Applications of Information Complexity II

David Woodruff
IBM Almaden

Outline

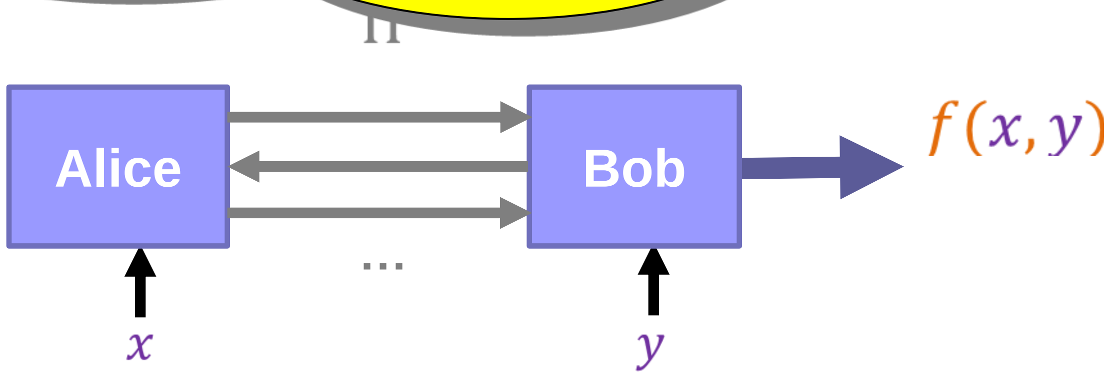
New Types of Direct Sum Theorems

1. Direct Sum with Aborts

2. Direct Sum of Compositions

- $D_{\mu,\delta}(f)$ = minimum communication over all correct protocols Π

Can't we fix the private randomness?



- Distribution μ on inputs (x, y)
- Correctness: $\Pr_{(X,Y) \sim \mu, \text{ private randomness}} [\Pi(X,Y) = f(X,Y)] \geq 1-\delta$
- Communication: $\max_{x,y, \text{ private randomness}} |\Pi(x,y)|$

Distributional Communication Complexity vs. Information Complexity

- By averaging, there is a good fixing of the randomness:

$$D_{\mu,\delta}(f) = \min_{\text{correct deterministic } \Pi} \max_{x,y} |\Pi(x,y)|$$

- However, we'll use the notion of information complexity:

$$IC^{ext}_{\mu,\delta}(f) = \min_{\text{correct } \Pi} I(X, Y ; \Pi) = H(X, Y) - H(X, Y \mid \Pi)$$

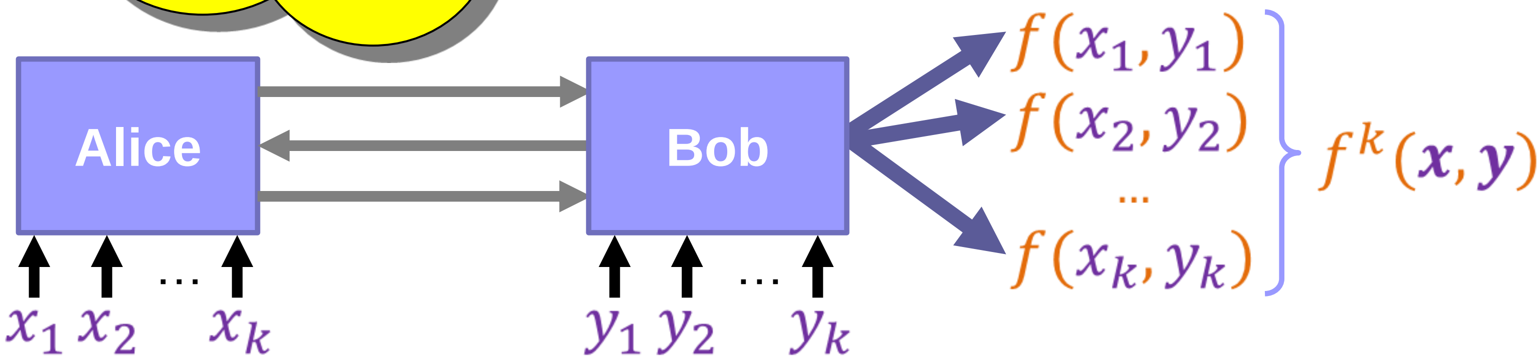
Can't fix private randomness of Π and preserve $I(X, Y ; \Pi)$

Input X
Randomness R

$X \oplus R$ →

Mu… a
Co… roblem

Require that the protocol solves all k instances



- Distribution μ^k on inputs $(x,y) = (x_1, y_1), \ldots, (x_k, y_k)$
- Correctness: $\Pr_{(X,Y) \sim \mu^k, \text{ private randomness}} [\Pi(X,Y) = f^k(X,Y)] \geq 1-\delta$
- $D_{\mu^k, \delta}(f^k) = \min_{\text{correct } \Pi} \max_{x,y, \text{ private randomness}} |\Pi(x,y)|$

How Hard is Solving all k Copies?

Main question: Is solving instances independently the best we can do?

$$D_{\mu^k,\delta}(f^k) \overset{?}{\geq} \Omega(k).D_{\mu,\frac{\delta}{k}}(f)$$

- **Direct sum theorems**
  - $D_{\mu^k,\delta}(f^k) \geq \Omega(\sqrt{k}).D_{\mu,\delta}(f)$ [BBCR 10]
  - $D^r_{\mu^k,\delta}(f^k) \geq \Omega(k).\left(D^r_{\mu,\delta}(f) - r - \sqrt{r.D^r_{\mu,\delta}(f)}\right)$ [BR 11]
  - ...

- **Direct product theorems**
  - $D_{\mu^k,1-\left(1-\frac{1}{3}\right)^k}(f^k) \geq \Omega(\sqrt{k}).D_{\mu,\frac{1}{3}}(f)$ [BRWY]

Jain et al (bounded rounds)

How Hard is Solving all k Copies?

Main question: Is solving instances independently the best we can do?

$$D_{\mu^k,\delta}(f^k) \overset{?}{\geq} \Omega(k).D_{\mu,\frac{\delta}{k}}(f)$$

- Direct
  - [BCR 10]
  - [BR 11]
  - ...

None attains above bound!

Impossible for general problems [FKNN95]

However, general but relaxed statement (with aborts) is true for information complexity

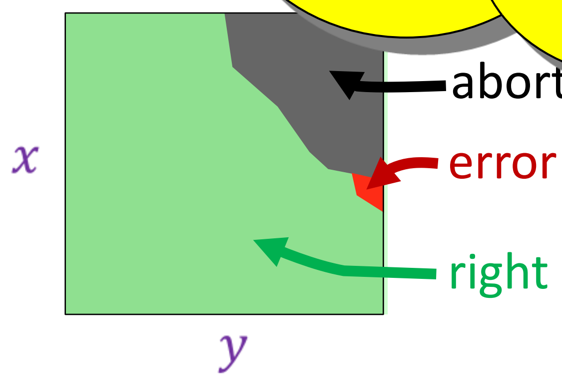
Protocols with A

Definition: A randomized

with probability 1-α over

- (Abort) $\Pr_{(X,Y) \sim \mu}$ [Π
- (Error) $\Pr_{(X,Y) \sim \mu}$ [Π($] \leq \delta$

We fix $\alpha = 1/20$ and write $IC^{ext}_{\mu, 1/20, \beta, \delta}(f) = IC^{ext}_{\mu, \beta, \delta}(f)$



- When protocol aborts, it "knows it is wrong"

- $IC^{ext}_{\mu, \alpha, \beta, \delta}(f) = \min I(X, Y ; \Pi)$, over Π that (μ, α, β, δ)-compute f

Direct Sum with Aborts

Main result: Stronger direct sum for every communication problem via protocols with abortion

solving k instances with error δ
is as hard as
solving each instance with error $\frac{\delta}{k}$ and constant **abortion**

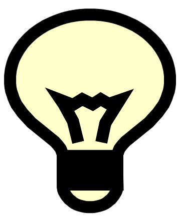
- Formally, $IC^{ext}_{\mu^k, \delta}(f^k) = \Omega(k) \cdot IC^{ext}_{\mu, 1/10, \delta/k}(f)$

- Number r of communication rounds is preserved
  $IC^{ext, r}_{\mu^k, \delta}(f^k) = \Omega(k) \cdot IC^{ext, r}_{\mu, 1/10, \delta/k}(f)$

Proof Idea

- $I(X^1, Y^1, \ldots, X^k, Y^k ; \Pi) = \sum_i I(X^i, Y^i ; \Pi \mid X^{<i}, Y^{<i})$

$$= \sum_i \sum_{x,y} I(X^i, Y^i ; \Pi \mid X^{<i} = x, Y^{<i} = y) \cdot \Pr[X^{<i} = x, Y^{<i} = y]$$



B…)

$$1 - \delta \leq \Pr(\text{all } k \text{ correct})$$

$$= \prod_{i=1..k} \Pr(i \text{ correct} \mid \text{correct up to } i-1)$$

Create…

- R… ,B), and…

- *Check* if Π's output is correct on first i-1 instances

- If correct, then $\Pi_{x,y}$ (A,B) outputs the i-th output of Π

- Else, Abort

Application: Direct Sum for Equality

- For strings x and y, $EQ(x,y) = 1$ if $x = y$, else $EQ(x,y) = 0$

$x_1, \ldots, x_k \in \{0,1\}^n$   $y_1, \ldots, y_k \in \{0,1\}^n$

- $EQ^k = (EQ(x_1, y_1), EQ(x_2, y_2), \ldots, EQ(x_k, y_k))$

- Standard direct sum:

  $IC^{ext,1}_{\mu^k, 1/3}(EQ^k) = \Omega(k) \cdot IC^{ext,1}_{\mu, 1/3}(EQ)$

  For any μ, $IC^{ext,1}_{\mu, 1/3}(EQ) = O(1)$, so LB is $\Omega(k)$

- Direct sum with aborts:

  $IC^{ext,1}_{\mu^k, 1/3}(EQ^k) = \Omega(k) \cdot IC^{ext,1}_{\mu, 1/10, 1/(3k)}(EQ) = \Omega(k \log k)$

Sketching Applications

Optimal lower bounds (improve by a log k factor)

- Sketching a sequence $u_1, \ldots, u_k$ of vectors, and sequence $v_1, \ldots, v_k$ of vectors in a stream to $(1+\varepsilon)$-approximate all distances $|u_i - v_j|_p$

- Sketching matrices A and B in a stream so that for all i, j, $(A \cdot B)_{i,j}$ can be approximated with additive error $\varepsilon |A_i|*|B_j|$

Set Intersection Application

$S \subseteq [n]$
$|S| = k$

$T \subseteq [n]$
$|T| = k$

Each party should locally output $S \cap T$

- Randomized protocol with $O(k)$ bits of communication.

- In $O(r)$ rounds, obtain $O(k \text{ ilog}^{(r)} k)$ communication [WY]
    - $\text{ilog}^{(1)} k = \log k$, $\text{ilog}^{(2)} k = \log \log k$, etc.

- Combining [BCK] and direct sum with aborts, any r-round protocol w. pr. $\geq 2/3$ requires $\Omega(k \text{ ilog}^{(r)} k)$ communication

Outline

New Types of Direct Sum Theorems

1. Direct Sum with Aborts

2. Direct Sum of Compositions

Composing Functions

- 2-Party Communication Complexity
  - Alice has input x. Bob has input y

- Consider $x = (x^1, \ldots, x^r)$ and $y = (y^1, \ldots, y^r) \in (\{0,1\}^n)^r$

- $f(x,y) = h(g(x^1, y^1), \ldots, g(x^r, y^r))$ for Boolean function g

h Outer function

Inner function

$g(x^1, y^1)$ $g(x^2, y^2)$ $\ldots$ $g(x^{r-1}, y^{r-1})$ $g(x^r, y^r)$

Given information complexity lower bounds for h and g, when is there an information complexity lower bound for f?

Composing Functions

- ALL-COPIES = (g(x^1, y^1), …, g(

- DISJ(x,y) = $\vee_{i=1}^{r} (x^i \wedge y^i)$

- TRIBES(x,y) = $\wedge_{i=1}^{r}$ DISJ(x^i, y^i)

- Key to information complexity lower bo… is an *embedding step*
  - Lower bound I(X, Y ; Π) = Σ_i I(X… protocol for each i to solve the…
  - Lower bound I(X^i, Y^i ; Π | $X^{<i}$, $Y^{<i}$) b… y of g
  - Combining function h needs to be sensitive to individual coordinates (under appropriate distribution)

OR function *sensitive* to individual coordinates

AND function *sensitive* to individual instances of DISJ

Composing Functions

- What if outer function h is not sensitive to individual coordinates?

-

Can we use
IC^{ext}(Gap-Thresh(AND)) to bound
IC^{ext}(Gap-Thresh(g)) for other
functions g?

- No obvious embeddi... p-Thresh!

- For specific choices of inner function [BGPW, CR]:

  $IC^{ext}_{\mu, \delta}(\text{Gap-Thresh}(a^1 \wedge b^1, \ldots, a^r \wedge b^r)) = \Omega(r)$ for μ a product uniform distribution on $a=(a^1, \ldots, a^r)$, $b=(b^1, \ldots, b^r)$

E

- On other coordinates j, choose a random player $P^i_j \in$ {Alice, Bob}
- If P^i_j = Alice, $X^i_j \in_R \{0,1\}$, $Y^i_j = 0$
- If P^i_j = Bob, $X^i_j = 0$, $Y^i_j \in_R \{0,1\}$

Call this distribution μ

Say μ is collapsing

- Kno ...)), distribution

What if we could embed

so that inner function g(

Embed into random special coordinate S^i

- Example embedding for g = DISJ

0	1	0	a_i	0	0	0	0	0	1

0	0	0	b_i	1	1	0	0	1	0

Analyzing th...

- Let Π be protoco... ...)))
- By ch...
- Maximum likelih... $B^{<i}$, **P, S,** there is a predictor θ w... Ω(1)
  - Holds for ...
- I(Π ; X^1,...,X... | $X^{<i}$, $Y^{<i}$, **P, S**)
  ... | $X^{<i}$,$Y^{<i}$,$A^{<i}$,$B^{<i}$, **P, S**)
  $\geq \Sigma_{i=1}$ I(Π ; X^i, Y^i | $A^{<i}$, $B^{<i}$, **P, S**)

..., B^r | **P, S**)

...(AND)) = Ω(r)

$X^{<i}$, $Y^{<i}$ determines $A^{<i}$, $B^{<i}$ given **P, S**

Chain rule

- Ind... ..., $Y^{<i}$
- Conditioning ... entropy

Now let's look at the information cost

Normally we would look at information cost.

Here we look at an intermediate measure

Guessing Game

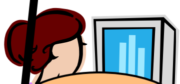
Protocol Ψ is correct if can guess (A,B) w. pr. $1/4 + \Omega(1)$ given Ψ(U,V), S, P

CIC^{ext}(Guessing Game) = $\min_{\text{correct } \Psi} I(\Pi ; U, V \mid S, P)$

- I(
- Embe
  - Consider a protocol Ψ for the

...ndom coordinate S
- Embed random bits A, B on S

u ∈ {0,1}



Lower bounds for this problem imply lower bound for DISJ

...psing distribution μ:

0	1	0	A	0	0	0	0	0	1

0	0	0	B	1	1	0	0	1	0

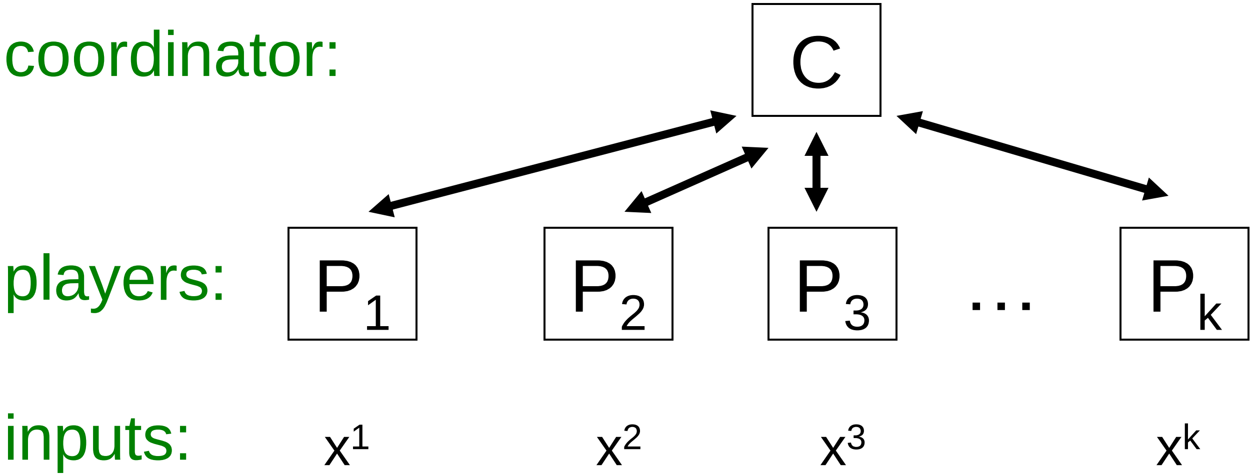
Show CIC^{ext}(Guessing Game) = $\Omega(n)$

Proof related to DISJ lower bound

$= \Omega(r) \cdot \mathrm{CIC}^{ext}$(Guessing Game)

- *Embedding Step*
  - Create a protocol $\Pi_{i,a,b,p,s}$ for Guessing Game on inputs (U,V)
  - Use private randomness, a, b, p, s to sample X^j, Y^j for $j \neq i$
  - Set $(X^i, Y^i) = (U,V)$
  - Let the transcript of $\Pi_{i,a,b,p,s}$ equal the transcript of Π
  - Use predictor θ, given a, b, p, s, P^i, S^i, and the transcript Π, to guess A^i, B^i w. pr. $1/4 + \Omega(1)$, so solve Guessing Game

Distributed Streaming Model



- Each $x^i \in \{-M, -M+1, \ldots, M\}^n$

- Problems on $x = x^1 + x^2 + \ldots + x^k$: sampling, p-norms, heavy hitters, compressed sensing, quantiles, entropy

- Direct Sum of Compositions (generalized to k players): tight bounds for approximating $|x|_2$ and additive ε approx. to entropy

Open Questions

- Direct Sum with Aborts:

  $D_{\mu, \delta}(f^k) = \Omega(k) \cdot IC^{ext}_{\mu, 1/10, \delta/k}(f)$

Instead of f^k, when can we improve the standard direct sum theorem for combining operators such as MAJ, OR, etc.?

- Direct Sum of Compositions: for which functions g is

  $IC^{ext}(\text{Gap-Thresh}(g(x^1, y^1), \ldots, g(x^r, y^r))) = \Omega(r \cdot n)$?

- See Section 4 of http://arxiv.org/abs/1112.5153 for work on a related Gap-Thresh(XOR) problem (a bit different than Gap-Thresh(DISJ))

- Gap-Thresh(DISJ) problem in followup work [WZ]