Lecture 14

Topics in Complexity Theory and Pseudorandomness (Spring 2013)
Rutgers University
Swastik Kopparty
Scribes: Amey Bhangale, Mrinal Kumar

In this lecture, we will discuss following topics

- 2-source extractors
- Pseudorandom generator for polynomials over $\mathbb{F}_2$
- Super concentrators

1 2-Source Extractor

Definition 1. *(k, ϵ) 2-source extractor is a function*

$$E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^m$$

such that if X and Y are independent random variables with $H_\infty(X) \geq k$, $H_\infty(Y) \geq k$ then $E(X, Y)$ is ϵ-close to the uniform distribution U_m

The important difference between 2-source extractor and the extractor which we saw in previous lectures is that 2-source extractor doesn't require any *pure* random bits as an input.

Exercise : (Existence of a 2-source extractor) Show that a random function E with $m \leq k - O\left(\log \frac{1}{\epsilon}\right)$ is a (k, ϵ) 2-source extractor.

1.1 2-source extractor with $m = 1$

A 2-source extractor for $m = 1$ case is itself very interesting. In this lecture we will only study these extractors. We can associate a bipartite graph with a 2-source extractor as follows:

Let $G(L, R, E)$ be a bipartite graphs on $L = R = \{0,1\}^n$ with an edge between $X \in L$ and $Y \in R$ if and only if $E(X, Y) = 1$. Then the 2-source extractor condition says that for all $S \subseteq L, T \subseteq R$, $|S| = |T| = 2^k$, the number of edges between S and T is roughly half the total number of edges i.e. $\left(\frac{1}{2} \pm \epsilon\right)|S||T|$.

Theorem 2. *(Chor,Goldreich 80's) Define $E : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ as:*

$$E(X, Y) = \langle x, y \rangle_{\mathbb{F}_2}$$

Then E is an $\left(\frac{n}{2} + 2\log \frac{1}{\epsilon}, \epsilon\right)$ extrator.

Proof. We will only prove the theorem with $2\log\frac{1}{\epsilon}$ replaced by $3\log\frac{1}{\epsilon}$, because someone suggested a pretty good proof strategy in class, and this is what that proof strategy gave. Oh well.

To begin, keeping the associated biaprtite graph picture in mind, we would be interested in knowing given $T \subseteq \{0,1\}^n$, how many $x's$ are there such that the following condition holds:

$$\Pr_{y\in T}\left[\langle x, y\rangle = 0\right] \text{ is } \epsilon \text{ far from } \frac{1}{2} \tag{1}$$

Above condition is equivalent to asking, how many $x's$ are there satisfying:

$$\left|\frac{1}{|T|}\sum_{y\in T}(-1)^{\langle x,y\rangle}\right| > \epsilon \tag{2}$$

Define a function f as follows:

$$f : \{0,1\}^n \to \mathbb{R}$$
$$f = \begin{cases} 1 & \text{if } y \in T \\ 0 & \text{otherwise} \end{cases}$$

Then, equation 2 becomes,

$$\frac{1}{|T|}\left|\langle f, \chi_x\rangle\right| > \epsilon$$
$$|\hat{f}(x)| > \epsilon|T| \tag{3}$$

We know that $\|f\|^2 = |T|$, so expanding L.H.S by using Parseval's Identity, we get

$$\frac{1}{2^n}\sum_{x\in\{0,1\}^n}\langle f, \chi_x\rangle^2 = |T|$$

$$\sum_{x\in\{0,1\}^n}\hat{f}(x)^2 = 2^n|T| \tag{4}$$

Hence using equation 3 and 4, we get number of $x \in \{0,1\}^n$ satisfying equation 3 which is same as number of $x's$ satisfying equation 2 is atmost $\frac{2^n|T|}{\epsilon^2|T|^2} = \frac{2^n}{\epsilon^2|T|}$. Reformulating L.H.S of equation 2, we get

$$\left|\left\{x : \left|\mathop{\mathbb{E}}_{y\in T}(-1)^{\langle x,y\rangle}\right| > \epsilon\right\}\right| \leq \frac{2^n}{\epsilon^2|T|} \tag{5}$$

Now we are ready to prove the theorem. Suppose for contradiction, assume that E is not an $\left(\frac{n}{2} + 2\log\frac{1}{\epsilon}, \epsilon\right)$ extractor i.e. assume that there exists $S, T \subseteq \{0,1\}^n$ such that

$$|S| = |T| \geq 2^{\frac{n}{2}+2\log\frac{1}{\epsilon}} = \frac{2^{n/2}}{\epsilon^2} \tag{6}$$

and

$$\left|\mathop{\mathbb{E}}_{x\in S, y\in T}\left[(-1)^{\langle x,y\rangle}\right]\right| > \epsilon \tag{7}$$

Now, we claim that

$$\left|\mathop{\mathbb{E}}_{x\in S,y\in T}\left[(-1)^{\langle x,y\rangle}\right]\right| > \epsilon \implies \Pr_{x\in S}\left[\left|\mathop{\mathbb{E}}_{y\in T}(-1)^{\langle x,y\rangle}\right| > \frac{\epsilon}{2}\right] \geq \frac{\epsilon}{2} \tag{8}$$

If not,

$$\Pr_{x\in S}\left[\left|\mathop{\mathbb{E}}_{y\in T}(-1)^{\langle x,y\rangle}\right| > \frac{\epsilon}{2}\right] < \frac{\epsilon}{2}$$

Therefore, using the fact that for any $x \in \{0,1\}^n$ and $T \subseteq \{0,1\}^n$, $\left|\mathbb{E}_{y\in T}(-1)^{\langle x,y\rangle}\right| \leq 1$

$$\begin{aligned}\mathop{\mathbb{E}}_{x\in S}\left[\left|\mathop{\mathbb{E}}_{y\in T}(-1)^{\langle x,y\rangle}\right|\right] &\leq \Pr_{x\in S}\left[\left|\mathop{\mathbb{E}}_{y\in T}(-1)^{\langle x,y\rangle}\right| > \frac{\epsilon}{2}\right]\cdot\left|\mathop{\mathbb{E}}_{y\in T}(-1)^{\langle x,y\rangle}\right| + \Pr_{x\in S}\left[\left|\mathop{\mathbb{E}}_{y\in T}(-1)^{\langle x,y\rangle}\right| \leq \frac{\epsilon}{2}\right]\frac{\epsilon}{2}\\ &< \frac{\epsilon}{2}\cdot 1 + 1\cdot\frac{\epsilon}{2}\\ &= \epsilon\end{aligned}$$

contradicting equation 7. Therefore, we now have

$$\Pr_{x\in S}\left[\left|\mathop{\mathbb{E}}_{y\in T}(-1)^{\langle x,y\rangle}\right| > \frac{\epsilon}{2}\right] \geq \frac{\epsilon}{2}$$

From the above equation,

$$|S| \leq \frac{2}{\epsilon}\left|\left\{x : \left|\mathop{\mathbb{E}}_{y\in T}(-1)^{\langle x,y\rangle}\right| > \frac{\epsilon}{2}\right\}\right|$$

Using equation 5,

$$|S| \leq \frac{2}{\epsilon}\cdot\frac{4}{\epsilon^2}\cdot\frac{2^n}{|T|}$$

which implies $|S||T| \leq \frac{8}{\epsilon^3}2^n$, contradicting 6. Hence E is an $\left(\frac{n}{2} + 3\log\frac{1}{\epsilon}, \epsilon\right)$ extrator. □

2 Pseudorandom generators for polynomials over $\mathbb{F}_2$

Let P_d be the set of multilinear polynomials over $\mathbb{F}_2$ of degree atmost d in n variables.

Definition 3. PRG against P_d: *Pseudorandom generator against P_d is a function*

$$G : \{0,1\}^s \to \{0,1\}^n$$

such that

$$\forall f \in \mathsf{P}_d, \left|\Pr_{x\in\{0,1\}^s}[f(G(x)) = 1] - \Pr_{y\in\{0,1\}^n}[f(y) = 1]\right| < \epsilon$$

Following fact gives a motivation for studying PRGs against such polynomials.

Claim 4. *PRG against P_d for $d = (\log n)^{\omega(1)}$ is a PRG againct poly(n) sized $\mathsf{AC}^0(\oplus)$ circuits.*

Proof. We know from previous lectures (lecture 4, lemma 3 - Razborov's lemma) that for all circuits $C \in \mathsf{AC}^0(\oplus)$, there exists a distribution of polynomials $\mathcal{F}$ of degree $(\log n)^{O(1)}$ such that

$$\forall z, \Pr_{f \in \mathcal{F}}[f(z) = C(z)] \geq 1 - \epsilon$$

We want,

$$\Pr_{x \in \{0,1\}^s}[C(G(x)) = 1] \approx_\epsilon \Pr_{y \in \{0,1\}^n}[C(y) = 1]$$

But since distribution of polynomials $\mathcal{F}$ fools circuit C, we have

$$\Pr_{x \in \{0,1\}^s}[C(G(x)) = 1] \approx_\epsilon \Pr_{\substack{x \in \{0,1\}^s \\ f \in \mathcal{F}}}[f(G(x)) = 1]$$

Now, using the fact that G is a PRG against polynomials of degree $(\log n)^{\omega(1)}$,

$$\Pr_{\substack{x \in \{0,1\}^s \\ f \in \mathcal{F}}}[f(G(x)) = 1] \approx_\epsilon \Pr_{\substack{y \in \{0,1\}^n \\ f \in \mathcal{F}}}[f(y) = 1]$$

Finally, since $\mathcal{F}$ was a distribution that fools C,

$$\Pr_{\substack{y \in \{0,1\}^n \\ f \in \mathcal{F}}}[f(y) = 1] \approx_\epsilon \Pr_{y \in \{0,1\}^n}[C(y) = 1]$$

From above three equations we have,

$$\Pr_{x \in \{0,1\}^s}[C(G(x)) = 1] \approx_\epsilon \Pr_{y \in \{0,1\}^n}[C(y) = 1]$$

as needed.

□

In previous lectures, we constructed ϵ-biased spaces with seed length $\log n + O(\log \frac{1}{\epsilon})$. Following theorem gives a pseudoramdom generator for P_d which requires seed length $d \log n + O(d \log \frac{1}{\epsilon})$

Theorem 5. *(Viola) If we pick $x_1, x_2, \ldots, x_d$ from an ϵ-biased distribution then $x_1 + x_2 + \cdots + x_d$ fools degree d polynomials.*

Proof. The difference between linear function and degree d polynomial is that for all linear functions f, if we evaluate f on some x which is choosen uniformly at random from the set $\{0,1\}^n$ then the quantity $\Pr[f(x) = c]$ where $c \in \{0,1\}$ is either $0, 1$ or $1/2$. But for a general class of polynomials of higher degree, it can have many different values for the above quantity. So, in order to prove this claim we somehow need to take care of all these polynomials.

We will prove the claim by induction on degree d of polynomials. Consider the case when $d = 1$, in this case P_1 is a set of non-constant linear functions. For a non-constant linear function l,

$$\left| \mathop{\mathbb{E}}_{y \in U_n} \left[(-1)^{l(y)} \right] \right| = 0$$

Also, if μ is an ϵ-biased distribution,

$$\left| \mathop{\mathbb{E}}_{y \in \mu} \left[(-1)^{l(y)} \right] \right| \leq \epsilon$$

Hence,

$$\left| \mathop{\mathbb{E}}_{y \in U_n} \left[(-1)^{l(y)} \right] - \mathop{\mathbb{E}}_{y \in \mu} \left[(-1)^{l(y)} \right] \right| \leq \epsilon$$

as required.

Assume that the claim is true for $d-1$. We will now prove the claim for degree d polynomials. We will classify the degree d polynomials into two classes and prove the claim for both of them.

- **Case 1:** $Pr_{y \in \{0,1\}^n}[f(y) = 1]$ is *close* to $\frac{1}{2}$
- **Case 2:** $Pr_{y \in \{0,1\}^n}[f(y) = 1]$ is *far away* from $\frac{1}{2}$

First, consider case 1. The condition is equivalent to saying

$$\left| \mathop{\mathbb{E}}_{y \in \{0,1\}^n} \left[(-1)^{f(y)} \right] \right| \text{ is small}$$

We want to show that if this is the case then

$$A = \left| \mathop{\mathbb{E}}_{x_1, x_2, \ldots, x_d} \left[(-1)^{f(x_1 + x_2 + \cdots + x_d)} \right] \right| \text{ is also small}$$

Consider,

$$\begin{aligned}
A^2 &= \left(\mathop{\mathbb{E}}_{x_1, x_2, \ldots, x_d} \left[(-1)^{f(x_1 + x_2 + \cdots + x_d)} \right] \right)^2 \\
&= \left(\mathop{\mathbb{E}}_{x_1, x_2, \ldots, x_{d-1}} \left[\mathop{\mathbb{E}}_{x_d} \left[(-1)^{f(x_1 + x_2 + \cdots + x_d)} \right] \right] \right)^2 \\
&\leq \mathop{\mathbb{E}}_{x_1, x_2, \ldots, x_{d-1}} \left[\mathop{\mathbb{E}}_{x_d} \left[(-1)^{f(x_1 + x_2 + \cdots + x_d)} \right]^2 \right] \\
&= \mathop{\mathbb{E}}_{x_1, x_2, \ldots, x_{d-1}} \left[\mathop{\mathbb{E}}_{x_d, x_{d'}} \left[(-1)^{f(x_1 + x_2 + \cdots + x_d)} \cdot (-1)^{f(x_1 + x_2 + \cdots + x_{d'})} \right] \right] \\
&= \mathop{\mathbb{E}}_{x_1, x_2, \ldots, x_{d-1}} \left[\mathop{\mathbb{E}}_{x_d, x_{d'}} \left[(-1)^{f(x_1 + x_2 + \cdots + x_d) + f(x_1 + x_2 + \cdots + x_{d'})} \right] \right]
\end{aligned}$$

We need following fact

Fact 6. *If f is a degree d polynomial in n variables and*

$$g_{b,b'}(a_1, a_2, \ldots, a_n) = f(a+b) \pm f(a+b')$$

for fixed vectors $b, b' \in \{0,1\}^n$ then degree of $g_{b,b'}$ is $d-1$

Using above fact, we can write $f(x_1+x_2+\cdots+x_d)+f(x_1+x_2+\cdots+x_{d'})$ as $h_{x_d,x_{d'}}(x_1+x_2+\cdots+x_{d-1})$ where $h_{x_d,x_{d'}}$ is a degree $d-1$ polynomial. Therefore,

$$A^2 \leq \mathop{\mathbb{E}}_{x_d,x_{d'}} \left[\mathop{\mathbb{E}}_{x_1,x_2,\ldots,x_{d-1}} \left[(-1)^{h_{x_d,x_{d'}}(x_1+x_2+\cdots+x_{d-1})} \right] \right]$$

By induction hypothesis, $x_1 + x_2 + \cdots + x_{d-1}$ fools $h_{x_d,x_{d'}}$ as $h_{x_d,x_{d'}}$ is a degree $d-1$ polynomial. So we can write,

$$A^2 \leq_\epsilon \mathop{\mathbb{E}}_{x_d,x_{d'}} \left[\mathop{\mathbb{E}}_{U\in U_n} \left[(-1)^{h_{x_d,x_{d'}}(U)} \right] \right]$$

$$= \mathop{\mathbb{E}}_{x_d,x_{d'}} \left[\mathop{\mathbb{E}}_{U\in U_n} \left[(-1)^{f(U+x_d)+f(U+x_{d'})} \right] \right]$$

$$= \mathop{\mathbb{E}}_{U\in U_n} \left[\mathop{\mathbb{E}}_{x_d,x_{d'}} \left[(-1)^{f(U+x_d)+f(U+x_{d'})} \right] \right]$$

We will now use Fourier analysis to bound the above quantity. Consider the function g defined as follows,

$$g : \{0,1\}^n \to \{-1,+1\}$$
$$g(x) = (-1)^{f(x)}$$

We can write g in fourier expansion as

$$g(x) = \frac{1}{2^n} \sum_{\alpha\in\{0,1\}^n} \hat{g}(\alpha)\chi_\alpha(x)$$

Now,

$$A^2 \leq_\epsilon \mathop{\mathbb{E}}_{U\in U_n} \left[\mathop{\mathbb{E}}_{x_d,x_{d'}} \left[(-1)^{f(U+x_d)+f(U+x_{d'})} \right] \right]$$

$$= \mathop{\mathbb{E}}_{U\in U_n} \left[\mathop{\mathbb{E}}_{x_d,x_{d'}} \left[g(U+x_d) \cdot g(U+x_{d'}) \right] \right]$$

$$= \mathop{\mathbb{E}}_{U\in U_n} \left[\mathop{\mathbb{E}}_{x_d,x_{d'}} \left[\frac{1}{2^n} \sum_\alpha \hat{g}(\alpha)\chi_\alpha(U+x_d) \cdot \frac{1}{2^n} \sum_\beta \hat{g}(\beta)\chi_\beta(U+x_{d'}) \right] \right]$$

$$= \frac{1}{2^{2n}} \mathop{\mathbb{E}}_{U \in U_n, x_d, x_{d'}} \left[\sum_{\alpha,\beta} \hat{g}(\alpha)\hat{g}(\beta)\chi_{\alpha+\beta}(U)\chi_\alpha(x_d)\chi_\beta(x_{d'}) \right]$$

$$= \frac{1}{2^{2n}} \mathop{\mathbb{E}}_{x_d, x_{d'}} \left[\sum_{\alpha,\beta} \hat{g}(\alpha)\hat{g}(\beta)\chi_\alpha(x_d)\chi_\beta(x_{d'}) \mathop{\mathbb{E}}_{U \in U_n} [\chi_{\alpha+\beta}(U)] \right]$$

Using the fact that

$$\mathop{\mathbb{E}}_{U \in U_n} [\chi_{\alpha+\beta}(U)] = \begin{cases} 1 & \text{if } \alpha \neq \beta \\ 0 & \text{if } \alpha = \beta \end{cases}$$

$$A^2 \leq_\epsilon \frac{1}{2^{2n}} \mathop{\mathbb{E}}_{x_d, x_{d'}} \left[\sum_{\alpha} \hat{g}(\alpha)^2 \chi_\alpha(x_d)\chi_\alpha(x_{d'}) \right]$$

$$= \frac{1}{2^{2n}} \mathop{\mathbb{E}}_{x_d, x_{d'}} \left[\sum_{\alpha} \hat{g}(\alpha)^2 \chi_\alpha(x_d)\chi_\alpha(x_{d'}) \right]$$

$$= \frac{1}{2^{2n}} \mathop{\mathbb{E}}_{x_d, x_{d'}} \left[\hat{g}(\emptyset)^2 \cdot \chi_\emptyset(x_d)\chi_\emptyset(x_{d'}) + \sum_{\alpha \in \{0,1\}^n \setminus \{0\}^n} \hat{g}(\alpha)^2 \chi_\alpha(x_d)\chi_\alpha(x_{d'}) \right]$$

$$= \frac{1}{2^{2n}} \left(\hat{g}(\emptyset)^2 \cdot \mathop{\mathbb{E}}_{x_d, x_{d'}} [\chi_\emptyset(x_d)\chi_\emptyset(x_{d'})] + \sum_{\alpha \in \{0,1\}^n \setminus \{0\}^n} \hat{g}(\alpha)^2 \mathop{\mathbb{E}}_{x_d, x_{d'}} [\chi_\alpha(x_d)\chi_\alpha(x_{d'})] \right)$$

$$= \frac{1}{2^{2n}} \left(\hat{g}(\emptyset)^2 \cdot \mathop{\mathbb{E}}_{x_d, x_{d'}} [1] + \sum_{\alpha \in \{0,1\}^n \setminus \{0\}^n} \hat{g}(\alpha)^2 \mathop{\mathbb{E}}_{x_d} [\chi_\alpha(x_d)] \mathop{\mathbb{E}}_{x_{d'}} [\chi_\alpha(x_{d'})] \right)$$

Since χ_α is -1 raised to the power a non-constant linear function for $\alpha \neq \emptyset$ and x_d, x'_d coming from ϵ-biased spaces, we have

$$\mathop{\mathbb{E}}_{x_d} [\chi_\alpha(x_d)] \mathop{\mathbb{E}}_{x_{d'}} [\chi_\alpha(x_{d'})] \leq \epsilon \cdot \epsilon$$

Therefore,

$$A^2 \leq_\epsilon \frac{1}{2^{2n}} \left(\hat{g}(\emptyset)^2 + \sum_{\alpha \in \{0,1\}^n \setminus \{0\}^n} \hat{g}(\alpha)^2 \epsilon^2 \right)$$

$$= \frac{\hat{g}(\emptyset)^2}{2^{2n}} + \epsilon^2 \sum_{\alpha \in \{0,1\}^n \setminus \{0\}^n} \frac{\hat{g}(\alpha)^2}{2^{2n}}$$

Now,

$$\begin{aligned} \frac{\hat{g}(\emptyset)}{2^n} &= \frac{1}{2^n} \sum_y g(y) \cdot \chi_\emptyset(y) \\ &= \frac{1}{2^n} \sum_y g(y) \cdot 1 \\ &= \frac{1}{2^n} \sum_y (-1)^{f(y)} \\ &= \mathop{\mathbb{E}}_{y \in U} [(-1)^{f(y)}] \end{aligned}$$

Also, since

$$g(x) = \frac{1}{2^n} \sum_{\alpha \in \{0,1\}^n} \hat{g}(\alpha) \chi_\alpha(x)$$

By Parseval's Identity,

$$\sum_{\alpha \in \{0,1\}^n} \frac{\hat{g}(\alpha)^2}{2^{2n}} = 1$$

Therefore,

$$A^2 \leq_\epsilon \left| \mathop{\mathbb{E}}_{y \in U} [(-1)^{f(y)}] \right|^2 + \epsilon^2$$

Since we are in the case where $\left| \mathbb{E}_{y \in U} [(-1)^{f(y)}] \right|$ is small, A is also small as required.

Now, consider case 2 where the quantity

$$\left| \mathop{\mathbb{E}}_{y \in U} [(-1)^{f(y)}] \right|$$

is big. Consider the quantity,

$$B = \left| \mathop{\mathbb{E}}_{y \in U_n} \left[(-1)^{f(y)} \right] \right| \left| \left(\mathop{\mathbb{E}}_{x_1, x_2, \ldots, x_d} \left[(-1)^{f(x_1 + x_2 + \cdots + x_d)} \right] - \mathop{\mathbb{E}}_{U} \left[(-1)^{f(U)} \right] \right) \right| \tag{9}$$

$$B = \left| \mathop{\mathbb{E}}_{\substack{x_1,x_2,\ldots,x_d \\ y\in U_n}} \left[(-1)^{f(x_1+x_2+\cdots+x_d)+f(y)}\right] - \mathop{\mathbb{E}}_{U,y\in U_n} \left[(-1)^{f(U)+f(y)}\right] \right|$$

$$= \left| \mathop{\mathbb{E}}_{\substack{x_1,x_2,\ldots,x_d \\ y\in U_n}} \left[(-1)^{f(x_1+x_2+\cdots+x_d)+f(y+x_1+\cdots+x_d)}\right] - \mathop{\mathbb{E}}_{U,y\in U_n} \left[(-1)^{f(U)+f(y+U)}\right] \right|$$

Define, $f_y^{'}(z) = f(z) + f(z+y)$, by fact 6, $f_y^{'}$ is a degree $d-1$ polynomial in z. So,

$$B \leq \mathop{\mathbb{E}}_{y} \left[\left| \mathop{\mathbb{E}}_{x_1,x_2,\ldots,x_d} \left[(-1)^{f_y^{'}(x_1+x_2+\cdots+x_d)}\right] - \mathop{\mathbb{E}}_{U} \left[(-1)^{f_y^{'}(U)}\right] \right| \right]$$

Since $f_y^{'}$ is a degree $d-1$ polynomial, using induction hypothesis,

$$\left| \mathop{\mathbb{E}}_{x_1,x_2,\ldots,x_d} \left[(-1)^{f_y^{'}(x_1+x_2+\cdots+x_d)}\right] - \mathop{\mathbb{E}}_{U} \left[(-1)^{f_y^{'}(U)}\right] \right| \leq \epsilon$$

Therefore,

$$B \leq \mathop{\mathbb{E}}_{y}[\epsilon] = \epsilon$$

Hence, using equation 9,

$$\left| \left(\mathop{\mathbb{E}}_{x_1,x_2,\ldots,x_d} \left[(-1)^{f(x_1+x_2+\cdots+x_d)}\right] - \mathop{\mathbb{E}}_{U} \left[(-1)^{f(U)}\right] \right) \right| \tag{10}$$

is small as required. □

3 Superconcentrators and Circuit Lower Bounds

In this section, we will look at an approach to proving circuit lower bounds developed by Valiant. In process, we will define explore some graph theoretic properties of the circuits and show that the approach is not strong enough to imply non trivial lower bounds.

Model: We will be looking at circuits which have gates with unbounded fan-in which compute a certain linear combination of its input. The constant involved in the linear combination would be given by the label of the gate. In the course of this discussion, the circuits will have multiple outputs.

Function: We will look at circuits which compute the vector Mx for some fixed $n \times n$ matrix M and a n dimensional vector x as input. We will assume the we are working over some field $\mathcal{F}$.

It is easy to observe that for any matrix M, a corresponding circuit for Mx can be constructed using n^2 wires. It can also be shown via a counting argument that a random matrix requires $n^{2-o(1)}$ wires. Our aim here is to show that there is an explicit matrix M for which any circuit C of the above type requires $\omega(n)$ wires. To this end, let us observe some properties of the circuit as a graph. In the rest of the lecture, for some matrix M, we will denote a circuit of the above described type, computing Mx by C_M.

3.1 Properties of the circuit

The first observation below tells us that any circuit computing Mx for an invertible matrix M cannot have a very small cut that separates all output gates from all the input gates.

Observation 7. *Any set of edges which disconnects the input gates from the output gates in circuit C_M has size at least n for an invertible $N \times n$ matrix M.*

Proof. Let E be a cut in the circuit C. The circuit can be decomposed into two parts, C_1 computing the function on wires in E from the input gates I and C_2 computing the output functions O from inputs E. Both these computations, by C_1 and C_2 correspond to linear transformations by some matrices M_1 and M_2 respectively, such that $Mx = M_2(M_1x)$. Now, if $|E| < n$, then dimension of M_1x is strictly less than n, which in turn implies that the dimension of $M_2(M_1x)$ is less than n, which contradicts the fact that M is invertible. Hence, E has at least n edges. □

A circuit which has this property that any set of edges disconnecting the input from the output has at least n edges is called a concentrator. The above observation just implies that if M is invertible, then C_M is a concentrator. It might seem at this point that this observation is sufficient for showing a superlinear lower bound on the number of edges in C_M. The claim is not true as for the identity matrix, there is a trivial linear sized circuit. So, we need to observe stronger connectivity properties in the circuit C to be able to obtain superlinear lower bound. Let us define the notion of *Superconcentrator* for this purpose.

Definition 8. *A directed acyclic graph with nodes I of indegree 0 and nodes O of outdegree 0 is said to be a superconcentrator if $\forall S \subseteq I, T \subseteq O$ such that $|S| = |T|$, there exist $|S|$ edge disjoint paths from S to T.*

Let us now argue that for a matrix M such that every square submatrix for it has full rank, any circuit C_M is a superconcentrator.

Claim 9. *If for an $n \times n$ matrix M, every square submatrix is of full rank, then the circuit C_M is a superconcentrator.*

Proof. Let there be subsets S of inputs, T of outputs such that $|S| = |T| = r$. Let us fix all the inputs outside S to some constants. Let us now look at the coordinates in T as functions of the inputs in S. Observe that $(Mx)_T = u + M^{'}x_S$ for a constant vector u of dimension r and $M^{'}$ is a submatrix of M of dimension r. Since $M^{'}$ has full rank, so the subcircuit between S and T of C is computing an invertible map. Therefore, by an argument similar to that in Observation 7 tells us that any cut between S and T must be of size at least $|S|$. From here, the claim follows via Menger's Theorem. □

The requirements for M are more strict than that required in the concentrator approach discussed above. But, there are matrices satisfying the hypothesis of this claim. One such matrix is the Cauchy matrix. Valiant conjectured that any such superconcentrator with n inputs and n outputs must have $\omega(n)$ edges. We will now discuss a construction which shows that the conjecture is false.

Theorem 10 (Valiant)**.** *There exist circuits with n inputs, n outputs, depth $\log n$ and $O(n)$ edges which are superconcentrators.*

Proof. In this construction, we will be using bipartite expanders $G = (L, R, E)$ of constant degree such that $|L| = l, |R| = \frac{3l}{4}$ and for each $S \subseteq L$ of size at most $\frac{l}{2}$, $|N(S)| \geq |S|$. Such expanders can be shown to exist via probabilistic method and we skip this argument here. Let us now construct the graph G iteratively starting with a trivial graph $G_1 = (L_1, R_1, E_1)$ which is just complete graph $K_{1,1}$. In general, G_i has $2i$ layers of vertices $L_i, L_{i-1}, \ldots, R_{i-1}, R_i$, satisfying $\frac{|L_j|}{|L_{j-1}|} = \frac{|R_j|}{|R_{j-1}|} = 4/3$ for $j \geq 2$ and the graph between L_j and L_{j-1} is a bipartite expander of the type described above. Similarly, for R_j and R_{j-1}. In addition to these edges which come from the expanders, we add a perfect matching between L_j and R_j for each j. Clearly, at the end of i steps, we have $O((4/3)^i)$ nodes and the number of edges is linear in the number of vertices as we have used constant degree expanders at each step, and adding a perfect matching just increases their degree by 1. In particular, the number of edges in an n vertex graph constructed by this method, $e(n)$ will be given by the recurrence $e(n) \leq c.n + n + e(\frac{3n}{4})$ for a constant c. This solves to a linear number of edges.

We now need to argue that this is indeed a superconcentrator. Let us consider any set $S \subseteq L_i$, $T \subseteq R_i$ with $|S| = |T|$. We need to show that there are at least $|S|$ edge disjoint paths from S to T. We will argue this via induction on j. The base case for $j = 1$ is easy to verify. Let us assume that the claim holds G_{i-1} and argue that it holds for G_i. Now, let us consider two cases based on the size of $|S|$.

1. $|S| = |T| \leq |L_i|/2$.
   In this case, observe that the bipartite graph between S and neighbourhood of S in L_{i-1} satisfies the hypothesis of the Hall's Theorem. So, there is a perfect matching between S and $N(S) \cap L_{i-1}$. Similarly, for R_i. Therefore, it is enough to show that the number of edge disjoint paths between $N(S) \cap L_{i-1}$ and $N(T) \cap R_{i-1}$ is at least $|S| = |T| = |N(S) \cap L_{i-1}| = |N(T) \cap R_{i-1}|$; which in turn is true from the induction hypothesis.

2. $|S| = |T| > |L_i|/2$.
   Let $S_1 = N(S) \cap R_i$ and $R_1 = N(S) \cap L_i$. Since the edges between L_i and R_i formed a perfect matching, we know that $|S_1| = |S|$. Now, $|S_1| = |T| > \frac{n}{2}$ tells us that $S_1 \cap T \neq \phi$. So, the vertices corresponding to S in $S_1 \cap T$ are matched via edges of the perfect matching. Let us label the set of these vertices by M. So, all that remains to show is that $S \setminus M$ and $T \setminus S_1 \cap T$ have at least $|S \setminus M|$ edge disjoint paths between them. But, now observe that $|S \setminus M| = |T \setminus S_1 \cap T| \leq |L_i|/2$ and we are in the case in item 1 of this proof.

□

Thus, this approach of superconcentrators in the present form is not strong enough to show superlinear lower bounds for linear transformations. The problem continues to remain open today.