New bounds for Betti numbers of semi-algebraic sets and algorithms for computing them

> Saugata Basu School of Mathematics Georgia Institute of Technology.

Semi-algebraic Sets

 Subsets of R^k defined by a formula involving a finite number of polynomial equalities and inequalities.

Semi-algebraic Sets

- Subsets of R^k defined by a formula involving a finite number of polynomial equalities and inequalities.
- A basic semi-algebraic set is one defined by a conjunction of weak inequalities of the form P ≥ 0.

Semi-algebraic Sets

- Subsets of R^k defined by a formula involving a finite number of polynomial equalities and inequalities.
- A basic semi-algebraic set is one defined by a conjunction of weak inequalities of the form P ≥ 0.
- They arise as configurations spaces (in robotic motion planning, molecular chemistry etc.), CAD models and many other applications in computational geometry.

Uniform bounds on the number of connected components, Betti numbers etc., in terms of the degrees and number of polynomials used to describe them.

Parameters of complexity:

Uniform bounds on the number of connected components, Betti numbers etc., in terms of the degrees and number of polynomials used to describe them.

Parameters of complexity: Number of polynomials : *n* (controls the *combinatorial complexity*)

Uniform bounds on the number of connected components, Betti numbers etc., in terms of the degrees and number of polynomials used to describe them.

Parameters of complexity: Number of polynomials : n (controls the *combinatorial complexity*) Degree bound : d (controls the *algebraic complexity*)

Uniform bounds on the number of connected components, Betti numbers etc., in terms of the degrees and number of polynomials used to describe them.

Parameters of complexity: Number of polynomials : n (controls the *combinatorial complexity*) Degree bound : d (controls the *algebraic complexity*) Dimension of the ambient space : k

Classical Result on the Topology of Semi-algebraic Sets

Theorem 1. (Oleinik and Petrovsky, Thom, Milnor) Let $S \subset \mathbb{R}^k$ be the set defined by the conjunction of n inequalities,

$$P_1 \ge 0, \dots, P_n \ge 0, P_i \in \mathbb{R}[X_1, \dots, X_k],$$

 $deg(P_i) \leq d, 1 \leq i \leq n$. Then,

$$\sum_{i} \beta_{i}(S) = nd(2nd - 1)^{k-1} = O(nd)^{k}.$$

Tightness

The above bound is actually quite tight. Example: Let

$$P_i = L_{i,1}^2 \cdots L_{i,\lfloor d/2 \rfloor}^2 - \epsilon,$$

where the L_{ij} 's are generic linear polynomials and $\epsilon > 0$ and sufficiently small. The set S defined by $P_1 \ge 0, \ldots, P_n \ge 0$ has $\Omega(nd)^k$ connected components and hence $\beta_0(S) = \Omega(nd)^k$.

What about the higher Betti Numbers ?

- Cannot construct examples such that $\beta_i(S) = \Omega(nd)^k$ for i > 0.
- The technique used for proving the above result does not help:

Replace the semi-algebraic set S by another set bounded by a smooth algebraic hypersurface of degree nd having the same homotopy type as S.

Then bound the Betti numbers of this hypersurface using Morse theory and the Bezout bound on the number of solutions of a system of polynomial equations.



Graded Bounds

Theorem 2. (*B*, 2001) Let $S \subset \mathbb{R}^k$ be the set defined by the conjunction of n inequalities,

$$P_1 \ge 0, \ldots, P_n \ge 0, P_i \in \mathbb{R}[X_1, \ldots, X_k],$$

 $deg(P_i) \le d, 1 \le i \le n.$

contained in a variety Z(Q) of real dimension k', and $deg(Q) \leq d$. Then,

$$\beta_i(S) \le \binom{n}{k'-i} (2d)^k.$$

The case of the union

Theorem 3. (B, 2001) Let $S \subset \mathbb{R}^k$ be the set defined by the disjunction of n inequalities,

$$P_1 \ge 0, \ldots, P_n \ge 0, P_i \in \mathbb{R}[X_1, \ldots, X_k],$$

 $deg(P_i) \leq d, 1 \leq i \leq n$. Then,

$$\beta_i(S) \le \binom{n}{i+1} (2d)^k.$$

Sets defined by Quadratic Inequalities

Sets defined by Quadratic Inequalities

Theorem 4. (B, 2001) Let ℓ be any fixed number and let $S \subset \mathbb{R}^k$ be defined by $P_1 \ge 0, \ldots, P_n \ge 0$ with $\deg(P_i) \le 2$. Then, $\beta_{k-\ell}(S) \le {n \choose \ell} k^{O(\ell)}$.

Sets defined by Quadratic Inequalities

Theorem 4. (B, 2001) Let ℓ be any fixed number and let $S \subset \mathbb{R}^k$ be defined by $P_1 \ge 0, \ldots, P_n \ge 0$ with $\deg(P_i) \le 2$. Then, $\beta_{k-\ell}(S) \le {n \choose \ell} k^{O(\ell)}$.

This bound is polynomial.

Notice that the lowest Betti numbers of S better not be polynomially bounded. Example:

S defined by $X_1(X_1 - 1) \ge 0, \dots, X_k(X_k - 1) \ge 0$. Clearly, $\beta_0(S) = 2^k$.

Betti Numbers of Sign Patterns I

Betti Numbers of Sign Patterns I

Let Q and P be finite subsets of ℝ[X₁,...,X_k]. A sign condition on P is an element of {0,1,-1}^P.

Betti Numbers of Sign Patterns I

- Let Q and P be finite subsets of ℝ[X₁,...,X_k]. A sign condition on P is an element of {0,1,-1}^P.
- Let b_i(σ) denote the *i*-th Betti number of the realization of σ, and let b_i(Q, P) = Σ_σ b_i(σ).

Betti Numbers of Sign Patterns II

Betti Numbers of Sign Patterns II

 Let b_i(d, k, k', n) be the maximum of b_i(Q, P) over all Q, P where Q and P are finite subsets of of R[X₁,...,X_k], whose elements have degree at most d, #(P) = n and the algebraic set Z(Q) has dimension k'.

Betti Numbers of Sign Patterns II

- Let b_i(d, k, k', n) be the maximum of b_i(Q, P) over all Q, P where Q and P are finite subsets of of R[X₁,...,X_k], whose elements have degree at most d, #(P) = n and the algebraic set Z(Q) has dimension k'.
- Previously known (B, Pollack, Roy (1995))

$$b_0(d,k,k',n) = \binom{4n}{k'} d(2d-1)^{k-1} = \binom{n}{k'} O(d)^k.$$

Betti Numbers of Sign Patterns III

Theorem 5. (*B*, *Pollack*, *Roy*, 2002)

$$b_i(d,k,k',n) \le \sum_{0 \le j \le k'-i} \binom{n}{j} 4^j d(2d-1)^{k-1} = \binom{n}{k'-i} O(d)^k.$$

Betti Numbers of Sign Patterns III

Theorem 5. (*B*, *Pollack*, *Roy*, 2002)

$$b_i(d,k,k',n) \le \sum_{0 \le j \le k'-i} \binom{n}{j} 4^j d(2d-1)^{k-1} = \binom{n}{k'-i} O(d)^k.$$

Applications ?

Generalized Mayer-Vietoris Exact Sequence

- Let A₁,..., A_n be subcomplexes of a finite simplicial complex A such that A = A₁∪···∪A_n. Let Cⁱ(A) denote the ℝ-vector space of i co-chains of A, and C^{*}(A) = ⊕_iCⁱ(A).
- We will denote by $A_{\alpha_0,...,\alpha_p}$ the subcomplex $A_{\alpha_0} \cap \cdots \cap A_{\alpha_p}$.
- The following sequence of homomorphisms is exact.

$$0 \longrightarrow C^*(A) \xrightarrow{r} \prod_{\alpha_0} C^*(A_{\alpha_0}) \xrightarrow{\delta} \prod_{\alpha_0 < \alpha_1} C^*(A_{\alpha_0,\alpha_1})$$
$$\cdots \xrightarrow{\delta} \prod_{\alpha_0 < \dots < \alpha_p} C^*(A_{\alpha_0,\dots,\alpha_p}) \cdots \xrightarrow{\delta} \prod_{\alpha_0 < \dots < \alpha_{p+1}} C^*(A_{\alpha_0,\dots,\alpha_{p+1}}) \cdots \xrightarrow{\delta} \cdots$$

Mayer-Vietoris Double Complex I

We now consider the following bigraded double complex $\mathcal{M}^{p,q}$, with a total differential $D = \delta + (-1)^p d$, where

$$\mathcal{M}^{p,q} = \prod_{\alpha_0,\dots,\alpha_p} C^q(A_{\alpha_0,\dots,\alpha_p}).$$

 $lpha_0, ..., lpha_p$



Double Complex



The Associated Total Complex



• A sequence of vector spaces progressively approximating the homology of the total complex. More precisely,

- A sequence of vector spaces progressively approximating the homology of the total complex. More precisely,
- a sequence of bi-graded vector spaces and differentials $(E_r, d_r : E_r^{p,q} \rightarrow E_r^{p+r,q-r+1}),$

- A sequence of vector spaces progressively approximating the homology of the total complex. More precisely,
- a sequence of bi-graded vector spaces and differentials $(E_r, d_r : E_r^{p,q} \rightarrow E_r^{p+r,q-r+1}),$
- $E_{r+1} = H(E_r, d_r),$

- A sequence of vector spaces progressively approximating the homology of the total complex. More precisely,
- a sequence of bi-graded vector spaces and differentials $(E_r, d_r : E_r^{p,q} \rightarrow E_r^{p+r,q-r+1}),$
- $E_{r+1} = H(E_r, d_r),$
- $E_{\infty} = H^*$ (Associated Total Complex).



Figure 1: The differentials d_r in the spectral sequence (E_r,d_r)
Two Spectral Sequences

• There are two spectral sequences associated with $\mathcal{M}^{p,q}$ both converging to $H^*_D(\mathcal{M})$. The first terms of these are:

Two Spectral Sequences

• There are two spectral sequences associated with $\mathcal{M}^{p,q}$ both converging to $H^*_D(\mathcal{M})$. The first terms of these are:

$$E_1 = H_{\delta}(\mathcal{M}), E_2 = H_d H_{\delta}(\mathcal{M})$$

Two Spectral Sequences

• There are two spectral sequences associated with $\mathcal{M}^{p,q}$ both converging to $H^*_D(\mathcal{M})$. The first terms of these are:

$$E_1 = H_{\delta}(\mathcal{M}), E_2 = H_d H_{\delta}(\mathcal{M})$$

$$E_1' = H_d(\mathcal{M}), E_2' = H_\delta H_d(\mathcal{M})$$



a a a



The degeneration of this sequence at E_2 shows that $H_D^*(\mathcal{M}) \cong H^*(A)$.

 $E_{1}' = \begin{array}{cccc} & & & & & & & & \\ & & & & & & \\ & & & & & \\ & & & & \\ & & & & \\ & &$

Lemma 1

Lemma 6. Let A be a finite simplicial complex and A_1, \ldots, A_n subcomplexes of A such that $A = A_1 \cup \cdots \cup A_n$. Suppose that for every ℓ , $0 \leq \ell \leq i$, and for every $(\ell + 1)$ tuple $A_{\alpha_0}, \ldots, A_{\alpha_\ell}$, $\beta_{i-\ell}(A_{\alpha_0,\ldots,\alpha_\ell}) \leq M$. Then, $\beta_i(A) \leq \sum_{0 < \ell < i} {n \choose \ell+1} M$.

Lemma 2

Lemma 7. Let $P_1, \ldots, P_l \in R[X_1, \ldots, X_k]$, $deg(P_i) \leq d$, and $l \leq k$. Let S be the set defined by the conjunction of the inequalities $P_i \geq 0$. Assume that S is bounded. Then, $\sum_i \beta_i(S) = (2d)^k$.

Theorem 3 follows. Theorem 2 follows by a dual argument. Theorem 4 follows using a result of Barvinok (1995).

An arrangement in \mathbb{R}^k is a collection of n objects in \mathbb{R}^k each of constant description complexity.

An arrangement in \mathbb{R}^k is a collection of n objects in \mathbb{R}^k each of constant description complexity.

• Arrangements of lines in the plane, or more generally hyperplanes in \mathbb{R}^k .

An arrangement in \mathbb{R}^k is a collection of n objects in \mathbb{R}^k each of constant description complexity.

- Arrangements of lines in the plane, or more generally hyperplanes in \mathbb{R}^k .
- Arrangements of balls or simplices in \mathbb{R}^k .

An arrangement in \mathbb{R}^k is a collection of n objects in \mathbb{R}^k each of constant description complexity.

- Arrangements of lines in the plane, or more generally hyperplanes in \mathbb{R}^k .
- Arrangements of balls or simplices in \mathbb{R}^k .
- Arrangements of semi-algebraic objects in \mathbb{R}^k , each defined by a fixed number of polynomials of constant degree.

Arrangements of lines in the \mathbb{R}^2



Arrangement of circles in \mathbb{R}^2



Arrangement of tori in \mathbb{R}^3



 Schwartz and Sharir, in their seminal papers on the Piano Mover's Problem (Motion Planning).

- Schwartz and Sharir, in their seminal papers on the Piano Mover's Problem (Motion Planning).
- Computing the Betti numbers of arrangements of balls by Edelsbrunner et al (Molecular Biology).

- Schwartz and Sharir, in their seminal papers on the Piano Mover's Problem (Motion Planning).
- Computing the Betti numbers of arrangements of balls by Edelsbrunner et al (Molecular Biology).
- Computing the Betti numbers of triangulated manifolds (Edelsbrunner, Dey, Guha et al).

Complexity of Algorithms

Complexity of Algorithms

 In computational geometry it is customary to study the combinatorial complexity of algorithms. The algebraic complexity (dependence on the degree) is considered to be a constant.

Complexity of Algorithms

- In computational geometry it is customary to study the combinatorial complexity of algorithms. The algebraic complexity (dependence on the degree) is considered to be a constant.
- We only count the number of algebraic operations and ignore the cost of doing linear algebra.

Two Approaches

Two Approaches

Global vs Local

First Approach (Global): Using Triangulations

First Approach (Global): Using Triangulations



Triangulation via Cylindrical Algebraic Decomposition



Computing Betti Numbers using Global Triangulations

• Compact semi-algebraic sets are finitely triangulable.

Computing Betti Numbers using Global Triangulations

- Compact semi-algebraic sets are finitely triangulable.
- First triangulate the arrangement using *Cylindrical algebraic decomposition* and then compute the Betti numbers of the corresponding simplicial complex.
- But ...

Computing Betti Numbers using Global Triangulations

- Compact semi-algebraic sets are finitely triangulable.
- First triangulate the arrangement using *Cylindrical algebraic decomposition* and then compute the Betti numbers of the corresponding simplicial complex.
- But ... CAD produces $O(n^{2^k})$ simplices in the worst case.

Second Approach (Local): Using the Nerve Complex

Second Approach (Local): Using the Nerve Complex

 If the sets have the special property that all their non-empty intersections are contractible we can use the *nerve lemma* (Leray, Folkman).

Second Approach (Local): Using the Nerve Complex

- If the sets have the special property that all their non-empty intersections are contractible we can use the *nerve lemma* (Leray, Folkman).
- The homology groups of the union are then isomorphic to the homology groups of a combinatorially defined complex called the *nerve complex*.

The Nerve Complex



Figure 2: The nerve complex of a union of disks

Computing the Betti Numbers via the Nerve Complex (local algorithm)

• The nerve complex has *n* vertices, one vertex for each set in the union, and a simplex for each *non-empty* intersection among the sets.

Computing the Betti Numbers via the Nerve Complex (local algorithm)

- The nerve complex has *n* vertices, one vertex for each set in the union, and a simplex for each *non-empty* intersection among the sets.
- Thus, the $(\ell+1)$ -skeleton of the nerve complex can be computed by testing for non-emptiness of each of the possible $\sum_{1 \le j \le \ell+2} {n \choose j} = O(n^{\ell+2})$ at most $(\ell+2)$ -ary intersections among the n given sets.
If the sets are such that the topology of the "small" intersections are controlled, then

- If the sets are such that the topology of the "small" intersections are controlled, then
- we can use the Leray spectral sequence as a substitute for the nerve lemma.

- If the sets are such that the topology of the "small" intersections are controlled, then
- we can use the Leray spectral sequence as a substitute for the nerve lemma.
- The algorithmic version gives the first efficient algorithm for computing the Betti numbers, without the double-exponential complexity entailed in CAD.

Main Result

Theorem 8. Let $S_1, \ldots, S_n \subset \mathbb{R}^k$ be compact semi-algebraic sets of constant description complexity and let $S = \bigcup_{1 \leq i \leq n} S_i$, and $0 \leq \ell \leq k-1$. Then, there is an algorithm to compute $\beta_0(S), \ldots, \beta_\ell(S)$, whose complexity is $O(n^{\ell+2})$.

• Compute the spectral sequence (E'_r, d_r) of the Mayer-Vietoris double complex.

- Compute the spectral sequence (E'_r, d_r) of the Mayer-Vietoris double complex.
- In order to compute β_{ℓ} , we only need to compute upto $E'_{\ell+2}$.

- Compute the spectral sequence (E'_r, d_r) of the Mayer-Vietoris double complex.
- In order to compute β_{ℓ} , we only need to compute upto $E'_{\ell+2}$. But the punchline is that:

- Compute the spectral sequence (E'_r, d_r) of the Mayer-Vietoris double complex.
- In order to compute β_{ℓ} , we only need to compute upto $E'_{\ell+2}$. But the punchline is that:
- In order to compute the differentials $d_r, 1 \le r \le \ell + 1$, it suffices to have independent triangulations of the different unions taken upto $\ell + 2$ at a time.

• For instance, it should be intuitively clear that in order to compute $\beta_0(\cup_i S_i)$ it suffices to triangulate pairs.

• Truly polynomial time algorithms for computing the highest Betti numbers of sets defined by quadratic inequalities ?

- Truly polynomial time algorithms for computing the highest Betti numbers of sets defined by quadratic inequalities ?
- Truly single exponential time algorithm for computing all the Betti numbers of semi-algebraic sets.

- Truly polynomial time algorithms for computing the highest Betti numbers of sets defined by quadratic inequalities ?
- Truly single exponential time algorithm for computing all the Betti numbers of semi-algebraic sets.
- Same idea is applicable as a divide-and-conquer tool for computing the homology of arbitrary simplicial complexes, given a covering. What kind of efficiency do we derive ?

- Truly polynomial time algorithms for computing the highest Betti numbers of sets defined by quadratic inequalities ?
- Truly single exponential time algorithm for computing all the Betti numbers of semi-algebraic sets.
- Same idea is applicable as a divide-and-conquer tool for computing the homology of arbitrary simplicial complexes, given a covering. What kind of efficiency do we derive ?

• To what extent does topological simplicity aid algorithms in computational geometry ?