{\mathcal{O}}(nd^{n-1}\tau + n^2)).$$

There is an randomized algorithm that approximate f^* in $\widetilde{\mathcal{O}}_B(d^{3n}\lg(\frac{1}{\epsilon})(nd^{2n-2}\tau + \lg(\frac{1}{\epsilon})))$ bit operations, with probability of success $1 - \frac{1}{\epsilon}$.

Notice the lower bound on f^* in the unconstrained case, has an exponent n instead of n^2. This is similar to the bounds in Eqs. (1) and (2) where the infimum is attained, and hence it is asymptotically optimal. We emphasize that Theorem 2 holds without any assumptions on the input. If the infimum is attained, then it is among the critical values of f, which we can compute by solving the system of the partial derivatives of f. The complexity of this step is dominated by the complexity of the computation of the asymptotic critical value; hence we do not detail on this further.

Newton non-degenerate polynomial functions

Whenever $\mathcal{X} =: X$ is a (real) algebraic set, we present an algorithm for computing the optimal value f^*, under some genericity conditions on the input polynomials with respect to their Newton polytopes. We refer the reader to §4.1 for detailed presentation of polynomial sparsity and Newton polytopes.

A polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$ is a linear combination $\sum c_{\boldsymbol{a}}\, \boldsymbol{x}^{\boldsymbol{a}}$ of monomials $\boldsymbol{x}^{\boldsymbol{a}} := x_1^{a_1} \cdots x_n^{a_n}$, where the exponent vectors $\boldsymbol{a}$ run over a finite subset A of $\mathbb{N}^n$, and $c_{\boldsymbol{a}} \in \mathbb{K}^*$. Then A is the *support* of f and the *Newton polytope*, $\mathsf{NP}(f)$, of f is the convex hull of A in $\mathbb{R}^n$. Given a collection $\boldsymbol{\Delta} := (\Delta_0, \Delta_1, \ldots, \Delta_k)$ of integer polytopes in $\mathbb{R}^n_{\geq 0}$, we use $\mathbb{R}^{\boldsymbol{\Delta}}$ to denote the space of all tuples of polynomials whose *Newton tuple* is $\boldsymbol{\Delta}$:

$$\mathbb{R}^{\boldsymbol{\Delta}} := \{(f_0, f_1, \ldots, f_k) \;\mid\; f_i \in \mathbb{K}[z_1, \ldots, z_n],\ \mathsf{NP}(f_i) = \Delta_i,\ i = 0, 1, \ldots, k\}\,.$$

We can identify this space with the space of polynomial functions $X \longrightarrow \mathbb{R}$, with $\boldsymbol{x} \longmapsto f_0(\boldsymbol{x})$, where X is the real algebraic set $\mathbb{V}_{\mathbb{R}}(f_1, \ldots, f_k) \subset \mathbb{R}^n$, and $\mathsf{NP}(f_i) = \Delta_i$, $i = 0, 1, \ldots, k$.

The following theorem states that, for polynomials belonging to a Zariski open set, we can compute their bifurcation set at infinity $\mathcal{B}_f^\infty$ by exploiting their Newton polytopes.

Theorem 3 (Theorem 18)**.** *For every collection $\boldsymbol{\Delta}$ of lattice polytopes in $(\mathbb{R}_{\geq 0})^n$, there is a Zariski open subset $\Omega \subset \mathbb{K}^{\boldsymbol{\Delta}}$, such that for every $\boldsymbol{f} \in \Omega$, corresponding to a real polynomial function $f : X \longrightarrow \mathbb{R}$, the set $\mathcal{B}_f^\infty$ can be computed effectively. Furthermore, the values in $\mathcal{B}_f^\infty$ depend only on the coefficients of the polynomials whose exponent vectors appear at the faces of polytopes in $\boldsymbol{\Delta}$.*

The polynomial functions belonging to the open set Ω of the previous theorem are called *Newton non-degenerate*. Theorem 18 supports an algorithm for computing $\mathcal{B}_f^\infty$ that has several advantages compared with other "classical" methods. First, we compute a smaller superset of the bifurcation set at infinity. Secondly, the resulting algorithm applies for a dense family of polynomial functions sharing a collection of Newton polytopes. Finally, since the theorem and the supporting algorithms depends on the Newton polytopes of the input polynomials, we exploit the sparsity of the input and on our way to estimate the bifurcation set we compute with smaller polynomial systems. Ultimately, we obtain the following result.

Theorem 4 (§4.3)**.** *Let $\boldsymbol{\Delta}$ be a tuple of integer polytopes in $(\mathbb{R}_{\geq})^n$, and let $\boldsymbol{f} \in \mathbb{R}^{\boldsymbol{\Delta}}$ be a Newton non-degenerate element corresponding to a real polynomial function $f : X \longrightarrow \mathbb{R}$. Assume furthermore that all polynomials involved in X have degree at most d. Then,*

$$2^{-\eta} \leq |f^*| \leq 2^{\eta}, \quad \textit{where } \eta := \mathcal{O}(n^2 d^{n-1}(n+\tau)).$$

Similarly to Theorem 2, the bound on the infimum for Newton non-degenerate functions is single exponential in n and matches the bounds in Eqs. (1) and (2); hence it is asymptotically optimal.

1.2 Organization of the paper and proof strategies

The rest of the paper is structured as follows. The next section presents (some of) the notation that we use in throughout and some necessary preliminaries.

In §2, we will describe a decomposition $\mathcal{S}(\mathcal{X})$ of a complete semi-algebraic set, which we call *algebraic stratification*. We then show Theorem 8: For any semi-algebraic function $f : \mathcal{X} \longrightarrow \mathbb{R}$, one can construct a finite set of complex polynomial functions $\{f_s : V_s \longrightarrow \mathbb{C}\}_{s \in \mathcal{S}(\mathcal{X})}$ such that for each $s \in \mathcal{S}(\mathcal{X})$, we have $s \subset V_s$, and $f^* \in \mathcal{K}_{f_t}$ for some stratum t above. Now, since the subsets V_s are algebraic, the values in $\mathcal{K}_{f_s}$ can be computed effectively thanks to results of Jelonek and Kurdyka in [21]. These effective expressions will become useful in §3 for showing Theorem 1.

In §3 we consider the constrained optimization problem, Problem 1, under the assumption that the feasible region is a *complete* semi-algebraic set. We employ various tools to obtains precise bitsize and degree estimates for the infimum.

§4 is devoted to proving Theorem 4; we first describe the types of coherent faces in the tuple $\boldsymbol{\Delta}$ that are key for computing the bifurcation set. The proof relies on classical results on A-discriminants and face-resultants of polynomial tuples over the real and complex fields [18, 14, 15]. Once the expressions of the infimum are well-established in terms of an elimination ideal, we can then prove upper bounds on its absolute value.

In §5 we consider the unconstrained polynomial optimization problem. We present bounds on the infimum, an algorithm to compute it, and precise bit complexity estimates; under no assumptions on the input.

We mention auxiliary results that we need for the proofs of various results in the Appendix.

1.3 Notation

We denote by $\mathcal{O}$, resp. $\mathcal{O}_B$, the arithmetic, resp. bit, complexity and we use the soft-O notations, $\widetilde{\mathcal{O}}$, respectively $\widetilde{\mathcal{O}}_B$, to ignore (poly-)logarithmic factors. We denote by $\mathbb{R}^n_{\geq 0}$ the positive orthant and we use the abbreviation $[n]$ for $\{1, 2, \ldots, n\}$. We use bold letters to denote vectors.

A Monte Carlo algorithm is a randomized algorithm that its output might not be correct with a certain probability. A Las Vegas randomized algorithm always outputs the correct result, but its runtime is not always the same, even for the same input.

Algebraic varieties and smooth maps Consider $\boldsymbol{x} = (x_1, \ldots, x_n)$ and let $\boldsymbol{x}_{-i}$ denote all the variables except the variable x_i. For a polynomial $f \in \mathbb{K}[x_1, \ldots, x_n] = \mathbb{K}[\boldsymbol{x}]$, where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, we denote its zero set by $\mathbb{V}_{\mathbb{K}}(f) \subset \mathbb{K}^n$ and by $\mathbb{V}_{\mathbb{K}^*}(f) \subset (\mathbb{K}^*)^n$ its zero set over the corresponding torus. We use the same notation after replacing f by its bold form $\boldsymbol{f}$ if it is a tuple of polynomials, that is $\boldsymbol{f} = (f_i)_{i \in I}$ for some finite subset $I \subset \mathbb{N}$.

Let $f : X \longrightarrow Y$ be any smooth, or analytic, map between two manifolds. Let $\mathcal{G}$ be the graph in $X \times Y$ of f and consider the map

$$P := \pi_{\mathcal{G}} : \begin{array}{ccc} \mathcal{G} & \longrightarrow & Y \\ (\boldsymbol{x},\, \boldsymbol{y}) & \longmapsto & \boldsymbol{y}\ , \end{array}$$

where $\pi_{\mathcal{G}}$ is the restriction of $\pi : X \times Y \longrightarrow Y$, $(\boldsymbol{x},\ \boldsymbol{y}) \longmapsto \boldsymbol{y}$ on $\mathcal{G}$. We say that f is a $\mathcal{C}^\infty$-*fibration*, if for every $\boldsymbol{y} \in \pi(\mathcal{G})$, the set $Y \times f^{-1}(\boldsymbol{y})$ is diffeomorfic to $P^{-1}(\pi(\mathcal{G}))$. The map f is said to be a *locally trivial fibration* if, for each $\boldsymbol{x} \in X$, there exists a neighborhood U of x for which the restricted map $f|_U : U \longrightarrow Y$ is a $\mathcal{C}^\infty$-fibration.

A point $\boldsymbol{x} \in X$ is said to be a *critical point of* f if the *co-rank* $\operatorname{corank}(df)_{\boldsymbol{x}} := \min(\dim X,\ \dim Y) - \operatorname{rank}(df)_{\boldsymbol{x}}$ is positive. We denote by $\operatorname{Crit}(f)$ the set of critical points of f and by $\mathcal{K}^0_f(f)$ the corresponding critical values.

If, instead, X is a smooth variety defined by polynomials $g_1, \ldots, g_p \in \mathbb{K}[x_1, \ldots, x_n]$, then the critical locus of f consists of all points $\boldsymbol{x} \in X$ for which the rank of the following *Jacobian matrix*

$$\operatorname{Jac} \boldsymbol{f} := \begin{bmatrix} \partial f_1/\partial x_1 & \cdots & \partial f_k/\partial x_1 & \partial g_1/\partial x_1 & \cdots & \partial g_p/\partial x_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ \partial f_1/\partial x_n & \cdots & \partial f_k/\partial x_n & \partial g_1/\partial x_n & \cdots & \partial g_p/\partial x_n \end{bmatrix}$$

is lower than the codimension of X.

About polynomials and their roots For a polynomial $f \in \mathbb{Z}[\boldsymbol{x}] = \mathbb{Z}[x_1, \ldots, x_n]$ its infinity norm $\|f\|_\infty$ equals the maximum of absolute values of its coefficients. The bitsize of a polynomial is the logarithm of its infinity norm. We also call the latter the bitsize of the polynomial, that is a shortcut for the maximum bitsize of all its coefficients. A univariate (multivariate) polynomial is of size (d, τ) when its (total) degree is at most d and has bitsize τ. We represent a real algebraic number $\alpha \in \mathbb{R}$ using the *isolating interval representation*; it includes a square-free polynomial, A, which vanishes at α and an interval with rational endpoints that contains α and no other root of A. If $\alpha \in \mathbb{C}$, then instead of an interval we use a rectangle in $\mathbb{R}^2$ where the coordinates of its vertices are rational numbers.

2 Critical values at infinity and bounds on the infimum

We present a decomposition of a semi-algebraic set, under some transversality conditions. Then, we relate the infimum of a polynomial function, say f, restricted on this semi-algebraic set, with the bifurcation set(s) of f, restricted to the strata of the decomposition.

2.1 The bifurcation and Rabier sets

Let $X \subset \mathbb{K}^n$ be a smooth affine variety, where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, and let $f : X \longrightarrow \mathbb{K}$ be a polynomial function on X. That is, if $X = \mathbb{V}_{\mathbb{K}}(\mathcal{I})$ for some ideal $\mathcal{I}$, then $f \in \mathbb{K}[x_1, \ldots, x_n]/\mathcal{I}$. A generalization [52, 49, 50, 46] of Thom's result [45] indicates that outside a finite set $S \subset \mathbb{K}$, the following restricted function is a $\mathcal{C}^\infty$-fibration:

$$f|_{f^{-1}(\mathbb{K} \setminus S)} : f^{-1}(\mathbb{K} \setminus S) \longrightarrow \mathbb{K} \setminus S. \tag{5}$$

The smallest subset $S \subset \mathbb{K}$ for which the function in (5) is a $\mathcal{C}^\infty$-fibration is the *bifurcation set* of f and we denote it by $\mathcal{B}_f$. The *bifurcation values* $z \in \mathcal{B}_f$ are of two types: (i) $z = f(\boldsymbol{x})$ for some $\boldsymbol{x} \in \operatorname{Crit}(f)$; we denote this set by $\mathcal{B}^0_f$

(also $\mathcal{K}_f^0$), or (ii) z is such that for an arbitrarily large compact subset $K \subset \mathbb{C}^n$ and any small disc D containing z, the restricted map

$$f|_{f^{-1}(D)\setminus K} : f^{-1}(D) \setminus K \longrightarrow D$$

is not a $\mathcal{C}^\infty$-fibration (see e.g., [46, Definition 2.1]). The set of all such values is denoted by $\mathcal{B}_f^\infty$, and we call it the *bifurcation set at infinity*. Then, we get $\mathcal{B}_f^0 = \mathcal{B}_f^\infty \cup \mathcal{B}_f^0$. We notice that the inclusion $\mathcal{B}_f^0 \subset \mathcal{B}_f$ can be strict. For example, the function $(x,\ y) \longmapsto x + x^2y$ has no critical points [11], whereas 0 is a bifurcation value as $f^{-1}(0)$ is the only fiber with more than one connected component in $\mathbb{K}^2$.

Recall the Rabier set defined at the beginning in (3). We define $\mathcal{K}_f := \mathcal{K}_f^\infty \cup \mathcal{K}_f^0$ to be the set of *generalized critical values* of f. Consequently, f is locally trivial fibration over the $\mathbb{K} \setminus \mathcal{K}_f$.

In §3 and §5 we compute $\mathcal{K}_f$ using algebraic elimination.

2.2 A decomposition of a (complete) semi-algebraic set

Consider a basic closed semi-algebraic set

$$\mathcal{X} := \{g_1 = 0, \ldots, g_r = 0,\ g_{r+1} \geq 0, \ldots, g_s \geq 0,\ g_1, \ldots g_s \in \mathbb{R}[x_1, \ldots, x_n]\}, \tag{6}$$

and let $f : \mathcal{X} \longrightarrow \mathbb{R}$ be a semi-algebraic function given as the restriction of a polynomial function $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ on $\mathcal{X}$, that is

$$f :=\ F|_{\mathcal{X}} : \mathcal{X} \longrightarrow \mathbb{R}.$$

Theorem 8 demonstrates that the infimum f^* of f over $\mathcal{X}$ is in the union of bifurcation sets of restricted functions $F|_S$, for some algebraic sets $S \subset \mathbb{R}^n$.

This leads us to consider the *complexification* of these functions $\mathbb{C}F|_{\mathbb{C}S} : \mathbb{C}S \longrightarrow \mathbb{C}$, where we extend the domain and range of $F|_S$ to $\mathbb{C}S \subset \mathbb{C}^n$ and $\mathbb{C}$, respectively. Subsequently, we use the inclusion of Eq. (4) to show that f^* lies in the union of the corresponding Rabier sets (Corollary 9). In this way we can compute and/or approximate f^*.

Let $I = \{1, \ldots, r\} \cup I_0$, where I_0 is a subset of $s + [r - s]$; that is I is a subset of $[r]$ that always contains the set $\{1, \ldots, r\}$. For any such I, we denote by $\mathbb{V}_{\mathbb{K}}(\boldsymbol{g}_I)$ the common zero locus in $\mathbb{K}^n$ of g_i, for all $i \in I$, where $\mathbb{K} \in \{\mathbb{C}, \mathbb{R}\}$.

Definition 5 (Complete semialgebraic set)**.** *A semialgebraic set $\mathcal{X}$, as in* (6)*, is* complete *if, for every set of indices I, as above, the set $\mathbb{V}_{\mathbb{R}}(\boldsymbol{g}_I)$ is a smooth complete intersection of codimension $\#I$. Then, there exists $\alpha \in \mathbb{N}$ and a filtration*

$$\varnothing =: \mathcal{X}_{-1} \subset \mathcal{X}_\alpha \subset \mathcal{X}_{\alpha+1} \subset \cdots \subset \mathcal{X}_{n-r} := \mathcal{X}, \tag{7}$$

satisfying the following properties for every $i = \alpha, \ldots, n - r$:

- *(Π_1) The set $\mathcal{X}_i$ is a basic closed semi-algebraic set with pure dimension i, and*
- *(Π_2) the Zariski closure of any connected component of $\mathcal{X}_i \setminus \mathcal{X}_{i-1}$, is $\mathbb{V}_{\mathbb{R}}(\boldsymbol{g}_I)$, where I is a subset of $[s]$, such that $[r] \subset I$ and $\#I = n - i$.*

Since $\mathcal{X}$ is complete, we have $\alpha = \max(0, n - s)$.

Definition 6 (Stratification of a complete set $\mathcal{X}$)**.** *If $\mathcal{X}$ is complete semi-algebraic subset, then $\mathcal{S}(\mathcal{X}) := \{\mathcal{X}_i \mid \alpha \leq i \leq n - r\}$ is called an* algebraic stratification *of $\mathcal{X}$; we call the elements of $\mathcal{S}(\mathcal{X})$* strata.[3]

Notation 7. *For each such $I \subset [s]$, we use the shorthand notation $F|_I$ for the restricted function*

$$F|_{\mathbb{V}_{\mathbb{R}}(\boldsymbol{g}_I)} : \mathbb{V}_{\mathbb{R}}(\boldsymbol{g}_I) \longrightarrow \mathbb{R}.$$

Theorem 8. *Let $\mathcal{X}$, as in Eq.* (6)*, be a complete semialgebraic subset of $\mathbb{R}^n$. Let $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a polynomial function, and assume that the infimum f^* of the function $f := F|_{\mathcal{X}} : \mathcal{X} \longrightarrow \mathbb{R}$ is not attained. Then, there exists a set of indices I, such that $[r] \subset I \subset [s]$, for which*

$$f^* \in \mathcal{B}_{F|_I}. \tag{8}$$

Proof. Assume without loss of generality that $f^* = 0$. We will prove (8) by finding I above for which the zero betti number of $F|_I^{-1}(0)$ differs from that of any $F|_I^{-1}(z)$ with $z \in \mathbb{R}$ close enough to 0.

[3] To the best of our knowledge, in its current form, the definition of a complete stratification does not appear in the literature.

We start by choosing I above. Since f^* is not attained, and $\mathcal{X}$ is closed (in the Euclidean topology), there exists $\varepsilon > 0$ such that all components of $f^{-1}(]0,\ \varepsilon[)$ are unbounded. We choose ε small enough so that $f^{-1}(]0,\ z[)$ has the same number of connected components in $\mathbb{R}^n$ for any $z \in]0,\ \varepsilon[$. Furthermore, one can set ε even smaller if necessary so that each connected component C of $f^{-1}(]0,\ \varepsilon[)$ coincides with a connected component of $\sigma \cap F^{-1}(]0,\ \varepsilon[)$ for some stratum $\sigma \in \mathcal{S}(X)$.

Pick one such stratum σ above, together with a component C, and let $J \subset [s]$ be the index set containing $[r]$ such that $\mathbb{V}_{\mathbb{R}}(\boldsymbol{g}_J)$ is the Zariski closure of σ (c.f. (Π_2)). Since $\mathcal{X}$ is complete, we can choose a sub-stratum $\delta \subset \sigma$ and a superset I containing J such that a connected component of $\mathbb{V}_{\mathbb{R}}(\boldsymbol{g}_I) \cap F^{-1}(]0,\ \varepsilon[)$ coincides with $\delta \cap F^{-1}(]0,\ \varepsilon[)$. In what follows, we set $Z := \mathbb{V}_{\mathbb{R}}(\boldsymbol{g}_I)$ and $S := \delta \cap F^{-1}(]0,\ \varepsilon[)$.

Now, we use betti numbers counting of preimages under $F|_Z$ to finish the proof. Recall that $\mathcal{B}_{F|_Z}$ is finite in $\mathbb{R}$. Then, we may take a smaller $\varepsilon > 0$ if necessary so that

$$]0,\ \varepsilon[\ \cap\ \mathcal{B}_{F|_Z} = \varnothing \tag{9}$$

is satisfied. Hence, there exists $k \in \mathbb{N}$, such that for every $z \in]0,\ \varepsilon[$, the zero betti number satisfies

$$b_0\big(F|_Z^{-1}(z)\big) = k. \tag{10}$$

We finish the proof by contradiction; assume that $0 \notin \mathcal{B}_{F|_Z}$. Then, we have $b_0\big(F|_Z^{-1}(0)\big) = k$. Since S is a connected component of $F|_Z^{-1}([0,\ \varepsilon[)$, for each z in the half-closed interval $[0,\ \varepsilon[$, we get

$$b_0\big(F|_Z^{-1}(z)\big) = b_0\big(F|_{Z\setminus S}^{-1}(z)\big) + b_0\big(F|_S^{-1}(z)\big).$$

Then, Equality (10) implies that

$$b_0\big(F|_{Z\setminus S}^{-1}(z)\big) + b_0\big(F|_S^{-1}(z)\big) = b_0\big(F|_{Z\setminus S}^{-1}(0)\big) + b_0\big(F|_S^{-1}(0)\big).$$

Since $F|_S^{-1}(0) = f|_S^{-1}(0) = \varnothing$, we get that for each $z \in]0,\ \varepsilon[$, the below equality holds:

$$b_0\big(F|_{Z\setminus S}^{-1}(0)\big) > b_0\big(F|_{Z\setminus S}^{-1}(z)\big). \tag{11}$$

Furthermore, from the definitions of S and Z, we get that $F|_{Z\setminus S}^{-1}(z)$ is algebraic for each $z \in\ [0,\ \varepsilon[$. Therefore, inequality (11) implies that there exists $\boldsymbol{x} \in F|_{Z\setminus S}^{-1}(0)$ such that $\boldsymbol{x} \in \mathrm{Crit}(F|_Z)$. This contradicts $0 \notin \mathcal{B}_{F|_Z}$. □

In what follows, for every $I \subset [s]$, we use the shorthand notation $F|_{\mathbb{C}I}$ for the complexification of the real restricted map $F|_I$. That is,

$$F|_{\mathbb{C}I} := \mathbb{C}F|_{\mathbb{V}_{\mathbb{C}}(\boldsymbol{g}_I)} : \mathbb{V}_{\mathbb{C}}(\boldsymbol{g}_I) \longrightarrow \mathbb{C}.$$

Corollary 9. *Let $\mathcal{X}$, F, and f be as in Theorem 8. Then, there exists a set of indices I, such that $[r] \subset I \subset [s]$, for which it holds*

$$f^* \in \mathcal{K}_{F|_{\mathbb{C}I}}.$$

Proof. Theorem 8, together with the Rabier property (3) imply that

$$f^* \in \mathcal{B}_{F|_I} \subset \mathcal{K}_{F|_I},$$

for some I satisfying $[r] \subset I \subset [s]$. Finally, since we have $\|z\|_{\mathbb{C}} = \|z\|_{\mathbb{R}}$ for any $z \in \mathbb{R}^n$, we get

$$\mathcal{K}_{F|_I} \subset \mathcal{K}_{F|_{\mathbb{C}I}}, \tag{12}$$

for every $I \subset [s]$. This completes the proof. □

3 Constrained optimization and asymptotic critical values

We provide bounds on the absolute values of points in the Rabier set of polynomial functions $f : X \longrightarrow \mathbb{C}$ over smooth affine varieties $X \subset \mathbb{C}^n$. Whenever f is given by real polynomials, this eventually leads to bounds for the optimization problem $\inf_{\boldsymbol{x} \in X \cap \mathbb{R}^n} f(\boldsymbol{x})$.

3.1 Asymptotic critical values over a smooth variety $X \subset \mathbb{C}^n$

Consider a polynomial $F \in \mathbb{Q}[x_1, \ldots, x_n]$ and a smooth algebraic variety $X \subseteq \mathbb{C}^n$ of dimension $\delta = n - r$, that is the zero set of the polynomials $\boldsymbol{g} = \{g_1, \ldots, g_r\} \in \mathbb{Q}[\boldsymbol{x}]$. As our goal is to bound the asymptotic critical values of the function $f := F|_X$, it is without loss of generality to consider X to be complete intersection; see Theorem 29 and Remark 4. If we have more than r polynomials, then we can consider r generic linear combinations of them to construct a "complete intersection" system, at the expense of adding some additional component(s) in X. The original components of the initial system can be identified by considering several generic linear combinations if necessary. As this technique does not change a lot the bitsize of the polynomial to compute with, and we only interested to bound (and not compute) the related quantities, we do not detail further. We refer the reader to [9] for additional details and an application in computing the Chow form of a variety.

Consider the $(r+1) \times n$ matrix

$$C = \begin{pmatrix} \nabla F \\ \mathrm{Jac}(\boldsymbol{g}) \end{pmatrix},$$

where ∇F is the gradient of F and has dimension $1 \times n$, while $\mathrm{Jac}(\boldsymbol{g})$ is the Jacobian matrix of $\boldsymbol{g}$ and has dimension $r \times n$. Consider the set of all subsets of $[n]$ of cardinality $r+1$; we denoted by $\left[\begin{smallmatrix} n \\ r+1 \end{smallmatrix}\right]$. Then, $I \in \left[\begin{smallmatrix} n \\ r+1 \end{smallmatrix}\right] \Leftrightarrow I \subseteq [n]$ and $|I| = r+1$. There are $s = \binom{n}{r+1}$ such subsets I and to each of them, we associate an integer i, where $1 \leq i \leq s$.

Let M_i be the $(r+1) \times (r+1)$ square submatrix of C, obtained by selecting the columns of C with indices in I. For $j \in I$, let $M_{i,j}$ be the square $r \times r$ submatrix of M_i, obtained by omitting the first row (which corresponds to ∇F) and the j-th column; there are $r+1$ such matrices, for every i.

For a specific i and $j \in I$, let m_i and $m_{i,j}$ be the determinants of M_i and $M_{i,j}$, respectively; they are polynomials in $\mathbb{Z}[\boldsymbol{x}]$. We also need the definition of the rational function

$$w_{i,j} = \frac{m_i}{m_{i,j}} = \frac{\det M_i}{\det M_{i,j}} \ .$$

Consider the vector $\boldsymbol{j} = (j_1, \ldots, j_s) \in \mathbb{N}^s$, where $j_i \in I$. There are $(r+1)^s$ such vectors. For each $\boldsymbol{j}$, we consider the map $\Phi_{\boldsymbol{j}}$, which is

$$\begin{array}{rccl} \Phi_{\boldsymbol{j}} & : & X & \longrightarrow & \mathbb{C}^n \times \mathbb{C}^{(n+1)s} \\ & & \boldsymbol{x} & \longmapsto & (F(\boldsymbol{x}), \{w_{i,j_i}(\boldsymbol{x}), x_1 w_{i,j_i}(\boldsymbol{x}), \ldots, x_n w_{i,j_i}(\boldsymbol{x})\}_{i \in [s]}). \end{array} \tag{13}$$

If $\mathcal{G}_{\boldsymbol{j}} := \overline{\mathtt{Im}(\Phi_{\boldsymbol{j}})}$, and $\mathcal{G} = \bigcup_{\boldsymbol{j}} \mathcal{G}_{\boldsymbol{j}}$, then $\mathcal{K}(f) = L \cap \Gamma$, where $L = \underbrace{(0, \ldots, 0)}_{(n+1)s \text{ times}} \times \mathbb{C}$ (see [21]).

Upper and lower bounds on the roots of these polynomials, will also hold truer for the asymptotic critical values.

3.1.1 Resultant systems using determinants

Our presentation is based on van der Waerden [48] and Yap [53]. Let $\boldsymbol{A} = \{A_1, \ldots, A_p\} \subseteq (\mathbb{Z}[\boldsymbol{c}])[\boldsymbol{x}]$ be a system of polynomial equations in the variables $\boldsymbol{x}$, the coefficients of which are polynomials with integer coefficients in the additional set of variables $\boldsymbol{c}$. Let the degree of each polynomial A_i with respect to $\boldsymbol{x}$ be d_i and $d = \max_{i \in [p]} d_i$.

The resultant system is a set of polynomials in the coefficients $\boldsymbol{c}$ having the property that all of them vanish if the polynomials in $\boldsymbol{A}$ have a common non-trivial solution. In the case where $p = n$, then the resultant system consists of a single polynomial, which we call the (homogeneous) resultant of the system.

Let $\nu_m = \binom{m+n-1}{n-1} \leq m^{n-1}$, where m is a positive integer that we will specify in the sequel. Let $\mathtt{P}^m = \{\boldsymbol{x}^{\boldsymbol{\alpha}} \,|\, |\boldsymbol{\alpha}| = m\}$, that is the set of all monomials in $\boldsymbol{x}$ of degree m. Next, we consider the set

$$\boldsymbol{A}_m = \{\boldsymbol{x}^{\boldsymbol{\alpha}} A \mid \deg(\boldsymbol{x}^{\boldsymbol{\alpha}} A) = m, A \in \boldsymbol{A}, \boldsymbol{x}^{\boldsymbol{\alpha}} \in \mathtt{P}^m\} \ .$$

It holds $|\boldsymbol{A}_m| \geq \nu_m$. Let M_m be the $(|\boldsymbol{A}_m| \times \nu_m)$-matrix, the rows of which correspond to the polynomials in $\boldsymbol{A}_m$ and its columns correspond to the elements of $\mathtt{P}^m$.

Let $R_m \subseteq \mathbb{Z}[\boldsymbol{c}]$ be the set of all the $(\nu_m \times \nu_m)$-minors of the matrix M_m.

Theorem 10 ([53, XI Thm. 9 & 13])**.** *If $m \geq 1 + n(13\, d^n - 1)$, then R_m is a resultant system for $\boldsymbol{A}$.*

The $13\, d^n$ factor in the bound for m corresponds to an effective bound on the Nullstellensatz. We note that we can also use improved bounds of effective Nullstellensatz, but this bound suffices for our purposes. For our purposes, we do not need to count or manipulate with the polynomials in R_m. We just need to bound their degree, as polynomials in $\boldsymbol{c}$, and their bitsize.

3.1.2 Bounding the asymptotic critical values using resultant systems

To obtain worst case bound on the asymptotic critical values, it suffices to consider worst case bounds for $\mathcal{G}_{\boldsymbol{j}}$, for a fixed pair $(\boldsymbol{j}, I)$, with $I \in [\,_{r+1}^{n}]$ and $\boldsymbol{j} = (j_1, \ldots, j_s) \in \mathbb{N}^s$ satisfying $j_i \in I$. We emphasize that our goal is *not* to compute these values, but rather to bound them effectively.

Let $X_{\boldsymbol{j}} = \mathbb{V}(\{m_{i,j_i}\}_{i\in[s]}) \subseteq \mathbb{C}^n$ be the zero set of all denominators in $\Phi_{\boldsymbol{j}}$, Eq. (13). Let $\delta_i = \dim(X_{\boldsymbol{j}})$ be the dimension $X_{\boldsymbol{j}}$. As it is not a complete intersection, we can consider $n - \delta_i$ generic linear combinations of the polynomials m_{i,j_i}. Let these new polynomials be $\{h_1, \ldots, h_{n-\delta_i}\}$.

Consider the set of polynomials

$$J = \Big\{ \{g_i(\boldsymbol{x})\}_{i\in[r]}, F(\boldsymbol{x}) - z, \\ \{m_{i,j_i}(\boldsymbol{x})y_i - m_i(\boldsymbol{x}), m_{i,j_i}(\boldsymbol{x})y_i - x_1 m_i(\boldsymbol{x}), \ldots, m_{i,j_i}(\boldsymbol{x})y_i - x_n m_i(\boldsymbol{x})\}_{i\in[s]}, \\ t\,h(\boldsymbol{x}) - 1 \Big\} \subseteq \mathbb{Z}[\boldsymbol{x}, t, y_1, \ldots, y_s, z], \quad (14)$$

where $h(\boldsymbol{x}) = \prod_{i\in[n-\delta_i]} h_i(\boldsymbol{x})$.

This set contains $r + 1 + (n + 1)\binom{n}{r+1} + 1$ equations. We should eliminate the variable $(\boldsymbol{x}, t)$; this would result in polynomials in $\mathbb{Z}[y_1, \ldots, y_s, z]$. Then, if we set $y_i = 0$, we obtain univariate polynomials in z, the roots of which contains the asymptotic critical values of f.

Let F be of size (d, τ) and g_i of size (d_1, τ). Standard calculations based on Claim 26 result that $m_{i,j}$ is of size $(\mathcal{O}(rd_1), \widetilde{\mathcal{O}}(r(d + n + \tau)))$, m_i is of size $(\mathcal{O}(d + rd_1), \widetilde{\mathcal{O}}(r(d + n + \tau) + d_1))$.

Based on the results in the appendix, the polynomials h_i have the same size as the polynomials $m_{i,j}$. Thus, $h(\boldsymbol{x})$ is of size $(\mathcal{O}(nrd_1), \widetilde{\mathcal{O}}(nr(d + n + \tau)))$,

To eliminate the variables $\boldsymbol{x}$ from the polynomials in J, we will use resultant systems from Sec. 3.1.1. The maximum degree of the involved polynomials is $\mathcal{O}(nrd_1)$, thus the Nullstellensatz bound becomes $\mathcal{O}((nrd_1)^n)$ which in turn implies that the various matrices are of dimension $\mathcal{O}((nrd_1)^{n^2})$.

The elements of the matrix (or matrices) are polynomials in the variables $y_1, \ldots, y_s, z$. Their degree with respect to y_i and z is one. Their maximum bitsize is $\widetilde{\mathcal{O}}(r(d + n + \tau) + d_1))$. Hence, their determinant is a polynomial in $(\mathbb{Z}[y_1, \ldots, y_s])[z]$ of degree $\mathcal{O}((nrd_1)^{n^2})$. and bitsize

$$\eta = \widetilde{\mathcal{O}}(r(d + n + \tau) + d_1)(nrd_1)^{n^2}).$$

If we set the variables y_i to zero, then we obtain a univariate polynomial in $\mathbb{Z}[z]$ of the same size, the roots of which contain the asymptotic critical values.

The arguments above yield the following result.

Theorem 11. *Let $F \in \mathbb{Z}[\boldsymbol{x}]$ be of degree d and bitsize τ. Let $X = \mathbb{V}(g_1, \ldots, g_r) \cap \mathbb{R}^n$ be a smooth real algebraic variety and g_i be of size (d_1, τ). The optimum value of the problem $f^* = \inf_{\boldsymbol{x}\in X} F(\boldsymbol{x})$ is an algebraic number of degree $\mathcal{O}((nrd_1)^{n^2})$ such that $2^{-\eta} \leq |f^*| \leq 2^{\eta}$, where $\eta = \widetilde{\mathcal{O}}(r(d + n + \tau) + d_1)(nrd_1)^{n^2})$.*

Finally, thanks to Corollary 9, we deduce our min Theorem 1.

4 The Newton non-degenerate case

As Theorem 8 presents, in order to locate the infimum of a semi-algebraic function, one can consider a polynomial function instead, and compute its bifurcation set. The latter, however, is intractable for arbitrary functions. Accordingly, we can use the Rabier set to effectively compute a superset containing it. Albeit the computation itself has high complexity as we will see in §3.

In this section, we introduce a large family of polynomial functions, and show that one can effectively approximate their bifurcation set using less expensive methods than the Rabier set.

4.1 Preliminaries on polytopes

A subset $\Pi \subset \mathbb{R}^n$ is called a *polyhedron* if it is the intersection of finitely-many closed half-spaces. The boundary of one such half-space is called a *supporting hyperplane* of Π. We say that a set Φ is a *face* of Π, we indicate this using

the notation $\Phi \prec \Pi$, if it is the intersection of a supporting hyperplane H of Π with its boundary, i.e., $\Phi = H \cap \partial\Pi$. We say that Φ is *origin* if Φ contains the point $\mathbf{0} := (0, \ldots, 0) \in \mathbb{R}^n$. A *polytope* is a bounded polyhedron.

A *tuple of polytopes*, or a *tuple* for short, is a map $\boldsymbol{\Delta}$ from a finite set $K \subset \mathbb{N}$ to the set of polytopes in $\mathbb{R}^n$. We call K the *support* of $\boldsymbol{\Delta}$ and we denote it by $[\boldsymbol{\Delta}]$. The Minkowski sum of elements in $\boldsymbol{\Delta}$

$$\left\{ \sum\nolimits_{k \in K} \boldsymbol{a}_k \,|\, \boldsymbol{a}_k \in \Delta_k \right\},$$

is also a polytope; we denote it by $\sum \boldsymbol{\Delta}$. In our setting, the *Minkowski sum* of any two subsets $X, Y \subset \mathbb{R}^n$ is the coordinate-wise sum $X + Y := \{x + y \mid x \in X,\ y \in Y\}$. The *dimension* of $\boldsymbol{\Delta}$ is defined as

$$\dim \boldsymbol{\Delta} := \dim(\sum \boldsymbol{\Delta}) - \#[\boldsymbol{\Delta}]. \tag{15}$$

For any $i \in [\boldsymbol{\Delta}]$, the i-th polytope in the tuple is Δ_i. For any $I \subset [\boldsymbol{\Delta}]$, $\boldsymbol{\Delta}_I := \{\Delta_i \mid i \in I\}$ is the *sub-tuple* (of $\boldsymbol{\Delta}$ associated to I). A tuple $\boldsymbol{\Gamma}$ is a *tuple-face* of $\boldsymbol{\Delta}$ (or, simply, *face* whenever it is clear from the context) if $[\boldsymbol{\Gamma}] = [\boldsymbol{\Delta}]$, $\Gamma_i \prec \Delta_i$, for each $i \in [\boldsymbol{\Gamma}]$, and $\sum \boldsymbol{\Gamma} \prec \sum \boldsymbol{\Delta}$.

We say that $\boldsymbol{\Gamma}$ is a *facing* of $\boldsymbol{\Delta}$, if there is a face $\boldsymbol{\Gamma}'$ of $\boldsymbol{\Delta}$ such that $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}'_I$ and $I = [\boldsymbol{\Gamma}]$. We also use the notation $\boldsymbol{\Gamma} \prec \boldsymbol{\Delta}$ for the facings.

Definition 12. *A facing $\boldsymbol{\Gamma} \prec \boldsymbol{\Delta}$ is said to be* important, *if there exists a face $\boldsymbol{\Gamma}' \prec \boldsymbol{\Delta}$ such that $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}'_{[\boldsymbol{\Gamma}]}$ and*

$$\dim \boldsymbol{\Gamma} \leq \dim \boldsymbol{\Gamma}'_I, \quad \forall\, I \supset [\boldsymbol{\Gamma}].$$

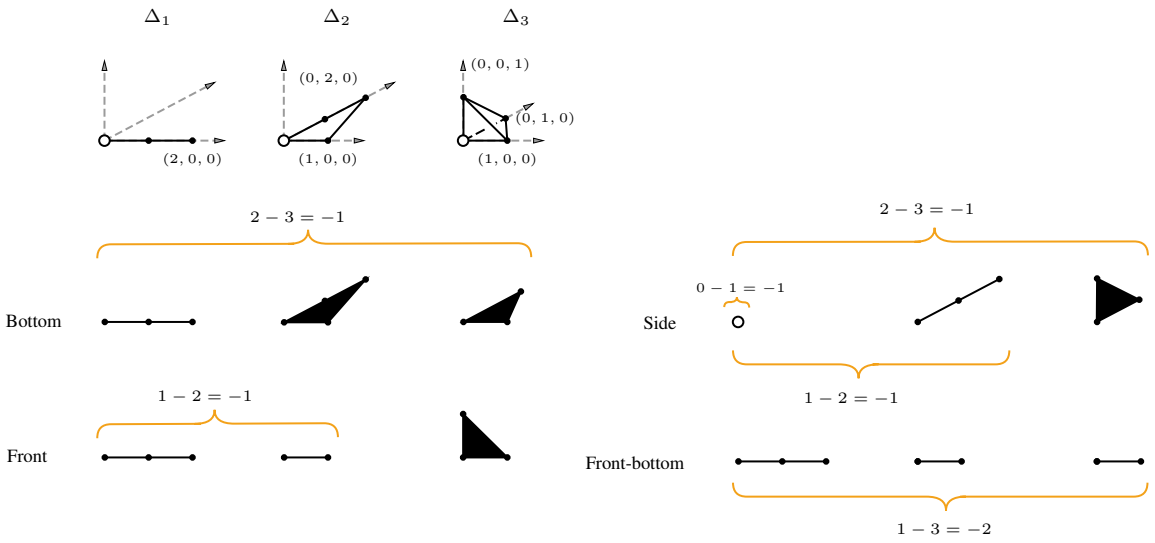

Figure 1: An example of a triple $\boldsymbol{\Delta}$. For each of the four selected faces of $\boldsymbol{\Delta}$, we indicate which are the important facings, and we compute their dimensions according to the formula (15).

For any polytope $\Pi \subset \mathbb{R}^n_{\geq 0}$, either $\{\mathbf{0}\}$ is a vertex of Π, or $\mathbf{0} \notin \Pi$.

Definition 13. *A tuple $\boldsymbol{\Delta}$ is said to be* origin *if its support $[\boldsymbol{\Delta}]$ contains the index 0 and the polytope Δ_0 contains the origin $\{\mathbf{0}\}$.*

For example, all important facings illustrated in Figure 1 are origin.

4.2 Application to polynomial functions

Let $\mathbb{K} \in \{\mathbb{C}, \mathbb{R}\}$. We consider the polynomial function $f : X \longrightarrow \mathbb{K}$, for some $X := \mathbb{V}_{\mathbb{K}}(g_1, \ldots, g_r) = \mathbb{V}_{\mathbb{K}}(\boldsymbol{g})$. If f is obtained as the restriction $F|_X = f$, for some polynomial $F : \mathbb{K}^n \longrightarrow \mathbb{K}$, then, we identify f with the collection of polynomials $\boldsymbol{f} := (F, g_1, \ldots, g_r) = (F, \boldsymbol{g})$.

Let $\boldsymbol{\Delta}$ denote the tuple of lattice polytopes given by $(\Delta_0, \Delta_1, \ldots, \Delta_r)$, where

$$\Delta_i := \mathsf{NP}(g_i),\ i = 1, \ldots, r,$$

and $\Delta_0 := \mathsf{NP}(F - t)$ for some generic $t \in \mathbb{K}$. That is, it holds that

$$\Delta_0 := \mathrm{conv}(\mathrm{supp}(F) \cup \{\mathbf{0}\}).$$

Notation 14. *For any subset $\sigma \subset \mathbb{R}^n$ and for any polynomial $P \in \mathbb{K}[x_1, \ldots, x_n]$ the* restriction of P to σ*, denoted by P_σ, is the polynomial*

$$\sum_{\boldsymbol{a} \in \sigma \cap \mathrm{supp}(P)} c_{\boldsymbol{a}}\, \boldsymbol{x}^{\boldsymbol{a}}.$$

For any facing $\boldsymbol{\Gamma} \prec \boldsymbol{\Delta}$, and any $\boldsymbol{f} \in \mathbb{K}^{\boldsymbol{\Delta}}$, we define the collection of polynomials $\boldsymbol{g}_{\boldsymbol{\Gamma}} := (\boldsymbol{g}_{\Gamma_i})_{i \in [\boldsymbol{\Gamma}] \setminus \{0\}}$, where for any $i \in [\boldsymbol{\Gamma}] \setminus \{0\}$, the expression $\boldsymbol{g}_{\Gamma_i}$ denotes the restricted polynomial g_{i,Γ_i}. Similarly, we use the notation $\boldsymbol{f}_{\boldsymbol{\Gamma}}$ to refer to the tuple $(\boldsymbol{f}_{\Gamma_i})_{i \in [\boldsymbol{\Gamma}]}$ which includes F_{Γ_0} whenever $0 \in [\Gamma]$.

Let $\boldsymbol{\Gamma}$ be an origin facing for which $\mathbb{V}_{\mathbb{K}^*}(\boldsymbol{g}_{\boldsymbol{\Gamma}})$ is non-empty. Then it defines a function

$$f_\Gamma := F_{\Gamma_0}|_{\mathbb{V}^*_{\mathbb{K}}(\boldsymbol{g}_{\boldsymbol{\Gamma}})} : \mathbb{V}^*_{\mathbb{K}}(\boldsymbol{g}_{\boldsymbol{\Gamma}}) \longrightarrow \mathbb{K},$$

where for any subset $S \subset \mathbb{K}^n$, we use the notation S^* to denote the intersection $S \cap (\mathbb{K}^*)^n$. We use this notation to define the set $\mathcal{D}_\Gamma(f) \subset \mathbb{K}$, called the *face-discriminant* of f, defined as

$$\mathcal{D}_\Gamma(f) := f_\Gamma(\mathrm{Crit}(f_\Gamma)) \tag{16}$$

In other words, the locus $\mathcal{D}_\Gamma(f)$ is the discriminant of the function f_Γ. By Bertini Theorem [25], face-resultants are finite sets in $\mathbb{K}$.

Example 15. *Consider the polynomial function $F : \mathbb{R}^3 \longrightarrow \mathbb{R}$, $(x, y, z) \longmapsto 1 + x + x^2$, and let $X \subset \mathbb{R}^3$ be the curve $\mathbb{V}_{\mathbb{R}}(g_1,\ g_2)$, where $g_1 := -2 + x + 2y - y^2$ and $g_2 := 1 + 2x - 3y + 4z$. The triple $\boldsymbol{f} := (F, g_1, g_2)$ has Newton tuple $\boldsymbol{\Delta}$ illustrated in Figure 1. Let us compute some of the face-discriminants of f.*

If $\boldsymbol{\Gamma} = \boldsymbol{\Delta}$, then we have $\mathcal{D}_\Gamma(f) = F(\mathrm{Crit}\ F|_{X^*}) = \mathbb{V}_{\mathbb{R}}(h)$, where $h = \langle F - t,\ g_1,\ g_2,\ \det \mathrm{Jac}_{(x,y,z)}\, \boldsymbol{f} \rangle$. Another trivial case is whenever $\boldsymbol{\Gamma}$ is any one of the important facings of the "Side" face of $\boldsymbol{\Delta}$ (see Figure 1); since $\Gamma_0 = \{(0, 0, 0)\}$, we always get $F_{\Gamma_0} \equiv 1$, and thus $\mathcal{D}_\Gamma(f) = \{1\}$.

Now let $\boldsymbol{\Gamma} \prec \boldsymbol{\Delta}$ be the important origin face that is the "Bottom" triple of Figure 1. Hence, we have $F_{\Gamma_0} = F_{\Delta_0} = F$. Then, the domain of $\boldsymbol{f}_{\boldsymbol{\Gamma}}$ is the union of two veritcal lines $\mathbb{V}_{\mathbb{R}}(g_1,\ 1 + 2x - 3y)$. A generic $t \in \mathbb{R}$, has no preimages under $\boldsymbol{f}_{\boldsymbol{\Gamma}}$. This shows that

$$\mathcal{D}_\Gamma(f) = F(\mathbb{V}_{\mathbb{R}}(g, 1 + 2x - 3y)) = \{1,\ 205/16\}.$$

Remark 1. *Note that, if $X = \mathbb{K}^n$, then we set $[\boldsymbol{\Delta}] = \{0\}$. Hence, every facing $\boldsymbol{\Gamma} \prec \boldsymbol{\Delta}$ corresponds to the first polytope Γ_0 that is a face of Δ_0. By definition, we get*

$$\mathcal{D}_\Gamma(f) = F_{\Gamma_0}(\mathrm{Crit}\, F_{\Gamma_0}).$$

Notation 16. *Let $\mathcal{I}(\boldsymbol{\Delta})$ and $\mathcal{O}(\boldsymbol{\Delta})$ denote the set of all important facings of $\boldsymbol{\Delta}$ and origin ones respectively. The intersection of the above two sets is denoted by $\mathcal{IO}(\boldsymbol{\Delta})$.*

We have the following classical result.

Theorem 17 ([32])**.** *Let $\boldsymbol{\Delta}$ be a tuple consisting of a single integer polytope in $\mathbb{R}^n_{\geq 0}$. Then, there exists a Zariski open subset $\Upsilon \subset \mathbb{C}^{\boldsymbol{\Delta}}$, such that for every polynomial $f \in \Upsilon$ it holds that*

$$\mathcal{B}_f \setminus \{f(\mathbf{0})\} \subset \bigcup_{\boldsymbol{\Gamma} \in \mathcal{IO}(\boldsymbol{\Delta})} \mathcal{D}_\Gamma(f).$$

The detailed version of Theorem 17 illustrates a description of the set Υ, together with all the origin faces of Δ that do not contribute to the infinity bifurcation set. This demonstrates that face-resultants are useful for effectively computing the bifurcation set of a large family of complex polynomials.

Let us state a generalization of Theorem 17, and postpone its proof to the end of this section.

Theorem 18. *Let $\boldsymbol{\Delta}$ be a tuple of lattice polytopes in $\mathbb{R}^n_{\geq 0}$ satisfying $[\boldsymbol{\Delta}] = \{0, 1, \ldots, r\}$, and assume that $\{\mathbf{0}\}$ is a vertex of Δ_0. Then, there exists a Zariski open subset $\Omega \subset \mathbb{K}^{\boldsymbol{\Delta}}$, such that for every $\boldsymbol{f} \in \Omega$, the set $\mathbb{V}_{\mathbb{K}}(g_1, \ldots, g_r)$ is smooth, and the correcponding polynomial function $f : \mathbb{V}_{\mathbb{K}}(g_1, \ldots, g_r) \longrightarrow \mathbb{K}$ satisfies*

$$\mathcal{B}_f \setminus \{f(\mathbf{0})\} \subset \bigcup_{\boldsymbol{\Gamma} \in \mathcal{IO}(\boldsymbol{\Delta})} \mathcal{D}_\Gamma(f). \tag{17}$$

In what follows, we use $f|_*$ to denote the function

$$f|_* := f|_{X^*} : X^* \longrightarrow \mathbb{K}.$$

Similarly to $\mathcal{B}_f$, the set $B_{f|_*}$ is also finite [45].

Lemma 19. *We have*

$$\mathcal{B}_f = \bigcup_H \mathcal{B}_{f|_H} \cup \mathcal{B}^{\infty}_{f|_*} \cup \mathcal{B}^0_f, \tag{18}$$

where H runs over all coordinate hyperplanes in $\mathbb{K}^n \setminus (\mathbb{K}^*)^n$, and $f|_H$ are the restricted functions $X \cap H \longrightarrow \mathbb{K}$.

Proof. We consider only the case where $X \neq (\mathbb{K}^*)^n$ as the complementary case is similar. Let $\lambda \in \mathcal{B}_f$, and assume that $\lambda \notin \mathcal{B}^{\infty}_{f|_*} \cup \mathcal{B}^0_f$. Without loss of generality we set $\lambda = 0$. Then, for every arbitrary large compact subset $K \subset \mathbb{K}^n$, there is a small enough neighborhood $U \ni 0$, for which the function $h : f|_*^{-1}(U) \setminus K \longrightarrow U$, obtained by restricting $f|_*$ to $f|_*^{-1}(U) \setminus K$, is a $\mathcal{C}^\infty$-fibration. Consequently, for each $z \in U$, we have $h^{-1}(0)$ and $h^{-1}(z)$ are diffeomorphic. Then, from $0 \notin \mathcal{B}^0_f := f(\operatorname{Crit} f)$, the two preimages are diffeomorphic:

$$\overline{h^{-1}}^X(0) \cong \overline{h^{-1}}^X(z), \tag{19}$$

where the closures are taken in the Eucludean topology induced on X. However, there exists a value $z \in U$ for which the following holds

$$f^{-1}(0) \setminus K \ncong f^{-1}(z) \setminus K. \tag{20}$$

We deduce from (19) and (20) that for some coordinate hyperplane H, the preimage $(f^{-1}(0) \setminus K) \cap H$ is not diffeomorphic to $(f^{-1}(z) \setminus K) \cap H$ for K large enough. This shows that $z \in \mathcal{B}_h$. □

In order to prove Theorem 18, we furthermore require the following notion for tuples $\boldsymbol{f} \in \mathbb{K}^{\boldsymbol{\Delta}}$. Taking the coefficients of $\boldsymbol{f}$ in the parameter space $\mathbb{K}^{\boldsymbol{\Delta}}$ as additional variables, we denote by $\{\boldsymbol{f} = 0\}$ the locus $\{F = g_1 = \cdots = g_r = 0\} \subset (\mathbb{K}^*)^n \times \mathbb{K}^{\boldsymbol{\Delta}}$. The *Bertini discriminant* $\boldsymbol{B}$ of $\boldsymbol{\Delta}$ (see e.g., [15, Definition 3.5]) is the bifurcation set of the map $\pi|_{\{\boldsymbol{f}=\boldsymbol{0}\}}$ by restricting to $\{F = g_1 = \cdots = g_r = 0\}$ the projection

$$\pi : (\mathbb{K}^*)^n \times \mathbb{K}^{\boldsymbol{\Delta}} \longrightarrow \mathbb{K}^{\boldsymbol{\Delta}}.$$

The Bertini discriminant can be computed as follows. For any facing $\boldsymbol{\Gamma} \prec \boldsymbol{\Delta}$, we define $D_{\boldsymbol{\Gamma}} \subset \mathbb{K}^{\boldsymbol{\Delta}}$ as the closure of

$$\left\{ \boldsymbol{f} \;\middle|\; \exists \boldsymbol{x}_0 \in \mathbb{V}^*_{\mathbb{K}}(\boldsymbol{f}_{\boldsymbol{\Gamma}}),\ \operatorname{Jac}_{\boldsymbol{x}}(\boldsymbol{f}_{\boldsymbol{\Gamma}})|_{\boldsymbol{x}=\boldsymbol{x}_0} \text{ does not have full rank} \right\} \tag{21}$$

In other words, the set $D_{\boldsymbol{\Gamma}}$ consists of all $\boldsymbol{f} \in \mathbb{K}^{\boldsymbol{\Delta}}$ for which there exists a point $\boldsymbol{x}_0 \in (\mathbb{K}^*)^n$ at which $\mathbb{V}^*_{\mathbb{K}}(\boldsymbol{f}_{\boldsymbol{\Gamma}})$ is not a complete intersection. It was shown in [15, Theorem 1.1 and Proposition 4.10] (see also [18, Section 5]) that $\operatorname{cdim} \boldsymbol{B} = 1$, and it holds that

$$\boldsymbol{B} = \bigcup_{\boldsymbol{\Gamma} \in \mathcal{I}(\boldsymbol{\Delta})} D_{\boldsymbol{\Gamma}}. \tag{22}$$

Remark 2. *The mentioned results in [15] are formulated for the case where $\mathbb{K} = \mathbb{C}$. For the real case, the statement assumes that $\boldsymbol{f}$ consists of only one polynomial F [18, Section 5.A]. Transitioning to arbitrary tuples $\boldsymbol{f}$ with $\mathbb{K} = \mathbb{R}$ can be done via the Cayley Trick (see e.g., [14, Section 6.2]).*

Proof of Theorem 18. Let $\boldsymbol{f} \in \mathbb{K}^{\boldsymbol{\Delta}} \setminus \boldsymbol{B}$, and let $\Lambda(\boldsymbol{f}) \subset \mathbb{K}^{\boldsymbol{\Delta}}$ be the line passing through $\boldsymbol{f}$ with direction $(1, 0, \ldots, 0)$, where the first coordinate corresponds to the constant term of F. In other words, we have

$$\Lambda(\boldsymbol{f}) := \{(F - z, g_1, \ldots, g_r) \mid z \in \mathbb{K}\}.$$

On the one hand, if $\Theta \subset \mathbb{K}^{\Delta}$ is the set of all $\boldsymbol{f}$ at which $\Lambda(\boldsymbol{f})$ intersects $\boldsymbol{B}$ transversally, then for each $\boldsymbol{f} \in \Theta$, the relation

$$\mathcal{B}_{f|_*} = \boldsymbol{B} \cap \Lambda(\boldsymbol{f}), \tag{23}$$

follows from the definition of $\boldsymbol{B}$. Note that if $\widetilde{\boldsymbol{\Delta}}$ is the tuple of polytopes $(\widetilde{\Delta}_0, \Delta_1, \ldots, \Delta_r)$, where

$$\widetilde{\Delta}_0 := \operatorname{conv}(\Delta_0 \cap \mathbb{N}^n \setminus \{\boldsymbol{0}\}),$$

then Θ is the preimage, under Π, of the bifurcation set of the map $\Pi|_{\boldsymbol{B}}$, where $\Pi := \mathbb{K}^{\boldsymbol{\Delta}} \longrightarrow \mathbb{K}^{\widetilde{\boldsymbol{\Delta}}}$ is the projection taking a tuple $\boldsymbol{f}$ to $\tilde{\boldsymbol{f}}$ by forgetting the constant term of the first polynomial F. Thanks to a theorem of Verdier [50], the set Θ contains a Zariski open subset $\Omega \subset \mathbb{K}^{\Delta}$.

On the other hand, for any $\boldsymbol{f} \in \mathbb{K}^{\Delta}$, and each facing $\boldsymbol{\Gamma} \prec \boldsymbol{\Delta}$, we have

$$\mathcal{D}_{\Gamma}(f) = D_{\boldsymbol{\Gamma}} \cap \Lambda(\boldsymbol{f}). \tag{24}$$

Therefore, since $\mathcal{I}(\boldsymbol{\Delta})$ contains the tuple $\boldsymbol{\Delta}$, it follows from (22), (23), and (24) that

$$\mathcal{B}_{f|_*} = \bigcup_{\boldsymbol{\Gamma} \in \mathcal{I}(\boldsymbol{\Delta})} \mathcal{D}_{\Gamma}(f). \tag{25}$$

Next, we show that we can replace the set $\mathcal{I}(\boldsymbol{\Delta})$ in (25) by $\mathcal{IO}(\boldsymbol{\Delta})$. Recall that $\boldsymbol{g}_{\boldsymbol{\Gamma}}$ refers to $(\boldsymbol{g}_{\Gamma_i})_{i \in [\boldsymbol{\Gamma}] \setminus \{0\}}$, and $\boldsymbol{f}_{\boldsymbol{\Gamma}}$ to $(\boldsymbol{f}_{\Gamma_i})_{i \in [\boldsymbol{\Gamma}]}$ which includes F_{Γ_0} whenever $0 \in [\Gamma]$.

The set $D_{\boldsymbol{\Gamma}}$ is a fiber bundle $\mathrm{Proj}^{-1}(\mathrm{Proj}(D_{\boldsymbol{\Gamma}}))$, where $\mathrm{Proj} : \mathbb{K}^{\boldsymbol{\Delta}} \longrightarrow \mathbb{K}^{\boldsymbol{\Gamma}}$, $\boldsymbol{f} \longmapsto \boldsymbol{f}_{\boldsymbol{\Gamma}}$. Hence, since $\boldsymbol{f} \notin D_{\boldsymbol{\Gamma}}$, we have $\boldsymbol{f}_{\boldsymbol{\Gamma}} \notin \mathrm{Proj}(D_{\boldsymbol{\Gamma}})$. This shows that $\mathrm{Proj}^{-1}(\boldsymbol{f}_{\boldsymbol{\Gamma}}) \cap D_{\boldsymbol{\Gamma}} = \varnothing$. Then, if $\boldsymbol{\Gamma}$ is not origin, we get $\Lambda(\boldsymbol{f}) \subset \mathrm{Proj}^{-1}(\boldsymbol{f}_{\boldsymbol{\Gamma}})$. We conclude that $\mathcal{D}_{\Gamma}(f) = \varnothing$ if $\boldsymbol{\Gamma}$ is not origin, and thus we get

$$\mathcal{B}_{f|_*} = \bigcup_{\boldsymbol{\Gamma} \in \mathcal{IO}(\boldsymbol{\Delta})} \mathcal{D}_{\Gamma}(f). \tag{26}$$

It remains to prove that for any $z \in \mathcal{B}_f \setminus (\mathcal{B}^{\infty}_{f|_*} \cup \mathcal{B}^0_f)$, we have $z \in \mathcal{D}_{\Gamma}(f)$ for some $\boldsymbol{\Gamma} \in \mathcal{IO}(\boldsymbol{\Delta})$. By Lemma 19, we have $z \in \mathcal{B}^{\infty}_h$ for the restricted function $h := f|_{\{x_i = 0\}} : X \cap \{x_i = 0\} \longrightarrow \mathbb{K}$ for some $i \in [n]$. Without loss of generality, we assume that $i = n$. Notice that g can also be expressed as the function

$$f_{\Gamma'} := F_{\Gamma'_0}\big|_{X \cap \{x_n = 0\}} \mathbb{V}_{\mathbb{K}}(\boldsymbol{g}_{\boldsymbol{\Gamma}'}) \longrightarrow \mathbb{K}, \tag{27}$$

where $\boldsymbol{\Gamma}' \prec \boldsymbol{\Delta}$ is the facing whose Minkowski sum spans the coordinate hyperplane $\{x_n = 0\} \subset \mathbb{R}^n$. Indeed, we deduce this by plugging $x_n = 0$ into the equations of $\boldsymbol{f}$.

By computing the Jacobian matrix, we deduce that $z \notin \mathcal{B}^0_f \Longrightarrow z \notin \mathcal{B}^0_h$. Then, thanks to Lemma 19 and (26) applied to h, if $z \notin \mathcal{R}_{\Gamma}(h)$ for some important origin facing $\boldsymbol{\Gamma} \prec \boldsymbol{\Gamma}'$, then $z \in \mathcal{B}_{\ell}$, where $\ell := h|_{\{x_j = 0\}} : X \cap \{x_j x_n = 0\} \longrightarrow \mathbb{K}$ for some $j \in [n-1]$. Notice that $\boldsymbol{\Gamma}$ is an important facing of $\boldsymbol{\Delta}$, and that $\mathcal{D}_{\Gamma}(f) = \mathcal{D}_{\Gamma}(h)$.

Applying recursively the above arguments, we deduce that either $z \in \mathcal{D}_{\Gamma}(f)$ for some $\boldsymbol{\Gamma} \in \mathcal{IO}(\boldsymbol{\Delta})$, or there exists $\widetilde{\boldsymbol{\Gamma}} \in \mathcal{IO}(\boldsymbol{\Delta})$ such that $z \in \mathcal{B}_{f_{\widetilde{\Gamma}}}$, where $\widetilde{\Gamma}_0 = \{\mathbf{0}\}$, and

$$f_{\widetilde{\Gamma}} := F_{\widetilde{\Gamma}_0}\Big|_{X \cap \{x_1 \cdots x_n = 0\}} \mathbb{V}_{\mathbb{K}}(\boldsymbol{g}_{\widetilde{\boldsymbol{\Gamma}}}) \longrightarrow \mathbb{K}. \tag{28}$$

Clearly, we have $f_{\widetilde{\Gamma}}$ is the trivial map whose target space is the point $\{F(\mathbf{0})\}$. This finishes the proof. □

Remark 3. *From the proof of Theorem 18, in combination with Eq.* (26)*, we can deduce that* $\mathcal{B}_f \subset \mathcal{B}_{f|_*}$ *if* $\boldsymbol{f} \in \Omega$.

4.3 Bounding the infimum

To obtain all the asymptotic critical values it suffices to go through all the face lattice of $\boldsymbol{\Delta}$ and compute the corresponding face discriminant of f, see Eq. (16). As we aim for worst case bounds, we consider the worst case where all the polynomials are of degree d. That is we ignore their restriction to facings. We emphasize, once more, that is suffices for our purposes as we are targeting an algorithm but bounds on the bitsize on the corresponding quantities.

The face discriminant results in a polynomial system of at most $\mathcal{O}(s + n^n)$ polynomials; that is the s polynomials g_i, the $\mathcal{O}(n^n)$ polynomials coming from the rank condition on the Jacobian, and the polynomial $F - z$. The bitsize of the polynomials is at most $\widetilde{\mathcal{O}}(n\tau)$; dominated by the polynomials from the rank condition on the Jacobian. Our goal is to bound the roots of this system, which is 0-dimensional, due to Bertini's theorem. In particular, we want to bound z, that corresponds to the asymptotic critical values and consequently to f^*. Following [13], the bounds appearing in Theorem 4 hold true for any coordinate of the roots of the system and for f^*.

5 Unconstrained optimization and asymptotic critical values

We present an algorithm (Alg. 1) to compute (and to obtain precise bounds on) the elements of the Rabier set $\mathcal{K}(f)$, that is the set of generalized critical values, of a polynomial function $f : \mathbb{C}^n \to \mathbb{C}$, where $f \in \mathbb{Z}[\boldsymbol{x}]$. One of these values corresponds to the global infimum of the (polynomial) optimization problem $f^* = \inf_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x})$. We assume that f has degree d and maximum coefficient bitsize τ, or in short that f is of size (d, τ).

5.1 Asymptotic critical values over $\mathbb{C}^n$

The (pseudo-code of the) algorithm in Alg. 1 supports the computations of the asymptotic critical values of f and is based on [22]. The main difference is the way that we perform the elimination. Instead of Gröbner basis we use resultant matrices to control better the bitsize of the various objects and the complexity of the overall algorithm. The algorithm relies on the following notion.

Definition 20. *For any $f \in \mathbb{C}[x_1, \ldots, x_n]$, we define a set $\{g_1, \ldots, g_{n-1}\} \subset \mathbb{C}[x_1, \ldots, x_n]$ of polynomials given by*

$$g_k := \sum_{i=1}^{n} a_i^{(k)} \frac{\partial f}{\partial x_i} + \sum_{i,j=1}^{n} b_{i,j}^{(k)} x_i \frac{\partial f}{\partial x_j}, \quad \text{for } k \in [n-1],$$

for some choice of coefficients $(\boldsymbol{a}, \boldsymbol{b}) := (\boldsymbol{a}^{(1)}, \boldsymbol{b}^{(1)}, \ldots, \boldsymbol{a}^{(n-1)}, \boldsymbol{b}^{(n-1)}) \in \mathbb{C}^{n^3-n}$. Then, the super polar curve *of f is the set*

$$\mathcal{G}_f(\boldsymbol{a}, \boldsymbol{b}) := \overline{\mathbb{V}(g_1, \ldots, g_{n-1}) \setminus \mathrm{Crit}(f)},$$

where we use here the Zariski closure. The super polar curve is said to be non-degenerate *if $\mathcal{G}_f(\boldsymbol{a}, \boldsymbol{b}) \subset \mathbb{C}^n$ is indeed a curve, i.e., of dimension 1.*

Roughly speaking, super polar curves encapsulate the behaviour at infinity of the image of (sequence of) points in $\mathbb{C}^n$ under f when their norm tends to infinity. A key point of the algorithm is to ensure that the linear combinations g_k are such that the resulting set $\mathcal{G}_f(\boldsymbol{a}, \boldsymbol{b})$ is non-degenerate.

Algorithm 1: ASYMPTOTICCRITICALVALUES(f)

Input : $f \in \mathbb{Z}[x_1, \ldots, x_n]$
Output : A set $K \in \mathbb{C}$ such that $\mathcal{K}_\infty(f) \subset K$.

1 Choose generic $a_i^{(k)}$ and $b_{i,j}^{(k)}$, for $k \in [n-1]$ and $i, j \in [n]$
2 **for** $k \in [n-1]$ **do**
3 $\quad g_k(\boldsymbol{x}) = \sum_{i=1}^{n} a_i^{(k)} \frac{\partial}{\partial x_i} f(\boldsymbol{x}) + \sum_{i,j=1}^{n} b_{i,j}^{(k)} x_i \frac{\partial}{\partial x_j} f(\boldsymbol{x})$;
4 **for** $i \in [n]$ **do**
5 $\quad \bar{h}_i \leftarrow \mathtt{Elim}(\{g_1(\boldsymbol{x}), \ldots, g_{n-1}(\boldsymbol{x}), f(\boldsymbol{x}) - z\}, \boldsymbol{x}_{-i}) \in \mathbb{Z}[x_i, z]$;
6 $\quad h_i \leftarrow \mathtt{lc}(\bar{h}_i, x_i)$;
7 $h \leftarrow \prod_{i=1}^{n} h_i(z) \in \mathbb{Z}[z]$;
8 RETURN $K := \{\gamma \in \mathbb{C} \,|\, h(\gamma) = 0\}$ `/* The distinct roots of` $h \in \mathbb{Z}[z]$ `*/`

To study the complexity of Alg. 1, the bitsize of the elements of a linear transformation that we can apply to f and its derivatives to obtain the polynomials g_k, such that the corresponding super-polar curve is non-degenerate. This bound permits us to estimate the probability of success when we perform random linear combinations.

Nevertheless, if needed, we can always obtain a Las Vegas algorithm, that is an algorithm that always returns a correct output, if we allow ourselves the cost of certifying the result; that is to check if the resulting super-polar curve is indeed non-degenerate. It was shown in [22] that Alg. 1 computes correctly the Rabier set of a polynomial f. Proving our main result, Theorem 25, relies on the complexity analysis of Alg. 1. For this we need two technical results. The first is the following proposition, the proof of which is in the Appendix (Lemma 28).

Proposition 21. *Consider an affine variety $V \in \mathbb{C}^n$ of dimension δ and degree D. Let S be finite set of numbers such that $0 \notin S$. Also consider the linear form $h(\boldsymbol{x}) = c_0 + \sum_{i=1}^{n} c_i x_i$ and $H = \mathbb{V}(h)$. If we choose the c_i's independently and uniformly at random from S, then*

$$\Pr[\dim(V \cap H) = \delta - 1] \geq 1 - \tfrac{D}{|S|}.$$

Moreover, there is a specialization of the c_i's to integers such that their bitsize is $\lceil \lg(2\kappa D) \rceil + 1$, where $\kappa \geq 2$ is a constant, so that it holds $\dim(V \cap H) = \delta - 1$.

The the second technical result is as follows.

Lemma 22 (Genericity of the super polar curve)**.** *For any $f \in \mathbb{Z}[x_1, \ldots, x_n]$, there are integers $(\boldsymbol{a}, \boldsymbol{b}) \in \mathbb{Z}^{n^3-n}$ of bitsize $\mathcal{O}(n^2 + \lg d)$ such that the super-polar curve $\mathcal{G}_f(\boldsymbol{a}, \boldsymbol{b})$ is non-degenerate.*

In addition, if we pick the elements of $(\boldsymbol{a}, \boldsymbol{b})$ uniformly at random from a set S containing $2\, cnd^{n-1}$ elements, where c is a constant, then the super-polar curve is non-degenerate with probability $\geq 1 - \frac{1}{c}$.

Proof. Assume that each of the polynomials The polynomials $g_1, \ldots, g_{n-1} \in (\mathbb{Z}[\boldsymbol{a}, \boldsymbol{b}])[\boldsymbol{x}]$ are of degree at most d with respect to $\boldsymbol{x}$.

Following [22], we know that if $(\boldsymbol{a}, \boldsymbol{b})$ lies in a Zariski open set $\mathcal{Z} \subset \mathbb{C}^{n^3-n}$, then $\mathcal{G}_f(\boldsymbol{a}, \boldsymbol{b})$ is non-degenerate.

Then, thanks to Proposition 21, there are linear polynomials, say $L_0, L_n \in \mathbb{Z}[\boldsymbol{x}]$, where $L_i = \ell_{i,0} + \sum_{j=1}^{n} \ell_{i,j} x_j$ for $i \in \{0, n\}$, such that zero set of the following polynomial system

$$\{L_0 = g_1 = \cdots = g_{n-1} = L_n = 0\}, \tag{29}$$

has no solutions for any $(\boldsymbol{a}, \boldsymbol{b}) \in \mathcal{Z}$.

The coefficients of the linear polynomials L_0 and L_n are integers, but for the moments we consider them as (new) parameters. The system (29) consists of $n+1$ polynomials in n variables, i.e., $\{x_1, \ldots, x_n\}$.

Let $\boldsymbol{\ell}$ denote the set of cofficients of L_0 and L_n. Eliminating the variables $\boldsymbol{x}$ from the system (29) we obtain an ideal generated by a polynomial $R \in \mathbb{C}[\boldsymbol{\ell}, \boldsymbol{a}, \boldsymbol{b}]$, called the *resultant*. Note that R is not identically zero. Indeed, from $(\boldsymbol{a}, \boldsymbol{b}) \in \mathcal{Z}$, we get $\dim(\mathcal{G}_f(\boldsymbol{a}, \boldsymbol{b})) = 1$ and thus Proposition 21 guarantees that there are (two) linear polynomials, hence a specialization of $\boldsymbol{\ell}$, such that the system in (29) has no solutions. Therefore, for any choice $(\boldsymbol{\ell}, \boldsymbol{a}, \boldsymbol{b})$ such that $(\boldsymbol{a}, \boldsymbol{b}) \in \mathcal{Z}$ and $R(\boldsymbol{\ell}, \boldsymbol{a}, \boldsymbol{b}) \neq 0$, we have $\dim \mathcal{G}_f(\boldsymbol{a}, \boldsymbol{b}) = 1$ and (29) has no solutions.

Notice that $R \in (\mathbb{Z}[\boldsymbol{c}, \boldsymbol{\ell}])[\boldsymbol{a}, \boldsymbol{b}]$ and as a polynomials in $\boldsymbol{a}$, or $\boldsymbol{b}$ it has at most $n^3 - n$ variables and its total degree is at most nd^{n-1}.

Following the DeMillo-Lipton-Schwartz-Zippel (DLSZ) lemma [29, 39, 36], if S is a finite set of integers, then R has at most $2nd^{n-1}|S|^{n^3-n-1}$ solutions in the grid S^{n^3-n}.

If S contains the first $2\ cnd^{n-1}$ integers (where c is a constant greater than 2), then we deduce that there is a specialization of $(\boldsymbol{a}, \boldsymbol{b})$ such that $\dim(\Gamma) = 1$. Also the bitsize of the elements of $\boldsymbol{a}$ and $\boldsymbol{b}$ is $\mathcal{O}(n^2 + \lg(d))$.

Moreover, thanks to (DLSZ), if we pick a specialization for $(\boldsymbol{a}, \boldsymbol{b})$ uniformly at random in the grid S^{n^3-n}, then $\dim(\Gamma) = 1$ with probability $\geq 1 - \frac{1}{c}$. □

Theorem 23. *Let $f \in \mathbb{Z}[\boldsymbol{x}]$ be of (total) degree at most d and bitsize τ. Then, Alg. 1 is a Monte Carlo algorithm that computes the asymptotic critical values in $\widetilde{\mathcal{O}}_B(n^{\omega n-\omega+1} d^{(\omega+3)n-\omega-1}\tau)$ bit operations, where ω is the exponent of the complexity of matrix multiplication.*

The asymptotic critical values are algebraic numbers of degree $\mathcal{O}(d^{n-1})$ and the (sum of the) bitsize(s) of their isolating intervals (or boxes) is $\widetilde{\mathcal{O}}(nd^{2n-2}\tau + n^2 d^{n-1})$.

Proof. First we construct the polynomials g_k, for $k \in [n-1]$; they are of degree d. Following Lemma 22 the bitsize of elements of $\boldsymbol{a}$ and $\boldsymbol{b}$ is $\mathcal{O}(n^2 + \lg d)$, hence the bitsize of g_k is $\mathcal{O}(\tau + n^2 + \lg d)$.

Then we consider the set of polynomials

$$\{g_1, \ldots, g_{n-1}, f - z\}, \tag{30}$$

and we eliminate the variables $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$, for $i \in [n]$. Thus, we perform the elimination n times.

Without loss of generality, we may assume that $i = n$.

After eliminating the variables $x_1, \ldots, x_{n-1}$ from the set of polynomials in (30), we obtain a polynomial $R_n \in \mathbb{Z}[x_n, z]$; this is the resultant of the polynomials in (30). We compute R_n using the Macaulay matrix, say M, corresponding to the system in (30). In particular we express R_n as the ratio of two determinant of (sub)matrices of M. It might happen that the determinant in the denominator is zero. To avoid this problem and still compute R_n we use the technique of generalized characteristic polynomial [6]. For this we perturb symbolically the polynomials g_i to become $G_i = g_i + sx_i^d$, where s is a new variable and $i \in [n-1]$; and $G_n = f - z$. In this way, the ratio of determinants from M result in a polynomial $S_n \in \mathbb{Z}[x_n, z][s]$. The non-vanishing coefficient of $S_n \in \mathbb{Z}[s]$, in front of the monomial with the smallest degree in s, is a power of R_n.

To obtain bounds on S_n and R_n we rely on [13]. Following [8], S_n is homogeneous in the coefficients of the each polynomial G_i. In particular, each term of S_n is as follows

$$\varrho_k |\boldsymbol{c}_1|^{d^{n-1}} \cdots |\boldsymbol{c}_n|^{d^{n-1}},$$

where the semantics of $|\boldsymbol{c}_i|^{d^{n-1}}$ are that it represents a product of coefficients of the polynomial G_i all of which have total degree d^{n-1}. The coefficients of G_i are in $\mathbb{Z}[x_n, z, s]$, they have multidegree $(d, 0, 1)$, for $i \in [n-1]$ and G_n

has multidegree $(d,1,0)$; their bitsize τ. Hence, $|c_i|^{d^{n-1}}$ is of multidegree $(d^n,0,d^{n-1})$ for $i\in[n-1]$ and G_n has multidegree $(d^n,d^{n-1},0)$; their bitsize $\mathcal{O}(d^{n-1}\tau+d^{n-1}\lg(d))$. Also $\lg|\varrho_k|=\mathcal{O}(nd^{n-1}\lg(d))$. Consequently, each term is a polynomial in $\mathbb{Z}[x_n,z,s]$ of multidegreee (d^n,d^{n-1},d^{n-1}) and bitsize $\mathcal{O}(nd^{n-1}\tau+nd^{n-1}\lg(d)+n^2\lg(nd))$, using Claim 26. As there are at most $\mathcal{O}(d^{3n})$ terms, the bitsize of S_n and R_n is $\mathcal{O}(nd^{n-1}\tau+nd^{n-1}\lg(d)+n^2\lg(nd))$.

To actually compute S_n and R_n we use perform the elimination using the Macaulay matrix. We consider the homogenization of the polynomials $G_1,\ldots,G_n$. be introducing a new variable x_0.

Let $m=\sum_{i=1}^m(d-1)+1=n(d-1)+1\leq nd$ and $N=\binom{m+n-1}{n-1}\leq(nd)^{d-1}$. The latter corresponds to all monomials in n variables of degree $\leq m$.

As the elements of M are polynomials in $\mathbb{Z}[x_n,z,s]$ we apply Kronecker's trick to obtain univariate polynomials. In particular, we perform the transformation $x_n\leftarrow s^{d^{n-1}+1}$ and $z\leftarrow s^{(d^{n-1}+1)d^n}$. Then, the elements of M become polynomials in $\mathbb{Z}[s]$ of degree $\leq d^{2n-1}+d^n+1=\mathcal{O}(d^{2n})$. Now, we can compute the determinants $\mathcal{O}(N^\omega d^{2n})=n^{\omega(n-1)}d^{(\omega+2n)-\omega}$ arithmetic operations.

If we account on the bitsize of S_n (and R_n) we deduce that we can compute them $\widetilde{\mathcal{O}}_B(n^{\omega n-\omega+1}d^{(\omega+3)n-\omega-1}\tau)$ bit operations.

Hence, after we arrange the terms of the resultant, we can recover the polynomial and $\bar{h}_n$ and thus $h_n\in\mathbb{Z}[z]$. The latter has size $(d^{n-1},\widetilde{\mathcal{O}}(nd^{n-1}\tau+n^2))$.

We isolate its real (or complex) real roots in $\widetilde{\mathcal{O}}_B(nd^{3n-3}\tau+n^2d^{2n-2})$ [33]. The (sum of the) bitsize of the isolating interval(s) (or boxes if we consider the complex roots) of the roots is $\widetilde{\mathcal{O}}(nd^{2n-2}\tau+n^2d^{n-1})$, e.g., [13]. □

5.2 Polynomial optimization over $\mathbb{R}^n$

In this part, we show how to use the Rabier set of f (3), together with Alg. 1, for computing the global infimum over $\mathbb{R}^n$. Following Theorem 8 (see also, [38]), it holds that

$$\inf_{x\in\mathbb{R}^n}f(x)\subset\mathcal{K}(f)\cap\mathbb{R},\tag{31}$$

where $\mathcal{K}(f)=\mathcal{K}_0(f)\cup\mathcal{K}_\infty(f)$ is the Rabier set of critical values of f. It remains to identify which of these values corresponds to the infimum of f.

If the Rabier values are $e_1,\ldots,e_m$, then we need to compute rational numbers that interlace them [38], that is r_i for $0\leq i\leq m$, such that

$$r_0<e_1<r_1<\cdots<r_{m-1}<e_m<r_m$$

and test if the real hypersurfaces $H_i:=f^{-1}(r_i)\cap\mathbb{R}^n$ are empty or not. The bitsize of corresponding polynomial defining H_i is dominated by the bitsize of r_i, let it be σ_i. Thanks to Theorem 23, we have

$$\sum_{i=0}^m\sigma_i=\widetilde{\mathcal{O}}(nd^{2n-2}\tau+n^2d^{n-1}).$$

Next, we need the following theorem.

Theorem 24 ([12]). *For $f\in\mathbb{Z}[\boldsymbol{x}]$ of size (d,τ) there is a randomized algorithm that with probability $1-\epsilon$ decides if $\mathbb{V}(f)\cap\mathbb{R}^n$ is empty or not using $\widetilde{\mathcal{O}}_B(d^{3n}\lg(\frac{1}{\epsilon})(\tau+\lg(\frac{1}{\epsilon})))$ bit operations.*

Using Theorem 24 we can test the emptiness of all the real hypersurfaces in

$$\sum_{i=0}^m\widetilde{\mathcal{O}}_B(d^{3n}\lg(\tfrac{1}{\epsilon})(\sigma_i+\lg(\tfrac{1}{\epsilon})))=\widetilde{\mathcal{O}}_B(d^{3n}\lg(\tfrac{1}{\epsilon})(nd^{2n-2}\tau+\lg(\tfrac{1}{\epsilon})))$$

bit operations with probability of success $1-\frac{1}{\epsilon}$.

This bounds is the complexity of a Monte Carlo algorithm for optimizing a multivariate polynomial over $\mathbb{R}^n$.

By combining all the previous results, we have the following theorem.

Theorem 25. *Let $f\in\mathbb{Z}[x]$ be of degree d and bitsize τ. We can compute the optimum value of the problem $f^*=\inf\boldsymbol{x}\in\mathbb{R}^nf(\boldsymbol{x})$ in $\widetilde{\mathcal{O}}_B(d^{3n}\lg(\frac{1}{\epsilon})(nd^{2n-2}\tau+\lg(\frac{1}{\epsilon})))$ bit operations, with probability of success $1-\frac{1}{\epsilon}$.*

In addition, it holds $2^{-\eta}\leq|f^*|\leq2^\eta$, where $\eta=\widetilde{\mathcal{O}}(nd^{n-1}\tau+n^2))$. To distinguish f^* from the other generalized critical values, we need approximate it with up to precision of at most $\widetilde{\mathcal{O}}(nd^{2n-2}\tau+n^2d^{n-1})$ bits.

References

[1] C. Bajaj. The algebraic degree of geometric optimization problems. *Discrete & Computational Geometry*, 3:177–191, 1988.

[2] S. Basu, R. Pollack, and M-F.Roy. *Algorithms in Real Algebraic Geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer-Verlag, 2003.

[3] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry*. Springer Berlin Heidelberg, 2006.

[4] C. W. Brown and J. H. Davenport. The complexity of quantifier elimination and cylindrical algebraic decomposition. In *Proceedings of the 2007 international symposium on Symbolic and algebraic computation*, pages 54–60, 2007.

[5] M. Burr, S. Gao, and E. Tsigaridas. The complexity of subdivision for diameter-distance tests. *Journal of Symbolic Computation*, 101:1–27, 2020.

[6] J. Canny. Generalised characteristic polynomials. *Journal of Symbolic Computation*, 9(3):241–250, Mar. 1990.

[7] B. F. Caviness and J. R. Johnson. *Quantifier elimination and cylindrical algebraic decomposition*. Springer Science & Business Media, 2012.

[8] D. A. Cox, J. Little, and D. O'shea. *Using algebraic geometry*, volume 185. Springer, 2006.

[9] M. L. Dogan, A. A. Ergür, and E. Tsigaridas. On the complexity of chow and hurwitz forms. *ACM Communications in Computer Algebra*, 57(4):167–199, 2024.

[10] J. Draisma, E. Horobeţ, G. Ottaviani, B. Sturmfels, and R. R. Thomas. The euclidean distance degree of an algebraic variety. *Foundations of Computational Mathematics (FoCM)*, 16:99–149, 2016.

[11] A. H. Durfee. Five definitions of critical point at infinity. pages 345–360. Springer, 1998.

[12] J. Elliott, M. Giesbrecht, and É. Schost. On the bit complexity of finding points in connected components of a smooth real hypersurface. In *Proc. 45th International Symposium on Symbolic and Algebraic Computation (ISSAC)*, pages 170–177, 2020.

[13] I. Emiris, B. Mourrain, and E. Tsigaridas. Separation bounds for polynomial systems. *Journal of Symbolic Computation*, 101:128–151, 2020.

[14] A. Esterov. Newton polyhedra of discriminants of projections. *Discrete Comput. Geom.*, 44(1):96–148, 2010.

[15] A. Esterov. The discriminant of a system of equations. *Advances in Mathematics*, 245:534–572, 2013.

[16] G. Farin and D. Hansford. *Curves and surfaces for CAGD: a practical guide*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 5 edition, 2002.

[17] A. Garg. Special case algorithms for Nullstellensatz and transcendence degree. Master's thesis, IIT Kanpur, 2020.

[18] I. M. Gel'fand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, resultants, and multidimensional determinants*. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1994.

[19] M. Giusti and J. Heintz. La détermination de la dimension et des points isolées d'une variété algébrique peuvent s' effectuer en temps polynomial. In *Computational Algebraic Geometry and Commutative Algebra, Cortona*, volume 34, pages 216–256, 1991.

[20] Z. Jelonek and K. Kurdyka. On asymptotic critical values of a complex polynomial. *J. Reine Angew. Math.*, 565:1–11, 2003.

[21] Z. Jelonek and K. Kurdyka. Quantitative generalized bertini-sard theorem for smooth affine varieties. *Discrete Comput. Geom.*, 34(4):659–678, 2005.

[22] Z. Jelonek and M. Tibăr. Detecting asymptotic non-regular values by polar curves. *International Mathematics Research Notices*, 2017(3):809–829, 2017.

[23] G. Jeronimo and D. Perrucci. A probabilistic symbolic algorithm to find the minimum of a polynomial function on a basic closed semialgebraic set. *Discrete & Computational Geometry*, 52(2):260–277, 2014.

[24] G. Jeronimo, D. Perrucci, and E. Tsigaridas. On the minimum of a polynomial function on a basic closed semialgebraic set and applications. *SIAM Journal on Optimization*, 23(1):241–255, 2013.

[25] J.-P. Jouanolou. *Théorèmes de Bertini et applications*. Birkhäuser, Boston, 1983.

[26] H. K. Khalil. *Nonlinear systems*, volume 3. Prentice hall Upper Saddle River, NJ, 2002.

[27] E. Kunz. *Introduction to commutative algebra and algebraic geometry*. Springer Science & Business Media, 1985.

[28] C. La Valle and J. Tonelli-Cueto. Some lower bounds on the reach of an algebraic variety. *arXiv e-prints*, pages arXiv–2402, 2024.

[29] R. J. Lipton. The curious history of Schwarz-Zippel Lemma.
https://rjlipton.wordpress.com /2009/11/30/the-curious-history-of -the-schwartz-zippel-lemma/, 2009.

[30] A. Mantzaflaris, É. Schost, and E. Tsigaridas. Sparse rational univariate representation. In *Proceedings of the 2017 ACM on International Symposium on Symbolic and Algebraic Computation*, pages 301–308, 2017.

[31] A. Némethi. Theorie de Lefschetz pour variétés algebriques affines. *CR Acad. Sc. Paris*, 303:567–570, 1986.

[32] A. Némethi and A. Zaharia. On the bifurcation set of a polynomial function and newton boundary. *Publications of the Research Institute for Mathematical Sciences*, 26(4):681–689, 1990.

[33] V. Y. Pan. Univariate polynomials: nearly optimal algorithms for numerical factorization and root-finding. *Journal of Symbolic Computation*, 33(5):701–733, 2002.

[34] T. S. Pham and H. H. Vui. *Genericity in polynomial optimization*, volume 3. World Scientific, 2016.

[35] P. J. Rabier. Ehresmann fibrations and Palais–Smale conditions for morphisms of Finsler manifolds. *Ann. of Math.*, pages 647–691, 1997.

[36] O. E. Raz, M. Sharir, and J. Solymosi. "Polynomials vanishing on grids: The Elekes-Rónyai problem revisited". In *Proceedings of the thirtieth annual symposium on Computational geometry*, page 251. ACM, 2014.

[37] M. Safey El Din. Testing sign conditions on a multivariate polynomial and applications. *Mathematics in Computer Science*, 1(1):177–207, 2007.

[38] M. Safey El Din. Computing the global optimum of a multivariate polynomial over the reals. In *Proc. ACM 21st International Symposium on Symbolic and Algebraic Computation (ISSAC)*, pages 71–78, 2008.

[39] N. Saxena. Progress on polynomial identity testing. *Bulletin of the EATCS*, 99:49–79, 2009.

[40] M. Schaefer and D. Štefankovič. Fixed points, nash equilibria, and the existential theory of the reals. *Theory of Computing Systems*, 60(2):172–193, 2017.

[41] J. T. Schwartz and M. Sharir. On the "piano movers'" problem i. the case of a two-dimensional rigid polygonal body moving amidst polygonal barriers. *Communications on pure and applied mathematics*, 36(3):345–398, 1983.

[42] I. R. Shafarevich and M. Reid. *Basic algebraic geometry*, volume 2. Springer, 1994.

[43] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo. *Robotics: modelling, planning and control*. Springer Science and Business Media, 2009.

[44] Z. Song, D. P. Woodruff, and P. Zhong. Low rank approximation with entrywise l1-norm error. In *Proc. 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 688–701, 2017.

[45] R. Thom. Ensembles et morphismes stratifiés. *Bull. Am. Math. Soc.*, 75(2):240–284, 1969.

[46] M. Tibăr. Regularity at infinity of real and complex polynomial functions. *London Mathematical Society Lecture Note Series*, pages 249–264, 1999.

[47] M. Tibăr. *Polynomials and vanishing cycles*, volume 170. Cambridge University Press, 2007.

[48] B. L. Van der Waerden. Moderne algebra. 1937.

[49] A. N. Varchenko. Theorems on the topological equisingularity of families of algebraic varieties and families of polynomial mappings. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, 36(5):957–1019, 1972.

[50] J.-L. Verdier. Stratifications de Whitney et théoreme de Bertini-Sard. *Invent. Math.*, 36(1):295–312, 1976.

[51] H. H. Vui and P. T. Son. Critical values of singularities at infinity of complex polynomials. *Vietnam Journal of Mathematics*, 36(1):1–38, 2008.

[52] A. Wallace. Linear sections of algebraic varieties. *Indiana Univ. Math. J.*, 20(12):1153–1162, 1971.

[53] C.-K. Yap. *Fundamental problems of algorithmic algebra*. Oxford University Press, New York, 2000.

[54] A. Zaharia. On the bifurcation set of a polynomial function and Newton boundary. II. *Kodai Math. J.*, 19(2):218–233, 1996.

A Useful results and bounds

A.1 Multivariate polynomial multipication

We need the following result(s) on multivariate polynomial multiplication. For the rather straightforward proofs we refer the reader to [30].

Claim 26 (Output bounds on polynomial multiplication)**.** *The following bounds holds:*

(i) Consider two multivariate polynomials, f_1 and f_2, in ν variables of total degrees δ, having bitsize τ_1 and τ_2, respectively. Then $f = f_1 f_2$ is a polynomial in ν variables, of total degree 2δ and bitsize $\tau_1 + \tau_2 + 2\,\nu\,\lg(\delta)$.

(ii) Using induction, the product of m polynomials in ν variables, $f = \prod_{i=1}^{m} f_i$, each of total degree δ_i and bitsize τ_i, is a polynomial of total degree $\sum_{i=1}^{m} \delta_i$ and bitsize $\sum_{i=1}^{m} \tau_i + 12\,\nu\,m\,\lg(m)\,\lg(\sum_{i=1}^{m} \delta_i)$.

(iii) Let f be a polynomial in ν variables of total degree δ and bitsize τ. The m-th power of f, f^m, is a polynomial of total degree $m\delta$ and bitsize $m\tau + 12\nu m \lg(\delta)$.

A.2 Intersecting a variety

Theorem 27 ([42, Cor. 1.13, Sec. 6, Chapter 1])**.** *Consider an irreducible projective variety $V \subset \mathbb{P}^n$ of dimension δ. If a homogeneous polynomial f does not vanish on V, then $\dim(V \cap \mathbb{V}(f)) = \delta - 1$. In addition, all the components of the intersection have the same dimension.*

The following lemma is similar to the ones in [17, Chapter 3].

Lemma 28. *Consider a projective variety $V \in \mathbb{P}^n$ of dimension δ and degree D. Let S be finite set of integers such that $0 \notin S$. Also consider the linear form $h(\boldsymbol{x}) = \sum_{i=0}^{n} c_i x_i$ and $H = \mathbb{V}(h)$. If we choose the c_i's independently and uniformly at random from S, then*

$$\Pr[\dim(V \cap H) = \delta - 1] \geq 1 - \tfrac{D}{|S|}.$$

Moreover, there is a specialization of the c_i's such that their bitsize is $\lceil \lg(\kappa D) \rceil + 1$, where $\kappa \geq 2$ is a constant, so that it holds $\dim(V \cap H) = \delta - 1$.

Proof. Let $V = \bigcup_{j=1}^{r} V_j$ be the irreducible decomposition of V, where $\deg(V_j) \geq 1$ and $\dim(V_j) = \delta_j$. Notice that $r \leq D$, as the number of components is bounded by the degree of the variety; in addition, $D = \sum_{i=j}^{r} \deg(V_j)$. Let $h(\boldsymbol{x}) = \sum_{i=0}^{n} c_i x_i$, where we consider the c_i's as intermediates that we will specialize in the sequel.

First, we consider the component V_j. Due to Theorem 27, $\dim(V_j \cap H) = \delta_j - 1$ or $\dim(V_j \cap H) = \delta_j$. In the latter case $V_j \subseteq H$ and so if $\boldsymbol{p}_j = (p_{j0}, \ldots, p_{jn}) \in V_j$, then it holds $\boldsymbol{p}_j \in H$.

We evaluate h at $\boldsymbol{p}_j$, that is $h_j = \sum_{i=0}^{n} c_i p_{ji}$ and we consider h_j as a polynomial in the c_i's. Then h_j has at most $|S|^n$ solutions in the grid $\underbrace{S \times \cdots \times S}_{n+1} = S^{n+1}$. In other words there, are at most $|S|^n$ specializations of the c_i's in the grid S^{n+1} such that $\dim(V_j \cap H) = \delta_j$. Consequently,

$$\Pr[\dim(V_j \cap H) = \delta_j] \leq \tfrac{|S|^n}{|S|^{n+1}} = 1/|S|.$$

Thus

$$\Pr[\bigcup_{j=1}^{r} \{\dim(V_j \cap H) = \delta_j\}] \leq \sum_{j=1}^{r} \Pr[\dim(V_j \cap H) = \delta_j] \leq \tfrac{r}{|S|} \leq \tfrac{D}{|S|},$$

which implies

$$\Pr[\dim(V \cap H) = \delta - 1] = 1 - \Pr[\bigcup_{j=1}^{r} \{\dim(V_j \cap H) = \delta_j\}] \geq 1 - \frac{D}{|S|}.$$

If we choose as S the first κD positive integers, where $\kappa \geq 2$ is a constant, then there is at least one specialization of the c_i's in the grid S^{n+1} such that $\dim(V_j \cap H) = \delta_j - 1$, for all $j \in [r]$; and so the bound on the bitsize follows. □

Proof of Proposition 21. In the affine case, we have to also consider the case $V \cap H = \varnothing$, in addition to the two cases of Theorem. 27. Let $\bar{V}$ and $\bar{H}$ be the projective closures of V and H, respectively. It holds $\dim(\bar{V}) = \delta$ and $\deg(\bar{V}) = D$; also $\bar{V} \cap \bar{H} \neq \varnothing$, expect if $\dim(\bar{V}) = 0$.

If $V \cap H = \varnothing$, then intersection $\bar{V} \cap \bar{H}$ occurs at the hyperplane at infinity, say L_∞.

If we want to hold $\dim(V \cap H) = \delta = 1$, then (i) $\dim(\bar{V} \cap \bar{H}) = \delta - 1$ and (ii) $\bar{V} \cap \bar{H} \not\subseteq L_\infty$, unless $\bar{V} \cap \bar{H} = \varnothing$.

The second condition implies if $V \cap H = \varnothing$, then $\bar{V} \cap \bar{H}$ is contained at infinity. Thus, the dimension of $\bar{V} \cap \bar{H}$ is $\dim(\bar{V})$ and because affine varieties and their projective closures have the same dimension, it also holds that $\dim(V \cap H) = \delta$. We want to avoid this situation.

For the first condition, as $\bar{V}$ and $\bar{H}$ are projective varieties, we can use Lemma 28. Thus

$$\Pr \geq 1 - \frac{D}{|S|}$$

and we can also select c_i's.

For the second condition, we notice that $\bar{V} \not\subseteq L_\infty$. Hence, by Theorem 27, $\dim(\bar{\cap} L_\infty) = \delta - 1$. Now, consider the projective varieties $\bar{V} \cap L_\infty$ and $\bar{H}$.

If it happens, following Lemma 28 that

$$\Pr[\dim(\bar{V} \cap L_\infty \cap \bar{H}) = \delta - 2] \geq 1 - \frac{D}{|S|}$$

and we can also find appropriate c_i's.

If it happens that $\dim(\bar{V} \cap L_\infty \cap \bar{H}) = \delta - 2$, then $\bar{V} \cap \bar{H} \not\subseteq L_\infty$. This is so, because otherwise $\bar{V} \cap L_\infty \cap \bar{H} = \bar{V} \cap \bar{H}$, where the dimensions are $\delta - 2$ and $\delta - 1$, respectively. This is a contradiction.

Using the union bound, we deduce the required probability bound. □

A.3 Reduction to a square system

Our techniques on bounding the bitsize and the degrees of the various polynomials appearing in the polynomial optimization problems rely on exploiting the properties of the resultant. However, to use the resultant the corresponding polynomial systems have to have as many equations as variables; that is we need to compute with square systems. In the optimization problems that we are interested in, in almost all the cases, the number of polynomials is bigger than the number of unknowns.

We are able to treat this case using a result due to Giusti and Heinz [19]. Let $f_1, \ldots, f_p \in \mathbb{Q}[x_1, \ldots, x_n]$ be polynomials of positive degree, bounded by d. Denote by V the algebraic variety $\mathbb{V}(f_1, \ldots, f_p) \subseteq \mathbb{C}^n$. Given $\eta \in \overline{\mathbb{Q}}$, we denote by $\hat{f}_\eta$ the linear combination $f_1 + \eta^1 f_2 + \cdots + \eta^{p-1} f_p$.

Theorem 29. *[19, Section 3.4.1] Let $\Gamma \subset \mathbb{Z}$ of cardinal $p\, d^n + 1$. There exists $\gamma = (\gamma_1, \ldots, \gamma_n) \in \Gamma^n$ such that each irreducible component of $\hat{V} = \mathbb{V}(\hat{f}_{\gamma_1}, \ldots, \hat{f}_{\gamma_n})$ is either a component of V or a point.*

Proof. First, it is clear that for all $x \in V$, any linear combination of the polynomials $f_1, \ldots, f_p$ vanishes at x.

Then for $1 \leq i \leq n$ and $(\gamma_1, \ldots, \gamma_i) \in \Gamma^i$, let us denote by $\hat{V}_i$ the variety $\mathbb{V}(\hat{f}_{\gamma_1}, \ldots, \hat{f}_{\gamma_i})$. We prove by induction that for $1 \leq i \leq n$, there exist points $\gamma_1, \ldots, \gamma_i \in \Gamma$ such that the dimension of any irreducible component of $\hat{V}_i$ not contained in V is $n - i$.

The case $i = 1$ is obvious. Let $i > 1$ and assume that the result is proved for $i - 1$. Let $\hat{V}_{i-1} = \mathbb{V}(\hat{f}_{\gamma_1}, \ldots, \hat{f}_{\gamma_{i-1}})$ and let C be an irreducible component of $\hat{V}_{i-1}$ such that $C \not\subset V$. Let x_C be a point in $C \setminus V$, so that at least one polynomial among $f_1, \ldots, f_p$ does not vanish at x_C. Let T be a new indeterminate and let

$$w_C(T) = f_1(x_C) + T f_2(x_C) + \cdots + T^{p-1} f_p(x_C) \in \overline{\mathbb{Q}}[T].$$

By construction, w_C is not identically zero. Hence, so is the product $P_i(T) = \prod_C w_C(T)$, for all irreducible component C of $\hat{V}_{i-1}$ not included in V. By the Bézout bound, there are at most d^n such components. Hence, the polynomial $P_i(T)$ has degree at most pd^n so that it has at most pd^n roots. Since Γ has cardinal $pd^n + 1$, there exists $\gamma_i \in \Gamma$ such that $P_i(\gamma_i) \neq 0$. Let $\hat{f}_{\gamma_i} = f_1 + \gamma_i f_2 + \cdots + \gamma_i^{s-1} f_p$.

Then for any x_C as above, $\hat{f}_{\gamma_i}(x_C) \neq 0$. In particular, this means that $C \not\subset \mathbb{V}(\hat{f}_{\gamma_i})$.

By Krull's Principal Ideal Theorem [27, Corollary 3.2 p. 131], this implies that the intersection $C \cap \mathbb{V}(\hat{f}_{\gamma_i})$ is either empty or equidimensional of dimension $\dim C - 1$. By induction assumption, C has dimension $n - i + 1$. Hence,

$C \cap \mathbb{V}(\hat{f}_{\gamma_i})$ is either empty or of equidimensional of dimension $n-i$. Since this is true for any irreducible component C of $\hat{V}_{i-1}$ not contained in V, this proves that the dimension of any irreducible component of $\hat{V}_i$ not contained in V is $n-i$.

In particular for $i=n$, this proves that the irreducible components of $\hat{V}=\hat{V}_n$ not contained in V have dimension 0, so that it is necessarily a point. □

Remark 4. *The previous theorem implies that if we are given a polynomial system that it is not necessarily square, then we can make square by considering a generic linear combination. The price that we pay for this is that the bitsize of the polynomials becomes $\mathcal{O}(\tau + n\lg(pd)) = \widetilde{\mathcal{O}}(\tau+n)$ from τ and that we add additional points to the variety. For out purposes, the increase in the in bitsize is negligible.*