A Proof of Strassen's Degree Bound for Homogeneous Arithmetic Circuits

Senior Thesis
Princeton University
Department of Mathematics

Ben Edelman

May 7, 2018

Abstract

The field of algebraic complexity theory is concerned with the amount of fundamental resources needed to perform various algebraic computations. One of the central challenges of algebraic complexity theory is to find explicit polynomials that cannot be computed by small arithmetic circuits. Strassen's degree bound [1] on the complexity of circuits computing various natural explicit collections of polynomials – such as ${x_1}^k, {x_2}^k, \ldots, {x_n}^k$, as well as the elementary symmetric polynomials – remains unsurpassed in this regard 45 years after its publication. In this thesis, I introduce arithmetic circuits and the degree bound, and I provide an alternate proof for the special class of arithmetic circuits known as homogeneous arithmetic circuits.

Contents

1 Introduction **2**
1.1 Algebraic computation 2
1.1.1 Arithmetic circuits 2
1.2 Strassen's degree bound 5
1.3 Other proofs of the degree bound 9
1.4 Main result 10

2 Proof of the main theorem **11**
2.1 A few more preliminaries 11
2.2 Full proof 12

3 Conclusion **14**

4 Acknowledgments **15**

1 Introduction

1.1 Algebraic computation

Much of computer science nowadays fundamentally deals with not bits and bytes and ANDs and ORs, but with numbers and variables and additions and multiplications. Since Turing it has been known that both ways of thinking about computation are in the end equivalent, but when—as is very common—we want to multiply matrices, or perform a Fourier transform, or implement an algebraic error-correcting code, the algebraic way of thinking can often be a more natural way of thinking about the problem at hand systematically. In so many cases, we want to be able to study the computation of *polynomials*. This is especially true for the study of the hardness, or 'complexity', of algebraic computation.

In order to rigorously study how hard polynomials are to compute, we must fix a model of computation. Even though we can technically model the computation of polynomials with Turing machines or boolean circuits, arithmetic circuits are the dominant model for this sort of computation because they make analysis much simpler and essentially capture the full complexity of algebraic computation.[1]

1.1.1 Arithmetic circuits

An arithmetic circuit C over a field $\mathbb{F}$ is a directed acyclic graph: in other words, it consists of vertices, and edges (with direction indicated) that connect pairs of vertices. Each vertex v computes a polynomial $[v]$ over $\mathbb{F}$ based on what type of vertex it is:

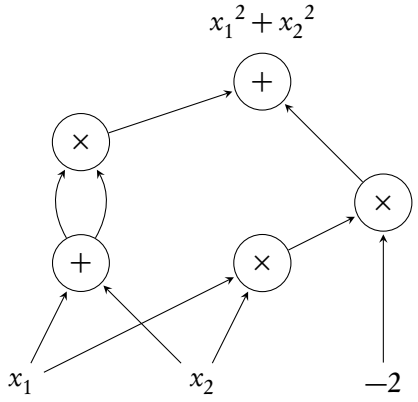
- *Input vertices* are vertices that have no in-edges. Each input vertex is labeled with either a field element $f \in \mathbb{F}$ or a variable x_i, and we say that the vertex computes its label.
- *Gates* are vertices that have in-edges. We require of C that every gate has in-degree 2 (i.e. 'fan-in' 2). For a given gate g, let u and v be the two vertices that have edges leading into g. There are two types of gates:
  - *Plus gates*, labeled with $+$, satisfy $[g] = [u] + [v]$.

[1] Discussion of alternative models for algebraic computation can be found in chapter 4 of [2].

– *Product gates*, labeled with $\times$, satisfy $[g] = [u] \cdot [v]$.

In this manner, every vertex inductively computes a polynomial. Some of the vertices are *output vertices*. Let Out be the set of output vertices. Then we say that C computes the polynomials computed by the vertices in Out.[2]

Here's an example. Let C_1 be the following circuit over $\mathbb{C}$:



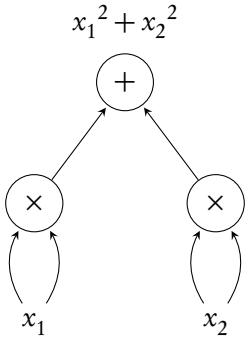
Let's say the topmost plus gate in C_1 is the only output gate. C_1, thus, computes the polynomial $x_1^2 + x_2^2$.

The measure of complexity of algebraic computation we will use is *circuit size*. We will define the size of a circuit C, denoted $S(C)$, as the number of vertices in C. [3] Looking at our example, we see that $S(C_1) = 8$.

The arithmetic circuit complexity of a polynomial f, denoted $S(f)$, is given by the size of the smallest circuit that computes f. More generally, the complexity of the collection of polynomials $f_1, \ldots, f_k$, is defined similarly and denoted $S(f_1, \ldots, f_k)$. Suppose $f = x_1^2 + x_2^2$. Because C_1 computes f, we know $S(f) \leq 8$. But, as is easy to see, C_1 is a rather inefficient way of computing f. We can compute it more efficiently with the following circuit, C_2:

[2] One fine but important distinction to make is that if C computes the polynomial $f(x_1, \ldots, x_n)$, and $f(b_1, \ldots, b_n) = g(b_1, \ldots, b_n)$ for all $b_1, \ldots, b_n \in \mathbb{F}^n$, this does not necessarily imply that C computes g. We only say C computes g if the form of f and g as formal polynomials is the same (they have the same terms). In jargon, we are dealing with *syntactic*, not *semantic* computation.

[3] The edge count will differ from the vertex count by a factor of ≤ 2 because the fan-in is 2.



$S(C_2) = 5$, so $S(f) \leq 5$. In fact, it turns out that $S(f) = 5$. In this simple case, the lower bound $S(f) \geq 5$ can be deduced through specific means, but it isn't particularly interesting or useful to prove circuit lower bounds laboriously one at a time: the real task is to find lower bounds on the arithmetic circuit complexity of infinite families of desired outputs.

Before we touch on the state of the art on arithmetic circuit lower bounds, I'll present two somewhat trivial lower bounds.

Remark 1 (simple dimension bound)**.** *If the variables $x_1, x_2, \ldots, x_n$ all appear in the formula for f, then $S(f) \geq n$.*

Proof. If C computes f, then C must contain n input vertices labeled $x_1, \ldots, x_n$ respectively: otherwise there would be no way for any of these variables to enter the computation. □

In other words, for a polynomial over $\mathbb{F}[x_1, \ldots, x_n]$, it is often easy to demonstrate a arithmetic circuit lower bound of n. What we tend to be interested in, then, is superlinear (in n) lower bounds.

Remark 2 (simple degree bound [4])**.** *If $\deg f \geq 2^d$, then $S(f) \geq d$.*

Proof. The proof is by induction on d. If $d = 0$, then the bound is trivial. Now suppose we know that any polynomial of degree $\geq 2^{s-1}$ can only be computed by circuits of size $\geq s - 1$. In any circuit C computing a polynomial f of degree $d \geq 2^s$, the output gate g is a plus gate or a product gate. In either case, at least one

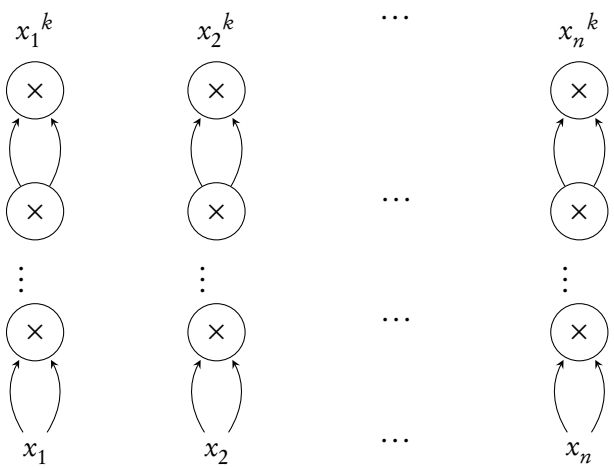
[4] When I refer to the degree of a polynomial, I am always talking about the total degree: the maximum of the degrees of each term, where the degree of a term is the sum of the exponents of its variables.

of its predecessor vertices v must compute a polynomial $[v]$ of degree $\geq \frac{d}{2} \geq 2^{s-1}$. Since C is acyclic, there are no paths from g to v, so the portion of the circuit that computes v does not include g. This portion must have size $\geq s-1$ by the induction hypothesis, so C, which additionally includes at least g, must have size $\geq s$. □

In other words, for a polynomial of degree D, it is easy to demonstrate a lower bound of $\log D$. Typically, we care about polynomials with degree at most polynomial in the number of variables, so this remark doesn't give us superlinear lower bounds.

1.2 Strassen's degree bound

Suppose we want to compute ${x_1}^k, \ldots, {x_n}^k$. By Remark 2, each polynomial ${x_i}^k$ has complexity $\log k$. And it seems intuitively plausible that the ability to compute all these monomials together in parallel shouldn't allow us to use fewer resources than computing them separately: in other words, that the following circuit C of size $n \log k$ is the smallest one we can hope for:



I encourage you to try playing around with small cases and seeing whether you share this intuition.

In a landmark 1973 paper [1] that remains the state of the art in general arithmetic circuit lower bounds, Strassen proved that the above intuition is correct: no

circuit smaller than C computes the n monomials.[5]

Proposition 1. $S({x_1}^k, \ldots, {x_n}^k) \geq n \log k$

In fact, this proposition is just one case of Strassen's result, which is known as the degree bound. The full statement of the degree bound (and its proof) relies upon some basic algebraic geometry, beginning with the concept of an algebraic variety.[6] In this section we will work over the algebraically closed field $\mathbb{C}$.

We define the *variety* of a set of polynomials $f_1, \ldots, f_k$ over $\mathbb{C}[x_1, \ldots, x_n]$, denoted $V(f_1, \ldots, f_k)$, as the set of points in $\mathbb{C}^n$ that are mapped to zero by all the polynomials f_i when considered as functions from $\mathbb{C}^n \to \mathbb{C}$:

$$V(f_1, \ldots, f_k) = \{(a_1, \ldots, a_n) \in \mathbb{C}^n : f_1((a_1, \ldots, a_n)) = \ldots = f_k((a_1, \ldots, a_n)) = 0\}$$

Varieties have two properties that we will be concerned with: *dimension* and *degree*.

The *dimension* of a variety $V(f_1, \ldots, f_k)$ roughly corresponds to the intuitive conception of the dimension of the set V: if V contains a finite number of points, its dimension of 0; if V forms a curve, then its dimension is 1; and so on. Formally, $\dim V(f_1, \ldots, f_k)$ is the minimum size of a set of hyperplanes such that the intersection of these hyperplanes with V is finite.[7]

The *degree* of $V(f_1, \ldots, f_k)$, denoted $\deg V(f_1, \ldots, f_k)$, is the maximum finite number of points we can obtain by intersecting the variety with $\dim V(f_1, \ldots, f_k)$ hyperplanes. Thus, dimension and degree are intimately related. Strassen's proof relies upon the notion of degree, while dimension has relevance for my contribution.

Strassen's proof hinges on a classical theorem from algebraic geometry, Bezout's theorem, which I will state without proof.

Theorem 1 (Bezout's theorem)**.** $\deg V(f_1, \ldots, f_k) \leq \prod_{i=1}^{k} \deg f_i$

Now we are ready to see Strassen's theorem.

[5] Strassen's proof only deals with algebraically closed fields.

[6] My presentation of the degree bound is indebted to the recent survey by Chen, Kayal, and Wigderson [3].

[7] Hyperplanes are linear subspaces of dimension $n-1$, which can be defined by equations of the form $a_1 x_1 + \ldots + a_n x_n = b$.

Theorem 2 (Strassen's degree bound)**.** *For any collection of polynomials* $f_1, \ldots, f_k \in \mathbb{C}[x_1, \ldots, x_n]$,

$$S(f_1, \ldots, f_k) \geq \log \deg V(y_1 - f_1(\mathbf{x}), \ldots, y_k - f_k(\mathbf{x}))$$

where the variables $y_1, \ldots, y_k$ *are introduced as auxiliaries.*

(In particular, if there exist field elements $a_1, \ldots, a_n \in \mathbb{F}$ such that the number of solutions for $f_1(\mathbf{x}) = a_1, \ldots, f_n(\mathbf{x}) = a_n$ is N, then $S(f_1, \ldots, f_k) \geq \log N$.)

Proof sketch. Consider any circuit C that computes $f_1, \ldots, f_k$, and let $s = S(C)$. The main idea of the proof is to introduce an auxiliary variable z_{v_i} for each vertex v_i in C and encode the circuit into a variety over $\mathbb{C}[x_1, \ldots, x_n, z_{v_1}, \ldots, z_{v_s}]$.

We introduce one polynomial h_v for each vertex v in C:

- If v is an input vertex labeled by x_i, let $h_v(x_1, \ldots, x_n, z_{v_1}, \ldots, z_{v_s}) = z_v - x_i$. It v is labeled by the field element c, let $h_v(x_1, \ldots, x_n, z_{v_1}, \ldots, z_{v_s}) = z_v - c$.
- If v is a plus gate, let $h_v(x_1, \ldots, x_n, z_{v_1}, \ldots, z_{v_s}) = z_v - (z_u + z_w)$, where u and w are the predecessors of v.
- If v is a product gate, let $h_v(x_1, \ldots, x_n, z_{v_1}, \ldots, z_{v_s}) = z_v - z_u z_w$.

The condition $h_v = 0$ asserts that the output of gate v follows from the inputs in the proper manner. The set of conditions $\{h_v = 0 \colon v \in C\}$, then, asserts that the outputs of C (the variables $\{z_v : v \in \mathsf{Out}\}$) follow from the inputs $x_1, \ldots, x_n$ in the proper manner. If we recall the auxiliarly variables y_i from the theorem statement, we can restate this as the observation that mapping the output variables $\{z_v : v \in \mathsf{Out}\}$ to the variables $\{y_i : i \in [k]\}$ induces a bijection from $V(\{h_v : v \in C\})$ to $V(\{y_i - f_i(x_1, \ldots, x_n) : i \in [k]\}$.

By Bezout's theorem, $\deg V(\{h_v : v \in C\}) \leq \prod_{v \in C} \deg h_v \leq 2^{S(C)}$. Hence, $\deg V(\{y_i - f_i(x_1, \ldots, x_n) : i \in [k]\} \leq 2^{S(C)}$, i.e.:

$$S(C) \geq \log \deg V(\{y_i - f_i(x_1, \ldots, x_n) : i \in [k]\})$$

. □

Note that the degree bound is in fact a lower bound on the number of product gates, because in the proof the input vertices and plus gates yield polynomials of degree 1, while product gates yield polynomials of degree 2, and only the degree 2 factors are responsible for the magnitude of $\prod_{v \in C} \deg h_v$.

Now that we have seen the degree bound, we can apply it to to obtain Proposition 1: $S({x_1}^k, \ldots, {x_n}^k) \geq n \log k$.

Proof. By the degree bound, $S({x_1}^k, \ldots, {x_n}^k) \geq \log \deg V(y_1 - {x_1}^k, \ldots, y_n - {x_n}^k)$. If we intersect $V(y_1 - {x_1}^k, \ldots, y_n - {x_n}^k)$ with the hyperplanes $\{y_i = 1 : i \in [n]\}$, then we obtain the set $\{{\zeta_k}^j : j \in [k]\}^n$, where ζ_k is a primitive kth root of unity and the exponentiation of the set indicates taking the cartesian product. This set has size k^n, so $S({x_1}^k, \ldots, {x_n}^k) \geq \log k^n = n \log k$. □

Let's briefly see another important example application of the degree bound: the *elementary symmetric polynomials*. We define the elementary symmetric polynomials $\{\sigma_j : j \in [n]\}$ as:

$$
\begin{aligned}
\sigma_1(x_1, \ldots, x_n) &= x_1 + \ldots + x_n \\
\sigma_2(x_1, \ldots, x_n) &= \sum_{1 \leq i < j \leq n} x_i x_j \\
&\vdots \\
\sigma_j(x_1, \ldots, x_n) &= \sum_{1 \leq i_1 < \ldots < i_j \leq n} x_{i_1} \cdots x_{i_j} \\
&\vdots \\
\sigma_n(x_1, \ldots, x_n) &= x_1 \cdots x_n
\end{aligned}
$$

Corollary 1. $S(\sigma_1, \ldots, \sigma_n) \geq \log(n!) = \Omega(n \log n)$

I leave the proof of this corollary as an interesting exercise.

In 1983, Baur and Strassen proved a remarkable result that allows the application of the degree bound to lower-bounding the complexity of single polynomials [4]. They proved that computing a polynomial f and all the partial derivatives of f only incurs a constant factor blowup in the size of the required circuit versus just computing f alone:

Remark 3 (Baur-Strassen). *For any* $f \in \mathbb{C}[x_1, \ldots, x_n]$,

$$S(f) = \Omega\left(S(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})\right)$$

Consider now the polynomial $f(x_1, \ldots, x_n) = {x_1}^k + \ldots + {x_n}^k$. We have

$$S(f) = \Omega\left(S(f, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})\right) = \Omega\left(S(k{x_1}^{k-1}, \ldots, k{x_n}^{k-1})\right) = \Omega(n \log k)$$

where the last equality follows from the degree bound. A subtler application of the same technique yields $S(\sigma_j) = \Omega(n \log \min(j, n-j))$. Both this bound and the bound on f are tight up to constant factors.

At this point I would like to emphasize again that the degree bound remains the state of the art 45 years after Strassen proved it. Nobody knows how to prove a lower bound better than $\Omega(n \log n)$ on a polynomial of degree $O(n)$, even though it can be shown that most polynomials have complexity that isn't even polynomial in n [3].

1.3 Other proofs of the degree bound

Even though improving on the degree bound appears to be beyond the reach of the field at the moment, several researchers over the years have found new proofs of the classic result.

Notably, in 1976, Schönhage derived with elementary means a lemma on algebraic dependence, and used this lemma to prove the degree bound without relying on any theorems from algebraic geometry such as Bezout's theorem [5].

Later, in 1996, Smolensky [6] provided a different elementary proof of the degree bound for the special case ${x_1}^k, \ldots, {x_n}^k$ by giving a reduction from arithmetic circuits to a novel circuit model invented specifically for this proof.

Most relevant to this paper is the recent proof by Kumar in 2017 for the case ${x_1}^k + \ldots + {x_n}^k$ for homogeneous circuits, not all arithmetic circuits [7]. A homogeneous polynomial is a polynomial in which each term has the same degree. Homogeneous polynomials are quite common in algebraic computation; all the examples I've used of polynomials have been homogeneous. A *homogeneous circuit* is an arithmetic circuit in which $[v]$ is homogeneous for every vertex v. An equivalent definition is that a homogeneous circuit is an arithmetic circuit for

which the two predecessor polynomials of any plus gate must have the same degree. Kumar's paper mainly deals with proving a new lower bound for algebraic branching programs, another model of algebraic computation, but it mentions without a full proof that its methods can be extended to arithmetic circuits. While I didn't know about this paper until this thesis was almost completed, its methods are in fact very similar to mine. My main result won't just cover the special case ${x_1}^k + \ldots + {x_n}^k$, and my proof uses different language than Kumar's, but both proofs are quite similar. It is my hope that my exposition can make it even clearer how the case of homogeneous circuits is fruitful to methods rather different from those of Strassen (and Schönhage, and Smolensky). Moreover, my proof will involve a new lemma (Lemma 1) that sheds light on the structural properties of homogeneous arithmetic circuits.

1.4 Main result

A few algebraic (but not algebraic geometric) definitions are in order before my theorem can be properly stated.

An *ideal* I of the polynomial ring $R = \mathbb{F}[x_1, \ldots, x_n]$ is a subset of R that is a group under addition and satisfies $fh \in I$ for all $f \in I, h \in R$. The ideal generated by the polynomials $f_1, \ldots, f_k$, is defined as $\{h_1 f_1 + \ldots + h_n f_n : h_1, \ldots, h_n \in R\}$ and is denoted by $(f_1, \ldots, f_k)$. Note that $(f_1, \ldots, f_k)$ is the smallest ideal containing $f_1, \ldots, f_k$.

An ideal I is *proper* if $I \subsetneq R$.

A proper ideal I is *prime* if, for any $f, h \in R$, whenever $fh \in R$ then either $f \in R$ or $h \in R$.

The *codimension* (also called the *height*) of a prime ideal I is the length of the longest chain[8] of prime ideals leading up to I [8]. By this I mean that $\operatorname{codim} I$ is the length k of the largest sequence of prime ideals $I_1, \ldots, I_k$ over R such that $I_1 \subsetneq \ldots \subsetneq I_k = I$. In general, the codimension of any proper ideal I is defined as the smallest codimension of any prime ideal containing I. Interestingly, connecting back to the discussion of algebraic geometry, it turns out that $\operatorname{codim}(f_1, \ldots, f_k) = n - \dim V(f_1, \ldots, f_k)$.

[8] For the existence of a longest chain I am relying on the fact that any polynomial ring is Noetherian, which is a consequence of Hilbert's basis theorem.

There's one last notation to introduce: $S_H(f_1,\ldots,f_k)$ is the size of the smallest homogeneous arithmetic circuit computing $f_1,\ldots,f_k$. We are now ready to see the main result.

Theorem 3 (Main theorem)**.** *If $f_1,\ldots,f_k$ are homogeneous polynomials each of degree $\geq D = 2^d$, and* $\operatorname{codim}(f_1,\ldots,f_k) \geq r$*, then $S_H(f_1,\ldots,f_n) \geq rd$.*

Proof sketch. The full proof is in Section 2, but I can sketch the basic idea beforehand. Essentially, for any circuit C computing $f_1,\ldots,f_k$, we partition C into d levels: the vertices computing polynomials with degree in $[2^{d-1},2^d)$, those with degree in $[2^{d-2},2^{d-1})$, and so on, all the way down to those with degree in $[1,2)$. We will prove that the polynomials computed by the vertices in any level 'span the space' of the polynomials computed in all the higher levels (to be precise, we'll look at the ideals generated by these polynomials). Because we are given that the polynomials in the highest level 'span a big space' (i.e., generate an ideal with co-dimension at least r), this will imply that each of the lower levels will also need to span this big space and thus have lots of vertices (r vertices each, to be precise). C must therefore contain at least rd vertices. □

2 Proof of the main theorem

2.1 A few more preliminaries

Before I give the full proof, there are a few more facts it will be helpful to have in hand.

First of all, note the somewhat trivial fact that if $I \subset J$, then $\operatorname{codim} I \leq \operatorname{codim} J$. This will come in handy.

The second fact is the only outside theorem I need for my proof. It is (the generalized version of) Krull's principal ideal theorem, a classic result from commutative algebra.

Theorem 4 (General form of Krull's principal ideal theorem)**.** *If R is a Noetherian ring (so, for example, if it is any polynomial ring over a field), and $h_1,\ldots,h_r \in R$, then*

$$\operatorname{codim}(h_1,\ldots,h_r) \leq r$$

as long as $(h_1,\ldots,h_r)$ is a proper ideal.

A proof of the theorem can be found in pretty much any primer on commutative algebra, such as [8].

We can use Krull's principal ideal theorem to help answer a simple question: what is the codimension of $(x_1, \ldots, x_n)$? Well, it is easy to see that the ideals $(x_1), (x_1, x_2), \ldots, (x_1, \ldots, x_n)$ are all prime, so $\operatorname{codim}(x_1, \ldots, x_n) \geq n$. By Krull's principal ideal theorem, $\operatorname{codim}(x_1, \ldots, x_n) \leq n$, so $\operatorname{codim}(x_1, \ldots, x_n) = n$.

For our setting, it will be particularly useful that if an ideal is generated by homogeneous polynomials of degree ≥ 1, then it is proper. This fact can be easily verified.

One more thing: For brevity of presentation, if V is a set of vertices in an arithmetic circuit, let $(V) := (\{[v] : v \in V\})$.

2.2 Full proof

Lemma 1. *Given $d \geq 1$ and a homogeneous circuit C, let U_1 be the set of vertices in C computing polynomials of degree $\geq 2^d$ and let U_2 be the set of vertices in C computing polynomials of degree in $[2^{d-1}, 2^d)$. Then $(U_1) \subset (U_2)$.*

Proof. It is sufficient to show that for any individual vertex v with $\deg[v] \geq 2^d$, it is true that $[v] \in (U_2)$. To prove this, it is helpful to use the fact that the vertices of any directed acyclic graph (such as C) can be topologically ordered (given a labeling $v_1, \ldots, v_{S(C)}$) such that all edges respect the ordering: for any $i, j \in [S(C)]$, if $i < j$ then there is no edge from v_j to v_i.

Let m be the lowest index such that $\deg[v_m] \geq 2^d$. For every $i \geq m$, $v_i \in U_1$. The proof will be by strong induction on the index i of the sorted vertices.

The base case is when $i = m$. v_i cannot be a plus gate because if it were, the polynomials computed by its predecessors would have the same degree as $[v]$, which is impossible by definition of m. It also can't be an input vertex because $\deg[v] \geq 2$. Thus, v_i is a product gate. The degrees of its two predecessors add up to $\deg[v]$, so at least one of its predecessors–call it u–must have degree $\geq \deg[v]/2 \geq 2^{d-1}$. We also know by definition of m that $\deg[u] < 2^d$, so $u \in U_2$. $[v_m]$ is the product of $[u]$ with some other polynomial, so $[v_m] \in ([u]) \in (U_2)$.

Now, for the induction step, we want to prove $[v_j] \in (U_2)$ given that $[v_i] \in (U_2)$ for all $m \leq i < j$. If v_j is a plus gate, it is in the ideal generated by its two predecessors, which have degree $\deg[v_j]$ and thus have indices in $[m, j)$, implying

they are in (U_2). Thus, in this case, $v_j \in (U_2)$. If v_j is a product gate, as above one of its predecessors must have degree $\geq 2^{d-1}$, and $[v_j]$ is generated by this predecessor, so again $v_j \in (U_2)$, and we are done. □

As an aside, it follows easily from this lemma that if we fix the value of every vertex in U_2 to 0, then all the vertices in U_1 will evaluate to 0.

Theorem 3 (Main theorem)**.** *If $f_1, \ldots, f_k$ are homogeneous polynomials each of degree $\geq D = 2^d$, and* $\operatorname{codim}(f_1, \ldots, f_k) \geq r$*, then $S_H(f_1, \ldots, f_n) \geq rd$.*

Proof. The proof is by induction on d.

If $d = 0$, then the bound trivially holds.

Now suppose we know the statement is true whenever $d = s-1$, and we want to prove it for $d = s$. In other words, we need to prove that if a homogeneous circuit C computes a collection of polynomials $f_1, \ldots, f_k$ that satisfies $\deg f_i \geq 2^s$ for all i and $\operatorname{codim}(f_1, \ldots, f_k) \geq r$, then $S(C) \geq rs$. In order to prove this, we need only show that there exists a set of vertices W in C such that $2^{s-1} \leq \deg[w] < 2^s$ for all $w \in W$ and $\operatorname{codim}(W) \geq r$. Because C is homogeneous, directed edges in C never point in the direction of lower degree, so the portion of C leading up to W is completely disjoint from Out. This implies $S(C) \geq |\mathsf{Out}| + S_H(W)$. Furthermore, $|\mathsf{Out}| \geq r$ by Krull's principal ideal theorem and $S_H(W) \geq r(s-1)$ by the induction hypothesis, so $S(C) \geq r + r(s-1) = rs$. Hence, once we demonstrate W exists we will be done. And we can do this by applying Lemma 1, by which there must exist a set of vertices W such that $2^{s-1} \leq \deg[w] < 2^s$ for all $w \in W$, and $(f_1, \ldots, f_k) \subset (W)$. This last condition implies $\operatorname{codim} W \geq \operatorname{codim}(f_1, \ldots, f_k) \geq r$, so the proof is complete. □

Corollary 2.

$$S_H({x_1}^k, \ldots, {x_n}^k) \geq n \log k$$

Proof. In order to apply Theorem 3 and conclude the result, we only need to show that $\operatorname{codim}({x_1}^k, \ldots, {x_n}^k) = n$. If we are working over the field $\mathbb{C}$, then the easiest way to do this is through the equivalent claim that $\dim V({x_1}^k, \ldots, {x_n}^k) = 0$, which is true because $V({x_1}^k, \ldots, {x_n}^k) = \{0\}$. But we can prove $\operatorname{codim}({x_1}^k, \ldots, {x_n}^k) = n$ for general fields using just the definition of codimension, not any algebraic geometry, as follows.

First, we compute the smallest prime ideal I containing $({x_1}^k, \ldots, {x_n}^k)$, because by definition $\operatorname{codim}({x_1}^k, \ldots, {x_n}^k) = \operatorname{codim} I$. Since ${x_1}^k = x_1 \cdot {x_1}^{k-1} \in I$, then by the definition of prime ideals, if $x_1 \notin I$ then ${x_1}^{k-1} \in I$. And ${x_1}^{k-1} = x_1 \cdot {x_1}^{k-2}$, so if $x_1 \notin I$ then ${x_1}^{k-2} \in I$. We can continue inducting in this manner to show that if $x_1 \notin I$ then $x_1 \in I$, a contradiction, so it must be true that $x_1 \in I$. The same argument implies $x_2 \in I, \ldots, x_n \in I$. Thus, $(x_1, \ldots, x_n) \in I$. $(x_1, \ldots, x_n)$ is prime because if $f \in (x_1, \ldots, x_n)$ then $\deg f > 0$, so for any f_1, f_2 such that $f_1 f_2 = f$ it is true that either f_1 or f_2 has positive degree and so is also in $(x_1, \ldots, x_n)$. Thus, $I = (x_1, \ldots, x_n)$. We proved earlier that $\operatorname{codim}(x_1, \ldots, x_n) = n$, so we are done. □

Corollary 3. $S_H(\sigma_1, \ldots, \sigma_n) = \Omega(n \log n)$

Proof. $S_H(\sigma_1, \ldots, \sigma_n) \geq S_H(\sigma_{n/2+1}, \ldots, \sigma_n)$. And for $n/2 \leq j \leq n$, $\deg \sigma_j = j \geq n/2$. In order to apply Theorem 3, we must show $\operatorname{codim}(\sigma_{n/2+1}, \ldots, \sigma_n) \geq n/2$. Let I be a prime ideal containing $(\sigma_{n/2+1}, \ldots, \sigma_n)$. Since I contains $\sigma_n = x_1 \cdots x_n$, if it doesn't contain x_1 then it must contain $x_2 \cdots x_n$. If the latter is true, then if it doesn't contain x_2 it must contain $x_3 \cdots x_n$. We can continue this line of reasoning to determine that I must contain x_i for some i. Suppose without loss of generality that $x_1 \in I$. Then, by the definition of an ideal, we can take σ_{n-1}, subtract out all the terms with x_1, and what remains, $x_2 \cdots x_n$, must still be in I. From this we can deduce as above without loss of generality that $x_2 \in I$, and apply the same trick to σ_{n-2} and so on to eventually find that without loss of generality $x_1, \ldots, x_{n/2} \in I$. $\operatorname{codim}(x_1, \ldots, x_{n/2}) = n/2$, so by Theorem 3:

$$S_H(\sigma_{n/2+1}, \ldots, \sigma_n) \geq \frac{n}{2} \log \frac{n}{2} = \Omega(n \log n)$$

□

3 Conclusion

Kumar has already shown that the type of methods I use here to reprove Strassen's bound for the homogeneous case can be used to obtain new results about algebraic branching programs. Perhaps these methods, especially Lemma 1, can find further applications for other problems and models of computation.

Moreover, Strassen's original proof and Schönhage's proof only established the degree bound directly for infinite fields (the same applies for Kumar's proof,

because it uses algebraic geometric techniques). My proof does not require that the field be infinite.

While my proof is more general in this respect, it relies on at least three properties of homogeneous circuits not shared by other arithmetic circuits:

- In a homogeneous circuit, if u is a predecessor of v, then $\deg[u]$ is never greater than $\deg[v]$.
- In a homogeneous circuit, both the predecessors of a plus gate v compute polynomials with the same degree as $[v]$.
- Finally, it is crucial that any ideal generated by homogeneous polynomials of degree ≥ 1 is proper (in other words, it doesn't include 1).

4 Acknowledgments

Thank you to Ran for being a great advisor and for suggesting in class that the degree bound ought to have a more intuitive proof, to Enric for listening to my inchoate ravings and pointing out the inevitable flaws in my reasoning, and to all the artichokes of 2 Dickinson Street for making it the best place in the galaxy to be semi-productive.

References

[1] Volker Strassen. "Die Berechnungs komplexität von elementarsymmetrischen Funktionen und von Interpolationskoeffizienten." *Numerische Mathematik* 20, no. 3 (1973): 238-251.

[2] Peter Bürgisser, Michael Clausen, and Amin Shokrollahi. *Algebraic Complexity Theory*. Vol. 315. Springer Science & Business Media, 1996.

[3] Xi Chen, Neeraj Kayal, and Avi Wigderson. "Partial derivatives in arithmetic complexity and beyond." *Foundations and Trends in Theoretical Computer Science* 6, no. 1–2 (2011): 1-138.

[4] Walter Baur and Volker Strassen. "The complexity of partial derivatives." *Theoretical Computer Science* 22, no. 3 (1983): 317-330.

[5] Arnold Schönhage. "An elementary proof for Strassen's degree bound." *Theoretical Computer Science* 3, no. 2 (1976): 267-272.

[6] Roman Smolensky. "Easy lower bound for a strange computational model." *Computational Complexity* 6, no. 3 (1996): 213-216.

[7] Mrinal Kumar. "A Quadratic Lower Bound for Homogeneous Algebraic Branching Programs." In *32nd Computational Complexity Conference*. 2017.

[8] David Eisenbud. *Commutative Algebra: with a view toward algebraic geometry*. Graduate Texts in Mathematics Vol. 150. Springer-Verlag, 1995.