Computing the first Betti number and the connected components of semi-algebraic sets *

Saugata Basu[†]
Richard Pollack[‡]
Marie-Françoise Roy[§]

November 3, 2004

Abstract

In this paper we describe the first singly exponential algorithm for computing the first Betti number of a given semi-algebraic set. We also describe algorithms for obtaining semi-algebraic descriptions of the semi-algebraically connected components of any given real algebraic or semi-algebraic set. Singly exponential algorithms for computing the zero-th Betti number, and the Euler-Poincaré characteristic, were known before. No singly exponential algorithm was known for computing any of the individual Betti numbers other than the zero-th one.

1 Introduction

Let R be a real closed field and $S \subset \mathrm{R}^k$ a semi-algebraic set defined by a quantifier-free Boolean formula with atoms of the form $P > 0, P < 0, P = 0$ for $P \in \mathcal{P} \subset \mathrm{R}[X_1, \ldots, X_k]$. We call S a $\mathcal{P}$-semi-algebraic set. If, instead, the Boolean formula has atoms of the form $P = 0, P \geq 0, P \leq 0,\ P \in \mathcal{P}$, and additionally contains no negation, then we will call S a $\mathcal{P}$-closed semi-algebraic set. It is well known [24, 25, 23, 30, 1] that the topological complexity of S (measured by the various Betti numbers of S) is bounded by $O(sd)^k$, where $s = \#(\mathcal{P})$ and $d = \max_{P \in \mathcal{P}} \deg(P)$. More precise bounds on the individual Betti numbers of S appear in [2]. Even though the Betti numbers of S are bounded singly exponentially in k, there is no singly exponential algorithm for computing the Betti numbers of S. This absence is related to the fact that there is no known algorithm for producing a singly exponential sized triangulation of S (which would immediately imply a singly exponential algorithm for computing the Betti numbers of S). In fact, the existence of a singly exponential sized triangulation, is considered to be a major open question in algorithmic real algebraic geometry. Moreover, determining the exact complexity of computing the Betti numbers of semi-algebraic sets is an area of active research in computational complexity theory, for instance in investigating counting versions of complexity classes in the Blum-Shub-Smale model of computation (see [14]).

Doubly exponential algorithms (with complexity $(sd)^{2^{O(k)}}$) for computing all the Betti numbers are known, since it is possible to obtain a triangulation of S in doubly exponential time using cylindrical algebraic decomposition [16, 9]. In the absence of a singly exponential time algorithm for computing triangulations of semi-algebraic sets, algorithms with single exponential complexity are known only for the problems of testing emptiness [27, 4], computing the zero-th Betti number (i.e. the number of semi-algebraically connected components of S) [20, 15, 19, 7], as well as the Euler-Poincaré characteristic of S [1].

In this paper we describe the first singly exponential algorithm for computing the first Betti number of a given semi-algebraic set $S \subset \mathrm{R}^k$. In the process, we also give efficient algorithms for obtaining semi-algebraic descriptions of the semi-algebraically connected components of a given real algebraic or semi-algebraic set.

*2000 Mathematics Subject Classification 14P10, 14P25

[†]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, U.S.A., saugata@math.gatech.edu. Supported in part an NSF Career Award 0133597 and an Alfred P. Sloan Foundation Fellowship

[‡]Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, U.S.A., pollack@cims.nyu.edu. Supported in part by NSA grant MDA904-01-1-0057 and NSF grants CCR-9732101 and CCR-0098246.

[§]IRMAR (URA CNRS 305), Université de Rennes, Campus de Beaulieu 35042 Rennes cedex FRANCE, mfroy@maths.univ-rennes1.fr.

These algorithms have complexity bounds which improve the complexity of the best previously known algorithm [22].

The rest of the paper is organized as follows.

There are several ideas involved in the design of our algorithm for computing the first Betti number of a given semi-algebraic set, which corresponds to the main steps in our algorithm. We describe each of them separately in different sections.

In Section 3 we recall the notion of a roadmap of an algebraic set [7] and indicate how to use it to construct connecting paths in basic semi-algebraic sets.

In Section 4 we define certain semi-algebraic sets which we call parametrized paths and prove that these sets are semi-algebraically contractible. We also outline the input, output, and complexity of an algorithm computing a covering of a given basic semi-algebraic set, $S \subset \mathrm{R}^k$, by a singly exponential number of parametrized paths.

The tool from algebraic topology that we use to compute the first Betti number of a given semi-algebraic set from a covering of the set by acyclic sets requires that the sets occurring in the covering be closed. One difficulty with the covering referred to above is that the sets occurring in the covering are not necessarily closed. In Section 5 we show how to overcome this difficulty using a technique introduced by Gabrielov and Vorobjov [18]. Using their method we are able to convert the covering into one where the sets remain acyclic but have the additional property of being closed.

In Section 7 we show how to compute the first Betti number of a given closed and bounded semi-algebraic set using a covering by closed acyclic sets. The main tool here is a spectral sequence associated to the Mayer-Vietoris double complex. We show how to compute the first Betti number once we have computed a contractible closed covering and the number of connected components of their pair-wise and triple-wise intersections of the sets in this covering and their incidences. If the size of the acyclic covering is singly exponential, this yields a singly exponential algorithm for computing the first Betti number. Extensions of these ideas for computing a fixed number of higher Betti numbers in singly exponential time is possible and is reported on in a subsequent paper [3]. Finally, in Section 8 we indicate that the algorithms described in Section 4 actually produces descriptions of the connected components of a given algebraic or semi-algebraic set in an efficient manner.

2 Preliminaries

Let R be a real closed field. For an element $a \in \mathrm{R}$ we let

$$\mathrm{sign}(a) = \begin{cases} 0 & \text{if } a = 0, \\ 1 & \text{if } a > 0, \\ -1 & \text{if } a < 0. \end{cases}$$

If $\mathcal{P}$ is a finite subset of $\mathrm{R}[X_1, \ldots, X_k]$, we write the **set of zeros** of $\mathcal{P}$ in R^k as

$$\mathrm{Z}(\mathcal{P}, \mathrm{R}^k) = \{x \in \mathrm{R}^k \mid \bigwedge_{P \in \mathcal{P}} P(x) = 0\}.$$

We denote by $B(0, r)$ the open ball with center 0 and radius r.

Let $\mathcal{Q}$ and $\mathcal{P}$ be finite subsets of $\mathrm{R}[X_1, \ldots, X_k]$, $Z = \mathrm{Z}(\mathcal{Q}, \mathrm{R}^k)$, and $Z_r = Z \cap B(0, r)$. A *sign condition* on $\mathcal{P}$ is an element of $\{0, 1, -1\}^{\mathcal{P}}$. The *realization of the sign condition σ over Z*, $\mathcal{R}(\sigma, Z)$, is the basic semi-algebraic set

$$\{x \in \mathrm{R}^k \mid \bigwedge_{Q \in \mathcal{Q}} Q(x) = 0 \wedge \bigwedge_{P \in \mathcal{P}} \mathrm{sign}(P(x)) = \sigma(P)\}.$$

The *realization of the sign condition σ over Z_r*, $\mathcal{R}(\sigma, Z_r)$, is the basic semi-algebraic set $\mathcal{R}(\sigma, Z) \cap B(0, r)$. For the rest of the paper, we fix an open ball $B(0, r)$ with center 0 and radius r big enough so that, for every sign condition σ, $\mathcal{R}(\sigma, Z)$ and $\mathcal{R}(\sigma, Z_r)$ are homeomorphic. This is always possible by the local conical structure at infinity of semi-algebraic sets ([12], page 225).

A closed and bounded semi-algebraic set $S \subset \mathrm{R}^k$ is semi-algebraically triangulable (see [9]), and we denote by $H_i(S)$ the i-th simplicial homology group of S with rational coefficients. The groups $H_i(S)$ are invariant under semi-algebraic homeomorphisms and coincide with the corresponding singular homology groups when $\mathrm{R} = \mathbb{R}$. We denote by $b_i(S)$ the i-th Betti number of S (that is, the dimension of $H_i(S)$ as a vector space), and $b(S)$ the sum $\sum_i b_i(S)$. For a closed but not necessarily bounded semi-algebraic set $S \subset \mathrm{R}^k$, we will denote by $H_i(S)$ the i-th simplicial homology group of $S \cap \overline{B(0,r)}$, where r is sufficiently large. The sets $S \cap \overline{B(0,r)}$ are semi-algebraically homeomorphic for all sufficiently large $r > 0$, by the local conical structure at infinity of semi-algebraic sets, and hence this definition makes sense.

The definition of homology groups of arbitrary semi-algebraic sets in R^k requires some care and several possibilities exist. In this paper, we define the homology groups of realizations of sign conditions as follows.

Let R denote a real closed field and R' a real closed field containing R. Given a semi-algebraic set S in R^k, the *extension* of S to R', denoted $\mathrm{Ext}(S, \mathrm{R}')$, is the semi-algebraic subset of R'^k defined by the same quantifier free formula that defines S. The set $\mathrm{Ext}(S, \mathrm{R}')$ is well defined (i.e. it only depends on the set S and not on the quantifier free formula chosen to describe it). This is an easy consequence of the transfer principle [9].

Now, let $S \subset \mathrm{R}^k$ be a $\mathcal{P}$-semialgebraic set, where $\mathcal{P} = \{P_1, \ldots, P_s\}$ is a finite subset of $\mathrm{R}[X_1, \ldots, X_k]$. Let $\phi(X)$ be a quantifier-free formula defining S. Let $P_i = \sum_\alpha a_{i,\alpha} X^\alpha$ where the $a_{i,\alpha} \in \mathrm{R}$. Let $A = (\ldots, A_{i,\alpha}, \ldots)$ denote the vector of variables corresponding to the coefficients of the polynomials in the family $\mathcal{P}$, and let $a = (\ldots, a_{i,\alpha}, \ldots) \in \mathrm{R}^N$ denote the vector of the actual coefficients of the polynomials in $\mathcal{P}$. Let $\psi(A, X)$ denote the formula obtained from $\phi(X)$ by replacing each coefficient of each polynomial in $\mathcal{P}$ by the corresponding variable, so that $\phi(X) = \psi(a, X)$. It follows from Hardt's triviality theorem for semi-algebraic mappings [21], that there exists, $a' \in \mathbb{R}^N_{\mathrm{alg}}$ such that denoting by $S' \subset \mathbb{R}^k_{\mathrm{alg}}$ the semi-algebraic set defined by $\psi(a', X)$, the semi-algebraic set $\mathrm{Ext}(S', \mathrm{R})$, has the same homeomorphism type as S. We define the homology groups of S to be the singular homology groups of $\mathrm{Ext}(S', \mathbb{R})$. It follows from the Tarski-Seidenberg transfer principle, and the corresponding property of singular homology groups, that the homology groups defined this way are invariant under semi-algebraic homotopies. It is also clear that this definition is compatible with the simplicial homology for closed, bounded semi-algebraic sets, and the singular homology groups when the ground field is $\mathbb{R}$. Finally it is also clear that, the Betti numbers are not changed after extension: $b_i(S) = b_i(\mathrm{Ext}(S, \mathrm{R}'))$.

3 Roadmap of a semi-algebraic set

We first define a roadmap of a semi-algebraic set. Roadmaps are crucial ingredients in all singly exponential algorithms known for computing connectivity properties of semi-algebraic sets such as computing the number of connected components, as well as testing whether two points of a given semi-algebraic set belong to the same semi-algebraically connected component.

We use the following notations. Given $x = (x_1, \ldots, x_k)$ we write $\bar{x}_i$ for $(x_1, \ldots, x_i)$, and $\tilde{x}_i$ for $(x_{i+1}, \ldots, x_k)$. We also denote by $\pi_{1\ldots j}$ the projection, $x \mapsto \bar{x}_j$. Given a set $S \subset \mathrm{R}^k$, $y \in \mathrm{R}^j$ we denote by $S_y = S \cap \pi^{-1}_{1\ldots j}(y)$.

Let $S \subset \mathrm{R}^k$ be a semi-algebraic set. A **roadmap** for S is a semi-algebraic set M of dimension at most one contained in S which satisfies the following roadmap conditions:

RM_1 For every semi-algebraically connected component D of S, $D \cap M$ is semi-algebraically connected.

RM_2 For every $x \in \mathrm{R}$ and for every semi-algebraically connected component D' of S_x, $D' \cap M \neq \emptyset$.

We describe the construction of a roadmap M for a bounded algebraic set $\mathrm{Z}(Q, \mathrm{R}^k)$ which contains a finite set of points $\mathcal{N}$ of $\mathrm{Z}(Q, \mathrm{R}^k)$. A precise description of how the construction can be performed algorithmically can be found in [9].

A key ingredient of the roadmap is the construction of a particular finite set of points having the property that, they intersect every connected component of $\mathrm{Z}(Q, \mathrm{R}^k)$. We call them X_1-pseudo-critical points, since they are obtained as limits of the critical points of the projection to the X_1 coordinate of a bounded nonsingular algebraic hypersurface defined by a particular infinitesimal deformation of the polynomial Q. Their projections on the X_1-axis are called pseudo-critical values. These points are obtained as follows.

We denote by $\mathrm{R}\langle\zeta\rangle$ the real closed field of algebraic Puiseux series in ζ with coefficients in R [9]. The sign of a Puiseux series in $\mathrm{R}\langle\zeta\rangle$ agrees with the sign of the coefficient of the lowest degree term in ζ. This induces a unique order on $\mathrm{R}\langle\zeta\rangle$ which makes ζ infinitesimal: ζ is positive and smaller than any positive element of R. When $a \in \mathrm{R}\langle\zeta\rangle$ is bounded by an element of R, $\lim_\zeta(a)$ is the constant term of a, obtained by substituting 0 for ζ in a. We now define the deformation $\bar{Q}$ of Q as follows. Suppose that $\mathrm{Z}(Q, \mathrm{R}^k)$ is contained in the ball of center 0 and radius $1/c$. Let $\bar{d}$ be an even integer bigger than the degree d of Q,

$$G_k(\bar{d}, c) = c^{\bar{d}}(X_1^{\bar{d}} + \cdots + X_k^{\bar{d}} + X_2^2 + \cdots + X_k^2) - (2k-1), \tag{1}$$

$$\bar{Q} = \zeta G_k(\bar{d}, c) + (1-\zeta)Q. \tag{2}$$

The algebraic set $\mathrm{Z}(\bar{Q}, \mathrm{R}\langle\zeta\rangle^k)$ is a bounded and non-singular hypersurface lying infinitesimally close to $\mathrm{Z}(Q, \mathrm{R}^k)$, and the critical points of the projection map onto the X_1 co-ordinate restricted to $\mathrm{Z}(\bar{Q}, \mathrm{R}\langle\zeta\rangle^k)$ form a finite set of points. We take the images of these points under $\lim_\zeta$ and we call the points obtained in this manner the X_1-pseudo-critical points of $\mathrm{Z}(Q, \mathrm{R}^k)$. Their projections on the X_1-axis are called pseudo-critical values.

The construction of the roadmap of an algebraic set containing a finite number of input points $\mathcal{N}$ of this algebraic set is as follows. We first construct X_2-pseudo-critical points on $\mathrm{Z}(Q, \mathrm{R}^k)$ in a parametric way along the X_1-axis, by following continuously, as x varies on the X_1-axis, the X_2-pseudo-critical points on $\mathrm{Z}(Q, \mathrm{R}^k)_x$. This results in curve segments and their endpoints on $\mathrm{Z}(Q, \mathrm{R}^k)$. The curve segments are continuous semi-algebraic curves parametrized by open intervals on the X_1-axis, and their endpoints are points of $\mathrm{Z}(Q, \mathrm{R}^k)$ above the corresponding endpoints of the open intervals. Since these curves and their endpoints include, for every $x \in \mathrm{R}$, the X_2-pseudo-critical points of $\mathrm{Z}(Q, \mathrm{R}^k)_x$, they meet every connected component of $\mathrm{Z}(Q, \mathrm{R}^k)_x$. Thus the set of curve segments and their endpoints already satisfy RM_2. However, it is clear that this set might not be semi-algebraically connected in a semi-algebraically connected component, so RM_1 might not be satisfied. We add additional curve segments to ensure connectedness by recursing in certain distinguished hyperplanes defined by $X_1 = z$ for distinguished values z.

The set of **distinguished values** is the union of the X_1-pseudo-critical values, the first coordinates of the input points $\mathcal{N}$ and the first coordinates of the endpoints of the curve segments. A **distinguished hyperplane** is an hyperplane defined by $X_1 = v$, where v is a distinguished value. The input points, the endpoints of the curve segments and the intersections of the curve segments with the distinguished hyperplanes define the set of **distinguished points** .

Let the distinguished values be $_1 < \ldots < v_\ell$. Note that amongst these are the X_1-pseudo-critical values. Above each interval (v_i, v_{i+1}), we have constructed a collection of curve segments $\mathcal{C}_i$ meeting every semi-algebraically connected component of $\mathrm{Z}(Q, \mathrm{R}^k)_v$ for every $v \in (v_i, v_{i+1})$. Above each distinguished value v_i, we have a set of distinguished points $\mathcal{N}_i$. Each curve segment in $\mathcal{C}_i$ has an endpoint in $\mathcal{N}_i$ and another in $\mathcal{N}_{i+1}$. Moreover, the union of the $\mathcal{N}_i$ contains $\mathcal{N}$.

We then repeat this construction in each distinguished hyperplane H_i defined by $X_1 = v_i$ with input $Q(v_i, X_2, \ldots, X_k)$ and the distinguished points in $\mathcal{N}_i$. Thus, we construct distinguished values, $v_{i,1}, \ldots, v_{i,\ell(i)}$ of $\mathrm{Z}(Q(v_i, X_2, \ldots, X_k), \mathrm{R}^{k-1})$ (with the role of X_1 being now played by X_2) and the process is iterated until for $I = (i_1, \ldots, i_{k-2}), 1 \le i_1 \le \ell, \ldots, 1 \le i_{k-2} \le \ell(i_1, \ldots, i_{k-3})$, we have distinguished values $v_{I,1} < \ldots < v_{I,\ell(I)}$ along the X_{k-1} axis with corresponding sets of curve segments and sets of distinguished points with the required incidences between them.

The following proposition is proved in [7] (see also [9]).

Proposition 3.1 *The semi-algebraic set* $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k), \mathcal{N})$ *obtained by this construction is a roadmap for* $\mathrm{Z}(Q, \mathrm{R}^k)$ *containing* $\mathcal{N}$*.*

Note that if $x \in \mathrm{Z}(Q, \mathrm{R}^k)$, $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k), \{x\})$ contains a path, $\gamma(x)$, connecting a distinguished point p of $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k))$ to x.

Later in this paper we shall examine the properties of parametrized paths which are the unions of connecting paths starting at a given p and ending at x, where x varies over a certain semi-algebraic subset of $\mathrm{Z}(Q, \mathrm{R}^k)$. In order to do so it is useful to have a better understanding of the structure of these connecting paths – especially, of their dependence on x.

We first note that for any $x = (x_1, \ldots, x_k) \in \mathrm{Z}(Q, \mathrm{R}^k)$, we have by construction that, $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k))$ is contained in $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k), \{x\})$. In fact, $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k), \{x\}) = \mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k)) \cup \mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k)_{x_1}, \mathcal{M}_{x_1})$, where $\mathcal{M}_{x_1}$ consists of $\tilde{x}_1$ and the finite set of points obtained by intersecting the curves in $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k))$ parametrized by the X_1-coordinate, with the hyperplane $\pi_1^{-1}(x_1)$.

A connecting path $\gamma(x)$ joining a distinguished point p of $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k))$ to x can be extracted from $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k), \{x\})$. The connecting path $\gamma(x)$ consists of two consecutive parts, $\gamma_0(x)$ and $\Gamma_1(x)$. The path $\gamma_0(x)$ is contained in $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k))$ and the path $\Gamma_1(x)$ is contained in $\mathrm{Z}(Q, \mathrm{R}^k)_{x_1}$. Moreover, $\Gamma_1(x)$ can again be decomposed into two parts, $\gamma_1(x)$ and $\Gamma_2(x)$ with $\Gamma_2(x)$ contained in $\mathrm{Z}(Q, \mathrm{R}^k)_{\bar{x}_2}$ and so on.

If $y = (y_1, \ldots, y_k) \in \mathrm{Z}(Q, \mathrm{R}^k)$ is another point such that $x_1 \neq y_1$, then since $\mathrm{Z}(Q, \mathrm{R}^k)_{x_1}$ and $\mathrm{Z}(Q, \mathrm{R}^k)_{y_1}$ are disjoint, it is clear that $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k), \{x\}) \cap \mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k), \{y\}) = \mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k))$. Now consider a connecting path $\gamma(y)$ extracted from $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k), \{y\})$. The images of $\Gamma_1(x)$ and $\Gamma_1(y)$ are disjoint. If the image of $\gamma_0(y)$ (which is contained in $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k))$ follows the same sequence of curve segments as $\gamma_0(x)$ starting at p, then it is clear that the images of the paths $\gamma(x)$ and $\gamma(y)$ has the property that they are identical upto a point and they are disjoint after it. We call this the *divergence property.*

More generally, if the points x and y are such that, $x_i = y_i, 1 \leq i \leq j$ and $x_{j+1} \neq y_{j+1}$, then the paths $\Gamma_{j+1}(x)$ and $\Gamma_{j+1}(y)$, contained in $\mathrm{Z}(Q, \mathrm{R}^k)_{\bar{x}_{j+1}}$ and $\mathrm{Z}(Q, \mathrm{R}^k)_{\bar{y}_{j+1}}$ respectively, will be disjoint. Moreover paths $\gamma_0(x), \ldots, \gamma_j(x)$ and $\gamma_0(y), \ldots, \gamma_j(y)$ are composed of the same sequence of curve segments, then $\gamma(x)$ and $\gamma(y)$ will also have the divergence property.

We now consider connecting paths in the semi-algebraic setting. We are given a polynomial $Q \in \mathrm{R}[X_1, \ldots, X_k]$ such that $\mathrm{Z}(Q, \mathrm{R}^k)$ is bounded and a finite set of polynomials $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$ in strong ℓ-general position with respect to Q. This means that any $\ell + 1$ polynomials belonging to $\mathcal{P}$ have no zeros in common with Q in R^k, and any ℓ polynomials belonging to $\mathcal{P}$ have at most a finite number of zeros in common with Q in R^k.

For every point x of $\mathrm{Z}(Q, \mathrm{R}^k)$, we denote by $\sigma(x)$ the sign condition on $\mathcal{P}$ at x. Let $\mathcal{R}(\overline{\sigma}(x), \mathrm{Z}(Q, \mathrm{R}^k)) = \{x \in \mathrm{Z}(Q, \mathrm{R}^k) \mid \bigwedge_{P \in \mathcal{P}} \ \mathrm{sign}(P(x)) \in \overline{\sigma}(x)(P)\}$, where $\overline{\sigma}$ is the relaxation of σ defined by

$$\begin{cases} \overline{\sigma} = \{0\} & \text{if} \quad \sigma = 0, \\ \overline{\sigma} = \{0, 1\} & \text{if} \quad \sigma = 1, \\ \overline{\sigma} = \{0, -1\} & \text{if} \quad \sigma = -1. \end{cases}$$

We say that $\overline{\sigma}(x)$ is the weak sign condition defined by x on $\mathcal{P}$. We denote by $\mathcal{P}(x)$ the union of $\{Q\}$ and the set of polynomials in $\mathcal{P}$ vanishing at x.

The connecting algorithm associates to $x \in \mathrm{Z}(Q, \mathrm{R}^k)$ a path entirely contained in the realization of $\overline{\sigma}(x)$ connecting x to a distinguished point of the roadmap of some $\mathrm{Z}(\mathcal{P}', \mathrm{R}^k)$, with $\mathcal{P}(x) \subset \mathcal{P}'$. The connecting algorithm proceeds as follows: construct a path γ connecting a distinguished point of $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k))$ to x contained in $\mathrm{RM}(\mathrm{Z}(Q, \mathrm{R}^k), x)$. If no polynomial of $\mathcal{P} \setminus \mathcal{P}(x)$ vanishes on γ, we are done. Otherwise let y be the last point of γ such that some polynomial of $\mathcal{P} \setminus \mathcal{P}(x)$ vanishes at y. Now keep the part of γ connecting y to x as end of the connecting path, and iterate the construction with y, noting that the realization of $\overline{\sigma}(y)$ is contained in the realization of $\overline{\sigma}(x)$, and $\mathcal{P} \setminus \mathcal{P}(y)$ is in $\ell - 1$ strong general position with respect to $\mathrm{Z}(\mathcal{P}(y), \mathrm{R}^k)$.

As in the algebraic case, two such connecting paths which start with the same sequence of curve segments will have the divergence property. This follows from the divergence property in the algebraic case and the recursive definition of connecting paths.

Formal descriptions and complexity analysis of the algorithms described above for computing roadmaps and connecting paths of algebraic and basic semi-algebraic sets can be found in [9] (Algorithm 15.12 and Algorithm 16.8).

4 Constructing Contractible Coverings

In this section, we show how to obtain a covering of a given semi-algebraic set by a family of semi-algebraically contractible subsets. The construction is based on a parametrized version of the connecting algorithm: we find quantifier free formulas on the parameters for which the description of the connecting path does not change. For each such formula we take as an element of the cover, the semi-algebraic set consisting of the union of images of the corresponding connecting paths.

We first define parametrized paths. A parametrized path is a semi-algebraic set which is a union of semi-algebraic paths having the divergence property.

More precisely,

Definition 4.1 *A parametrized path γ is a continuous semi-algebraic mapping from $V \subset \mathrm{R}^{k+1} \to \mathrm{R}^k$, such that, denoting by $U = \pi_{1\ldots k}(V) \subset \mathrm{R}^k$, there exists a semi-algebraic continuous function ℓ from $\ell : U \to [0, +\infty)$ and there exists a point a in R^k, such that*

1. $V = \{(x, t) \mid x \in U, 0 \leq t \leq \ell(x)\}$,
2. $\forall\, x \in U,\ \gamma(x, 0) = a$,
3. $\forall\, x \in U,\ \gamma(x, \ell(x)) = x$,
4. $\forall\, x \in U, \forall\, y \in U, \forall\, s \in [0, \ell(x)], \forall\, t \in [0, \ell(y)]\ (\gamma(x, s) = \gamma(y, t) \Rightarrow s = t)$,
5. $\forall\, x \in U, \forall\, y \in U, \forall\, s \in [0, \min(\ell(x), \ell(y))]\ (\gamma(x, s) = \gamma(y, s) \Rightarrow \forall\, t \leq s\ \gamma(x, t) = \gamma(y, t))$.

Proposition 4.2 *The image of a parametrized path γ is semi-algebraically contractible.*

Proof : Let $W = \mathrm{Im}(\gamma)$. The semi-algebraic continuous mapping $\phi : W \times [0,1] \to W$ sending

- $(\gamma(x,t), s)$ to $\gamma(x, s \cdot \ell(x))$ if $t \geq s \cdot \ell(x)$,
- $(\gamma(x,t), s)$ to $\gamma(x,t)$ if $t < s \cdot \ell(x)$.

satisfies

$$\phi(\gamma(x,t),0) = a,$$
$$\phi(\gamma(x,t),1) = \gamma(x,t)$$

and is thus a semi-algebraic contraction from W to $\{a\}$. □

We now describe how to compute parametrized paths in single exponential time using a parametrized version of the connecting algorithm.

We describe the input, output and complexity of the algorithm which appears in [9] (Algorithm 16.15).

Algorithm 4.3 (Parametrized Bounded Connecting)

Input : *a polynomial* $Q \in \mathrm{D}[X_1, \ldots, X_k]$, *such that* $\mathrm{Z}(Q, \mathrm{R}^k) \subset B(0, 1/c)$,
a finite set of polynomials $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$ *in strong* ℓ*-general position with respect to* Q.

Output : *a finite set* Θ *of quantifier free formulas such that for every semi-algebraically connected component* S *of the realization of every weak sign condition on* $\mathcal{P}$ *on* $\mathrm{Z}(Q, \mathrm{R}^k)$, *there exists a subset* $\Theta(S) \subset \Theta$ *such that* $S = \bigcup_{\theta \in \Theta(S)} \mathcal{R}(\theta, \mathrm{Z}(Q, \mathrm{R}^k))$,
for every $\theta \in \Theta$, *a semi-algebraic set* $\gamma(\theta) \subset \mathrm{R}^{2k}$ *such that for every* $y \in \mathcal{R}(\theta, \mathrm{R}^k)$, $\gamma(\theta)(y)$ *is a semi-algebraic path which connects the point* y *to a distinguished point* $a(\theta)$ *of some roadmap* $\mathrm{RM}(\mathrm{Z}(\mathcal{P}' \cup \{Q\}, \mathrm{R}^k))$ *where* $\mathcal{P}' \subset \mathcal{P}$, *staying inside* $\mathcal{R}(\overline{\sigma}(y), \mathrm{Z}(Q, \mathrm{R}^k))$.

Complexity : $s^{k+1} d^{O(k^4)}$, *where* s *is a bound on the number of elements of* $\mathcal{P}$ *and* d *is a bound on the degrees of* Q *and the elements of* $\mathcal{P}$.

The only extra property of the output of Algorithm 16.15 in [9] that we need is that for each $\theta \in \Theta$ output by the algorithm, $\gamma(\theta)$ is a parametrized path.

Using the notation introduced above, we have:

Proposition 4.4 *For each* $\theta \in \Theta$ *output by Algorithm 4.3,* $\gamma(\theta)$ *is a parametrized path.*

Proof: From the divergence property of the paths $\gamma(y, \cdot)$ (see discussion in Section 3), it is easy to see that $\gamma(\theta)$ satisfies the conditions of Definition 4.1. □

5 Replacement by closed sets without changing homology

In this section, we describe a construction due to Gabrielov and Vorobjov [18] for replacing any given semi-algebraic subset of a bounded semi-algebraic set by a closed bounded semi-algebraic subset having isomorphic homology groups. Moreover, the polynomials defining the bounded closed semi-algebraic subset are closely related (by infinitesimal perturbations) to the polynomials defining the original subset. In particular, their degrees do not increase, while the number of polynomials used in the definition of the new set is at most the square of the number used in the definition of the original set. We observe a couple of extra properties of these sets, which are not noted in [18], that are useful later in the paper.

Given a finite family of polynomials $\mathcal{C} \subset \mathrm{R}[X_1, \ldots, X_k]$ with t elements, let S be a bounded $\mathcal{C}$-closed set. We denote by $\mathrm{Sign}(S, \mathcal{C})$ the set of realizable sign conditions of $\mathcal{C}$ whose realization is contained in S.

Let $S_i, 1 \leq i \leq N$ be $\mathcal{C}$-semi-algebraic subsets of S such that $\cup_{1 \leq i \leq N} S_i = S$.

Notice that the S_i's are not necessarily closed.

We will now adapt a novel construction due to Gabrielov and Vorobjov [18], using which we can replace the sets S_i by closed and bounded semi-algebraic sets S_i' with isomorphic homology groups.

More precisely, let $\epsilon_t \gg \epsilon_{t-1} \gg \cdots \gg \epsilon_1 > 0$ be infinitesimals. We denote by R' the field of Puiseux series $\mathrm{R}\langle\epsilon_t\rangle \cdots \langle\epsilon_1\rangle$. We are going to replace the sets S_i by sets S_i' such that $S_i' \subset \mathrm{R}'^k$ such that,

1. The homology groups $H_*(S'_i)$ and $H_*(S_i)$ are isomorphic for $1 \leq i \leq N$,

2. $\cup_{1 \leq i \leq N} S'_i = \mathrm{Ext}(S, \mathrm{R}')$.

We now describe the construction due to Gabrielov and Vorobjov.

For $\sigma \in \mathrm{Sign}(S, \mathcal{C})$ we define the level of σ as $\#\{P \in \mathcal{C} | \sigma(P) = 0\}$.

For each level $m, 0 \leq m \leq t$, we denote by

$$\sigma_{m,1}, \ldots \sigma_{m,n(m)}$$

the elements of $\mathrm{Sign}(S, \mathcal{C})$ of level m. We denote by $\mathcal{R}^c_+(\sigma_{m,j})$ the intersection of $\mathrm{Ext}(S, \mathrm{R}')$ with the semi-algebraic set defined by the conjunction of the inequalities,

$$\begin{cases} -\epsilon_m \leq P \leq \epsilon_m, & \text{for each } P \in \mathcal{C} \text{ such that } \sigma_{m,j}(P) = 0, \\ P \geq 0, & \text{for each } P \in \mathcal{C} \text{ such that } \sigma_{m,j}(P) = 1, \\ P \leq 0, & \text{for each } P \in \mathcal{C} \text{ such that } \sigma_{m,j}(P) = -1. \end{cases}$$

We denote by, $\mathcal{R}^o_+(\sigma_{m,j})$ the intersection of $\mathrm{Ext}(S, \mathrm{R}')$ with the semi-algebraic set defined by the conjunction of the inequalities,

$$\begin{cases} -\epsilon_m < P < \epsilon_m, & \text{for each } P \in \mathcal{C} \text{ such that } \sigma_{m,j}(P) = 0, \\ P > 0, & \text{for each } P \in \mathcal{C} \text{ such that } \sigma_{m,j}(P) = 1, \\ P < 0, & \text{for each } P \in \mathcal{C} \text{ such that } \sigma_{m,j}(P) = -1. \end{cases}$$

Notice that $\mathcal{R}^o_+(\sigma_{m,j})$ (resp. $\mathcal{R}^c_+(\sigma_{m,j})$) is an open (resp. closed) semi-algebraic subset of $\mathrm{Ext}(S, \mathrm{R}')$ and they both contain $E(\mathcal{R}(\sigma_{m,j}), \mathrm{R}')$.

Let $X \subset S$ be a $\mathcal{C}$-semi-algebraic set such that $X = \cup_{\sigma \in \Sigma} \mathcal{R}(\sigma)$ with $\Sigma \subset \mathrm{Sign}(S, \mathcal{C})$.

We define a sequence of sets, $X^{m,j}$, $0 \leq m \leq t, 0 \leq j \leq n(m)$ inductively.

- Define $X^{0,0} = \mathrm{Ext}(X, R')$.
- For $0 \leq m \leq t$,
  - For $1 \leq j \leq n(m)$ construct $X^{m,j}$ from $X^{m,j-1}$ as follows.
  - If $\sigma_{m,j} \in \Sigma$ then $X^{m,j} = X^{m,j-1} \cup \mathcal{R}^c_+(\sigma_{m,j})$.
  - If $\sigma_{m,j} \notin \Sigma$ then $X^{m,j} = X^{m,j-1} \setminus \mathcal{R}^o_+(\sigma_{m,j})$.
- Define $X^{m+1,0} = X^{m,n(m)}$.

We denote by X' the set $X^{t+1,0}$.

The following theorem is a slight generalization of a result in [18] and is proved by similar techniques (see [10]).

Theorem 5.1 *(Gabrielov-Vorobjov) If X is acyclic, then X' is also acyclic.*

We now a prove a useful property of the sets S'_i, where S'_i are the closed semi-algebraic sets obtained from the sets S_i using the construction described above.

Proposition 5.2 *(Covering property)*

$$\mathrm{Ext}(S, R') = \bigcup_{1 \leq i \leq N} \mathrm{Ext}(S_i, R') = \bigcup_{1 \leq i \leq N} S'_i.$$

It is useful for the proof to introduce a partial order on the set of realizable sign conditions $\mathrm{Sign}(S, \mathcal{C})$ by, $\tau \prec \sigma$ if $\tau(P) = 0 \Rightarrow \sigma(P) = 0$, and $\sigma(P) \neq 0 \Rightarrow \tau(P) = \sigma(P)$. Note that if $\sigma_{m,j} \prec \sigma_{m',j'}$ then $(m,j) = (m', j')$, or $m < m'$.

Proof: Since, each $\mathrm{Ext}(S_i, \mathrm{R}')$ and S'_i are contained in $\mathrm{Ext}(S, R')$ it is clear that $\cup_{1 \leq i \leq N} S'_i \subset \mathrm{Ext}(S, R')$. We now prove the reverse inclusion by contradiction.

Let $x \in \mathrm{Ext}(S, R')$ but $x \notin \cup_{1 \leq i \leq N} S'_i$. Since by definition $S_i^{0,0} = \mathrm{Ext}(S_i, \mathrm{R}')$ for each i, there must exist m, j, $0 \leq m \leq t$, $0 \leq j \leq n(m)$, such that,

1. $x \notin \cup_{1\le i\le N} S_i^{m',j'}$ for all m', j' such that $m' = m, j' \ge j$, or $m' > m$, and

2. $x \in \cup_{1\le i\le N} S_i^{m,j-1}$.

Since $x \in \cup_{1\le i\le N} S_i^{m,j-1}$, there exists ℓ, $1 \le \ell \le N$ such that $x \in S_\ell^{m,j-1}$, and since $x \notin \cup_{1\le i\le N} S_i^{m,j}$, $x \notin S_\ell^{m,j}$. Consider the sign condition $\sigma_{m,j}$. Clearly $x \in \mathcal{R}_+^o(\sigma_{m,j})$ and hence must satisfy the following:

$$\begin{aligned}
-\epsilon_m < P(x) < \epsilon_m, &\ \text{ for } P \in \mathcal{C} \text{ such that } \sigma_{m,j}(P) = 0,\\
P(x) > 0, &\ \text{ for } P \in \mathcal{C} \text{ such that } \sigma_{m,j}(P) = +1,\\
P(x) < 0, &\ \text{ for } P \in \mathcal{C} \text{ such that } \sigma_{m,j}(P) = -1.
\end{aligned}$$

Let $y = \lim_{\varepsilon_m} x$. Since S is closed and bounded, and $x \in \mathrm{Ext}(S, R')$, y must also be in $\mathrm{Ext}(S, R')$. Let the sign condition of $\mathcal{C}$ at the point y be $\sigma_{m',j'}$, so that $y \in \mathcal{R}_+^c(\sigma_{m',j'})$. Clearly $\sigma_{m,j} \prec \sigma_{m',j'}$.

We prove now that $x \in \mathcal{R}_+^c(\sigma_{m',j'})$. If $(m, j) = (m', j')$, it is clear that $x \in \mathrm{R}_+^o(\sigma_{m',j'}) \subset \mathrm{R}_+^c(\sigma_{m',j'})$. Else, if $m' > m$, for $P \in \mathcal{C}$, $P(y) = 0 \Rightarrow -\varepsilon_{m'} \le P(x) \le \varepsilon'_m$ since $y = \lim_{\varepsilon_m} x$ and $\epsilon_{m'} \gg \epsilon_m$. Moreover, from the definition of y and the fact that $P \in \mathrm{R}[X_1, \ldots, X_k]$, we have that $P(y) > 0 \Rightarrow P(x) > 0$, and $P(y) < 0 \Rightarrow P(x) < 0$, for $P \in \mathcal{C}$. So, if $m' > m$, then $x \in \mathrm{R}_+^c(\sigma_{m',j'})$.

Now, there exists $\ell', 1 \le \ell' \le N$, such that $\mathcal{R}(\sigma_{m',j'}) \subset S_{\ell'}$. Consider the set $S_{\ell'}^{m',j'}$. We have that $\mathcal{R}_+^c(\sigma_{m',j'}) \subset S_{\ell'}^{m',j'}$. Thus, $x \in S_{\ell'}^{m',j'}$, with either $(m, j) = (m', j')$ or $m' > m$. In either case, this contradicts the choice of m and j. □

6 Constructing closed acyclic coverings of closed semi-algebraic sets

In this section we describe an algorithm for computing closed contractible coverings of $\mathcal{P}$-closed semi-algebraic sets, combining the results of Section 4 and Section 5.

Algorithm 6.1 (Acyclic Covering)

Input : *a finite set of s polynomials $\mathcal{P} \subset \mathrm{D}[X_1, \ldots, X_k]$ in strong ℓ-general position on R^k*

a $\mathcal{P}$-closed semi-algebraic set S contained in the sphere of center 0 and radius r.

Output : *a finite family of t polynomials $\mathcal{C} \subset \mathrm{D}[X_1, \ldots, X_k]$, such that*

the finite family $\mathcal{C}' \subset \mathrm{D}'[X_1, \ldots, X_k]$ (with $\mathrm{D}' = \mathrm{D}[\epsilon_t, \ldots, \epsilon_1]$) defined by $\mathcal{C}' = \{Q \pm \epsilon_i \mid Q \in \mathcal{C}, 1 \le i \le t\}$.

a set of $\mathcal{C}'$-closed formulas $\{\phi_1, \ldots, \phi_M\}$ such that

each $\mathcal{R}(\phi_i, \mathrm{R}'^k)$ is acyclic,

their union $\cup_{1\le i\le M} \mathcal{R}(\phi_i, \mathrm{R}'^k) = \mathrm{Ext}(S, \mathrm{R}')$.

Procedure :

Step 1 *Let $Q = X_1^2 + \ldots + X_k^2 - r^2$. Call Algorithm 4.3 (Parametrized Bounded Connecting) with input $Q, \mathcal{P}$. For each $\theta_i \in \Theta$ in the output, let S_i denote the semi-algebraic set $\bigcup_{y\in\mathcal{R}(\theta_i,\mathrm{R}^k)} \mathrm{Image}(\gamma(\theta_i)(y, \cdot))$. Let ψ_i denote the formula defining S_i. Define $\mathcal{C}$ as the set of all polynomials used in the definition of all the S_i's and compute $\mathrm{Sign}(S, \mathcal{C})$. .*

Step 2 *For each of the $\mathcal{C}$-semi-algebraic sets S_i obtained in Step 1, output the formula ϕ_i describing the $\mathcal{C}'$-closed semi-algebraic set S'_i, using the notation in the Gabrielov-Vorobjov construction.*

Proof of correctness : The correctness of the algorithm is a consequence of Propositions 4.4, 4.2, 5.2, and Theorem 5.1. □

Complexity analysis: The complexity of Step 1 of the algorithm is bounded by $s^{k+1} d^{O(k^4)}$, where s is a bound on the number of elements of $\mathcal{P}$ and d is a bound on the degrees of the elements of $\mathcal{P}$, using the complexity analysis of Algorithm 4.3 (Parametrized Bounded Connecting). The number of polynomials in $\mathcal{C}$ is $s^{k+1} d^{O(k^4)}$ and their degree is $d^{O(k^3)}$. Thus the complexity of computing $\mathrm{Sign}_m(S, \mathcal{C})$ is bounded by $s^{(k+1)^2} d^{O(k^5)}$ using Algorithm 13.37 (Sampling on an Algebraic Set) in [9].

□

7 Computing the first Betti number of $\mathcal{P}$- semi-algebraic sets

Let $A_1, \ldots, A_n$ be sub-complexes of a finite simplicial complex A such that $A = A_1 \cup \cdots \cup A_n$. Note that the intersections of any number of the sub-complexes, A_i, is again a sub-complex of A. We will denote by $A_{i_0,\ldots,i_p}$ the sub-complex $A_{i_0} \cap \cdots \cap A_{i_p}$.

Let $C^i(A)$ denote the $\mathbb{Q}$-vector space of i co-chains of A, and $C^\bullet(A)$, the complex

$$\cdots \to C^{q-1}(A) \xrightarrow{d} C^q(A) \xrightarrow{d} C^{q+1}(A) \to \cdots$$

where $d : C^q(A) \to C^{q+1}(A)$ are the usual co-boundary homomorphisms. More precisely, given $\omega \in C^q(A)$, and a $q+1$ simplex $[a_0, \ldots, a_{q+1}] \in A$,

$$d\omega([a_0, \ldots, a_{q+1}]) = \sum_{0 \le i \le q+1} (-1)^i \omega([a_0, \ldots, \hat{a}_i, \ldots, a_{q+1}]) \tag{3}$$

(here and everywhere else in the paper $\hat{}$ denotes omission). Now extend $d\omega$ to a linear form on all of $C_{q+1}(A)$ by linearity, to obtain an element of $C^{q+1}(A)$.

The generalized Mayer-Vietoris sequence is the following:

$$0 \longrightarrow C^\bullet(A) \xrightarrow{r} \prod_{i_0} C^\bullet(A_{i_0}) \xrightarrow{\delta_1} \prod_{i_0 < i_1} C^\bullet(A_{i_0,i_1})$$

$$\cdots \xrightarrow{\delta_{p-1}} \prod_{i_0 < \cdots < i_p} C^\bullet(A_{i_0,\ldots,i_p}) \xrightarrow{\delta_p} \prod_{i_0 < \cdots < i_{p+1}} C^\bullet(A_{i_0,\ldots,i_{p+1}}) \cdots$$

where r is induced by restriction and the connecting homomorphisms δ are described below.

Given an $\omega \in \prod_{i_0 < \cdots < i_p} C^q(A_{i_0,\ldots,i_p})$ we define $\delta(\omega)$ as follows:
first note that $\delta(\omega) \in \prod_{i_0 < \cdots < i_{p+1}} C^q(A_{i_0,\ldots,i_{p+1}})$, and it suffices to define $\delta(\omega)_{i_0,\ldots,i_{p+1}}$ for each $(p+2)$-tuple $0 \le i_0 < \cdots < i_{p+1} \le n$. Note that, $\delta(\omega)_{i_0,\ldots,i_{p+1}}$ is a linear form on the vector space, $C_q(A_{i_0,\ldots,i_{p+1}})$, and hence is determined by its values on the q-simplices in the complex $A_{i_0,\ldots,i_{p+1}}$. Furthermore, each q-simplex, $s \in A_{i_0,\ldots,i_{p+1}}$ is automatically a simplex of the complexes $A_{i_0,\ldots,\hat{i}_i,\ldots i_{p+1}}$, $0 \le i \le p+1$.

We define,

$$(\delta\omega)_{i_0,\ldots,i_{p+1}}(s) = \sum_{0 \le i \le p+1} (-1)^i \omega_{i_0,\ldots,\hat{i}_i,\ldots,i_{p+1}}(s),$$

(here and everywhere else in the paper $\hat{}$ denotes omission). The fact that the generalized Mayer-Vietoris sequence is exact is classical (see [2] for example).

The cohomology groups $H^0(A_{i_0,\ldots,i_p})$ are isomorphic to the $\mathbb{Q}$-vector space of locally constant functions on $A_{i_0,\ldots,i_p}$ and the induced homomorphisms, $\delta_p : H^*(A_{i_0,\ldots,i_p}) \to H^*(A_{i_0,\ldots,i_{p+1}})$ are then given by generalized restrictions, i.e. for

$$\phi \in \oplus_{1 \le i_0 < \cdots < i_p \le n} H^0(A_{i_0,\ldots,i_p}),$$

a locally constant function on $A_{i_0,\ldots,i_p}$,

$$\delta_p(\phi)_{i_0,\ldots,i_{p+1}} = \sum_{i=0}^{p} (-1)^i \phi_{i_0,\ldots,\hat{i}_i,\ldots,i_{p+1}}|_{A_{i_0,\ldots,i_{p+1}}}.$$

The following proposition provides the key tool for computing the first Betti number.

Proposition 7.1 *Let $A_1, \ldots, A_n$ be sub-complexes of a finite simplicial complex A such that $A = A_1 \cup \cdots \cup A_n$ and each A_i is acyclic, that is $H^0(A_i) = \mathbb{Q}$ and $H^q(A_i) = 0$ for all $q > 0$. Then, $b_1(A) = \dim(\mathrm{Ker}(\delta_2)) - \dim(\mathrm{Im}(\delta_1))$, with*

$$\prod_i H^0(A_i) \xrightarrow{\delta_1} \prod_{i<j} H^0(A_{i,j}) \xrightarrow{\delta_2} \prod_{i<j<\ell} H^0(A_{i,j,\ell})$$

Proof: See Appendix A. □

Algorithm 7.2 (First Betti Number of a $\mathcal{P}$- Semi-algebraic Set)

Input : *a finite set of polynomials* $\mathcal{Q} \subset \mathrm{D}[X_1, \ldots, X_k]$,

a formula defining a $\mathcal{Q}$*-semi-algebraic set,* T.

Output : $b_1(T)$.

Procedure :

Step 1 *Let* ε *be an infinitesimal. Define* $\tilde{T}$ *as the intersection of* $\mathrm{Ext}(T, \langle\varepsilon\rangle)$ *with the ball of center* 0 *and radius* $1/\varepsilon$*. Define* $\mathcal{P}$ *as* $\mathcal{Q} \cup \{\varepsilon^2(X_1^2 + \ldots + X_k^2 + X_{k+1}^2) - 4, X_{k+1}\}$ *Replace* $\tilde{T}$ *by the* $\mathcal{P}$*- semi-algebraic set* S *defined as the intersection of the cylinder* $\tilde{T} \times \mathrm{R}\langle\varepsilon\rangle$ *with the upper hemisphere defined by* $\varepsilon^2(X_1^2 + \ldots + X_k^2 + X_{k+1}^2) = 4, X_{k+1} \geq 0$.

Step 2 *Using the Gabrielov-Vorobjov construction described above, replace* S *by a* $\mathcal{P}'$*-closed set,* S'*. Note that* $\mathcal{P}'$ *is in general position with respect to the sphere of center* 0 *and radius* $2/\varepsilon$*.*

Step 3 *Use Algorithm 6.1 (Acyclic Covering) with input* $\varepsilon^2(X_1^2 + \ldots + X_k^2 + X_{k+1}^2) - 4$ *and* $\mathcal{P}'$*, to compute a covering of* S' *by closed, bounded and acyclic sets,* S_i*, described by formulas* ϕ_i*.*

Step 4 *Use Algorithm 16.27 (General Roadmap) in [9] to compute sample points of the connected components of the pairwise and triplewise intersections of the* S_i*'s and compute their incidences.*

Step 5 *Using linear algebra compute* $\dim(\mathrm{Ker}(\delta_2)) - \dim(\mathrm{Im}(\delta_1))$*, with*

$$\prod_i H^0(S_i) \xrightarrow{\delta_1} \prod_{i<j} H^0(S_i \cap S_j) \xrightarrow{\delta_2} \prod_{i<j<\ell} H^0(S_i \cap S_j \cap S_\ell)$$

Proof of correctness : The correctness of the algorithm is a consequence of Theorem 5.1, Algorithm 6.1 (Acyclic Covering), Algorithm 16.21 (Connected Components of a Semi-algebraic Set) in [9], and Proposition 7.1. □

Complexity analysis: Each step is clearly singly exponential from the complexity analysis of Algorithm 6.1 (Acyclic Covering), and Algorithm 16.27 (General Roadmap) in [9] and the fact that the linear algebra in Step 5 can also be performed in singly exponential time. □

8 Computing Connected Components

If one is interested in computing semi-algebraic descriptions of the connected components of a given semi-algebraic set, then using Algorithm 4.3 (Parametrized Bounded Connecting) it is possible to do so with a complexity making precise the one of previously known algorithms, whose complexities were of the form $(sd)^{k^{O(1)}}$ (see [22]). We have the following theorems (we refer the reader to [9] for details of the proof).

Theorem 8.1 *If* $\mathrm{Z}(Q, \mathrm{R}^k)$ *is an algebraic set defined as the zero set of a polynomial* $Q \in \mathrm{D}[X_1, \ldots, X_k]$ *of degree* $\leq d$*, then there is an algorithm that outputs quantifier free formulas whose realizations are the semi-algebraically connected components of* $\mathrm{Z}(Q, \mathrm{R}^k)$*. The complexity of the algorithm in the ring generated by the coefficients of* Q *is bounded by* $d^{O(k^3)}$ *and the degrees of the polynomials that appear in the output are bounded by* $O(d)^{k^2}$*. Moreover, if* $\mathrm{D} = \mathbb{Z}$*, and the bitsizes of the coefficients of the polynomials are bounded by* τ*, then the bitsizes of the integers appearing in the intermediate computations and the output are bounded by* $\tau d^{O(k^2)}$*.*

Theorem 8.2 *Let* $\mathcal{P} = \{P_1, \ldots, P_s\} \subset \mathrm{D}[X_1, \ldots, X_k]$ *with* $\deg(P_i) \leq d, 1 \leq i \leq s$ *and a semi-algebraic set* S *defined by a* $\mathcal{P}$ *quantifier-free formula. There exists an algorithm that outputs quantifier-free semi-algebraic descriptions of all the semi-algebraically connected components of* S*. The complexity of the algorithm is bounded by* $s^{k+1}d^{O(k^4)}$*. The degrees of the polynomials that appear in the output are bounded by* $d^{O(k^3)}$*. Moreover, if the input polynomials have integer coefficients whose bitsize is bounded by* τ *the bitsize of coefficients output is* $d^{O(k^3)}\tau$*.*

References

[1] S. Basu, *On Bounding the Betti Numbers and Computing the Euler Characteristics of Semi-algebraic Sets*, Discrete and Computational Geometry, 22 1-18 (1999).

[2] S. Basu, *On different bounds on different Betti numbers*, Discrete and Computational Geometry, Vol 30, No. 1 (2003).

[3] S. Basu, *Computing the first few Betti numbers of semi-algebraic sets in singly exponential time*, preprint (2004).
(Available at `www.math.gatech.edu/~saugata/bettifew.ps`.)

[4] S. Basu, R. Pollack, M.-F. Roy, *On the Combinatorial and Algebraic Complexity of Quantifier Elimination*, Journal of the ACM , 43 1002–1045, (1996).

[5] S. Basu, R. Pollack, M.-F. Roy, *On the number of cells defined by a family of polynomials on a variety*, Mathematika, 43 (1996) 120-126.

[6] S. Basu, R. Pollack, M.-F. Roy, *On Computing a Set of Points meeting every Semi-algebraically Connected Component of a Family of Polynomials on a Variety*, Journal of Complexity, March 1997, Vol 13, Number 1, 28-37.

[7] S. Basu, R. Pollack, M.-F. Roy, *Computing Roadmaps of Semi-algebraic Sets on a Variety*, Journal of the AMS, vol 3, 1 55-82 (1999).

[8] S. Basu, R. Pollack, M.-F. Roy, *On the Betti numbers of sign conditions*, Proceedings of the AMS (2004).

[9] S. Basu, R. Pollack, M.-F. Roy, *Algorithms in Real Algebraic Geometry*, Springer-Verlag, 2003.

[10] S. Basu, R. Pollack, M.-F. Roy, *Computing the first Betti number and the connected components of semi-algebraic sets*, (full paper in preparation).

[11] M. Ben-Or, D. Kozen , J. Reif, *The complexity of elementary algebra and geometry*, J. of Computer and Systems Sciences, 18:251– 264, (1986).

[12] J. Bochnak, M. Coste, M.-F. Roy, *Géométrie algébrique réelle,* Springer-Verlag (1987). *Real algebraic geometry*, Springer-Verlag (1998).

[13] A. Borel, J.C. Moore, *Homology theory for locally compact spaces*, Mich. Math. J., 7:137– 159, (1960).

[14] P. Burgisser, F. Cucker, *Counting Complexity Classes for Numeric Computations II: Algebraic and Semi-algebraic Sets*, preprint.

[15] J. Canny, *Computing road maps in general semi-algebraic sets*, The Computer Journal, 36: 504–514, (1993).

[16] G. Collins, *Quantifier elimination for real closed fields by cylindric algebraic decomposition*, In Second GI Conference on Automata Theory and Formal Languages. Lecture Notes in Computer Science, vol. 33, pp. 134-183, Springer- Verlag, Berlin (1975).

[17] D. Cox, J. Little, D. O'Shea, *Ideals, varieties and algorithms: an introduction to computational algebraic geometry and commutative algebra*, Undergraduate Texts in Mathematics, Springer-Verlag, New York (1997).

[18] A. Gabrielov, N. Vorobjov Betti Numbers for Quantifier-free Formulae, to appear in Discrete and Computational Geometry.

[19] L. Gournay, J. J. Risler, *Construction of roadmaps of semi-algebraic sets*, Appl. Algebra Eng. Commun. Comput. 4, No.4, 239-252 (1993).

[20] D. Grigor'ev, N. Vorobjov, *Counting connected components of a semi-algebraic set in subexponential time*, Comput. Complexity 2, No.2, 133-186 (1992).

[21] R. M. Hardt, *Semi-algebraic Local Triviality in Semi-algebraic Mappings*, Am. J. Math. 102, 291-302 (1980).

[22] J. Heintz, M.-F. Roy, P. Solernò, *Description of the Connected Components of a Semialgebraic Set in Single Exponential Time*, Discrete and Computational Geometry, 11, 121-140 (1994).

[23] J. Milnor, *On the Betti numbers of real varieties*, Proc. AMS 15, 275-280 (1964).

[24] O. A. Oleĭnik, *Estimates of the Betti numbers of real algebraic hypersurfaces*, Mat. Sb. (N.S.), 28 (70): 635–640 (Russian) (1951).

[25] O. A. Oleĭnik, I. B. Petrovskii, *On the topology of real algebraic surfaces*, Izv. Akad. Nauk SSSR 13, 389-402 (1949).

[26] P. Pedersen, M.-F. Roy, A. Szpirglas, *Counting real zeroes in the multivariate case*, Computational algebraic geometry, Eyssette et Galligo ed. Progress in Mathematics 109, 203-224, Birkhauser (1993).

[27] J. Renegar. *On the computational complexity and geometry of the first order theory of the reals*, Journal of Symbolic Computation, 13: 255–352 (1992).

[28] M.-F. Roy, A. Szpirglas *Complexity of computation on real algebraic numbers*, Journal of Symbolic Computation 10, No.1, 39-51 (1990).

[29] E. H. Spanier Algebraic Topology, McGraw-Hill Book Company, 1966.

[30] R. Thom, *Sur l'homologie des variétés algébriques réelles*, Differential and Combinatorial Topology, 255–265. Princeton University Press, Princeton (1965).

9 Appendix A

To prove Proposition 7.1, we consider the following bi-graded double complex $\mathcal{M}^{p,q}$, with a total differential $D = \delta + (-1)^p d$, where

$$\mathcal{M}^{p,q} = \prod_{i_0,\dots,i_p} C^q(A_{i_0,\dots,i_p}).$$

$$\begin{array}{ccccccccc}
 & & \vdots & & \vdots & & \vdots & & \\
 & & \uparrow d & & \uparrow d & & \uparrow d & & \\
0 & \longrightarrow & \prod_{i_0} C^3(A_{i_0}) & \xrightarrow{\delta} & \prod_{i_0<i_1} C^3(A_{i_0,i_1}) & \xrightarrow{\delta} & \prod_{i_0<i_1<i_2} C^3(A_{i_0,i_1,i_2}) & \longrightarrow & \\
 & & \uparrow d & & \uparrow d & & \uparrow d & & \\
0 & \longrightarrow & \prod_{i_0} C^2(A_{i_0}) & \xrightarrow{\delta} & \prod_{i_0<i_1} C^2(A_{i_0,i_1}) & \xrightarrow{\delta} & \prod_{i_0<i_1<i_2} C^2(A_{i_0,i_1,i_2}) & \longrightarrow & \\
 & & \uparrow d & & \uparrow d & & \uparrow d & & \\
0 & \longrightarrow & \prod_{i_0} C^1(A_{i_0}) & \xrightarrow{\delta} & \prod_{i_0<i_1} C^1(A_{i_0,i_1}) & \xrightarrow{\delta} & \prod_{i_0<i_1<i_2} C^1(A_{i_0,i_1,i_2}) & \longrightarrow & \\
 & & \uparrow d & & \uparrow d & & \uparrow d & & \\
0 & \longrightarrow & \prod_{i_0} C^0(A_{i_0}) & \xrightarrow{\delta} & \prod_{i_0<i_1} C^0(A_{i_0,i_1}) & \xrightarrow{\delta} & \prod_{i_0<i_1<i_2} C^0(A_{i_0,i_1,i_2}) & \longrightarrow & \\
 & & \uparrow d & & \uparrow d & & \uparrow d & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

There are two spectral sequences (corresponding to taking horizontal or vertical filtrations respectively) associated with $\mathcal{M}^{p,q}$ both converging to $H^*_D(\mathcal{M})$. The first terms of these are ${}'E_1 = H_\delta \mathcal{M}, {}'E_2 = H_d H_\delta \mathcal{M}$, and $''E_1 = H_d\mathcal{M}, ''E_2 = H_\delta H_d \mathcal{M}$. Because of the exactness of the generalized Mayer-Vietoris sequence, we have that,

$$'E_1 = \left| \begin{array}{cccccc} \vdots & \vdots & \vdots & \vdots & \vdots & \\ \uparrow d & \uparrow 0 & \uparrow 0 & \uparrow 0 & \uparrow 0 & \\ C^3(A) & 0 & 0 & 0 & 0 & \cdots \\ \uparrow d & \uparrow 0 & \uparrow 0 & \uparrow 0 & \uparrow 0 & \\ C^2(A) & 0 & 0 & 0 & 0 & \cdots \\ \uparrow d & \uparrow 0 & \uparrow 0 & \uparrow 0 & \uparrow 0 & \\ C^1(A) & 0 & 0 & 0 & 0 & \cdots \\ \uparrow d & \uparrow 0 & \uparrow 0 & \uparrow 0 & \uparrow 0 & \\ C^0(A) & 0 & 0 & 0 & 0 & \cdots \\ \hline \end{array} \right.$$

and

$$'E_2 = \left| \begin{array}{ccccccc} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ H^3(A) & 0 & 0 & 0 & 0 & 0 & \cdots \\ H^2(A) & 0 & 0 & 0 & 0 & 0 & \cdots \\ H^1(A) & 0 & 0 & 0 & 0 & 0 & \cdots \\ H^0(A) & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline \end{array} \right.$$

The degeneration of this sequence at E_2 shows that $H^*_D(\mathcal{M}) \cong H^*(A)$.
The initial term $''E_1$ of the second spectral sequence is given by,

$$''E_1 = \left| \begin{array}{cccccc} \vdots & & \vdots & & \vdots & \\ \prod_{i_0} H^3(A_{i_0}) & \xrightarrow{\delta} & \prod_{i_0<i_1} H^3(A_{i_0,i_1}) & \xrightarrow{\delta} & \prod_{i_0<i_1<i_2} H^3(A_{i_0,i_1,i_2}) & \longrightarrow \\ \prod_{i_0} H^2(A_{i_0}) & \xrightarrow{\delta} & \prod_{i_0<i_1} H^2(A_{i_0,i_1}) & \xrightarrow{\delta} & \prod_{i_0<i_1<i_2} H^2(A_{i_0,i_1,i_2}) & \longrightarrow \\ \prod_{i_0} H^1(A_{i_0}) & \xrightarrow{\delta} & \prod_{i_0<i_1} H^1(A_{i_0,i_1}) & \xrightarrow{\delta} & \prod_{i_0<i_1<i_2} H^1(A_{i_0,i_1,i_2}) & \longrightarrow \\ \prod_{i_0} H^0(A_{i_0}) & \xrightarrow{\delta} & \prod_{i_0<i_1} H^0(A_{i_0,i_1}) & \xrightarrow{\delta} & \prod_{i_0<i_1<i_2} H^0(A_{i_0,i_1,i_2}) & \longrightarrow \\ \hline \end{array} \right.$$

The cohomology groups $H^0(A_{i_0,\ldots,i_p})$ occurring as summands in the bottom row of $''E_1$ are isomorphic to the $\mathbb{Q}$-vector space of locally constant functions on $A_{i_0,\ldots,i_p}$ and the homomorphisms, $''d_1 : ''E_1^{p,0} \to ''E_1^{p+1,0}$ are then given by generalized restrictions, i.e. for

$$\phi \in \oplus_{1 \le i_0 < \cdots < i_p \le n} H^0(A_{i_0,\ldots,i_p}),$$

a locally constant function on $A_{i_0,\ldots,i_p}$,

$$''d_1(\phi)_{i_0,\ldots,i_{p+1}} = \sum_{i=0}^{p} (-1)^i \phi_{i_0,\ldots,\hat{i_i},\ldots,i_{p+1}}|_{A_{i_0,\ldots,i_{p+1}}}.$$

Proof of Proposition 7.1: Since, $H^q(A_i) = 0$ for all $q > 0$, all the terms in the first column of $''E_1$ are zero except the bottom term, and clearly $''d_2^{0,1} = ''d_2^{1,0} = 0$. Thus, $''E_\infty^{1,0} = ''E_2^{1,0}$ and $''E_\infty^{0,1} = 0$. Thus, $H^1(A) \cong ''E_\infty^{1,0} \oplus ''E_\infty^{0,1} \cong ''E_2^{1,0}$. □