ON TOPOLOGICAL LOWER BOUNDS FOR ALGEBRAIC COMPUTATION TREES

ANDREI GABRIELOV AND NICOLAI VOROBJOV

Abstract. We prove that the height of any algebraic computation tree for deciding membership in a semialgebraic set $\Sigma \subset \mathbb{R}^n$ is bounded from below by

$$\frac{c_1 \log(\mathrm{b}_m(\Sigma))}{m+1} - c_2 n,$$

where $\mathrm{b}_m(\Sigma)$ is the m-th Betti number of Σ with respect to "ordinary" (singular) homology, and $c_1,\ c_2$ are some (absolute) positive constants. This result complements the well known lower bound by Yao [7] for *locally closed* semialgebraic sets in terms of the total *Borel-Moore* Betti number.

We also prove that if $\rho : \mathbb{R}^n \to \mathbb{R}^{n-r}$ is the projection map, then the height of any tree deciding membership in Σ is bounded from below by

$$\frac{c_1 \log(\mathrm{b}_m(\rho(\Sigma)))}{(m+1)^2} - \frac{c_2 n}{m+1}$$

for some positive constants $c_1,\ c_2$.

We illustrate these general results by examples of lower complexity bounds for some specific computational problems.

1. Introduction

The *algebraic computation tree* is a standard sequential model of computation for deciding membership problems for semialgebraic sets. Among various general methods for obtaining lower complexity bounds for this model, one of the most efficient uses homotopy invariants, Euler characteristic and Betti numbers, as arguments for the bounding functions. The history of this approach started probably in mid 70s with the work of Dobkin and Lipton [3], and features such prominent results as Ben-Or's [1] in 1983 and Yao's [7] in 1997. The present paper is inspired by the latter. We will discuss the results of [7] in some detail.

We assume that the reader is familiar with the concept of the algebraic computation tree, so we give here just a brief formal description, closely following [2].

Definition 1.1. An algebraic computation tree $\mathcal{T}$ with input variables $X_1, \ldots, X_n$ taking real values, is a tree having three types of vertices: computation (outdegree 1), branch (outdegree 3), and leaves (outdegree 0). To each vertex v of $\mathcal{T}$ a variable Y_v is assigned.

With each computation vertex v an expression $Y_v = a * b$ is associated, where $* \in \{+, -, \times, /\}$, and a, b are either real constants, or input variables, or variables associated with predecessor vertices of v, or a combination of these.

At each branch vertex v, the variable Y_v is assigned the value which is either a real constant, or an input variable, or a variable associated with a predecessor vertex. The three outgoing edges of v correspond to signs $Y_v > 0$, $Y_v = 0$, $Y_v < 0$.

With each leaf w a basic semialgebraic set (called *leaf set*) is associated, defined by equations of the kind either $Y_v > 0$, or $Y_v = 0$, or $Y_v < 0$, for all variables Y_v

associated with predecessor branch vertices v of w along the branch leading from the root to w. The sign of each Y_v is determined by the outgoing edge in the branch vertex. In addition, each leaf carries a label "Yes" or "No". The tree $\mathcal{T}$ *tests membership* in the union of all Yes leaf sets.

The semantics of this model of computation is straightforward. On an input $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the input variables X_i get the corresponding values x_i, the arithmetic operations are executed in computation vertices v and the real values are obtained by variables Y_v. At branch vertices v the sign of the value of Y_v is determined, and the corresponding outgoing edge is chosen. As a result a certain branch ending up in a leaf w is selected. The input $\mathbf{x}$ belongs to the semialgebraic set assigned to w. If w is a Yes leaf, then $\mathbf{x}$ is said to be accepted by the tree $\mathcal{T}$.

It can be assumed without loss of generality [2], that there are no divisions used in a tree.

We will be interested in lower bounds on the *heights* of algebraic computation trees testing membership in a given semialgebraic set Σ. A detailed outline of the development of lower bounds that depend on topological characteristics of a set can be found in [7] (see also [2]). We mention here just two highlights.

The first most important achievement was the proof by Ben-Or [1] of the bound $c_1 \log(\mathrm{b}_0(\Sigma)) - c_2 n$, where $\mathrm{b}_0(\Sigma)$ is the number of connected components of Σ, and $c_1,\ c_2$ are some absolute positive constants. This general bound implies non-trivial, and sometimes tight, lower bounds for specific computational problems, such as *Distinctness* and *Knapsack*.

One of the most general results so far in this direction belongs to Yao [7]. Suppose a semialgebraic set Σ is locally closed and bounded. Let $\mathrm{b}^{BM}(\Sigma)$ be the total Betti number (the sum of all Betti numbers) of Σ with respect to the Borel-Moore homology $H_*^{BM}(\Sigma)$. Yao proved the lower bound

$$c_1 \log(\mathrm{b}^{BM}(\Sigma)) - c_2 n, \tag{1.1}$$

where $c_1,\ c_2$ are some absolute positive constants. From this he deduced a tight lower bound for *k-Distinctness* problem, and other non-trivial lower bounds for specific problems.

The use of Borel-Moore homology is a very strong condition, which implies subadditivity of the total Betti number. Subadditivity is the property on which the whole of the Yao's argument depends. It is natural to ask whether an analogous bound can be found for the usual, singular, homology theory, which is applicable to arbitrary (not necessarily locally closed) semialgebraic sets. Of course, in this case subadditivity is not necessarily true. Observe that for compact sets Borel-Moore Betti numbers coincide with singular Betti numbers, while for non-compact locally closed sets these two types of Betti numbers can be incomparable.

In this paper we prove two main theorems. Firstly, we prove the lower bound

$$\frac{c_1 \log(\mathrm{b}_m(\Sigma))}{m+1} - c_2 n,$$

where $\mathrm{b}_m(\Sigma)$ is the m-th Betti number of an arbitrary semialgebraic set Σ with respect to singular homology, and $c_1,\ c_2$ are some absolute positive constants. Note that this bound depends on an individual Betti number rather than on the sum of Betti numbers. For Betti numbers of a small (fixed) index m the bound turns into $c_1 \log(\mathrm{b}_m(\Sigma)) - c_2 n$ which is similar to Yao's bound. The proof is based on a construction from [5] which transforms Σ into a compact semialgebraic set Σ'

having the same Betti numbers as Σ up to a given index m. We then prove that for any algebraic computation tree $\mathcal{T}$ for Σ there is an algebraic computation tree $\mathcal{T}'$ for Σ' having the height not exceeding, up to a multiplicative constant, m times the height of $\mathcal{T}$. It remains to apply Yao's bound to Σ'.

Our second main result is a lower bound in terms of Betti numbers of the projection $\rho(\Sigma)$ of Σ to a subspace, rather than Betti numbers of Σ itself. Note that the topology of the image under a projection may be much more complex than the topology of the set being projected. We are not aware of previous lower bounds of this sort. The bound is

$$\frac{c_1 \log(\mathrm{b}_m(\rho(\Sigma)))}{(m+1)^2} - \frac{c_2 n}{m+1}$$

for some positive constants $c_1,\ c_2$, which again should be applied for small (fixed) values of m. The proof uses a spectral sequence associated with the projection map [6], which allows to bound from above Betti numbers of the projection of Σ in terms of Betti numbers of fiber products by itself of the compactification of Σ.

We illustrate these general results by examples of lower complexity bounds for some specific computational problems.

2. Topological tools

In this section we formulate the results from [4, 5, 6] which are used further in this paper.

In what follows, for a topological space X, let $\mathrm{b}_m(X) := \operatorname{rank} H_m(X)$ be its m-th Betti number with respect to the singular homology group $H_m(X)$ with coefficients in some fixed Abelian group. By $\mathrm{b}(X)$ we denote the *total* Betti number of X, i.e., the sum $\sum_{i\geq 0} \mathrm{b}_i(X)$.

2.1. Upper bounds on Betti numbers.

Consider a semialgebraic set $S = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathcal{F}(\mathbf{x})\}$, where $\mathcal{F}$ is a Boolean combination of polynomial equations and inequalities of the kind $h(\mathbf{x}) = 0$ or $h(\mathbf{x}) > 0$, and $h \in \mathbb{R}[x_1, \ldots, x_n]$. Suppose that the number of different polynomials h is s and that their degrees do not exceed d.

Proposition 2.1 ([5], Theorem 6.3)**.** *The m-th Betti number of S satisfies*

(1) $\mathrm{b}_m(S) = O(s^2 d)^n$*;*
(2) $\mathrm{b}_m(S) = O((m+1)sd)^n$*.*

Remark 2.2. Unlike classical Petrovski-Oleinik-Thom-Milnor bound for basic semialgebraic sets, used in [7], the bounds in Proposition 2.1 are applicable to arbitrary semialgebraic sets defined by a quantifier-free formulae. They are slightly weaker than the classical bound $O(sd)^n$ by a multiplicative factor at the base of the exponent, namely s in (1) and $m+1$ in (2).

Further, in the proof of Theorem 3.1, we will need the bound (1) from this proposition. We won't need bound (2) as such but we shall use in an essential way the constructions from [5] used for proving this bound. We now proceed to describing this technique.

2.2. Approximation by monotone families.

Definition 2.3. Let G be a compact semialgebraic set. Consider a semialgebraic family $\{S_\delta\}_{\delta>0}$ of compact subsets of G, such that for all $\delta, \delta' \in (0,1)$, if $\delta' > \delta$, then $S_{\delta'} \subset S_\delta$. Denote $S := \bigcup_{\delta>0} S_\delta$.

For each $\delta > 0$, let $\{S_{\delta,\varepsilon}\}$ be a semialgebraic family of compact subsets of G such that:

(i) for all $\varepsilon, \varepsilon' \in (0,1)$, if $\varepsilon' > \varepsilon$, then $S_{\delta,\varepsilon} \subset S_{\delta,\varepsilon'}$;
(ii) $S_\delta = \bigcap_{\varepsilon>0} S_{\delta,\varepsilon}$;
(iii) for all sufficiently small $\delta' > 0$ and for all $\varepsilon' > 0$, there exists an open in G set $U \subset G$ such that $S_\delta \subset U \subset S_{\delta',\varepsilon'}$.

We say that S is *represented* by the families $\{S_\delta\}$ and $\{S_{\delta,\varepsilon}\}$ in G.

Consider the following two particular cases.

Case 1. Let a semialgebraic set S be given as a disjoint union of *basic* semialgebraic sets (i.e., sets each defined by a conjunction of equations and strict inequalities). (Note that an algebraic computation tree represents the corresponding set in exactly this way.) Let δ and ε be some positive constants.

Suppose first that S is bounded in $\mathbb{R}^n$, and take as G a closed ball of a sufficiently large radius centered at 0. The set S_δ is the result of the replacement, independently in each basic set in the union, of all inequalities $h > 0$ and $h < 0$ by $h \geq \delta$ and $h \leq -\delta$ respectively. The set $S_{\delta,\varepsilon}$ is obtained by replacing, independently in each basic set, all expressions $h > 0$, $h < 0$ and $h = 0$ by $h \geq \delta$, $h \leq -\delta$ and $h^2 - \varepsilon \leq 0$, respectively. One can easily verify (see [5]) that the set S, is represented by families $\{S_\delta\}$ and $\{S_{\delta,\varepsilon}\}$ in G.

Now suppose that S is not necessarily bounded. In this case one can take the semialgebraic one-point (Alexandrov) compactification of $\mathbb{R}^n$ as G. Define sets S_δ and $S_{\delta,\varepsilon}$ as in the bounded case, replacing equations and inequalities independently in each basic set, and then taking the conjunction of the resulting formula with $|\mathbf{x}|^2 \leq 1/\delta$. Again, S is represented by $\{S_\delta\}$ and $\{S_{\delta,\varepsilon}\}$ in G.

Case 2. Let $\rho : \mathbb{R}^{n+r} \to \mathbb{R}^n$ be the projection map, and $S \subset \mathbb{R}^{n+r}$ be a semialgebraic set, given as a disjoint union of *basic* semialgebraic sets. The set S is represented by families $\{S_\delta\}$, $\{S_{\delta,\varepsilon}\}$ in the compactification of $\mathbb{R}^{n+r}$ as described in **Case 1**. One can easily verify (see [5]), that the projection $\rho(S)$ is represented by families $\{\rho(S_\delta)\}$, $\{\rho(S_{\delta,\varepsilon})\}$ (in the Alexandrov compactification of $\mathbb{R}^n$ if necessary).

Returning to the general case, suppose that a semialgebraic set S is *represented* by families $\{S_\delta\}$ and $\{S_{\delta,\varepsilon}\}$ in G.

For a sequence $\varepsilon_0, \delta_0, \varepsilon_1, \delta_1, \ldots, \varepsilon_m, \delta_m$, where $m \geq 0$, introduce the compact set

$$T_m(S) := S_{\delta_0,\varepsilon_0} \cup S_{\delta_1,\varepsilon_1} \cup \cdots \cup S_{\delta_m,\varepsilon_m}.$$

Observe that in **Case 2**, we have the equality

$$T_m(\rho(S)) = \rho(T_m(S)). \tag{2.1}$$

In what follows, for two real numbers a and b we write $a \ll b$ to mean "a is sufficiently smaller than b" (see formal Definition 1.7 in [5]).

Proposition 2.4 ([5], Theorem 1.5)**.** *For any $m \geq 0$, and*

$$0 < \varepsilon_0 \ll \delta_0 \ll \varepsilon_1 \ll \delta_1 \ll \cdots \ll \varepsilon_m \ll \delta_m \ll 1$$

we have

(i) *for every $0 \leq k \leq m$, there is an epimorphism $\varphi_k : H_k(T_m(S)) \to H_k(S)$, in particular, $\mathrm{b}_k(S) \leq \mathrm{b}_k(T_m(S))$;*

(ii) *in* **Case 1**, *for every* $k \leq m-1$, *the epimorphism* φ_k *is an isomorphism, in particular,* $\mathrm{b}_k(S) = \mathrm{b}_k(T_m(S))$. *Moreover, if* $m \geq \dim(S)$, *then* $T_m(S)$ *is homotopy equivalent to* S.

2.3. Betti numbers of projections.

Definition 2.5. For two maps $f_1 : X_1 \to Y$ and $f_2 : X_2 \to Y$, the *fibered product* of X_1 and X_2 is defined as

$$X_1 \times_Y X_2 := \{(\mathbf{x}_1, \mathbf{x}_2) \in X_1 \times X_2 |\ f_1(\mathbf{x}_1) = f_2(\mathbf{x}_2)\}.$$

Proposition 2.6 ([6], Theorem 1)**.** *Let* $f : X \to Y$ *be a closed surjective semialgebraic map (in particular,* f *can be the projection map to a subspace, with a compact* X*). Then*

$$\mathrm{b}_m(Y) \leq \sum_{p+q=m} \mathrm{b}_q(W_p),$$

where

$$W_p := \underbrace{X \times_Y \cdots \times_Y X}_{(p+1) \text{ times}}.$$

3. General lower bounds

We start with a theorem which immediately follows from an upper bound on the total Betti number of an arbitrary semialgebraic set in Proposition 2.1.

Theorem 3.1. *Let* k *be the height of an algebraic computation tree* $\mathcal{T}$ *testing membership in a semi-algebraic set* $\Sigma \subset \mathbb{R}^n$. *Then*

$$k = \Omega\left(\frac{\log(\mathrm{b}(\Sigma))}{n}\right),$$

where $\mathrm{b}(\Sigma)$ *is the total Betti number of* Σ.

Proof. Since in each computation vertex at most one multiplication can be performed, every polynomial occurring in the disjunctive normal form defining Σ has a degree at most 2^k. The number of polynomials in the conjunction defining the set attached to a Yes leaf is at most k, while the number of Yes leaves does not exceed 3^k. It follows that the total number of polynomials defining Σ is at most $k3^k$. Then, according to Proposition 2.1, (1), $\mathrm{b}(\Sigma) \leq O((k3^k)^2 2^k)^n$. Taking logarithms we get the result. □

Remark 3.2. The bound in the theorem is significantly weaker than Yao's bound (1.1). However, as explained in the introduction, it is applicable to any semialgebraic set, not necessarily a locally closed one. The upper bound on the total Betti number, used in the proof, is applicable to arbitrary semialgebraic set, unlike classical Petrovski-Oleinik-Thom-Milnor bounds employed in [7].

Theorem 3.3. *Let* k *be the height of an algebraic computation tree* $\mathcal{T}$ *testing membership in a semi-algebraic set* $\Sigma \subset \mathbb{R}^n$. *Then*

$$k \geq \frac{c_1 \log(\mathrm{b}_m(\Sigma))}{m+1} - c_2 n,$$

where $\mathrm{b}_m(\Sigma)$ *is the* m*-th Betti number of* Σ, *and* $c_1,\ c_2$ *are some positive constants.*

Lemma 3.4. *Let $\mathcal{T}_1, \mathcal{T}_2$ be algebraic computation trees testing membership in semi-algebraic sets Σ_1 and Σ_2 respectively, and having heights k_1 and k_2 respectively. Then there is a tree $\mathcal{T}_\cup$ testing membership in $\Sigma_1 \cup \Sigma_2$, and a tree $\mathcal{T}_\cap$ testing membership in $\Sigma_1 \cap \Sigma_2$, both having heights at most $O(k_1 + k_2)$.*

Proof. To construct $\mathcal{T}_\cup$, attach a copy of $\mathcal{T}_2$ to each No leaf of the tree $\mathcal{T}_1$. For $\mathcal{T}_\cap$, attach a copy of $\mathcal{T}_2$ to each Yes leaf of the tree $\mathcal{T}_1$. □

Lemma 3.5. *Let $\mathcal{T}$ be a tree for Σ, having height k. Then for any $\ell \geq 0$ there exists a tree $\mathcal{T}_\ell$ for $T_\ell(\Sigma)$ whose height does not exceed $k' \leq c((\ell+1)k+n)$ for some positive constant c.*

Proof. The plan of the proof is as follows. The construction of $\mathcal{T}_\ell$ consists of two stages. On the first stage we perform the construction for $\ell = 0$ and arbitrary ε, δ, and get the tree $\mathcal{T}_{\varepsilon,\delta}$. The height of $\mathcal{T}_{\varepsilon,\delta}$ is not larger than c times the height of $\mathcal{T}$ for a constant $c > 0$. On the second stage we construct $\mathcal{T}_\ell$ for an arbitrary ℓ by induction. On the base step, construct the tree $\mathcal{T}_0 = \mathcal{T}_{\varepsilon_0,\delta_0}$. Suppose we constructed the tree $\mathcal{T}_{\ell-1}$. The tree $\mathcal{T}_\ell$ is obtained from $\mathcal{T}_{\ell-1}$ by attaching to *each* No leaf of the latter, a copy of the tree $\mathcal{T}_{\varepsilon_\ell,\delta_\ell}$, considering the leaf as the root of $\mathcal{T}_{\varepsilon_\ell,\delta_\ell}$.

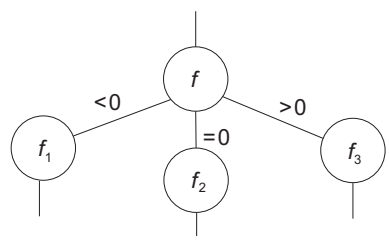

FIGURE 1. Tree $\mathcal{S}$.

Now we proceed to a more detailed proof.

Let $h := X_1^2 + \cdots + X_n^2$. The root r of the tree $\mathcal{T}_{\varepsilon,\delta}$ is a branch vertex with the polynomial $h - 1/\delta$ attached. The child of r, corresponding to > 0, is a No leaf. Take the other two children as roots of two copies of the tree $\mathcal{T}$. The construction of $\mathcal{T}_{\varepsilon,\delta}$ now continues identically for both copies, by induction, as follows. In $\mathcal{T}$, let v be the closest branch vertex to its root, and let f be the polynomial attached to v (if such branch vertex does not exist, then the construction of $\mathcal{T}_{\varepsilon,\delta}$ is completed). Then the neighbourhood of v in $\mathcal{T}$ looks like the tree $\mathcal{S}$ on Figure 1. Here f_1, f_2 and f_3 are polynomials attached to children v_1, v_2, v_3 of v. Replace this neighbourhood by the tree $\mathcal{S}'$ on Figure 2. Notice that the leaves of $\mathcal{S}'$, are labelled again by f_1, f_2, f_3 while one of the leaves is a No leaf. Attach to each leaf of $\mathcal{S}'$, labelled by f_i, the subtree of $\mathcal{T}$ rooted at v_i (unless v_i is a leaf of $\mathcal{T}$). Denote the resulting tree by $\mathcal{T}'$. This completes the base of the induction.

On the next induction step perform the same replacement operation, as on the base step, for each subtree of $\mathcal{T}'$ rooted at a leaf of $\mathcal{S}'$ which is not a No leaf. If for a leaf of $\mathcal{S}'$ no such subtree exists, i.e., vertex v_i is a leaf of $\mathcal{T}$, then this vertex is taken as a leaf of $\mathcal{T}_{\varepsilon,\delta}$, it is a Yes leaf if and only if v_i is a Yes leaf in $\mathcal{T}$. Denote the results of replacements again by $\mathcal{S}'$, and the resulting tree again by $\mathcal{T}'$.

Further induction steps are performed in the same fashion, by applying the replacement operation, described at the base step, to subtrees of $\mathcal{T}'$ rooted at

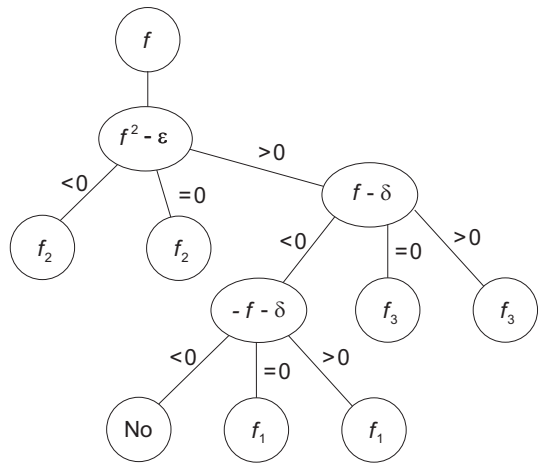

FIGURE 2. Tree $\mathcal{S}'$.

leaves of the trees $\mathcal{S}'$ obtained on the previous induction step. The construction of $\mathcal{T}_{\varepsilon,\delta}$ is completed when all leaves of trees $\mathcal{S}'$ become leaves of $\mathcal{T}_{\varepsilon,\delta}$.

Note that the height of $\mathcal{T}_{\varepsilon,\delta}$ is not larger than c times the height of $\mathcal{T}$ for a constant $c > 0$.

Now we prove by induction on the construction that $\mathcal{T}_{\varepsilon,\delta}$ is a tree testing membership in $\Sigma_{\varepsilon,\delta}$ (recall the notation from Case 1, Section 2.2). Observe that either $h - 1/\delta = 0$ or $h - 1/\delta < 0$ is present in the definition of any Yes leaf set. Assume, as before, that in $\mathcal{T}$ the vertex v is the closest branch vertex to the root, and f is the polynomial attached to v. Observe that each leaf set of $\mathcal{T}$, in particular each Yes leaf set, is of the kind either $\{f = 0, \ldots\}$, or $\{f > 0, \ldots\}$, or $\{f < 0, \ldots\}$. In the tree $\mathcal{T}'$, on the base step of the construction of $\mathcal{T}_{\varepsilon,\delta}$, the leaf $\{f = 0, \ldots\}$ will be replaced by two leaves, $\{f^2 - \varepsilon = 0, \ldots\}$ and $\{f^2 - \varepsilon < 0, \ldots\}$, the leaf $\{f > 0, \ldots\}$ – by two leaves, $\{f - \delta = 0, \ldots\}$ and $\{f - \delta > 0, \ldots\}$, while the leaf $\{f < 0, \ldots\}$ – by two leaves, $\{-f - \delta = 0, \ldots\}$ and $\{-f - \delta > 0, \ldots\}$. It follows that $\{f = 0, \ldots\} \subset \Sigma$ if and only if $\{f^2 - \varepsilon \leq 0, \ldots\}$ is a subset of the set tested by $\mathcal{T}'$, and similar for sets $\{f > 0, \ldots\}$ and $\{f < 0, \ldots\}$. Proceeding by induction, we conclude that $\Sigma_{\varepsilon,\delta}$ is the set tested by $\mathcal{T}_{\varepsilon,\delta}$.

Now construct $\mathcal{T}_\ell$ for arbitrary ℓ by induction. On the base step, start with the path of $O(n)$ computation vertices at the end of which the polynomial $h = X_1^2 + \cdots + X_n^2$ is computed. Continue with the tree $\mathcal{T}_{\varepsilon_0,\delta_0}$. The result of these two steps is the tree $\mathcal{T}_0$. Suppose we constructed the tree $\mathcal{T}_{\ell-1}$ for $\ell \geq 1$. The tree $\mathcal{T}_\ell$ is obtained from $\mathcal{T}_{\ell-1}$ by attaching to *each* No leaf of the latter, the tree $\mathcal{T}_{\varepsilon_\ell,\delta_\ell}$, considering the leaf as the root of $\mathcal{T}_{\varepsilon_\ell,\delta_\ell}$. By Lemma 3.4, the result is indeed $\mathcal{T}_\ell$.

Obviously the height of $\mathcal{T}_\ell$ does not exceed $k' \leq c((\ell+1)k + n)$ for a constant $c > 0$. □

Proof of Theorem 3.3. By Lemma 3.5, $k' \leq c((m+1)k + n)$ for a constant $c > 0$. By (1.1), since $T_m(\Sigma)$ is compact,

$$c((m+1)k + n) \geq c_1 \log(\mathrm{b}((T_m(\Sigma)) - c_2 n$$

for some positive constants $c_1,\ c_2$. Hence, for the m-th Betti number,

$$c((m+1)k+n) \geq c_1 \log(\mathrm{b}_m((T_m(\Sigma) - c_2 n.$$

It follows, by Proposition 2.4, that

$$c((m+1)k+n) \geq c_1 \log(\mathrm{b}_m(\Sigma)) - c_2 n.$$

Hence, the theorem. □

4. Projections

Theorem 4.1. *Let k be the height of an algebraic computation tree $\mathcal{T}$ testing membership in a semi-algebraic set $\Sigma \subset \mathbb{R}^n$. Let $\rho:\ \mathbb{R}^n \to \mathbb{R}^{n-r}$ be the projection map. Then*

$$k \geq \frac{c_1 \log(\mathrm{b}_m(\rho(\Sigma)))}{(m+1)^2} - \frac{c_2 n}{m+1} \tag{4.1}$$

for some positive constants $c_1,\ c_2$.

Let

$$W_p := \underbrace{T_m(\Sigma) \times_{\rho(T_m(\Sigma))} \cdots \times_{\rho(T_m(\Sigma))} T_m(\Sigma)}_{\text{(p+1) times}}.$$

Lemma 4.2. *Let $\mathcal{T}$ be a tree for Σ, having height k. Then there exists a tree $\mathcal{T}_m^W$ for W_p whose height does not exceed $c(p+1)((m+1)k+n)$ for some positive constant c.*

Proof. Lemma 3.5 implies that there is a tree $\mathcal{T}_m$ for $T_m(\Sigma)$ having the height not exceeding $c((m+1)k+n)$. The problem of membership in W_p has input variables

$$X_1, \ldots, X_{n-r}, Y_{1,n-r+1}, \ldots, Y_{1,n}, \ldots, Y_{p,n-r+1}, \ldots, Y_{p,n}.$$

Construct the tree $\mathcal{T}_m^W$ inductively, starting with a copy of $\mathcal{T}_m$ with input variables $X_1, \ldots, X_{n-r}, Y_{1,n-r+1}, \ldots, Y_{1,n}$. Then, using Lemma 3.4, attach to each Yes leaf another copy of $\mathcal{T}_m$ with input variables $X_1, \ldots, X_{n-r}, Y_{2,n-r+1}, \ldots, Y_{2,n}$, and so on. The height of the resulting tree $\mathcal{T}_m^W$ is at most $p+1$ times the height of the tree $\mathcal{T}_m$, as required. □

Proof of Theorem 4.1. According to Proposition 2.4,

$$\mathrm{b}_m(\rho(T_m(\Sigma))) \leq \sum_{p+q=m} \mathrm{b}_q(W_p). \tag{4.2}$$

Let $\mathrm{b}(W_\nu) := \max_{0 \leq p \leq m} \mathrm{b}(W_p)$, and k' be the height of a tree for W_ν.

Since W_ν is compact, by (1.1), we have

$$k' \geq c_1' \log(\mathrm{b}(W_\nu)) - c_2'(n + \nu r)$$

for some positive constants $c_1',\ c_2'$, thus, replacing k' by a larger number according to Lemma 4.2, and using $m \geq \nu$, we get

$$(m+1)^2 k + (m+1)n \geq c_1'' \log(\mathrm{b}(W_\nu)) - c_2''(n + (m+1)r)$$

for some positive constants $c_1'',\ c_2''$. But

$$\sum_{p+q=m} \mathrm{b}_q(W_p) \leq m\ \mathrm{b}(W_\nu),$$

so using (4.2) we have

$$(m+1)^2k + (m+1)n \geq c_1''(\log(\mathrm{b}_m(\rho(T_m(\Sigma))) - \log m) - c_2''(n + (m+1)r).$$

Hence,

$$k \geq \frac{c_1 \log(\mathrm{b}_m(\rho(T_m(\Sigma)))}{(m+1)^2} - \frac{c_2 n}{m+1}$$

for some positive constants $c_1,\ c_2$.

According to (2.1), $\rho(T_m(\Sigma)) = T_m(\rho(\Sigma))$, while, by Proposition 2.4,

$$\mathrm{b}_m(T_m(\rho(\Sigma))) \geq \mathrm{b}_m(\rho(\Sigma)).$$

It follows that

$$k \geq \frac{c_1 \log(\mathrm{b}_m(\rho(\Sigma)))}{(m+1)^2} - \frac{c_2 n}{m+1}.$$

□

5. Applications

In this section we apply the general bounds from Theorems 3.3 and 4.1 to examples of specific computational problems. These problems admit obvious variations.

"Parity of integers". This is the following computational problem.

Let m be a positive integer. For given n real numbers $x_1, \ldots, x_n$ such that $1 \leq x_i \leq m$ for all i, decide whether the following property is true: either all x_i are integer or exactly two of them are not integer.

Observe that complexity of this problem has an *upper bound* $O(n \log m)$: the computation tree for each x_i checks whether it coincides with one of the numbers $1, \ldots, m$ using binary search.

To obtain a lower bound, consider the integer lattice $\{1, \ldots, m\}^n$ in $\mathbb{R}^n$ and let Σ be the union of all open 2-dimensional squares and all vertices. Then the problem is equivalent to deciding membership in Σ. Observe that Σ is not locally closed. It is homotopy equivalent to a 2-plane with $\Omega(m^n)$ punctured points, so $b_1(\Sigma) = \Omega(m^n)$. By Theorem 3.3, the height of any algebraic computation tree testing membership in Σ is $\Omega(n \log m)$.

"Crossing number". Let Σ be a smooth connected bounded semialgebraic curve in $\mathbb{R}^n$. Then Σ is a (smooth) embedding of either the circle S^1 or the interval $(0, 1)$ into $\mathbb{R}^n$ (for $n = 3$ and a circle this is a knot). The total Betti number of Σ is at most 2.

Observe that the image under the projection of Σ onto a generic 2-dimensional linear subspace has only double points as singular points.

The *crossing number* $C(\Sigma)$ of Σ is the maximal number of singular points of the image of the projection over all generic 2-dimensional linear subspaces.

Theorem 5.1. *The complexity of membership in Σ is at least $c_1 \log C(\Sigma) - c_2 n$ for some positive constants $c_1,\ c_2$.*

Proof. Let $\rho(\Sigma)$ be the image of Σ under the projection to the plane on which the crossing number is realized. Then $C(\Sigma)$ is less by 2 (if Σ is an embedding of S^1), or otherwise by 1, than the number of connected components of the complement to $\rho(\Sigma)$ in the plane. By Alexander duality, the number of connected components is the same as $\mathrm{b}_1(\rho(\Sigma))$, hence the lower bound follows from Theorem 4.1. □

References

[1] M. Ben-Or, Lower Bounds For Algebraic Computation Trees, *Proceedings of 15th Annual ACM Symposium on Theory of Computing*, (1983), 80-86.

[2] P. Bürgisser, F. Cucker, Variations by complexity theorists on three themes of Euler, Bézout, Betti, and Poincaré, in: Complexity of Computations and Proofs, Jan Krajicek (ed.), *Quaderini di Matematica*, 13 (2005), 73-152.

[3] D. Dobkin and R.J. Lipton, On the complexity of computations under varying sets of primitives, *J. Comput. Syst. Sci.*, 18 (1979), 86-91.

[4] A. Gabrielov, N. Vorobjov, Betti numbers of semialgebraic sets defined by quantifier-free formulae, *Discrete Comput. Geom.*, 33 (2005), 395-401.

[5] A. Gabrielov, N. Vorobjov, Approximation of definable sets by compact families, and upper bounds on homotopy and homology, *J. London Math. Soc.*, 80 (2009), 35-54.

[6] A. Gabrielv, N. Vorobjov, T. Zell, Betti numbers of semialgebraic and sub-Pfaffian sets, *J. London Math. Soc.*, 69, part 1 (2004), 27-43.

[7] A.C.C. Yao, Decision tree complexity and Betti numbers, *J. Comput. Syst. Sci.*, 55 (1997), 36-43.

Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA
E-mail address: `agabriel@math.purdue.edu`

Department of Computer Science, University of Bath, Bath BA2 7AY, England, UK
E-mail address: `nnv@cs.bath.ac.uk`