Presentations

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

* You must submit your topic presentation by FEB 21
by email (tonipitassi@gmail.com)

- List 2 topics (1st and second choice)
including papers you will present
- List partner – you can work alone or with
a partner
- You may select up to 2 dates that are bad for you

Presentations cont'd

- Come to office hrs or make an appt if you want to discuss possible topics before Feb 21

Your Presentation should include

- Presentation in class ~45 mins (60 if 2)
- Lecture notes
- Questions for discussion, Open problems

- I am available to meet with you before your presentation date to go over your presentation – email me.

Last class

- $P^{cc}(f) \leq NP^{cc}(f) \cdot coNP^{cc}(f)$ / $\log$ (Partition number)

  min # of partitions of M_f into monochrom. subrectangles

  ?? poly dec cc

- Log Rank conjecture

Last class

- $P^{cc}(f) \le NP^{cc}(f) \cdot coNP^{cc}(f)$
- Log Rank conjecture

TODAY:

- Randomized CC
- Relationship between Randomized + Distributional
- Lower bounds via Discrepancy

RANDOMIZED CC

Recall: for $0 < \varepsilon < \frac{1}{2}$

a 2-sided cc protocol for f is a protocol Π such that:

$$\forall (x,y) \quad \Pr\left[\Pi(x,y) = f(x,y)\right] \geq 1-\varepsilon$$

$$BPP^{cc}_{\varepsilon}(f) = \min_{\substack{\text{protocols } \Pi \\ \text{with error } \varepsilon}} \ \max_{\substack{(x,y) \\ |x|=|y|=n}} \left[\#\text{bits sent by } \Pi \text{ on } (x,y)\right]$$

(public coin)

DISTRIBUTIONAL COMPLEXITY

Let μ be a probability distribution over $X \times Y$,
$X, Y = \{0,1\}^n$.

A deterministic protocol Π computes $f : X \times Y \to \{0,1\}$
with error $\leq \varepsilon$ wrt μ if: $\Pr_{(x,y)\sim\mu}[\Pi(x,y) = f(x,y)] \geq 1-\varepsilon$

The (μ, ε)-distributional cc of f, $D^{\mu}_{\varepsilon}(f)$,
is the minimum cost over all deterministic protocols
that compute f over μ with error $\leq \varepsilon$

DISTRIBUTIONAL COMPLEXITY

Let μ be a probability distribution over $X \times Y$, $X, Y = \{0,1\}^n$.

A deterministic protocol Π computes $f : X \times Y \to \{0,1\}$ with error $\leq \varepsilon$ wrt μ if: $\Pr_{(x,y)\sim\mu}\left[\Pi(x,y) = f(x,y)\right] \geq 1-\varepsilon$

The (μ, ε)-distributional cc of f, $D^{\mu}_{\varepsilon}(f)$, is the minimum cost over all deterministic protocols that compute f over μ with error $\leq \varepsilon$

Theorem $BPP^{cc}_{\varepsilon}(f) = \max_{\mu} D^{\mu}_{\varepsilon}(f)$

Theorem $BPP^{cc}_{\varepsilon}(f) = \max_{\mu} D^{\mu}_{\varepsilon}(f)$

Proof

① $BPP_{\varepsilon}(f) \geq \max_{\mu} D^{\mu}_{\varepsilon}(f)$:

Let Π be a BPP^{cc}_{ε} protocol for f of cost c.

$\therefore \Pr_{r}[\Pi(x,y,r) = f(x,y)] \geq 1-\varepsilon$

$\therefore$ by averaging, there exists some r^* st

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

✓

MiniMax Theorem

- A 2-player zero sum game:

  Player I has a finite set $A = \{a_1 \dots a_m\}$ of (pure) strategies

  Player II " " " " $B = \{b_1, .., b_n\}$ " " "

- Each player has a utility $U_I(a_i, b_j)$, $U_{II}(a_i, b_j)$

  for each pair (a_i, b_j) of strategies

- Zero sum: $\forall (i,j)\ U_I(a_i\, b_j) + U_{II}(a_i\, b_j) = 0$

  Example: $\forall (a_i, b_j)$ one player wins (utility $= +1$)
  and the other player loses (utility $= -1$)

Example

U_I:

	b_1	b_2	b_3	b_4	b_5
a_1	-1	-1	-1	-1	-1
a_2	-1	1	1	1	-1
a_3	-1	1	1	1	-1
a_4	-1	-1	1	1	-1
a_5	-1	-1	1	1	-1

U_{II}:

	b_1	b_2	b_3	b_4	b_5
a_1	1	1	1	1	1
a_2	1	-1	-1	-1	1
a_3	1	-1	-1	-1	1
a_4	1	1	-1	-1	1
a_5	1	1	-1	-1	1

Mixed strategies (p,q)

Player I: p is a probability distribution over $\{a_1, \dots, a_m\}$

$p_i = \text{Prob}(a_i)$

Player II: q is a probability distrib over $\{b_1, \dots, b_n\}$

Payoff for player I on (p,q) $\stackrel{\text{def}}{=} \displaystyle\sum_{i=1}^{m} \sum_{j=1}^{n} p_i \, U_I(a_i, b_j)\, q_j$

Example

U_I:

	b_1	b_2	b_3	b_4	b_5
a_1	-1	-1	-1	-1	-1
a_2	-1	1	1	1	-1
a_3	-1	1	1	1	-1
a_4	-1	-1	1	1	-1
a_5	-1	-1	1	1	-1

U_{II}:

	b_1	b_2	b_3	b_4	b_5
a_1	1	1	1	1	1
a_2	1	-1	-1	-1	1
a_3	1	-1	-1	-1	1
a_4	1	1	-1	-1	1
a_5	1	1	-1	-1	1

Mixed strategies (p,q)

Player I: p is a probability distribution over $\{a_1, \dots, a_m\}$
$p_i = \text{Prob}(a_i)$

Player II: q is a probability distrib over $\{b_1, \dots, b_n\}$

Payoff for player I on (p,q) $\triangleq \sum_{i=1}^{m} \sum_{j=1}^{n} p_i U_I(a_i, b_j) q_j$ ← $M(p,q)$

Player I wants to find mixed strategy p to maximize $\min_q M(p,q)$

Player II wants to find strategy q to minimize $\max_p M(p,q)$

MINIMAX THEOREM (Linear Programming Duality)

For every 2-person zero-sum game there exists an equilibrium strategy

That is, $\exists$ value v such that

$$v = \max_p \min_q M(p,q) = \min_q \max_p M(p,q)$$

Theorem $BPP^{cc}_{\varepsilon}(f) = \max_{\mu} D^{\mu}_{\varepsilon}(f)$

Proof

(2) $BPP_{\varepsilon}(f) \leq \max_{\mu} D^{\mu}_{\varepsilon}(f)$ (minimax argument)

Player I (protocol designer)

pure strategies : all c-bit deterministic protocols

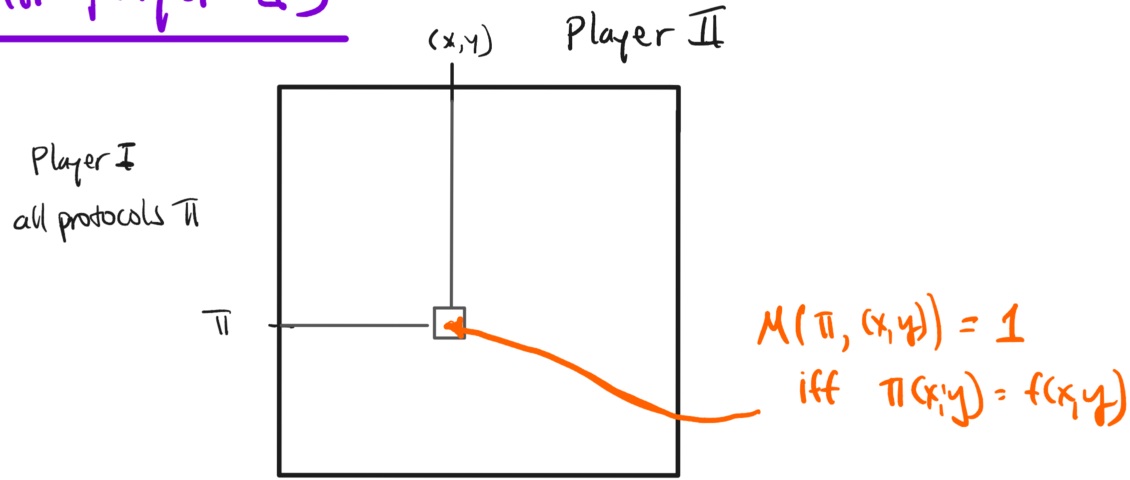
mixed strategies : distrib over all c-bit deterministic protocols

Player II (adversary)

Pure strategies : all inputs (x,y)

mixed strategies : all distributions over $X \times Y$

Payoff M (for player I)



Assume: $\forall \mu$ over $X \times Y$ $\exists$ a deterministic protocol (of cost c) such that

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

$\therefore$ $\min_{\mu} \max_{\Pi} M(\mu, \Pi) \geq 1-\varepsilon$

$\therefore$ By Minimax theorem, $\max_{\Pi} \min_{\mu} M(\mu, \Pi) \geq 1-\varepsilon$

$\therefore$ $\exists$ randomized protocol s.t. for all fixed (x,y), payoff is $\geq 1-\varepsilon$

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 <u>discrepancy</u> measure.

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

Let R be a subrectangle of M_f

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

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

small discrepancy means large rectangles balanced

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

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

says low discrepancy implies high cc wrt μ

Claim for every μ, $D^\mu_\varepsilon(f) \geq \log\left(\frac{1}{3\,\mathrm{Disc}_\mu(f)}\right)$

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

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

Claim for every μ, $D_\varepsilon^\mu(f) \geq \log\left(\frac{1}{3\,\mathrm{Disc}_\mu(f)}\right)$

says low discrepancy implies high cc wrt μ

Intuition behind claim for μ = uniform distribution

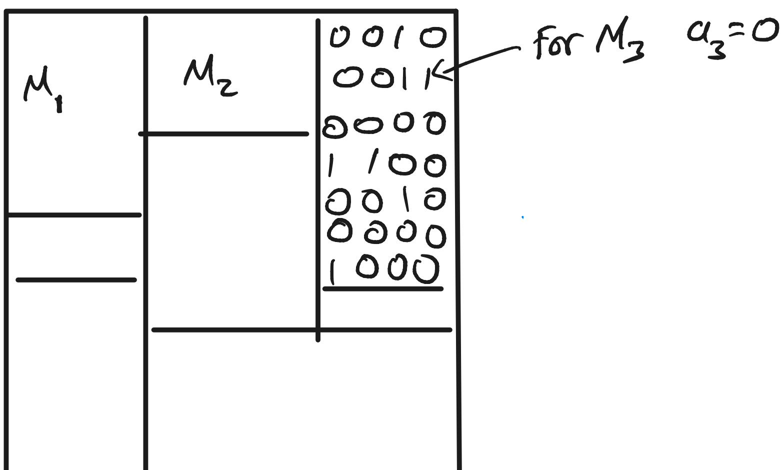
(same intuition for any μ)

- 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.
- Since all large R's are nearly balanced, Π must make a lot (too many) errors.

Claim If $D^{\mu}_{1/3}(f) \leq c$ then $\mathrm{DISC}_{\mu}(f) \geq \frac{1}{2^c}$

Pf Let Π be a c-cost deterministic protocol. So Π partitions M_f into disjoint subrectangles $M_1, M_2, \ldots, M_{2^c}$

We can assume wlog that $\forall i \in [2^c]$, Π returns most common answer, call it a_i, in M_i (most common wrt μ)



Claim If $D^{\mu}_{1/3}(f) \leq c$ then $\mathrm{Disc}_{\mu}(f) \leq \frac{1}{2^c}$

Pf Let Π be a c-cost deterministic protocol. So Π partitions M_f into disjoint subrectangles $M_1, M_2, \ldots, M_{2^c}$

We can assume wlog that $\forall i \in [2^c]$, Π returns most common answer, call it a_i, in M_i (most common wrt μ)

Since Π has error $\leq \frac{1}{3}$ wrt μ:

$$2^c \, \mathrm{Disc}_{\mu}(f) \geq \sum_{i=1}^{2^c} \mathrm{Disc}_{\mu}(M_i) = \sum_{i=0}^{2^c} \left| \mu(M_i \cap f^{-1}(a_i)) - \mu(M_i \cap f^{-1}(1-a_i)) \right|$$

$$\geq \sum_{i=1}^{2^c} \mu(M_i \cap f^{-1}(a_i)) - \sum_{i=1}^{2^c} \mu(M_i \cap f^{-1}(1-a_i))$$

$$\geq \frac{1}{3}$$

$$\therefore c \geq \log\left(\frac{1}{3\,\mathrm{Disc}_{\mu}(f)}\right)$$

Relating Discrepancy to λ_{max}

all eigenvalues small
$\Rightarrow$ pseudorandom
so large rectangles are
pretty balanced

<u>Lemma</u> (Eigenvalue Bound)
Let f be symmetric, with range $\{-1, 1\}$

Then $\mathrm{Disc}(f, A \times B) \leq 2^{-2n} \lambda_{max} \sqrt{|A| \times |B|}$

(λ_{max} is largest eigenvalue of M_f)

Lemma (Eigenvalue Bound) Let f be symmetric. μ = unif distrib.

Then $\mathrm{Disc}_\mu(f, A \times B) \le 2^{-2n} \lambda_{max} \sqrt{|A| \times |B|}$

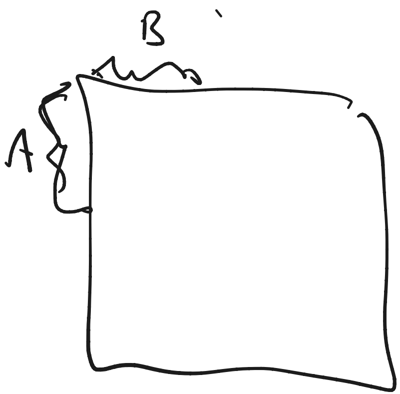
Proof Since M_f symmetric, eigenvectors v_i of M_f form orthonormal basis for $\mathbb{R}^N$ $(N = 2^n)$. Let λ_i be corresponding eigenvalue.

So $M_f v_i = \lambda_i v_i$

Note: entries of M_f in $\{-1, 1\}$

Write 1_A, 1_B in this basis:

$$1_A = \sum_i \alpha_i v_i \qquad \underbrace{1_B}_{} = \underbrace{\sum \beta_i v_i}_{}$$



$\underbrace{11111}_{}\ 0 \ldots$

Lemma (Eigenvalue Bound) Let f be symmetric. Then

$$\mathrm{Disc}(f, A\times B) \le 2^{-2n}\,\lambda_{max}\sqrt{|A|\times|B|}$$

sums up all ± 1 entries in $A \times B$

Pf $\underbrace{2^{2n}}\,\mathrm{Disc}(f, A\times B) = \left| 1_A M_f 1_B \right|$

$$= \left| \left(\sum \alpha_i v_i\right)^T \left(\sum \beta_i \lambda_i v_i\right) \right|$$

$v_1, v_2, \ldots$ an orthonormal basis

$$= \left| \sum \alpha_i \beta_i \lambda_i \right|$$

$$\le \lambda_{max} \left| \sum \alpha_i \beta_i \right|$$

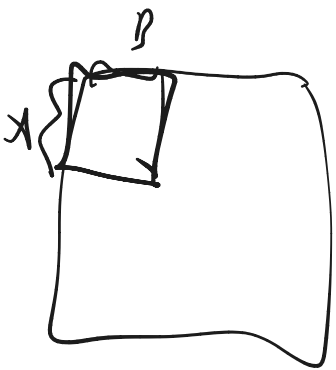
cauchy-Swartz

$$\le \lambda_{max} \sqrt{\sum \alpha_i^2}\sqrt{\sum \beta_i^2}$$

Parseval identity

$$\sum \alpha_i^2 = \|1_A\|^2 = |A|$$

$$= \lambda_{max}\sqrt{|A|\cdot|B|}$$



sum of squares of fourier coeffs of f $= f^2$

Randomized Lower Bound for IP (inner product)

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

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:

$$\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1
\end{bmatrix}$$

Randomized Lower Bound for IP (inner product)

$$H_0 = \square$$

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

$$H_0 : \begin{bmatrix} 1 \end{bmatrix}$$

$$H_1 : \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H_2 : \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$H_3 : \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

Facts about H_n:

1. every pair of rows is orthogonal

   $\therefore H_n^2 = N \cdot I$ $\qquad (N = 2^n)$

2. Rows correspond to all 2^n parity functions

3. Matrix is symmetric

4. $H_n^2 = H_n H_n^T = 2^n I_N$ (proof by induction)

5. By 4, $\forall v \quad v H_n H_n^T = 2^n v$

   $\therefore 2^n$ is the only eigenvalue of HH^T

   $\therefore$ the eigenvalues of H are $\pm 2^{n/2}$

$\therefore \lambda_{max}(H_n) = 2^{n/2}$

$\therefore$ By eigenvalue bound (μ = unif distrib)

$$2^{2n} \, \mathrm{Disc}(IP_n, A \times B) \leq \lambda_{max} \sqrt{|A||B|} = \sqrt{2^n |A| \cdot |B|}$$

Let $A = X$, $B = Y$

Then $\mathrm{Disc}(IP_n, A \times B) \leq \sqrt{2^n 2^n 2^n} \cdot 2^{-2n}$

$$= 2^{-n/2}$$

$\therefore BPP^{cc}_{1/3}(IP_n) \geq \log\left(\frac{1}{3\,\mathrm{Disc}(IP_n)}\right) = \Omega(n)$

Set Disjointness: KS Razborov - - - - - - To come soon

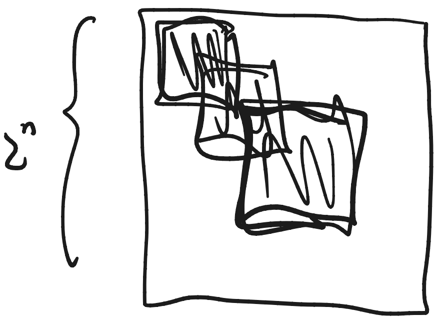
$\mathrm{DISJ}(x, y) = 1$ iff $|x \cap y| \neq \phi$

(ie $\exists i \; x_i = y_i = 1$)

DISJ easy for NP^{cc}

Theorem $\exists \mu$ such that $D^{\mu}_{\varepsilon}(\mathrm{DISJ}) = \Omega(n)$.

$\therefore \; BPP_{\varepsilon}(\mathrm{DISJ}) = \Omega(n)$



Can cover all 1's of DISJ by n 1-mono. subrectangles

$R_i = \{(x, y) \mid x_i = y_i = 1\}$

A related $\overset{\text{cc-like}}{\text{measure}}$ (related to discrepancy)

Information complexity of f

given a distribution $\underline{\mu}$ on M_f, given a protocol Π

we define

$$IC_\mu(X \mid Y, \Pi)$$ = "How much info Bob (holds y) learns about x from protocol

r.v. ~~the~~ marginal of μ on X

$$IC_\mu(Y \mid X, \Pi)$$

external info → $IC(X, Y \mid \Pi)$

Another related method to lower Bound randomized cc via discrepancy (that generalizes to NOF)

Called "BNS" method after seminal paper by Babai, Nisan, Szegedy introducing NOF cc and this method

Theorem [BNS bound]

Let $f: X \times Y \to \{-1,1\}$, μ a distribution over $X \times Y$. Then

$$\mathrm{DISC}_\mu(f)^2 \leq |Y| \cdot \sum_{x,x' \in X} \left| \sum_{y \in Y} \mu(x,y)\,\mu(x',y)\,f(x,y)\,f(x',y) \right|$$

Replacing sums by expectations can also write as:

$$\frac{\mathrm{Disc}_\mu(f)^2}{|X|^2 \times |Y|^2} \leq \mathop{\mathbb{E}}_{x,x'} \left| \mathop{\mathbb{E}}_{y} \mu(x,y)\,\mu(x',y)\,f(x,y)\,f(x',y) \right|$$

Theorem [BNS bound]

Replacing sums by expectations can also write as:

$$\frac{\mathrm{Disc}_\mu(f)^2}{|X|^2 \times |Y|^2} \leq \mathop{\mathbb{E}}_{x,x'} \left| \mathop{\mathbb{E}}_{y} \mu(x,y)\,\mu(x',y)\,f(x,y)\,f(x',y) \right|$$

Proof (μ = unif. distrib)

$$\mathrm{disc}(f, A\times B) = \left| \mathop{\mathbb{E}}_{x,y} 1_A(x)\, 1_B(x)\, f(x,y) \right|$$

so

$$\mathrm{disc}(f, A\times B)^2 = \left(\mathop{\mathbb{E}}_{y} 1_B(y) \mathop{\mathbb{E}}_{x} 1_A(x)\, f(x,y) \right)^2$$

$$\leq \mathop{\mathbb{E}}_{y} \left(\mathop{\mathbb{E}}_{x} 1_A(x)\, f(x,y) \right)^2$$

$$= \mathop{\mathbb{E}}_{y} \left(\mathop{\mathbb{E}}_{x,x'} 1_A(x)\, 1_A(x')\, f(x,y)\, f(x',y) \right)$$

$$= \mathop{\mathbb{E}}_{x,x'} 1_A(x)\, 1_A(x') \left(\mathop{\mathbb{E}}_{y} f(x,y)\, f(x',y) \right)$$

$$\leq \mathop{\mathbb{E}}_{x,x'} \left| \mathop{\mathbb{E}}_{y} f(x,y)\, f(x',y) \right|$$

Cauchy Swartz

$$\left(\mathbb{E}[Z]\right)^2 \leq \mathbb{E}[Z^2]$$

Theorem [BNS bound]

Let $f: X \times Y \to \{-1,1\}$, μ a distribution over $X \times Y$. Then

$$\mathrm{DISC}_\mu(f)^2 \leq |Y| \cdot \sum_{x,x' \in X} \left| \sum_{y \in Y} \mu(x,y)\,\mu(x',y)\,f(x,y)\,f(x',y) \right|$$

Proof (μ arbitrary distrib). Let $R = A \times B$

Let $\alpha_x = 1$ $\forall x \in A$, $\beta_y = 1$ $\forall y \in B$, and for all other x, y

Let α_x, β_y be independent r.v.'s uniformly distributed over $\{-1,1\}$.

Then

$$\mathrm{Disc}_\mu(R) = \left| \sum_{(x,y) \in R} \mu(x,y)\, f(x,y) \right|$$

$$= \left| \sum_{(x,y) \in R} \mathbb{E}[\alpha_x \beta_y]\, \mu(x,y) f(x,y) + \sum_{(x,y) \notin R} \mathbb{E}[\alpha_x \beta_y]\, \mu(x,y) f(x,y) \right|$$

$$= \left| \mathbb{E}\Big[\sum_{(x,y)} \alpha_x \beta_y\, \mu(x,y)\, f(x,y) \right|$$

So there is a fixed assignment $\alpha_x, \beta_y \in \{-1,1\}$ for all x,y such that

$$\mathrm{Disc}_\mu(f) \leq \left| \sum_{(x,y)} \alpha_x \beta_y\, \mu(x,y) f(x,y) \right|$$

Theorem [BNS bound]

Let $f: X \times Y \to \{-1,1\}$, μ a distribution over $X \times Y$. Then

$$\mathrm{DISC}_\mu(f)^2 \leq |Y| \cdot \sum_{x,x' \in X} \left| \sum_{y \in Y} \mu(x,y)\mu(x',y) f(x,y) f(x',y) \right|$$

Proof cont'd

So there is a fixed assignment $\alpha_x, \beta_y \in \{-1,1\}$ for all x,y such that

$$\mathrm{Disc}_\mu(f) \leq \left| \sum_{(x,y)} \alpha_x \beta_y \mu(x,y) f(x,y) \right|$$

$$\therefore \quad \mathrm{Disc}_\mu(F)^2 \leq |Y| \sum_y \left(\beta_y \sum_x \alpha_x \mu(x,y) f(x,y) \right)^2$$

Squaring both sides plus cauchy schwartz

$$\left(\mathbb{E}[Z]\right)^2 \leq \mathbb{E}[Z^2]$$

$$= |Y| \sum_{x,x'} \alpha_x \alpha_{x'} \sum_y \mu(x,y)\mu(x',y) f(x,y) f(x',y)$$

$$\leq |Y| \sum_{x,x'} \left| \sum_y \mu(x,y)\mu(x',y) f(x,y) f(x',y) \right|$$