Communication Complexity with Small Advantage

Thomas Watson

University of Memphis

Classic results

Classic results

$$
\begin{aligned}
\mathsf{R}(\textsc{Inner-Product}) &= \Theta(n) \\
\mathsf{R}(\textsc{Set-Intersection}) &= \Theta(n) \\
\mathsf{R}(\textsc{Gap-Hamming}) &= \Theta(n)
\end{aligned}
$$

Classic results

$$\begin{aligned}
\mathsf{R}(\textsc{Inner-Product}) &= \Theta(n)\\
\mathsf{R}(\textsc{Set-Intersection}) &= \Theta(n)\\
\mathsf{R}(\textsc{Gap-Hamming}) &= \Theta(n)
\end{aligned}$$

$\mathsf{R} =$ randomized c.c., success probability $\geq 3/4$

Classic results

$$\begin{aligned} \mathsf{R}(\textsc{Inner-Product}) &= \Theta(n) \\ \mathsf{R}(\textsc{Set-Intersection}) &= \Theta(n) \\ \mathsf{R}(\textsc{Gap-Hamming}) &= \Theta(n) \end{aligned}$$

R = randomized c.c., success probability $\geq 3/4$

Small advantage:
$\mathsf{R}_{1/2+\epsilon}$ = randomized c.c., success probability $\geq 1/2 + \epsilon$

Classic results — revisited

$$
\begin{aligned}
\mathsf{R}_{1/2+\epsilon}(\textsc{Inner-Product}) &= \Theta(n) \quad (\text{unless } \epsilon \leq 2^{-\Omega(n)}) \\
\mathsf{R}_{1/2+\epsilon}(\textsc{Set-Intersection}) &= \Theta(\epsilon \cdot n) \\
\mathsf{R}_{1/2+\epsilon}(\textsc{Gap-Hamming}) &= \Theta(\epsilon^2 \cdot n)
\end{aligned}
$$

Classic results — revisited

$$\begin{aligned}
\mathsf{R}_{1/2+\epsilon}(\textsc{Inner-Product}) &= \Theta(n) \quad (\text{unless } \epsilon \leq 2^{-\Omega(n)}) \\
\mathsf{R}_{1/2+\epsilon}(\textsc{Set-Intersection}) &= \Theta(\epsilon \cdot n) \\
\mathsf{R}_{1/2+\epsilon}(\textsc{Gap-Hamming}) &= \Theta(\epsilon^2 \cdot n)
\end{aligned}$$

$\vdots$

other functions?

Climbing the polynomial hierarchy

NP:

$\mathsf{R}_{1/2+\epsilon}(\textsc{Set-Intersection}) = \Theta(\epsilon \cdot n)$

Climbing the polynomial hierarchy

NP:

$\mathsf{R}_{1/2+\epsilon}(\textsc{Set-Intersection}) = \Theta(\epsilon \cdot n)$

[Braverman–Moitra STOC'13, Göös–Watson RANDOM'14]

(information complexity) (corruption)

Climbing the polynomial hierarchy

NP:

$\mathrm{R}_{1/2+\epsilon}(\textsc{Set-Intersection}) \;=\; \Theta(\epsilon \cdot n)$

[Braverman–Moitra STOC'13, Göös–Watson RANDOM'14]
(information complexity) (corruption)

Σ_2P, Π_2P:

Climbing the polynomial hierarchy

NP:

$R_{1/2+\epsilon}(\textsc{Set-Intersection}) = \Theta(\epsilon \cdot n)$

[Braverman–Moitra STOC'13, Göös–Watson RANDOM'14]
(information complexity) (corruption)

Σ_2P, Π_2P:

$R_{1/2+\epsilon}(\textsc{Tribes}) = \Theta(\epsilon \cdot n)$

Climbing the polynomial hierarchy

NP:

$R_{1/2+\epsilon}(\textsc{Set-Intersection}) = \Theta(\epsilon \cdot n)$

[Braverman–Moitra STOC'13, Göös–Watson RANDOM'14]
(information complexity) (corruption)

Σ_2P, Π_2P:

$R_{1/2+\epsilon}(\textsc{Tribes}) = \Theta(\epsilon \cdot n)$

Higher levels? (read-once AC^0 formulas)

Climbing the polynomial hierarchy

NP:

$\mathsf{R}_{1/2+\epsilon}(\textsc{Set-Intersection}) = \Theta(\epsilon \cdot n)$

[Braverman–Moitra STOC'13, Göös–Watson RANDOM'14]
(information complexity) (corruption)

Σ_2P, Π_2P:

$\mathsf{R}_{1/2+\epsilon}(\textsc{Tribes}) = \Theta(\epsilon \cdot n)$

Higher levels? (read-once AC^0 formulas)

Constant advantage: well-understood
[Jayram–Kopparty–Raghavendra/Leonardos–Saks CCC'09]

Climbing the polynomial hierarchy

NP:

$R_{1/2+\epsilon}(\textsc{Set-Intersection}) = \Theta(\epsilon \cdot n)$

[Braverman–Moitra STOC'13, (information complexity) Göös–Watson RANDOM'14] (corruption)

Σ_2P, Π_2P:

$R_{1/2+\epsilon}(\textsc{Tribes}) = \Theta(\epsilon \cdot n)$

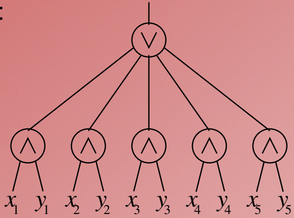
Higher levels? (read-once AC^0 formulas)

Constant advantage: well-understood
[Jayram–Kopparty–Raghavendra/Leonardos–Saks CCC'09]

Small advantage: open

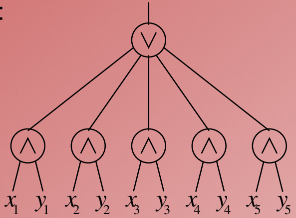
Function definitions

SET-INTERSECTION:

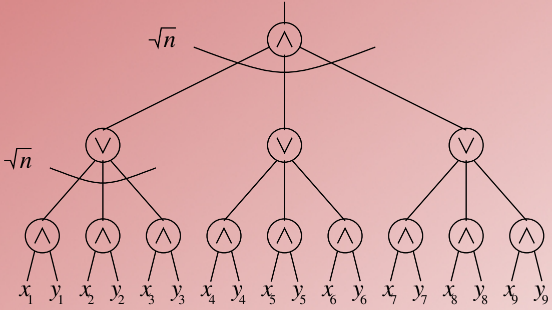


Function definitions

Set-Intersection:



Tribes:



What's known about TRIBES?

$$\mathrm{R}(\textsc{Tribes}) = \Theta(n)$$

What's known about TRIBES?

$\mathsf{R}(\textsc{Tribes}) = \Theta(n)$

[Jayram–Kumar–Sivakumar STOC'03, Harsha–Jain FSTTCS'13]
(information complexity) (smooth rectangle bound)

What's known about $\textsc{Tribes}$?

$\mathsf{R}(\textsc{Tribes}) \; = \; \Theta(n)$

[Jayram–Kumar–Sivakumar STOC'03, Harsha–Jain FSTTCS'13]
(information complexity) (smooth rectangle bound)

$R_{1/2+\epsilon}(\textsc{Tribes}) \; = \; ??$

What's known about TRIBES?

$\mathrm{R}(\textsc{Tribes}) \; = \; \Theta(n)$

[Jayram–Kumar–Sivakumar STOC'03, Harsha–Jain FSTTCS'13]
(information complexity) (smooth rectangle bound)

$R_{1/2+\epsilon}(\textsc{Tribes}) \; = \; ??$

[Göös–Watson] trick: $\mathrm{R}_{1/2+\epsilon} \; \geq \; \Omega(\epsilon \cdot \text{corruption bound})$

What's known about TRIBES?

$\mathsf{R}(\textsc{Tribes}) = \Theta(n)$

[Jayram–Kumar–Sivakumar STOC'03, Harsha–Jain FSTTCS'13]
(information complexity) (smooth rectangle bound)

$R_{1/2+\epsilon}(\textsc{Tribes}) = ??$

[Göös–Watson] trick: $\mathsf{R}_{1/2+\epsilon} \geq \Omega(\epsilon \cdot \text{corruption bound})$

- Doesn't work for TRIBES: corruption bound $\approx \sqrt{n}$

What's known about $\textsc{Tribes}$?

$\mathsf{R}(\textsc{Tribes}) \; = \; \Theta(n)$

[Jayram–Kumar–Sivakumar STOC'03, Harsha–Jain FSTTCS'13]
(information complexity) (smooth rectangle bound)

$R_{1/2+\epsilon}(\textsc{Tribes}) \; = \; ??$

[Göös–Watson] trick: $\mathsf{R}_{1/2+\epsilon} \; \geq \; \Omega(\epsilon \cdot \text{corruption bound})$

- Doesn't work for $\textsc{Tribes}$: corruption bound $\approx \sqrt{n}$

?? Similar trick: $\mathsf{R}_{1/2+\epsilon} \; \geq \; \Omega(\epsilon \cdot \text{smooth rectangle bound})$??

What's known about $\textsc{Tribes}$?

$\mathsf{R}(\textsc{Tribes}) = \Theta(n)$

[Jayram–Kumar–Sivakumar STOC'03, Harsha–Jain FSTTCS'13]
(information complexity) (smooth rectangle bound)

$R_{1/2+\epsilon}(\textsc{Tribes}) = ??$

[Göös–Watson] trick: $\mathsf{R}_{1/2+\epsilon} \geq \Omega(\epsilon \cdot \text{corruption bound})$

- Doesn't work for $\textsc{Tribes}$: corruption bound $\approx \sqrt{n}$

?? Similar trick: $\mathsf{R}_{1/2+\epsilon} \geq \Omega(\epsilon \cdot \text{smooth rectangle bound})$??

- Fails in general ($\textsc{Gap-Hamming}$)

Our approach for TRIBES

Information complexity:

Our approach for TRIBES

Information complexity:

- $\Omega(1)$-advantage for TRIBES [JKS'03]

Our approach for TRIBES

Information complexity:

- $\Omega(1)$-advantage for TRIBES [JKS'03]
- ϵ-advantage for SET-INTER [BM'13]

Our approach for TRIBES

Information complexity:

- $\Omega(1)$-advantage for TRIBES [JKS'03]
- ϵ-advantage for SET-INTER [BM'13]
- Combine?

Our approach for TRIBES

Information complexity:

- $\Omega(1)$-advantage for TRIBES [JKS'03]
- ϵ-advantage for SET-INTER [BM'13]
- Combine?

4-step approach:

Our approach for TRIBES

Information complexity:

- $\Omega(1)$-advantage for TRIBES [JKS'03]
- ϵ-advantage for SET-INTER [BM'13]
- Combine?

4-step approach:

1. Conditioning and direct sum
2. Uniformly covering a pair of gadgets
3. Relating information and probabilities for inputs
4. Relating information and probabilities for transcripts

Preliminaries

Idea from [BM'13]:

Preliminaries

Idea from [BM'13]:

Suffices to use $3\textsc{Eq}$ gadget instead of $\textsc{And}$ gadget

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Preliminaries

Idea from [BM'13]:

Suffices to use $3\textsc{Eq}$ gadget instead of $\textsc{And}$ gadget

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Usual info complexity proofs:

- mutual info $\rightarrow$ Hellinger distance $\rightarrow$ statistical distance

Preliminaries

Idea from [BM'13]:

Suffices to use 3EQ gadget instead of AND gadget

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Usual info complexity proofs:

- mutual info $\rightarrow$ Hellinger distance $\rightarrow$ statistical distance
- quadratic loss—very roughly:

Preliminaries

Idea from [BM'13]:

Suffices to use $3\textsc{Eq}$ gadget $\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$ instead of $\textsc{And}$ gadget $\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$

Usual info complexity proofs:

- mutual info $\to$ Hellinger distance $\to$ statistical distance
- quadratic loss—very roughly:
  - info cost: quadratic terms (in small parameters)

Preliminaries

Idea from [BM'13]:

Suffices to use $3\textsc{Eq}$ gadget instead of $\textsc{And}$ gadget

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Usual info complexity proofs:

- mutual info $\to$ Hellinger distance $\to$ statistical distance
- quadratic loss—very roughly:
  - info cost: quadratic terms (in small parameters)
  - probabilities: quadratic and linear terms

Preliminaries

Idea from [BM'13]:

Suffices to use $3\textsc{Eq}$ gadget instead of $\textsc{And}$ gadget

$$\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} \qquad \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$$

Usual info complexity proofs:

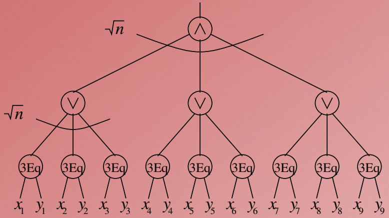
- mutual info $\rightarrow$ Hellinger distance $\rightarrow$ statistical distance
- quadratic loss—very roughly:
  - info cost: quadratic terms (in small parameters)
  - probabilities: quadratic and linear terms

$3\textsc{Eq}$ has nice symmetry properties

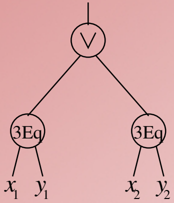
- Exploit to get linear terms to perfectly cancel

1. Conditioning and direct sum

info cost $o(\epsilon \cdot n)$ $\rightsquigarrow$ info cost $o(\epsilon)$

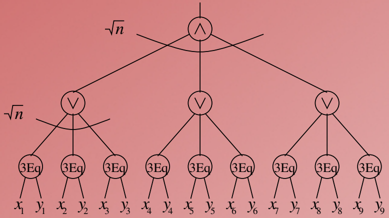


$\rightsquigarrow$

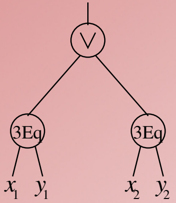


1. Conditioning and direct sum

info cost $o(\epsilon \cdot n)$ $\rightsquigarrow$ info cost $o(\epsilon)$



$\rightsquigarrow$



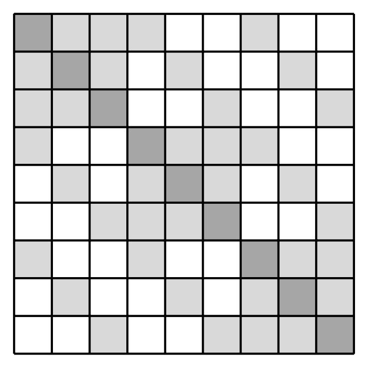
Want to show:

advantage $\leq O$(info cost)

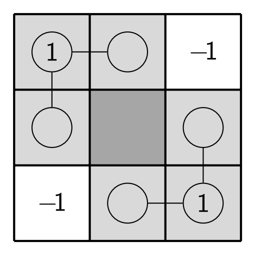
2. Uniformly covering a pair of gadgets

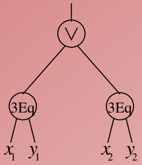
2. Uniformly covering a pair of gadgets

Uniformly cover



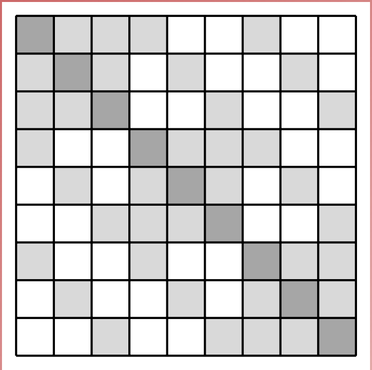
with



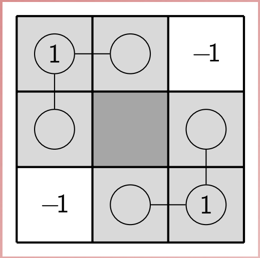


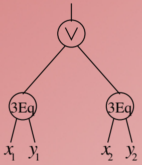
2. Uniformly covering a pair of gadgets

Uniformly cover



with

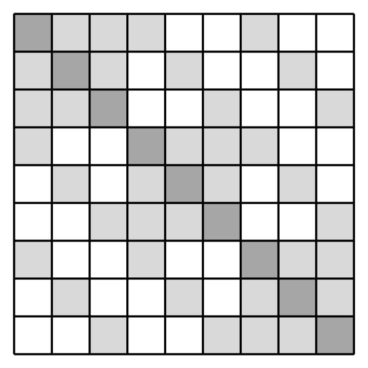




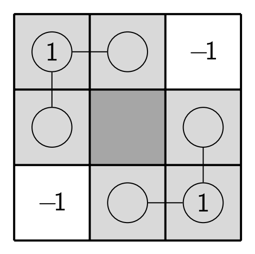
Lemma: Linear combination of acceptance probabilities $\leq O(\sum$ four contributions to info cost$)$

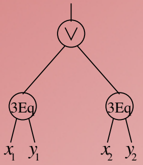
2. Uniformly covering a pair of gadgets

Uniformly cover



with



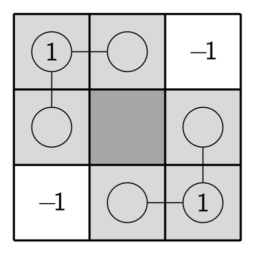


Lemma: Linear combination of acceptance probabilities
$\leq O(\sum$ four contributions to info cost$)$

$\Downarrow$

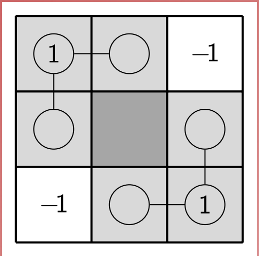
adv $\leq$ $\left(\sum_{\text{light gray}} \text{acc prob}\right) - \left(\sum_{\text{white}} \text{acc prob}\right)$
$\leq O(\text{info cost})$

3. Relating information and probabilities for inputs



Lemma: Linear combination of acceptance probabilities
$\leq O(\sum$ four contributions to info cost)

3. Relating information and probabilities for inputs

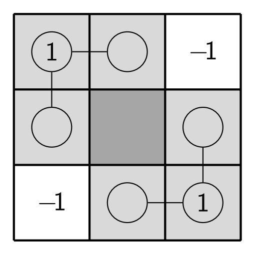


Lemma: Linear combination of acceptance probabilities
$\leq O(\sum$ four contributions to info cost)

Prove for individual transcripts?

contribution to lin comb of acc prob $\leq$ O(contribution to info costs)

3. Relating information and probabilities for inputs



Lemma: Linear combination of acceptance probabilities
$\leq O(\sum$ four contributions to info cost$)$

Prove for individual transcripts?

contribution to lin comb of acc prob $\leq$ $O($contribution to info costs$)$

[BM'13] setting: yes

Our setting: ....

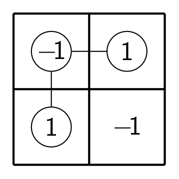
3. Relating information and probabilities for inputs

[BM'13] transcript lemma:

Our transcript lemma:

3. Relating information and probabilities for inputs

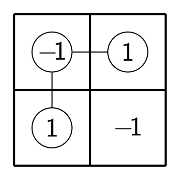
[BM'13] transcript lemma:



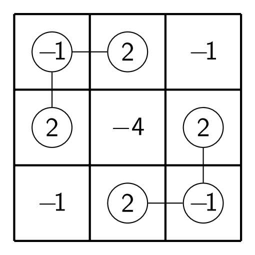
Our transcript lemma:

3. Relating information and probabilities for inputs

[BM'13] transcript lemma:

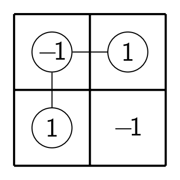


Our transcript lemma:

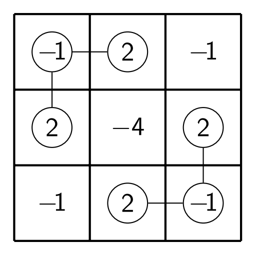


3. Relating information and probabilities for inputs

[BM'13] transcript lemma:



Our transcript lemma:

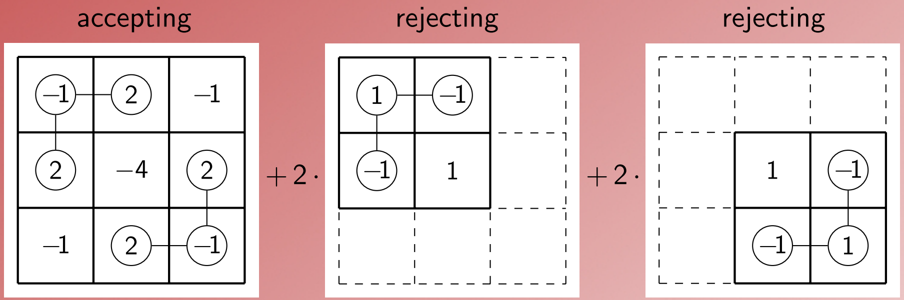


∀ transcript:

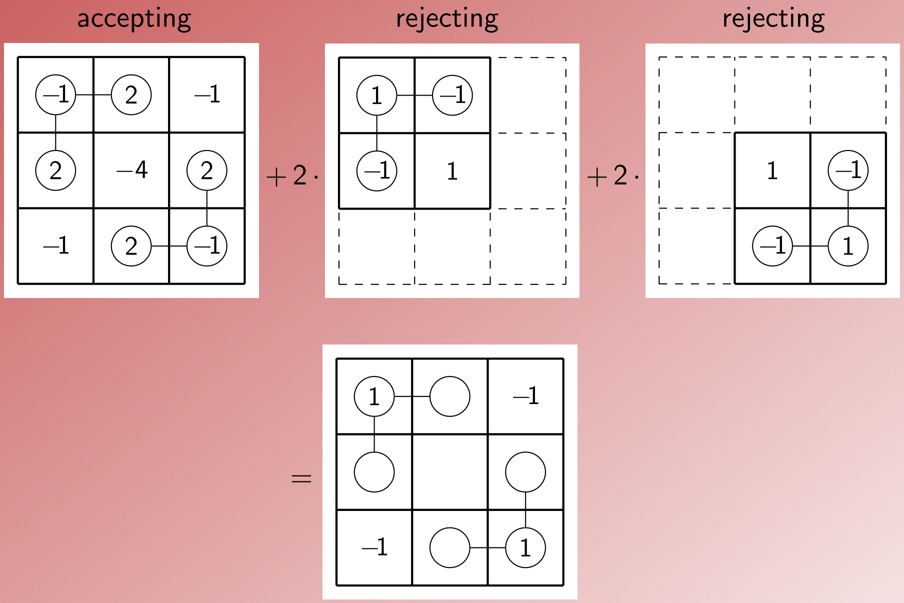
contribution to lin comb of prob $\leq$ O(contribution to info costs)

3. Relating information and probabilities for inputs

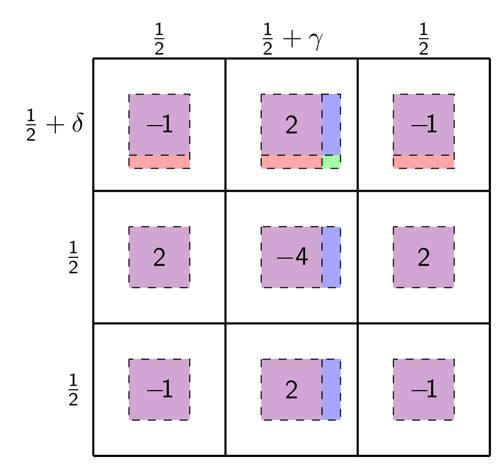
3. Relating information and probabilities for inputs



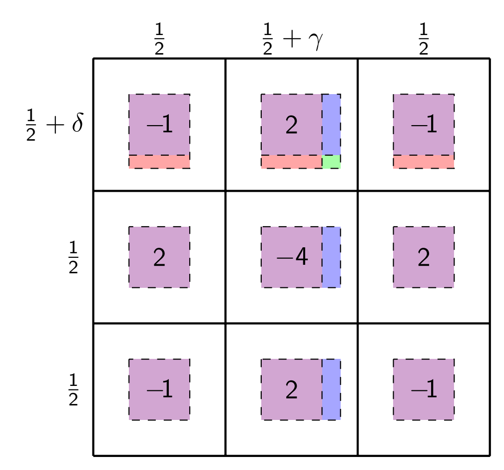
3. Relating information and probabilities for inputs



4. Relating information and probabilities for transcripts

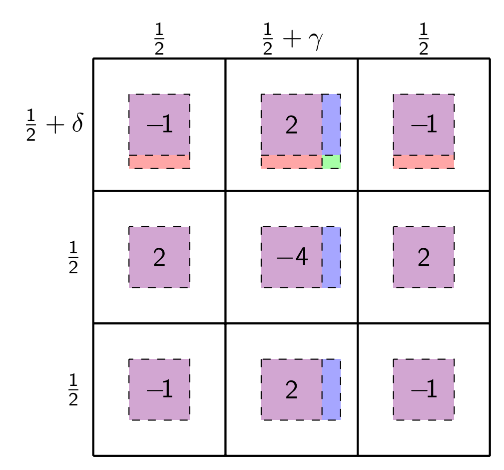


4. Relating information and probabilities for transcripts



lin comb of probabilities
$= 2 \cdot$ green area
$= \Theta(\delta\gamma)$

4. Relating information and probabilities for transcripts

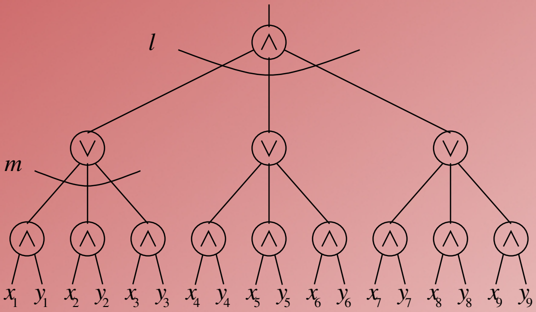


lin comb of probabilities
$= 2 \cdot$ green area
$= \Theta(\delta\gamma)$

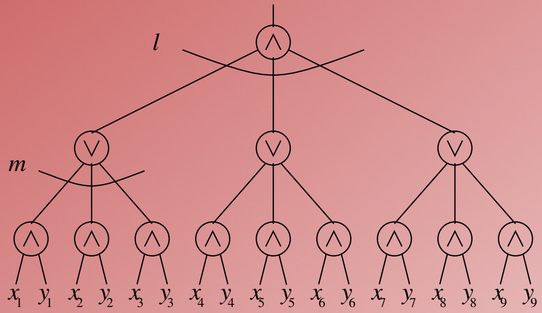
$\leq$

contribution to info costs
$= \Theta(\delta^2 + \gamma^2)$

Generalized TRIBES

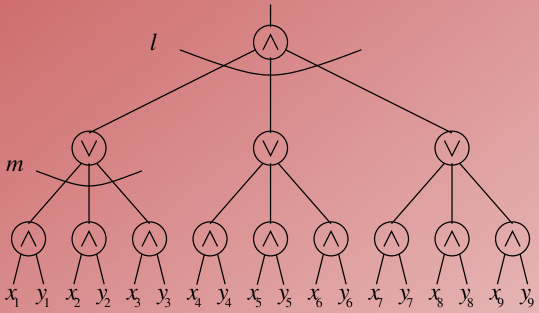


Generalized TRIBES



$\Omega(\epsilon \cdot \ell m)$: still holds

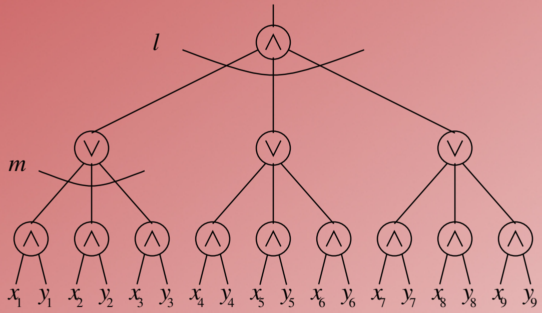
Generalized TRIBES



$\Omega(\epsilon \cdot \ell m)$: still holds

$O(\epsilon \cdot \ell m)$: if $\epsilon \geq \Omega(1/\ell)$

Generalized TRIBES

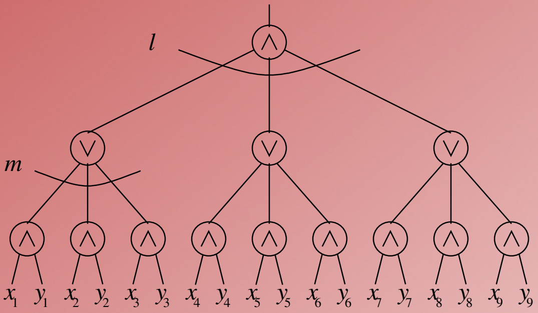


$\Omega(\epsilon \cdot \ell m)$: still holds

$O(\epsilon \cdot \ell m)$: if $\epsilon \geq \Omega(1/\ell)$

What if $\epsilon \leq o(1/\ell)$?

Generalized Tribes



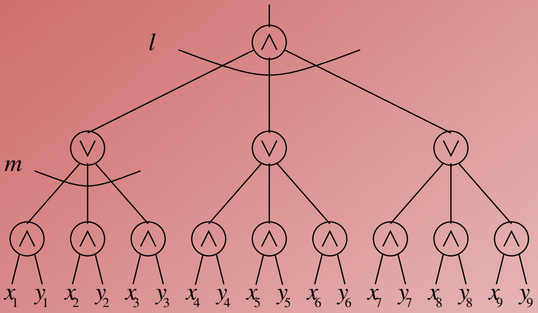
$\Omega(\epsilon \cdot \ell m)$: still holds

$O(\epsilon \cdot \ell m)$: if $\epsilon \geq \Omega(1/\ell)$

What if $\epsilon \leq o(1/\ell)$?

$\ell = 2$: $O(\sqrt{\epsilon} \cdot m)$

Generalized TRIBES



$\Omega(\epsilon \cdot \ell m)$: still holds

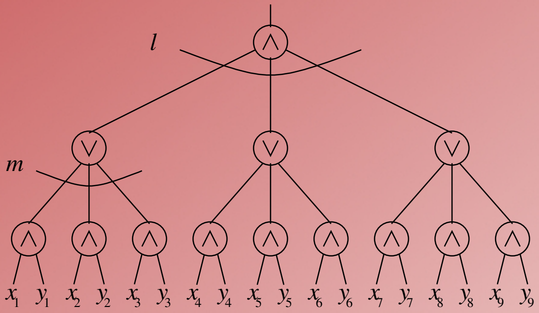
$O(\epsilon \cdot \ell m)$: if $\epsilon \geq \Omega(1/\ell)$

What if $\epsilon \leq o(1/\ell)$?

$\ell = 2$: $O(\sqrt{\epsilon} \cdot m)$

$\Omega(\sqrt{\epsilon} \cdot m)$
for decision trees

Generalized Tribes



$\Omega(\epsilon \cdot \ell m)$: still holds

$O(\epsilon \cdot \ell m)$: if $\epsilon \geq \Omega(1/\ell)$

What if $\epsilon \leq o(1/\ell)$?

$\ell = 2$: $O(\sqrt{\epsilon} \cdot m)$

$\Omega(\sqrt{\epsilon} \cdot m)$
for decision trees

open:
$\Omega(\sqrt{\epsilon} \cdot m)$
for communication?

Which part contains the intersecting coordinate?

Which part contains the intersecting coordinate?

Inputs: Uniquely intersecting subsets of $[n]$
$[n]$ is partitioned into ℓ equal-size parts
Output: Which part contains the intersection?

Which part contains the intersecting coordinate?

Inputs: Uniquely intersecting subsets of $[n]$
$[n]$ is partitioned into ℓ equal-size parts
Output: Which part contains the intersection?

$\mathsf{R}_{1/\ell+\epsilon}(\text{this problem}) = \Theta(\epsilon \cdot n)$

Which part contains the intersecting coordinate?

Inputs: Uniquely intersecting subsets of $[n]$
$[n]$ is partitioned into ℓ equal-size parts
Output: Which part contains the intersection?

$R_{1/\ell+\epsilon}$(this problem) $= \Theta(\epsilon \cdot n)$

Proof: Combine [BM'13], and direct sum for info complexity under promise that exactly one input evaluates to 1

Which part contains the intersecting coordinate?

Inputs: Uniquely intersecting subsets of $[n]$
$[n]$ is partitioned into ℓ equal-size parts
Output: Which part contains the intersection?

$\mathsf{R}_{1/\ell+\epsilon}(\text{this problem}) \;=\; \Theta(\epsilon \cdot n)$

Proof: Combine [BM'13], and direct sum for info complexity under promise that exactly one input evaluates to 1

Also: Simplified proof of $\mathsf{UP} \cap \mathsf{coUP} \not\subseteq \mathsf{BPP}$ [Klauck, CCC'03]

The end