Complexity Lower Bounds for Computation Trees with Elementary Transcendental Function Gates

(Extended abstract)

Dima Grigoriev* Nicolai Vorobjov
Departments of Computer Science and Mathematics
Penn State University
University Park, PA 16802
dima@cse.psu.edu vorobjov@cse.psu.edu

Abstract

We consider computation trees which admit as gate functions along with the usual arithmetic operations also algebraic or transcendental functions like exp, log, sin, *square root (defined in the relevant domains) or much more general, Pfaffian functions. A new method for proving lower bounds on the depth of these trees is developed which allows to prove a lower bound* $\Omega(\sqrt{\log N})$ *for testing membership to a convex polyhedron with* N *facets of all dimensions, provided that* N *is large enough. This method differs essentially from the approaches adopted for algebraic computation trees ([1], [4], [26], [13]).*

1 Pfaffian computation trees

We consider the following computation model, a generalization of the algebraic computation trees (see, e.g., [1], [26]).

Definition 1. *Pfaffian computation tree* $\mathcal{T}$ *is a tree at every node* v *of which a Pfaffian function* f_v *in variables* $X_1, \dots, X_n$ *is attached, which satisfies the following properties. Let* $f_{v_0}, \dots, f_{v_l}, f_{v_{l+1}} = f_v$ *be the functions attached to all the nodes along the branch* $\mathcal{T}_v$ *of* $\mathcal{T}$ *leading from the root* v_0 *to* $v_{l+1} = v$. *We assume that the Pfaffian function* f_v *satisfies the following differential equation (see [20]):*

$$df_v = \sum_{1 \le j \le n} g_{v,j}(X_1, \dots, X_n, f_{v_0}, \dots, f_{v_l}, f_v) dX_j,$$

where $g_{v,j}$ *are polynomials with real coefficients.*

The tree $\mathcal{T}$ *branches at* v *to its three sons according to the sign of* f_v *(cf. [1]). Thereby, to each node* v *one can assign (by induction on the depth* $l+1$ *of* v*) a set* $U_v \subset \mathbf{R}^n$ *consisting of all the points for which the sign conditions for the functions* $f_{v_0}, \dots, f_{v_l}$ *along the branch* $\mathcal{T}_v$ *are valid. Thus, at the induction step, one assigns to three sons of* v *the sets*

$$U_v \cap \{f_v > 0\},\ U_v \cap \{f_v = 0\},\ U_v \cap \{f_v < 0\},$$

respectively. We assume also that the function f_v *is real analytic in* U_v. *To each leaf* w *of* $\mathcal{T}$ *an output "yes" or "no" is assigned, we call the set* U_w *accepting set if to* w *"yes" is assigned. We say that* $\mathcal{T}$ *tests the membership problem to the union of all accepting sets (sf. [1]).*

Taking polynomials as the gate functions f_v in $\mathcal{T}$, we come to the algebraic computation trees. The examples of other gate Pfaffian functions f_v for $0 \le q \le l$ are:
(1) $\exp(f_{v_q})$;
(2) $1/f_{v_q}$, defined for $f_{v_q} \neq 0$;
(3) $\log(f_{v_q})$, with log defined on the positive half-line;
(4) $\sin(f_{v_q})$, with sin defined on an interval $(-\pi + 2\pi r, \pi + 2\pi r)$;
(5) $\tan(f_{v_q})$, with tan defined on $(-\pi/2 + \pi r, \pi/2 + \pi r)$;
(6) $\sqrt{f_{v_q}}$, with square root defined on the positive half-line.

We suppose that the degrees $\deg(g_{v,j})$ of the polynomials occuring in the definition of the gate functions f_v in $\mathcal{T}$, are less than d.

Now let us formulate the main result.

Theorem. *Let a Pfaffian computation tree* $\mathcal{T}$ *test*

*Supported in part by Volkswagen-Stiftung

the membership problem to a closed convex polyhedron $P \subset \mathbf{R}^n$, having N facets of all dimensions from zero to n. Then the depth k of $\mathcal{T}$ is greater than $\Omega(\sqrt{\log N})$, provided that $N \geq (nd)^{\Omega(n^4 \log d)}$.

The complete proof of the theorem one can find in [16]. Here we outline some of its ideas.

A special case of the theorem, when $n = 2$, so P is a polygon, was proved in [12].

Several methods, based on topological characteristics, are known for obtaining complexity lower bounds for algebraic computation trees testing membership to a *semialgebraic* set $S \subset \mathbf{R}^n$. In [1] the bound $\Omega(\log C)$ was proved, where C is the number of connected components of S or of its complement, in [2], [4], [25] the bound $\Omega(\log \chi)$ for Euler characteristic χ of S was obtained. A stronger lower bound $\Omega(\log B)$ was proved in [2], [3], [26], where B is the sum of Betti numbers of S. Actually, one can directly extend these results to Pfaffian computation trees, replacing in the proofs the references to Milnor's bound [22] for B by the references to Khovanskii's bound [19], [20] for the sum of Betti numbers of a semi-Pfaffian set.

This leads to the following proposition [12].

Proposition. *If a Pfaffian computation tree tests the membership problem to a semi-Pfaffian set with the sum of Betti numbers $\mathcal{B}$, then the depth of the tree is greater than $\Omega(\sqrt{\log \mathcal{B}})$.*

There is a conjecture that the bound in [20] can be improved. If this were true, it would lead to the bound $\Omega(\log N)$ in the theorem and $\Omega(\log \mathcal{B})$ in the proposition.

Observe that because the sum of Betti numbers of a convex polyhedron is 1, the theorem does not, apparently, follow from the proposition. In [13] the bound $\Omega(\log N)$ was proved for testing membership to a polyhedron with N facets by an algebraic decision tree (for N large enough, cf. the theorem). In [27] a similar lower bound was shown for a weaker model of linear decision trees. One cannot directly extend the method from [13] to Pfaffian computation trees since in [13] the effective quantifier elimination procedure for the first-order theory of reals (see [15], [10], [17], [23]) was essentially used, while for the theories involving Pfaffian functions (for instance, exp), the quantifier elimination does not exist [7], [8].

The computations involving other functions, besides arithmetic, were considered in several papers. In [18] for the circuits involving root extractions a complexity lower bound for computing an algebraic function was obtained, in [14] this result was extended to the circuits involving exp and log. In [11] lower bounds on parallel complexity for Pfaffian sigmoids were obtained.

Let us mention that for testing membership to a polyhedron an *upper* complexity bound $O(\log N)n^{O(1)}$ was shown in [21] for linear decision trees.

2 Nonstandard fields and angle points

Fix an accepting set $U_w \subset P$ and let $v_0, v_1, \dots, v_k = w$ be all the nodes of the branch in $\mathcal{T}$ leading from the root to w. Then $U_w = \{f_{v_0}\sigma_0 0, f_{v_1}\sigma_1 0, \dots, f_{v_k}\sigma_k 0\}$ for suitable signs $\sigma_0, \dots, \sigma_k \in \{<, =, >\}$. Rename the functions $\pm f_{v_0}, \dots, \pm f_{v_k}$ by $u_0, \dots, u_k$ in such a way that $U_w = \{u_0 = \dots = u_{k_1} = 0, u_{k_1+1} > 0, \dots, u_k > 0\}$. Denote $f = u_0^2 + \dots + u_{k_1}^2$.

Because for each i-dimensional facet P_i of P there exists an accepting set U_{w_1} such that $\dim(U_{w_1} \cap P_i) = i$, for proving the theorem it is sufficient to bound from above the number ν_i of i-dimensional facets P_i for which $\dim(U_w \cap P_i) = i$.

Estimation of ν_i uses essentially the notion of i-angle points. In order to define this notion, we have to invoke nonstandard extensions of reals, which we now briefly describe following [24]. The details could be found in [6]. The nonstandard extensions were used in [24], their algebraic version was essentially involved in [15], [10], [17], [23] for effective solving systems of inequalities and deciding Tarski algebra.

There exists a sequence of ordered fields

$$\mathbf{R}_0 = \mathbf{R} \subset \mathbf{R}_1 \subset \mathbf{R}_2 \subset \cdots \subset \mathbf{R}_j \subset \cdots$$

in which the field $\mathbf{R}_j$, $j \geq 1$ contains an element $\varepsilon_j > 0$ infinitesimal relative to the elements of $\mathbf{R}_{j-1}$ (i.e., for every positive element $a \in \mathbf{R}_{j-1}$ the inequality $\varepsilon_j < a$ is true). In addition, for every function

$$\varphi : \mathbf{R}_{j-1}^n \longrightarrow \mathbf{R}_{j-1}$$

there exists a natural extension of φ which is a function from $\mathbf{R}_j^n$ to $\mathbf{R}_j$. We say that $\mathbf{R}_i$ is a nonstandard extension of $\mathbf{R}_j$ for $0 \leq j < i$.

Consider the language $\mathcal{L}_j$, $j \geq 0$ of the first order predicate calculus, in which the set of all function symbols is in a bijective correspondence with the set of all functions of several arguments from $\mathbf{R}_j$ taking values in $\mathbf{R}_j$, and the only predicate is the equality relation. We shall say that a closed (i.e., containing no free variables) formula Φ of the language $\mathcal{L}_j$ *is true in* $\mathbf{R}_j$, $j \geq 0$, if and only if the statement expressed by this formula with respect to $\mathbf{R}_j$ is true. The following "transfer principle" is valid: for all integers $0 \le j < i$ the closed formula Φ of $\mathcal{L}_j$ is true in $\mathbf{R}_j$ if and only if it is true in $\mathbf{R}_i$.

An element $z \in \mathbf{R}_i,\ i \ge 0$ is called *infinitesimal relative to* $\mathbf{R}_j,\ 0 \le j < i$ if for every $0 < w \in \mathbf{R}_j$ the inequality $|z| < w$ is valid. An element $z \in \mathbf{R}_i$ is called infinitely large, if $z = 1/z_0$, where z_0 is infinitesimal. If $z \in \mathbf{R}_i$ not infinitely large relative to $\mathbf{R}_j$, z is called $\mathbf{R}_j$-finite.

One can prove [6], that if an element $z \in \mathbf{R}_i$ is $\mathbf{R}_j$-finite then there exist unique elements $z_1 \in \mathbf{R}_j$ and $z_2 \in \mathbf{R}_i$, where z_2 is infinitesimal relative to $\mathbf{R}_j$, such that $z = z_1 + z_2$. In this case z_1 is called the *standard part* of z (relative to $\mathbf{R}_j$) and is denoted by $z_1 = \mathrm{st}_j(z)$. One can extend the operation st_j (componentwise) to vectors from $\mathbf{R}_i^n$ and (elementwise) to subsets of $\mathbf{R}_i^n$.

Denote $m = n^3 - n^2 + n$ and standard part st_m we denote for brevity by st.

Definition 2. *A point $x \in U_w \subset \mathbf{R}^n_{m+2}$ is called 0-quasiangle if the inequalities*

$$u_{k_1+1}(x) \ge \varepsilon_1, \dots, u_k(x) \ge \varepsilon_1$$

are valid, there exist the points $y_1, \dots, y_n \in \{f = \varepsilon_{m+2}\}$ such that the distances $\|y_i - x\| \le \varepsilon_{m+1},\ 1 \le i \le n$ and

$$\det \left| \frac{\mathrm{grad}_{y_1}(f)}{\|\mathrm{grad}_{y_1}(f)\|}, \dots, \frac{\mathrm{grad}_{y_n}(f)}{\|\mathrm{grad}_{y_n}(f)\|} \right| > \varepsilon_1,$$

where $\mathrm{grad}_y(f) = \left(\frac{\partial f}{\partial X_1}, \dots, \frac{\partial f}{\partial X_n}\right)^T (y)$.

One can prove with the help of the transfer principle that ε_{m+2} is not a critical value of f, hence $\mathrm{grad}_{y_j}(f) \ne 0,\ 1 \le j \le n$.

A *semi-Pfaffian set* (see, e.g., [8], [9]) is defined, roughly speaking, as a set of points in $\mathbf{R}^n$ satisfying a Boolean formula with the atomic subformulas of the form $(g > 0)$, where g is a Pfaffian function.

A *sub-Pfaffian set* is defined as a set of points in $\mathbf{R}^n$ satisfying a formula (called *Pfaffian formula over* $\mathbf{R}$) with quantifiers $\forall,\ \exists$ restricted to bounded intervals, where the quantifier-free part is a Boolean formula with the atomic subformulas of the form $(g > 0)$.

A sub-Pfaffian set has a finite number of connected components each being, in its turn, a sub-Pfaffian set [8]. On the other hand, a theorem of Gabrielov [7], [8] states that each sub-Pfaffian set can be represented by a formula with solely existential quantifiers (unfortunately the bounds for resulting formula are not efficient).

In [19], [20] an explicit efficient bound is proved on the number of all connected components of a semi-Pfaffian set. This obviously implies the same bound on the number of the connected components for a sub-Pfaffian set given by a formula with solely existential quantifiers, because a projection of a connected set is connected.

One can extend the definitions of semi-Pfaffian and sub-Pfaffian sets to nonstandard fields. If a sub-Pfaffian set is defined by a Pfaffian formula over a field $\mathbf{R}_j$, then the same formula defines a sub-Pfaffian set over a field $\mathbf{R}_i,\ i \ge j$; we call the latter set the *completion* of the former one, and use for it the same notation.

The bound for the number of all connected components of a sub-Pfaffian set (in particular, finiteness of this number) holds also over nonstandard fields due to the transfer principle and the theorem of Gabrielov.

Note that U_w is semi-Pfaffian, while the set of all 0-quasiangle points is sub-Pfaffian.

In [5] (see also [13])it was proved that for each $1 \le j \le n$ there exists a family $\mathcal{A}_j$ consisting of $j(n-j)+1$ j-dimensional subspaces in $\mathbf{R}^n$ such that for any $(n-j)$-dimensional subspace $Q \subset \mathbf{R}^n$ there is a certain element $R \in \mathcal{A}_j$ for which $(R \cap Q) = \{0\}$.

Definition 3. *A point $x \in U_w$ is called i-quasiangle $(0 \le i < n)$ if for each subspace $\Pi \in \mathcal{A}_{n-i}$ the point x is a 0-quasiangle point in the semi-Pfaffian set $U_w \cap \Pi(x)$ where $\Pi(x)$ is the $(n-i)$-dimensional plane parallel to Π and passing through x (here we apply the Definition 2 of 0-quasiangle points to the restriction of f on $\Pi(x)$).*

The set of all i-quasiangle points we denote by $\tilde{A}_i \subset U_w$. Observe that $\tilde{A}_i$ is sub-Pfaffian.

Definition 4. *The points of the set $A_i = st(\tilde{A}_i) \subset \mathbf{R}^n_m$ are called i-angle.*

Lemma 1. *The set A_i is sub-Pfaffian and $A_i \subset U_w$.*

The next lemma shows, informally speaking, that if for i-dimensional facet P_i of P we have $\dim(U_w \cap P_i) = i$ (recall that our purpose is to estimate the number ν_i of such P_i), then $U_w \cap P_i$ lies both in $\tilde{A}_i$ and in A_i.

Lemma 2. *Assume that $\dim(U_w \cap P_i) = i$. If for two points $\tilde{x} \in U_w \cap P_i \cap \mathbf{R}^n,\ x \in P_i \cap \mathbf{R}^n_m$ the distance $\|\tilde{x} - x\|$ is infinitesimal relative to $\mathbf{R}$ then $x \in \tilde{A}_i$ and $x \in A_i$.*

Lemma 3. $\dim(A_i) \le i$.

Lemmas 2 and 3 allow to reduce the estimating of ν_i to the problem of estimating the number of i-dimensional facets P_i of P which have full i-dimensional intersections with at most i-dimensional set A_i. This problem is treated in the next section.

3 Flat points

Definition 5. *Let $0 \le i \le n-1$. A point $x \in A_i$ is called i-flat if there exists an i-plane Π, passing through x such that $\dim(\Pi \cap A_i) = i$.*

Denote by $\Phi_i \subset A_i$ the set of all i-flat points. Note that for $i = 0$ Lemma 3 entails that $\dim(A_0) \le 0$, i.e. A_0 consists of at most a finite number of points, therefore $\Phi_0 = A_0$.

Lemma 4. *There is at most a finite number of i-planes Π for which $\dim(\Pi \cap \Phi_i) = i$, and Φ_i is contained in the union of all such i-planes.*

Lemma 5. *If a connected component ϕ of Φ_i has a nonempty intersection $\phi \cap P_i \ne \emptyset$ with an i-facet P_i of P, then $\phi \subset P_i$.*

Lemma 5 allows to reduce the problem under consideration, of estimating the number of P_i such that $\dim(A_i \cap P_i) = i$, to estimating the number of all connected components of Φ_i. Because we are able (see Section 2) to bound the number of all connected components of a sub-Pfaffian set given by a Pfaffian formula with solely existential quantifiers (this is *not* the case for the sub-Pfaffian set Φ_i), we introduce the following notion.

Definition 6. *A point $y \in \tilde{A}_i$ is called i-pseudoflat if there exist the points $v_1, \dots, v_i \in \tilde{A}_i$ such that*

$$|\det(v_1 - y, \dots, v_i - y)^T (v_1 - y, \dots, v_i - y)| > \varepsilon_1$$

and the points

$$y + \sum_{1 \le l \le i} \varepsilon_{ji+l+1}(v_l - y) \in \tilde{A}_i$$

for all $0 \le j \le n^2$.

Denote the sub-Pfaffian set of all i-pseudoflat points by $\tilde{\Phi}_i$. Observe that $\tilde{\Phi}_i$ can be defined by a Pfaffian formula with solely existential quantifiers.

The following lemma justifies the introduction of infinitesimals ε_{ji+l+1} in Definition 6.

Lemma 6. *Let the points $x, v_1, \dots, v_i \in A_i$ be such that the vectors $v_1 - x, \dots, v_i - x$ are linearly independent. Denote by Π the unique i-plane passing through $x, v_1, \dots, v_i$. If the points*

$$x + \sum_{1 \le l \le i} \varepsilon_{ji+l+1}(v_l - x) \in A_i$$

for all $0 \le j \le n^2$, then $\dim(A_i \cap \Pi) = i$.

The proof of the next inclusion relies on Lemma 6.

Lemma 7. $\mathrm{st}(\tilde{\Phi}_i) \subset \Phi_i$.

Using Lemma 2 we obtain the following lemma.

Lemma 8. *If $\dim(U_w \cap P_i) = i$ then $U_w \cap P_i \cap \mathbf{R}^n \subset \tilde{\Phi}_i$.*

One can prove (cf. Lemma 1 in [15]) that for any connected sub-Pfaffian set V its standard part $\mathrm{st}(V)$ is also connected. Together with Lemmas 7, 8 this allows to reduce the estimating the number of all connected components ϕ of Φ_i such that $\dim(\phi \cap P_i) = i$ for some i-facet P_i of P (see Lemma 5), to estimating the number of connected components of $\tilde{\Phi}_i$. The latter follows from Khovanskii's bound [19], [20].

Lemma 9. *The number of all connected components of the set $\tilde{\Phi}_i$ does not exceed $2^{k^2}(ndk)^{O(k+n^4)}$.*

We conclude that $\nu_i \le 2^{k^2}(ndk)^{O(k+n^4)}$. Hence, $N < 3^k 2^{k^2}(ndk)^{O(k+n^4)}$ (see the beginning of Section 2). This implies the theorem.

References

[1] M. Ben-Or, *Lower bounds for algebraic computation trees,* Proceedings of ACM Symposium on Theory of Computing, 1983, 80-86.

[2] A. Björner, *Subspace arrangements,* Proceedings of 1st European Congress of Mathematicians, Paris, 1992.

[3] A. Björner, and L. Lovasz, *Linear decision trees, subspace arrangements and Möbius functions,* Preprint, 1992.

[4] A. Björner, L. Lovasz, and A. Yao, *Linear decision trees: volume estimates and topological bounds,* Proceedings of ACM Symposium on Theory of Computing,1992, 170-177.

[5] A. Chistov, *Fast parallel calculation of the rank of matrices over a field of arbitrary characteristics,* Lecture Notes in Computer Science, V. 199, 1985, 63-69.

[6] M. Davis, *Applied Nonstandard Analysis,* John Wiley, 1977.

[7] A. Gabrielov, *Projections of semi-analytic sets,* Functional Anal. Appl., V. 2, 1968, 282-291.

[8] A. Gabrielov, *Existential formulas for analytic functions,* Preprint 93-60, Math. Sciences Inst., Cornell University, 1993.

[9] A. Gabrielov, and N. Vorobjov, *Complexity of stratifications of semi-Pfaffian sets,* Preprint 94-4, Math. Sciences Inst., Cornell University, 1994.

[10] D. Grigoriev, *Complexity of deciding Tarski algebra,* Journal Symbolic Comp., V. 5, 1988, 65-108.

[11] D. Grigoriev, *Deviation theorems for Pfaffian sigmoids,* St. Petersburg Math. Journal, V. 6, 1994, 127-131.

[12] D. Grigoriev, and M. Karpinski, *Lower bounds on complexity of testing membership to a polygon for algebraic, randomized and* $\exp - \log$ *trees,* Technical report TR-93-042, International Computer Science Institute, Berkeley, 1993.

[13] D. Grigoriev, M. Karpinski, and N. Vorobjov, *Lower bounds on testing membership to a polyhedron by algebraic decision trees,* Proceedings of ACM Symposium on Theory of Computing, 1994, 635-644.

[14] D. Grigoriev, M. Singer, and A. Yao, *On computing algebraic functions using logarithms and exponentials,* Technical report 93-07, DIMACS, 1993.

[15] D. Grigoriev, and N. Vorobjov, *Solving systems of polynomial inequalities in subexponential time,* Journal Symbolic Comp., V. 5, 1988, 37-64.

[16] D. Grigoriev, and N. Vorobjov, *Complexity lower bounds for computation trees with elementary transcendental function gates,* Preprint, 1994.

[17] J. Heintz, M.-F. Roy, and P. Solernó, *Sur la complexite du principe de Tarski-Seidenberg,* Bull. Soc. Math. France, V. 118, 1990, 101-126.

[18] J. Já Já, *Computation of algebraic functions with root extractions,* Proceedings of IEEE Symposium on Foundations of Computer Science, 1981, 95-100.

[19] A. Khovanskii, *Fewnomials and Pfaff manifolds,* Proceedings of International Congress of Mathematicians, Warszawa, 1983, 549-564.

[20] A. Khovanskii, *Fewnomials,* Translations of Mathematical Monographs, V. 88, Amer. Math. Soc., 1991.

[21] F. Meyer auf der Heide, *Fast algorithms for n-dimensional restrictions of hard problems,* Proceedings of ACM Symposium on Theory of Computing, 1985, 413-420.

[22] J. Milnor, *On the Betti numbers of real varieties,* Proc. Amer. Math. Soc., V. 15, 1964, 275-280.

[23] J. Renegar, *On the computational complexity and the first order theory of the reals. Parts I-III,* Journal Symbolic Comp., V. 13, 1992, 255-352.

[24] N. Vorobjov, *The complexity of deciding consistency of systems of polynomial in exponential inequalities,* Journal Symbolic Comp., V. 13, 1992, 139-173.

[25] A. Yao, *Algebraic decision trees and Euler characteristics,* Proceedings of IEEE Symposium on Foundations of Computer Science, 1992, 268-277.

[26] A. Yao, *Decision tree complexity and Betti numbers,* Proceedings of ACM Symposium on Theory of Computing, 1994, 615-624.

[27] A. Yao, and R. Rivest, *On the polyhedral decision problem,* SIAM J. Comput., V. 9, 1980, 343-347.