Presentations

See pdf file "Course Presentation Topics"
to view suggested list of topics with links to
papers/talks.

* Sign up for your presentations: Feb 20 - 23
  Pick 3 topics:
  (i) Main presenter
  (ii) Helper
  (iii) Reviewer

sign up for these
among the topics
with main presenters

Presentations cont'd

- Talk to me if you want to discuss possible topics before deadline

Your Presentation should include

- Presentation in class ~50 mins
- Lecture notes / slides
- Questions for discussion, open problems

- I will meet with main presenter & helper ahead of time to review your presentation

Last Class

1. Deterministic $CC(f) \le$ Nondet $CC(f) \cdot$ co-Nondet $CC(f)$

   and similar result for decision tree complexity

2. Randomized CC vs Distributional CC

   and Yao Minimax Thm

<u>Today</u> : $BPP^{cc}(IP_n) = \Omega(n)$

- Discrepancy as a measure that lower bounds randomized CC

- show IP has low discrepancy via eigenvalue analysis of IP comm. complexity matrix

Fourier Analysis of Boolean functions

Define $\chi_y : \{0,1\}^n \to \{-1,1\}$: $\chi_y(x) = (-1)^{x \cdot y} = \begin{cases} -1 & \text{if parity of } x_y \text{ odd} \\ 1 & \text{" " " " even} \end{cases}$] Fourier basis

Fact $\forall y \neq y'$: $\langle \chi_y, \chi_{y'} \rangle \stackrel{d}{=} \mathbb{E}_x [\chi_y(x) \cdot \chi_{y'}(x)] = 0$

$y = y'$: $\langle \chi_y, \chi_y \rangle = 1$

Representation of $f : \{0,1\}^n \to \mathbb{R}$ over Basis $\{\chi_y\}_{y \in \{0,1\}^n}$:

$$f(x) = \sum_y \underbrace{\widehat{f(y)}}_{\text{Fourier coeff of } \chi_y} \cdot \chi_y(x) \quad , \quad \widehat{f(y)} = \underbrace{\langle f, \chi_y \rangle}_{\text{"correlation of } f \text{ with } \chi_y\text{"}} := \mathbb{E}_x [f(x) \cdot \chi_y(x)]$$

Parseval Identity:

$$\langle f, f \rangle \stackrel{d}{=} \mathbb{E}_x [f(x)^2] = \sum_y \widehat{f(y)}^2$$

Parseval Identity:

$$\langle f, f \rangle \triangleq \mathop{\mathbb{E}}_{x \in \{-1,1\}}[f(x)^2] = \sum_{y} \widehat{f(y)}^2$$

Pf

$$\mathop{\mathbb{E}}_{x}[f(x)^2] = \mathop{\mathbb{E}}_{x}\left[\left(\sum_{y} \widehat{f(y)} \chi_y\right) \cdot \left(\sum_{y'} \widehat{f(y')} \chi_{y'}\right)\right]$$ Write f over Fourier basis

$$= \mathop{\mathbb{E}}_{x}\left[\sum_{y,y'} \widehat{f(y)}\, \widehat{f(y')}\, \chi_y \cdot \chi_{y'}\right]$$

$$= \sum_{y,y'} \mathop{\mathbb{E}}_{x}\left[\widehat{f(y)}\, \widehat{f(y')}\, \chi_y \cdot \chi_{y'}\right]$$ Linearity of expectation

$$= \sum_{y,y'} \widehat{f(y)}\, \widehat{f(y')} \mathop{\mathbb{E}}_{x} \langle \chi_y, \chi_{y'} \rangle$$ orthogonality of $\chi_y, \chi_{y'}$ $y \neq y'$ and $\langle \chi_y \chi_y \rangle = 1$

$$= \sum_{y} \widehat{f(y)}^2$$

Randomized Lower Bound for IP (inner product)

$$IP_n(x,y) = \sum_{i=1}^{n} x_i y_i \pmod 2$$

M_{IP} is the Hadamard matrix $(+1 \approx 0,\ -1 \approx 1)$:

$H_0 = \boxed{1}$

$$H_n = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}$$

H_3:

1	1	1	1	1	1	1	1
1	-1	1	-1	1	-1	1	-1
1	1	-1	-1	1	1	-1	-1
1	-1	-1	1	1	-1	-1	1
1	1	1	1	-1	-1	-1	-1
1	-1	1	-1	-1	1	-1	1
1	1	-1	-1	-1	-1	1	1
1	-1	-1	1	-1	1	1	-1

Thm Let μ = unif dist. $D_\mu^{1/3}(IP_n) = \Omega(n)$

Randomized Lower Bound for IP (inner product)

$$H_0 = \boxed{1}$$

$$H_n = \begin{array}{|c|c|} \hline H_{n-1} & H_{n-1} \\ \hline H_{n-1} & -H_{n-1} \\ \hline \end{array}$$

H_0 : $\begin{array}{|c|} \hline 1 \\ \hline \end{array}$

H_1 : $\begin{array}{|c|c|} \hline 1 & 1 \\ \hline 1 & -1 \\ \hline \end{array}$

H_2 : $\begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 1 \\ \hline 1 & -1 & 1 & -1 \\ \hline 1 & 1 & -1 & -1 \\ \hline 1 & -1 & -1 & 1 \\ \hline \end{array}$

H_3 :

$$\begin{array}{|c|c|c|c|c|c|c|c|} \hline 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ \hline 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ \hline 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ \hline 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \hline 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ \hline 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ \hline 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ \hline \end{array}$$

Rows correspond to
all parity functions
which are pairwise orthogonal

$\therefore$ RANK $(H_n) = 2^n$

Randomized CC Lower Bounds

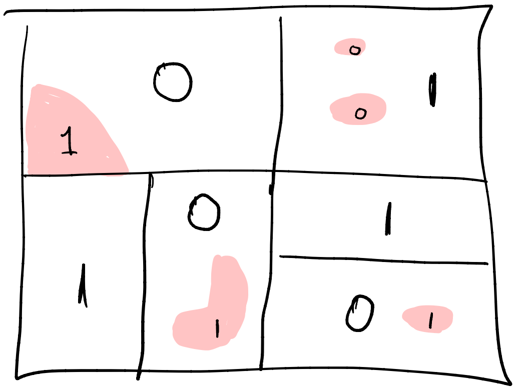
To prove a randomized CC LB of c for $f: \underbrace{\{0,1\}^n}_{X} \times \underbrace{\{0,1\}^n}_{Y} \to \{0,1\}$

it is nec. & sufficient to give a distrib μ over $X \times Y$

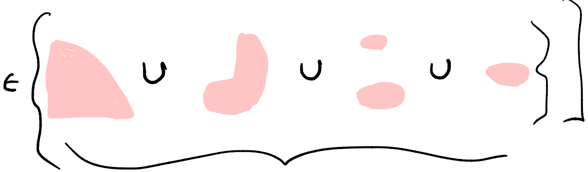
s.t. $\underbrace{D_\varepsilon^\mu(f)}_{\substack{\varepsilon\text{-error} \\ \text{det. cc of } f \\ \text{wrt } \mu}} \geq c$

For $\varepsilon = 0$ we saw that a zero error protocol of cost c partitioned M_f into $\leq 2^c$ mono. rectangles.

For $\varepsilon > 0$, the <u>average</u> error over the rectangles is ε.



Protocol Π for f with overall error $\frac{1}{10}$ wrt μ = uniform dist

$\Pr_{(x,y)\sim\mu}\Big[(x,y)\in$  $\Big]$

error set

$$\Pr_{(x,y)}\left[(x,y)\in \text{Error Set}\right] = \sum_{R_i} \Pr\left[(x,y)\in R_i\right]\cdot \Pr\left[(x,y) \text{ in error set} \mid (x,y)\in R_i\right]$$

Randomized CC Lower Bounds via Discrepancy

By minimax, to prove lower bounds for BPP_ε^{cc} protocols for f it suffices to find a distribution μ s.t. $D_\varepsilon^{\mu}(f)$ is large.

We will prove LBs wrt μ via the discrepancy measure.

Discrepancy of M_f wrt a distribution μ over $X \times Y$

[special case to think about: μ = unif distrib]

Let R be a subrectangle of M_f

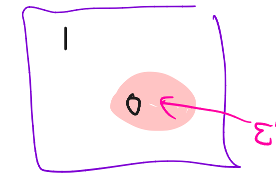
Defn $\mathrm{Disc}_\mu(R) = |\mu(R \cap f^{-1}(1)) - \mu(R \cap f^{-1}(0))|$

measures $|\#1\text{'s} - \#0\text{'s in } R| / 2^{2n}$

Defn $\mathrm{Disc}_\mu(f) = \max\limits_{R \subseteq M_f} \mathrm{Disc}_\mu(R)$

small discrepancy means all large rectangles are balanced

Defn $\mathrm{Disc}_\mu(f,R) = \left| \mu(R \cap f^{-1}(1)) - \mu(R \cap f^{-1}(0)) \right|$



$$\mu(R \cap f^{-1}(1)) = \Pr_{(x,y)\sim\mu}\left[(x,y) \in R \wedge (x,y) \in f^{-1}(1)\right]$$

$$= \Pr\left[(x,y) \in R\right] \cdot \Pr\left[(x,y) \in f^{-1}(1) \mid (x,y) \in R\right]$$

$$= \mu(R) \cdot \frac{\#\text{ 1's in } R}{|R|}$$

$$\therefore \mathrm{Disc}_\mu(f,R) = \left| \mu(R) \cdot \frac{\#\text{ 1's in } R}{|R|} - \mu(R) \frac{\#\text{ 0's in } R}{|R|} \right|$$

$$= \mu(R) \cdot \frac{\left|\#\text{ 1's in } R - \#\text{ 0's in } R\right|}{|R|} = \frac{1}{2^{2n}} \left|\#\text{ 1's in } R - \#\text{ 0's in } R\right|$$

Defn $\mathrm{Disc}_\mu(f,R) = |\mu(R \cap f^{-1}(1)) - \mu(R \cap f^{-1}(0))|$

Defn $\mathrm{Disc}_\mu(f) = \max_{R \subseteq M_f} \mathrm{Disc}_\mu(R)$

Claim for every μ, $D^\varepsilon_\mu(f) \geq \log\left(\frac{1-2\varepsilon}{\mathrm{Disc}_\mu(f)}\right)$

says low discrepancy implies high cc wrt μ

Intuition behind claim for μ = uniform distribution

(same intuition for any μ)

- Low discrepancy says that all large $R \subseteq M_f$ are nearly balanced
- A low cost deterministic protocol Π for f partitions M_f into few subrectangles, so most of them are large
- Low discrepancy $\Rightarrow$ Large R's are nearly balanced, Π must $\Rightarrow$ any low cc Π will make a lot of errors.

Theorem 1 Let $f : X \times Y \to \{0,1\}$, μ distrib over $X \times Y$.

Then $\underbrace{D_\mu^\varepsilon(f)}_{\varepsilon} \geq \log\left(\frac{1-2\varepsilon}{\mathrm{Disc}_\mu(f)}\right)$

For $\varepsilon = 1/3$: $D_\mu^{1/3}(f) \geq \log\left(\frac{1}{3\,\mathrm{Disc}_\mu(f)}\right)$

Theorem 1 Let $f: X \times Y \to \{0,1\}$, μ distrib over $X \times Y$.

Then $\underbrace{D_\mu^\varepsilon(f)}_{c} \geq \log\left(\frac{1-2\varepsilon}{\mathrm{Disc}_\mu(f)}\right)$

Pf Let Π be cost c det protocol with error ε wrt μ.

we will show: $\mathrm{Disc}_\mu(f) \geq (1-2\varepsilon)2^{-c}$ which implies thm

Let leaves of Π correspond to the partition of M_f into subrectangles

$R_1, \ldots, R_t$, $t \leq 2^c$

Let a_i be the value (most popular value in R_i) that Π outputs on R_i.

Since Π has error $\leq \varepsilon$:

$$1-\varepsilon \leq \Pr_{(x,y)\sim\mu}[\Pi(x,y) = f(x,y)]$$

$$= \sum_{i=1}^{t} \Pr_{(x,y)\sim\mu}[(x,y) \in R_i \wedge f(x,y) = a_i]$$

$$= \sum_{i=1}^{t} \mu(R_i \cap \underbrace{f^{-1}(a_i)}) \leftarrow \text{correct value}$$

Similarly

$$\varepsilon \geq \Pr_{(x,y)\sim\mu} [\, \Pi(x,y) \neq f(x,y) \,]$$

$$= \sum_{i=1}^{t} \mu(R_i \cap f^{-1}(1-a_i))$$

incorrect value

combining the 2 inequalities (1st − second)

$$1 - 2\varepsilon \leq \sum_{i=1}^{t} \mu(R_i \cap f^{-1}(a_i)) - \mu(R_i - f^{-1}(1-a_i))$$

$$\leq \sum_{i=1}^{t} |\mu(R_i \cap f^{-1}(0)) - \mu(R_i \cap f^{-1}(1))|$$

$$\leq \sum_{i=1}^{t} \max_{R} \operatorname{Disc}_\mu(f, R)$$

$$\leq 2^c \cdot \operatorname{Disc}_\mu(f)$$

Let μ = uniform distrib, $M = IP_n$ matrix (= Hadamard / Walsh matrix; $-1,1$ valued)

$$\mathrm{Disc}_\mu(IP_n, R = S \times T) = \frac{1}{2^{2n}} \left| \sum_{\substack{x \in S \\ y \in T}} M(x,y) \right|$$

$$\mathrm{Disc}_\mu(IP_n) = \max_{S,T} \mathrm{Disc}_\mu(IP_n, S \times T)$$

$\boxed{\underline{\text{Theorem 2}} \quad \mathrm{Disc}_\mu(IP_n) \le 2^{-n/2}}$

$\boxed{\underline{\text{Theorem 1}} \text{ Let } f: X \times Y \to \{0,1\},\ \mu \text{ distrib over } X \times Y. \text{ Then } \underbrace{D^\varepsilon_\mu(f)}_{c} \ge \log\left(\frac{1-2\varepsilon}{\mathrm{Disc}_\mu(f)}\right)}$

Combining Theorems 1, 2 we get:

$\boxed{\underline{\text{Theorem}} \quad D^{1/3}_\mu(IP_n) = \Omega(n)}$ (LB on distributional complexity of IP_n over unif dist μ)

and by Yao, $BPP^{1/3}(IP_n) = \Omega(n)$

Let μ = uniform distrib, $M = IP_n$ matrix (= Hadamard / Walsh matrix)

$$\mathrm{Disc}_\mu(f, R = S\times T) = \frac{1}{2^{2n}} \left| \sum_{\substack{x\in S\\ y\in T}} M(x,y) \right|$$

$$\mathrm{Disc}_\mu(f) = \max_{S,T} \mathrm{Disc}_\mu(f, S\times T)$$

$$\boxed{\underline{\text{Theorem 2}} \quad \mathrm{Disc}_\mu(IP_n) \le 2^{-n/2}}$$

$$\boxed{\underline{\text{Claim}} \quad \forall S,T \quad \left| \sum_{\substack{x\in S\\ y\in T}} M(x,y) \right| \le 2^{n/2} \sqrt{|S|\,|T|}}$$

<u>Pf of Thm 2 (Assuming claim):</u>

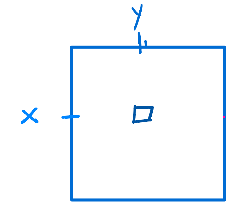
$$\mathrm{Disc}_\mu(f) = \max_{S,T} \frac{1}{2^{2n}} \left| \sum_{\substack{x\in S\\ y\in T}} M(x,y) \right|$$

$$\le \max_{S,T} \frac{1}{2^{2n}} \cdot 2^{n/2} \sqrt{|S|\cdot|T|}$$

$$= \max_{S,T} \frac{\sqrt{|S|\cdot|T|}}{2^{3n/2}} \le \frac{\sqrt{2^n\cdot 2^n}}{2^{3n/2}} = \frac{2^n}{2^{3n/2}} = 2^{-n/2}$$

Claim $\forall S,T \quad \left|\sum_{\substack{x\in S\\ y\in T}} M(x,y)\right| \le 2^{\frac{n}{2}}\sqrt{|S|\,|T|}$

IP Matrix M :



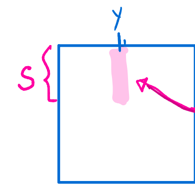
$\forall x,y \in \{-1,1\}^n$

$M(x,y) = (-1)^{x\cdot y} = \chi_y(x)$

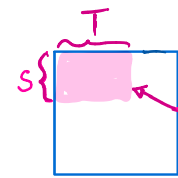
Pf Fix $R = S\times T$.

Let $A(y) = \sum_{x\in S} M(x,y)$

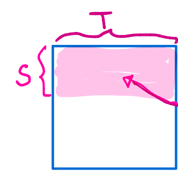
Then $\left|\sum_{\substack{x\in S\\ y\in T}} M(x,y)\right| = \left|\sum_{y\in T} A(y)\right|$



$A(y)$ is sum of these values



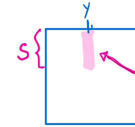
$\left|\sum_{y\in T} A(y)\right|$ is sum of these values.



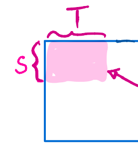
when $T = \mathcal{Y}$ (all columns)

$\left|\sum_{y\in\mathcal{Y}} A(x)\right|$ is sum of these values

Claim $\forall S,T \quad \left|\sum_{\substack{x\in S\\ y\in T}} M(x,y)\right| < 2^{\frac{n}{2}}\sqrt{|S|\,|T|}$



$A(y)$ is sum of these values



$\left|\sum_{y\in T} A(y)\right|$ is sum of these values.

Pf Fix $R = S\times T$.

Let $A(y) = \sum_{x\in S} M(x,y)$

Then $\left|\sum_{\substack{x\in S\\ y\in T}} M(x,y)\right| = \left|\sum_{y\in T} A(y)\right|$

Let $g^S(x)$ = char. function of S (in 0/1)

$$g^S(x) = \sum_y \underbrace{\widehat{g^S}(y)} \cdot \chi_y(x) \quad \leftarrow \text{Fourier rep'n}$$

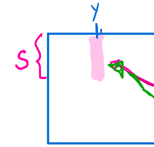
$$\widehat{g^S}(y) := \mathop{\mathbb{E}}_x\left[g^S(x)\cdot\chi_y(x)\right]$$

$$= \frac{1}{2^n}\sum_{x\in S}\chi_y(x)$$

$$= \frac{1}{2^n}\underbrace{\sum_{x\in S} M(x,y)}_{A(y)}$$

$$\therefore A(y) = 2^n\cdot\widehat{g^S}(y)$$

$$\boxed{\underline{\text{Claim}} \quad \forall S,T \quad \left| \sum_{\substack{x\in S\\ y\in T}} M(x,y) \right| < 2^{\frac{n}{2}} \sqrt{|S|\,|T|}}$$



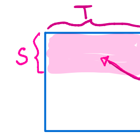
$A(y)$ is sum of these values

equals $\widehat{g^S(y)}$ up to normalization factor 2^n

$\underline{\text{Pf}}$ Fix $R = S \times T$.

Let $A(y) = \sum_{x\in S} M(x,y)$

Then $\left| \sum_{\substack{x\in S\\ y\in T}} M(x,y) \right| = \left| \sum_{y\in T} A(y) \right|$



when $T = X$ (all columns)

$\left| \sum_{y\in T} A(x) \right| = 2^n \sum_{y} \widehat{g^S(y)}$ sum of all Fourier coeff's of g^S

Let $g^S(x)$ = char. function of S $\left(g^S(x) = \begin{matrix} 1 & \text{if } x\in S \\ 0 & \text{ow} \end{matrix} \right)$

$g^S(x) = \sum_{y} \underbrace{\widehat{g^S(y)}} \cdot \chi_y(x)$ ← Fourier rep'n

$$\widehat{g^S(y)} := \mathbb{E}_{x}\left[g^S(x) \cdot \chi_y(x) \right]$$

$$= \frac{1}{2^n} \sum_{x\in S} \chi_y(x)$$

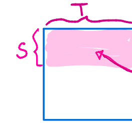
$$= \frac{1}{2^n} \underbrace{\sum_{x\in S} M(x,y)}_{A(y)}$$

$$\therefore A(y) = 2^n \cdot \widehat{g^S(y)}$$

This suggests we upper bound the square of $\left| \sum_{y\in T} A(y) \right|$

By Parseval, $\sum_{Y} \widehat{g^S(y)}^{\,2} = \mathbb{E}_{x}\left[g^S(x)^2 \right] = \frac{|S|}{2^n}$

$$\boxed{\text{Claim } \forall S,T \quad \left|\sum_{\substack{x\in S\\ y\in T}} M(x,y)\right| < 2^{n/2}\sqrt{|S|}\sqrt{|T|}}$$



when $T = X$ (all columns)

$\left|\sum_{y\in T} A(x)\right| = 2^n \sum_y \widehat{g^S}(y)$ sum of all Fourier coeff's of g^S

Parseval: $\sum_y \widehat{g^S}(y)^2 = \mathbb{E}_x[g^S(x)^2] = \frac{|S|}{2^n}$

Pf Fix $R = S\times T$.

Let $A(y) = \sum_{x\in S} M(x,y)$

Then $\left|\sum_{\substack{x\in S\\ y\in T}} M(x,y)\right| = \left|\sum_{y\in T} A(y)\right|$

$$\therefore \left(\frac{1}{|T|}\sum_{y\in T} A(y)\right)^2 \le \frac{1}{|T|}\left(\sum_{y\in T} A(y)^2\right)$$

multiply by $1/|T|$ and square. Then apply Jensen ineq: $\mathbb{E}[X]^2 \le \mathbb{E}[X^2]$

$$= \frac{1}{|T|}\, 2^{2n} \sum_y \widehat{g^S}(y)^2$$

Rewrite using $A(y) = 2^n \cdot \widehat{g^S}(y)$

$$\therefore \left|\sum_{y\in T} A(y)\right| \le \sqrt{|T|\, 2^{2n} \cdot \sum_y \widehat{g^S}(y)^2}$$

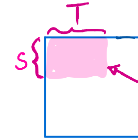
$$\le \sqrt{|T|\, 2^{2n}\, |S|\, \tfrac{1}{2^n}}$$

$$= \sqrt{|T|\cdot|S|}\cdot 2^{n/2}$$

Parseval:

$$\sum_{y\in T} \widehat{g^S}(y)^2 \le \sum_{y\in Y} \widehat{g^S}(y)^2 \le |S|/2^n$$

Claim $\forall S,T \quad \left| \sum_{\substack{x\in S \\ y\in T}} M(x,y) \right| \leq 2^{n/2} \sqrt{|S|\,|T|}$



$\left| \sum_{y\in T} A(y) \right|$ is sum of these values $= \frac{1}{2^n} \sum_{y} g^S(y)$

Parseval: $\sum_y \widehat{g^S}(y)^2 = \mathbb{E}_x [g^S(x)^2] = \frac{|S|}{2^n}$

Pf Fix $R = S \times T$.

Let $A(y) = \sum_{x\in S} M(x,y)$

Then $\left| \sum_{\substack{x\in S \\ y\in T}} M(x,y) \right| = \left| \sum_{y\in T} A(y) \right|$

multiply by $1/|T|$ and square. $\mathbb{E}[X]^2 \leq \mathbb{E}[X^2]$

$\therefore \left(\frac{1}{|T|} \sum_{y\in T} A(y) \right)^2 \leq \frac{1}{|T|} \left(\sum_{y\in T} A(y)^2 \right)$

Rewrite using Fourier coeff's

$= \frac{1}{|T|} 2^{2n} \sum_y \widehat{g^S}(y)^2$

$\therefore \left| \sum_{y\in T} A(y) \right| \leq \sqrt{|T|\, 2^{2n} \cdot \sum_y \widehat{g^S}(y)^2}$

Parseval

$\leq \sqrt{|T|\, 2^{2n}\, |S| \frac{1}{2^n}}$

$= \sqrt{|T| \cdot |S|} \cdot 2^{n/2}$

Intuition

- When $T = Y$, $\left| \sum_{\substack{x\in S \\ y\in T}} M(x,y) \right|^2 =$ weighted sum of squares of Fourier coeff's of f^S which is bounded by $\frac{|S|}{2^n}$ (Parseval)
- Why is $T = Y$ worst case? For smaller T, there can't be too many large fourier coeffs since sum of squares is bounded (2nd moment method)