Ecole Normale Supérieure de Lyon

M2 Internship Report

Advisor : Pascal Koiran

Lower bounds for univariate polynomials : a Wronskian approach

Timothée Pecatte

02/03/2014 to 07/11/2014

MC2 Team - Ecole Normale Supérieure de Lyon

Abstract

This is the final report of an internship in algebraic complexity. First, we give an introduction to algebraic complexity and we give some motivations for the study of the main model. Then we present two different tools we studied during this internship and use them to establish some lower bounds on this model. We finally discuss whether those bounds could be improved or not.

Contents

1 Algebraic complexity: an introduction

In algebraic complexity, the objects that are studied are no longer words over finite alphabet but polynomials over a field $\mathbb{F}$. However, the question remains the same: is a polynomial f hard to compute ? More precisely, we need to define a model of computation for polynomials and associated complexity measures. Arithmetic circuits are the most natural and standard model to compute polynomials. In this model, the inputs are variables $x_1, \ldots, x_n$, and the computation is performed using arithmetic operations $+, \times$, and may involve constants from the underlying field $\mathbb{F}$. The output of an arithmetic circuit is thus a polynomial (or a set of polynomial) in the input variable. The complexity measures associated are *size* and *depth* of the circuit which capture the number of operations and the maximal distance between an input gate and an output gate, respectively.

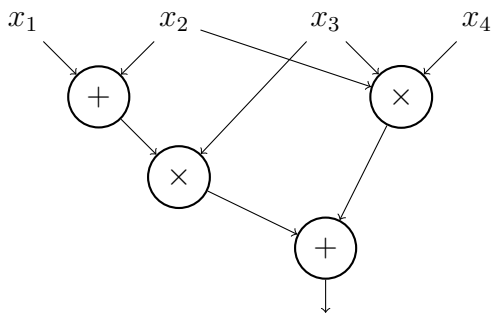

Figure 1: An arithmetic circuit computing $(x_1 + x_2)x_3 + x_2x_3x_4$, of depth 3 and of size 4.

There are two main kinds of problems in arithmetic complexity: find upper bounds and lower bounds on the complexity of a family of polynomials. Upper bounds usually consists of an explicit construction with controlled complexity whereas lower bounds are usually more complex to establish, involve advanced and new tools, and are often related to other interesting problems. Proving such bounds would allow to separate some *complexity classes*, in the same manner as in boolean complexity.

1.1 Valiant complexity classes

Arithmetic classes were first define in work of Valiant [Val79], in which he gave analogous definition for the classes P and NP in the algebraic world, and showed a complete problem for the later class. We now give some definitions to show the motivations of the problem we studied, more material about basic arithmetic complexity can be found in [BCS97, Bü00].

Definition 1 (VP). *A family of polynomials $\{f_n\}$ over $\mathbb{F}$ is* p-bounded *if there exists some polynomial $p : \mathbb{N} \to \mathbb{N}$ such that the number of variables and the degree of f_n are bounded by $p(n)$, and there is an arithmetic circuit of size at most $t(n)$ computing f_n. The class $\mathsf{VP}_{\mathbb{F}}$ consists of all p-bounded families over $\mathbb{F}$.*

The polynomial $f_n = x^{2^n}$, for example, is not in VP, even though it has $O(n)$ size circuits, as its degree is not polynomial. This restriction on the degree make in fact VP

more analogous to NC^2 than to P. Indeed, a depth reduction theorem proved in [VSBR83] states that any polynomial size algebraic circuit computing a polynomial of degree d can be turned into an algebraic circuit of polynomial size and depth $O(\log d \log n)$ (in fact we even have stronger depth reduction theorem, as we will see in next subsection). If the degree d if polynomially bounded, we end up with circuits of depth $O(\log^2 n)$, giving the analogy with NC^2.
A natural family in VP is the family of determinants:

$$\mathsf{DET}_n(X) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} x_{i,\sigma(i)}$$

An easy way to see that $\mathsf{DET} \in \mathsf{VP}$ is to compute it using Gauss pivot algorithm, which yields n^3 depth circuits, and then to get rid of the division gates using method described by Strassen [Str73].

Definition 2 (VNP)**.** *A family of polynomials $\{f_n\}$ over $\mathbb{F}$ is* p-definable *if there exists a family $\{g_n\}$ in $\mathsf{VP}_{\mathbb{F}}$ and two polynomially bounded functions $p, k : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$:*

$$f_n(x_1, \ldots, x_{k(n)}) = \sum_{w \in \{0,1\}^{p(n)}} g_{p(n)}\left(x_1, \ldots, x_{k(n)}, w_1 \ldots, w_{p(n)}\right)$$

The class $\mathsf{VNP}_{\mathbb{F}}$ consists of all p-definable families over $\mathbb{F}$.

The link between VNP and NP is harder to catch: the variables $(w_1, \ldots, w_{p(n)})$ can be seen as the "witness" and the summation is the algebraic equivalent of the existential quantifier for NP problems (in fact VNP is more analogous to $\#\mathsf{P}$ than to NP).
A natural family in VNP is the family of permanents:

$$\mathsf{PERM}_n(X) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} x_{i,\sigma(i)}$$

By definition of VP and VNP, it directly follows that $\mathsf{VP} \subseteq \mathsf{VNP}$. As an analogue to the P vs NP question, Valiant's conjectured that the inequality is strict:

Conjecture 3. $\mathsf{VP} \neq \mathsf{VNP}$

Arithmetic circuits have a lot of structure, so one could hope Valiant's conjecture to be easier than it's classical counterpart. In order to prove it, Valiant also defined a notion of completeness to capture the hardness of some polynomials in a same class. First, he defined a notion of reduction for two families of polynomials:

Definition 4. *The family $\{f_n\}$ is a* p-projection *of $\{g_n\}$ if there exists a polynomially bounded $p : \mathbb{N} \to \mathbb{N}$ such that for all n, f_n can be derived from $g_{p(n)}$ by a substitution of the variables by other variables or constants in $\mathbb{F}$.*

As one would expect, both VP and VNP are closed under p-projections. Moreover, Valiant showed in [Val79] PERM is complete for VNP, ie any family in VNP is a p-projection of the permanent. In particular, this implies that conjecture 3 is equivalent to prove an super-polynomial lower bound on the size of the circuits computing the permanent.

1.2 Restricted arithmetic circuit classes and depth reduction

General circuits are still too complicated to handle, very few lower bounds are known for them. Instead, people consider restricted circuit classes by adding constraint on the circuits computing the family of polynomials. We focus on depth restriction:

Definition 5 (Bounded-depth circuits)**.** *A family of circuits $\{C_i\}$ is of bounded depth if there exists a constant $d \in \mathbb{N}$ such that for any n, C_n has depth at most d.*

In particular, we will consider the case of depth-4 circuits, also known as $\Sigma\Pi\Sigma\Pi$ circuits. A $\Sigma\Pi\Sigma\Pi$ circuit is a depth-4 circuit with an addition gate at the top (output) then a layer of multiplication gates, then a layer of additive gates, then multiplication gates at bottom, i.e. it computes a polynomial of the form:

$$\sum_{i=1}^{k}\prod_{j=1}^{m}\sum_{l=1}^{t}\prod_{p\in S_{i,j,l}} x_p$$

where x_p is either an input variable or a constant in $\mathbb{F}$.
We usually denote $f_{i,j}(x_1,\ldots,x_n) = \sum_{l=1}^{t}\prod_{p\in S_{i,j,l}} x_p$ and thus $\Sigma\Pi\Sigma\Pi$ circuits compute polynomials of the form $\sum_{i=1}^{k}\prod_{j=1}^{m} f_{i,j}(x_1,\ldots,x_n)$, where the $f_{i,j}$'s are t-sparse multivariate polynomials.

Remark 6. *The choice of $\Sigma\Pi\Sigma\Pi$ rather than $\Pi\Sigma\Pi\Sigma$ isn't arbitrary: when we consider depth-d circuits, it's usually more interesting to consider circuits with an additive output gate. Indeed, if a polynomial f is computed by a circuit of depth d with a multiplicative output gate, we can always consider sub-circuits of depth $d-1$ which computes some factor of f. In the case of the additive output gate, it's more difficult to do the same because of possible cancellation: the sub-circuits of depths $d-1$ may compute polynomials of degree $> d$ and the final addition may cancel the term of too high degrees.*

The importance of $\Sigma\Pi\Sigma\Pi$ circuit comes from two main reasons. First, exponential lower bounds for DET and PERM for depth-3 circuits have already been proved in [GK98]. Second, Agrawal and Vinay [AV08] and subsequent strengthenings of Koiran [Koi12] and Tavenas [Tav13] showed that depth-4 circuits are as interesting as general circuits:

Theorem 7 ([AV08] [Koi12] [Tav13] Depth-reduction)**.** *Let f be an n-variate polynomial computed by a circuit of size s and of degree d. Then f is computed by a $\Sigma\Pi^{[O(\alpha)]}\Sigma\Pi^{[\beta]}$ circuit C of size $2^{O\left(\sqrt{d\log(ds)\log n}\right)}$ where $\alpha = \sqrt{d\frac{\log n}{\log ds}}$ and $\beta = \sqrt{d\frac{\log ds}{\log n}}$.*

In the particular case where $s, d = n^{O(1)}$, f is computed by a $\Sigma\Pi^{[O(\sqrt{d})]}\Sigma\Pi^{[c\cdot\sqrt{d}]}$ circuit C of size $n^{O(\sqrt{d})}$.

This depth-reduction theorem implies that lower bounds for the depth-4 arithmetic circuit model will give lower bounds for general arithmetic circuits. Recent results of [GKKS13, KSS13, FLMS13] gave lower bound that comes very close to the required threshold for different polynomial. For instance, Gupta, Kamath, Kayal and Saptharishi [GKKS13] showed the following lower bound for DET and PERM:

Theorem 8 ([GKKS13])**.** *Any $\Sigma\Pi^{[O(\sqrt{d})]}\Sigma\Pi^{[\sqrt{d}]}$ circuit computing DET_n or PERM_n has top fanin $2^{\Omega(\sqrt{n})}$.*

1.3 The univariate case: the real τ-conjecture

Univariate polynomials are just a special case of multivariate polynomials, but Koiran showed in [Koi10] that results on univariate polynomials would yield strong lower bounds for the multivariate case. The main advantage of the univariate approach is that univariate polynomials are well-known objects, and one could hope to use some real analysis tools. Multivariate and univariate circuits are connected by a simple transformation: replace inputs $x_1, \ldots, x_n$ of a general arithmetic circuits by some power of X, eg $x_i = X^{2^i}$. Starting from $\Sigma\Pi\Sigma\Pi$ circuits, the transformation results in *sum of products of sparse polynomials* (SPS) of the form: $\sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}(X)$ with $f_{i,j}$ t-sparse. Theorem 6 of [Koi10] shows that under the assumption that $\mathsf{PERM} \in \mathsf{VP}$, polynomials with "reasonable" coefficients can be efficiently represented by sum of product of sparse polynomials.

Example 9. *The family of polynomials $g_n(X) = \prod_{i=1}^{2^n}(x - i)$, called the Pochhammer-Wilkinson polynomials of order 2^n, have "reasonable" coefficients. Hence, if $\mathsf{PERM} \in \mathsf{VP}$, we could write g_n as $\sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}(x)$, with $k = 2^{O(\sqrt{n}\log^2 n)}, m = O(\sqrt{n})$, and where the $f_{i,j}$'s are t-sparse polynomials, with $t = 2^{O(\sqrt{n}\log n)}$.*

The ultimate goal being to prove $\mathsf{PERM} \notin \mathsf{VP}$, it would suffice to exhibit an univariate polynomial that doesn't admit efficient any SPS representation. The difficult part is to prove this later statement for a given polynomial. In the case of the Pochhammer-Wilkinson polynomials, one can remark that they have a lot of integer roots, it would hence suffice to prove that SPS polynomials have a relatively small number of integer roots. This leads to the real τ-conjecture, first defined in [Koi10]:

Conjecture 10. *Consider a nonzero polynomial of the form:*

$$\sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}(X)$$

where each $f_{i,j}$ has at most t monomials. The number of real roots of f is bounded by a polynomial function of kmt.

This conjecture directly implies that $\mathsf{PERM} \notin \mathsf{VP}$, using Example 9. Indeed, SPS with parameters $k = 2^{\sqrt{n}}, m = \sqrt{n}, t \leq 2^{\sqrt{n}}$ would only have $2^{o(n)}$ roots, whereas Pocchammer-Wilkinson polynomial of order 2^n has exactly 2^n roots. Notice that even a bound polynomial in $kt2^m$ is enough to have the contradiction.

A first naive bound follows from Descartes' rule of signs:

Descartes' Rule of Signs. *Given a polynomial $f(X) = \sum_{i=0}^{d} a_i X^i$, with $a_i \in \mathbb{R}$, the number of positive roots of the polynomial is either equal to the number of sign differences between consecutive nonzero coefficients, or is less than it by an even number.*

Corollary 11. *Given a t-sparse polynomial f, the number of real roots of f is bounded by $2t - 1$.*

This corollary directly implies that the number of real roots of a sum of product of t-sparse polynomial is bounded by $2kt^m - 1$, since it has at most kt^m different monomials.

No better bound is known so far but some results have been found in [GKPS11, KPT12] for a similar model: sums of products of powers of sparse polynomials. During the internship, we studied another variant of this model: sums of powers of (sparse) polynomials. This isn't really a restriction since we can transform the product by a sum of power using the following formula [Fis94]:

$$x_1 \cdot \ldots \cdot x_d = \frac{1}{d!} \sum_{\varepsilon \in \{-1,1\}^{d-1}} (x_1 + \varepsilon_1 x_2 + \ldots + \varepsilon_{d-1} x_d)^d$$

Example 12. *We saw that if* $\mathsf{PERM} \in \mathsf{VP}$*, then the Pochhammer-Wilkinson polynomial of order* 2^n *could be written as* $\sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}(x)$*, with* $k = 2^{O(\sqrt{n}\log^2 n)}, m = O(\sqrt{n})$*, and where the* $f_{i,j}$*'s are* t*-sparse polynomials, with* $t = 2^{O(\sqrt{n}\log n)}$*. Now, using formula above, the Pochhammer-Wilkinson polynomial of order* 2^n *could be written as* $\sum_{i=1}^{k} (h_i(x))^{O(\sqrt{n})}$*, with* $k = 2^{O(\sqrt{n}\log^2 n)}$ *and* h_i*'s some* t*-sparse polynomials, with* $t = 2^{O(\sqrt{n}\log n)}$*.*

We could formulate another real τ-conjecture for sums of powers of polynomial: an upper bound polynomial in $kt2^{\alpha}$, where α is the maximal exponent of the expression, is enough to implies $\mathsf{VP} \neq \mathsf{VNP}$. In this internship, we investigate other ways than the number of real roots to show that some univariate polynomials don't have any efficient representation as a sum of powers, using two main tools: the Wronskian and the shifted derivatives space.

2 The model and the tools

The model is the univariate analog of the model investigated in [Kay12], so we consider representation of the form:

$$f(x) = Q_1(x)^{e_1} + \ldots + Q_s(x)^{e_s}$$

where the e_i's are positive integer, and the Q_i's are arbitrary univariate polynomials of bounded degree. This model is an interesting one because it is quite a simple model but a lot of problems still remain open. Better lower bounds for those problems may involve new techniques which can be used for other open problems of algebraic complexity.

When the Q_i's have degree t, any such representation of the random polynomial of degree d must satisfy $s \geq \frac{d+1}{t+1}$. Indeed, as one Q_i has $t+1$ coefficients, the model, for a fixed s, has $s \cdot (t+1)$ degrees of freedom. On the other side, the polynomials of degree d have $d+1$ degrees of freedom, so we know that there are some polynomials that requires s to be greater that $\frac{d+1}{t+1}$. However, we are looking for "explicit" polynomials.

The strongest result we proved for an explicit polynomial is the following bound:

Theorem 13. *For any* $d, t \geq 2$ *such that* $t < \frac{d}{4}$*, the polynomial* $f(x) = \sum_{i=1}^{m} (x - a_i)^d$*, with distinct* a_i*'s and* $m = \left\lfloor \sqrt{\frac{d}{t}} \right\rfloor$*, is hard in the following sense: any representation of* f *of the form* $f = \sum_{i=1}^{s} \alpha_i Q_i^{e_i}$*, with each* Q_i *of degree* $\leq t$*,* $\alpha_i \in \mathbb{F}$*, must satisfy* $s = \Omega\left(\sqrt{\frac{d}{t}}\right)$*.*

Even in the case $t = 2$, it is still an open problem to find a better bound than $\Omega\left(\sqrt{d}\right)$. Nevertheless, we proved the linear bound in the case $t = 1$, but all the exponents must be the same:

Theorem 14. *For any d, the polynomial $f(x) = \sum_{i=1}^{m}(x - a_i)^d$, with distinct a_i's and $m = \left\lfloor \frac{d}{2} \right\rfloor$, is optimally hard in the following sense: any representation of f of the form $f = \sum_{i=1}^{s} \alpha_i l_i^d$, with each l_i of degree 1, must satisfy $s \geq \left\lfloor \frac{d}{2} \right\rfloor$.*

In the next section, we will prove several lower bounds using two different tools, but they will both use the same key ingredient, summarized in this observation:

Observation 15. *Given a polynomial f of the form $f(x) = Q^e(x)$, with Q of degree t, we can factorize $f^{(k)}$ by some powers of Q: for any $i \leq k$, we have $f^{(k)} = Q^{e-i}R$, with $\deg R \leq it - k$.*

This describes the strong structure of the model we work on and we will use it to show that some polynomials can't be written efficiently with such a nice structure. We now define and describe the two tools we used, the Wronskian and the shifted derivatives space.

2.1 The Wronskian

In mathematics, the *Wronskian* is a tool mainly used in the study of differential equations, where it can be used to show that a set of solutions is linearly independent.

Definition 16 (Wronskian). *For n real functions $f_1, \ldots, f_n$, which are $n-1$ times differentiable, the Wronskian $\mathrm{W}(f_1, \ldots, f_n)$ is defined by*

$$\mathrm{W}(f_1, \ldots, f_n)(x) = \begin{vmatrix} f_1(x) & f_2(x) & \ldots & f_n(x) \\ f_1'(x) & f_2'(x) & \ldots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)} & f_2^{(n-1)} & \ldots & f_n^{(n-1)} \end{vmatrix}$$

We will use the following formulas about the Wronskian whose proofs can be found in [PS76] (and which are known since the 19^{th} century). For any $f_1, \ldots, f_k, g$ which are $k-1$ times differentiable, we have $\mathrm{W}(gf_1, \ldots, gf_k) = g^k \mathrm{W}(f_1, \ldots, f_k)$. As a corollary, we have the following formula:

$$\mathrm{W}(f_1, \ldots, f_k) = (f_1)^k \mathrm{W}\left(\left(\frac{f_2}{f_1}\right)', \ldots, \left(\frac{f_k}{f_1}\right)'\right) \tag{1}$$

The basic property of the Wronskian about linear independence is that for any linearly dependent functions $f_1, \ldots, f_n$, the Wronskian $\mathrm{W}(f_1, \ldots, f_n)$ vanishes everywhere. The converse is false in general, Peano then Bôcher found counterexamples (see [EP] for a history of these results). However, several conditions sufficient to ensure that the vanishing of the Wronskian everywhere implies linear dependence were found. For instance, Bôcher proved [Bô00] that if the f_i's are analytic, then the converse holds. We will adapt the proof of this result to make it work with functions in $\mathbb{C}(X)$ (the field of complex rational functions), and in particular with polynomials in $\mathbb{C}[X]$.

Proposition 17. *For $f_1, \ldots, f_n \in \mathbb{C}(X)$, the functions are linearly dependent if and only if the Wronskian $\mathrm{W}(f_1, \ldots, f_n)$ vanishes everywhere.*

Proof. The proof goes on by induction on n. The basic case $n = 1$ is trivial.
Suppose that $f_1, \ldots, f_n \in \mathbb{C}(X)$ has a null Wronskian. Then, using Formula 1, if we denote $g_i = \left(\frac{f_i}{f_1}\right)'$, the Wronskian $\mathrm{W}(g_2, \ldots, g_n)$ is also null (in fact, it's null except on a finite number of complex values, but since the Wronskian is also a rational function, this implies its nullity everywhere). By induction hypothesis, the g_i's are linearly dependent, i.e. there exists $a_2, \ldots, a_n$ such that:

$$a_2 g_2 + \ldots a_n g_n = 0$$

By integrating this equality, since $g_i = \left(\frac{f_i}{f_1}\right)'$, there exists a_1 such that:

$$a_1 + a_2 \frac{f_2}{f_1} + \ldots + a_n \frac{f_n}{f_1} = 0$$

But then by multiplying by f_1, we obtain:

$$a_1 f_1 + a_2 f_2 + \ldots + a_n f_n = 0$$

Thus the family (f_i) is linearly dependent. □

Example 18. *Using proposition above, we can show that, for any distinct a_i's in $\mathbb{C}$, the family $S = \{(x-a_1)^d, \ldots, (x-a_{d+1})^d\}$ is a basis of $\mathbb{C}_d[X]$, the vector space of polynomials of degree less or equal than d. Since $\dim \mathbb{C}_d[X] = d + 1 = |S|$, we only have to show that S is linearly independent. Consider the Wronskian of the polynomials in $|S|$:*

$$Wr(x) = \mathrm{W}\left((x-a_1)^d, \ldots, (x-a_{d+1})^d\right) = \begin{vmatrix} (x-a_1)^d & \ldots & (x-a_{d+1})^d \\ d(x-a_1)^{d-1} & \ldots & d(x-a_{d+1})^{d-1} \\ \vdots & \ddots & \vdots \\ d! & \ldots & d! \end{vmatrix}$$

It's enough to show that the Wronskian is not the null polynomial. In fact, we will show that it's a (non-zero) constant polynomial. For any $z \in \mathbb{C}$, define $b_i = z - a_i$ and we have:

$$Wr(z) = \begin{vmatrix} b_1^d & \ldots & b_{d+1}^d \\ d \cdot b_1^{d-1} & \ldots & d \cdot b_{d+1}^{d-1} \\ \vdots & \ddots & \vdots \\ d! & \ldots & d! \end{vmatrix} = c \cdot \begin{vmatrix} b_1^d & \ldots & b_{d+1}^d \\ b_1^{d-1} & \ldots & b_{d+1}^{d-1} \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{vmatrix}$$

for some non-zero $c \in \mathbb{N}^*$ which only depends on d.
The last matrix is a Vandermonde matrix, so its determinant is equal to the product $\prod_{i \neq j}(b_i - b_j) = \prod_{i \neq j}(a_j - a_i)$, which is a non-zero constant, since all a_i's are distinct. The determinant is hence non-zero and so we have $Wr(z) \neq 0$ for any $z \in \mathbb{C}$, thus the family S is linearly independent.

This tool was used first to establish a bound in [KPT12] for sums of products of powers of sparse polynomials. They used some results from [VP75] which give a link between polynomials of the form $F = \sum_{i=1}^{m} Q_i$ and the Wronskian $\mathrm{W}(Q_1, \ldots, Q_m)$ concerning their number of roots. This gives them a non-trivial upper bound on the number of roots of a sum of products of powers of sparse polynomials.

In our case, in order to prove lower bounds, we will use another result from [VP75] which gives a bound on the multiplicity of a root depending on the Wronskian:

Lemma 19. *Let $Q_1, \ldots, Q_m$ be some linearly independent polynomial and $z_0 \in \mathbb{C}$, and let $F(z) = \sum_{i=1}^{m} Q_i(z)$. Then:*

$$\mathrm{N}_{z_0}(F) \leq m - 1 + \mathrm{N}_{z_0}(\mathrm{W}(Q_1, \ldots, Q_m))$$

where $\mathrm{N}_{z_0}(\mathrm{W}(Q_1, \ldots, Q_m))$ is finite since $\mathrm{W}(Q_1, \ldots, Q_m) \not\equiv 0$.

Proof. By multilinearity of the determinant, we have: $\mathrm{W}(Q_1, \ldots, Q_m) = \mathrm{W}(Q_1, \ldots, Q_{m-1}, F)$. Hence:

$$\mathrm{W}(Q_1, \ldots, Q_{m-1}, F) = \sum_{i=0}^{m-1} B_i F^{(i)}$$

where the B_i's are some co-factor of the Wronskian matrix. Then:

$$\mathrm{N}_{z_0}(\mathrm{W}(Q_1, \ldots, Q_m)) = \mathrm{N}_{z_0}\left(\sum_{i=0}^{m-1} B_i F^{(i)}\right) \geq \mathrm{N}_{z_0}(F) - (m-1)$$

proving the lemma. □

Remark 20. *The assumption about the independence of the Q_i's is not that restrictive in the following sense. If the Q_i's are dependent, take a subset of them which form a basis: $\{Q_{i_1}, \ldots Q_{i_n}\}$. By rewriting the other Q_i's in this basis, we obtain something of the following form: $F(z) = \sum_{j=1}^{n} a_j Q_{i_j}(z)$, with $a_j \in \mathbb{R}$. We then extract a subfamily of the basis by keeping only the Q_{i_j}'s with $a_j \neq 0$: $\{Q_{k_1}, \ldots Q_{k_p}\}$. The polynomial F can now be written as $F(z) = \sum_{j=1}^{p} b_j Q_{k_j}(z)$, with $b_j \in \mathbb{R}^*$. Now, we can use the lemma with those linearly independent polynomial to obtain:*

$$\mathrm{N}_{z_0}(F) \leq p - 1 + \mathrm{N}_{z_0}\left(\mathrm{W}\left(b_1 Q_{k_1}, \ldots, b_p Q_{k_p}\right)\right)$$

We factorize the Wronskian by the b_i's to obtain:

$$\mathrm{N}_{z_0}(F) \leq p - 1 + \mathrm{N}_{z_0}\left(\mathrm{W}\left(Q_{k_1}, \ldots, Q_{k_p}\right)\right)$$

Thus the result holds even for linearly dependent Q_i's, but we have to take the Wronskian of some particular subfamilies of the Q_i's in this case.

This lemma allows us to "transform" the sum into a Wronskian. It's quite useful in our model, since we have to deal with sums of powers of polynomial. Indeed, we can't factorize a sum, but we can factorize the corresponding Wronskian, and this will give us a lower bound for specific polynomials.

2.2 The space of shifted derivatives

This tool is used in more "natural" proofs than the ones involving the Wronskian. "Natural" proof strategies follow the same plan (outlined in [KS]):

Step 1: (normal form) For every circuit in the circuit class C of interest, express the polynomial computed as a *small sum of simple building blocks.*

Step 2: (complexity measure) Build a map $\Gamma : \mathbb{F}[X] \to \mathbb{Z}^+$ that is *sub-additive*, i.e. $\Gamma(f+g) \leq \Gamma(f) + \Gamma(g)$.

Step 3: (potential usefulness) Show that if B is a simple building block, then $\Gamma(B)$ is small. Further, check if $\Gamma(f)$ for a random polynomial f is large.

Step 4: (explicit lower bound) Find an explicit polynomial f for which $\Gamma(f)$ is large.

These are usually the steps taken in most of the existing arithmetic circuit lower bound proofs. The hard part is to build a useful complexity measure, which is small on the building blocks but large for a random polynomial. In our case, our model consists of sums of power of polynomial, hence a simple block is a power of a polynomial.

Our complexity measure is inspired by the one first defined in [Kay12]: the space of shifted partial derivatives. Using this complexity measure, Kayal proved exponential lower bounds on a similar multivariate model. The key intuition follows from Observation 15: derivatives of Q^e of order $\leq k$ all share a large common factor, namely Q^{e-k}. We try to capture this property with the following complexity measure:

Definition 21 (Shifted derivatives space)**.** *Let $f(x) \in \mathbb{F}[x]$ be a polynomial. The* span of the l-shifted k-th order derivatives *of f, denoted by $\left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k}$, is defined as:*

$$\left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} \quad \stackrel{def}{=} \quad \mathbb{F}\text{-}span\left\{ x^j \cdot f^{(i)}(x) \quad : \quad i \leq k,\ j \leq i+l \right\}$$

$\left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k}$ forms an $\mathbb{F}$-vector space and we denote by $\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k}$ the dimension of this space.

Remark 22. *We have two trivial upper bounds on the dimension of the shifted derivatives space. First, for any polynomial f of degree d, the degree of any polynomial in $\left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k}$ is less than $d+l$, hence $\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} \leq d+l+1$. Second, the dimension is less or equal than the cardinality of a generating family, thus $\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} \leq \sum_{i=0}^{k}(l+i+1)$. Thus, we have:*

$$\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} \leq \min\left(d+l+1,\ \ (k+1)l + \binom{k+2}{2} \right)$$

We will see in the next section some polynomials that achieve those bounds and thus have a full shifted derivative space.

Notice that since $\left\langle x^{\leq i+l} \cdot (f+g)^{(i)} \right\rangle_{i \leq k} \subseteq \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} + \left\langle x^{\leq i+l} \cdot g^{(i)} \right\rangle_{i \leq k}$, the measure we defined is sub-additive. Now that we have defined the complexity measure, we have to show that in our model, polynomials have a small complexity according to this measure:

Proposition 23. *For any polynomial f of degree d of the form $f = \sum_{i=1}^{s} \alpha_i Q_i^{e_i}$, with $\deg Q_i \leq t$ we have:*

$$\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} \leq s \cdot (l + kt + 1)$$

Proof. Since the measure is sub-additive, we only have to show that for a simple building block f of the form Q^{e_i}, with $\deg Q \leq t$, we have $\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} \leq l + kt + 1$. To do so, we use observation 15: any $g \in \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k}$ is of the form $g = Q^{e_i - k} \cdot R$, with $\deg R \leq l + kt$, since $\deg g \leq e_i \cdot t + l$. This directly gives the bound on the dimension. □

In the last part of the proof, in the next section, we will give an explicit lower bound on the dimension of shifted derivatives space of some explicit polynomial.

3 Lower bounds for sums of powers of polynomials

In this section, we give proofs of several lower bounds on the model for the same type of polynomials: $f(x) = \sum_{i=1}^{m} (x - a_i)^d$, for distinct a_i's and for some well chosen parameter m. Of course, there is a trivial upper bound on the number s of summands needed to express f as a sum of powers, since f is already in this form. Most of the result proved next will show that this is the "optimal" representation for f: if we have another representation of f as a sum of s powers, then m and s must have the same order.

The results differs in several ways: we will work with constant t or not, and we will allow the exponents to be unbounded or not. This later restriction is an important one since allowing unbounded exponents means there can be cancellations of terms of too high degree. We know that cancellations can add power to a model in general, for instance the most efficient circuits computing the determinant family use them, and we still don't know if it's the case with this model.

First, we will see the case $t = 2$ to provide some intuition on how the Wronskian works to prove lower bounds. Then, we will try to adapt the proof for the general case and it will give two lower bounds whether the exponents are unbounded or not. Next, we will study the case $t = 1$, with the only linear lower bound we known, to see what it's different and may allow us to improve lower bounds for general case. Finally, we will show the best bound we have found, using shifted derivatives.

3.1 Sums of powers of quadratic polynomials

As we will prove later, we have the optimal bound for $t = 1$ but already for $t = 2$, we only have a lower bound in $\Omega\left(\sqrt{d}\right)$. However, the proof for $t = 1$ only work with all exponents equal to d, whereas this bound holds even for unbounded exponents.

Theorem 24. *For any d, the polynomial $f(x) = \sum_{i=1}^{m} (x - a_i)^d$ with distinct a_i's and $m = \left\lfloor \frac{\sqrt{d}}{2} \right\rfloor$ is hard in the following sense: any representation of f of the form $f = \sum_{i=1}^{s} \alpha_i Q_i^{e_i}$, with each Q_i's of degree ≤ 2, must satisfy $s = \Omega\left(\sqrt{d}\right)$.*

Proof. For contradiction, assume that $f = \sum_{i=1}^{s} \alpha_i Q_i^{e_i}$ with $s < \frac{m}{2}$. Since we have $\deg(Q_i) \leq 2$, the Q_i's have at most $2s$ roots, and thus the inequality $s < \frac{m}{2}$ implies

that one $(x-a_i)$ does not divide any Q_j, without loss of generality $(x-a_1)$. We set $l = s+m-1$ and define for $i > s$:

$$\begin{cases} \alpha_i & = -1 \\ Q_i(x) & = (x - a_{i-s+1}) \\ e_i & = d \end{cases}$$

so that: $(x-a_1)^d = \sum_{i=1}^{l} \alpha_i Q_i^{e_i}(x)$.
Now, using lemma 19, for a certain subfamily $Q_{i_1}^{e_{i_1}}, \ldots, Q_{i_p}^{e_{i_p}}$ of the $Q_i^{e_i}$'s, we obtain:

$$d = \mathrm{N}_{a_1}\left((x-a_1)^d\right) \leq p - 1 + \mathrm{N}_{a_1}\left(\mathrm{W}\left(Q_{i_1}^{e_{i_1}}, \ldots, Q_{i_p}^{e_{i_p}}\right)\right) \tag{2}$$

We denote the Wronskian $\mathrm{W}\left(Q_{i_1}^{e_{i_1}}, \ldots, Q_{i_p}^{e_{i_p}}\right)$ simply by Wr and, using Observation (15), we factorize it by $Q_{i_j}^{e_{i_j}-(p-1)}$ for every j such that $e_j > p-1$. Since $\mathrm{N}_{a_1}(Q_i) = 0$ holds for any i, we obtain:

$$\mathrm{N}_{a_1}(\mathrm{Wr}) = \mathrm{N}_{a_1}\begin{vmatrix} R_{1,1} & \cdots & R_{1,p} \\ \vdots & \ddots & \vdots \\ R_{p,1} & \cdots & R_{p,p} \end{vmatrix} \quad \text{with} \quad \deg(R_{i,j}) \leq 2(p-1) - (i-1)$$

When e_{i_j} is smaller than $p-1$, the corresponding column is unchanged, and thus we even have $\deg(R_{i,j}) \leq e_j - (i-1) \leq p-1-(i-1)$. For indices i such that $Q_i(x) = (x - a_{i-k+1})$, we also have this same tighter bound because in this case we have $\deg(Q_i) = 1$.
The determinant above has degree $\leq 2p(p-1) - \binom{p}{2} = \frac{3}{2}p(p-1)$. Inequality (2) becomes:

$$d \;\leq\; \frac{3p^2}{2} - \frac{p}{2} - 1 \;\text{ hence }\; d \leq \frac{3p^2}{2} \leq \frac{3l^2}{2}$$

Now, using the assumption that $s < \frac{m}{2}$, we have $l^2 < \frac{9}{4}m^2$, and we finally obtain the contradiction:

$$d \;<\; \frac{27}{8}m^2 \quad \text{implies} \quad m \;>\; \frac{2\sqrt{2}}{3\sqrt{3}}\sqrt{d} \;>\; \frac{\sqrt{d}}{2}$$

□

3.2 The general case: unbounded exponents

When one tries to adapt directly the previous proof for the general case, it ends up with a weaker lower bound which is still interesting because it still makes no assumption about the value of the e_i's:

Theorem 25. *For any $t \geq 2, d$, the polynomial $f(x) = \sum_{i=1}^{m}(x-a_i)^d$, with distinct a_i's and $m = \left\lfloor \frac{2}{3}\sqrt{\frac{d}{t}} \right\rfloor$ is hard in the following sense: any representation of f of the form $f = \sum_{i=1}^{s} \alpha_i Q_i^{e_i}$, with each Q_i of degree $\leq t$ and $\alpha_i \in \mathbb{R}$, must satisfy $s = \Omega\left(\frac{1}{t}\sqrt{\frac{d}{t}}\right)$.*

Proof. For the sake of contradiction, assume that $f = \sum_{i=1}^{s} \alpha_i Q_i^{e_i}$ with $s < \frac{m}{t}$. Again, without loss of generality, $(x - a_1)$ does not divide any Q_i. We set $l = s + m - 1$ and define for $i > s$:

$$\begin{cases} \alpha_i & = -1 \\ Q_i(x) & = (x - a_{i-k+1}) \\ e_i & = d \end{cases}$$

so that: $(x - a_1)^d = \sum_{i=1}^{l} \alpha_i Q_i^{e_i}(x)$.
Now, using lemma 19, for a certain subfamily S of size p of the $Q_i^{e_i}$'s, we obtain:

$$d = \mathrm{N}_{a_1}\left((x - a_1)^d\right) \leq p - 1 + \mathrm{N}_{a_1}\left(\mathrm{W}\left(S\right)\right) \tag{3}$$

We now factorize it by $Q_{i_j}^{e_{i_j}-(p-1)}$ for every j such that $e_j > p - 1$. Since $\mathrm{N}_{a_1}(Q_i) = 0$ holds for any i, we obtain:

$$\mathrm{N}_{a_1}(\mathrm{Wr}) = \mathrm{N}_{a_1} \begin{vmatrix} R_{1,1} & \dots & R_{1,p} \\ \vdots & \ddots & \vdots \\ R_{p,1} & \cdots & R_{p,p} \end{vmatrix} \quad \text{with} \quad \deg(R_{i,j}) \leq t(p-1) - (i-1)$$

The determinant above has degree $\leq tp(p-1) - \binom{p}{2} = (t - 1/2)p(p-1)$ The inequality (3) becomes:

$$d \;\leq\; \left(t - \frac{1}{2}\right)p^2 - \left(t - \frac{3}{2}\right)p - 1$$

We drop the negative terms to obtain: $d \;\leq\; tp^2$. Now, using the assumption that $s < \frac{m}{t}$, we have $p^2 \leq l^2 < \left(1 + \frac{1}{t}\right)^2 m^2$, and we obtain:

$$d \;<\; t\left(1 + \frac{1}{t}\right)^2 m^2$$

We finally obtain the contradiction, since $t \geq 2$:

$$m \;>\; \frac{1}{1 + \frac{1}{t}}\sqrt{\frac{d}{t}} \;\geq\; \frac{2}{3}\sqrt{\frac{d}{t}}$$

This implies that:

$$s \geq \frac{1}{t}\left\lfloor \frac{2}{3}\sqrt{\frac{d}{t}} \right\rfloor = \Omega\left(\frac{1}{t}\sqrt{\frac{d}{t}}\right)$$

□

3.3 The general case: bounded exponents

In this section we will prove a stronger lower bound by weakening the hypothesis $s < \frac{m}{t}$. We reach the bound $\Omega\left(\sqrt{\frac{d}{t}}\right)$ but we need to have a bound on the e_i's, which means we don't allow cancellation.

Theorem 26. *For any t, d, the polynomial $f(x) = \sum_{i=1}^{m}(x-a_i)^d$, with distinct a_i's and $m = \left\lfloor \frac{\sqrt{2}}{3}\sqrt{\frac{d}{t}} \right\rfloor$, is hard in the following sense : any representation of f of the form $f = \sum_{i=1}^{s} \alpha_i Q_i^{e_i}$, with each Q_i of degree $\leq t$, $e_i \leq d/t$ and $\alpha_i \in \mathbb{C}$, must satisfy $s = \Omega\left(\sqrt{\frac{d}{t}}\right)$.*

Proof. Fix a constant $c \in \mathbb{R}^+$ which will be later set to $\frac{1}{2}$. Assume for contradiction that $s < c \cdot m$. We can't say any longer that one $(x-a_i)$ divides no Q_i's, but since the Q_i's have at most $st < c \cdot mt$ roots, a weaker statement still holds: if we denote by $\mu_{i,j}$ the multiplicity of a_i in Q_j, and define $\mu_i = \sum_j \mu_{i,j}$, there exists a_i such that $\mu_i < ct$, without loss of generality a_1. As before, we set $l = s + m - 1$ and define for $i > s$:

$$\begin{cases} \alpha_i & = -1 \\ Q_i(x) & = (x - a_{i-k+1}) \\ e_i & = d \end{cases}$$

so that: $(x-a_1)^d = \sum_{i=1}^{l} \alpha_i Q_i^{e_i}(x)$.
Now, using lemma 19, for a certain subfamily S of size p of the $Q_i^{e_i}$'s, we obtain:

$$d = \mathrm{N}_{a_1}\left((x-a_1)^d\right) \leq p - 1 + \mathrm{N}_{a_1}\left(\mathrm{W}\left(S\right)\right) \tag{4}$$

Now, when we factorize by $Q_{i_j}^{e_{i_j}-(p-1)}$, we add $\mu_{i_j,1}(e_{i_j}-(p-1))$ to the multiplicity. When we sum up, by hypothesis on a_1, we obtain a new term which is smaller than $ct(d/t-(p-1))$. Thus we have the following bound on the multiplicity of a_1 in the Wronskian:

$$\mathrm{N}_{a_1}\left(\mathrm{Wr}\right) \leq ct\left(\frac{d}{t} - (p-1)\right) + \mathrm{N}_{a_1}\begin{vmatrix} R_{1,1} & \dots & R_{1,p} \\ \vdots & \ddots & \vdots \\ R_{p,1} & \cdots & R_{p,p} \end{vmatrix}$$

with $\deg(R_{i,j}) \leq t(p-1) - (i-1)$.
The degree of this determinant is bounded by $tp(p-1) - \binom{p}{2} = (t-1/2)p(p-1)$. Equation (4) hence gives:

$$d \leq p - 1 + ct\left(\frac{d}{t} - (p-1)\right) + \left(t - \frac{1}{2}\right)p(p-1)$$

$$(1-c)\cdot d \leq \left(t-\frac{1}{2}\right)p^2 - \left(t-\frac{3}{2}\right)p - ct(p-1) - 1$$

We drop the negative terms to obtain:

$$(1-c)\cdot d \leq tp^2$$

Using the fact that $s < cm$, we have $p < (1+c)m$, which gives:

$$(1-c)\cdot d < t(1+c)^2 m^2$$

$$m^2 > \frac{1-c}{(1+c)^2}\cdot\frac{d}{t}$$

Taking $c = \frac{1}{2}$, we finally have the contradiction:

$$m^2 > \frac{2}{9} \cdot \frac{d}{t}$$

This implies that:

$$s \geq \frac{1}{3\sqrt{2}} \sqrt{\frac{d}{t}} = \Omega\left(\sqrt{\frac{d}{t}}\right)$$

□

3.4 Linear bound for degree 1

In the case $t = 1$, we have an optimal linear lower bound for the same type of polynomial. However, the exponents must all be equal to d to obtain the bound. We give here the proof of theorem 14, restated here:

Theorem 14. *For any d, the polynomial $f(x) = \sum_{i=1}^{m}(x - a_i)^d$, with distinct a_i's and $m = \left\lfloor \frac{d}{2} \right\rfloor$, is optimally hard in the following sense: any representation of f of the form $f = \sum_{i=1}^{s} \alpha_i l_i^d$, with each l_i of degree 1, must satisfy $s \geq \left\lfloor \frac{d}{2} \right\rfloor$.*

Proof. Assume for contradiction that there exists a representation of f of the form $\sum_{i=1}^{s} \alpha_i l_i^d$ with $s < m$. As before, we set $n = s + m$ and define for $i > s$:

$$\begin{cases} \alpha_i & = -1 \\ l_i(x) & = (x - a_{i-k+1}) \end{cases}$$

so that $\sum_{i=1}^{n} \alpha_i l_i^d = 0$.
Some l_i's could be the same, so we rewrite the sum as $\sum_{j=1}^{p} \beta_{i_j} l_{i_j}^d = 0$, where all l_{i_j}'s are different. Since $s < m$, there is at least one β_{i_j} which is non-zero, and this means that the family $(l_{i_j}^d)$ is linearly dependent. However, since $m \leq \frac{d}{2}$, we have $p \leq n \leq d$, and using Example 18, the family is hence linearly independent. This gives the contradiction. □

3.5 The lower bound using shifted derivatives

In this section, we will prove theorem 13 using shifted derivatives. The proof will consist in a lower bound on the dimension shifted derivatives space of the polynomials of the form $f(x) = \sum_{i=1}^{m}(x - a_i)^d$. To do so, we will show that f does not satisfy a particular kind of differential equations, under some conditions.

Definition 27. Shifted Differential Equations *(SDE) are a particular kind of differential equations of the form*

$$\sum_{i=0}^{k} P_i(x) f^{(i)}(x) = 0$$

for some polynomials $P_i \in \mathbb{C}[X]$ with $\deg P_i \leq i + l$.
The quantity k is called the order *and the quantity l is called the* shift.

This kind of differential equations is directly linked with the notion of shifted derivatives:

Proposition 28. *For any $f \in \mathbb{R}[X]$, if f doesn't satisfy any SDE of order k and of shift l, then $\left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k}$ is full, ie:*

$$\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} = \sum_{i=0}^{k} (l+i+1) = (k+1)l + \binom{k+2}{2}$$

In order to prove some conditions on the SDE satisfied by f, we first need to prove that the polynomials $(x-a_1)^d, \ldots, (x-a_m)^d$ cannot satisfy simultaneously a SDE if the order is not big enough:

Lemma 29. *For any $d, m \leq d$, for any distinct $(a_i) \in \mathbb{R}^m$, the following property holds for the family $S = \{(x-a_1)^d, \ldots, (x-a_m)^d\}$: if a SDE is satisfied by every polynomial in S, then the order of the SDE must be greater than m.*

Proof. Assume the family $S = \{(x-a_1)^d, \ldots, (x-a_m)^d\}$ satisfies the following SDE, with $k < m$:

$$\sum_{i=0}^{k} P_i(x) f^{(i)}(x) = 0 \tag{5}$$

We can factorize $(x-a_i)^{d-k}$ for any i in a same manner to obtain a new SDE satisfied by the family $S' = \{(x-a_1)^k, \ldots, (x-a_m)^k\}$:

$$\sum_{i=0}^{k} Q_i(x) f^{(i)}(x) = 0 \tag{6}$$

with $Q_i(x) = \frac{d!}{k!} \frac{(k-i)!}{(d-i)!} P_i(x)$.

But now, since $k < m$, the family S generate $\mathbb{R}_k[X]$, and thus this imply that every polynomial of degree $\leq k$ should satisfy the SDE (6). We obtain the contradiction by taking x^{i_0}, where i_0 is the smallest integer such that $Q_{i_0}(x) \not\equiv 0$. □

We can now prove the lower bound on the parameters of a SDE that f could satisfied, which will directly give the result.

Lemma 30. *For any $d, m \leq d$, for any distinct $(a_i) \in \mathbb{R}^m$, if the polynomial $f(x) = \sum_{i=1}^{m} (x-a_i)^d$ satisfies a SDE of parameters k, l then at least one of the two following conditions holds:*

i) $k \geq m$

ii) $l > \frac{d}{m} - 2m$

Proof. We will prove the result by showing that if f satisfies a SDE and i) doesn't hold, then ii) must hold. Assume that f satisfies a differential equation of the following form:

$$\sum_{i=0}^{k} P_i(x) f^{(i)}(x) = 0 \tag{7}$$

with $k < m$ and $\deg(P_i) \leq i + l$.
For every i, we denote by Q_i the unique polynomial such that:

$$\sum_{j=0}^{k} P_j(x) \left((x-a_i)^d\right)^{(i)}(x) = Q_i(x)(x-a_i)^{d-k}$$

Notice that Q_i is of degree at most $k+l$. Using lemma 29, since $k < m$, not all Q_i's can be 0, without loss of generality we have $Q_1 \not\equiv 0$. We set $f_i(x) = Q_i(x)(x-a_i)^{d-k}$ and, using linearity of differentiation, we rewrite differential equation (7) as:

$$-f_1(x) = \sum_{i=2}^{m} f_i(x)$$

Using Lemma 19, for a certain subset $I = \{i_1, \ldots, i_p\} \subseteq [\![2; m]\!]$, we obtain

$$d - k \leq \mathrm{N}_{a_1}(f_1) \leq p - 1 + \mathrm{N}_{a_1}(\mathrm{W}((f_i)_{i\in I})) \tag{8}$$

We can factorize the Wronskian by $(x-a_i)^{d-k-(p-1)}$ for any $i \in I$:

$$\mathrm{N}_{a_1}(\mathrm{Wr}) = \mathrm{N}_{a_1} \begin{vmatrix} R_{1,1} & \cdots & R_{1,p} \\ \vdots & \ddots & \vdots \\ R_{p,1} & \cdots & R_{p,p} \end{vmatrix}$$

with $\deg(R_{i,j}) \leq l + k + p - i$.
The determinant has degree $\leq p(l+k) + \binom{p}{2}$. Hence, inequality (8) becomes:

$$d - k \leq p - 1 + p(l+k) + \binom{p}{2}$$

Using the fact that $p \leq m-1$, we obtain:

$$d \leq (m-1)\cdot l + m \cdot k + \frac{(m-2)(m+1)}{2}$$

Divide by m and drop negatives terms to obtain:

$$\frac{d}{m} \leq l + k + \frac{m}{2}$$

Using the hypothesis that $k < m$, we finally have:

$$l > \frac{d}{m} - \frac{3}{2}m$$

□

We can now prove again the main result, restated here:

Theorem 13. *For any $d, t \geq 2$ such that $t < \frac{d}{4}$, the polynomial $f(x) = \sum_{i=1}^{m}(x-a_i)^d$, with distinct a_i's and $m = \left\lfloor \sqrt{\frac{d}{t}} \right\rfloor$, is hard in the following sense: any representation of f of the form $f = \sum_{i=1}^{s} \alpha_i Q_i^{e_i}$, with each Q_i of degree $\leq t$, $\alpha_i \in \mathbb{F}$, must satisfy $s = \Omega\left(\sqrt{\frac{d}{t}}\right)$.*

Proof. We take k and l small enough to ensure that f doesn't satisfy any SDE of parameters k and l. Using lemma 30, it is enough to take:

- $k = m - 1 = \left\lfloor \sqrt{\frac{d}{t}} \right\rfloor - 1$ so that $k < m$
- $l = \left\lfloor \sqrt{dt} - \frac{3}{2}\sqrt{\frac{d}{t}} \right\rfloor$ so that $l \leq \frac{d}{m} - \frac{3}{2}m$

Using proposition 28, we thus establish a lower bound on the dimension of the shifted derivatives space:

$$\begin{aligned}
\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} &= (k+1)l + \binom{k+2}{2} \\
&\geq \left(\sqrt{\frac{d}{t}} - 1\right)\left(\sqrt{dt} - \frac{3}{2}\sqrt{\frac{d}{t}} - 1\right) + \frac{1}{2}\left(\sqrt{\frac{d}{t}}\right)^2 \\
&= d\left(1 - \frac{1}{t} - \sqrt{\frac{t}{d}} + \frac{1}{2\sqrt{dt}} + \frac{1}{d}\right) \\
&\geq d\left(1 - \frac{1}{t} - \sqrt{\frac{t}{d}}\right)
\end{aligned}$$

Now, assume that $f = \sum_{i=1}^{s} \alpha_i Q_i^{e_i}$, for some Q_i's with $\deg Q_i \leq t$. Proposition 23 gives the following upper bound on the dimension:

$$\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} \leq s \cdot (l + kt + 1) \leq s \cdot 2\sqrt{dt}$$

Hence:

$$s \geq \frac{1 - \frac{1}{t} - \sqrt{\frac{t}{d}}}{2} \cdot \frac{d}{\sqrt{dt}}$$

Now, since $t < \frac{d}{4}$, we have $\sqrt{\frac{t}{d}} < \frac{1}{2}$ and thus:

$$s = \Omega\left(\sqrt{\frac{d}{t}}\right)$$

□

4 Discussion

In the previous section, we presented several proofs of lower bounds for the model of sums of powers of univariate polynomials. The best one we proved in Theorem 13 was established using shifted derivatives, and it is a stronger lower bound than the ones obtained using the Wronskian alone in Theorems 24, 25 and 26. However, the proofs involving the Wronskian alone are still interesting because it seems that we cannot improve the lower bound with the shifted derivatives. On the other hand, we propose one possible way to improve lower bounds using the Wronskian alone.

4.1 Multiplicity of the Wronskian

The proofs involving Wronskian followed the same pattern:

Step 1: Isolate a polynomial which has a root with high multiplicity

Step 2: Use Lemma 19 to "transform" the sum into a Wronskian

Step 3: Factorize the Wronskian by powers of polynomials

Step 4: Bound the multiplicity of the resulting determinant by its degree

The step that costs the most is the last one since we have to deal with a determinant of $k \times k$ matrix, with polynomials of degree $O(tk)$ in it, so the degree of the determinant will always be in $O(tk^2)$. And since the multiplicity cannot be greater that d, we will always have something like $d \leq c \cdot tk^2$, which gives the already known lower bound $\Omega\left(\sqrt{\frac{d}{t}}\right)$. For instance, if one could show that the multiplicity is at most $O(tk)$ (the degree of the polynomials in the determinant) instead of $O(tk^2)$, it will directly gives the desired linear lower bound.

4.2 Limitation of the method of shifted derivatives

The best lower bound we get is using shifted derivatives. However, it is probably not possible to improve it using this tool. Indeed, as we saw in Remark 22, the trivial upper bound on the dimension is:

$$\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} \leq \min \left(d+l+1, \;\; (k+1)l + \binom{k+2}{2} \right)$$

This upper bound is tight since we have proved in Section 3.5 that the equality holds for the polynomials of the form $f(x) = \sum_{i=1}^{m} (x - a_i)^d$. On the other hand, Proposition 23 states the upper bound on the model:

$$\dim \left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k} \leq s \cdot (l + kt + 1)$$

Thus, the best bound possible on the number of summands is:

$$s \geq \frac{\min \left(d+l+1, \;\; (k+1)l + \binom{k+2}{2} \right)}{l + kt + 1}$$

The choice of parameters $l = O(\sqrt{dt})$ and $k = O\left(\sqrt{\frac{d}{t}}\right)$ gives the lower bound of Theorem 13, $s = \Omega\left(\sqrt{\frac{d}{t}}\right)$, and we claim that this is optimal with this tool.

Proof. Denote

$$f(k,l) = min \left(\frac{d+l+1}{l+kt+1}, \;\; \frac{(k+1)l + \binom{k+2}{2}}{l+kt+1} \right)$$

We want to show that for any k, l, $f(k,l) \leq \sqrt{\frac{d}{t}} + 1$. To do so, we work with fixed k, and we differentiate two cases :

- $k \geq \sqrt{\frac{d}{t}}$: the function $g : l \mapsto \frac{d+l+1}{l+kt+1}$ is monotonous on $[0; +\infty[$. Moreover we have $g(+\infty) = 1$ and $g(0) = \frac{d+1}{kt+1} < \sqrt{\frac{d}{t}} + 1$. Thus $g(l) < \sqrt{\frac{d}{t}} + 1$ for any l.
- $k \leq \sqrt{\frac{d}{t}}$: the function $h : l \mapsto \frac{(k+1)l+\binom{k+2}{2}}{l+kt+1}$ is monotonous on $[0; +\infty[$. Moreover we have $h(+\infty) = k + 1 \leq \sqrt{\frac{d}{t}} + 1$ and $h(0) = \frac{(k+2)(k+1)}{2(kt+1)} \leq \frac{k+2}{t} < \sqrt{\frac{d}{t}}$. Thus $h(l) \leq \sqrt{\frac{d}{t}} + 1$ for any l.

In both cases, $f(k, l) \leq \sqrt{\frac{d}{t}} + 1$, hence, for any l, k, $f(k, l) \leq \sqrt{\frac{d}{t}} + 1$.

□

4.3 Conclusion

Even simple models like sums of powers of univariate polynomial aren't still very well understood. Better lower bounds for the related problems may involve new techniques which can be used for other open problems of algebraic complexity.
The Wronskian is a tool mainly used in the study of differential equations but some of its properties are interesting for some problems of algebraic complexity. A better understanding of this object, for instance of its roots, may provide new and stronger lower bounds on those sums of powers of univariate polynomials, which might be transposed to the multivariate case.

Acknowledgments

Neeraj Kayal introduced the model of sums of powers of univariate polynomial and its related problem to us. He also showed us how the method of using shifted derivatives and explained what are the limitation of this tool.

References

[AV08] Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. *In Foundations of Computer Science (FOCS)*, pages 67–75, 2008.

[BCS97] P. Bürgisser, M. Clausen, and M. A. Shokrollahi. *Algebraic Complexity Theory*. Springer, 1997.

[Bô00] M. Bôcher. The theory of linear dependence. *The Annals of Mathematics*, 1900.

[Bü00] P. Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*. Springer, 2000.

[EP] S.M. Engdahl and A.E. Parker. Peano on wronskians: A translation. http://www.maa.org/publications/periodicals/convergence/peano-on-wronskians-a-translation-introduction.

[Fis94] I. Fischer. *Mathematics Magazine*, volume 67, chapter Sums of like powers of multivariate linear forms, pages 59–61. 1994.

[FLMS13] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinievasan. Lower bounds for depth 4 formulas computing iterated matrix multiplication. *Electronic Colloquium on Computational Complexity (ECCC)*, 20(100), 2013.

[GK98] Dima Grigoriev and Marek Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In *Symposium on Theory of Computing (STOC)*, pages 577–582, 1998.

[GKKS13] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. *Conference on Computational Complexity (CCC)*, 2013.

[GKPS11] Bruno Grenet, Pascal Koiran, Natacha Portier, and Yann Strozecki. The limited power of powering: Polynomial identity testing and a depth-four lower bound for the permanent. *CoRR*, abs/1107.1434, 2011. FSTTCS'11.

[Kay12] Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 2012.

[Koi10] Pascal Koiran. Shallow circuits with high-powered inputs. *CoRR*, abs/1004.4960, 2010. Innovations in Computer Science.

[Koi12] Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theoretical Computer Science*, 448:56–65, 2012.

[KPT12] Pascal Koiran, Natacha Portier, and Sébastien Tavenas. A wronskian approach to the real τ-conjecture. *CoRR*, abs/1205.1015, 2012. MEGA'13 (to appear in Journal of Symbolic Computation).

[KS] Neeraj Kayal and Ramprasad Saptharishi. A selection of lower bounds for arithmetic circuits (survey). To appear in a special issue in the event of Somenath Biswas' 60th Birthday.

[KSS13] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. *Electronic Colloquium on Computational Complexity (ECCC)*, 20(91), 2013.

[PS76] G. Polya and G. Szego. *Problems and Theorems in Analysis*, volume Volume II. Springer, 1976.

[Str73] V. Strassen. Vermeidung von divisionen. *J. reine u. angew. Math*, 264:182–202, 1973.

[Tav13] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. *Mathematical Foundations of Computer Science (MFCS)*, pages 813–824, 2013.

[Val79] L. G. Valiant. Completeness classes in algebra. In *Proceedings of the 11th Annual STOC*, pages 249–261, 1979.

[VP75] M Voorhoeve and A.J Van Der Poorten. Wronskian determinants and the zeros of certain functions. *Indagationes Mathematicae (Proceedings)*, 78(5):417 – 424, 1975.

[VSBR83] L. G. Valiant, S. Skyum, S. Berkowitz, and C. Rackoff. Fast parallel computation of polynomials using few processors. *SIAM*, 12, 1983.