Sampling smooth points on real algebraic sets using perturbations

Katherine E. Harris* Jonathan D. Hauenstein † Hoon Hong ‡

January 2, 2025

Abstract

Many algorithms in computational algebraic geometry for understanding properties of real algebraic sets depend upon computing smooth sample points, most notably algorithms for computing the dimension of a real algebraic set. An approach by the first two authors and Agnes Szanto (1966-2022) obtained smooth sample points by computing the critical points of a possibly high degree polynomial that vanishes on the singular set and but does not vanish on the real algebraic set identically. This paper shows that smooth sample points can be obtained by using limits of perturbations thereby reducing the degree of the objective polynomial under consideration. This approach is then applied to computing the dimension of a real algebraic set.

Keywords. Smooth points, real algebraic sets, polynomial systems, perturbations, real numerical algebraic geometry, homotopy continuation, numerical algebraic geometry

1 Introduction

Let $Y \subset \mathbb{C}^n$ be a complex algebraic set and let $X \subset \mathbb{R}^n$ be the corresponding real algebraic set, that is, $X = Y \cap \mathbb{R}^n$. A common approach to studying the real algebraic set X is to compute smooth sample points in each connected component of X. Our initial motivation for computing smooth sample points came from a long standing open challenge (originally stated in [44] in 1999) to close the complexity gap between computing the real and complex cases of the dimension problem, i.e., determining if the dimension of the real algebraic set X could be computed with the same complexity as computing the dimension of the associated complex real algebraic set Y.

The approach in [20] considered the case when X was compact and found smooth sample points by computing the critical points of a polynomial $g \in \mathbb{R}[x_1, \ldots, x_n]$ on X where g was chosen such that it vanishes on every singular point of X but it does not vanish on X identically. To ensure that g did not vanish on X identically, isosingular deflation [23] was used in [20], which potentially resulted in a high degree polynomial g. In this paper, using perturbations and a limiting approach, we show that smooth sample points can be computed without needing deflation, thereby reducing the degree of g under consideration.

Due to the ubiquity of computing real solutions to systems of polynomial equations, many approaches have been proposed. Symbolic approaches for the related problem of determining

*Department of Mathematics & Computer Science, Beloit College, Beloit, WI 53511 (harriske@beloit.edu, `https://www.beloit.edu/live/profiles/5199-katherine-harris`)

†Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, IN 46556 (hauenstein@nd.edu, `https://www.nd.edu/~jhauenst`)

‡Department of Mathematics, North Carolina State University, Raleigh, NC 27695 (hong@ncsu.edu, `https://hong.math.ncsu.edu`)

if X is empty and computing sample points in each connected component of X can be found in [14, 18, 26, 33]. The current state of the art deterministic real sampling algorithm is given in [9, Alg. 13.3]. Symbolic methods computing critical points or generalized critical points of functions utilizing infinitesimals and randomization can be found in [1, 16, 34, 35, 37]. Sampling algorithms utilizing polar varieties, which will also be used below, are introduced and developed in [2,3,5–7,31,36]. Numerical approaches based on homotopy continuation are presented in [21,45].

The difficulty of finding smooth points in each connected component of a real algebraic set was pointed out in [31] where they constructed a singular plane curve with four real connected components such that critical points of the distance function from any point in $\mathbb{R}^2$ does not simultaneously yield smooth points on all components. More recently, [38,39] give the state of the art randomized symbolic algorithm to compute the real radical of an algebraic set with their method also including a subroutine that computes smooth points on real algebraic sets. For finding the dimension of X, the current state of the art deterministic algorithms are in [10, 29] and [9, Alg. 14.10]. Similarly, the best known probabilistic algorithms are [8, 30] and these algorithms work better in practice.

The remainder of the paper is structured as follows. Section 2 provides a precise statement of the problem that will be tackled while Section 3 provides preliminary but essential items needed for developing our proposed approach. Section 4 develops the approach for the compact cases. Section 5 generalizes the approach to arbitrary cases (including unbounded cases). Section 6 summarizes the findings in the previous sections into an algorithm for computing smooth sample points, determining the real dimension, and illustrating the algorithm on a few examples. Section 7 provides bounds on total number of points which need to be computed by the algorithm. The paper ends with a short conclusion in Section 8.

2 Problem statement

The following formulates the problem under consideration throughout this paper. Suppose that $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_n]$ are non-constant polynomials and Let

$$X = V(f_1, \ldots, f_k) \cap \mathbb{R}^n = \{x \in \mathbb{R}^n \mid f_1(x) = \cdots = f_k(x) = 0\}.$$

As stated in Introduction, the overall goal is to compute smooth sample points in each connected component of X. Furthermore, we would like to use the computed sample points to determine the dimension of X. Thus, we would like to organize smooth sample points in a manner that facilitates the determination of the dimension of X. The following problem statement provide such an organization.

Problem: Devise an algorithm for the following specification:

Input: Non-constant polynomials $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_n]$

Output: $\mathcal{Z}_0, \ldots, \mathcal{Z}_{n-1} \subset \mathbb{R}^n$ such that, for each $j = 0, \ldots, n-1$,

1. $\mathcal{Z}_j$ consists of finitely many real points in X;
2. each point in $\mathcal{Z}_j$ belongs to a component of X having dimension at least j;
3. $\mathcal{Z}_j$ contains at least one smooth point on every j-dimensional component of X.

The following are some immediate observations from these three conditions. First, the union of the output sets is a finite list of real points in X, which is common among the real sample algorithms as mentioned in the Introduction. Second, such an organization facilitates computing the dimension of X, i.e., $\dim X = \max\{j \mid \mathcal{Z}_j \neq \emptyset\}$. Finally, there is a smooth point computed of every component of X so that, in particular, $\mathcal{Z}_j = \emptyset$ implies that X has no connected components of dimension j.

3 Preliminaries

In this section, we will provide preliminary but essential items needed for developing our proposed approach to solve the problem stated in Section 2. Let $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_n]$ be non-constant polynomials. To avoid triviality, we assume that $k \geqslant 1$ and $n \geqslant 2$.

We will always reduce down to the single nonnegative polynomial case by replacing $f_1, \ldots, f_k$ with $\hat{f} = f_1^2 + \cdots + f_k^2$. This does not change the real algebraic set under consideration as

$$V(f_1, \ldots, f_k) \cap \mathbb{R}^n = V(\hat{f}) \cap \mathbb{R}^n.$$

The next simplification is to assume that the coordinates are in general position by replacing x by Ax for generic $A \in \mathbb{R}^{n \times n}$. Thus, we will consider $X = V(f) \cap \mathbb{R}^n$ throughout where

$$f(x) = \hat{f}(Ax) = f_1(Ax)^2 + \cdots + f_k(Ax)^2 \in \mathbb{R}[x_1, \ldots, x_n] \tag{1}$$

is a nonnegative polynomial of even degree $d \geqslant 2$ with coordinates in general position.

For $i = 1, \ldots, n$, let $\partial_i f$ be the partial derivative of f with respect to x_i. The key sets of interest are the polar varieties

$$V_{\epsilon,j} = V(f - \epsilon, \partial_{j+2} f, \ldots, \partial_n f) \subset \mathbb{C}^n \tag{2}$$

for $j = 0, \ldots, n-1$ where $\epsilon \in \mathbb{C}$. Trivially, we define $V_{\epsilon,-1} = \emptyset$.

For an algebraic set $Y = V(h_1, \ldots, h_\ell) \subset \mathbb{C}^n$, Y is said to be smooth pure-m-dimensional if every irreducible component of Y has dimension m and the dimension of the null space of the Jacobian matrix of $h_1, \ldots, h_\ell$ evaluated at every $y \in Y$ is m.

Proposition 3.1 *For $j = 0, \ldots, n-1$ and for all but finitely many values of $\epsilon \in \mathbb{C}$, $V_{\epsilon,j}$ is either empty or is a smooth pure-j-dimensional algebraic set.*

Proof. By Sard's theorem, $V(f - \epsilon)$ is a smooth hypersurface for all but finitely many values of $\epsilon \in \mathbb{C}$. The result then follows from polar varieties of a smooth hypersurface, e.g., see [4, Prop. 3] and [20, Cor. 6.4]. □

If one treats ϵ as an infinitesimal, $V_j = \lim_{\epsilon \to 0} V_{\epsilon,j}$ is well-defined and is either empty or pure-j-dimensional. Of course, V_j need not be smooth, e.g., $V_{n-1} = V(f)$ need not be smooth.

Alternatively, we can compute V_j by treating $\epsilon \in \mathbb{C}$. For example, this can be accomplished by by tracking along rays emanating from the origin, say $\epsilon_\theta(t) = te^{\theta\sqrt{-1}}$ where $\theta \in [0, 2\pi)$ and $t \geqslant 0$. The infinitesimal limit shows that, for every $\theta \in [0, 2\pi)$, $V_j = \lim_{t \to 0^+} V_{\epsilon_\theta(t),j}$. Moreover, for all but finitely many $\theta \in [0, 2\pi)$, $V_{\epsilon_\theta(t),j}$ is either empty or is a smooth pure-j-dimensional algebraic set for all $t > 0$. Hence, one is able to track along all but finitely many rays emanating from the origin from any point on the ray and limit to 0 to compute V_j.

Another equivalent description of V_j arises from treating ϵ as a variable. Suppose that B_j is a finite set so that, for every $\epsilon \in \mathbb{C} \backslash B_j$, $V_{\epsilon,j}$ has the same behavior as in Proposition 3.1. Thus,

$$Y_j = \overline{\bigcup_{\epsilon \in \mathbb{C} \backslash B_j} V_{\epsilon,j}} \subset \mathbb{C}^{n+1} \tag{3}$$

is a flat family in $(x, \epsilon) \in \mathbb{C}^{n+1}$ and

$$V_j \times \{0\} = Y_j \cap V(\epsilon) \subset \mathbb{C}^{n+1}. \tag{4}$$

Hence, defining equations for V_j can be computed algebraically via

$$(\langle f - \epsilon, \partial_{j+2} f, \ldots, \partial_n f\rangle : \epsilon^\infty + \langle \epsilon \rangle) \cap \mathbb{R}[x], \tag{5}$$

that is, computing the saturation of the ideal $\langle f - \epsilon, \partial_{j+2} f, \ldots, \partial_n f\rangle$ with respect to ϵ, setting ϵ to 0, and eliminating ϵ.

Example 3.2 *Consider $f_1(x) = x_2^2 - x_1^2 + x_1^4$ which defines the leminscate of Gerono that will be considered again in Example 3.8. For illustration purposes, we take $\hat{f}(x) = f_1(x)^2$ and*

$$f(x) = \hat{f}(Ax) = \left((3x_1 + 2x_2)^2 - (3x_2 - x_1)^2 + (3x_2 - x_1)^4\right)^2 \quad \text{where} \quad A = \begin{bmatrix} -1 & 3 \\ 3 & 2 \end{bmatrix}.$$

For generic $\epsilon \in \mathbb{C}$, $V_{\epsilon,0}$ is zero-dimensional and smooth consisting of 12 points in $\mathbb{C}^2$. The set $B_0 \subset \mathbb{C}$ where this statement does not hold is

$$B_0 = \left\{0, \frac{1}{16}, \frac{225625}{76527504}\right\}.$$

Therefore, for every ray not along the positive real direction, i.e., for every $\theta \in (0, 2\pi)$ and $t > 0$, $V_{\epsilon_\theta(t),0}$ is smooth and consists of 12 points. Moreover,

$$V_0 = \lim_{t \to 0^+} V_{\epsilon_\theta(t),0}$$

consists of 5 points, four as the limit of 2 paths each and the origin as the limit of 4 paths. Using `Macaulay2` *[17] to perform the computation in (5) yields the following defining equations for V_0:*

$$\begin{aligned} &6x_1^3 - 54x_1^2x_2 + 162x_1x_2^2 - 162x_2^3 - 9x_1 + 5x_2 = \\ &514976x_1^2x_2^2 - 1875984x_1x_2^3 + 2047320x_2^4 + 9747x_1^2 + 108300x_1x_2 - 63175x_2^2 = 0. \end{aligned}$$

These equations define a zero-dimensional ideal of degree 12 whose radical is a zero-dimensional ideal of degree 5 defined by

$$\begin{aligned} &-1672x_1x_2^2 - 10956x_2^3 + 285x_2 = \\ &-1672x_1^2x_2 - 10956x_1x_2^2 + 285x_1 = \\ &-36784x_1^3x_2 - 241032x_1^2x_2^2 + 331056x_1x_2^3 + 2169288x_2^4 - 8360x_1x_2 - 54780x_2^2 = 0. \end{aligned}$$

For $p \in X = V(f) \cap \mathbb{R}^n$, let $\dim_{\mathbb{R}}(f, p)$ denote the real local dimension of p with respect to X. For $y \in \mathbb{R}^n$, define $\|y\|^2 = y_1^2 + \cdots + y_n^2$. For $r > 0$, let

$$B(p, r) = \{x \in \mathbb{R}^n \mid \|x - p\|^2 < r^2\}.$$

Thus, $\dim_{\mathbb{R}}(f, p)$ is the dimension of $X \cap B(p, r)$ for all sufficiently small $r > 0$. Moreover, when $X \neq \varnothing$, the dimension of X is

$$\dim_{\mathbb{R}} X = \max_{p \in X} \dim_{\mathbb{R}}(f, p).$$

If $X = \varnothing$, we take $\dim_{\mathbb{R}} X = -1$.

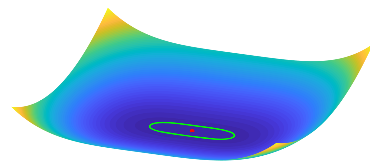
Treating $V_j \subset \mathbb{C}^n$ as a set, let $T_{V_j}(p)$ be the tangent space of V_j at $p \in V_j$. A point $p \in V_j$ is smooth if V_j and $T_{V_j}(p)$ have the same dimension. Otherwise, p is singular and the dimension of $T_{V_j}(p)$ is larger than the dimension of V_j. Let $\operatorname{Sing} V_j$ denote the set of singular points of V_j. The following shows that $V_j \cap \mathbb{R}^n$ must contain all points with real local dimension at most j and $\operatorname{Sing} V_j$ is contained in V_{j-1}.

Proposition 3.3 *For $j = 0, \ldots, n-1$:*

- *if $p \in V(f) \cap \mathbb{R}^n$ with $\dim_{\mathbb{R}}(f,p) \leqslant j$, then $p \in V_j \cap \mathbb{R}^n$, and*
- $\operatorname{Sing} V_j \subset V_{j-1}$ *where $V_{-1} = \varnothing$.*

Proof. Trivially, $V_{\epsilon,0} \subset V_{\epsilon,1} \subset \cdots \subset V_{\epsilon,n-1}$ implies $V_0 \subset V_1 \subset \cdots \subset V_{n-1}$. Suppose that $j \in \{0, \ldots, n-1\}$ and $p \in V(f) \cap \mathbb{R}^n$ with $0 \leqslant m = \dim_{\mathbb{R}}(f,p) \leqslant j$. The first result immediately follows by showing $p \in V_m$. The key additional property is $f \geqslant 0$ on $\mathbb{R}^n$ since f is a sum of squares. Let $\ell_{m,p}$ be a codimension m linear space passing through p defined by the vanishing of m linear polynomials $\ell_2, \ldots, \ell_{m+1}$ with real coefficients such that p is an isolated point of

$$X \cap \ell_{m,p} = V(f, \ell_2, \ldots, \ell_{m+1}) \cap \mathbb{R}^n.$$



Note that such a transverse intersection is a general property. Hence, p is a strict global minimum of f along the codimension m linear space $\ell_{m,p}$ so that f is strictly convex at p along $\ell_{m,p}$. For an infinitesimal $\epsilon > 0$, this implies that there is a smooth and bounded connected component of the level set $f - \epsilon = 0$ on $\ell_{m,p}$ of dimension $n-1-m$ infinitesimally close to p as illustrated by the figure on the right, that shows a surface plot of a strictly convex function with a smooth and bounded level set (green loop) near the strict global minimum (red point).

The addition of $n-1-m$ critical conditions applied to a smooth and bounded connected component of dimension $n-1-m$ ensures that there is a point infinitesimally close to p in

$$V_{\epsilon,m} \cap \ell_{m,p} = V(f - \epsilon, \ell_2, \ldots, \ell_{m+1}, \partial_{m+2} f, \ldots, \partial_n f) \text{ so that } p \in \lim_{\epsilon \to 0} (V_{\epsilon,m} \cap \ell_{m,p}) \subset V_m.$$

If $j = 0$, one trivially has $\operatorname{Sing} V_0 = \varnothing = V_{-1}$. Thus, assume that $1 \leqslant j \leqslant n-1$ and $p \in \operatorname{Sing} V_j$. Since $\operatorname{Sing} V_{j-1}$ has codimension at least one in V_j, we can take $\ell_{j-1,p}$ to be a codimension $j-1$ linear space passing through p defined by the vanishing of $j-1$ linear polynomials $\ell_2, \ldots, \ell_j$ with real coefficients such that $\ell_{j-1,p}$ intersects V_j and V_{j-1} transversely at p, i.e., $V_j \cap \ell_{j-1,p}$ is a curve containing p such that p is an isolated point in $\operatorname{Sing} V_j \cap \ell_{j-1,p}$ and, if $p \in V_{j-1}$, then p is an isolated point in $V_{j-1} \cap \ell_{j-1,p}$. Thus, as an isolated singularity of a curve, it follows from [43] that $V_{\epsilon,j} \cap \ell_{j-1,p}$ has a "bouquet of spheres" infinitesimally close to p that surround p expect possibly along the finitely many tangent cone directions of V_j at p. Thus, when p is a general point on an irreducible component of $\operatorname{Sing} V_j$, general position of the $(j+1)^{\text{st}}$ coordinate shows that there must be a point on the "bouquet of spheres" when applying one additional critical condition which yields a point in $V_{\epsilon,j-1} \cap \ell_{j-1,p}$ infinitesimally close to p and thus $p \in V_{j-1}$. Hence, the second result follows since V_{j-1} is an algebraic set by looping over each irreducible component of $\operatorname{Sing} V_j$. □

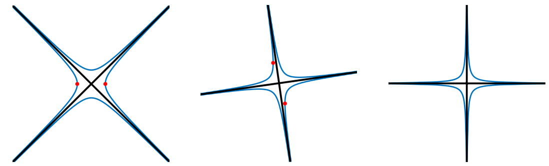
Example 3.4 *Consider $f_1(x) = x_1^2 - x_2^2$ which defines a pair of lines passing through the origin. Let $\hat{f}(x) = f_1(x)^2$ and*

$$f(x) = \hat{f}(Ax) = \left((x_1 - tx_2)^2 - (x_2 + tx_1)^2\right)^2 \quad \text{where} \quad A = \begin{bmatrix} 1 & -t \\ t & 1 \end{bmatrix}.$$

Locally at the origin, $V_1 = V(f) = V(f_1)$ is described by the pair of lines

$$(1-t)x_1 + (1+t)x_2 = 0,$$
$$(1+t)x_1 + (1-t)x_2 = 0.$$

Hence, for all $t \neq \pm 1$, there will be at least one point in $V_{\epsilon,0}$ infinitesimally close to the origin. In this case, there will be four such points with the figure on the right considering the cases for (a) $t = 0$, (b) $t = 3/4$, and (c) $t = 1$ plotted using $\epsilon = 10^{-3}$. In particular, the first two have two real (red points) and a pair of complex conjugate points infinitesimally close to the origin while the last has no points as expected.



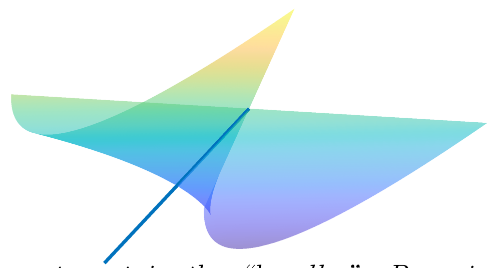
(a) (b) (c)

Example 3.5 *Consider $f_1(x) = x_1^2 - x_2^2 x_3$ which defines the Whitney umbrella. For illustration purposes, we take $\hat{f}(x) = f_1(x)^2$ and*

$$f(x) = \hat{f}(Ax) = \left((2x_1 - 3x_2 + x_3)^2 - (3x_1 + 2x_2 + 2x_3)^2(x_1 + 3x_2 + 2x_3)\right)^2 \quad \text{where} \quad A = \begin{bmatrix} 2 & -3 & 1 \\ 3 & 2 & 2 \\ 1 & 3 & 2 \end{bmatrix}.$$

The figure on the right illustrates $X = V(f) \cap \mathbb{R}^3$, which one observes that X has a singular set, namely the "handle" of the Whitney umbrella defined by

$$\{x \in \mathbb{R}^3 \mid 2x_1 - 3x_2 + x_3 = 3x_1 + 2x_2 + 2x_3 = 0\},$$



and points with both local dimension 2 and 1. In particular, the points on the "handle" with $x_1 + 3x_2 + 2x_3 < 0$ have local dimension 1. Proposition 3.3 shows that V_1 must contain the "handle." By using `Macaulay2` *[17] with (5), one obtains that V_1 indeed contains the "handle" along with a cubic curve.*

Example 3.6 *Consider $f_1(x) = x_2 - x_1^2$ and $f_2(x) = x_3^2 - x_1 x_2^2$ which together define a curve having a cusp at the origin. Let $f(x) = \hat{f}(Ax) = f_1(Ax)^2 + f_2(Ax)^2$ for a general $A \in \mathbb{R}^{3\times 3}$. The set $V_2 = V(f)$ is a hypersurface of degree 6, while V_1, as a set, is a curve of degree 15 with an isolated singularity at the origin. Moreover, $V_{\epsilon,0}$ contains 10 points infinitesimally close to the origin so that V_0 contains the origin. There are also 15 other points in V_0.*

Proposition 3.3 yields the following result regarding $\dim_{\mathbb{R}} X$.

Corollary 3.7 *Let $X = V(f) \cap \mathbb{R}^n$ and note that $V_{-1} = \varnothing$. Then,*

- *$X = \varnothing$, i.e., $\dim_{\mathbb{R}} X = -1$, if and only if $V_j \cap \mathbb{R}^n = \varnothing$ for all $j = 0, \ldots, n-1$; and*
- *$\dim_{\mathbb{R}} X = m \geqslant 0$ if and only if $V_{m-1} \cap \mathbb{R}^n \subsetneq V_m \cap \mathbb{R}^n = V_{m+1} \cap \mathbb{R}^n = \cdots = V_{n-1} \cap \mathbb{R}^n = X$.*

Proof. Since $V_j \cap \mathbb{R}^n \subset X$, $X = \varnothing$ implies $V_j \cap \mathbb{R}^n = \varnothing$ for $j = 0, \ldots, n-1$.

Conversely, if $X \neq \varnothing$, Proposition 3.3 shows there must exist $j \in \{0, \ldots, n-1\}$ with $V_j \cap \mathbb{R}^n \neq \varnothing$.

Suppose that $\dim_{\mathbb{R}} X = m \geqslant 0$. By Proposition 3.3, $X = V_j \cap \mathbb{R}^n$ for $j = m, m+1, \ldots, n-1$. Since the dimension of V_{m-1} is at most $m-1$, it can not contain X.

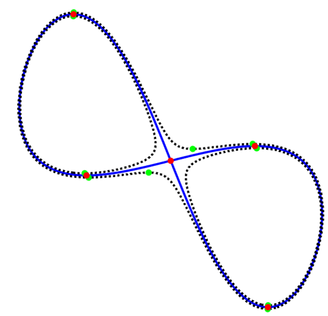
Conversely, suppose that $m \in \{0, \ldots, n-1\}$ such that

$$V_{m-1} \cap \mathbb{R}^n \subsetneq V_m \cap \mathbb{R}^n = V_{m+1} \cap \mathbb{R}^n = \cdots = V_{n-1} \cap \mathbb{R}^n = X.$$

Hence, V_m is not empty so that it must have dimension m yielding $\dim_{\mathbb{R}} X \leqslant m$. If $\dim_{\mathbb{R}} X < m$, then Proposition 3.3 yields $X = V_{m-1} \cap \mathbb{R}^n$ which is a contradiction. Hence, $\dim_{\mathbb{R}} X = m$. □

The following two examples consider a compact and unbounded real variety, respectively.

Example 3.8 *For a compact example, reconsider Example 3.2. For generic* $\epsilon \in \mathbb{C}$, $V_{\epsilon,1} = V(f-\epsilon)$ *is a smooth curve of degree* 8 *and* $V_{\epsilon,0} = V(f-\epsilon, \partial_2 f)$ *consists of* 12 *points.*

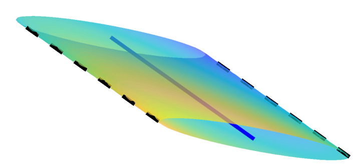


For the figure on the right, the solid blue curve is $X = V(f) \cap \mathbb{R}^2$. *For* $\epsilon = 3/20000$, $V_{\epsilon,1} \cap \mathbb{R}^2$ *is the dotted black curve while* $V_{\epsilon,0} \cap \mathbb{R}^2$ *consists of the 10 green points (the other two points are complex conjugates near the origin) in the figure. Now,* $V_1 = \lim_{\epsilon\to 0} V_{\epsilon,1} = V(f)$ *is a curve of degree* 4 *and multiplicity* 2 *with respect to* f *(which is not smooth in this case) while* $V_0 = \lim_{\epsilon\to 0} V_{\epsilon,0}$ *consists of 5 points (4 are the limit of two paths and the origin is the limit of 4 paths), which are the red points in the figure. Thus,* $V_0 \cap \mathbb{R}^2 \subsetneq V_1 \cap \mathbb{R}^2 = X$ *showing* $\dim_{\mathbb{R}} X = 1$.

Example 3.9 *For an unbounded example in* $\mathbb{R}^3$, *consider* $f_1(x) = x_1 + 1$ *and* $f_2(x) = x_2 - 1$ *which collectively defines a line in* $\mathbb{R}^3$. *Thus, we take* $\hat{f}(x) = f_1(x)^2 + f_2(x)^2$ *and, for illustration purposes,*

$$f(x) = \hat{f}(Ax) = (x_1 - x_2 + 2x_3 + 1)^2 + (-x_1 + 2x_2 - x_3 - 1)^2 \quad \textit{where} \quad A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & 1 & 2 \end{bmatrix}.$$

For generic $\epsilon \in \mathbb{C}$, $V_{\epsilon,2} = V(f-\epsilon)$ *is a smooth surface of degree* 2, $V_{\epsilon,1} = V(f-\epsilon, \partial_3 f)$ *consists of two lines, and* $V_{\epsilon,0} = \varnothing$.



Look at the figure on the right. The solid blue line in the figure on the right is $X = V(f) \cap \mathbb{R}^3$. *For* $\epsilon = 1$, $V_{\epsilon,2} \cap \mathbb{R}^3$ *is the surface and* $V_{\epsilon,1} \cap \mathbb{R}^3$ *is the union of the two dotted black lines. Since*

$$X = V_2 \cap \mathbb{R}^3 = V_1 \cap \mathbb{R}^3 \supsetneq V_0 \cap \mathbb{R}^3 = \varnothing,$$

$\dim_{\mathbb{R}} X = 1$. *Of course,* $V_2 = V(f)$ *is a union of two complex hyperplanes that intersect in the real line* X. *Since* X *is smooth and* $\dim_{\mathbb{R}}(f,p) = 1$ *for every* $p \in X$, *Proposition 3.3 provides no information regarding* $V_0 \cap \mathbb{R}^3$ *which, in this case, is empty (cf., [36]).*

As mentioned above, defining equations for V_j can be computed using (5) and thus one can test membership in V_j by simply evaluating the defining equations to determine if they vanish or not. On the other hand, numerical algebraic geometry also provides a membership test [41]. Given a point $q \in \mathbb{C}^n$, let $\ell_{j,q} \subset \mathbb{C}^n$ be a general linear space of codimension j passing through q. Thus, $q \in V_j$ if and only if $q \in V_j \cap \ell_{j,q}$ where $V_j \cap \ell_{j,q}$ contains at most finitely many points with $V_j \cap \ell_{j,q} = \lim_{\epsilon\to 0}(V_{\epsilon,j} \cap \ell_{j,q})$ facilitating computation via homotopy continuation. Moreover, [24] shows how one can apply isosingular deflation [23] to (4) to restore local quadratic convergence of Newton's method to compute the points in $V_j \cap \ell_{j,q}$ to arbitrary precision.

Example 3.10 *The symbolic computation of* V_0 *in Example 3.2 has a numerical counterpart with* $V_0 = \lim_{\epsilon\to 0} V(f-\epsilon, \partial_2 f)$. *Tracking 12 paths yield the 5 points in* V_0, *which, to 4 decimal places, are:*

$$(0,0), \quad (0.2772, -0.1838), \quad (-0.2772, 0.1838), \quad (0.0278, 0.1592), \quad (-0.0278, -0.1592).$$

As mentioned in Example 3.2, the origin is the limit of 4 paths and the other 4 points are the limit of 2 paths each with $12 = 1 \cdot 4 + 4 \cdot 2$.

4 Smooth points on compact real algebraic sets

We consider the problem stated in Section 2 for the compact case with the arbitrary case (including the unbounded case) considered in Section 5. Following Section 3, suppose that $f \in \mathbb{R}[x_1, \ldots, x_n]$ is a sum of squares polynomial of degree $d \geqslant 2$ with coordinates in general position and $X = V(f) \cap \mathbb{R}^n$ is compact. The key of the proposed approach to compute smooth points is to utilize limits of perturbations together with an adaptive objection function dependent upon f. That is, for $j = 0, \ldots, n-1$, let $C_{\epsilon,j}$ be the set consisting of

$$\text{critical points of } \ \partial_{j+1} f(x) \ \text{ with respect to } \ V_{\epsilon,j} = V(f - \epsilon, \partial_{j+2} f, \ldots, \partial_n f). \tag{6}$$

Note that $\partial_{j+1} f$ has degree $d-1$ while the objective function used in [20] has degree at most $n^{k+1} d$ where k is the number of iterations of isosingular deflation needed, which is dependent on the multiplicity structure. Let ∇h denote the gradient of h. Using the Fritz John condition [28],

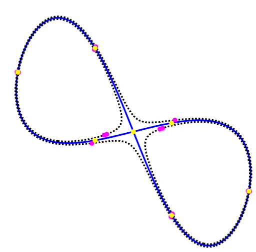
$$C_{\epsilon,j} = \left\{ x \in \mathbb{C}^n \;\middle|\; \exists \lambda \in \mathbb{P}^{n-j} \text{ such that } \begin{bmatrix} f(x) - \epsilon \\ \partial_{j+2} f(x) \\ \vdots \\ \partial_n f(x) \\ \lambda_0 \nabla f(x) + \lambda_1 \nabla \partial_{j+1} f(x) + \lambda_2 \nabla \partial_{j+2} f(x) + \cdots + \lambda_{n-j} \nabla \partial_n f(x) \end{bmatrix} = 0 \right\}.$$

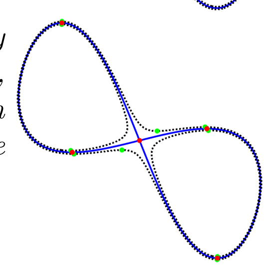
Note that, since $V_{\epsilon,0}$ is either empty or zero-dimensional, it is easy to see that $C_{\epsilon,0} = V_{\epsilon,0}$. In fact, for all $j = 0, \ldots, n-1$, $C_{\epsilon,j}$ is generically zero-dimensional.

Proposition 4.1 *For each $j = 0, \ldots, n-1$ and for all but finitely many values of $\epsilon \in \mathbb{C}$, $C_{\epsilon,j}$ is zero-dimensional.*

Proof. The result follows from Proposition 3.1. □

As with $V_{\epsilon,j}$ limiting to V_j, similar statements hold with $C_j = \lim_{\epsilon \to 0} C_{\epsilon,j}$ being well-defined.





Example 4.2 *Continuing with the setup from Examples 3.2 and 3.8, $C_{\epsilon,1}$ generically consists of 16 points while $C_{\epsilon,0} = V_{\epsilon,0}$ has 12 points. Look at the figure on the top right. It illustrates the critical points for $\epsilon = 3/20000$. The dotted black curve is $V_{\epsilon,1} \cap \mathbb{R}^2$ while the solid blue curve is $V_1 \cap \mathbb{R}^2$. The 14 magenta points form $C_{\epsilon,1} \cap \mathbb{R}^2$ (other two points in $C_{\epsilon,1}$ form a complex conjugate pair near the origin) while the 7 yellow points form $C_1 \cap \mathbb{R}^2$.*

Look at the figure on the bottom right. It is a copy of the figure already shown in Example 3.8, shown here again for convenience. Since $V_0 = C_0$, which are the five red points in the figure, one observes that the origin, which is the singular point of the lemniscate, is contained in both C_1 and V_0, while $(C_1 \backslash V_0) \cap \mathbb{R}^2$ consists of 6 smooth points.

The following shows $Z_j = (C_j \backslash V_{j-1}) \cap \mathbb{R}^n$ solves the problem in Section 2 in the compact case.

Theorem 4.3 *For $j = 0, \ldots, n-1$, the set $Z_j = (C_j \backslash V_{j-1}) \cap \mathbb{R}^n$ satisfies the following conditions:*

1. *$Z_j \subset X$ consists of finitely many real points;*

2. *for each* $p \in Z_j$, $\dim_{\mathbb{R}}(f,p) \geqslant j$;

3. Z_j *contains at least one smooth point on every* j*-dimensional component of* X.

Proof. The first statement follows from Proposition 4.1. The second statement follows from Proposition 3.3. Since the third statement is trivial if X does not have a j-dimensional component, let's assume that X has a j-dimensional component. If a polynomial g vanishes on the singular set of V_j and does not vanish identically on V_j, then [20] shows that the third statement holds for the limit of

$$\text{critical points of } \; g(x) \; \text{ with respect to } \; V_{\epsilon,j} = V(f-\epsilon, \partial_{j+2}f, \ldots, \partial_n f).$$

Let $o \in \mathbb{Q}_{\geqslant 0}$ and consider

$$\text{critical points of } \; \frac{\partial_{j+1}f(x)}{\epsilon^o} \; \text{ with respect to } \; V\left(\frac{f-\epsilon}{\epsilon^o}, \frac{\partial_{j+2}f}{\epsilon^o}, \ldots, \frac{\partial_n f}{\epsilon^o}\right). \tag{7}$$

Since the critical points of (7) satisfy the same conditions as $C_{\epsilon,j}$ for any $o \in \mathbb{Q}_{\geqslant 0}$, all that is left to show is that there exists an appropriate exponent o for each j-dimensional component of X.

Let Y_j be as in (3), which is not empty by assumption above. Let W be an irreducible component of V_j and let $x^* \in W$ be a general point. Hence, $(x^*, 0) \in Y_j$ and we want to consider the behavior of $g_{j,o}(x) := \frac{\partial_{j+1}f(x)}{\epsilon^o}$ for $(x,\epsilon) \in Y_j$ in a neighborhood of $(x^*, 0)$.

Let G_{j,x^*} be an irreducible germ of Y_j at $(x^*, 0)$. Hence, G_{j,x^*} is topologically unibranched at $(x^*, 0)$ and $\dim G_{j,x^*} = n - j + 1$. Let $\pi : \mathbb{C}^{n+1} \to \mathbb{C}^{n-j+1}$ be a general linear projection such that $\pi(x^*, 0) = 0$. Thus, π is a proper branched covering from G_{j,x^*} to a local neighborhood of the origin with, say, $\mu \geqslant 1$ sheets. In particular, the Fundamental Openness Principal [32, pg. 43] shows that π is an open map at $(x^*, 0)$. Let $\mathrm{Gr}(1, \mathbb{C}^{n-j+1})$ be the Grassmaninan of lines in $\mathbb{C}^{n-j+1}$ through the origin. Since π is an open map, the local behavior of the polynomial $\partial_{j+1}f(x)$ on the irreducible germ G_{j,x^*} can be completely characterized by computing $\partial_{j+1}f(x)$ on $G_{j,x^*} \cap \pi^{-1}(\ell)$ for general $\ell \in \mathrm{Gr}(1, \mathbb{C}^{n-j+1})$.

For a general $\ell \in \mathrm{Gr}(1, \mathbb{C}^{n-j+1})$, $G_{j,x^*} \cap \pi^{-1}(\ell)$ is a one-dimensional germ at $(x^*, 0)$ such that ϵ is nonconstant on each irreducible component of $G_{j,x^*} \cap \pi^{-1}(\ell)$. In particular, $G_{j,x^*} \cap \pi^{-1}(\ell)$ consists of precisely μ paths, say, $p_1(\epsilon), \ldots, p_\mu(\epsilon)$, such that, for each $k = 1, \ldots, \mu$, $\partial_{j+1}f(p_k(\epsilon))$ has a well-defined order in ϵ at 0, say $o_k \in \mathbb{Q}_{\geqslant 0}$. That is, $o_k \geqslant 0$ is the smallest number such that

$$\lim_{\epsilon \to 0} \frac{\partial_{j+1}f(p_k(\epsilon))}{\epsilon^{o_k}} \in \mathbb{C} \backslash \{0\}.$$

Therefore, one can take the minimum of o_k for $k = 1, \ldots, \mu$ resulting from G_{j,x^*} and then take the minimum over the irreducible germs G_{j,x^*} of Y_j at $(x^*, 0)$. Since this is a minimum of a finite number of rational numbers, this yields a rational number $o_W \geqslant 0$. Since $x^* \in W$ was general, this process shows that the limit of g_{j,o_W} is generically nonzero along W. The result now follows by repeating this process for each irreducible component W of V_j. □

Remark 4.4 *The key aspect of this proof is the exploitation of local irreducibility to obtain a proper branched covering over an affine space defined by an open map. If one does not have an open map, then such a minimum need not exist, e.g., for* $0 < k < 1$, $x(\epsilon) = (\epsilon^{k/2}, \epsilon^{1-k})$ *limits to the origin and satisfies* $x_1^2 x_2 - \epsilon = 0$, *but* $2x_1x_2 = 2e^{1-k/2}$ *has no minimum order for* $k \in (0,1)$. *An open map does exist at generic points along the two components arising in the limit, say at* $(\alpha, 0)$ *and* $(0, \beta)$ *for* $\alpha, \beta \neq 0$, *and thus yield a minimum order, namely* 1 *and* 1/2, *respectively.*

Example 4.5 *Continuing with Example 4.2, Y_1 as defined in (3) is irreducible of degree 8 and $V_1 = V(f)$ is irreducible of degree 4. For simplicity of presentation, let $x^* = (2/11, -3/11) \in V_1$,*

$$\pi(x, \epsilon) = (2(x_1 - 2/11) - 5(x_2 + 3/11) + 2\epsilon, -3(x_1 - 2/11) + (x_2 + 3/11) + 3\epsilon),$$

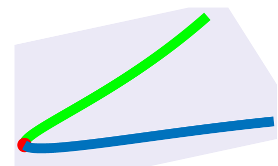
and $\ell = \{(-\alpha, 2\alpha) \mid \alpha \in \mathbb{C}\}$. Hence, $\pi(x^*, 0) = 0 \in \ell$.

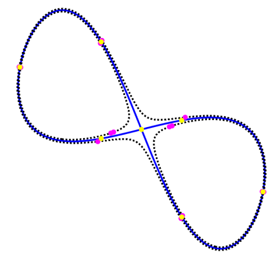
Look at the figure on the top right. It provides an illustration of

$$G_{1,x^*} \cap \pi^{-1}(\ell)$$

with each of the $\mu = 2$ paths $p_1(\epsilon)$ and $p_2(\epsilon)$ colored differently. In particular, the order of $\partial_2 f(p_k(\epsilon))$ is 1/2 for $k = 1, 2$. Hence, we can replace ϵ by t^2 and, for paths $p(t) \in C_{t^2,1}$, we can compute $v = \lim_{t \to 0} \partial_1 f(p(t))/t$.

Look at the figure on the bottom right. It is a copy of the figure already shown in Example 4.2 and provided again for convenience. For the six yellow dots not at the origin, each is obtained as a limit of two different paths for which the corresponding values of v are nonzero and negatives of each other. For the four paths leading to the origin, one has $v = 0$ as expected.





5 Smooth points on real algebraic sets

The following generalizes Theorem 4.3 for the compact case to arbitrary case (see Theorem 5.4). For unbounded components, there need not be critical points in $\mathbb{R}^n$ as illustrated in the following.

Example 5.1 *For the parabola from $f(x) = ((2x_2 - 3x_1) - (2x_1 + 5x_2)^2)^2$, $C_{\epsilon,1}$ is generically empty.*

This suggests that the definition of $C_{\epsilon,j}$ needs to be altered to ensure that critical points are obtained. Motivated by [15, 27], for general $c \in \mathbb{R}^n$, consider replacing (6) with

$$\text{critical points of } \frac{\partial_{j+1} f(x)}{D(x)} \text{ with respect to } V\left(\frac{f-\epsilon}{D}, \frac{\partial_{j+2} f}{D}, \ldots, \frac{\partial_n f}{D}\right). \tag{8}$$

where

$$D(x) = (\|x - c\|^2 + 1)^{d/2} \quad \text{and} \quad \|x - c\|^2 = (x_1 - c_1)^2 + \cdots + (x_n - c_n)^2.$$

In particular, since d is even, $\partial_{j+1} f$ has degree $d - 1$ while D has degree d and does not vanish on $\mathbb{R}^n$. Thus, $\partial_{j+1} f(x)/D(x)$ is continuous and bounded on $\mathbb{R}^n$ and zero when $\partial_{j+1} f(x)$ vanishes as well as at all real points at infinity. Since $\partial_{j+1} f$ does not vanish identically on $V_{\epsilon,j}$ for generic $\epsilon \in \mathbb{C}$, in the parlance of [15, 27], $\partial_{j+1} f(x)/D(x)$ is a routing function for general $c \in \mathbb{R}^n$. Since our primary concern here is computing sample points and not a full decomposition, we focus on properties of the critical points. In particular, the critical points satisfy

$$\mathcal{C}_{\epsilon,j} = \left\{ x \in \mathbb{C}^n \;\middle|\; \exists \lambda \in \mathbb{P}^{n-j}, \delta \in \mathbb{C} \text{ where } \begin{bmatrix} f(x) - \epsilon \\ \partial_{j+2} f(x) \\ \vdots \\ \partial_n f(x) \\ M(x, \lambda) \\ 1 - \delta(\|x - c\|^2 + 1) \end{bmatrix} = 0 \right\} \tag{9}$$

where

$$M(x, \lambda) = (\|x - c\|^2 + 1)\,(\lambda_0 \nabla f(x) + \lambda_1 \nabla \partial_{j+1} f(x) + \cdots + \lambda_{n-j} \nabla \partial_n f(x)) - d\lambda_1 \partial_{j+1} f(x)(x - c).$$

The last condition in (9) is Rabinowitz trick to ensure that critical points satisfy $D(x) \neq 0$. This could also be used to simplify $M(x, \lambda)$, i.e.,

$$M(x, \lambda) = \lambda_0 \nabla f(x) + \lambda_1 \nabla \partial_{j+1} f(x) + \cdots + \lambda_{n-j} \nabla \partial_n f(x) - d\delta\lambda_1 \partial_{j+1} f(x)(x - c).$$

Proposition 5.2 *For each $j = 0, \ldots, n-1$ and for all but finitely many values of $\epsilon \in \mathbb{C}$, $\mathcal{C}_{\epsilon,j}$ is zero-dimensional.*

Proof. The result follows from Proposition 3.1. □

As with $C_{\epsilon,j}$ limiting to C_j, similar statements hold with $\mathcal{C}_j = \lim_{\epsilon \to 0} \mathcal{C}_{\epsilon,j}$ being well-defined.

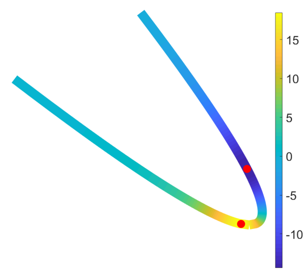
Example 5.3 *Let us continue with Example 5.1 where we considered*

$$f(x) = ((2x_2 - 3x_1) - (2x_1 + 5x_2)^2)^2.$$

For illustration, consider $c = 0$ where $\mathcal{C}_{\epsilon,1}$ has 8 points. Moreover, $\mathcal{C}_1$ consists of 4 points (each the limit of two points) with the figure on the right showing the two points in $\mathcal{C}_1 \cap \mathbb{R}^2$ along with value of the limit of

$$\partial_2 f(x)/(\epsilon^{1/2}(\|x\|^2 + 1)^2)$$

with the other limiting branch being the negative of this value.



Finally we are ready to state the main theorem of this paper.

Theorem 5.4 (Main Result) *For $j = 0, \ldots, n-1$, the set $\mathcal{Z}_j = (\mathcal{C}_j \backslash V_{j-1}) \cap \mathbb{R}^n$ satisfies the following conditions:*

1. *$\mathcal{Z}_j$ consists of finitely many real points in X;*
2. *each point in $\mathcal{Z}_j$ belongs to a component of X having dimension at least j;*
3. *$\mathcal{Z}_j$ contains at least one smooth point on every j-dimensional component of X.*

Proof. Since the proof is essentially the same as that of Theorem 4.3, it is omitted. In fact, we only have to replace $C_{\epsilon,j}$ with $\mathcal{C}_{\epsilon,j}$ in Theorem 4.3 to remove the compactness assumption of X. □

6 Algorithm

Theorem 5.4 immediately yields an algorithm that solves the problem posed in Section 2. By using the output of such an algorithm, one can immediately determine the dimension of X since

$$\dim X = \max\{j \mid \mathcal{Z}_j \neq \emptyset\}.$$

However, if one only wants to find the dimension X, then the algorithm can be structured so that $\mathcal{Z}_j$ are computed in the decreasing order on j starting form $j = n-1$ and returns j as soon as the first $\mathcal{Z}_j$ is determined to be nonempty yielding the dimension of X. Thus, we found that it is economical to formulate a single algorithm (Algorithm 1) for both situations:

Algorithm 1: Smooth Sample Points (SSP) or Dimension (DIM)

Input:

- Non-constant polynomials $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_n]$,
- $what \in \{\text{SSP, DIM}\}$.

Output:

- If $what = \text{SSP}$, then output $\mathcal{Z}_0, \ldots, \mathcal{Z}_{n-1} \subset \mathbb{R}^n$ such that
  1. $\mathcal{Z}_j$ consists of finitely many real points in X;
  2. each point in $\mathcal{Z}_j$ belong to a component of X having dimension at least j;
  3. $\mathcal{Z}_j$ contains at least one smooth point on every j-dimensional component of X.
- If $what = \text{DIM}$, then output the dimension of $X = V(f_1, \ldots, f_k) \cap \mathbb{R}^n$.

Randomly select $A \in \mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$, and $\theta \in [0, 2\pi)$;
$\epsilon(t) \leftarrow te^{\theta\sqrt{-1}}$;
for $j = n-1, n-2, \ldots, 0$ **do**
  $\mathcal{Z}_j \leftarrow \varnothing$;
  $\mathcal{C}_j \leftarrow \lim_{t \to 0^+} \mathcal{C}_{\epsilon(t),j}$ where $\mathcal{C}_{\epsilon(t),j}$ is given by (1) and (9);
  foreach $p \in \mathcal{C}_j \cap \mathbb{R}^n$ **do**
    if $p \in V_{j-1}$ **then** go to the next iteration on p;
    if $what = \text{SSP}$ **then** add $A \cdot p$ to $\mathcal{Z}_j$;
    if $what = \text{DIM}$ **then return** j;
  end
end
if $what = \text{SSP}$ **then return** $\mathcal{Z}_0, \ldots, \mathcal{Z}_{n-1}$;
if $what = \text{DIM}$ **then return** -1;

- computing the sets $\mathcal{Z}_0, \ldots, \mathcal{Z}_{n-1}$ as in the problem stated in Section 2;
- determining the dimension of X.

Theorem 6.1 *Assuming the randomly selected $A \in \mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$, and $\theta \in [0, 2\pi)$ are generic choices, Algorithm 1 is correct.*

Proof. Immediate from Theorem 5.4 and the obvious fact that $\dim X = \max\{j \mid \mathcal{Z}_j \neq \varnothing\}$. □

In the following examples, we make the choices in Algorithm 1 to be simple for illustrative purposes. In practice, one often generates each entry independently from a continuous distribution, e.g., each entry of A and c following a standard Gaussian distribution and θ from a uniform distribution on $(0, 2\pi)$.

From an implementation perspective, efficiencies can be obtained in Algorithm 1 by reusing some computations. Trivially, $\mathcal{C}_0 = V_0$. Since $V_{\epsilon,j-1} = V_{\epsilon,j} \cap V(\partial_{j+1} f)$, a witness set for $V_{\epsilon,j-1}$ can be obtained from a witness set for $V_{\epsilon,j}$ via intersection, e.g., see [25, Problem 1]. Finally, since $\mathcal{C}_{\epsilon,j}$ arises from intersecting $V_{\epsilon,j}$ with critical point conditions, $\mathcal{C}_{\epsilon,j}$ can be computed from $V_{\epsilon,j}$ as an extension computation [25, Problem 3].

Since the endpoints arising from the limits as $\epsilon \to 0$ are often singular, adaptive precision path tracking [12] and endgames (e.g., see [42, Chap. 10] and [13, Chap. 3]) can be used on the path tracking side and isosingular deflation [23] can be used to restore local quadratic convergence of Newton's method to refine the numerical approximations.

Example 6.2 *The following is a step by step trace of Algorithm 1 on Example 3.9 for computing the real dimension.*

Input:

- $f_1 = x_1 + 1$ *and* $f_2 = x_2 - 1$ *in* $\mathbb{R}[x_1, x_2, x_3]$
- $what = \text{DIM}$

Steps:

1. $A \leftarrow \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & 1 & 2 \end{bmatrix}$, $c \leftarrow (2, 3, 4)$, *and* $\theta \leftarrow \pi/3$
2. $\epsilon(t) \leftarrow te^{\pi/3\sqrt{-1}}$
3. $j \leftarrow 2$
4. $\mathcal{Z}_2 \leftarrow \emptyset$
5. $\mathcal{C}_2$ *consists of a real point* $(13/11, 8/11, -8/11)$ *and four nonreal points*
6. $p \leftarrow (13/11, 8/11, -8/11)$
7. $p \in V_1$ *so go to the next* p
6. *There is no more* p *so the loop ends*
3. $j \leftarrow 1$
4. $\mathcal{Z}_1 \leftarrow \emptyset$
5. $\mathcal{C}_1$ *consists of the real point* $(13/11, 8/11, -8/11)$
6. $p \leftarrow (13/11, 8/11, -8/11)$
7. $p \notin V_1$
9. *return* $j = 1$

Output: $dim X = 1$

Using a single core of Intel Xeon E5-2680 2.5GHz CPU with `Bertini` *[11], this computation took under* 3 *seconds. Note that* $A \cdot p$ *is the point* $(-1, 1, 18/11) \in V(f_1, f_2) \cap \mathbb{R}^3$.

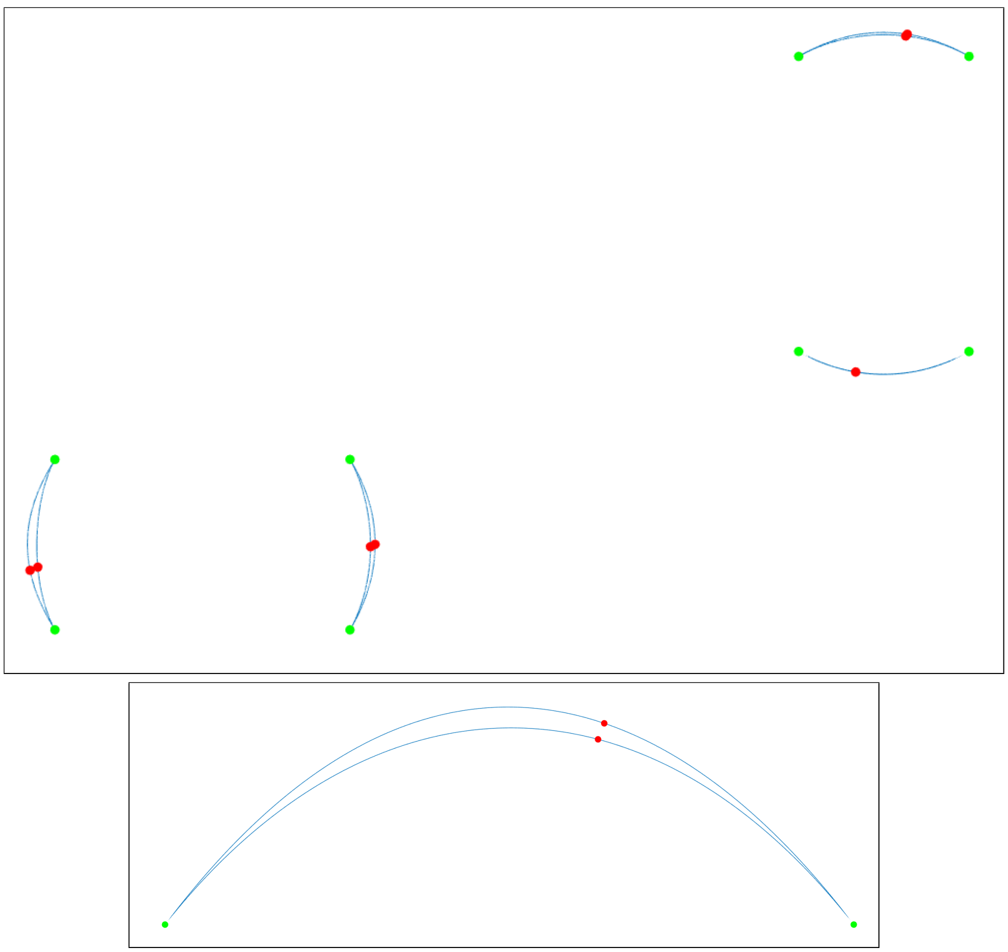
Example 6.3 *We next consider the following polynomial from [31]:*

$$f_1(x) = 100p_1(x)^2 + p_2(x)^3$$

where

$$p_1(x) = (x_1^2 + x_2^2 - 1)((x_1 - 4)^2 + (x_2 - 2)^2 - 1),$$
$$p_2(x) = (x_1 - 7/2)(x_1 - 9/2)(x_2 - 1/2)(x_2 + 1/2).$$

The figure below shows the set $X = V(f_1) \cap \mathbb{R}^2$ consisting of four connected components (blue curves), each having a thin crescent shape with two singular points (green points) along with a zoomed in version of the upper right connected component. This example demonstrates that a classical and often used approach of finding real sample points by computing critical points of the Euclidean distance function (first proposed by Seidenberg [40]) can not simultaneously compute a smooth point on each connected component [31, Prop. 3.2]. However, Algorithm 1 can be used to compute smooth sample points on each connected component of X. As with the previous illustrative example, we will provide a step by step trace of the algorithm.



Input:

- $f_1 = 100((x_1^2+x_2^2-1)((x_1-4)^2+(x_2-2)^2-1))^2+((x_1-7/2)(x_1-9/2)(x_2-1/2)(x_2+1/2))^3 \in \mathbb{R}[x_1, x_2]$
- $what$ = SSP

Steps:

1. $A = \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}$, $c = (1/2, 1/3)$, *and* $\theta = \pi/3$

2. $\epsilon(t) \leftarrow te^{\pi/3\sqrt{-1}}$

3. $j \leftarrow 1$

4. $\mathcal{Z}_1 \leftarrow \varnothing$

5. $\mathcal{C}_1$ *consists of* 16 *real points and* 156 *nonreal points. For the real points,* 8 *arise as the endpoint of a single path while the other* 8 *arise as the endpoint of* 4 *paths each.*

6./7./8. *The set* V_0 *consists of* 8 *real points and* 52 *nonreal points, the* 8 *real points in* $\mathcal{C}_1$ *arising from* 4 *paths each are contained in* V_0*. The* 8 *real points in* $\mathcal{C}_1$ *arising from a unique path are not contained in* V_0*. To* 4 *decimal places, this yields*

$$\mathcal{Z}_1 = \left\{ \begin{array}{cccc} (-1.0120, -0.1506), & (-0.9667, -0.1322), & (0.9863, -0.0122), & (1.0133, 0.0014) \\ (3.8346, 1.0135), & (3.8349, 1.0140), & (4.1287, 2.9859), & (4.1377, 2.9962) \end{array} \right\}.$$

3. $j \leftarrow 0$

4. $\mathcal{Z}_0 \leftarrow \varnothing$

5. $\mathcal{C}_0 = V_0$ *already computed above*

6./7./8. $V_{-1} = \varnothing$ *so the* 8 *real points in* $\mathcal{C}_0$ *form* $\mathcal{Z}_0$ *which, to* 4 *decimal places, is*

$$\mathcal{Z}_0 = \left\{ \begin{array}{cccc} (-0.8660, -0.5000), & (-0.8660, 0.5000), & (0.8660, -0.5000), & (0.8660, 0.5000) \\ (3.5000, 1.1340), & (3.5000, 2.8660), & (4.5000, 1.1340), & (4.5000, 2.8660) \end{array} \right\}.$$

Output: $\mathcal{Z}_0$ *(green points in figure),* $\mathcal{Z}_1$ *(red points in figure)*

Using a single core of Intel Xeon E5-2680 2.5GHz CPU with `Bertini` *[11], this computation took approximately 2 minutes.*

We conclude this section by comparing with [20].

Example 6.4 *For* $n \geqslant 3$ *and* $s \geqslant 2$ *even, consider*

$$\hat{f}(x) = \sum_{j=1}^{n} (x_1^s + x_2^s - 1 + p_j(x_3, \ldots, x_n))^2$$

where $p_j \in \mathbb{R}[x_3, \ldots, x_n]$ *is a general polynomial of degree* s *with* $p_j(0) = 0$*. Hence,*

$$V(\hat{f}) \cap \mathbb{R}^n = V(x_1^s + x_2^s - 1, x_3, \ldots, x_n) \cap \mathbb{R}^n$$

is a compact real curve. After performing a general change of coordinates $f(x) = \hat{f}(Ax)$*, there is an irreducible component of* V_1 *corresponding with* $V(y_1^s + y_2^s - 1, y_3, \ldots, y_n) \subset \mathbb{C}^n$ *after the appropriate change of coordinates. Since* $\partial_2 f$ *vanishes identically on this curve, the approach of [20] using isosingular deflation [23] to construct a polynomial which does not vanish identical on this curve as follows. Let* $F = \{f, \partial_3 f, \ldots, \partial_n f\}$*. Since the Jacobian matrix of* F*,* JF*, has rank* $n-2$ *on this curve, isosingular deflation will append* n *polynomials to* F*, namely the* $(n-1) \times (n-1)$ *minors of the* $(n-1) \times n$ *matrix* JF*, where each has degree* $(2s-2)(n-1)+1$*. Then,* g *is constructed by taking a random combination of the* $(n-1) \times (n-1)$ *minors of the Jacobian matrix of this deflated system, which has degree* $(2s-2)(n-1)^2$*. For comparison, the degree of* $\partial_2 f$ *is* $2s-1$*.*

7 Bounds

One way to analyze the computations performed is by bounding the total number of points that need to be computed. To that end, we consider bounds on the degrees of $V_{\epsilon,j}$, $C_{\epsilon,j}$, and $\mathcal{C}_{\epsilon,j}$ for generic values of $\epsilon \in \mathbb{C}$.

Proposition 7.1 *The degree of* $V_{\epsilon,j}$ *is at most*

$$d(d-1)^{n-1-j}. \tag{10}$$

Proof. Immediate from Bezout's bound. □

Proposition 7.2 *The degree of* $C_{\epsilon,j}$ *is at most*

$$d(d-1)^{n-1-j}\sum_{k=0}^{j}\binom{n-k-1}{j-k}(d-1)^k(d-2)^{j-k}. \tag{11}$$

Proof. One can view $C_{\epsilon,j}$ as the intersection of one polynomial of degree d, $n-1-j$ polynomials of degree at most $d-1$, and the rank deficiency set of an $(n-j+1)\times n$ matrix where each entry of the first row has degree at most $d-1$ and each entry of the remaining $n-j$ rows has degree at most $d-2$. Thus, the following bound arises from [22, Thm. 1]:

$$d(d-1)^{n-1-j}\sum_{\substack{k_1+\cdots+k_{n-j+1}=j\\ k_1,\ldots,k_{n-j+1}\geqslant 0}}(d-1)^{k_1}(d-2)^{j-k_1}.$$

This can be rewritten as

$$d(d-1)^{n-1-j}\sum_{k=0}^{j}A_k\ (d-1)^k(d-2)^{j-k}$$

where

$$A_k=\sum_{\substack{k_2+\cdots+k_{n-j+1}=j-k\\ k_2,\ldots,k_{n-j+1}\geqslant 0}}1=\sum_{\substack{\ell_1+\cdots+\ell_{n-j}=j-k\\ \ell_1,\ldots,\ell_{n-j}\geqslant 0}}1.$$

The claim immediately follows since

$$A_k=\binom{(n-j)+(j-k)-1}{(n-j)-1}=\binom{n-k-1}{j-k}.$$

□

The following example illustrates the quality of the above bounds.

Example 7.3 *Consider two nonnegative polynomials* P *and* Q *in* 5 *variables constructed as follows where* $V_{\mathbb{R}}(P)$ *and* $V_{\mathbb{R}}(Q)$ *are compact. Let* $x_6 = x_1$ *and define*

$$p(x)=\left(\sum_{i=1}^{5}x_i^2\right)^2-4\sum_{i=1}^{5}x_i^2x_{i+1}^2$$

arising from [19, Ex. 6.2] (see also [20, 30]). Let $q(x)$ be a sum of squares of 5 randomly selected quadratic forms. Since p and q are both homogeneous, we can compactify the real solution set by intersecting with the unit sphere by taking

$$P(x) = p(x) + \left(\sum_{i=1}^{5} x_i^2 - 1\right)^2 \quad \text{and} \quad Q(x) = q(x) + \left(\sum_{i=1}^{5} x_i^2 - 1\right)^2 \tag{12}$$

so that $\dim V_{\mathbb{R}}(P) = \dim V_{\mathbb{R}}(p) - 1$ *and* $\dim V_{\mathbb{R}}(Q) = \dim V_{\mathbb{R}}(q) - 1$. *In particular, P and Q both have $n = 5$ variables and degree $d = 4$.*

The following table compares the actual values for $\deg V_{\epsilon,j}$ *and* $\deg C_{\epsilon,j}$ *with the bounds provided in Propositions 7.1 and 7.2. Note that, for Q, the bounds are sharp.*

	j	(10)	$\deg V_{\epsilon,j}$	(11)	$\deg C_{\epsilon,j}$
	4	4	4	844	604
	3	12	12	1572	1332
P	2	36	36	1836	1596
	1	108	108	1188	868
	0	324	204	324	204
	4	4	4	844	844
	3	12	12	1572	1572
Q	2	36	36	1836	1836
	1	108	108	1188	1188
	0	324	324	324	324

The following bounds the total number of points that need to be computed in order to obtain $Z_0, \ldots, Z_{n-1}$ as in Theorem 4.3.

Proposition 7.4 *The total number of points that need to be computed in order to compute the collection of sets $Z_j = (C_j \backslash V_{j-1}) \cap \mathbb{R}^n$ for $j = 0, \ldots, n-1$ as in Theorem 4.3 is at most $(\sqrt{2}\, d)^{2n}$.*

Proof. We begin by simplifying the bounds given in Propositions 7.1 and 7.2.

$$\begin{aligned} \deg V_{\epsilon,j} &\leqslant d^{n-j}, \\ \deg C_{\epsilon,j} &\leqslant \sum_{k=0}^{j} \binom{n-k-1}{j-k} d(d-1)^{n-1} = \binom{n}{j} d(d-1)^{n-1} \leqslant \binom{n}{j} d^n. \end{aligned}$$

For $j = 0$, $Z_0 = C_0 \cap \mathbb{R}^n = V_0 \cap \mathbb{R}^n$ since $V_{-1} = \varnothing$. For $j = 1$, $Z_1 = (C_1 \backslash V_0) \cap \mathbb{R}^n$ where V_0 is zero-dimensional with degree at most $d(d-1)^{n-1}$ points. Thus, a positive-dimensional membership test is only needed for $j = 2, \ldots, n-1$. In particular, for each point $q \in C_j \cap \mathbb{R}^n$, one needs to determine if $q \in V_{j-1}$. Since the degree of C_j is at most $\binom{n}{j} d^n$ and the degree of V_{j-1} is at most d^{n-j+1}, the homotopy-based membership test [41] to determine $C_j \backslash V_{j-1}$ requires computing at most $\binom{n}{j} d^{2n-j+1}$ points. Hence the total number of points that need to be computed to yield $Z_0, \ldots, Z_{n-1}$ is at most

$$d^n + nd^n + \sum_{j=2}^{n} \binom{n}{j} d^{2n-j+1} \leqslant \sum_{j=0}^{n} \binom{n}{j} d^{2n} = 2^n d^{2n} = (\sqrt{2}\, d)^{2n}.$$

□

Finally the following extends this upper bound to the arbitrary case considered in Algorithm 1.

Theorem 7.5 *The total number of points computed in Algorithm 1 is at most* $(\sqrt{2}(d+1))^{2n}$.

Proof. The only difference with Proposition 7.4 is one now needs a bound on the number of points in $\mathcal{C}_{\epsilon,j}$. By using a 3-homogeneous Bezout bound, one obtains a bound of

$$\binom{n}{j} d(d-1)^{n-j-1}(d+1)^j \leqslant \binom{n}{j}(d+1)^n.$$

With this, the result follows with a similar proof of Proposition 7.4. □

8 Conclusion

By computing limits of critical points of a perturbation, a new approach based on Theorems 4.3 and 5.4 for obtaining smooth sample points on real algebraic sets was developed. In particular, perturbations eliminate the need for using isosingular deflation to construct a potentially high degree polynomial for computing critical points as in [20]. The use of a denominator in Section 5 ensures that both compact and unbounded components can be handled simultaneously as in Algorithm 1.

Complexity analysis was performed in terms of the number of points that need to be computed using homotopy continuation. In particular, without the denominator, the total is bounded above by $(\sqrt{2}\,d)^{2n}$ points (see Proposition 7.4) while using a denominator as in Algorithm 1 to handle the arbitrary case has at most $(\sqrt{2}\,(d+1))^{2n}$ points (see Theorem 7.5).

One application of computing smooth sample points is to compute dimension of the corresponding real algebraic set (see Algorithm 1). Since the initial motivation was to close the complexity gap between computing the real and complex cases of the dimension problem, future work involves analyzing the computational complexity of this perturbed approach using symbolic computation.

Acknowledgment

The idea of using perturbations to sample smooth points was discussed with our friend, collaborator, and colleague Agnes Szanto (1966-2022) before her untimely passing. The last two authors of this work were supported in part by the collaborative National Science Foundation grants CCF 2331400 and 2331401. The second author thanks Andrew Sommese for helpful comments and was also supported in part by the Robert and Sara Lumkins Collegiate Professorship and Simons Foundation SFM-00005696. The third author was also supported in part by National Science Foundation grant CCF 2212461.

References

[1] P. Aubry, F. Rouillier, and M. Safey El Din. Real solving for positive dimensional systems. *J. Symbolic Comput.*, 34(6):543–560, 2002.

[2] B. Bank, M. Giusti, J. Heintz, G. Lecerf, G. Matera, and P. Solerno. Degeneracy loci and polynomial equation solving. *Foundations of Computational Mathematics*, 15(1):159–184, 2015.

[3] B. Bank, M. Giusti, J. Heintz, R. Mandel, and G. M. Mbakop. Polar varieties and efficient real equation solving: the hypersurface case. In *Proceedings of the 3rd International Conference on Approximation and Optimization in the Caribbean (Puebla, 1995)*, page 13. Benemérita Univ. Autón. Puebla, Puebla, 1997.

[4] B. Bank, M. Giusti, J. Heintz, and G. Mbakop. Polar varieties, real equation solving, and data structures: The hypersurface case. *Journal of Complexity*, 13(1):5–27, 1997.

[5] B. Bank, M. Giusti, J. Heintz, and L. M. Pardo. Generalized polar varieties and an efficient real elimination procedure. *Kybernetika (Prague)*, 40(5):519–550, 2004.

[6] B. Bank, M. Giusti, J. Heintz, and L. M. Pardo. On the intrinsic complexity of point finding in real singular hypersurfaces. *Inform. Process. Lett.*, 109(19):1141–1144, 2009.

[7] B. Bank, M. Giusti, J. Heintz, M. Safey El Din, and E. Schost. On the geometry of polar varieties. *Applicable Algebra in Engineering, Communication and Computing*, 21(1):33–83, 2010.

[8] I. Bannwarth and M. Safey El Din. Probabilistic algorithm for computing the dimension of real algebraic sets. In *ISSAC'15—Proceedings of the 2015 ACM International Symposium on Symbolic and Algebraic Computation*, pages 37–44. ACM, New York, 2015.

[9] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in real algebraic geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[10] S. Basu, R. Pollack, and M.-F. Roy. Computing the dimension of a semi-algebraic set. *Journal of Mathematical Sciences*, 134(5):2346–2353, 2006.

[11] D. J. Bates, J. D. Hauenstein, A. J. Sommese, and C. W. Wampler. Bertini: Software for numerical algebraic geometry. Available at bertini.nd.edu with permanent doi: dx.doi.org/10.7274/R0H41PB5.

[12] D. J. Bates, J. D. Hauenstein, A. J. Sommese, and C. W. Wampler. Adaptive multiprecision path tracking. *SIAM Journal on Numerical Analysis*, 46(2):722–746, 2008.

[13] D. J. Bates, J. D. Hauenstein, A. J. Sommese, and C. W. Wampler. *Numerically Solving Polynomial Systems With Bertini*, volume 25 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2013.

[14] J. F. Canny. *The Complexity of Robot Motion Planning*. MIT Press, Cambridge, MA, USA, 1988.

[15] J. Cummings, J. D. Hauenstein, H. Hong, and C. D. Smyth. Smooth connectivity in real algebraic varieties. *Numerical Algorithms*, 2024.

[16] J.-C. Faugère, G. Moroz, F. Rouillier, and M. Safey El Din. Classification of the perspective-three-point problem, discriminant variety and real solving polynomial systems of inequalities. In *ISSAC'08 – Proceedings of the 2008 ACM International Symposium on Symbolic and Algebraic Computation*, pages 79–86, New York, NY, USA, 2008. ACM.

[17] D. R. Grayson and M. E. Stillman. Macaulay2, a software system for research in algebraic geometry. Available at `http://www.math.uiuc.edu/Macaulay2/`, 2022.

[18] D. Y. Grigor'ev and N. Vorobjov. Solving systems of polynomial inequalities in subexponential time. *Journal of Symbolic Computation*, 5(1):37–64, 1988.

[19] J. Han, L. Dai, and B. Xia. Constructing fewer open cells by gcd computation in cad projection. In *Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, ISSAC '14, page 240–247, New York, NY, USA, 2014. Association for Computing Machinery.

[20] K. Harris, J. D. Hauenstein, and A. Szanto. Smooth points on semi-algebraic sets. *Journal of Symbolic Computation*, 116:183–212, 2023.

[21] J. D. Hauenstein. Numerically computing real points on algebraic sets. *Acta Applicandae Mathematicae*, 125(1):105–119, 2013.

[22] J. D. Hauenstein, M. Safey El Din, E. Schost, and T. X. Vu. Solving determinantal systems using homotopy techniques. *Journal of Symbolic Computation*, 104:754–804, 2021.

[23] J. D. Hauenstein and C. W. Wampler. Isosingular sets and deflation. *Found. Comput. Math.*, 13(3):371–403, 2013.

[24] J. D. Hauenstein and C. W. Wampler. Numerically intersecting algebraic varieties via witness sets. *Appl. Math. Comput.*, 219(10):5730–5742, 2013.

[25] J. D. Hauenstein and C. W. Wampler. Unification and extension of intersection algorithms in numerical algebraic geometry. *Appl. Math. Comput.*, 293:226–243, 2017.

[26] J. Heintz, M.-F. Roy, and P. Solerno. Sur la complexité du principe de Tarski-Seidenberg. *Publications mathématiques et informatique de Rennes*, 4:103–120, 1989.

[27] H. Hong. Connectivity in semi-algebraic sets. In T. Ida, V. Negru, T. Jebelean, D. Petcu, S. M. Watt, and D. Zaharie, editors, *12th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2010, Timisoara, Romania, 23-26 September 2010*, pages 4–7. IEEE Computer Society, 2010.

[28] F. John. *Extremum problems with inequalities as subsidiary conditions*, pages 197–215. Springer Basel, Basel, 2014.

[29] P. Koiran. Randomized and deterministic algorithms for the dimension of algebraic varieties. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 36–45. IEEE, 1997.

[30] P. Lairez and M. Safey El Din. Computing the dimension of real algebraic sets. In *Proceedings of the 2021 on International Symposium on Symbolic and Algebraic Computation*, ISSAC '21, pages 257–264, New York, NY, USA, 2021. Association for Computing Machinery.

[31] H. C. Mork and R. Piene. Polars of real singular plane curves. In *Algorithms in Algebraic Geometry*, pages 99–115, New York, NY, 2008. Springer New York.

[32] D. Mumford. *Algebraic Geometry I Complex Projective Varieties*, volume No. 221 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin-New York, 1976.

[33] J. Renegar. A faster pspace algorithm for deciding the existential theory of the reals. In *[Proceedings 1988] 29th Annual Symposium on Foundations of Computer Science*, pages 291–295, 1988.

[34] F. Rouillier, M.-F. Roy, and M. Safey El Din. Finding at least one point in each connected component of a real algebraic set defined by a single equation. *Journal of Complexity*, 16(4):716–750, 2000.

[35] M. Safey El Din. Testing sign conditions on a multivariate polynomial and applications. *Math. Comput. Sci.*, 1(1):177–207, 2007.

[36] M. Safey El Din and E. Schost. Polar varieties and computation of one point in each connected component of a smooth algebraic set. In *Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation*, pages 224–231. ACM, New York, 2003.

[37] M. Safey El Din and P.-J. Spaenlehauer. Critical point computations on smooth varieties: degree and complexity bounds. In *Proceedings of the 2016 ACM International Symposium on Symbolic and Algebraic Computation*, pages 183–190. ACM, New York, 2016.

[38] M. Safey El Din, Z.-H. Yang, and L. Zhi. On the complexity of computing real radicals of polynomial systems. In *ISSAC'18—Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*, pages 351–358. ACM, New York, 2018.

[39] M. Safey El Din, Z.-H. Yang, and L. Zhi. Computing real radicals and s-radicals of polynomial systems. *Journal of Symbolic Computation*, 102:259–278, 2021.

[40] A. Seidenberg. A new decision method for elementary algebra. *Annals of Mathematics*, 60(2):365–374, 1954.

[41] A. J. Sommese, J. Verschelde, and C. W. Wampler. Numerical irreducible decomposition using projections from points on the components. In *Symbolic computation: solving equations in algebra, geometry, and engineering (South Hadley, MA, 2000)*, volume 286 of *Contemp. Math.*, pages 37–51. Amer. Math. Soc., Providence, RI, 2001.

[42] A. J. Sommese and C. W. Wampler. *The Numerical Solution Of Systems Of Polynomials Arising In Engineering And Science*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005.

[43] B. Teissier. A bouquet of bouquets for a birthday. In *Topological methods in modern mathematics (Stony Brook, NY, 1991)*, pages 93–122. Publish or Perish, Houston, TX, 1993.

[44] N. Vorobjov. Complexity of computing the local dimension of a semialgebraic set. *J. Symbolic Comput.*, 27(6):565–579, 1999.

[45] W. Wu and G. Reid. Finding points on real solution components and applications to differential polynomial systems. In *ISSAC 2013—Proceedings of the 38th International Symposium on Symbolic and Algebraic Computation*, pages 339–346. ACM, New York, 2013.