E0309 Topics in Complexity Theory **Spring 2015**

Lecture 12: March 4, 2015

Lecturer: Neeraj Kayal *Scribe: Abhijat Sharma*

12.1 Recap

In the last lecture, we proved the following result by Ellison[2]:

Theorem 12.1 *Any polynomial $f \in \mathbb{F}[X]$, where $\mathbb{F}$ is a field and $X = (x_1, x_2, ..., x_n)$ is a set of n formal variables, can be written as a sum of powers of affine forms, i.e*

$$f(X) = \alpha_1 l_1^{e_1} + \alpha_2 l_2^{e_2} + ... + \alpha_s l_s^{e_s} \tag{12.1}$$

where $l_1, l_2, ...l_s$ are affine polynomials in $x_1, x_2, ..., x_n$ (degree of each monomial in the l_i's is at-most one), for all $i = 1, 2, ..., s$, $e_i \leq d$ where $d \geq 1$ is the total degree of the polynomial f.

We eventually saw that without loss of generality, it is safe to assume that $e_1 = e_2 = ... = e_s = d$ and therefore $f(X) = \sum_{i=1}^{s} \alpha_i l_i^d$. We also used a previous result describing the minimum number of multiplications required to compute any random n-variate d-degree polynomial, to obtain a lower bound on the number of summands s needed to express any random n-variate d-degree polynomial as a sum of powers of affine forms (as in equation 12.1). We found that with high probability (probabilty $p = 1 - \frac{1}{2^{\frac{3}{4}\binom{n+d}{d}}}$), for any d-degree polynomial in n variables, the number of summands $s \geq \frac{1}{2 \log d}\sqrt{\binom{n+d}{d}}$, which can be approximated by Stirling's bounds (when $n \gg d$, $\binom{n+d}{d} \simeq n^d$) to $s \geq n^{d/2}$. Now, we proceed to obtain a tighter lower bound using a similar counting argument as used for the above bound, and explore more about the number of summands s for general polynomials.

12.2 A trivial upper bound on s

We claim that in any given polynomial, when it is expressed as in equation 12.1, the number of summands cannot exceed the number of monomials. More formally, as we know that there are at-most $\binom{n+d}{d}$ monomials in a d-degree n-variate polynomial,

Claim 12.2 *If $f(X) = \sum_{i=1}^{s'} \alpha_i l_i^d$ and $s' > \binom{n+d}{d}$, it is possible to "rewrite" $f = \sum_{i=1}^{s} \beta_i l_i^d$ such that $s \leq \binom{n+d}{d}$.*

Proof: Consider a vector space $V = \mathbb{F}_d[X]$, containing all possible n-variate d-degree polynomials, where every polynomial $f \in \mathbb{F}_d[X]$ is represented as a vector with $\binom{n+d}{d}$ components, each component corresponding to a particular monomial, the value of that component representing the coefficient of that monomial in the polynomial f. It can be easily observed that every possible polynomial in $\mathbb{F}_d[X]$ can be represented uniquely as the above described vector.

Now, suppose we are given a representation of a polynomial f as a sum of powers of affine forms, as follows:

$$f(X) = \alpha_1 l_1^d + \alpha_2 l_2^d + ...\alpha_{s'} l_{s'}^d$$

Then, $l_1^d, l_2^d, ...l_{s'}^d$ are, like f, all d-degree polynomials in n variables and thus belong to the vector space V. Thus, each of l_i^d for $i = 1, 2, ..., s'$ can be represented as a $\binom{n+d}{d}$-dimensional vector as described above. Now, we define a subspace $W = \mathbb{F} - span(l_1^d, l_2^d, ...l_{s'}^d)$, i.e W is the set of all linear combinations (referred as *linear span*) of the vectors representing $l_1^d, l_2^d, ...l_{s'}^d$, where the coefficients belong to a field $\mathbb{F}$. Clearly, the vector representing the polynomial $f(X) = \sum_{i=1}^{s'} \alpha_i l_i^d$ also belongs to W. By definition of a linear span, W is a subspace of V ($W \subseteq V$), which implies the dimension of W, $dim(W) \leq dim(V) = \binom{n+d}{d}$. Therefore, there exists a *basis* I of size equal to $dim(W)$. Let $k = |I|$ be the dimension of the subspace W. The set of s' vectors representing the polynomials l_i^d for $i = 1, 2, ..., s'$ cannot all be linearly independent, if $s\prime > \binom{n+d}{d} \geq k$. Hence, there exists a linearly dependent subset of these s' vectors, that forms the basis I of size k. Let that basis be $I = \{l_{j_1}^d, l_{j_2}^d, ..., l_{j_k}^d\}$ where $j_1, j_2, ..., j_k \in [s']$. Now, I is the basis for the subspace W, so every vector in W can be written as a linear combination of vectors in I. Earlier we stated that $f(X) \in W$ so $f(x)$ can be written as a linear combination of the basis vectors as follows:

$$f(X) = \beta_1 l_{j_1}^d + \beta_2 l_{j_2}^d + ... + \beta_k l_{j_k}^d$$

where $k = dim(W) \leq \binom{n+d}{d}$, which proves our claim. ■

12.3 Improving the lower bound on s

Consider all polynomials $f \in \mathbb{F}[X]$ in n variables and having degree d. Assuming that $\mathbb{F}$ is a finite field of cardinality $q = |\mathbb{F}|$. Then, total number of polynomials possible is equal to the total no. of ways of choosing $\binom{n+d}{d}$ coefficients for each of the monomials, which is equal to $q^{\binom{n+d}{d}}$.

Now, let us try to estimate the number of polynomials that can be written as a sum of dth powers of s affine forms, i.e $f = \sum_{i=1}^{s} \alpha_i l_i^d$.

Claim 12.3 *The number of polynomials g which can be expressed as $g = \alpha.l_d$, where l is an affine form, is $q^{(n+1)}$.*

Proof: Consider any arbitrary affine form $l(x_1, x_2, ..., x_n) = a_0 + a_1 x_1 + ...a_n x_n$ where $a_i \in \mathbb{F}$ for $i = 0, 1, 2, .., n$. So, if we choose $(n+1)$ values of the coefficients a_i, we fix the affine form l. This implies that there are $q^{(n+1)}$ possible choices for the affine form l. We argue that the number of polynomials $g = \alpha.l_d$ is equal to the number of possible affine forms l, because

$$g = \alpha.(a_0 + a_1 x_1 + ...a_n x_n)^d = (\sqrt[d]{\alpha}.a_0 + \sqrt[d]{\alpha}.a_1 x_1 + ... + \sqrt[d]{\alpha}.a_n x_n)^d$$

From the above equation, it is clear that we do not have to choose a value for the variable α, to fix the polynomial g. Therefore, g is uniquely defined by the $(q^{(n+1)}$ choices for the linear polynomial l. Observe that this argument is valid only if $\sqrt[d]{\alpha}$ is properly defined, like when $\mathbb{F} = \mathbb{C}$, the set of complex numbers. ■

So, to fix the polynomial f, we have to choose s such polynomials g as defined in the above claim. Thus, total number of polynomials that can be written as $f = \sum_{i=1}^{s} \alpha_i l_i^d$ are $(q^{(n+1)})^s = q^{s(n+1)}$. So, there are $s.(n+1)$ degrees of freedom to express polynomials as sum of powers of s affine powers, and this must at-least be equal to the degrees of freedom to pick a random n-variate d-degree polynomial. Therefore, with high

probability, the number of summands s would be such that

$$\begin{aligned} s.(n+1) &\geq \binom{n+d}{d} \\ s &\geq \frac{1}{n+1}\binom{n+d}{d} \\ &\simeq n^d \end{aligned}$$

The last approximation is using Stirling's formula, and it can be seen this is clearly a tighter lower bound than the earlier $n^{d/2}$. Now, we try to find an explicit polynomial f for which the minimum number of summands required, comes close to the above proved lower bound.

12.4 Finding the explicit polynomial f

Formally, the challenge is to find an explicit n-variate polynomial d-degree polynomial f such that any representation of the form $f = \sum_{i=1}^{s} \alpha_i l_i^d$ (where l_i's are affine forms), requires s to be "large" (at-least exponential in d).

In order to find the required polynomial we try to explore what are some ways in which a polynomial of the from $f = l_d$, where $l = a_0 + a_1x_1 + a_2x_2 + ... + a_nx_n$ is an affine form, differs from any other arbitrary n-variate d-degree polynomial R.

Question 12.4 *Given the random polynomial R, in n variables having degree d (the degree d is known), describe an efficient algorithm that can output YES or NO, whether the given polynomial R is of the form $R = l^d$, where l is an affine form. (input polynomial is given explicitly as coefficients of monomials)*

For simplicity, let us consider the case when R is an univariate polynomial of degree d, i.e $n = 1$ and we have to output whether R is of the form $R(x) = (ax+b)^d = a^dx^d + \binom{d}{1}a^{d-1}bx^{d-1} + ... + b^d$. The first strategy that we can think of is to look at the coefficients of x^d and the constant term in the polynomial R, calculate their dth roots to get probable values of a and b respectively, and then check if the other coefficients follow the required pattern, i.e whether coefficient of x^i is equal to $\binom{d}{i}a^ib^{d-i}x^i$. The problem with the above strategy is when the underlying field is $\mathbb{F} = \mathbb{C}$, there can be d possible dth roots of a number, which means that we would need to check all the coefficients for all possible values of a and b which would not be an efficient process. There exists an efficient algorithm that answers Question 12.4 and leads us to finding the required "hard" polynomial, eventually proving a strong lower bound, but first we look at a little stronger problem:

Question 12.5 *Given a univariate polynomial $f(x)$ with coefficients from the field $\mathbb{F} = \mathbb{C}$, and an integer s, is there an efficient algorithm that outputs YES or NO, whether f can be written as*

$$f(x) = (a_1x + b_1)^d + (a_2x + b_2)^d + ... + (a_sx + b_s)^d \tag{12.2}$$

A trivial method to solve this problem is to think of $a_1, a_2, ..., a_s, b_1, b_2, ..., b_s$ as unknowns, and let $f(x) = c_0 + c_1x + c_2x^2 + ...c_dx^d$. On comparing the coefficients of x^i on both sides of equation 12.2, we can see that every coefficient of c_i is equal to an explicit polynomial $g_i(\underline{a}, \underline{b})$, where $\underline{a} = (a_1, a_2, ..., a_s)$ and $\underline{b} = (b_1, b_2, ..., b_s)$. This gives us a system of $d+1$ polynomial equations in $2s$ unknowns $(a_1, .., a_s, b_1, ..., b_s)$, degree of each equation being at-most d. Now, the best known result of solving such a system of polynomial equations is given by the following theorem:

Theorem 12.6 *Any system of polynomial equations in m variables,*

$$\begin{aligned} g_1(z_1, z_2, ...z_m) &= 0 \\ g_2(z_1, z_2, ...z_m) &= 0 \\ &. \\ &. \\ g_r(z_1, z_2, ...z_m) &= 0 \end{aligned}$$

where $degree(g_i) \leq e$ *for all* $i = 1, 2, ..., r$*, can be solved in time polynomial in* $r.e^m$*.*

The above result gives us an exponential time algorithm for solving question 12.5, and unfortunately finding a more efficient algorithm, that answers the question in lesser time, is an open problem. It is also easy to see that an efficient solution to the multivariate analog of question 12.5 would also answer question 12.4, as well as lead us to the required explicit polynomial to prove a strong lower bound.

Now, coming back to question 12.4, we were trying to distinguish between a random d-degree univariate polynomial R and a polynomial f of the form $f(x) = (ax+b)^d$. Let us observe the behaviour of the first-order derivatives of these polynomials. When we differentiate $f(x)$ with respect to x, we get $f'(x) = d.a.(ax+b)^{d-1}$, which means that the greatest common divisor(GCD) of f and f' is the polynomial $(ax+b)^{d-1}$. Thus,the degree of the polynomial $GCD(f, f') = d-1$, which is not the case, in general, for any arbitrary polynomial R.

Claim 12.7 *With high probability, degree of the polynomial* $GCD(R, R')$ *is zero.*

Proof: Let us first look at the simple case when R is a univariate polynomial of degree d (say). Assuming, the underlying field is such that all the roots or zeroes of the polynomial R exist, and are well-defined (for example, complex numbers). Thus, in general, let $\alpha_1, \alpha_2, ..., \alpha_d$ be the roots, and

$$R = (x - \alpha_1)(x - \alpha_2)...(x - \alpha_d)$$

where $\alpha_1, \alpha_2, ..., \alpha_d$ may not all be distinct. Differentiating the above equation, we get

$$R' = \frac{\partial R}{\partial x} = \frac{R}{(x - \alpha_1)} + \frac{R}{(x - \alpha_2)} + ... + \frac{R}{(x - \alpha_d)}$$

It can be observed that the polynomial $GCD(R, R')$ will have a root α_i iff α_i is a repeated root of R, i.e R has a factor $(x - \alpha_i)^k$ for some $k > 1$. Clearly, R and R' have no common roots if all of $\alpha_1, \alpha_2, ..., \alpha_d$ are distinct, which happens with very high probability, $p = \frac{d!.\binom{q}{d}}{q^d}$ if q is the cardinality of the underlying field $\mathbb{F}$. Thus, with the high probability p, for an arbitrary polynomial R, the polynomial $GCD(R, R')$ is a constant and therefore has degree equal to zero. The above argument can be extended to multivariate polynomials, where instead of one polynomial R', we would consider the first-order partial derivatives of R with respect to each of the variables, as we describe in the following paragraph. ■

Thus, we have the answer to question 12.4. The algorithm, when given the arbitrary polynomial $R(\underline{X})$ as input (where $\underline{X} = (x_1, x_2, ..., x_n)$), just has to compute partial derivatives of R with respect to each of the n variables seperately and then compute the GCD of the polynomials, $R, \frac{\partial R}{\partial x_1}, \frac{\partial R}{\partial x_2}, ..., \frac{\partial R}{\partial x_n}$, and if this GCD polynomial has higher degree then with high probability, we can output that R is of the form l^d where l is an affine form. This algorithm is known to run efficiently because of the following claim.

Claim 12.8 *Given two polynomials* f_1 *and* f_2 *with degrees* d_1 *and* d_2 *respectively, it is possible to compute the polynomial* $GCD(f_1, f_2)$*, in time approximately* $O(poly(d_1, d_2))$*.*

Proof: The proof of the above claim follows from an algorithm very similar in operation to the Euclidean Algorithm for computing GCD of integers. Again, if we consider f_1 and f_2 to be univariate polynomials, we have the polynomial $GCD(f_1, f_2)$ defined upto multiplication by a field constant, and the algorithm follows straight from the analogy between n-digit integers and n-degree univariate polynomials. As long division of one univariate polynomial, by another univariate polynomial is clearly defined just like integers, the Euclid's algorithm correctly computes the required GCD in time $O(poly(d_1, d_2))$.

The greatest common divisor is defined and exists, more generally, for multivariate polynomials over a field or the ring of integers, and also over a unique factorization domain. There exist algorithms to compute them as soon as one has a GCD algorithm in the ring of coefficients. These algorithms proceed by a recursion on the number of variables to reduce the problem to a variant of Euclid's algorithm.[4] ■

Now, coming back to our quest of finding the explicit polynomial f in n-variables of degree d, which we use to prove a lower bound on the number of summands s, when f is expressed as a sum of powers of affine forms $f = \sum_{i=1}^{s} l_i^d$. For any given polynomial $h \in \mathbb{F}[x_1, x_2, ..., x_n]$, we define the set of first-order partial derivatives of h,

$$\boldsymbol{\partial}^{=1} h = \{\frac{\partial h}{\partial x_1}, \frac{\partial h}{\partial x_2}, ..., \frac{\partial h}{\partial x_n}\}$$

Similarly, we can define the set of k'th order partial derivatives denoted by $\boldsymbol{\partial}^{=k} h$. Consider the polynomial $g = l^d = (a_0 + a_1x_1 + a_2x_2 + ... + a_nx_n)^d$, then

$$\begin{aligned}\boldsymbol{\partial}^{=1} g &= \{\frac{\partial g}{\partial x_1}, \frac{\partial g}{\partial x_2}, ..., \frac{\partial g}{\partial x_n}\} \\ &= \{a_1.d.l^{d-1}, a_2.d.l^{d-1}, ..., a_n.d.l^{d-1}\}\end{aligned}$$

If each of the n polynomials of degree $d-1$, are expressed as vectors, with entries corresponding to coefficients of particular monomials, as it has been described earlier, we define the $\mathbb{F} - span(\boldsymbol{\partial}^{=1} g)$ as the set of linear combinations of polynomials from $\boldsymbol{\partial}^{=1} g$ (where $f = l^d$), with coefficients coming from the field $\mathbb{F}$.

Claim 12.9 *The dimension,* $dim(\mathbb{F} - span(\boldsymbol{\partial}^{=1} l^d)) \leq 1$.

Proof: We have computed and seen that every polynomial in $\boldsymbol{\partial}^{=1} l^d$ is a constant multiple of the polynomial l^{d-1}, and that would be true for any linear combination of these polynomials. Thus, every polynomial in $\mathbb{F} - span(\boldsymbol{\partial}^{=1} l^d)$ would be a constant multiple of l^{d-1} which would imply that the vectors representing these polynomials too would be a constant multiple of the vector representing the polynomial l^{d-1}, which proves the claim as the l^{d-1} would be the single vector that forms the basis of the set $\mathbb{F} - span(\boldsymbol{\partial}^{=1} l^d)$. ■

We can make the same claim for kth order partial derivatives, as every polynomial on the set $\boldsymbol{\partial}^{=k} l^d$ would be a constant multiple of the polynomial l^{d-k}. For example,

$$\frac{\partial(a_0 + a_1x_1 + a_2x_2 + ... + a_nx_n)^d}{\partial x_1 x_2 ... x_k} = a_1a_2...a_k.d(d-1)(d-2)...(d-k+1).l^{d-k}$$

Similarly, other such polynomials in $\boldsymbol{\partial}^{=1} f$ would only have different value of the outer constants being multiplied, but would all be multiples of l^{d-k}, making the vector representing l^{d-k} as the basis of $\mathbb{F} - span(\boldsymbol{\partial}^{=k} l^d)$, hence we claim:

Claim 12.10 *The dimension,* $dim(\mathbb{F} - span(\boldsymbol{\partial}^{=k} l^d)) \leq 1$.

We make another important observation about this quantity, the dimension of $\mathbb{F} - span$ of k'th order partial derivatives:

Lemma 12.11 (Subadditivity) *For any two polynomials g and h,*

$$dim(\mathbb{F} - span(\boldsymbol{\partial}^{=k}(g+h))) \leq dim(\mathbb{F} - span(\boldsymbol{\partial}^{=k}g)) + dim(\mathbb{F} - span(\boldsymbol{\partial}^{=k}h)) \tag{12.3}$$

Proof: To prove this lemma,let us look at any polynomial $p \in \mathbb{F} - span(\boldsymbol{\partial}^{=k}(g+h))$. The vector representing p would be a linear combination of vectors from $\boldsymbol{\partial}^{=k}(g+h)$, i.e

$$p = \alpha_1 q_1 + \alpha_2 q_2 + ... + \alpha_m q_m \tag{12.4}$$

where $q_1, q_2, ..., q_m \in \boldsymbol{\partial}^{=k}(g+h)$. By expanding the partial derivative over $(g+h)$, we can write for all i, $q_i = r_{i1} + r_{i2}$ where r_{i1}, r_{i2} are polynomials in the sets $\boldsymbol{\partial}^{=k}g, \boldsymbol{\partial}^{=k}h$ respectively.

Let $d_1 = dim(\mathbb{F} - span(\boldsymbol{\partial}^{=k}g))$ and $d_2 = dim(\mathbb{F} - span(\boldsymbol{\partial}^{=k}h))$. Then, for all i, r_{i1} and r_{i2} are linear combinations of d_1 and d_2 independent vectors respectively, which makes every q_i also a linear combination of at-most $d_1 + d_2$ linearly independent vectors. Substituting the q_i's in equation 12.4, we see that every polynomial p is a combination of at-most $d_1 + d_2$ independent polynomials. This implies that the basis of the set $\mathbb{F} - span(\boldsymbol{\partial}^{=k}(g+h))$ is of size $\leq d_1 + d_2$, which completes the proof of the lemma. ■

The results stated above in Claim 12.10 and Lemma 12.11 lead us to a direct relation between the number of summands when a polynomial f is expressed as $f = \alpha_1 l_1^d + \alpha_2 l_2^d + ...\alpha_s l_s^d$, where the l_i's are affine forms. Applying subaddditivity (Lemma 12.11) on the expansion,

$$dim(\mathbb{F} - span(\boldsymbol{\partial}^{=k}f)) \leq \sum_{i=1}^{s} dim(\mathbb{F} - span(\boldsymbol{\partial}^{=k}l_i^d)) \tag{12.5}$$

from Claim 12.10, every term on the right hand side is ≤ 1,

$$dim(\mathbb{F} - span(\boldsymbol{\partial}^{=k}f)) \leq s \tag{12.6}$$

We have seen that the partial derivatives of any order obey specific patterns when the polynomial is of the form $g = l^d$, where l is an affine form. However, for an arbitrary random polynomial R in n variables, it can be said as a rough *heuristic*, "the partial derivatives behave like independent random polynomials". For example, consider the polynomial $f = \prod_{i=1}^{n} x_i$. Then,$\boldsymbol{\partial}^{=1}f = \{\frac{f}{x_1}, \frac{f}{x_2}, ..., \frac{f}{x_n}\}$, Similarly, $\boldsymbol{\partial}^{=2}f$ would be the set $\{\frac{f}{x_i x_j}\}$ for all $1 \leq i, j \leq n$ $(i \neq j)$. Extending the observed pattern, $\boldsymbol{\partial}^{=k}f$ would be the set of all multi-linear monomials of degree $(n-k)$ over the set of variables $\underline{X} = \{x_1, x_2, ..., x_n\}$. Note that all such monomials would be linearly independent as each monomial corresponds to a different subset of $(n-k)$ variables from $\underline{X}$. Thus,

$$dim(\mathbb{F} - span(\boldsymbol{\partial}^{=k}f)) = \binom{n}{n-k} = \binom{n}{k} \tag{12.7}$$

As the number of k'th order partial derivatives is exactly equal to number of ways of choosing a subset of $(n-k)$ variables out of n. Thus, combining equations 12.6 and 12.7, we get $s \geq \binom{n}{k}$, and to obtain a lower bound on s, the maximum value of $\binom{n}{k}$ is chosen i.e $k = n/2$. Hence, $s \geq \binom{n}{n/2} \simeq \frac{2^n}{2\pi\sqrt{n}}$ using approximation from the binomial distribution.

Thus, we can conclude that the lower bound on the number of summands is exponential in n (and in d) which is pretty close to the bound expected for an arbitrary polynomial, from the counting argument. Also, it is a good exercise to consider f as the Determinant polynomial and execute the same argument to obtain a lower bound on s for the Determinant polynomial.

12.5 References

[1] Amir Shpilka and Amir Yehudayaoff, Arithmetic Circuits: A survey of recent results and open questions, 2010

[2] W.J. Ellison, A waring's problem' for homogeneous forms. *Proceedings of the Cambridge Philosophical Society*, 65:663–672, 1969.

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

[4] Wikipedia, Polynomial GCD, *http://en.wikipedia.org/wiki/Polynomial_greatest_common_divisor*