HARDNESS ESCALATION IN COMMUNICATION COMPLEXITY

Toniann Pitassi

INSTITUTE FOR ADVANCED STUDY · TRUTH · BEAUTY · FOUNDED 1930

UNIVERSITY OF TORONTO

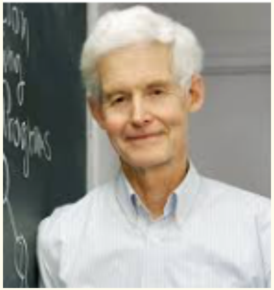
Joint work with:

Mika Göös

Thomas Watson

Robert Robere

Ben Rossman



Stephen Cook

HARDNESS ESCALATION IN COMMUNICATION ?

WHAT I WILL BE TALKING ABOUT:

- Models of interactive communication / information
- Their importance in so many areas of computer science ⟸ they capture information bottleneck
- A (relatively) new lower bound tool hardness escalation / lifting
  - applications
  - some ideas of proof

Communication complexity (Yao '79)

$x = 10111$

$y = 10110$

$1 \rightarrow$

$\leftarrow 0$

$0 \rightarrow$

$\leftarrow 1$

$0 \rightarrow$

last bit is $f(x,y)$

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

$$cc(f) = \min_{\pi \text{ computing } f} cc(\pi)$$

Example: EQUALITY(x,y)

$x = 10111$

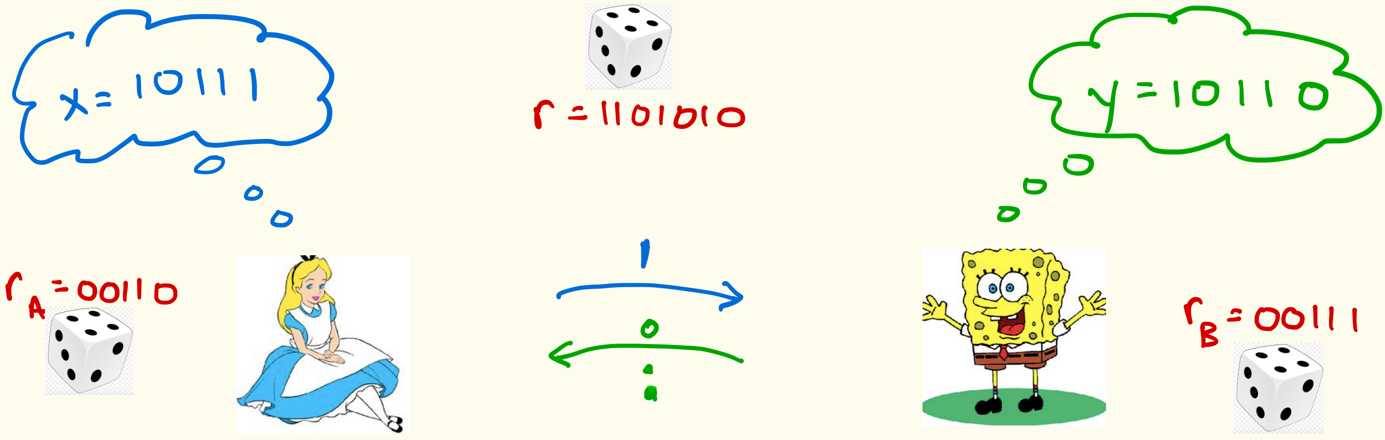
$x \stackrel{?}{=} y$

$y = 10110$

Deterministic CC $= \Omega(n)$

Randomized CC $= O(1)$

Randomized Communication Complexity



Π computes f if:

$$\forall x,y \quad \Pr_{r_A, r_B, r}\left[\Pi(x,y) \neq f(x,y)\right] \leq 1/3$$

$$rCC(f) = \min_{\Pi} \; CC(\Pi)$$

Equivalent Definition of Randomized CC

$rCC(f)$ = cost of best randomized protocol for f, with error $< 1/3$ on <u>every</u> input

$CC^{\mu}_{1/3}(f)$ = cost of best deterministic protocol for f over distribution μ, with error $< 1/3$

<u>Theorem</u> (Yao Min-max for cc)

$$rCC(f) = \max_{\mu} CC^{\mu}_{1/3}(f)$$

Rich History of Communication Complexity

- The most successful concrete model of computation for proving lower bounds

applications: streaming algorithms
cryptography (secret sharing schemes)
proof complexity
limitations of LP / SDP for NP-hard problems
game theory
distributed computing
graph theory
data structures,
circuit complexity,
...

Rich History of Communication Complexity

- The most successful concrete model of computation for proving lower bounds

- Many variants: communication <u>analog</u> of complexity classes

$$P^{cc}, \quad BPP^{cc}, \quad NP^{cc}$$

Rich History of Communication Complexity

- The most successful concrete model of computation for proving lower bounds
- Many variants: communication <u>analog</u> of complexity classes
- Tightly connected to <u>interactive</u> theory of communication

Classical Information Theory [Shannon '48]

$11010111 \sim Z$

transmitter

receiver

Data Compression Thm Every message can be compressed to its info-theoretic content

$$H(Z) \leq C(Z) < H(Z) + 1$$

[Noiseless, Huffman Coding]

(Interactive) Information complexity [CSWY '01, BJKS '04, BBCR '13]

$x = 10111$

$y = 10110$

$r_A = 1101011$

110

0

$\vdots$

$r_B = 0011011$

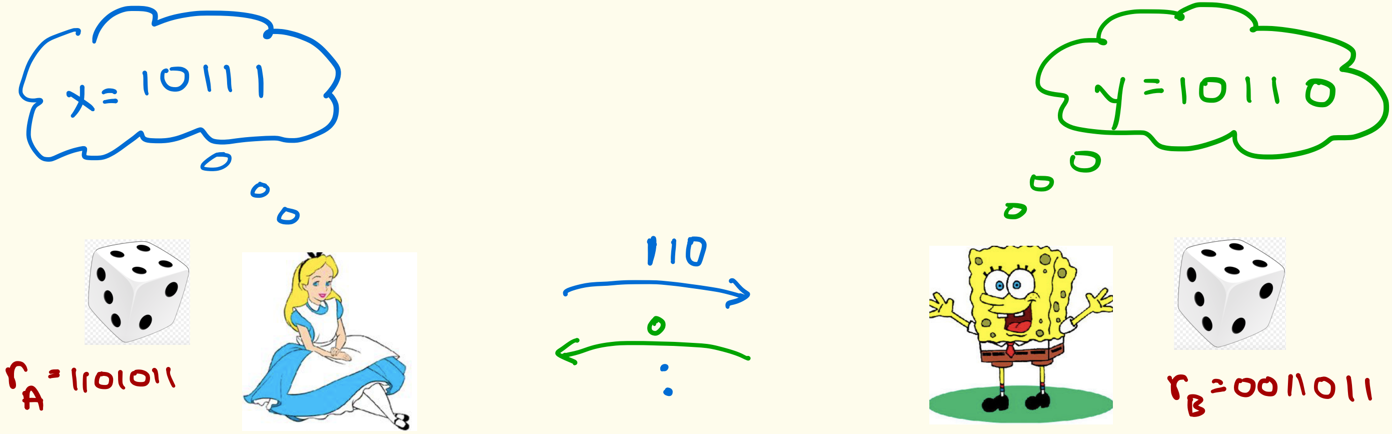
Internal Information Cost $IC_\mu(\Pi)$:

$$I_\mu(\Pi; Y \mid X) + I_\mu(\Pi; X \mid Y)$$

What Bob learns about X

What Alice learns about Y

(Interactive) Information complexity

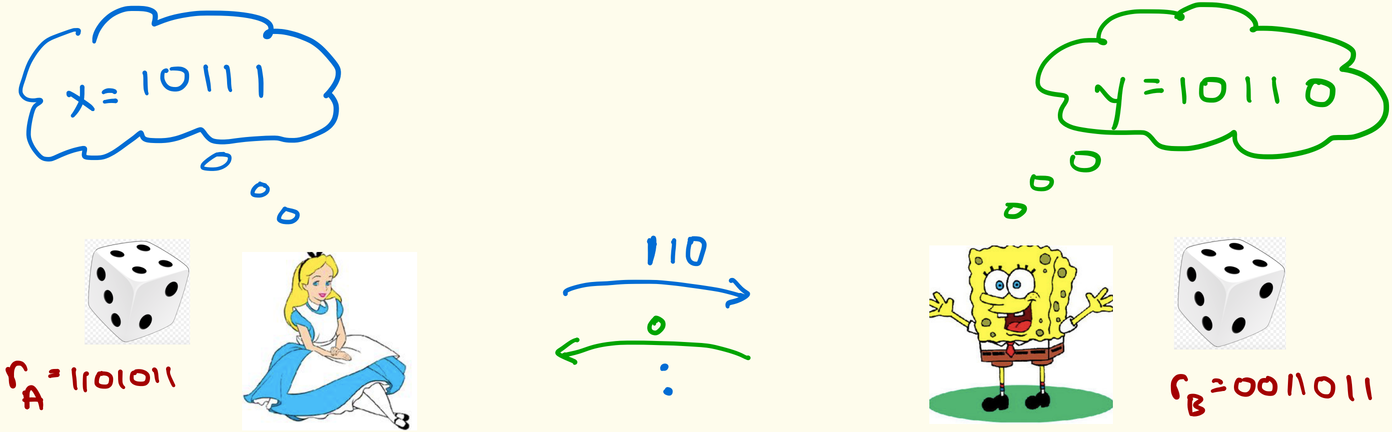


Internal Information Cost $IC_\mu(\Pi)$:

$$I_\mu(\Pi ; Y \mid X) + I_\mu(\Pi ; X \mid Y)$$

* $\Pi(x,y)$ is a random variable – distribution over transcript of messages sent on (x,y)
* $(x,y) \sim \mu$

(Interactive) Information complexity



$$IC_\mu(f) = \min_{\Pi \text{ computing } f} IC_\mu(\Pi)$$

$$IC(f) = \max_\mu IC_\mu(f)$$

<u>Big Q</u>: Can interactive communication be compressed? $[\mathrm{IC}(f) \approx \mathrm{rCC}(f)\,?]$

- Essentially all lower bound measures used to prove randomized CC LBs also give IC lower bounds

Big Q: Can interactive communication be compressed?

- Essentially all lower bound measures used to prove randomized CC LBs also give IC lower bounds

- Compression Results

  1. communication C, information I $\Longrightarrow \min\{2^{I}, \sqrt{I \cdot C}\}$ [BBCR'10]

  2. product distributions $\Longrightarrow \min\{I^2, I \cdot \mathrm{polylog}(C)\}$ [BBCR'10, KOL'16, S'16]

suppressing low order terms

Big Q: Can interactive communication be compressed?

- Essentially all lower bound measures used to prove randomized CC LBs also give IC lower bounds
- Compression Results
- Lower Bound [Ganor-Kol-Raz '16]

  There is a boolean function f with $I = \log\log\log n$ but $rCC \in \Omega(\log\log n)$

  exponential separation but in <u>very</u> low CC regime.

Big Q: Can interactive communication be compressed?

OPEN #1 $\quad I, C \overset{?}{\not\Longrightarrow} I \cdot \text{polylog}\, C$ communication

BIG QUESTION: $IC(f) = rCC(f)$?

[Can communication be compressed in the interactive setting?]

OPEN #1 $\quad I, C \overset{?}{\not\Rightarrow} I \cdot \mathrm{polylog}\, C$ communication

OPEN #2 $\quad I \in \Omega(\log n) \overset{?}{\Rightarrow} \mathrm{poly}(I)$ communication

LOWER BOUND METHODS IN CC

1. Find a matrix property implied by low cc protocol that f doesn't have

   [ex. rank, discrepancy, corruption]

2. Information theory proof (via direct sum property)

3. Hardness Escalation (Query to communication Lifting)

LOWER BOUND METHODS IN CC

1. Find a matrix property implied by low cc protocol that f doesn't have

   [ex. rank, discrepancy, corruption]

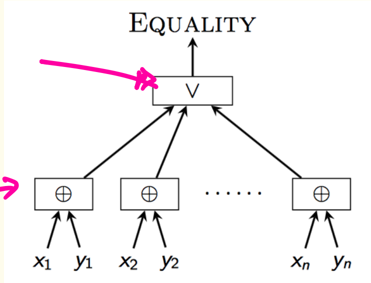
2. Information theory proof (via direct sum property)

3. Hardness Escalation (Query to communication Lifting)

INTUITION: MOST HARD COMMUNICATION PROBLEMS ARE COMPOSED FUNCTIONS $f \circ g^n$

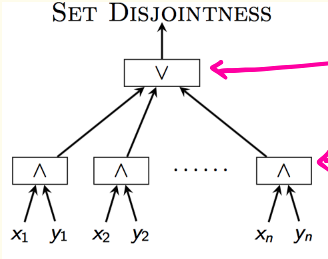
f : OR

g: =



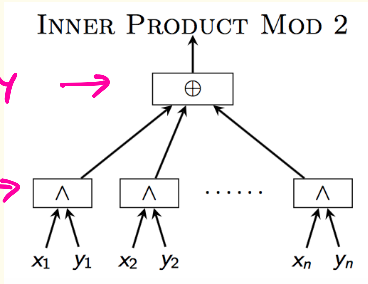
f : OR

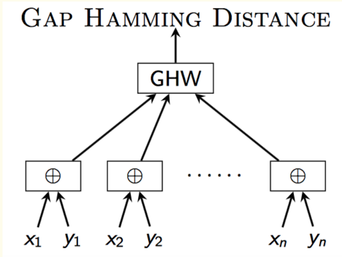
g: AND



f: PARITY

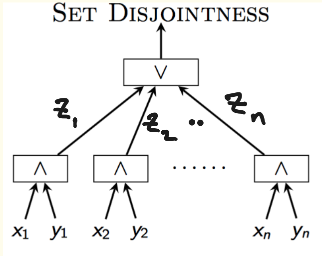
g: AND



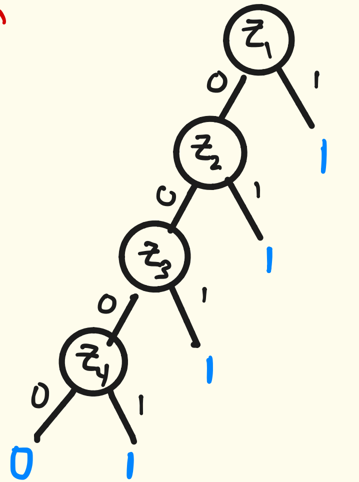


HARDNESS ESCALATION ("Lifting")

<u>Intuition</u> If g sufficiently hard, the best protocol for $f \circ g^n$ will simulate the best <u>decision tree</u> for f

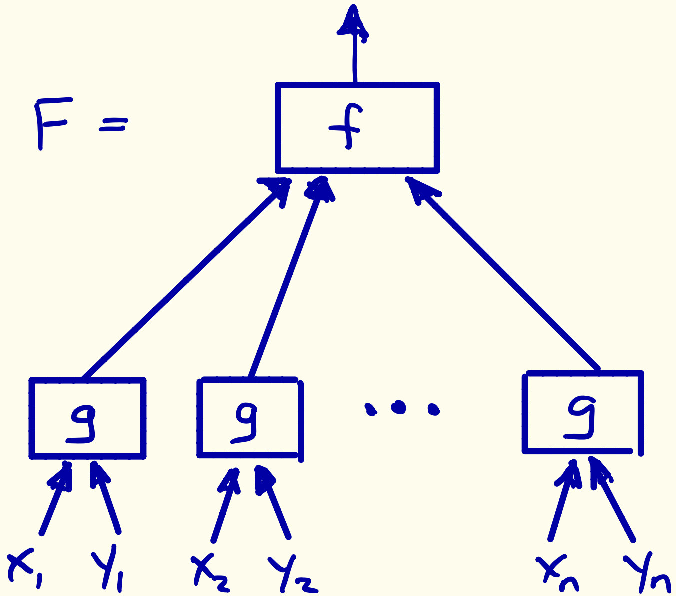


Query complexity of f $\approx$ Communication complexity of $f \circ g^n$



Query-to-communication Lifting

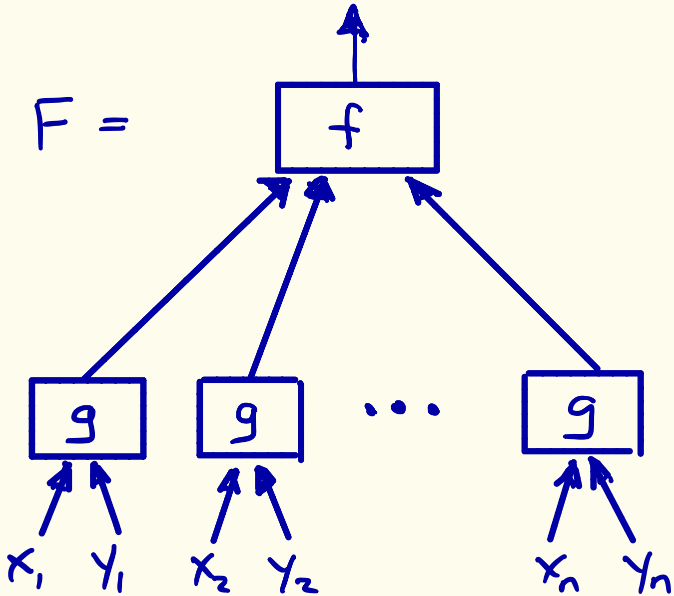
$f : \{0,1\}^n \to \{0,1\} \rightsquigarrow F =$



$g(x_i\, y_i) \to \{0,1\}$ small "gadget"

Query-to-communication Lifting

$f:\{0,1\}^n \to \{0,1\} \rightsquigarrow F =$

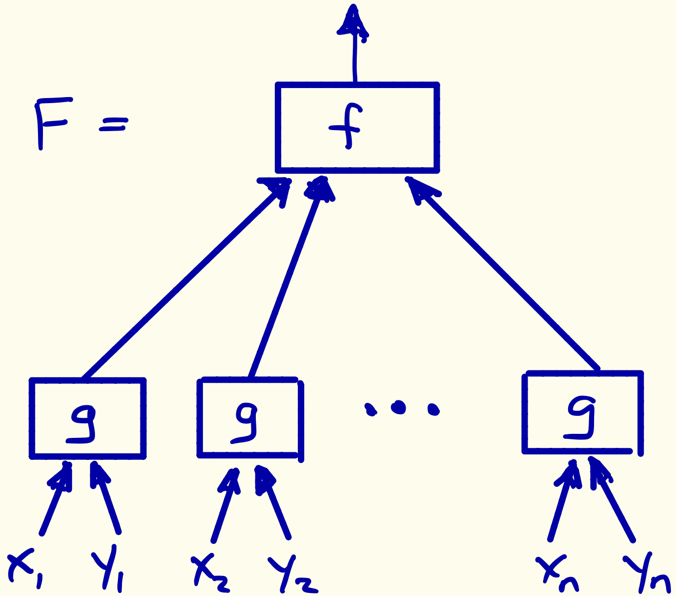


Easiness Escalation:

Decision tree for f $\rightsquigarrow$ Communication protocol for F

Query-to-communication Lifting

$f:\{0,1\}^n \to \{0,1\} \quad \Rrightarrow \quad F =$



Hardness Escalation

Decision tree for f $\Leftarrow$ communication protocol for F

An Abridged History

[Nisan, Wigderson FOCS '94]

Rank vs. communication

[Raz, McKenzie FOCS '97]

Monotone circuit depth (deterministic Lifting)

[Sherstov STOC '08]

Polynomial degree → rank (pattern matrix method)

[Göös, Lovett, Meka, Watson, Zuckerman STOC '15]

NonNeg Junta degree → NonNeg rank

[Göös, Pitassi, Watson FOCS '15]

Explicit/simplified deterministic lifting

[Göös, Pitassi, Watson FOCS '17]

BPP (randomized) lifting

Lifting Theorems Makes Lower Bounds Easy!

2 step Recipe :

1. Prove problem specific query lower bound

2. Apply Lifting theorem to obtain communication complexity lower bound

Applications

1. Monotone formula size / circuit lower bounds
2. Cryptography: Lower bounds for Linear secret sharing schemes (+ span programs)
3. Linear Programming: Extended formulations
4. Game Theory: Nash Equilibrium
5. Graph Theory: Alon-Saks-Seymour Conjecture
6. Proof complexity
7. Communication complexity separations
8. Quantum Lower Bounds

Applications

1. Monotone formula size / circuit lower bounds
2. Cryptography: Lower bounds for Linear secret sharing schemes (+ span programs)
3. Linear Programming: Extended formulations
4. Game Theory: Nash Equilibrium
5. Graph Theory: Alon-Saks-Seymour Conjecture
6. Proof Complexity
7. Communication complexity separations
8. Quantum Lower Bounds

[KW] COMMUNICATION COMPLEXITY LOWER BOUNDS $\Longrightarrow$ FORMULA SIZE LOWER BOUNDS

Let f be a Boolean function

$\boxed{KW_f}$: Alice gets $x \in f^{-1}(1)$ Bob gets $y \in f^{-1}(0)$

$KW_f(x,y)$: output i s.t. $x_i \neq y_i$

Theorem [KW]

The communication complexity of KW_f equals the minimum Boolean circuit depth for f!

[KW] COMMUNICATION COMPLEXITY LOWER BOUNDS $\Longrightarrow$ FORMULA SIZE LOWER BOUNDS

Let f be a Boolean function

$\boxed{mKW_f}$: Alice gets $x \in f^{-1}(1)$ Bob gets $y \in f^{-1}(0)$

$mKW_f(x,y)$: output i s.t. $x_i = 1$ $y_i = 0$

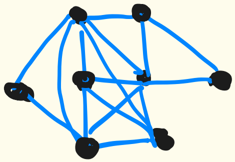
Theorem [KW]

The communication complexity of KW_f equals the minimum Boolean circuit depth for f

monotone

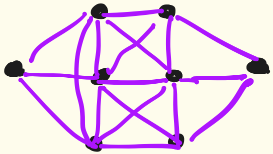
Example f = 5CLIQUE

Alice



$g \in f^{-1}(1)$

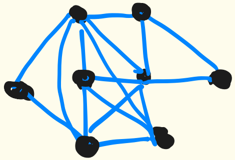
Bob



$g \in f^{-1}(0)$

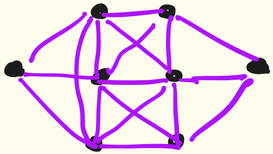
Example $f = 5\text{CLIQUE}$

Alice



$g \in f^{-1}(1)$

Bob



$g \in f^{-1}(0)$

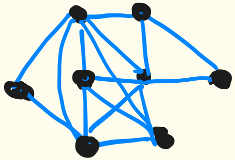
Theorem Monotone formulas for clique require size $2^{\Omega(n/\log n)}$

USES BPP LIFTING Theorem

Previous proofs: $2^{\Omega(n^{\varepsilon})}$

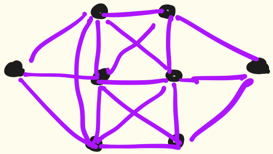
Example f = 5CLIQUE

Alice



$g \in f^{-1}(1)$

Bob



$g \in f^{-1}(0)$

Theorem [GGKS '18]

Monotone circuits for clique require exponential size

New proof using BPP LIFTING Theorem

Applications

1. Monotone formula size / circuit lower bounds
2. Cryptography: Lower bounds for Linear secret sharing schemes (+ span programs)
3. Linear Programming: Extended formulations
4. Game Theory: Nash Equilibrium
5. Graph Theory: Alon-Saks-Seymour Conjecture
6. Proof complexity
7. Communication complexity separations
8. Quantum Lower Bounds

(MONOTONE) SPAN PROGRAMS [KW '93]

$$
\begin{array}{c}
x_1 \\ x_2 \\ x_2 \\ x_3 \\ x_4 \\ x_5
\end{array}
\begin{array}{|c|c|c|c|c|c|c|c|}
\hline
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ \hline
0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline
0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline
\end{array}
\qquad M
$$

accept input α iff $\vec{1}$ in $\mathrm{span}(M|_\alpha)$

(MONOTONE) SPAN PROGRAMS

$$
\begin{array}{c}
x_1 \\ x_2 \\ x_2 \\ x_3 \\ x_4 \\ x_5
\end{array}
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}
\quad M
$$

Example: $\alpha = 11001$

is accepted !

IMPORTANCE OF SPAN PROGRAMS

- poly size monotone SPAN programs stronger than formulas (and incomparable to monotone circuits
- Monotone Span Programs characterize Linear Secret Sharing Schemes
- Close connection to Quantum Query Algorithms

Theorem [P-Robere '18] Clique requires exponential-size monotone span programs

(and hence expl size linear secret sharing schemes)

Applications

1. Monotone formula size / circuit lower bounds
2. Cryptography: Lower bounds for Linear secret sharing schemes (+ span programs)
3. Linear Programming: Extended formulations
4. Game Theory: Nash Equilibrium
5. Graph Theory: Alon-Saks-Seymour Conjecture
6. Proof complexity
7. Communication complexity separations
8. Quantum Lower Bounds

2-Player ε-Nash is Hard

2 players. Each has an N×N payoff matrix

.3	.6	.5
.2	.4	.1
.9	0	1

.1	.5	.9
.2	.4	.1
1	.9	0

2-Player ε-Nash is Hard

2 players. Each has an $N \times N$ payoff matrix

$$A = \begin{bmatrix} .3 & .6 & .5 \\ .2 & .4 & .1 \\ .9 & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} .1 & .5 & .9 \\ .2 & .4 & .1 \\ 1 & .9 & 0 \end{bmatrix}$$

$(\hat{x}, \hat{y})$ is an ε-Nash Equilibrium if:

$$\hat{x}^T A \hat{y} \geq x^T A \hat{y} - \varepsilon \qquad \forall x$$
$$\hat{x}^T B \hat{y} \geq \hat{x}^T B y - \varepsilon \qquad \forall y$$

Finding ε-Nash Equilibrium is Hard

Theorem [Göös-Rubinstein '18]

The randomized communication complexity of finding an ε-Nash equilibrium is $\geq N^{2-o(1)}$

.3	.6	.5
.2	.4	.1
.9	0	1

.1	.5	.9
.2	.4	.1
1	.9	0

Applications

1. Monotone formula size / circuit lower bounds
2. Cryptography: Lower bounds for Linear secret sharing schemes (+ span programs)
3. Linear Programming: Extended formulations
4. Game Theory: Nash Equilibrium
5. Graph Theory: Alon-Saks-Seymour Conjecture
6. Proof Complexity
7. Communication complexity separations
8. Quantum Lower Bounds

Clique vs Independent Set

Clique $x \subseteq [n]$
of G

Independent set $y \subseteq [n]$
of G

$$CIS_G(x,y) = \begin{cases} 0 & \text{if } |x \cap y| = 1 \\ 1 & \text{if } |x \cap y| = 0 \end{cases}$$

Clique vs Independent Set

Yannakakis Question:

Is there an $O(\log n)$ Nondeterministic cc protocol for CIS_g ?

$\Updownarrow$

Alon-Saks-Seymour Conjecture:

$$\forall g \quad \chi(g) \leq \mathrm{poly}(bp(g)) \,?$$

Clique vs Independent Set

Yannakakis Question:

Is there an $O(\log n)$ Nondeterministic cc protocol for CIS_g ?

$\Updownarrow$ BOTH ARE FALSE !

Alon-Saks-Seymour Conjecture:

$$\forall g \quad \chi(g) \leq \text{poly}(bp(g)) \ ?$$

Theorem [Göös '15]

$\exists g$, any Nondet. cc protocol for $CIS_g \in \Omega(\log^{1.13} n)$

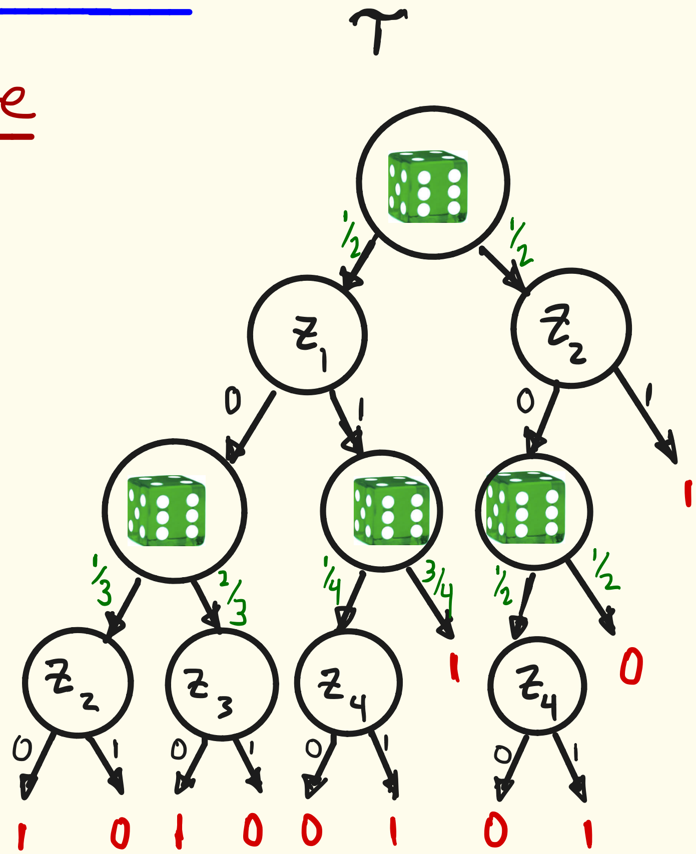
BPP LIFTING THEOREM

Theorem [Göös-P-Watson '17]

Randomized decision tree complexity of f $\approx$ Randomized communication complexity of $f \circ g^n$

Randomized Decision Trees

A **randomized decision tree** for $f(z_1 z_2 z_3 z_4)$:

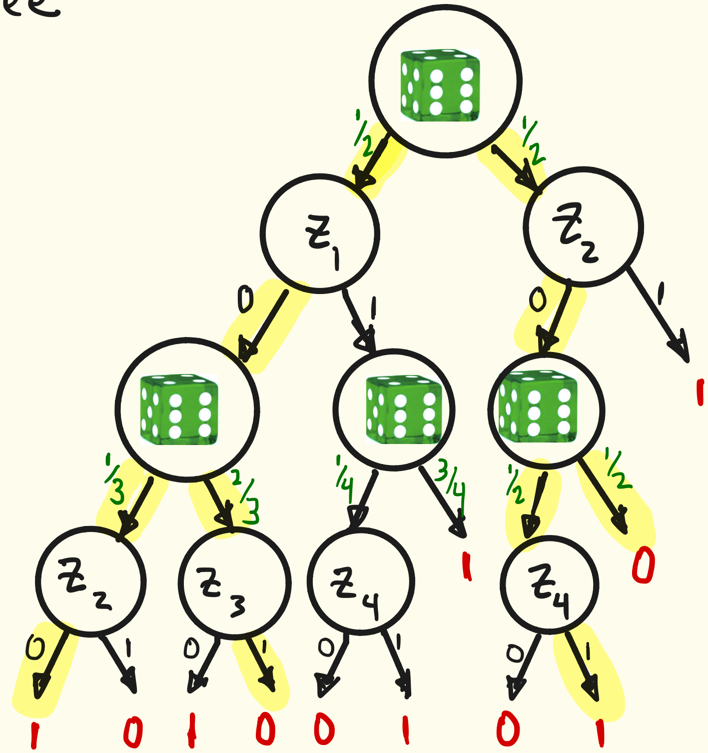


height(T) = max number of decision vertices

Randomized Decision Trees

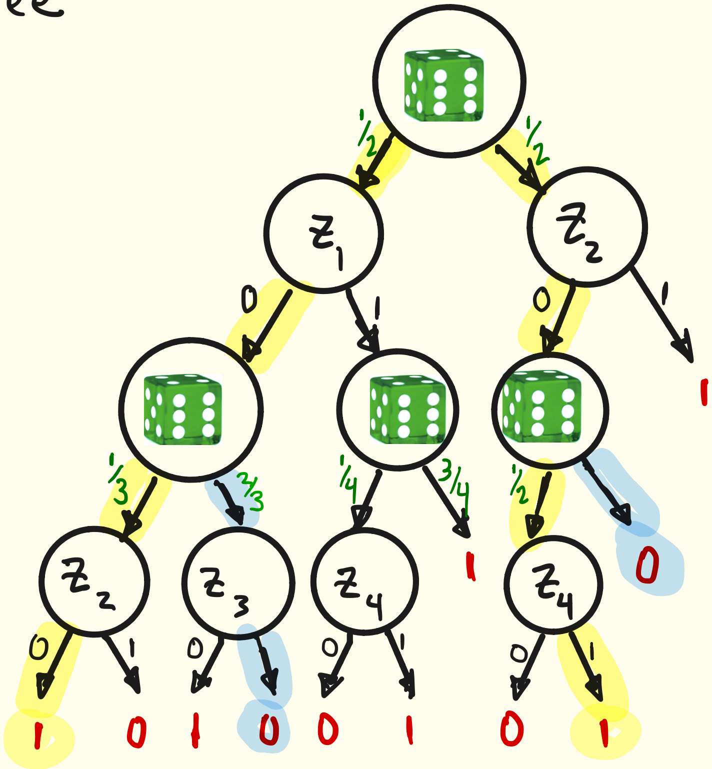
A randomized decision tree for $f(z_1 z_2 z_3 z_4)$:

$\Pr[T(0011) = 1]$



Randomized Decision Trees

A randomized decision tree for $f(z_1 z_2 z_3 z_4)$:



$$\Pr[T(0011) = 1]$$

$$= \frac{1}{2}\frac{1}{3} + \frac{1}{2}\frac{1}{2}$$

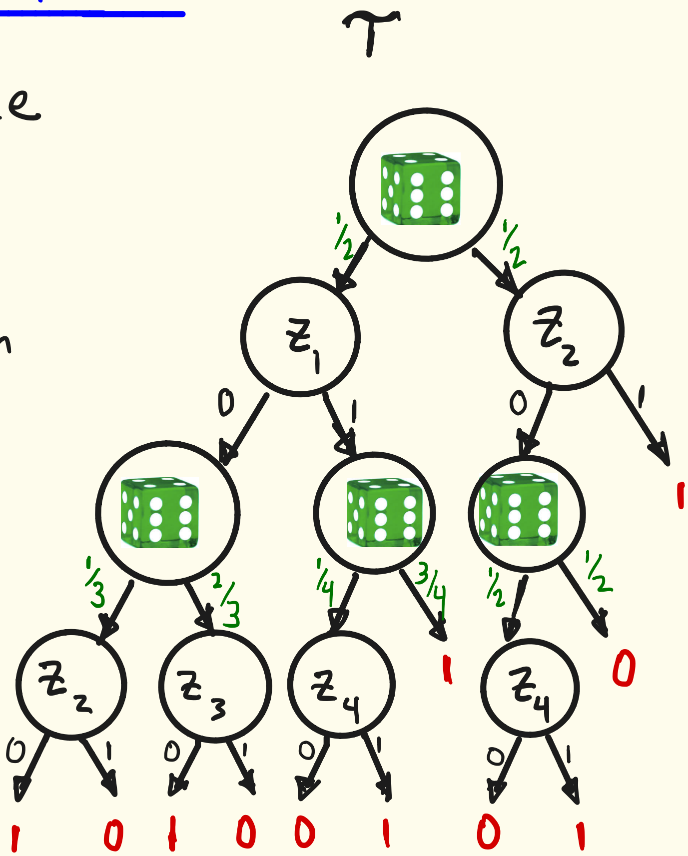
$$= \frac{5}{12}$$

Randomized Decision Trees

A randomized decision tree for $f(z_1 z_2 z_3 z_4)$:

$\mathcal{T}$ <u>computes</u> f if $\forall z \in \{0,1\}^n$

$$\Pr[\mathcal{T}(z) \neq f(z)] \leq 1/3$$



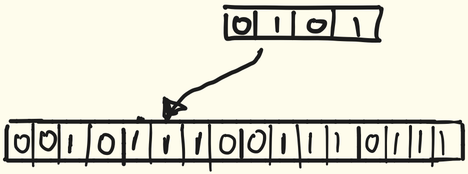
BPP LIFTING THEOREM

Theorem [GPW '17]

Randomized decision tree complexity of f $\approx$ Randomized communication complexity of $f \circ g^n$

g = "index gadget"

$g(x,y) = x|_y$

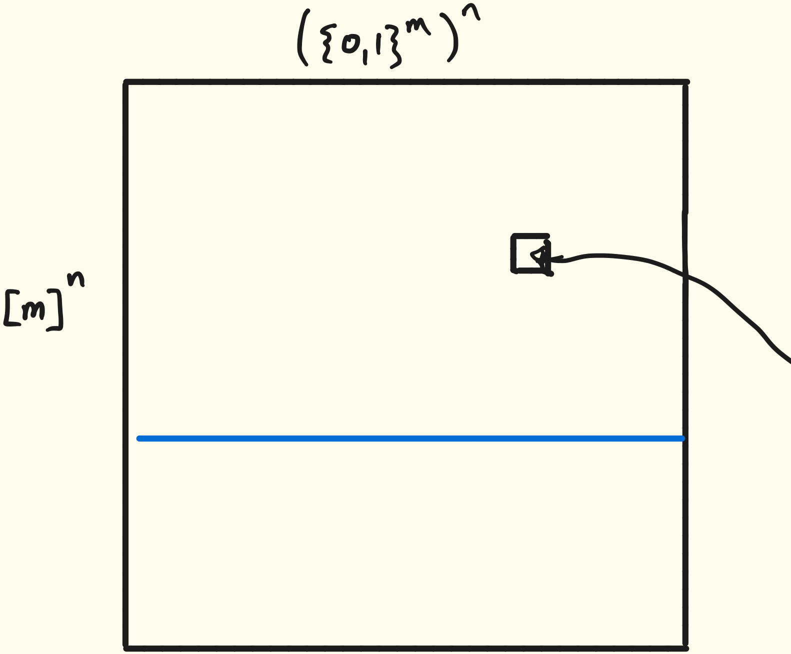


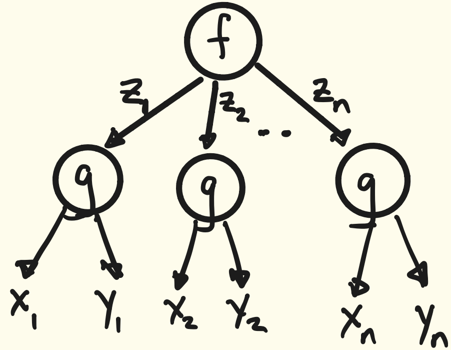
$x \in [m]$

$y \in \{0,1\}^m$

HOW TO PROVE BPP LIFTING THEOREM ?

wlog start with a deterministic cc protocol Π for $f \circ g^n$

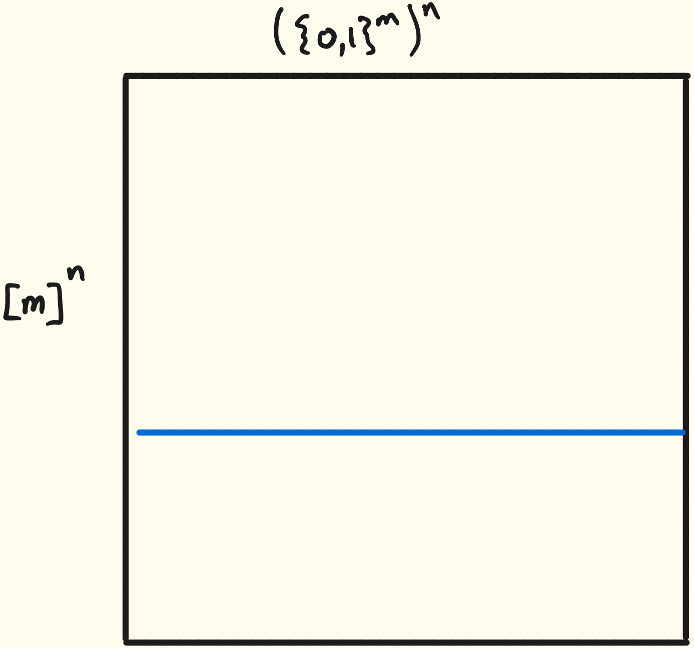


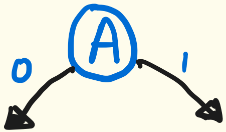


Each entry $((x_1..x_n),(y_1..y_n))$
labelled by $z \in \{0,1\}^n$

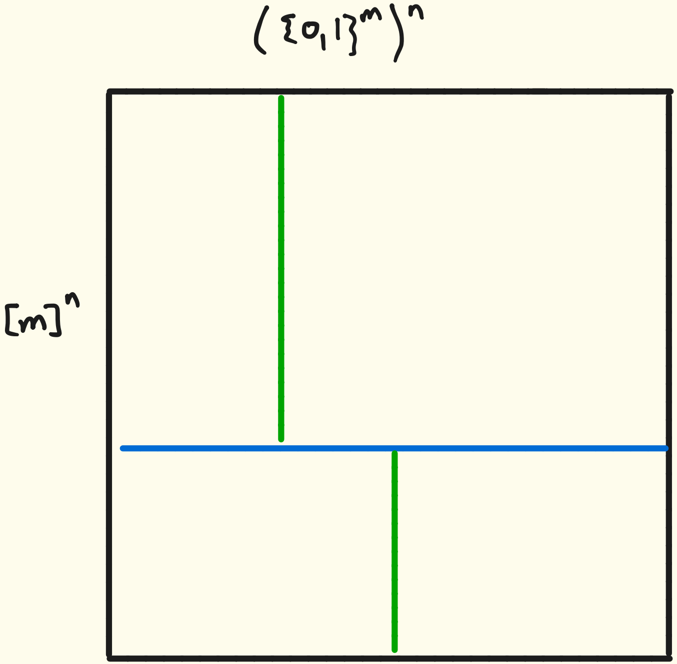
rows: all possible inputs for Alice
cols: " " " " Bob

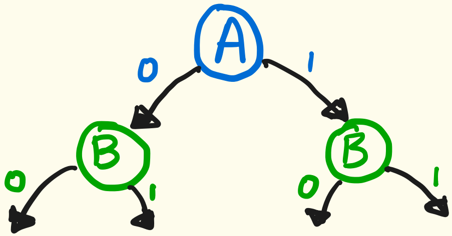
How to prove BPP LIFTING THEOREM ?



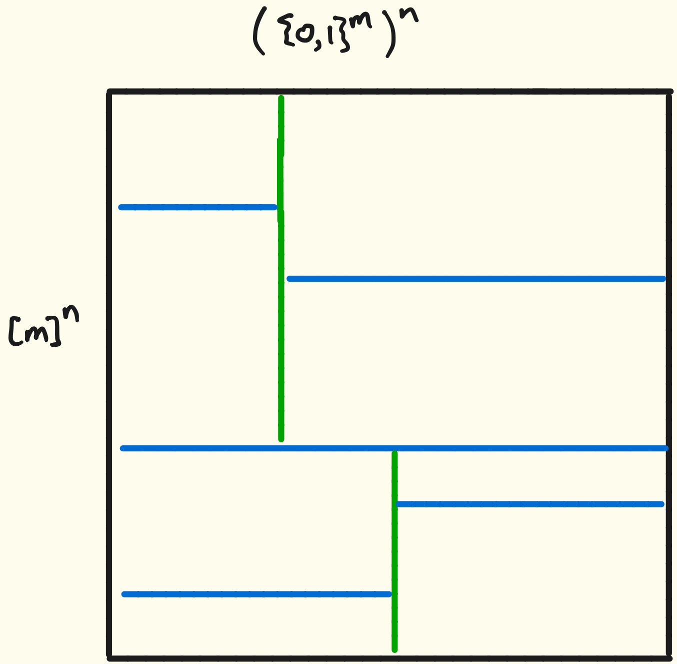


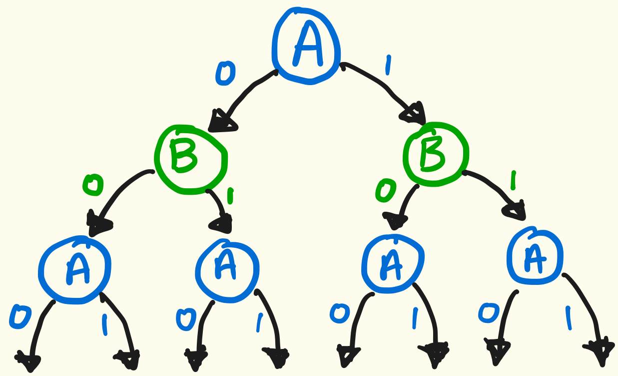
How to prove BPP LIFTING THEOREM ?



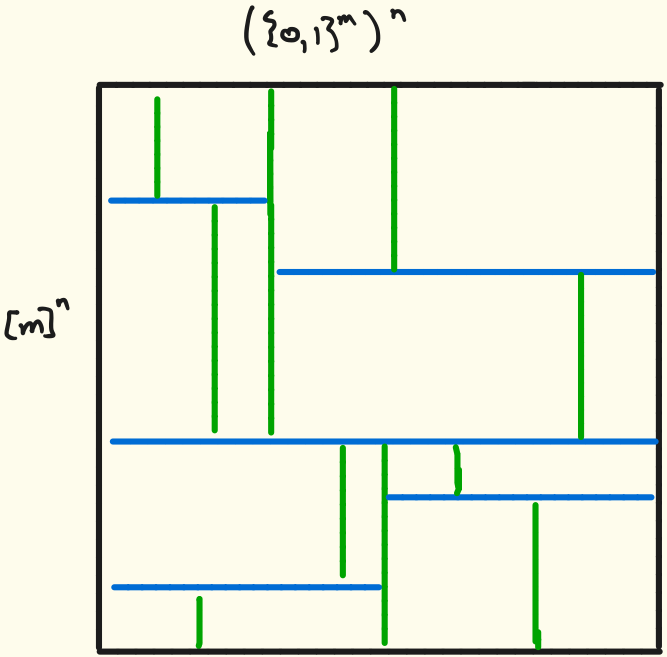


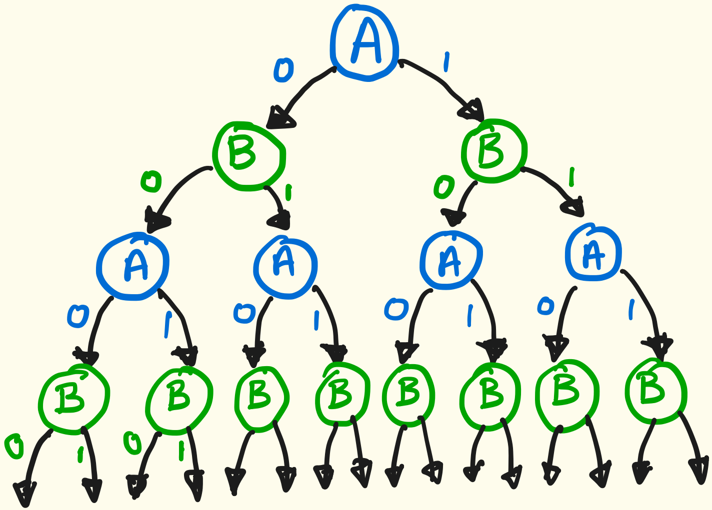
How to prove BPP LIFTING THEOREM ?



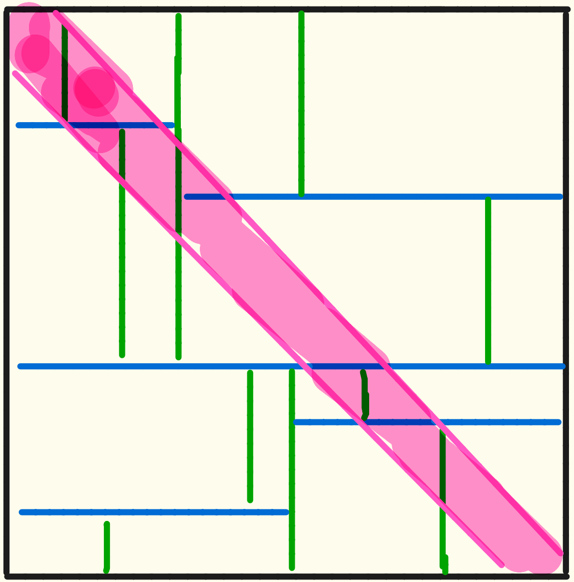


How to prove BPP LIFTING THEOREM ?





How to prove BPP LIFTING THEOREM ?

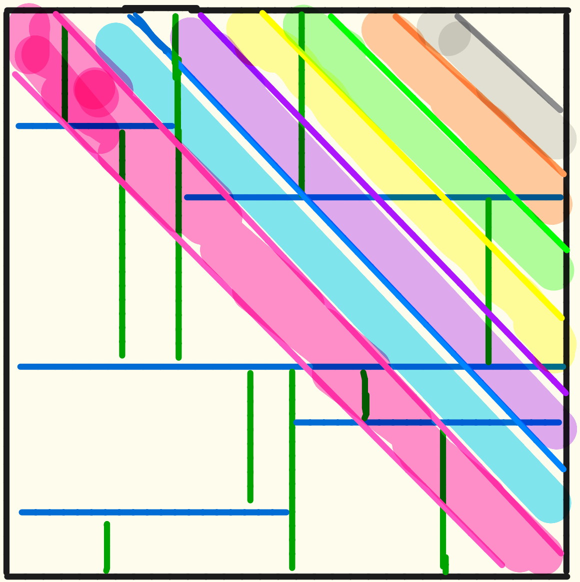


Let $z \in \{0,1\}^n$

The <u>z-slice</u> are the inputs $(x,y) \in (g^n)^{-1}(z)$

(the inputs to Alice/Bob consistent with z)

How to prove BPP LIFTING THEOREM?

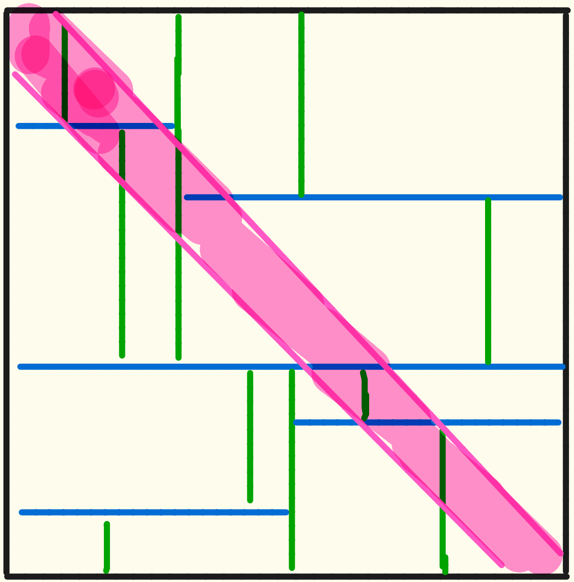


$[m]^n \times (\{0,1\}^m)^n$ partitioned into

2^n z-slices, one for each $z \in \{0,1\}^n$

SIMULATION

Π:

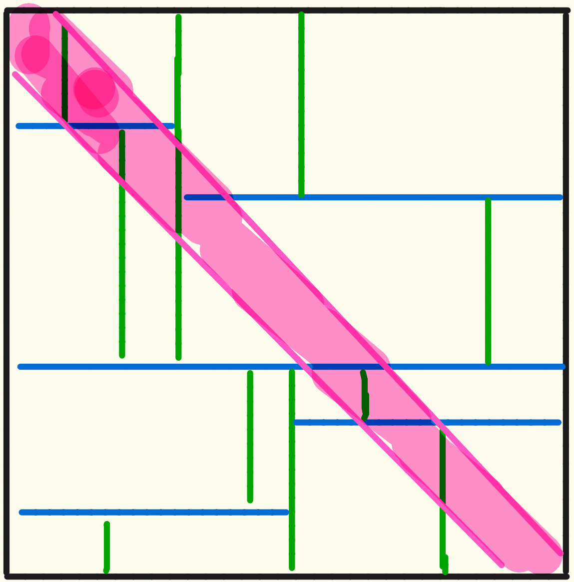


"Simulate Π" by a randomized decision tree $\mathcal{T}$

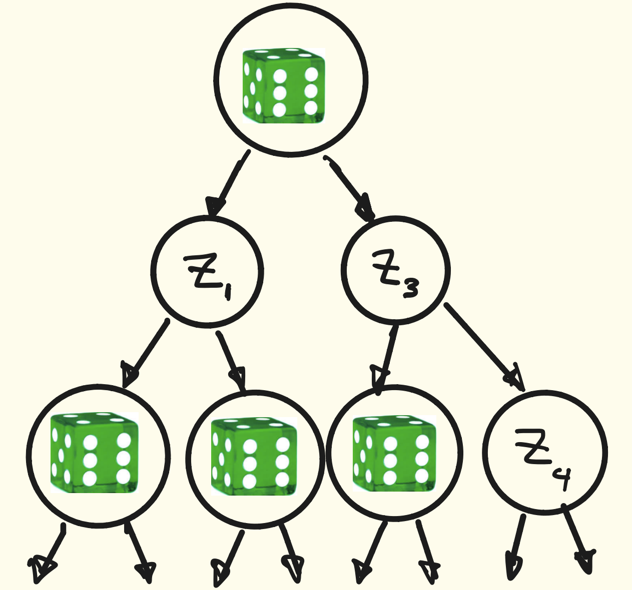
SIMULATION

"simulate Π" by a randomized decision tree $\mathcal{T}$

Π:



$\mathcal{T}$:



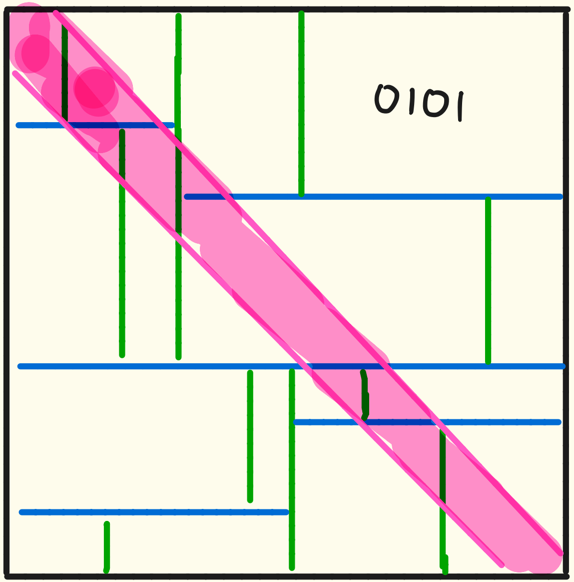
leaves of $\mathcal{T}$ labelled by transcripts (rectangles) of Π

GOAL $\forall z \in \{0,1\}^n$ $\boxed{1} \approx \boxed{2}$

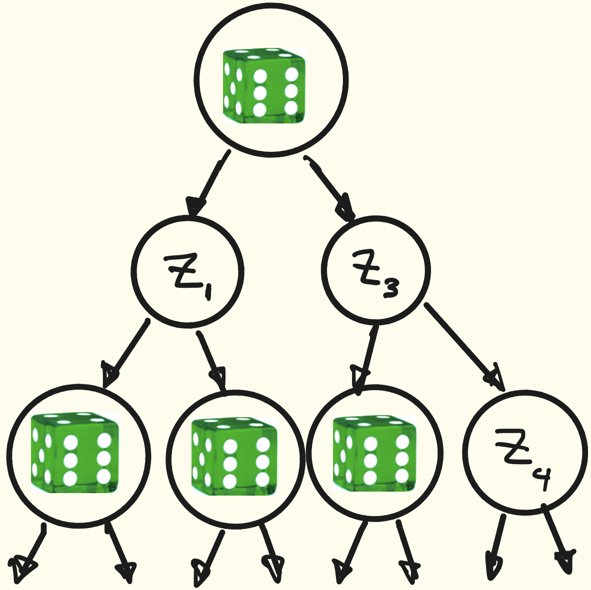
$\boxed{1}$ $\mathcal{T}(z)$ = output of randomized decision tree on z

$\boxed{2}$ Distribution over transcripts generated by Π on $(x,y) \sim (g^n)^{-1}(z)$

Π:



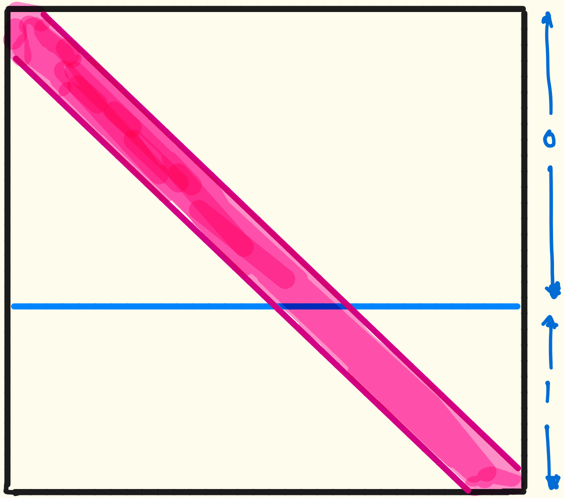
$\mathcal{T}$:

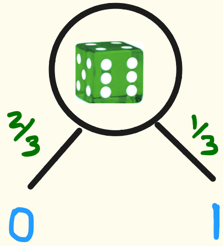


GOAL $\forall z$ $\boxed{1} \approx \boxed{2}$

$\boxed{1}$ $\tau(z)$

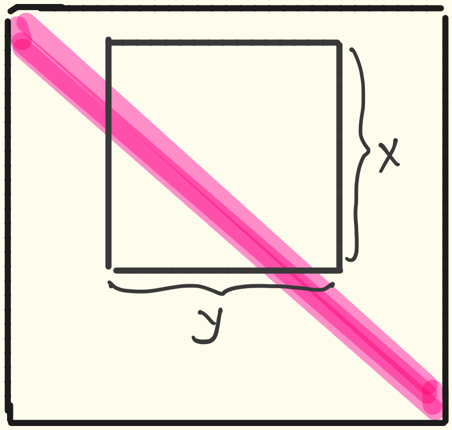
$\boxed{2}$ Distribution over transcripts generated by Π on $(x,y) \in (g^n)^{-1}(z)$





Idea:
Pretend marginals are uniform!

UNIFORM MARGINALS LEMMA



Let $X \subseteq [m]^n$ be DENSE

$Y \subseteq (\{0,1\}^m)^n$ be LARGE

Then $\forall z \in \{0,1\}^n$

$(x,y) \sim (g^n)^{-1}(z)$ has both marginals close to uniform on X and Y

DENSE: $H_\infty(X_I) \geq .9\,|I| \log m \;\; \forall I \subseteq [m]$

SIMULATION

I. If current X is DENSE AND Y IS LARGE:
simulate according to marginal probabilities

II Otherwise:

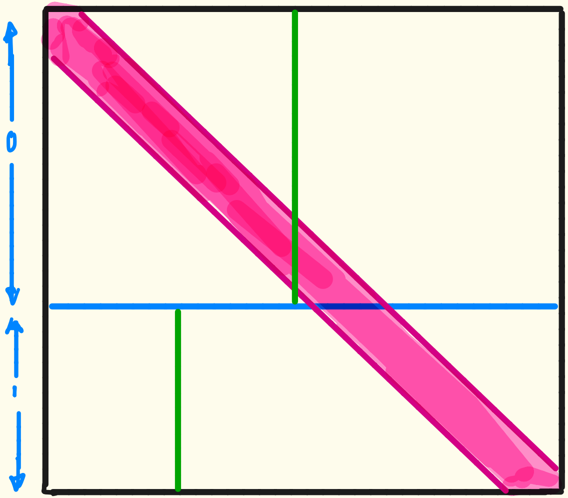
1. Compute partition $X = \bigcup_i X^i$ where each X^i is fixed on some $I \subseteq [n]$ and **dense** on $\bar{I}$
2. Update $X \leftarrow X^i$ with probability $|X^i|/|X|$
3. Query $z_I \in \{0,1\}^I$
4. Restrict Y so that $g^I(X_I, Y_I) = z_I$
5. Update $Y \leftarrow Y_{\bar{I}}$ and $X \leftarrow X_{\bar{I}}$ (which is **dense**)

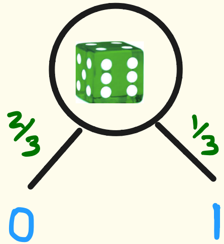
SIMULATION

⇒ I. If current X DENSE & Y LARGE, SIMULATE

II OTHERWISE:

1. Compute partition $X = \bigcup_i X^i$ where each X^i is fixed on some $I \subseteq [n]$ and **dense** on $\bar{I}$
2. Update $X \leftarrow X^i$ with probability $|X^i|/|X|$
3. Query $z_I \in \{0,1\}^I$
4. Restrict Y so that $g^I(X_I, Y_I) = z_I$
5. Update $Y \leftarrow Y_{\bar{I}}$ and $X \leftarrow X_{\bar{I}}$ (which is **dense**)



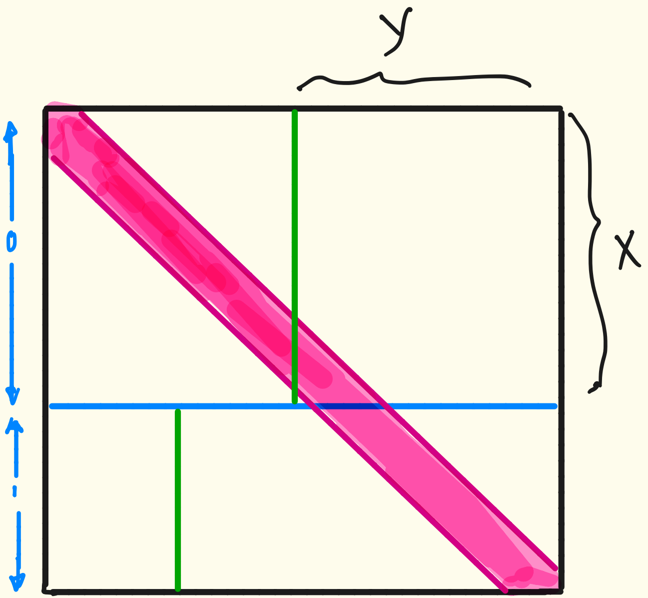


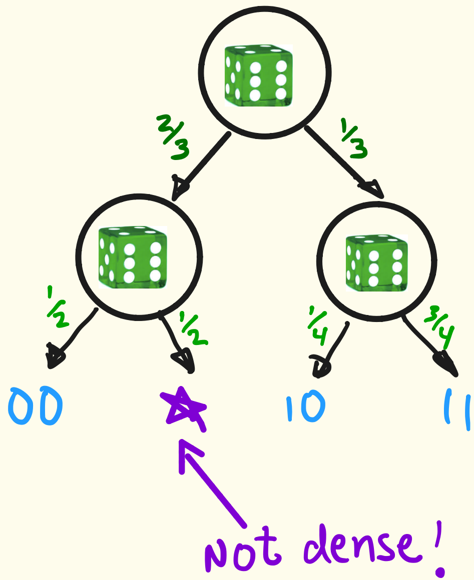
SIMULATION

I. If current x DENSE, y LARGE, SIMULATE

II. ELSE

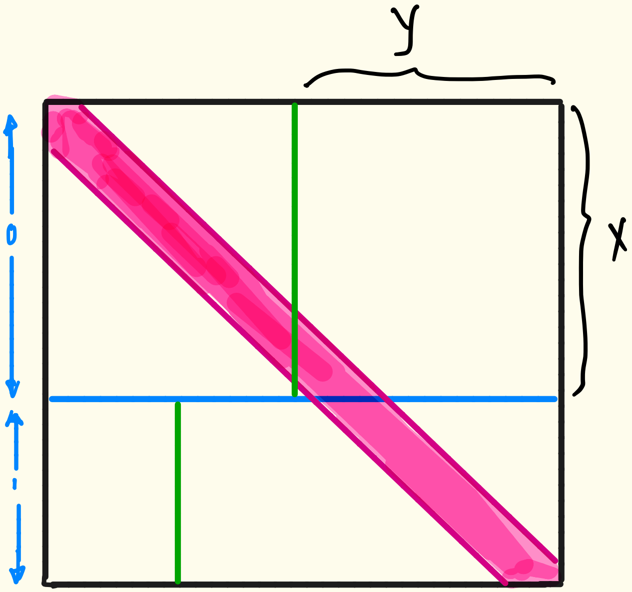
1. Compute partition $X = \bigcup_i X^i$ where each X^i is fixed on some $I \subseteq [n]$ and **dense** on $\bar{I}$
2. Update $X \leftarrow X^i$ with probability $|X^i|/|X|$
3. Query $z_I \in \{0,1\}^I$
4. Restrict Y so that $g^I(X_I, Y_I) = z_I$
5. Update $Y \leftarrow Y_{\bar{I}}$ and $X \leftarrow X_{\bar{I}}$ (which is **dense**)



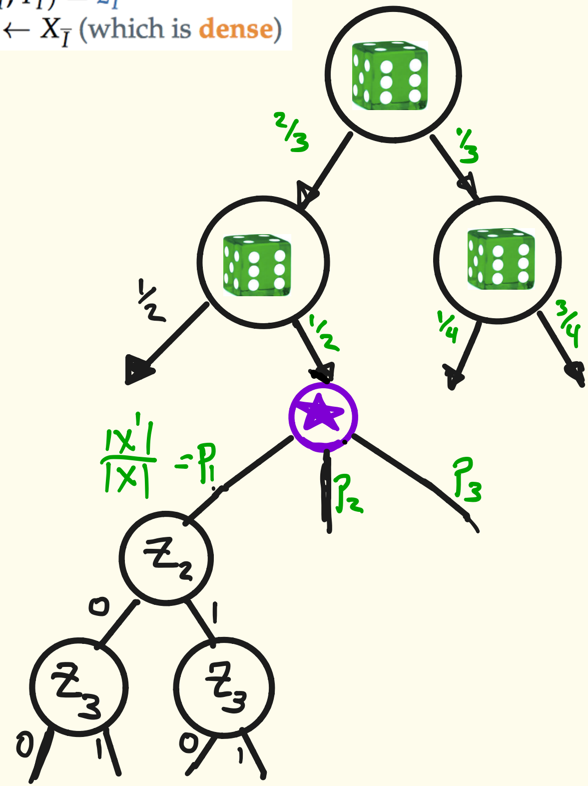


SIMULATION

1. Compute partition $X = \bigcup_i X^i$ where each X^i is fixed on some $I \subseteq [n]$ and **dense** on $\overline{I}$
2. Update $X \leftarrow X^i$ with probability $|X^i|/|X|$
3. Query $z_I \in \{0,1\}^I$
4. Restrict Y so that $g^I(X_I, Y_I) = z_I$
5. Update $Y \leftarrow Y_{\overline{I}}$ and $X \leftarrow X_{\overline{I}}$ (which is **dense**)



Say X' fixes coordinates 2,3



OPEN QUESTIONS

- Prove lifting theorems for information complexity
- Prove randomized Lifting theorem for constant-sized gadget g


Thanks!