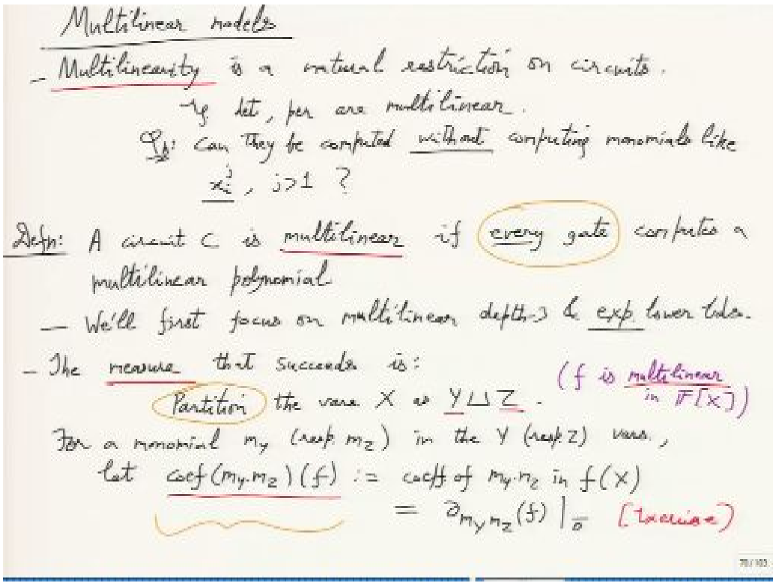
Arithmetic Circuit Complexity
Prof. Nitin Saxena
Department of Computer Science and Engineering
Indian Institute of Technology-Kanpur

Lecture - 16

So last time we started multilinear models, right.

(Refer Slide Time: 00:17)



Multilinear models

- Multilinearity is a natural restriction on circuits.
  - eg. det, per are multilinear.
  - Qn: Can they be computed without computing monomials like x_i^j, $j>1$?

Defn: A circuit C is multilinear if every gate computes a multilinear polynomial.

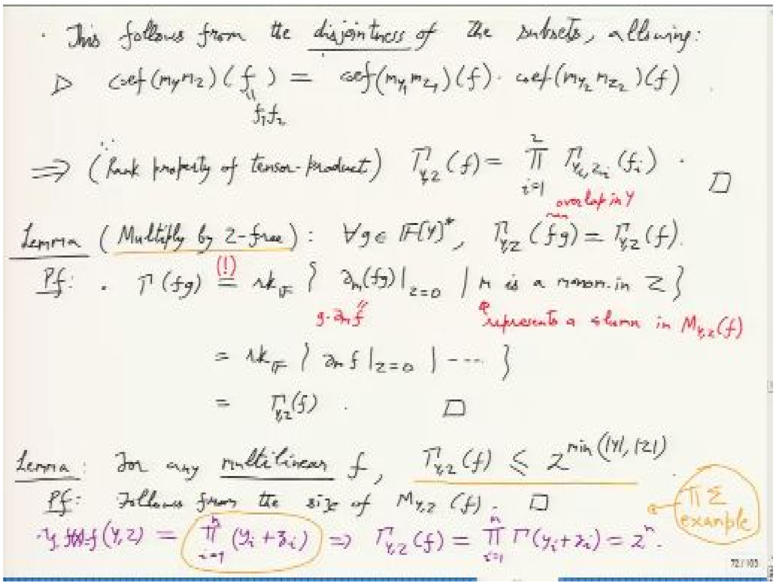
- We'll first focus on multilinear depth-3 & exp. lower bds.
- The measure that succeeds is:
  Partition the vars. X as $Y \sqcup Z$. (f is multilinear in $\mathbb{F}[X]$)

For a monomial m_y (resp. m_z) in the Y (resp. Z) vars.,
let $\mathrm{coef}(m_y \cdot m_z)(f) :=$ coeff of $m_y \cdot m_z$ in $f(X)$
$= \partial_{m_y m_z}(f)|_{\bar{0}}$ [exercise]

So this is, this will obviously compute a multilinear polynomial, but the condition is that every gate in the circuit should be computing a multilinear polynomial. So sometimes this is also called a syntactically multilinear circuit. So to study these models, the measure that we will look at is partition the variables x into y disjoint union z and then draw a matrix where you have rows y monomials columns z monomials and entry will be the coefficient; corresponding coefficient in the polynomial.

So this by the way does not need any circuit model. So this is really a definition for multilinear polynomials or even general polynomials. You can take any polynomial f and you can draw this matrix and then you look at the rank. So that rank is the measure - Γ.

(Refer Slide Time: 01:24)



· This follows from the disjointness of the subsets, allowing:

▷ $\text{coef}(m_y m_z)(\underset{f_1 f_2}{f}) = \text{coef}(m_{y_1} m_{z_1})(f) \cdot \text{coef}(m_{y_2} m_{z_2})(f)$

$\Rightarrow$ (Rank property of tensor-product) $\Gamma_{YZ}(f) = \prod_{i=1}^{2} \Gamma_{Y_i Z_i}(f_i)$. □

Lemma (Multiply by Z-free): $\forall g \in \mathbb{F}(Y)^*$, $\Gamma_{YZ}(\overset{\text{overlap in } Y}{fg}) = \Gamma_{YZ}(f)$.

Pf: · $\Gamma(fg) \overset{(!)}{=} rk_{\mathbb{F}}\{\underset{g\cdot\partial_m f}{\partial_m(fg)}|_{z=0} \mid m \text{ is a monom. in } Z\}$

represents a column in $M_{YZ}(f)$

$= rk_{\mathbb{F}}\{\partial_m f|_{z=0} \mid \cdots\}$

$= \Gamma_{YZ}(f)$. □

Lemma: For any multilinear f, $\Gamma_{YZ}(f) \leq 2^{\min(|Y|,|Z|)}$

Pf: Follows from the size of $M_{YZ}(f)$. □

· Eg. $f(Y,Z) = \prod_{i=1}^{n}(y_i + z_i) \Rightarrow \Gamma_{YZ}(f) = \prod_{i=1}^{n}\Gamma(y_i + z_i) = 2^n$.

ΠΣ example

72/103

So this is the measure, Γ with respect to a partition of f. So we have what we have shown is for two polynomials. Measure is upper bounded by the sum of the measures, respective measures. So that is sub additivity. We have shown that for two multilinear polynomials on disjoint variables there is exact multiplicative property. So the measure of the product is, product of the respective measures.

“Professor - student conversation starts” We have used multilinearity of the polynomial but we have not used multilinearity of the gate or something. No, these are not circuits. These are just polynomials. Till now the circuit has not been used. **“Professor - student conversation ends”.** So there is sub additivity, there is exact multiplicative property. And we showed that if we multiply by; if you multiply f by a polynomial that is either y-free or z-free. Then the measure does not change.

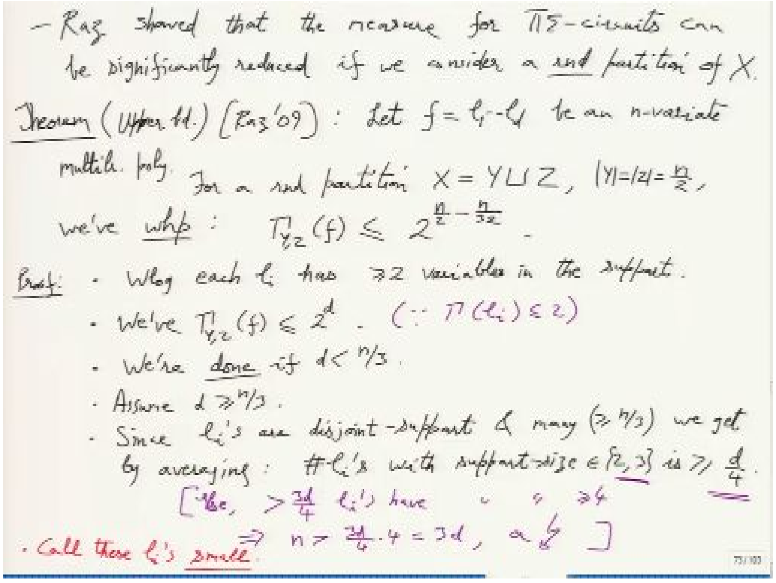
Yeah, this also did not need multilinearity. This is just, it is true for any polynomials. And finally, we showed that the measure never exceeds $2^{\min(|y|,|z|)}$ where $|y|, |z|$ are the number of rows and number of columns respectively. So rows and columns, minimum of that gives you the upper bound on the rank. And this is tight as we saw in this example.

In this product, you can see that the number of y monomials is 2^n and so is the z monomials and the rank is also that. So this is a depth - 2 example. That this is ΠΣ example. If you think in terms of circuits, then the circuit here is just product of sum and sum is just doing, adding two things two variables. Right, so what we now want to do is we want to show that although there is this $\Pi\Sigma$ example with measure as high as 2^n, maximum possible.

Still there is a way to, I mean there is a way to reduce this, okay. There is a trick by which we will use this measure on pi sigma circuits and show that the measure is small in some sense. So once we have shown this then we will the opposite of this will be that we take determinant and show that the similar trick gives a high measure, okay. So what is the trick? So the trick is that you do not partition in this way.

So the partition that we chose was actually the worst possible. So it gave you the maximum possible value of the measure. What if you change the partition? So do not follow these y_i, z_i's partitions, but let us say randomly pick a partition. So if you do that, then you will see that the measure actually falls. That is what made this useful for proving lower bounds.

(Refer Slide Time: 05:29)



– Raz showed that the measure for $\Pi\Sigma$-circuits can be significantly reduced if we consider a rnd partition of X.

Theorem (Upper bd.) [Raz'09]: Let $f = \ell_1 \cdots \ell_d$ be an n-variate multil. poly. For a rnd partition $X = Y \sqcup Z$, $|Y| = |Z| = \frac{n}{2}$, we've whp: $\Gamma_{Y,Z}(f) \leq 2^{\frac{n}{2} - \frac{n}{32}}$.

Proof:
- Wlog each ℓ_i has ≥ 2 variables in the support.
- We've $\Gamma_{Y,Z}(f) \leq 2^d$. ($\because \Gamma(\ell_i) \leq 2$)
- We're done if $d < n/3$.
- Assume $d \geq n/3$.
- Since ℓ_i's are disjoint-support & many ($\geq n/3$) we get by averaging: # ℓ_i's with support-size $\in \{2, 3\}$ is $\geq \frac{d}{4}$.
  [i.e., $> \frac{3d}{4}$ ℓ_i's have " " ≥ 4 $\Rightarrow n > \frac{3d}{4} \cdot 4 = 3d$, a $\lightning$]
- Call these ℓ_i's small.

73/103

So Ran Raz showed that the measure for $\Pi\Sigma$ circuits can be significantly reduced if we consider a random partition of the variable set. Remember that this matrix that we draw depends on how we partition the variables. There are exponentially many ways to partition the variables. So even if there is a single partition that reduces the measure it can be good. In fact, we will show that most of the partitions will be of this type. The measure will be quite small, it will not be sub exponentially small, but it will be much smaller than 2^n.

So that is the theorem - Upper Bound. So consider $\Pi\Sigma$ polynomial or circuit. So it is a product of linear polynomials, d many, on n variables and importantly they should have disjoint support, so that f is multilinear. Let f be an n-variate multilinear polynomial. For a random partition, and let us say these are equal parts, we have with high probability. The measure will be smaller than $2^{n/2 - n/32}$.

So the exponent falls just by a constant multiple. It is not a big fall, right. But since it is happening in the exponent, I mean overall it can be useful because you just have to now show that for some other polynomial like determinant, the measure is greater than this. So now you have some gap that you can utilize. With $2^{n/2}$ you had no gap. So now, this creates a gap that we can try to exploit later.

Once we have shown for product $\Pi\Sigma$ for $\Sigma\Pi\Sigma$ also we have this similar upper bound because of sub additivity right. So this upper bound is really for multilinear depth - 3 circuits, right. This measure is that powerful. It can actually analyze properties of multilinear depth - 3. So the way we will prove this is it will not be an easy proof.

We will essentially show that if you randomly pick y, this part y then many of the l_i's will either be y-free or z-free. If l_1 is y-free, then we know that it will not contribute anything to the rank to the measure. So if we can show that a large number of l_i's drop out of the race, that will give you the difference of $n/32$. So that can be shown probabilistically.

So let us assume first of all that each l_i has greater than equal to two variables in the support. If not, then they anyways do not play a role. So we do not consider univariate l_i's. So l_i's are at least bivariate or more. That we can assume. So now this implies that the measure for f is bounded by, this anyways I mean, this was always true. This is just because f is a product of d many. The rank of l_1 is at most 2.

It is a linear polynomial. So rank of each l_i is at most 2. So just product is 2^d. This does not need any assumption. This is true anyways. In fact I should say just we have this always since and then multiplicativity. So what this means is if d is too small, if it is smaller than this $n/2 - n/32$ then anyways we are done. So we are done if, in particular d is less than $n/3$.

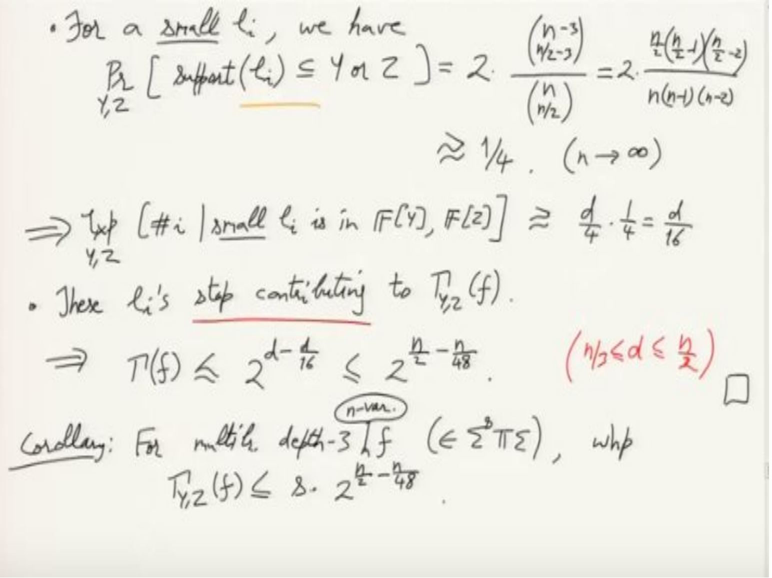
Because then you get an upper bound of $2^{n/3}$, which is even smaller than what you wanted to show. So we will assume that d is bigger than $n/3$ or at least $n/3$. And each l_i has at least two variables and l_i's have disjoint variables. So this gives you a constraint system. So since l_i's are disjoint support and many. So we get by an averaging argument that number of l_i's with support size 2 or 3.

So basically it is smaller than 4. The number of l_i's with support size less than 4 is at least d by 4. Why is that? Else you will get that greater than $3d/4$. So if it is less than 3d/4, then more than $3d/4$ many l_i's have support size 4 or more. Which would imply that the number of variables is greater than $3d/4 \cdot 4$ which is 3d. But we have assumed that these at least $n/3$.

So that is a contradiction. The lightning strikes. So since l_i's are many and they are disjoint they cannot each have many variables. So in other words, many of them many of the l_i's have minimum support possible which is 2. So we get for 2 and 3, we get this quantitative estimate of $d/4$. So we call these l_i's small. And we will only care about them in our probabilistic argument.

We know that these are present. So we can assume that the initial $d/4$ have support less than 4. They only have 2 or 3 variables.

(Refer Slide Time: 16:26)



- For a small l_i, we have

$$\Pr_{Y,Z}\left[\underline{\text{support}(l_i)} \subseteq Y \text{ or } Z\right] = 2 \cdot \frac{\binom{n-3}{n/2-3}}{\binom{n}{n/2}} = 2 \cdot \frac{\frac{n}{2}\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)}{n(n-1)(n-2)}$$

$$\gtrsim 1/4 \,. \quad (n \to \infty)$$

$$\Rightarrow \operatorname{Exp}_{Y,Z}\left[\#i \mid \underline{\text{small}}\ l_i \text{ is in } \mathbb{F}[Y], \mathbb{F}[Z]\right] \gtrsim \frac{d}{4} \cdot \frac{1}{4} = \frac{d}{16}$$

- These l_i's stop contributing to $T_{Y,Z}(f)$.

$$\Rightarrow T(f) \lessapprox 2^{d-\frac{d}{16}} \leqslant 2^{\frac{n}{2}-\frac{n}{48}} . \qquad (n/3 \leqslant d \leqslant \tfrac{n}{2}) \qquad \Box$$

Corollary: For multil. depth-3 $\overset{n\text{-var.}}{f}$ $(\in \Sigma^{s}\Pi\Sigma)$, whp

$$T_{Y,Z}(f) \leqslant s \cdot 2^{\frac{n}{2}-\frac{n}{48}} .$$

l_i we have, so what is the probability that when you randomly choose y of size, so y has $n/2$ variables. So that fixes z also. Then this particular l_i is either both all the three variables or all the two variables are in y, they fall in y or they fall in z. So whatever is the probability it is a constant.

So variable x_1 will fall in y or z with equal probability. **"Professor - student conversation starts"** Right, but then they would not be independent. Variable x_2 will be independent because we have the additional constraint that y has size $n/2$. So it is not like, it is not the case that for each variable, you are independently picking whether it falls in y or z. Because they are only in the expected size of x,y and z is $n/2$.

Maybe there are additional constraints. So you are undergoing independently another, amongst all subsets of size $n/2$ you pick one at random. It should be slightly different. No, but I want to write here, so all the three variables falling in y. So that will just be $n/2$ which is 3? Right, so I have to write that expression then. So maybe I give the exact expression. **"Professor - student conversation ends".**

So it will be what? For falling in y all three is; here I will the number of ways of choosing y is $\binom{n}{n/2}$ and if 3 fall in y then $n-3$ is it? But I do not like this. **"Professor - student conversation starts"** That expression is better, $\binom{n/2}{3}/\binom{n}{3}$.. No this seems to be the only expression. x_1, x_2, x_3 you have put in y and then what remains is just $n/2-3$.

That expression works. Because the total number of triples are y contains. No, but this argument has to come from this. So what is this? But this argument is right. The total number of triples that y contain. No, but I do not want to erase this. No, I do not want to erase this now. Forget this. This is the simplest expression. **"Professor - student conversation ends".**

So this denominator is $n(n-1)(n-2)$ and then we get numerator $n/2(n/2-1)(n/2-2)$. So this we can see is, is it more than $1/8$? Should be. Slightly smaller than half. Let me then not try to not venture into its exact calculation. Just say that this is very close to $1/4$. So n is asymptotically growing. So as n tends to infinity this is just $1/4$..

I mean, I we just wanted some constant probability. That we have got. And so this means well, I will kind of need this $1/4$ for the subsequent calculation. So let me continue with this $1/4$. So this gives me the expectation. So when you pick y randomly then l_i is either completely in y or completely in z. That probability comes out to be around $1/4$. And so the number of i's amongst these small l_i's.

So we have shown that the probability of l_i being y-free or z-free for a fixed l_i and small is at least $1/4$. That gives us the expectation of $1/4$, which is the probability times $d/4$, which is the number of small l_i's that we have seen before, this number. So the expectation of these small l_i's that ultimately will not contribute to the rank or the measure. This is, the number of these i's is expected to be d by 16.

So these l i's stop contributing to gamma y, z f, right. So essentially out of d, $1/16$ are not contributing. So that gives us the gap in the rank. So this implies that $\Gamma(f) \leq 2^{d-d/16}$. For now I have to use a fancier notation. Then this $2^{n/2-n/16}$ is at most, so $d \leq n/2$. This is because we have assumed every l_i has at least two variables and they are disjoint variables.

And $d \geq n/3$. So that gives us, so that will give us so, let us write down the properties of d.

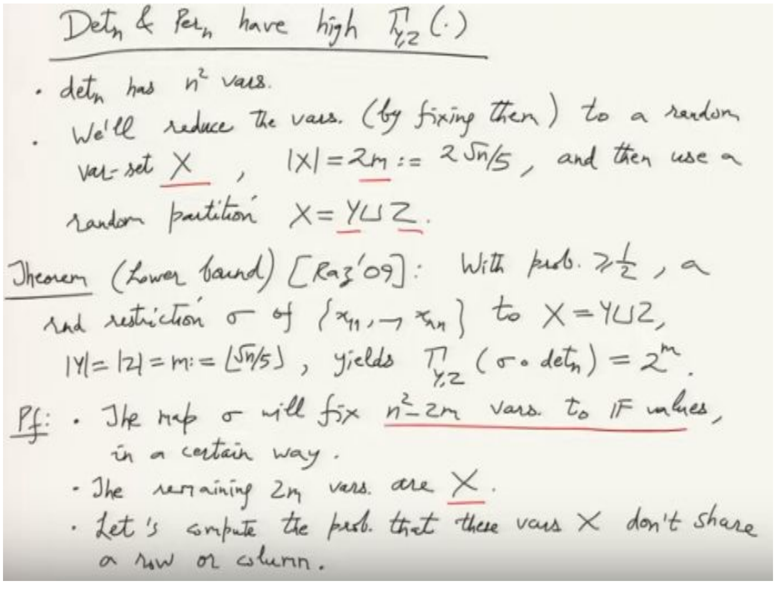
$$n/3 \leq d \leq n/2$$

So what do we get? We get $\Gamma(f) \leq 2^{n/2-n/48}$. Yeah, that is what we should get. So we have to change the theorem statement accordingly. But then we get, we get this gap of $n/48$.

And right so we can write a corollary here that for multilinear depth - 3 circuit f with s product gates. So let us just use $\Sigma\Pi\Sigma$, where the top fanin is s. So for this with high probability the measure is at most s times this. Right. So multilinear depth - 3 and n-variate which has s many gates then you get by sub additivity $s \cdot 2^{n/2-n/48}$.

So that is the upper bound and now if you pick your s suitably, which will be determinant and show that the measure is large. So that will give you s bigger than some bound which we will show it to be exponential. So let us now go to the other part, which is why does determinant have a large measure?

(Refer Slide Time: 27:08)



Det_n & Per_n have high $\Pi_{Y,Z}(\cdot)$

- $\det_n$ has n^2 vars.
- We'll reduce the vars. (by fixing them) to a random var-set $\underline{X}$, $|X| = \underline{2m} := 2\sqrt{n}/5$, and then use a random partition $X = \underline{Y} \sqcup \underline{Z}$.

Theorem (Lower bound) [Raz'09]: With prob. $\geq \frac{1}{2}$, a rnd restriction σ of $\{x_{11}, \ldots, x_{nn}\}$ to $X = Y \sqcup Z$, $|Y| = |Z| = m := \lfloor\sqrt{n}/5\rfloor$, yields $\Pi_{Y,Z}(\sigma \circ \det_n) = 2^m$.

Pf:
- The map σ will fix $\underline{n^2 - 2m \text{ vars. to } \mathbb{F} \text{ values}}$, in a certain way.
- The remaining $2m$ vars. are $\underline{X}$.
- Let's compute the prob. that these vars X don't share a row or column.

So determinant and permanent of $n \times n$ matrix have high Γ. So $\det_n$ has n^2 variables. Any ideas how to show that the measure is large. Again if you are not careful with the partition right so n^2 variables yeah if you are not careful then the measure can be really small. What do you mean symmetric? The partition is not symmetric. You are picking a subset of the variables.

So I mean, you can come up with small examples where when you pick different partitions, the measure will change wildly. So we will not go into those nitty gritties. We will just continue with our probabilistic argument that probabilistically if you choose y, then the measure will be high. I do not think it will be an optimal result. So what we will actually do is we will set a large number of the variables to zero.

And in the remaining variables, we will call the remaining variables x and that is where we will do the partitioning. So we will reduce the variables by fixing them mainly to zero. So let me not say where we are fixing. So we will reduce the number of variables to a random x, to a random variable set x which will be much smaller. So now this x will be of size 2m, which is around $\sqrt{n}$.

We will basically set a large number of variables to be zero and we will be left with around $\sqrt{n}$ variables out of n^2. It is a big fall, and then use a random partition. So this does not look optimal. Because we have killed many variables. Ideally we should have reduced to somewhat n many variables. But for the proof to work we will do we will go to $\sqrt{n}$.

It probably is related to the birthday paradox. So the probabilistic calculations will need $\sqrt{n}$. So that is the definition of x m, and y z. That is how we will do it and then we will show that the measure Γ with respect to this of evaluated determinant or restricted determinant is high. Determinant we started with n^2 variables, we reduced them to n to m or 2m.

And if the determinant had a depth - 3 circuit, then they are also the variables will reduce to 2m. So in the previous lemma, we are getting something like a gap of $m/48$ or whatever $n/24$. And that is what we are up against. So we have to show that the determinant has measure just more than that to get a lower bound on s. So we will prove this theorem, which will be enough. So this is due to Raz from same paper.

So with probability at least 50% a random restriction σ of the variables n^2 variables to x, x of size around $2^{\sqrt{n}/5}$ yields a large measure not determinant, but this restricted by Σ. So this measure will be exactly 2^m. So remember that there are 2m variables. So this is like in the exponent you are getting half of that. And so when you compare this with the previous theorem, the difference which you got that is basically giving you a lower bound on s.

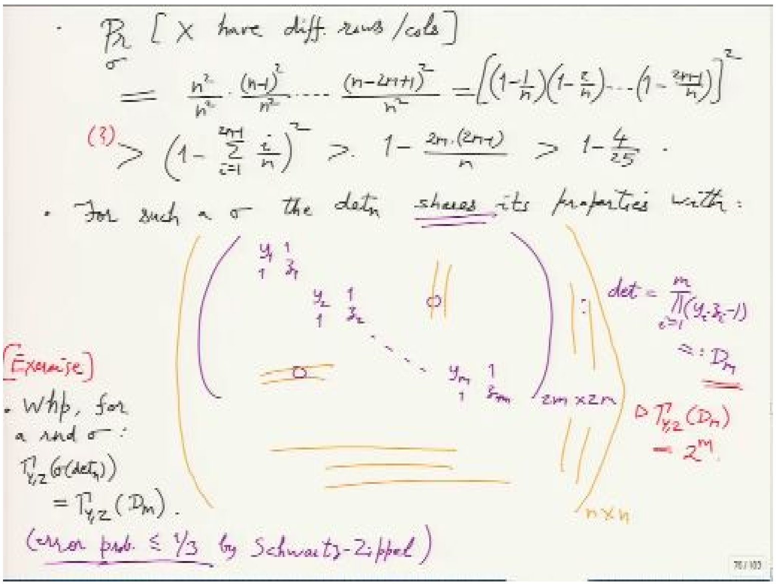
So it is exponential but since m is $\sqrt{n}$, you get that s is at least $2^{\sqrt{n}}$. So it is not, it does not seem optimal. But still it is kind of exponential, it is an exponential lower bound to compute $n \times n$ determinant k. Is the statement clear and the connection between the lower bound and the upper bound theorems.

So the way we will prove this is yeah the thing which we always do with determinant is to give a lower bound on the rank we somehow identify triangular or diagonal structures. And using that we show a rank lower bound. So that will be done essentially by σ. σ will fix the variables in such a way that you see a diagonal structure, diagonal block structure, which is easy to rank bound.

So the map σ will fix $n^2 - 2m$ variables to f values in a certain way. The remaining variables, we are calling them x. So let us compute the probability that these remaining variables, they do not share a row or a column. So if you have variables v_1 and v_2 in x, then we want v_1 v_2 to be in different rows and in different columns.

This is again the intuition is to get somewhat diagonal block structure. So let us compute the probability that these variables x do not share a row or column. So this is kind of a two dimensional version of birthday paradox.

(Refer Slide Time: 37:32)



- $\Pr_{\sigma}$ [X have diff. rows / cols]

$$= \frac{n^2}{n^2} \cdot \frac{(n-1)^2}{n^2} \cdots \frac{(n-2m+1)^2}{n^2} = \left[\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{2m-1}{n}\right)\right]^2$$

$$\overset{(?)}{>} \left(1-\sum_{i=1}^{2m-1}\frac{i}{n}\right)^2 > 1-\frac{2m\cdot(2m-1)}{n} > 1-\frac{4}{25}.$$

- For such a σ the $\det_n$ shares its properties with:

$$\det = \prod_{i=1}^{m}(y_i z_i - 1) =: D_m$$

$$\triangleright \; T_{y,z}(D_m) = 2^m.$$

[Exercise]

- Whp, for a and σ: $T_{y,z}(\sigma(\det_n)) = T_{y,z}(D_m)$.

(error prob. $\leq 1/3$ by Schwartz-Zippel)

So σ is random evaluation of that many variables and you want x, the remaining variables unfixed variables x, they have different rows and columns. So this probability exactly is well first so okay we are looking at this process as if picking the variables of x. Not fixing variables, but just picking these free variables. So the first free variable is free. You have n^2 favorable choices.

But once you have picked it then you have picked a row and a column. So they are gone. So now you are left with $(n-1)^2$ possibilities and this then $(n-2)^2$ and so on. And you want to pick no so in terms of m. So $(n-2m+1)^2$. So that will give you the 2m^th variable. All possibilities are m^2. So this is the same expression you get in birthday paradox squared right.

And so you have this intuition that m should be on the lesser side of $\sqrt{n}$. If it is on the larger side of $\sqrt{n}$, then this will be a problem. So that is why we needed m to be $\sqrt{m}/5$. So this is greater than $\left(1-\sum_{i=1}^{2n-1}\frac{i}{n}\right)^2$ well, this will need a slight calculation which I am skipping. But the intuition is that 1 and then you have a negative term, then you have a positive term and so on.

Yeah, so do that as an exercise. That is the question mark. And so the larger contribution comes from the first two terms and the remaining will only increase it. So it is the $\sum \frac{i}{n}$ that you have to estimate then, which is for which you know the lower bound. So that gives you an upper bound for which you know upper bound. So this will give you a lower bound $1 - \frac{2m(2m-1)}{n}$.

So for the square we take the two inside and then $\sum i$ is what we have estimated. And this comes out to be more than $1 - 4/25$. So this is a huge probability. With a very large probability x will have different rows and columns, variables in x. So what is the advantage of that? So for, so these are the free variables and the remaining are all fixed to constants. So how does your determinant matrix look like?

So for such a σ the determinant n shares its properties with the following structure. So this is only $2m \times 2m$. So this will need some explanation, but I will not go into all the details. So you started with the $n \times n$ matrix. You fixed a lot of variables leaving only 2m variables free.

So assuming that these variables, remaining variables or free variables are organized in a diagonal fashion right this is the $2m \times 2m$ matrix is the matrix which will give you which is the kind of the symbolic part. Everything else in the original $n \times n$ matrix is fixed to constant. And they are fixed in a random way. So that is because I just want to say that it shares its properties. It is not exactly this.

So the zeros are also random values in general, for general σ. But look at this matrix. So this matrix has determinant? So its determinant is equal to $\prod_{i=1}^{m}(y_i z_i - 1)$. So if you get to this matrix, then its determinant is something which we understand and that is

$\prod_{i=1}^{m}(y_i z_i - 1)$. So let us call it D$_m$. So this is really the variable part and for this random sigma what you will get is this multiplied by some constant and plus another constant.

So when you will look at the measure you are really computing the measure of this polynomial, this 2m variant polynomial. So to work out the details, you have to assume that σ was fixing the other variables to random points. That will be needed. So in general you will have I mean, thing that I have skipped is this. This is a slightly; this is a bigger matrix where here you have constants and here also you have constants.

And these are also constants and not zero. But when you will start differentiating this the measure that you will get that measure will be lower bounded by the measure of D$_m$. That is the thing to realize. From now on we will just discuss $\Gamma(D_m)$. This is $n \times n$ and we have identified a sub matrix and the determinant of that and let us look at the measure.

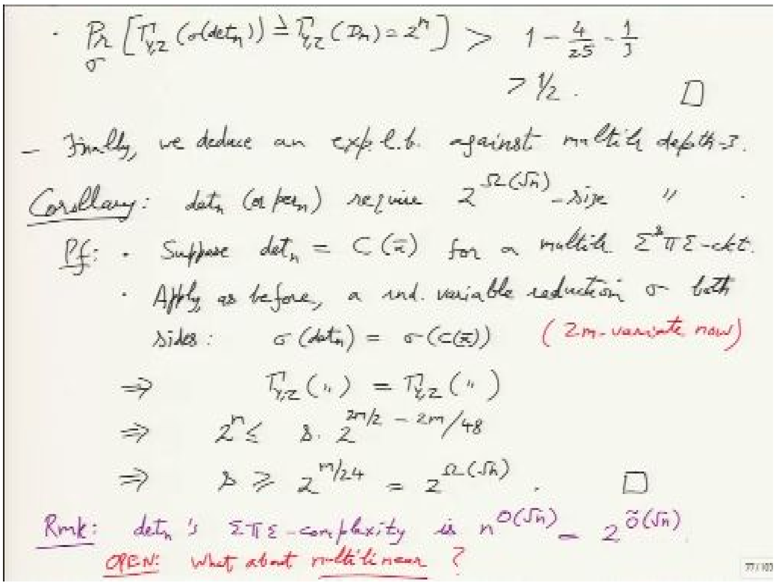
So this really is the reason this picture is the reason why we are doing all this fixing of variables and keeping 2m variables free in the end. So any question about the overall idea? The details I leave. This picture should be worth a thousand words. This may be the exercise that with high probability for a random σ, $\Gamma_{Y,Z}(\sigma(det_n)) = \Gamma_{Y,Z}(D_m)$.

And what is this high probability. So this how much is the error here? I think if you assume the field to be larger than n in size sorry, no for this for this event that $\Gamma(\sigma(det))$ is $\Gamma(D_m)$. The probability of this happening is given by the Schwartz–Zippel lemma. So basically n over the field size is the error. So if you assume field size to be bigger than n, 2n then you get probability half or probability one-third.

So we can assume here that error probability is less than equal to one-third by a Schwartz–Zippel kind of argument. Yeah, so whatever I have said this will happen by random fixing with probability at least two-third. So at least two-third probability we can reduce the study of $\Gamma(\sigma(det))$ to $\Gamma(D_m)$. And then $\Gamma(D_m)$ is something that we can calculate.

So what is $\Gamma(D_m)$? It is like y_i + z_i calculation that we did. So it is 2^m. Remember that y is the set of y_i's and z is the set of z_i's. So for this polynomial the partition is the worst possible. It maximizes the measure. So you get 2^m. So that is the lower bound for determinant, measure lower bound for the determinant. The last thing is we have to be sure about the probability that the error probability is not too large.

(Refer Slide Time: 49:24)



- $\Pr_{\sigma}\left[\Gamma_{Y,Z}(\sigma(det_n)) \geq \Gamma_{Y,Z}(D_m) = 2^m\right] > 1 - \frac{4}{25} - \frac{1}{3}$ $> 1/2$. □

– Finally, we deduce an exp. l.b. against multil. depth-3.

Corollary: det_n (or per_n) require $2^{\Omega(\sqrt{n})}$-size " .

Pf: • Suppose $det_n = C(\bar{x})$ for a multil. $\Sigma^s\Pi\Sigma$-ckt.

• Apply as before, a rnd. variable reduction σ on both sides: $\sigma(det_n) = \sigma(C(\bar{x}))$ (2m-variate now)

$\Rightarrow$ $\Gamma_{Y,Z}(\text{"}) = \Gamma_{Y,Z}(\text{"})$

$\Rightarrow$ $2^m \leq s \cdot 2^{2m/2 - 2m/48}$

$\Rightarrow$ $s \geq 2^{m/24} = 2^{\Omega(\sqrt{n})}$. □

Rmk: det_n's $\Sigma\Pi\Sigma$-complexity is $n^{O(\sqrt{n})} = 2^{\tilde{O}(\sqrt{n})}$

OPEN: What about multilinear ?

77/103

So let us check that.

$$Pr_{\sigma}\left[\Gamma_{Y,Z}(\sigma(det_n) \geq \Gamma_{Y,Z}(D_m) = 2^m\right] > 1 - \tfrac{4}{25} - \tfrac{1}{3} > 1/2.$$

To be sure I have put an inequality here. I only care about this. Do not want them to be exactly equal. So probability that the measure of determinant after restriction is at least 2^m. So that picture holds. So that depends on two things.

One is that choice of x variables should be in this diagonal fashion and second is fixing of the variables to constant should be good. If you do an unfortunate fixing for example, you set everything to zero, right then the rank will just fall, D_m will not help. So those two errors we have to check. So choice of x carefully is 4/25 error and fixing of the remaining variables is error one-third, right. So this is still greater than half.

So that is the probability of the lower bound, okay. So there are at least half of the choices of y, give you a large measure. Where by large we mean for n^2 variables we are just getting $2^{\sqrt{n}}$ lower bound. So this is not really optimal. I would have preferred 2^n. But the probabilistic argument goes via birthday paradox and so you get slightly worse.

You have to study probabilistically right; this the upper bound was a probabilistic upper bound. So then lower bound also should be probabilistic. Otherwise, how will you match the two? And if you try to do this probabilistically, then you get into the birthday paradox. So I am not sure whether this $2^{\sqrt{n}}$ is optimal. Is there an example somewhere? I do not think this is resolved.

Right, so now we are set for the, you can connect the two. So we deduce an exponential lower bound against multilinear depth - 3. So yeah, so we did only for determinants, but you can see that this argument does not really care about the difference. So permanent also will, once you have this diagonal block structure permanent will behave as well. It will give you just $y_i z_i$ + 1. So that will not change anything.

So this is really trivially the same, directly the same proof. So determinant or permanent require $2^{\Omega(\sqrt{n})}$ size multilinear depth – 3. So this is what you will get, but I do not think we know any such constructions. So we do not have a multilinear depth - 3 representation for determinant available. Neither for determinant nor for permanent, right over any characteristic field.

So for permanent what you know is multilinear depth – 3 of size 2^m. And for determinant I cannot even think below $n!$, which is n^n. So we do not really have such constructions. This most probably is not optimal. But still it is a strong lower bound. Now suppose determinant has such a representation, suppose determinant is $C(\bar{x})$ for multilinear $\Sigma^s\Pi\Sigma$ circuit.

So then apply as before as described before a random variable reduction sigma both sides, right. So you will get $\sigma(det_n) = \sigma(C(\bar{x}))$. So this will bring down the variables from n^2 to around $\sqrt{n}$, 2m. So this becomes 2m-variate now. And then you apply the measure, both sides. So on one hand you will see that Γ of determinant in the probabilistic space of choices, this is with high probability 2^m.

On the other hand $\Gamma(\sigma(C(\bar{x})))$. So note that $\sigma(C(\bar{x}))$ is still $\Sigma^s\Pi\Sigma$. So multilinear depth - 3, this is quite important that multilinear depth - 3 structure has not changed by σ because σ is either fixing the variables or it is keeping them free so this will not change. I mean, this can only reduce the product fanin. It cannot increase it and it will not change the $\Sigma\Pi\Sigma$ structure.

So we can apply the upper bound which with high probability comes out to be $s \cdot 2^{2m/2 - 2m/48}$, since the number of variables is 2m. Note that this is quite precarious, so the 2^m cancels both sides. So whatever is the minus part that is the lower bound on s. So this gives you $s \geq 2^{m/24}$, which is just $2^{\Omega(\sqrt{n})}$

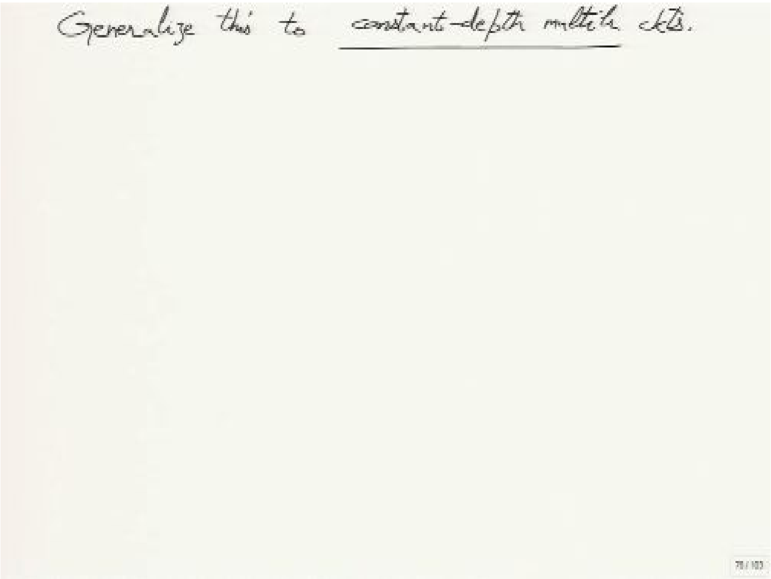
Did I make a mistake in the representation claim? Yeah, that was wrong. So for determinant actually, by depth - 3 reduction, you will get a depth - 3 circuit of size n to the $\sqrt{n}$. But it will not be multilinear. Yeah it is not multilinear but if you look at the size bound, then this $2^{\sqrt{n}}$ is close. What? No depth – 3, to come to depth – 3 you have to use a duality trick.

Those things are inherently non multilinear. Otherwise you will keep on hanging at depth – 4. So determinant it is $\Sigma\Pi\Sigma$ complexity is $n^{O(\sqrt{n})}$. So this we know over good fields assuming the characteristic large or zero. We know that there is a $\Sigma\Pi\Sigma$ representation but it is not multilinear. So that is open but still it is interesting that the quantitative bound you are getting is $2^{\sqrt{n}}$, which is kind of matching with n to the $\sqrt{n}$.

But qualitative difference is that this multilinear restriction is. Sorry? N to the log n? What is log n? In the exponent the gap is of log n. Yeah, so it is $\tilde{O}(\sqrt{n})$. Yeah, so if you can compare $\tilde{O}$ with Ω, then it is the same. But we do not have a multilinear representation. That seems very unlikely coming to depth - 3 with multilinear.

So anyways, this is the status. So next thing we should do next time, which is tomorrow. So tomorrow what we will do is we will generalize this to higher depths.

(Refer Slide Time: 1:01:24)



Generalize this to constant-depth multilin ckts.

78 / 103

To constant depth multilinear circuits, okay. So the point of doing $\Sigma\Pi\Sigma$ first was to give you a baby introduction to the proof technique, and then the real thing will be going to constant depth and then finally going to formulas, multilinear formula. So we will show that determinant if you try to express it as a constant depth multilinear circuit or like a multilinear formula, then it has to be super polynomially large.

Some more combinatorics will be required for that, basically in the upper bound part. The lower bound part is the same. In the upper bound part we have to show that even these constant depth multilinear model or multilinear formula models even they are weak in this measure sense.