Electronic Colloquium on Computational Complexity, Report No. 81 (2012)

An exponential lower bound for the sum of powers of bounded degree polynomials

Neeraj Kayal *

Dedicated to Somenath Biswas on his 60th birthday.

Abstract

In this work we consider representations of multivariate polynomials in $\mathbb{F}[\mathbf{x}]$ of the form

$$f(\mathbf{x}) = Q_1(\mathbf{x})^{e_1} + Q_2(\mathbf{x})^{e_2} + \ldots + Q_s(\mathbf{x})^{e_s},$$

where the e_i's are positive integers and the Q_i's are arbitary multivariate polynomials of bounded degree. We give an explicit n-variate polynomial f of degree n such that any representation of the above form for f requires the number of summands s to be $2^{\Omega(n)}$.

Motivation. Let $\mathbb{F}$ be a field, $\mathbb{F}[\mathbf{x}]$ be the set of n-variate polynomials over $\mathbb{F}$ and $d \geq 1$ be an integer. For a polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$, we consider representations of the form

$$f(\mathbf{x}) = Q_1^{e_1}(\mathbf{x}) + Q_2^{e_2}(\mathbf{x}) + \ldots + Q_s^{e_s}(\mathbf{x}), \tag{1}$$

where the $Q_i(\mathbf{x})$'s are polynomials of degree at most d. We do this with an eye towards proving lower bounds for the number of summands s required to write some explicit polynomial f in the above form. Our motivation for this line of inquiry stems from some recent results and problems posed in the field of arithmetic complexity. Agrawal and Vinay [AV08] showed that proving exponential lower bounds for depth four arithmetic circuits implies exponential lower bounds for arbitrary depth arithmetic circuits. In our case, a representation of the form (1) above corresponds to computing f via a depth four $\Sigma\Pi\Sigma\Pi$ arithmetic circuit where the bottommost layer of multiplication gates have fanin bounded by d and the second-last layer of multiplication gates actually consists of exponentiation gates of arbitrarily large degree (i.e. multiplication gates where all the incoming edges originate from a single node). Meanwhile Hrubes, Wigderson and Yehudayoff [HWY10] look at the situation where $d = e_1 = e_2 = \ldots = e_s = 2$ and ask for a superlinear lower bound on the number of summands s for an explicit n-variate biquadratic polynomial f. They show that such a superlinear lower bound implies an exponential lower bound on the size of arithmetic circuits computing the noncommutative permanent. Finally Chen, Kayal and Wigderson [CKW11] pose the problem of proving lower bounds for bounded depth arithmetic circuits with addition and exponentiation gates. Our main theorem is a lower bound on the number of summands in any representation of the form (1) for an explicit polynomial.

Theorem 1. (Lower bound for sum of powers). *Let $\mathbb{F}$ be any field and $\mathbb{F}[\mathbf{x}]$ be the ring of polynomials over the set of indeterminates $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Let $e_1, e_2, \ldots, e_s$ be positive integers and $Q_1, Q_2, \ldots, Q_s \in \mathbb{F}[\mathbf{x}]$ be multivariate polynomials each of degree at most d. If*

$$Q_1^{e_1} + Q_2^{e_2} + \ldots + Q_s^{e_s} = (x_1 \cdot x_2 \cdot \ldots \cdot x_n),$$

then we must have that $(\log s) = \Omega(\frac{n}{2^d})$. In particular, if d is a constant then $s = 2^{\Omega(n)}$.

Remark 2. 1. The fact that the f in the lower bound above consists of a single monomial indicates above all the severe limitation of representations of the form (1).

*Microsoft Research India, neeraka@microsoft.com

ISSN 1433-8092

2. An upper bound of $2^{n/d}$ is an easy corollary of Fischer [Fis94]. Specifically, let $\mathbb{F}$ be an algebraically closed field with $\mathrm{char}(\mathbb{F}) > n$. Then for all integers $d \geq 1$ there exist polynomials $Q_1, Q_2, \ldots, Q_s$ each of degree d such that
$$Q_1^{e_1} + Q_2^{e_2} + \ldots + Q_s^{e_s} = (x_1 \cdot x_2 \cdot \ldots \cdot x_n),$$
and the number of summands s is at most $2^{n/d}$. Fischer [Fis94] gives an explicit set of 2^{m-1} linear forms $\ell_1, \ell_2, \ldots, \ell_{2^m}$ such that
$$(y_1 \cdot y_2 \cdot \ldots \cdot y_m) = \sum_{i \in [2^{m-1}]} \ell_i(\mathbf{y})^m.$$
Replacing y_i by $(\prod_{j \in [d]} x_{(i-1)d+j})$ in the above equation we get a representation of $f = \prod_{i \in [n]} x_i$ as a sum of about $2^{n/d}$ powers of polynomials of degree d.

3. The problem posed by Chen, Kayal and Wigderson ([CKW11], section 10.1) remains open – even for the case of depth four circuits with addition and exponentiation gates, where all the exponentiation gates are allowed to raise their respective inputs to an arbitrary exponent.

Notation

$[n]$ denotes the set $\{1, 2, \ldots, n\}$. For an n-tuple of nonnegative integers $\mathbf{i} = (i_1, i_2, \ldots, i_n) \in \mathbb{Z}_{\geq 0}^n$, $|i|$ denotes the sum $\sum_{j \in [n]} |i_j|$.
Shorthand for partial derivatives. For a polynomial $f(\mathbf{x}) \in \mathbb{F}[x_1, x_2, \ldots, x_n]$, we use $\partial_i f$ as a shorthand for $\frac{\partial f}{\partial x_i}$, the formal partial derivatve of f with respect to the variable x_i. For $\mathbf{i} = (i_1, i_2, \ldots, i_n) \in \mathbb{Z}_{\geq 0}^n$, we use the following shorthand
$$\boldsymbol{\partial}^{\mathbf{i}} f \stackrel{\text{def}}{=} \frac{\partial^{i_1}}{\partial x_1^{i_1}}\left(\frac{\partial^{i_2}}{\partial x_2^{i_2}}\left(\cdots\left(\frac{\partial^{i_n} f}{\partial x_n^{i_n}}\right)\cdots\right)\right).$$
$\mathbb{F}$-linear dependence. We will say that polynomials $f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x})$ are $\mathbb{F}$-*linearly dependent* if there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathbb{F}$, not all zero, such that
$$\alpha_1 \cdot f_1 + \alpha_2 \cdot f_2 + \ldots + \alpha_m \cdot f_m = 0,$$
otherwise they are $\mathbb{F}$-*linearly independent*. For a set of polynomials $S \subseteq \mathbb{F}[\mathbf{x}]$, $\dim(S)$ is the size of a maximal $\mathbb{F}$-linearly independent subset of polynomials in S. The $\mathbb{F}-$span of a set S of polynomials is defined as the set of all possible $\mathbb{F}$-linear combinations of polynomials from S, i.e.
$$\mathbb{F}\text{-span}(S) \stackrel{\text{def}}{=} \{(\alpha_1 \cdot f_1 + \ldots + \alpha_m \cdot f_m) \ : \ f_i \in S \ \text{ and } \ \alpha_i \in \mathbb{F} \ \text{ for all } i \in [m].\}$$
Note that $\mathbb{F}-$span(S) forms a vector space and that $\dim(S)$ is the same as the dimension of this vector space.

Proof of the lower bound (theorem 1)

Definition 3. *For a polynomial $f \in \mathbb{F}[x_1, x_2, \ldots, x_n]$, let*
$$(\boldsymbol{\partial}^{\leq k} f) \stackrel{\text{def}}{=} \{\partial^{\mathbf{i}} f \ \ : \ \ \mathbf{i} \in \mathbb{Z}_{\geq 0}^n \ \text{ and } \ |\mathbf{i}| \leq k\}$$
For a set $S \subseteq \mathbb{F}[\mathbf{x}]$ let
$$\mathbf{x}^{\leq \ell} \cdot S \stackrel{\text{def}}{=} \{\mathbf{x}^{\mathbf{j}} \cdot f \ : \ f \in S, \ \ \mathbf{j} \in \mathbb{Z}_{\geq 0}^n \ \text{ and } \ |\mathbf{j}| \leq \ell\} \subseteq \mathbb{F}[\mathbf{x}].$$
In particular,
$$\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k} f) \stackrel{\text{def}}{=} \{\mathbf{x}^{\mathbf{j}} \cdot (\partial^{\mathbf{i}} f) \ \ : \ \ \mathbf{i} \in \mathbb{Z}_{\geq 0}^n, \ \mathbf{j} \in \mathbb{Z}_{\geq 0}^n \ \text{ where } \ |\mathbf{i}| \leq k, \ \ |\mathbf{j}| \leq \ell\}$$
Finally, for a polynomial $g \in \mathbb{F}[\mathbf{x}]$ we will often use $\mathbf{x}^{\leq \ell} \cdot g$ as a shorthand for $\mathbf{x}^{\leq \ell} \cdot \{g\}$.

In what follows we use the following convention to improve clarity: for an integer $t < 0$ and a polynomial $Q(\mathbf{x})$, Q^t stands for the zero polynomial.

Lemma 4. *Let*

$$f = Q_1^{e_1} + Q_2^{e_2} + \ldots + Q_s^{e_s},$$

where each $Q_j(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is of degree at most d. Then

$$\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k} f) \subseteq \mathbb{F}\text{−span}\left(\bigcup_{j\in[s]} \bigcup_{t\in[0..k]} \bigcup_{\substack{\mathbf{i}\in\mathbb{Z}^n_{\geq 0}\\ |\mathbf{i}|\leq \ell+(d-1)t}} \mathbf{x}^{\mathbf{i}} \cdot Q_j^{e_j - t} \right) \tag{2}$$

In particular,

$$\dim\left(\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k} f)\right) \leq s \cdot (k+1) \cdot \binom{n+\ell+(d-1)k}{\ell+(d-1)k}$$

Proof. By linearity of derivatives we have

$$\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k}(\sum_{j\in[s]} Q_j^{e_j})) \subseteq \mathbb{F}\text{−span}\left(\bigcup_{j\in[s]} (\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k} Q_j^{e_j}))\right)$$

and therefore it suffices to show that

$$\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k}(Q_j^{e_j})) \subseteq \mathbb{F}\text{−span}\left(\bigcup_{t\in[0..k]} \bigcup_{\substack{\mathbf{i}\in\mathbb{Z}^n_{\geq 0}\\ |\mathbf{i}|\leq \ell+(d-1)t}} \mathbf{x}^{\mathbf{i}} \cdot Q_j^{e_j - t} \right). \tag{3}$$

Now, by induction on k one can show that

$$\boldsymbol{\partial}^{\leq k}(Q_j^{e_j}) \subseteq \mathbb{F}\text{−span}\left(\bigcup_{t\in[0..k]} \bigcup_{\substack{\mathbf{i}\in\mathbb{Z}^n_{\geq 0}\\ |\mathbf{i}|\leq (d-1)t}} \mathbf{x}^{\mathbf{i}} \cdot Q_j^{e_j - t} \right). \tag{4}$$

Also note that for any polynomial g and any two nonnegative integers ℓ, r we have

$$\mathbf{x}^{\leq \ell} \cdot (\mathbf{x}^{\leq r} \cdot (g)) = \mathbf{x}^{\leq \ell + r} \cdot (g). \tag{5}$$

Thus applying (5) to (4) we get (3) and therefore (2) as well. Finally since the set of monomials $\mathbf{x}^{\leq r}$ is of size $\binom{n+r}{r}$ we have

$$\begin{aligned}
\dim\left(\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k} f)\right) &\leq \sum_{j\in[s]} \sum_{t\in[k]} \binom{n+\ell+(d-1)t}{\ell+(d-1)t} \\
&= s \cdot \sum_{t\in[k]} \binom{n+\ell+(d-1)t}{\ell+(d-1)t} \\
&\leq s \cdot (k+1) \cdot \binom{n+\ell+(d-1)k}{\ell+(d-1)k}
\end{aligned}$$

This proves the lemma. $\square$

Lemma 5. *Let $f = (x_1 \cdot x_2 \cdot \ldots \cdot x_n) \in \mathbb{F}[\mathbf{x}]$. Then*

$$\dim\left(\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k} f)\right) \geq \binom{n}{k} \cdot \binom{n-k+\ell}{\ell}.$$

Proof. Let $S = \left(\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k} f)\right) \subseteq \mathbb{F}[\mathbf{x}]$. Since f is a monomial, we have that all the polynomials in S are in fact monomials and therefore $\dim(S)$ is precisely the number of distinct monomials in S. Since monomials with distinct supports are distinct, it therefore suffices to show that for every set $T \subseteq [n]$ of size $(n-k)$, there are $\binom{n-k+\ell}{\ell}$ distinct monomials in S supported only on variables indexed by T; in other words there are $\binom{n-k+\ell}{\ell}$ monomials in S of the form $\prod_{i \in T} x_i^{e_i}$, where each $e_i \geq 1$. To see this consider the monomial $m = \prod_{i \in T} x_i$. Then $m \in \boldsymbol{\partial}^{\leq k} f$ as m can be obtained from f by taking the derivative with respect to the set of k variables with indices not in T, i.e.

$$m = \boldsymbol{\partial}^{\mathbf{i}}(x_1 \cdot x_2 \cdot \ldots \cdot x_n), \mathbf{i} = (i_1, i_2, \ldots, i_n) \text{ where } i_j = \begin{cases} 0 & \text{if } j \in T \\ 1 & \text{otherwise} \end{cases}$$

Thus the set of monomials in S supported on variables indexed by T is precisely the set of monomials of the form

$$\left(\prod_{i \in T} x_i^{e_i}\right) \cdot m, \quad \text{where each } e_i \geq 0 \quad \text{and } \sum_{i \in T} e_i \leq \ell.$$

There are exactly $\binom{n-k+\ell}{\ell}$ monomials of the above form. This proves the lemma. □

With these estimates in hand, we are ready to give a proof of theorem 1.
Proof of Theorem 1: Assume that

$$(x_1 \cdot x_2 \cdot \ldots \cdot x_n) = Q_1^{e_1} + Q_2^{e_2} + \ldots + Q_s^{e_s}.$$

Then for every $k, \ell \geq 0$ we must have

$$\dim\left(\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k}(\prod_{i \in [n]} x_i))\right) = \dim\left(\mathbf{x}^{\leq \ell} \cdot (\boldsymbol{\partial}^{\leq k}(\sum_{i \in [s]} Q_i^{e_i}))\right)$$

Using the estimates provided by lemmas 4 and 5 for all $k, \ell \geq 0$ we have

$$\binom{n}{k} \cdot \binom{n-k+\ell}{\ell} \leq s \cdot (k+1) \cdot \binom{n+\ell+(d-1)k}{\ell+(d-1)k}$$

and therefore

$$s \geq \frac{1}{k+1}\binom{n}{k} \cdot \binom{n-k+\ell}{\ell} \Big/ \binom{n+\ell+(d-1)k}{\ell+(d-1)k}.$$

Now setting $\ell = n$ and $k = c \cdot 2^{-d} \cdot n$ (for a suitable constant c) and using Stirling's approximation ($\ln n! = n \cdot \ln n - n + O(\ln n)$), we get an asymptotic lower bound on s:

$$\ln s = \Omega(\frac{n}{2^d}) + O(d \ln n).$$

In particular when d is a constant then $s = 2^{\Omega(n)}$. This proves the theorem. □

Acknowledgements

The author would like to thank Ankit Gupta and Pritish Kamath for many helpful discussions. In particular, the discussions with Ankit lead to the consideration of the property in definition 3 and discussions with Pritish helped in refining and getting a sense of the estimates of the sums of binomial coefficients involved.

References

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

[CKW11] Xi Chen, Neeraj Kayal, and Avi Wigderson. Partial derivatives in arithmetic complexity. *Foundations and Trends in Theoretical Computer Science*, 2011.

[Fis94] I. Fischer. Sums of like powers of multivariate linear forms. *Mathematics Magazine*, 67(1):59–61, 1994.

[HWY10] Pavel Hrubes, Avi Wigderson, and Amir Yehudayoff. Non-commutative circuits and the sum-of-squares problem. In *STOC*, pages 667–676, 2010.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de