Lower Bounds for sums of powers of low degree univariates.

Neeraj Kayal
Microsoft Research India
neeraka@microsoft.com

Pascal Koiran
Ecole Normale Supérieure de Lyon
Pascal.Koiran@ens-lyon.fr

Timothée Pecatte
Ecole Normale Supérieure de Lyon
timothee.pecatte@ens-lyon.fr

Chandan Saha
Indian Institute of Science
chandan@csa.iisc.ernet.in

April 29, 2015

Abstract

We consider the problem of representing a univariate polynomial $f(x)$ as a sum of powers of low degree polynomials. We prove a lower bound of $\Omega\left(\sqrt{\frac{d}{t}}\right)$ for writing an explicit univariate degree-d polynomial $f(x)$ as a sum of powers of degree-t polynomials.

1 Introduction

Valiant [Val79], defined the classes VP and VNP as the algebraic analogs of the classes P and NP. Informally, VP consists of (families of) efficiently computable (low-degree, multivariate) polynomials while VNP consists of (families of) explicit (low-degree, multivariate) polynomials. The problem of separating VNP from VP has since been one of the most important open problems in arithmetic complexity. Another basic question in complexity in general is whether computation can be efficiently parallelized. A seminal work by [VSBR83] showed that computation of low degree polynomials can indeed be efficiently parallelized - any *small* arithmetic circuit C computing a *low degree* multivariate polynomial $f(\mathbf{x})$ can be transformed to obtain another circuit C' of *low* depth and whose size is *not too large* computing the same polynomial $f(\mathbf{x})$. Subsequent refinements and improvements were obtained in a series of works [AJMV98, AV08, Koi12, Tav13, GKKS13b]. This line of work in particular yields the following depth reduction result which shows that if a polynomial can be efficiently computed then it has a *not too large* representation as a sum of powers of low degree polynomials. Specifically:

Proposition 1. (Implicit in [Tav13] and [GKKS13b]). *Let $\{f_n(\mathbf{x}) : n \geq 1\}$ be a family of n-variate polynomials of degree $d = d(n)$ over an underlying field $\mathbb{F}$ which is*

algebraically closed and has characteristic zero. If this family is in VP *then* $f_n(\mathbf{x})$ *admits a representation of the form*

$$f_n(\mathbf{x}) = \sum_{i=1}^{s} Q_i(\mathbf{x})^{e_i} \quad \text{where } \deg(Q_i) \leq \sqrt{d} \tag{1}$$

and where the number of summands s *is at most* $n^{O(\sqrt{d})}$.

Strong enough lower bounds for sums of powers imply general circuit lower bounds. These depth reduction results also provide a potential approach towards the VP versus VNP problem – via proving strong enough lower bounds for low depth circuits. In particular, the contrapositive version of Proposition 1 means that a strong enough (at least $n^{\omega(\sqrt{d})}$) lower bound for representing an explicit family of polynomials $\{f_n(\mathbf{x}) : n \geq 1\}$ in the form (1) above will imply that this family is not in VP, thereby separating VP and VNP. Promising progress along this direction has recently been obtained. [Kay12] considered representations of the form (1) above and introduced a complexity measure called *dimension of shifted partials* and obtained a $2^{\Omega(\sqrt{d})}$ lower bound for representations of the form (1) above. Follow-up work [GKKS13a, KSS14] obtained an $n^{\Omega(\sqrt{d})}$ lower bound for such representations, thereby coming tantalizingly close to the threshold required for obtaining superpolynomial lower bounds for general circuits. Since then, these techniques have been intensely investigated and followup work by [FLMS14, KS14b, KS14a] have used these techniques to obtain optimality of the known depth reduction results in many interesting cases. Some of these follow-up works also suggest that the dimension of shifted partials in itself might not be strong enough to separate VP from VNP. Further work [KLSS14, KS14c, KS14a] has suitably adapted and generalized the complexity measure to obtain lower bounds for more subclasses of arithmetic circuits.

Univariate sums of powers. Motivated by proposition 1, we introduce and study the problem of representing a *univariate* polynomial as a sum of powers of low-degree polynomials.

Definition 1. *Let* $t \geq 1$ *be an integer. For a polynomial* $f(x) \in \mathbb{F}[x]$*, define the sum of degree-*t*-powers complexity of* f*, denoted* $s_t(f)$*, as the smallest integer* s *such that* f *can be written as*

$$f(x) = \sum_{i=1}^{s} \alpha_i \cdot Q_i(x)^{e_i}, \quad \text{where } \forall i : \alpha_i \in \mathbb{F}, \deg(Q_i) \leq t.$$

We remark here that if the underlying field $\mathbb{F}$ is algebraically closed, we can assume without loss of generality that each scalar $\alpha_i = 1$. We seek to exhibit explicit polynomials $f(x)$ for which $s_t(f)$ is as large as possible. The motivation for this study is that univariate polynomials being much more well-known and easier to study than multivariate polynomials one can first try to develop proof techniques that yield improved lower bounds for the univariate case. In particular, the invariant theory of binary forms (aka univariate polynomials) is much better understood as compared to multivariate polynomials. One could also hope to apply some of the proof ideas from real/complex analysis or from the

vast literature on Waring's problem[1] to obtain improved lower bounds on $s_t(f)$. Our underlying hope is that some such improved proof technique or proof idea might admit a suitable generalization to the multivariate case as well. This could be one potential way to attack the VP versus VNP problem. We also note that there are formal results essentially following from the work of Koiran [Koi11] which imply that seemingly mild lower bounds for a slight variant of the model being considered here directly implies a separation of VP from VNP.

Proposition 2. [Implicit in [Koi11]]. *If there is an explicit family of* univariate *polynomials* $\{f_d(x) : d \geq 1\}$ *over an underlying algebraically closed field* $\mathbb{F}$ *of characteristic zero such that any representation of the form*

$$f_d(x) = \sum_{i=1}^{s} Q_i(x)^{e_i}, \quad \text{where Sparsity}(Q_i) \leq t,$$

requires the number of summands s *to be at least* $\left(\frac{d}{t}\right)^{\Omega(1)}$ *then* VP *is different from* VNP.

This means that proving relatively mild lower bounds on a similar model (but with the degree bound replaced by the corresponding sparsity bound) already implies that VP is different from VNP.

Our results. In describing our results, we avoid floor/ceil notations for ease of presentation. Throughout the rest of this paper, the underlying field $\mathbb{F}$ will be of characteristic zero. We first note that a standard dimension counting argument implies that for a random polynomial $f(x)$ of degree d it is almost surely the case that $s_t(f) \geq \frac{d+1}{t+1}$. In comparison to this benchmark, we prove a lower bound of $s_t(f) \geq \Omega\left(\sqrt{\frac{d}{t}}\right)$ for an explicit family of polynomials of degree d. Specifically, we have:

Theorem 3. *Let* $d, t \geq 2$ *be integers. Let* $a_1, \ldots, a_{2t}$ *be any* $2t$ *distinct elements of the underlying field* $\mathbb{F}$*. Assume* $\mathbb{F}$ *is of characteristic zero. Let*

$$g \stackrel{def}{=} \prod_{k=1}^{2t} (x - a_k).$$

Define the univariate polynomial,

$$f(x) \stackrel{def}{=} g(x)^{\frac{d}{2t}}. \tag{2}$$

Then

$$s_t(f) \geq \Omega\left(\sqrt{\frac{d}{t}}\right).$$

[1] Waring's problem asks whether each natural number k has an associated positive integer $s(k)$ such that every natural number is the sum of at most s k-th powers of natural numbers. For example, every natural number is the sum of at most 4 squares, 9 cubes. Many variants of Waring's problem for algebraic integers and polynomials have also been studied.

Our proof here employs the Wronskian[2] and is therefore quite different from the proof technique used in the recent works on homogeneous depth four circuits [Kay12, GKKS13a, KSS14]. These works employ a complexity measure called the dimension of shifted partials to obtain lower bounds for a similar multivariate model. We also show that a suitable variant of shifted partials does yield a similar lower bound albeit for a different target polynomial. Specifically, we have:

Theorem 4. *Let $d, t \geq 2$ be integers such that $t < \frac{d}{4}$. Let the polynomial $f(x) = \sum_{i=1}^{m}(x - a_i)^d$, with distinct a_i's and let $m = \left\lfloor \sqrt{\frac{d}{t}} \right\rfloor$. Then $s_t(f) \geq \Omega\left(\sqrt{\frac{d}{t}}\right)$.*

Remark 5.

1. **Optimality of the lower bound.** The polynomial $f(x)$ in theorem 4 has the nice feature that it can also be expressed as a sum of $O(\sqrt{d/t})$ summands, each of which is a power of a polynomial of degree at most t. So, in this sense theorem 4 gives an optimal lower bound. The *target polynomial* in theorem 3 does not seem to have this property.

2. **Methods.** In the proof of theorem 4, we show that the dimension of shifted derivatives of the polynomial $f(x)$ is the maximum possible (for the appropriate choice of parameters). Since the polynomial $f(x)$ of theorem 4 also satisfies $s_t(f) \leq O\left(\sqrt{\frac{d}{t}}\right)$, it indicates that a lower bound better than $\Omega\left(\sqrt{\frac{d}{t}}\right)$ probably cannot be obtained via shifted derivatives. It is currently conceivable that the Wronskian-based proof could yield better lower bounds. A more detailed discussion on this may be found in Pecatte's internship report [Pic14].

3. **On replacing the degree bound by the corresponding sparsity bound.** We also note that for multivariate polynomials, recent work by [KLSS14] successfully replaced the bound on the degrees of the Q_i's by the corresponding bound on the sparsity of the Q_i's. We note in passing that by proposition 2, if we could prove an analogous result as the one above but with the degree bound on the Q_i's replaced by a bound on their sparsities, then we would obtain a separation of VP from VNP. In this sparse setting, the best lower bound that is currently known is $\Omega\left(\sqrt{\frac{\log d}{\log t}}\right)$. It applies to any polynomial of degree d that has d distinct real roots [KPT15].

4. **Upper bounds.** While the focus of this paper is on lower bounds, it is also natural to ask about upper bounds on $s_t(f)$. As mentioned above, the lower bound $s_t(f) \geq \frac{d+1}{t+1}$ follows from a simple dimension counting argument. Recent work on the Waring problem for polynomials [FOS12] shows that this bound is tight for a generic polynomial of degree d when $t+1$ divides $d+1$. Moreover, a general result on "maximum rank versus generic rank" (Theorem 1 in [BT14]) shows that moving from a generic polynomial to a worst-case polynomial at most doubles $s_t(f)$. We conclude that the upper bound $s_t(f) \leq 2 \cdot \frac{d+1}{t+1}$ applies to *any* polynomial of degree d when $t+1$ divides $d+1$. Note that this upper boud is nonconstructive. A simple explicit construction shows that $s_t(f) = O((d/t)^2)$ for all f.

[2] The Wronskian has previously been employed in arithmetic complexity previously in [KPT15] to obtain nontrivial (but rather weak) lower bounds for writing a polynomial as a sum of powers of sparse polynomials. Indeed, [KPT15] manage to prove something stronger - they obtain weak (but still nontrivial and interesting) bounds on the number of real roots of sums of powers of sparse polynomials.

2 Preliminaries

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 2. [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 fact about the Wronskian whose proofs can be found in [PS76] (and which are known since the 19^{th} century).

Proposition 6. *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:*

$$\mathrm{W}(f_1, \ldots, f_k) = (f_1)^k \mathrm{W}\left(\left(\frac{f_2}{f_1}\right)', \ldots, \left(\frac{f_k}{f_1}\right)'\right). \tag{3}$$

Also, if f_1 is a perfect power say if $f_1 = Q^e$ where $e \geq k$ then Q^{e-k+1} divides $\mathrm{W}(Q^e, f_2, \ldots, f_k)$.

Another basic fact about the Wronskian is that it captures linear dependence of polynomials in $\mathbb{F}[x]$.

Proposition 7. *[Boc01] Let $\mathbb{F}$ be a field of characteristic zero. For univariate polynomials $f_1, \ldots, f_n \in \mathbb{F}[x]$, they are linearly dependent if and only if the Wronskian $\mathrm{W}(f_1, \ldots, f_n)$ vanishes everywhere.*

We will also use another result from [VP75] which gives a bound on the multiplicity of a root depending on the Wronskian. For a field element $\alpha \in \mathbb{F}$, and a polynomial $g(x) \in \mathbb{F}[x]$, let $\mathrm{N}_\alpha(g)$ denote the multiplicity of g at α, i.e. the highest power of $(x-\alpha)$ which divides $g(x)$.

Lemma 8. *Let $\mathbb{F}$ be a field of characteristic zero. Let $Q_1, \ldots, Q_m$ be some linearly independent polynomial and $\alpha \in \mathbb{F}$, and let $F(x) = \sum_{i=1}^m Q_i(x)$. Then:*

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

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

2.2 The space of shifted derivatives

In section 4 we give an alternate lower bound proof via a slight variant of a complexity measure first defined in [Kay12]: the space of shifted partial derivatives. Using this complexity measure, [Kay12] obtained exponential lower bounds on a similar multivariate model. The key intuition follows from the following simple observation : 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 3 (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:*

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

$\left\langle x^{\leq i+l} \cdot f^{(i)} \right\rangle_{i \leq k}$ 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 9. 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 the above 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.

3 Proof of theorem 3

Suppose $f = \sum_{i=1}^{s} \alpha_i \cdot Q_i^{e_i}$. Since degree of every Q_i is bounded by t and $\deg(f) = d$, $e_i \geq \frac{d}{t}$ for some $i \in [s]$. Without loss of generality, let $e_1 \geq \frac{d}{t}$. Also, we can assume that $Q_1^{e_1}, \ldots, Q_s^{e_s}$ are $\mathbb{F}$-linearly independent - if not, we work with a basis and a smaller value for s. By taking derivatives of both sides of the equation $f = \sum_{i=1}^{s} \alpha_i \cdot Q_i^{e_i}$ with respect to x for j times we have,

$$\sum_{i=1}^{s} \alpha_i \cdot [Q_i^{e_i}]^{(j)} = f^{(j)}, \qquad \text{for every } j \in \{0, \ldots, s-1\}\ ,$$

where $[Q_i^{e_i}]^{(j)}$ and $f^{(j)}$ are the j-th derivatives of $Q_i^{e_i}$ and f, respectively, with respect to x. The above equation defines a system of linear equations in $\alpha_1, \ldots, \alpha_s$. By applying Cramer's rule,

$$\alpha_1 = \frac{\mathrm{W}\left(f, Q_2^{e_2}, \ldots, Q_s^{e_s}\right)}{\mathrm{W}\left(Q_1^{e_1}, Q_2^{e_2}, \ldots, Q_s^{e_s}\right)}, \tag{4}$$

where $\mathrm{W}\left(g_1, \ldots, g_s\right)$ is the Wronskian determinant of the polynomials $g_1, \ldots, g_s$. Since $Q_1^{e_1}, Q_2^{e_2}, \ldots, Q_s^{e_s}$ are $\mathbb{F}$-linearly independent, $\mathrm{W}\left(Q_1^{e_1}, Q_2^{e_2}, \ldots, Q_s^{e_s}\right) \neq 0$. Observe that unless $s = \Omega\left(\frac{d}{t}\right)$, $Q_1^{e_1-(s-1)}$ divides $\mathrm{W}\left(Q_1^{e_1}, Q_2^{e_2}, \ldots, Q_s^{e_s}\right)$ and $g^{\frac{d}{2t}-(s-1)}$ divides $\mathrm{W}\left(f, Q_2^{e_2}, \ldots, Q_s^{e_s}\right)$. Let

$$\Delta \stackrel{\text{def}}{=} \{i \mid e_i \geq s \text{ and } 2 \leq i \leq s\}.$$

Then, $\prod_{i\in\Delta} Q_i^{e_i-(s-1)}$ divides both $\mathrm{W}\left(Q_1^{e_1}, Q_2^{e_2}, \ldots, Q_s^{e_s}\right)$ and $\mathrm{W}\left(f, Q_2^{e_2}, \ldots, Q_s^{e_s}\right)$. Thus, by analyzing the factors coming out common from the Wronskian determinants, we can express α_1 as

$$\begin{aligned}\alpha_1 &= \frac{g^{\frac{d}{2t}-(s-1)} \cdot \prod_{i\in\Delta} Q_i^{e_i-(s-1)} \cdot W_1}{Q_1^{e_1-(s-1)} \cdot \prod_{i\in\Delta} Q_i^{e_i-(s-1)} \cdot W_2} \\ &= \frac{g^{\frac{d}{2t}-(s-1)} \cdot W_1}{Q_1^{e_1-(s-1)} \cdot W_2}. \end{aligned} \tag{5}$$

Now observe that after taking $Q_1^{e_1-(s-1)}$ and $\prod_{i\in\Delta} Q_i^{e_i-(s-1)}$ common from $\mathrm{W}\left(Q_1^{e_1}, Q_2^{e_2}, \ldots, Q_s^{e_s}\right)$, every polynomial in the r-th row of the Wronskian matrix of $Q_1^{e_1}, Q_2^{e_2}, \ldots, Q_s^{e_s}$ has degree upper bounded by $(s-1)t-(r-1)$. Hence,

$$\deg(W_2) \leq s(s-1)t - \sum_{r=1}^{s}(r-1) \leq s^2 t.$$

Since α_1 is a field element, $g^{\frac{d}{2t}-(s-1)}$ must divide $Q_1^{e_1-(s-1)} \cdot W_2$ (by Equation 5). Polynomial g has $2t$ distinct roots, whereas polynomial Q_1 has at most t roots. Therefore, there are t distinct roots of g such that each of these roots divide W_2 with multiplicity $\frac{d}{2t}-(s-1)$. Since $\deg(W_2) \leq s^2 t$,

$$\begin{aligned} s^2 t &\geq t \cdot \left[\frac{d}{2t} - (s-1)\right] \\ \Rightarrow\ s^2 + s &\geq \frac{d}{2t} + 1 \\ \Rightarrow\ s &\geq \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{d}{t}} - \frac{1}{2}. \end{aligned}$$

The t=1 case. When t=1, the above argument can be strengthened to show the following: if $x^d + x^{d-1}$ is expressed as a sum of s-many d-th powers of linear polynomials then s is at least $d+1$. Such an optimum bound also follows from a work on representing homogeneous (multivariate) polynomials as sums of linear forms by Kleppe [Kle99].

4 An alternative proof using shifted partials

In this section, we will give a proof of theorem 4 using shifted derivatives. The proof will consist in first giving an upper bound on the dimension of shifted partials of a sum of powers of low degree polynomials. Thereafter, we give a lower bound on the dimension of 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.

4.1 Upper bounding the dimension of shifted partial derivatives.

Recall the complexity measure called the shifted partial dimension as defined in section 2. We first show that in our model, polynomials have a small complexity according to this measure:

Proposition 10. *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, with $\deg Q \leq t$, we have $\dim \left\langle x^{\leq i+l} \cdot Q^{e(i)} \right\rangle_{i \leq k} \leq l + kt + 1$. Now note that any $g \in \left\langle x^{\leq i+l} \cdot Q^{e(i)} \right\rangle_{i \leq k}$ is of the form $g = Q^{e-k} \cdot R$. Moreover $\deg(R) \leq l + kt$ (since $\deg g \leq e \cdot t + l$). This directly gives the bound on the dimension. □

4.2 Lower bounding the dimension of shifted derivatives for an explicit polynomial.

We now give an explicit lower bound on the dimension of shifted derivative space of our explicit polynomial.

Definition 4. 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 $P_i \in \mathbb{F}[x]$, not all zero, 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 11. *For any $h(x) \in \mathbb{F}[x]$, if h does not satisfy any SDE of order k and of shift l, then $\left\langle x^{\leq i+l} \cdot h^{(i)} \right\rangle_{i \leq k}$ is full, i.e. :*

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

In order to prove some conditions on the SDE satisfied by our target explicit polynomial $f(x)$, 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 12. *For any $d, m \leq d$, for any distinct $(a_1, a_2, \ldots, a_m) \in \mathbb{F}^m$, 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 $h \in S$, then the order of the SDE must be greater than or equal to m.*

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

$$\sum_{i=0}^{k} P_i(x) h^{(i)}(x) = 0 \quad \forall h \in S. \tag{6}$$

For all $j \in [m]$, we can factor out $(x-a_j)^{d-k}$ from the above equation to obtain a new SDE satisfied by the family $S' = \{(x-a_1)^k, \ldots, (x-a_m)^k\}$. i.e.:

$$\sum_{i=0}^{k} R_i(x) h^{(i)}(x) = 0 \quad \forall h \in S', \tag{7}$$

with $R_i(x) \stackrel{\text{def}}{=} \frac{d!}{k!}\frac{(k-i)!}{(d-i)!} P_i(x)$.

But now, since $k < m$, the family S' generate $\mathbb{F}_k[x]$ (the vector space of polynomials of degree at most k), and thus this implies that every polynomial of degree $\leq k$ should satisfy the SDE (7). We obtain the contradiction by plugging in $h(x) = x^{i_0}$ in SDE (7), where i_0 is the smallest integer such that $R_{i_0}(x) \not\equiv 0$. □

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

Lemma 13. *For any $d, m \leq d$, for any m distinct elements $a_1, a_2, \ldots, a_m \in \mathbb{F}$, 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) $k \geq m$, or,

ii) $l > \frac{d}{m} - \frac{3}{2} \cdot m$.

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:

$$\sum_{i=0}^{k} P_i(x) f^{(i)}(x) = 0, \tag{8}$$

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

$$\sum_{i=0}^{k} P_i(x) \left((x-a_j)^d\right)^{(i)}(x) = R_j(x)(x-a_j)^{d-k}.$$

Notice that R_j is of degree at most $k+l$. By lemma 12, since $k < m$, not all R_j's can be 0, without loss of generality we have $R_1 \not\equiv 0$. For $j \in [m]$, we set $f_j(x) = R_j(x)(x-a_j)^{d-k}$ and, using linearity of differentiation, we rewrite differential equation (8) as:

$$-f_1(x) = \sum_{j=2}^{m} f_j(x).$$

Using Lemma 8, for a certain subset $J = \{j_1, \ldots, j_p\} \subseteq [2..m]$, we obtain

$$d - k \leq \mathrm{N}_{a_1}(f_1) \leq p - 1 + \mathrm{N}_{a_1}(\mathrm{W}((f_j)_{j \in J})). \tag{9}$$

We can factorize the Wronskian by $(x-a_j)^{d-k-(p-1)}$ for any $j \in J$:

$$\mathrm{N}_{a_1}\left(\mathrm{W}\left((f_j)_{j\in J}\right)\right) = \mathrm{N}_{a_1}\begin{vmatrix} R_{1,1} & \dots & R_{1,p} \\ \vdots & \ddots & \vdots \\ R_{p,1} & \cdots & R_{p,p} \end{vmatrix},$$

with $\deg(R_{i,j}) \leq l+k+p-i$.
The determinant has degree $\leq p(l+k)+\binom{p}{2}$. Hence, inequality (9) 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 negative terms to obtain:

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

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

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

□

4.3 Putting things together

We are now ready to give a proof of theorem 4.

Proof. We take k and l small enough to ensure that f does not satisfy any SDE of parameters k and l. Using lemma 13, 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 11, 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 10 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:

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

□

5 Discussion

In this work, we introduce the model of sums of powers of univariates and gave a new proof technique (via the Wronskian) to prove a lower bound in this model. Even though the existing technique of shifted partials also yields a similar lower bound in this model, our proof (via the Wronskian) could nevertheless be interesting for it is different and perhaps some suitable generalization of it might yield improved lower bounds for some classes of multivariate circuits. In any case, we feel that the sum of powers of univariates model is easier to analyze and may serve as a testbed for other candidate techniques or complexity measures aiming to obtain improved circuit lower bounds. We conclude by mentioning a few open problems that are implicit in remark 5.

- Obtain a lower bound for sums of powers of t-sparse polynomials which is better than $\Omega(\sqrt{\frac{\log d}{\log t}})$.
- Obtain a $d^{O(1)}$-time algorithm for expressing a given degree d polynomial as a sum of $O(\frac{d}{t})$-many powers of degree-t polynomials.
- Improve the $\Omega(\sqrt{\frac{d}{t}})$ lower bound shown in this work.

Acknowledgments

A part of this work was done at the Dagstuhl workshop on algebra in computational complexity (seminar 14391) and at the WACT 2015 workshop in Saarbrücken. We also thank Michael Forbes for pointing out the results of Blekherman and Teitler [BT14] to us.

References

[AJMV98] Eric Allender, Jia Jiao, Meena Mahajan, and V. Vinay. Non-Commutative Arithmetic Circuits: Depth Reduction and Size Lower Bounds. *Theor. Comput. Sci.*, 209(1-2):47–86, 1998.

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

[Boc01] M. Bocher. The theory of linear dependence. *Annals of Mathematics*, 2(1/4):81–96, 1900-1901.

[BT14] Grigoriy Blekherman and Zach Teitler. On maximum, typical and generic ranks. *Mathematische Annalen*, pages 1–11, 2014.

[FLMS14] Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth 4 formulas computing iterated matrix multiplication. In *Symposium on Theory of Computing, STOC 2014*, pages 128–135, 2014.

[FOS12] Ralf Fröberg, Giorgio Ottaviani, and Boris Shapiro. On the Waring problem for polynomial rings. *Proceedings of the National Academy of Sciences*, 109(15):5600–5602, 2012.

[GKKS13a] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. In *Conference on Computational Complexity (CCC)*, pages 65–73, 2013.

[GKKS13b] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth three. In *Foundations of Computer Science (FOCS)*, pages 578–587, 2013.

[Kay12] Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:81, 2012.

[Kle99] Johannes Kleppe. Representing a Homogenous Polynomial as a Sum of Powers of Linear Forms. *Thesis for the degree of Candidatus Scientiarum (University of Oslo)*, 1999. Available at `http://folk.uio.no/johannkl/kleppe-master.pdf`.

[KLSS14] Neeraj Kayal, Nutan Limaye, Chandan Saha, and Srikanth Srinivasan. An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS*, pages 61–70, 2014.

[Koi11] Pascal Koiran. Shallow circuits with high-powered inputs. In *Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 7-9, 2011. Proceedings*, pages 309–320, 2011.

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

[KPT15] Pascal Koiran, Natacha Portier, and Sébastien Tavenas. A Wronskian approach to the real τ-conjecture. *J. Symb. Comput.*, 68:195–214, 2015.

[KS14a] Neeraj Kayal and Chandan Saha. Lower bounds for depth three arithmetic circuits with small bottom fanin. *Electronic Colloquium on Computational Complexity (ECCC)*, 21:89, 2014.

[KS14b] Mrinal Kumar and Shubhangi Saraf. The limits of depth reduction for arithmetic formulas: it's all about the top fan-in. In *Symposium on Theory of Computing, STOC*, pages 136–145, 2014.

[KS14c] Mrinal Kumar and Shubhangi Saraf. On the power of homogeneous depth 4 arithmetic circuits. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS*, pages 364–373, 2014.

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *Symposium on Theory of Computing, STOC 2014*, pages 146–153, 2014.

[Pic14] Timothée Picatte. Lower bounds for univariate polynomials: a Wronskian approach. *M2 Internship Report (Ecole Normale Supérieure de Lyon)*, 2014. Available at `http://perso.ens-lyon.fr/pascal.koiran/timothee_pecatte_master2report.pdf`.

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

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

[Val79] L. G. Valiant. Completeness Classes in Algebra. In *Symposium on Theory of computing STOC*, pages 249–261, 1979.

[VP75] M. Voorhoeve and A. J. Van Der Pooerten. Wronskian determinants and the zeros of certain functions. *Indagationes Mathematicae*, 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 Journal on Computing*, 12(4):641–644, 1983.