Beating the Direct Sum Theorem in Communication Complexity

Grigory Yaroslavtsev

Pennsylvania State University

Aarhus Universitet, Theory seminar

Joint work with Marco Molinaro (CMU) and David Woodruff (IBM)

SODA'13

Results

Stronger Direct Sum Theorem in communication complexity for equality-type functions

$$R_{\delta}\left(f^{k}\right) \geq \Omega(k) R_{\frac{\delta}{k}}(f)$$

$$D_{\mu^{k},\delta}\left(f^{k}\right) \geq \Omega(k)\, D_{\mu,\frac{\delta}{k}}(f)$$

} **Yao's principle**

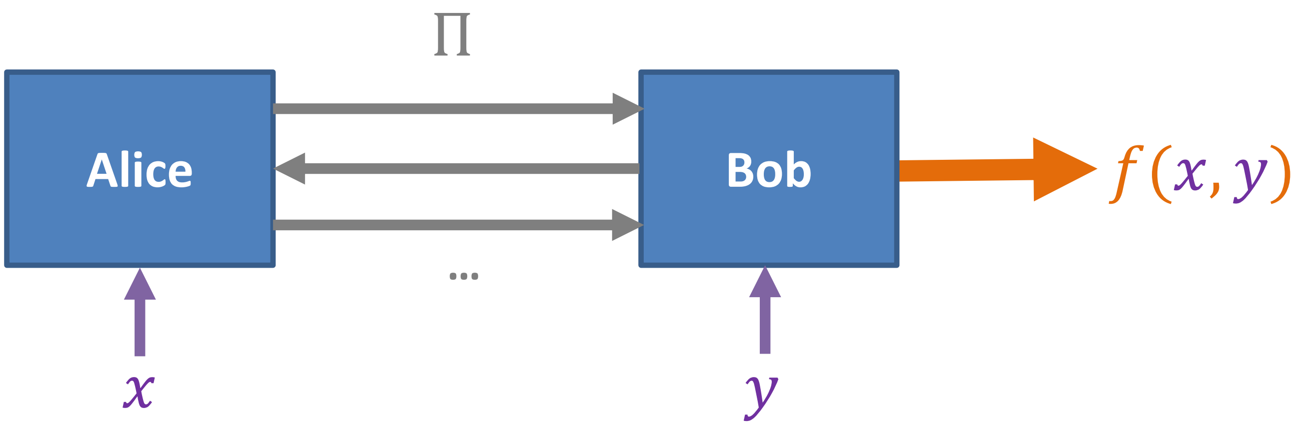
Optimal lower bounds for sketching problems:

- Johnson-Lindenstrauss transform for n vectors
- Pairwise ℓ_1- and ℓ_2-distance estimation
- Matrix multiplication
- Join size estimation of multiple databases

Communication Complexity

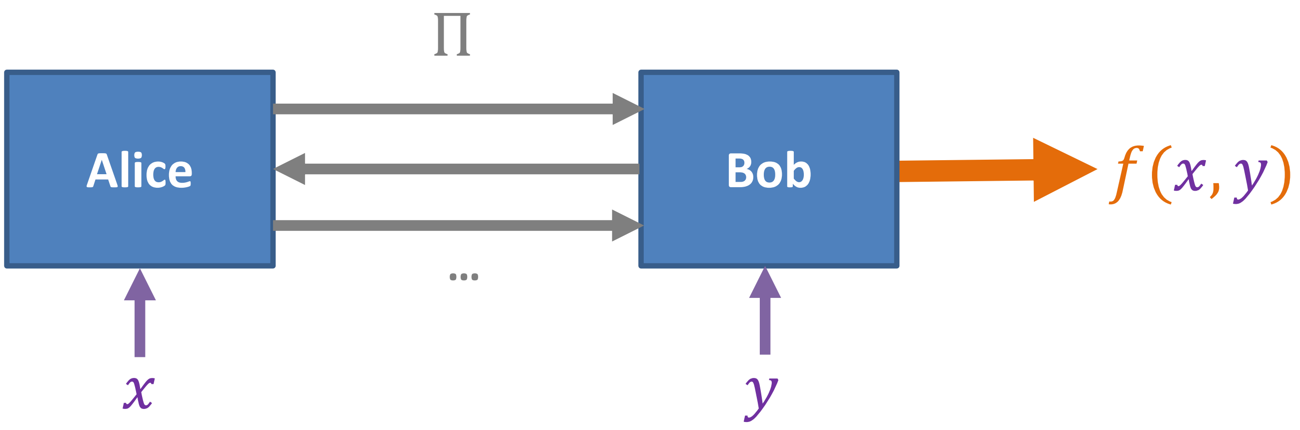
- 2 deterministic players: Alice and Bob
- Joint function f
- Communicate and compute

$\Pi(x, y)$ denotes **transcript** or **output**



Communication Complexity

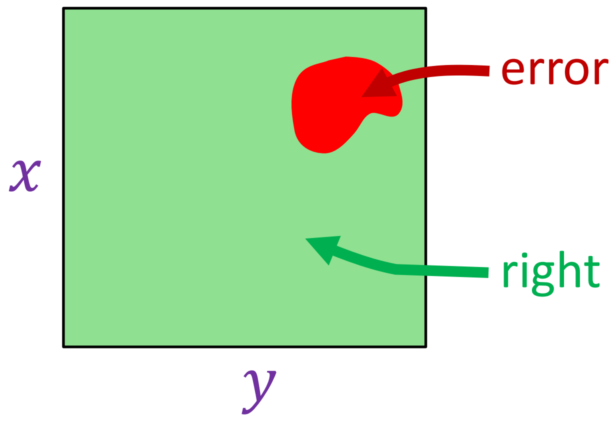
- 2 deterministic players: Alice and Bob
- Joint function f
- Communicate and compute f



- **Ex:** $x, y \in \{0,1\}^n$, want to output $x \stackrel{?}{=} y$

Communication Complexity

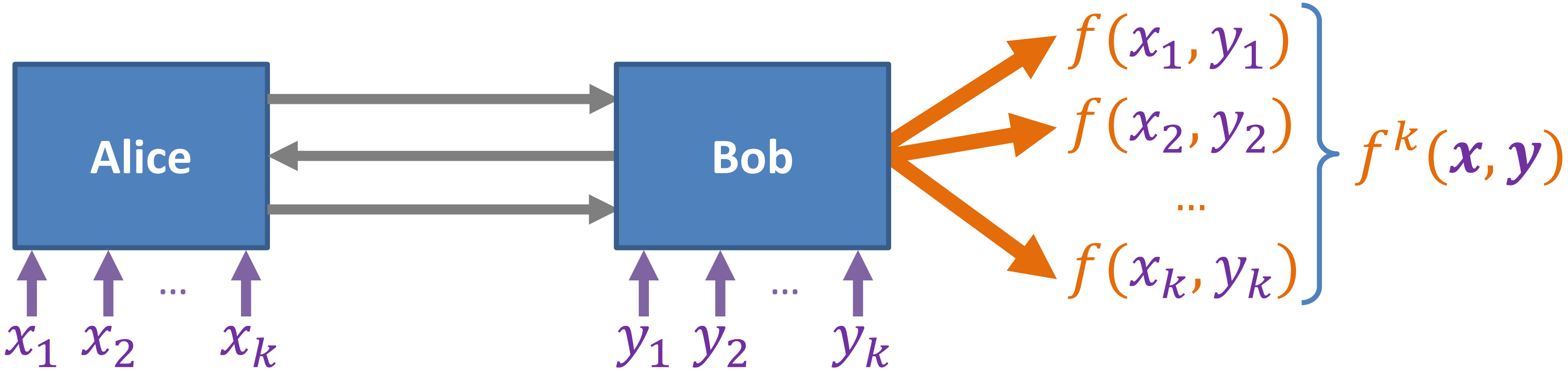
- Consider distribution μ over inputs
- **Goal:** Compute $f(x, y)$ for all but δ μ-fraction of inputs while minimizing longest communication



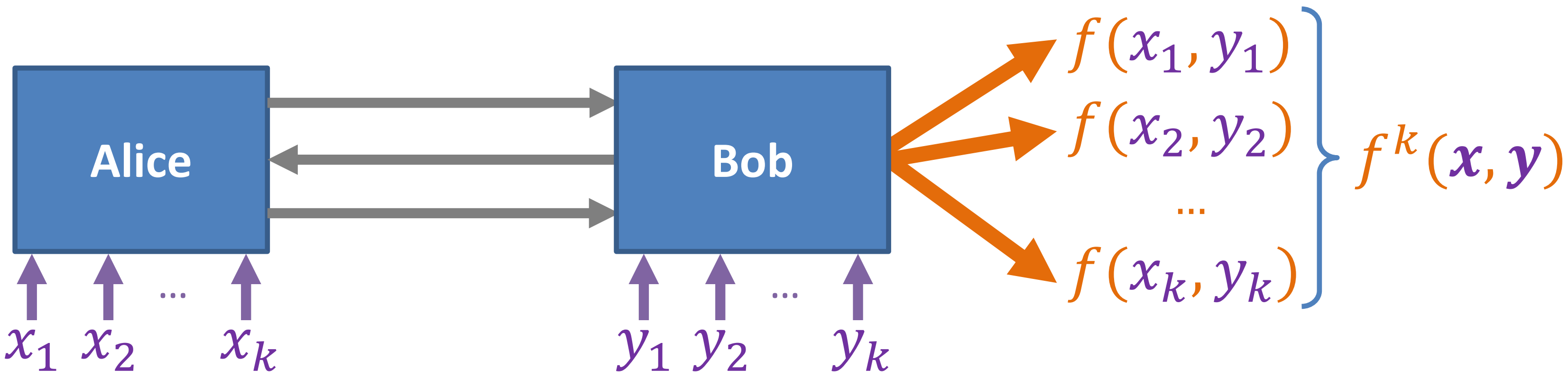
- **Distributional complexity**

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

Multiple Instances



Multiple Instances



- **Goal:** Compute $f^k(\boldsymbol{x}, \boldsymbol{y})$ for all but δ μ^k-fraction of inputs while minimizing longest communication
- **Distributional complexity:** $D_{\mu^k,\delta}(f^k)$

Multiple Instances

Main question: How much can we save against solving each independently?

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

- **Sometimes a bit:** in the private randomness model

$D_{\mu,\frac{1}{3}}(EQ_n) \geq \Omega(\log n)$ but $D_{\mu,\frac{1}{3}}(EQ_n^n) = O(n)$ [FKNN95]

Multiple Instances

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

- **Direct sum theorems**
  - D_{μ^k} … [BBCR 10]
  - D_{μ^k} … [BBCR 10]
  - $D^r_{\mu^k,\delta}$ … [BR 11]

None attains above bound

- **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]

Information Complexity

- **Information cost:** For protocol Π and $(X, Y) \sim \mu$, information revealed about input is

$$\mathrm{IC}_\mu(\Pi) = \mathrm{I}(\Pi(X, Y);\ X, Y) = H(X, Y) - H(X, Y|\Pi)$$

- **Information complexity:**

  $\mathrm{IC}_{\mu,\delta}(f)$ = min information cost over all δ-protocols

Connection: Communication is at least information

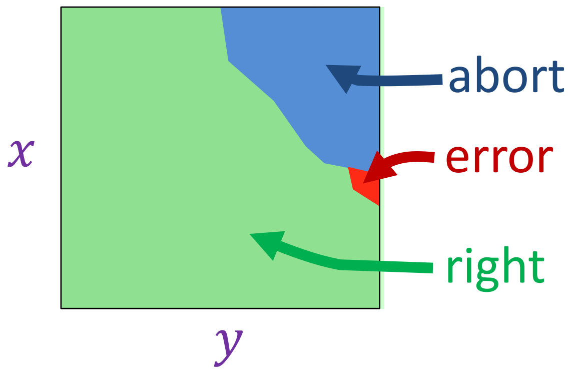
$$D_{\mu,\delta}(f) \geq \mathrm{IC}_{\mu,\delta}(f)$$

- For non-product μ we will work with conditional information complexity $\mathrm{IC}_{\mu,\delta}(f|\nu)$

Protocols with Abortion

Def: A protocol Π $(\boldsymbol{\beta}, \boldsymbol{\delta})$-computes f if

- (Abortion) $\Pr(\Pi(X,Y) = \boldsymbol{abort}) \leq \boldsymbol{\beta}$
- (Error) $\Pr(\Pi(X,Y) \neq f(X,Y) \mid \Pi(X,Y) \neq abort) \leq \boldsymbol{\delta}$



Obs: Stronger guarantee than being wrong with prob. $\approx \boldsymbol{\beta} + \boldsymbol{\delta}$

$\mathrm{IC}_{\mu,\beta,\delta}(f)$ = min information cost over all protocols that $(\boldsymbol{\beta}, \boldsymbol{\delta})$-compute f

Stronger Direct Sum Theorem

Theorem: For every communication problem

Solving k copies with error $\boldsymbol{\delta}$ requires solving each copy with **constant abortion and error $\boldsymbol{\frac{\delta}{k}}$**

Stronger Direct Sum Theorem

- The distribution μ of (X,Y) is a **product** distribution if $\mu(X,Y) = \mu_x(X)\mu_y(Y)$
- $(\mu, \boldsymbol{\nu})$ is a **mixture of product distributions**, if for every $\boldsymbol{t}$ the distribution $(\mu|\boldsymbol{\nu} = \boldsymbol{t})$ is a product distribution

Theorem: For every communication problem f, mixture of product distributions $(\mu, \boldsymbol{\nu})$and $\boldsymbol{\delta} > 0$

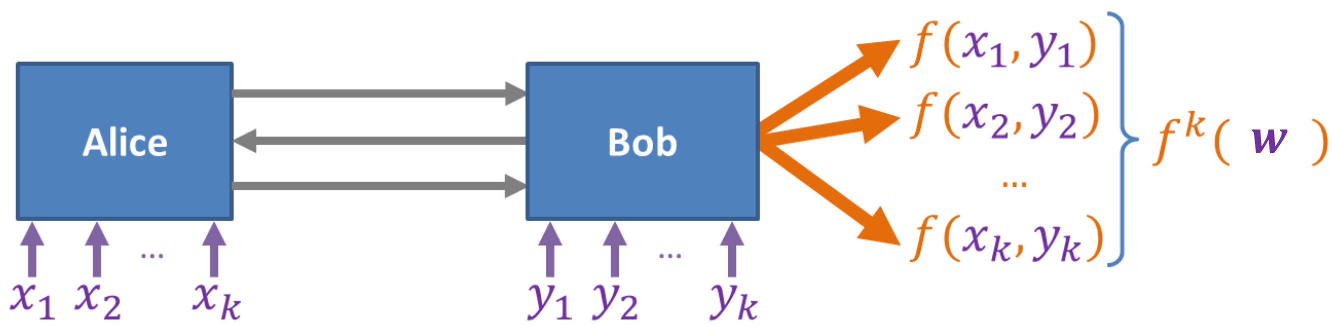
$$\mathrm{IC}_{\mu^{\boldsymbol{k}},\boldsymbol{\delta}}\left(f^{\boldsymbol{k}}|\boldsymbol{\nu}^{\boldsymbol{k}}\right) \geq \Omega(\boldsymbol{k})\ \mathrm{IC}_{\mu,\frac{\boldsymbol{\delta}}{10},O\left(\frac{\boldsymbol{\delta}}{\boldsymbol{k}}\right)}(f|\boldsymbol{\nu})$$

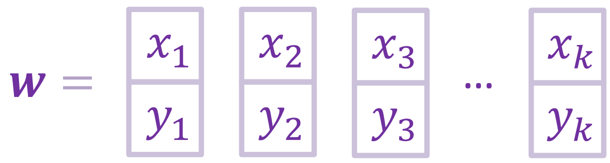
Also holds for one-way and bounded-round communication

Stronger Direct Sum Theorem

Theorem: For every communication problem and product μ

$$\mathrm{IC}_{\mu^k,\delta}(f^k) \geq \Omega(k)\,\mathrm{IC}_{\mu,\frac{\delta}{10},O\left(\frac{\delta}{k}\right)}(f)$$





Stronger Direct Sum Theorem

Proof: Consider protocol Π that computes $f^{\boldsymbol{k}}$ with prob $1-\boldsymbol{\delta}$. Want to show

$$\mathrm{I}(\Pi(\boldsymbol{W});\ \boldsymbol{W}) \geq \Omega(\boldsymbol{k})\mathrm{IC}_{\mu,\frac{\boldsymbol{\delta}}{10},O\left(\frac{\boldsymbol{\delta}}{\boldsymbol{k}}\right)}(f)$$

1) Chain rule:

$$\mathrm{I}(\Pi(\boldsymbol{W}); W) = \sum_{i=1}^{k} \mathrm{I}(\Pi(\boldsymbol{W});\ W_i | \boldsymbol{W}_{<i})$$

By averaging suffices to show that for at least $\Omega(\boldsymbol{k})$ values of $\boldsymbol{i}$

$$\mathrm{I}(\Pi(\boldsymbol{W}); W_i | \boldsymbol{W}_{<i}) \ \geq \mathrm{IC}_{\mu,\frac{\boldsymbol{\delta}}{10},O\left(\frac{\boldsymbol{\delta}}{\boldsymbol{k}}\right)}(f)$$

Want to obtain from Π a protocol **with abortion** to solve i-th copy $f_i^{\boldsymbol{k}}$ with error prob $\frac{\boldsymbol{\delta}}{\boldsymbol{k}}$ and information cost at most

$$\mathrm{I}(\Pi(\boldsymbol{W}); W_i | \boldsymbol{W}_{<i})$$

Stronger Direct Sum Theorem

2) Conditioning amplifies success: For typical i

- $\Pr\left(\Pi_{<i}(\boldsymbol{W}) \neq f_{<i}^{\boldsymbol{k}}(\boldsymbol{W})\right) \leq \boldsymbol{\delta}$
- $\Pr\left(\Pi_i(\boldsymbol{W}) \neq f_i^{\boldsymbol{k}}(\boldsymbol{W}) \mid \Pi_{<i}(\boldsymbol{W}) = f_{<i}^{\boldsymbol{k}}(\boldsymbol{W})\right) = O\left(\frac{\boldsymbol{\delta}}{\boldsymbol{k}}\right)$

This is because

$$1 - \boldsymbol{\delta} \leq \Pr\left(\Pi = f^{\boldsymbol{k}}\right) = \prod_{i=1..k} \Pr\left(\Pi_i = f_i^{\boldsymbol{k}} \mid \Pi_{<i} = f_{<i}^{\boldsymbol{k}}\right)$$

3) Theorem: For a typical i there exists a prefix $\boldsymbol{w}_{<i} \in \mathrm{X}^{\mathrm{i}-1} \times \mathrm{Y}^{\mathrm{i}-1}$ and a set **G** of fixings of the suffix of Π such that:

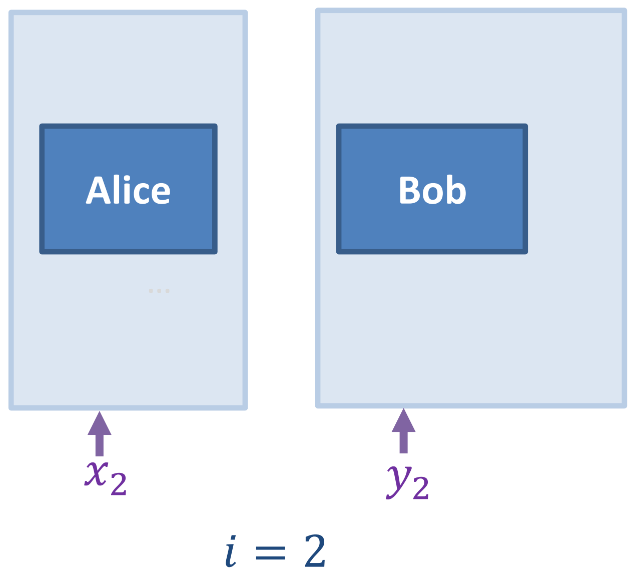
1. Information cost only changes by a constant factor after fixing $\boldsymbol{w}_{<i}$
2. **G** is a constant fraction of all fixings
3. For every fixing $(\boldsymbol{w}_{<i}, \boldsymbol{w}_{>i} \in \mathbf{G})$ the error probability on W_i is $\leq \frac{\boldsymbol{\delta}}{10}$
4. $\Pr\left(\Pi_i(\boldsymbol{w}_{<i} W_{\boldsymbol{i}} \boldsymbol{w}_{>i}) \neq f_i^{\boldsymbol{k}}(W_{\boldsymbol{i}}) \mid \Pi_{<i}(\boldsymbol{w}_{<i} W_{\boldsymbol{i}} \boldsymbol{w}_{>i}) = f_{<i}^{\boldsymbol{k}}(\boldsymbol{w}_{<i} W_{\boldsymbol{i}} \boldsymbol{w}_{>i})\right) = O\left(\frac{\boldsymbol{\delta}}{\boldsymbol{k}}\right)$

Stronger Direct Sum Theorem

Therorem: For a typical i there exists a prefix $\boldsymbol{w}_{<i} \in \mathrm{X}^{\mathrm{i}-1} \times \mathrm{Y}^{\mathrm{i}-1}$ and a set **G** of fixings of the suffix of Π such that:

1. Information cost only changes by a constant factor after fixing $\boldsymbol{w}_{<i}$
2. **G** is a constant fraction of all fixings
3. For every fixing $(\boldsymbol{w}_{<i}, \boldsymbol{w}_{>i} \in \mathbf{G})$ the error probability on $\boldsymbol{w}_{<i}$ is $\leq \frac{\delta}{10}$
4. $\Pr\left(\Pi_i(\boldsymbol{w}_{<i} W_{\boldsymbol{i}} \boldsymbol{w}_{>i}) \neq f_i^{k}(W_{\boldsymbol{i}}) \mid \Pi_{<i}(\boldsymbol{w}_{<i} W_{\boldsymbol{i}} \boldsymbol{w}_{>i}) = f_{<i}^{k}(\boldsymbol{w}_{<i} W_{\boldsymbol{i}} \boldsymbol{w}_{>i})\right) = O\left(\frac{\delta}{k}\right)$

Protocol with abortion for solving f_i^k



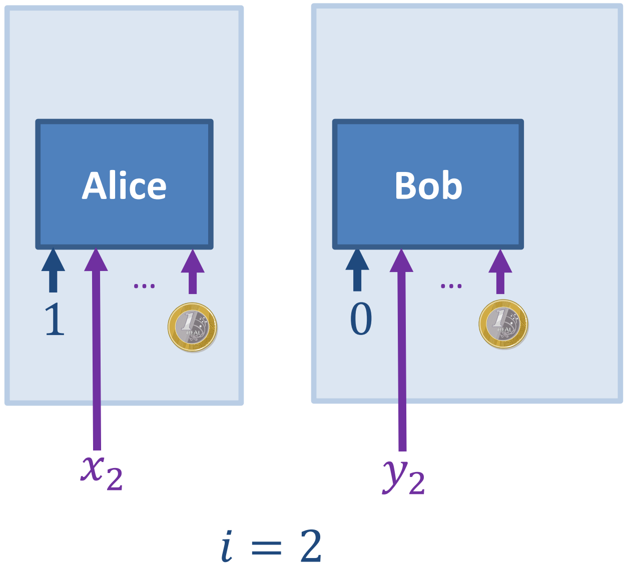
Stronger Direct Sum Theorem

Therorem: For a typical i there exists a prefix $\boldsymbol{w}_{<i} \in \mathrm{X}^{\mathrm{i}-1} \times \mathrm{Y}^{\mathrm{i}-1}$ and a set **G** of fixings of the suffix of Π and random seeds such that:

1. Information cost only changes by a constant factor after fixing $\boldsymbol{w}_{<i}$
2. **G** is a constant fraction of all fixings
3. For every fixing $(\boldsymbol{w}_{<i}, (\boldsymbol{w}_{>i,}\mathrm{r}) \in \mathbf{G})$ the error probability on $\boldsymbol{w}_{<i}$ is $\leq \frac{\delta}{10}$
4. $\Pr\left(\Pi_i(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i}) \neq f_i^k(W_i) \mid \Pi_{<i}(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i}) = f_{<i}^k(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i})\right) = O\left(\frac{\delta}{k}\right)$

Protocol with abortion for solving f_i^k

- Fix the prefix $\boldsymbol{w}_{<i}$ and sample $\boldsymbol{W}_{>i}$



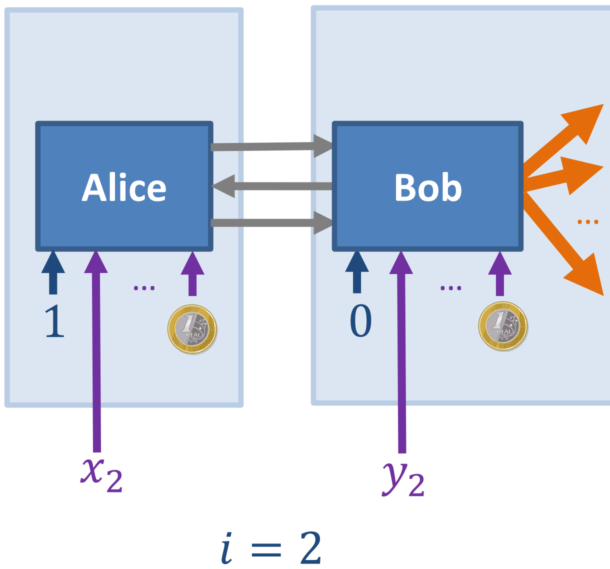
Stronger Direct Sum Theorem

Therorem: For a typical i there exists a prefix $\boldsymbol{w}_{<i} \in \mathrm{X}^{\mathrm{i-1}} \times \mathrm{Y}^{\mathrm{i-1}}$ and a set **G** of fixings of the suffix of Π and random seeds such that:

1. Information cost only changes by a constant factor after fixing $\boldsymbol{w}_{<i}$
2. **G** is a constant fraction of all fixings
3. For every fixing $(\boldsymbol{w}_{<i}, (\boldsymbol{w}_{>i,} \mathrm{r}) \in \mathbf{G})$ the error probability on $\boldsymbol{w}_{<i}$ is $\leq \frac{\delta}{10}$
4. $\Pr\left(\Pi_i(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i}) \neq f_i^k(W_i) \mid \Pi_{<i}(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i}) = f_{<i}^k(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i})\right) = O\left(\frac{\delta}{k}\right)$

Protocol with abortion for solving f_i^k

- Fix the prefix $\boldsymbol{w}_{<i}$ and sample $\boldsymbol{W}_{>i}$
- Run Π



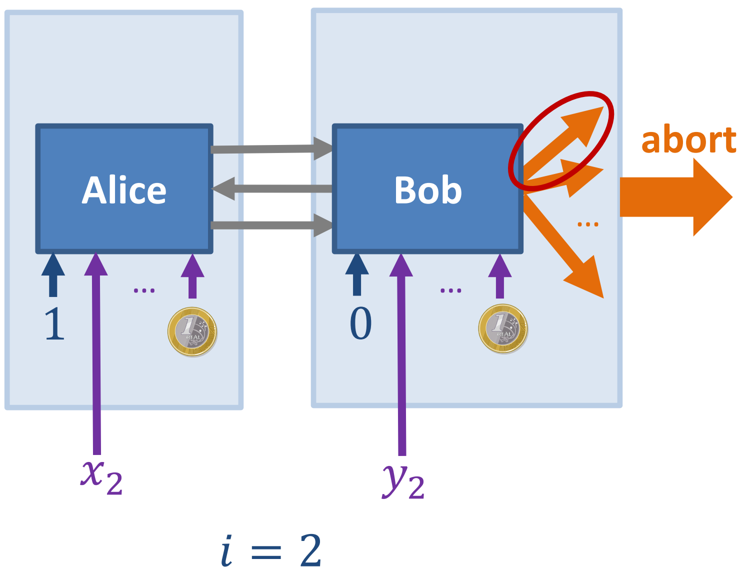
Stronger Direct Sum Theorem

Therorem: For a typical i there exists a prefix $\boldsymbol{w}_{<i} \in \mathrm{X}^{\mathrm{i}-1} \times \mathrm{Y}^{\mathrm{i}-1}$ and a set **G** of fixings of the suffix of Π and random seeds such that:

1. Information cost only changes by a constant factor after fixing $\boldsymbol{w}_{<i}$
2. **G** is a constant fraction of all fixings
3. For every fixing $(\boldsymbol{w}_{<i}, (\boldsymbol{w}_{>i}, \mathrm{r}) \in \mathbf{G})$ the error probability on $\boldsymbol{w}_{<i}$ is $\leq \frac{\delta}{10}$
4. $\Pr\left(\Pi_i(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i}) \neq f_i^k(W_i) \mid \Pi_{<i}(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i}) = f_{<i}^k(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i})\right) = O\left(\frac{\delta}{k}\right)$

Protocol with abortion for solving f_i^k

- Fix the prefix $\boldsymbol{w}_{<i}$ and sample $\boldsymbol{W}_{>i}$
- Run Π
- Verify if error on some copy $1, 2, \ldots i-1$
  - If so, “abort”



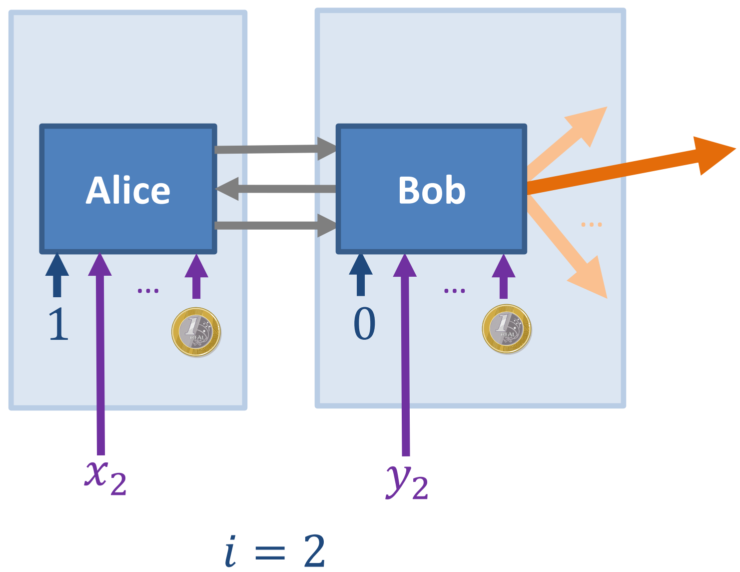
Stronger Direct Sum Theorem

Therorem: For a typical i there exists a prefix $\boldsymbol{w}_{<i} \in \mathrm{X}^{i-1} \times \mathrm{Y}^{i-1}$ and a set **G** of fixings of the suffix of Π and random seeds such that:

1. Information cost only changes by a constant factor after fixing $\boldsymbol{w}_{<i}$
2. **G** is a constant fraction of all fixings
3. For every fixing ($\boldsymbol{w}_{<i}$, ($\boldsymbol{w}_{>i}$,r) ∈ $\mathbf{G}$) the error probability on $\boldsymbol{w}_{<i}$ is $\leq \frac{\delta}{10}$
4. $\Pr\left(\Pi_i(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i}) \neq f_i^k(W_i) \mid \Pi_{<i}(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i}) = f_{<i}^k(\boldsymbol{w}_{<i} W_i \boldsymbol{w}_{>i})\right) = O\left(\frac{\delta}{k}\right)$

Protocol with abortion for solving f_i^k

- Fix a typical prefix $\boldsymbol{w}_{<i}$ and sample $\boldsymbol{W}_{>i}$
- Run Π
- Verify if error on some copy $1,2,\ldots i-1$
  - If so, "abort"
- Else report i-th output



Stronger Direct Sum Theorem

Therorem: For a typical i there exists a prefix $\boldsymbol{w}_{<i} \in \mathrm{X}^{\mathrm{i}-1} \times \mathrm{Y}^{\mathrm{i}-1}$ and a set **G** of fixings of the suffix of Π and random seeds such that:

1. Information cost only changes by a constant factor after fixing $\boldsymbol{w}_{<i}$
2. **G** is a constant fraction of all fixings
3. For every fixing $(\boldsymbol{w}_{<i}, (\boldsymbol{w}_{>i,}\mathrm{r}) \in \mathbf{G})$ the error probability on $\boldsymbol{w}_{<i}$ is $\leq \frac{\delta}{10}$
4. $\Pr\left(\Pi_i(\boldsymbol{w}_{<i} W_{\boldsymbol{i}} \boldsymbol{w}_{>i}) \neq f_i^{\boldsymbol{k}}(W_{\boldsymbol{i}}) \mid \Pi_{<i}(\boldsymbol{w}_{<i} W_{\boldsymbol{i}} \boldsymbol{w}_{>i}) = f_{<i}^{\boldsymbol{k}}(\boldsymbol{w}_{<i} W_{\boldsymbol{i}} \boldsymbol{w}_{>i})\right) = O\left(\frac{\delta}{k}\right)$

Protocol with abortion for solving $f_i^{\boldsymbol{k}}$

- Fix a typical prefix $\boldsymbol{w}_{<i}$ and sample $\boldsymbol{W}_{>i}$
- Run Π
- Verify if error on some copy $1,2,\ldots i-1$
  - If so, "abort"
- Else report i-th output

Protocol $\left(\frac{\delta}{10}, O\left(\frac{\delta}{k}\right)\right)$-computes f_i^k and has information cost exactly $\mathrm{I}(\Pi(\boldsymbol{w}_{<i}\boldsymbol{W}_{\geq \boldsymbol{i}}); W_i)$

Recap

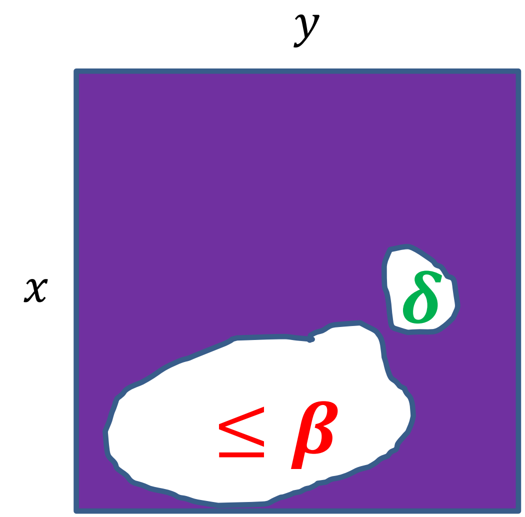
Theorem: For every communication problem

$$\mathrm{IC}_{\mu^{k},\delta}\left(f^{k}|\nu^{k}\right) \geq \Omega(k)\ \mathrm{IC}_{\mu,\frac{1}{20},\frac{\delta}{10},O\left(\frac{\delta}{k}\right)}(f|\nu)$$

Corollary For equality-type problems

$$D_{\mu^{k},\delta}\left(f^{k}\right) \geq \Omega(k)\ \mathrm{IC}_{\mu,\frac{1}{20},\frac{\delta}{10},O\left(\frac{\delta}{k}\right)}(f|\nu) \geq \Omega(k) D_{\mu,\frac{\delta}{k}}(f)$$

Protocols with abortion



- A protocol (α, β, δ)-computes f if with probability $\geq (1-\alpha)$ over its **randomness**
  - It **aborts** with probability $\leq \beta$
  - Conditioned on non-abortion is **correct** w.p. $\geq \mathbf{1-\delta}$

- (μ, ν) is a convex combination of product distributions over $((X \times Y) \times D)$ (marginals: μ over $(X \times Y)$ and ν over D)
- $IC_{\mu,\alpha,\beta,\delta}(f|\nu)$ = minimum information cost of a protocol which (α, β, δ)-computes over (μ, ν).
- $IC_{\mu,\delta}(f|\nu)$ = $IC_{\mu,0,0,\delta}(f|\nu)$

Strong direct sum

- **Strong direct sum:** For every function $\boldsymbol{f}$ and a convex combination of product distributions $(\boldsymbol{\mu}, \boldsymbol{\nu})$

$$\boldsymbol{IC}_{\boldsymbol{\mu}^k,\boldsymbol{\delta}}\left(\boldsymbol{f}^k \middle| \boldsymbol{\nu}^k\right) \geq \Omega(k)\ \boldsymbol{IC}_{\boldsymbol{\mu},\frac{\mathbf{1}}{\mathbf{20}},\frac{\mathbf{1}}{\mathbf{10}},\frac{\boldsymbol{\delta}}{k}}(\boldsymbol{f}|\boldsymbol{\nu})$$

- **Strong** because of high success probability $(1 - \frac{\boldsymbol{\delta}}{k})$
- Gives an extra $\mathbf{log}\,\boldsymbol{k}$ in the lower bound as compared to a weak direct sum [Bar-Yossef, Jayram, Kumar, Sivakumar]

$$\boldsymbol{IC}_{\boldsymbol{\mu}^k,\boldsymbol{\delta}}\left(\boldsymbol{f}^k \middle| \boldsymbol{\nu}^k\right) \geq \Omega(k)\ \boldsymbol{IC}_{\boldsymbol{\mu},\boldsymbol{\delta}}(\boldsymbol{f}|\boldsymbol{\nu})$$

One-way Equality with abortion

- $EQ^{\ell}(\boldsymbol{x},\boldsymbol{y})$ = 1 iff $\boldsymbol{x} = \boldsymbol{y}$, where $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^{\ell}$

- **Theorem:** For $\ell = \log(1/20\boldsymbol{\delta})$ there exists $(\boldsymbol{\mu}, \boldsymbol{\nu})$:

$$\boldsymbol{IC}^{\rightarrow}_{\mu,\frac{\mathbf{1}}{\mathbf{20}},\frac{\mathbf{1}}{\mathbf{10}},\delta}\left(EQ^{\ell}\middle|\boldsymbol{\nu}\right) = \boldsymbol{\Omega}(\log(\mathbf{1}/\boldsymbol{\delta}))$$

- **Corollary:** solving $\boldsymbol{k}$ copies of Equality with constant probability requires one-way communication $\boldsymbol{\Omega}(\boldsymbol{k} \log \boldsymbol{k})$ (for sufficiently long strings x_i, y_i)

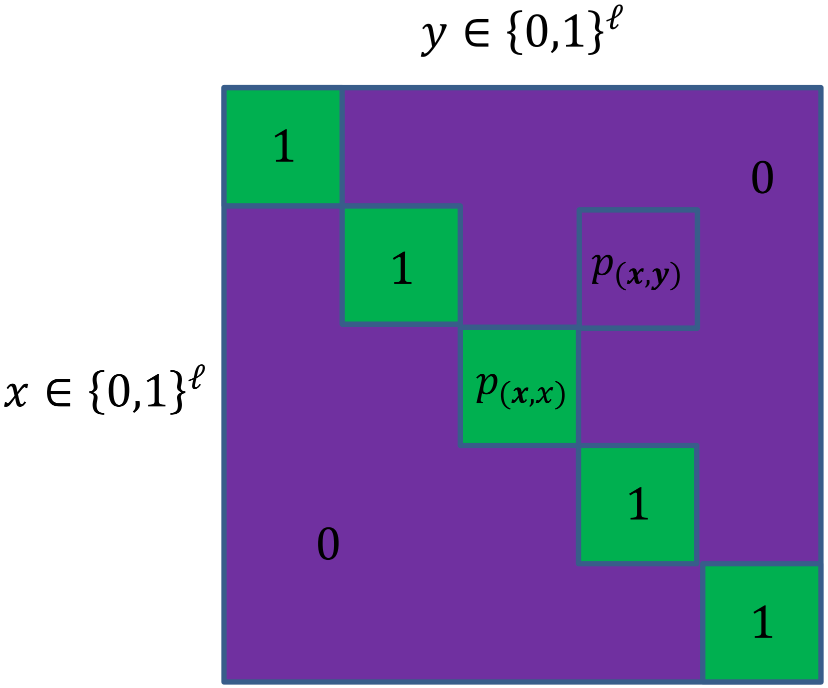
- Hard distribution $((\boldsymbol{X}\ \boldsymbol{Y})\ D_0 \boldsymbol{D})$
  - Random variable for conditioning: $(\mathrm{D}_0 \boldsymbol{D}) \sim U(\{0,1\}^{\ell+1})$
  - If $\mathrm{D}_0 = 0$ then $(\boldsymbol{X}\ \boldsymbol{Y}) \sim U\left(\{0,1\}^{\ell}\right) \times U\left(\{0,1\}^{\ell}\right)$
  - If $\mathrm{D}_0 = 1$ then $\boldsymbol{X} = \boldsymbol{Y} = \boldsymbol{D}$

Equality with abortion

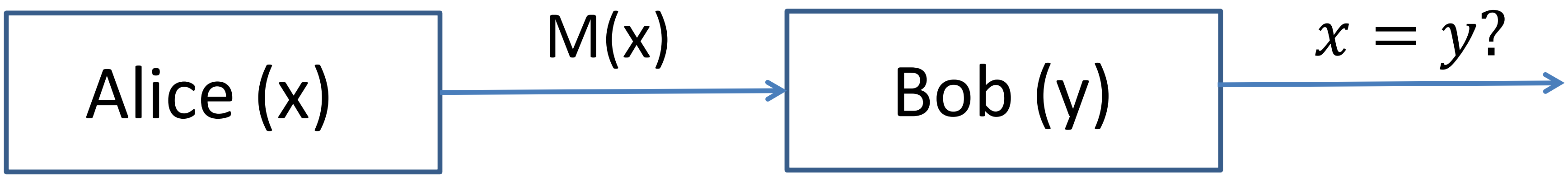
- Hard distribution$((\boldsymbol{X},\boldsymbol{Y}), D_0\boldsymbol{D})$:
  - $(\mathrm{D}_0\boldsymbol{D}) \sim U(\{0,1\}^{\ell+1})$
  - If $\mathrm{D}_0 = 0$ then $(\boldsymbol{X},\boldsymbol{Y}) \sim U(\{0,1\}^{\ell}) \times U(\{0,1\}^{\ell})$
  - If $\mathrm{D}_0 = 1$ then $\boldsymbol{X} = \boldsymbol{Y} = \boldsymbol{D}$ ($\geq \frac{1}{2}$ of the mass on the diagonal)

For $\ell = \log \frac{1}{20\delta}$

- $\boldsymbol{p}_{(\boldsymbol{x},\boldsymbol{x})} = 200\delta^2 + 10\delta$
- $\boldsymbol{p}_{(\boldsymbol{x},\boldsymbol{y})} = 200\delta^2$



Equality with abortion



- $\boldsymbol{IC}^{\rightarrow}_{\boldsymbol{\mu},\ldots}(\boldsymbol{f}|\boldsymbol{\nu}) = \min_{M} I(M(\boldsymbol{X}); \boldsymbol{X}, \boldsymbol{Y}|D_0\boldsymbol{D}) = \min_{M} I(M(\boldsymbol{X}); \boldsymbol{X}|D_0\boldsymbol{D})$
- $I(M(\boldsymbol{X}); \boldsymbol{X}|D_0\boldsymbol{D}) = \boldsymbol{H(X|D_0D)} - \boldsymbol{H(X|M(X), D_0D)}$

- $\boldsymbol{H(X|D_0D)} \geq \Pr[D_0 = 0] \cdot H(\boldsymbol{X}|D_0 = 0, \boldsymbol{D})) = 1/2 \log{(1/20\delta)}$

- By Fano's inequality ([Cover, Thomas]):

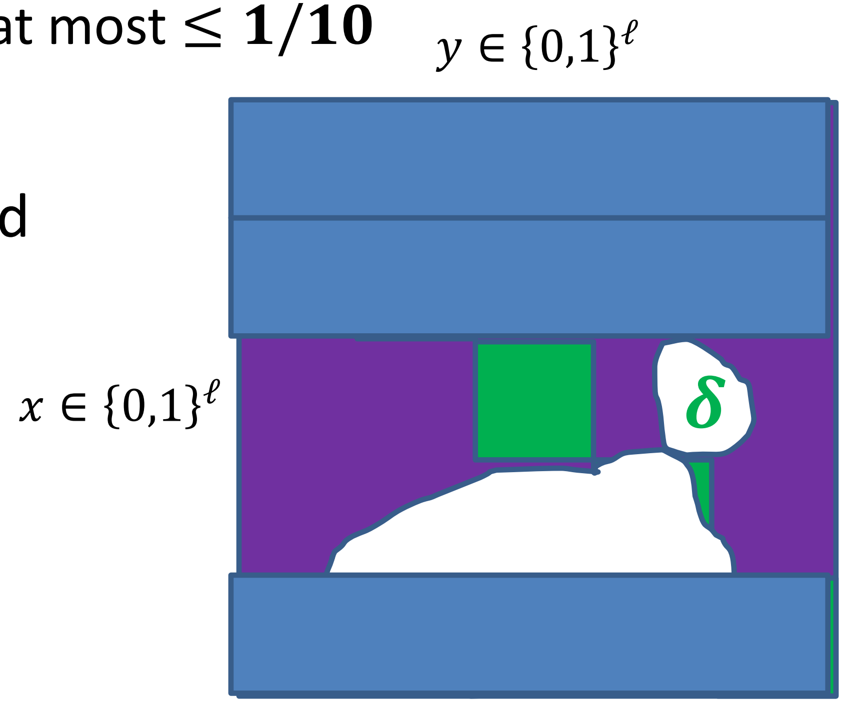
$\boldsymbol{H(X|M(X), D_0D)} \leq 1 + p_e \log(|supp(\boldsymbol{X})|) = 1 + \boldsymbol{p_e} \log{(1/20\delta)}$,

where $\boldsymbol{p_e} = \min_{g} \Pr[g\big(M(\boldsymbol{X}, D_0\boldsymbol{D})\big) \neq \boldsymbol{X}]$ and g is a deterministic function

- Suffices to show that there exists a predictor with error $p_e \leq \frac{2}{5} < \frac{1}{2}$

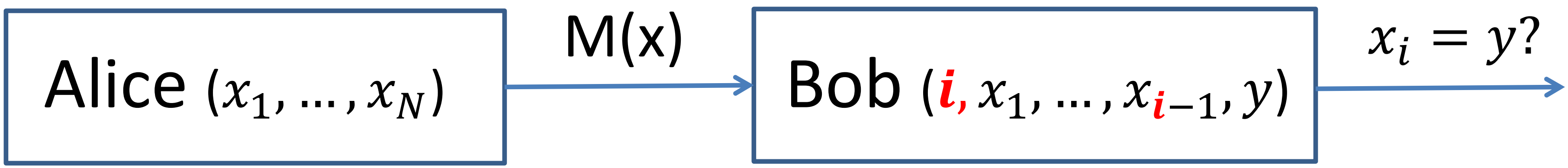
Predictor for Equality

- A row x is **good** if the protocol $\Pi(x, y) = 1$ **iff** $x = y$
- If the $\Pi\left(0, \mathbf{\frac{1}{10}}, \boldsymbol{\delta}\right)$-computes EQ^{ℓ} then $\leq \mathbf{\frac{3}{10}}$ fraction of rows is not **good**:
  - Fraction of rows with an abortion on the diagonal (x, x) is $\leq \mathbf{1/5}$
  - Fraction of rows with an error is at most $\leq \mathbf{1/10}$
- **Predictor**: If the row is **good** then Bob can simulate Π for every y and recover x!
- If $\Pi\left(\mathbf{\frac{1}{20}}\ \mathbf{\frac{1}{10}}, \boldsymbol{\delta}\right)$ -computes EQ^{ℓ} then $p_e \leq \frac{3}{10} + \mathbf{\frac{1}{20}} < \frac{2}{5}$



Augmented indexing

- Augmented indexing over large alphabet ($x_i, y \in [\boldsymbol{m}]$)



- **Theorem:** For sufficiently large $\boldsymbol{m}$ there exists $(\boldsymbol{\mu}, \boldsymbol{\nu})$:
  $$\boldsymbol{IC}^{\rightarrow}_{\boldsymbol{\mu}, \frac{\mathbf{1}}{\mathbf{20}}, \frac{\mathbf{1}}{\mathbf{10}}, \frac{\mathbf{1}}{\boldsymbol{m}}}(Augmented\ Index|\boldsymbol{\nu}) = \boldsymbol{\Omega}(\boldsymbol{N}\ log\ \boldsymbol{m})$$

- **Corollary:** Solving $\boldsymbol{k}$ copies of Augmented Indexing (with const. prob.) requires one-way communication $\boldsymbol{\Omega}(\boldsymbol{N}\ \boldsymbol{k} \log \boldsymbol{k})$ (for sufficiently large alphabet size)

Application: JL-transform of **n** vectors

- Let **S** be a distribution over $\boldsymbol{k} \times d$ matrices, such that for any $v_1, \ldots, v_n \in R^d$ with prob. $\geq 1 - \boldsymbol{\delta}$
$$\left\| \boldsymbol{S} v_i - \boldsymbol{S} v_j \right\|_2 = (1 \pm \boldsymbol{\epsilon}) \left\| v_i - v_j \right\|_2$$
- $\boldsymbol{k} = \#$ rows in $\boldsymbol{S} \geq \frac{\mathbf{1}}{\boldsymbol{\epsilon}^{\mathbf{2}}} \log \left(\frac{\boldsymbol{n}}{\boldsymbol{\delta}} \right)$, dependence on **n** is new
- Even if $\boldsymbol{S}$ is allowed to depend on the first **n**/2 points
- Any enconding $\phi(v_1), \ldots, \phi(v_n)$ that allows pairwise ℓ_p-distance estimation for $p \in \{1,2\}$ requires
$$\Omega\left(\boldsymbol{n}\, \boldsymbol{\epsilon}^{-2} \log \frac{\boldsymbol{n}}{\boldsymbol{\delta}} \, (\log d + \log M) \right) \text{ bits}$$
(M = max abs. value in v_i)

Other applications

- Sketching matrix products
  - Minimum number of columns in a $\boldsymbol{n} \times \boldsymbol{k}$ matrix **S** such that $C = A\boldsymbol{S}\boldsymbol{S}^{\boldsymbol{T}}B$ is a good approximation for AB, where A, B are $\boldsymbol{n} \times \boldsymbol{n}$ matrices?
  - $\left|(AB)_{i,j} - C_{i,j}\right| \leq \epsilon \left|\left|A_i\right|\right|_2 \left|\left|B^j\right|\right|_2$ => $\boldsymbol{k} = O(\epsilon^{-2}\log\frac{\boldsymbol{n}}{\boldsymbol{\delta}})$ [Sarlos]
  - $\left|\left|AB - C\right|\right|_F \leq \epsilon \left|\left|A\right|\right|_F \left|\left|B\right|\right|_F$ => $\boldsymbol{k} = O(\epsilon^{-2}\log\frac{1}{\boldsymbol{\delta}})$ [Clarkson,Woodruff]
  - **Our result:** entry-wise guarantee indeed requires $\boldsymbol{k} = \Omega(\epsilon^{-2}\log\frac{\boldsymbol{n}}{\boldsymbol{\delta}})$

- Optimality of database sketching [Alon, Gibbons, Matias, Szegedy] and mergeable summaries

Open problems

- **Strong direct sum:** For every function $\boldsymbol{f}$ and a convex combination of product distributions $(\boldsymbol{\mu}, \boldsymbol{\nu})$
$$\boldsymbol{IC}_{\mu^k,\delta}\left(\boldsymbol{f}^k \middle| \boldsymbol{\nu}^k\right) \geq \Omega(k)\ \boldsymbol{IC}_{\mu,\frac{1}{20},\frac{1}{10},\frac{\delta}{k}}(\boldsymbol{f}|\boldsymbol{\nu})$$
- More problems with low-error one-way lower bounds?
- Natural problems for low-error 2-way lower bounds (disjointness doesn't work)?
- Applications of direct sums to property testing? [Blais, Brody, Matulef '11, Goldreich '13]
- Strong direct sum for predicates $g\big(f(x_1, y_1), \ldots f(x_k, y_k)\big)$? For OR-EQUALITY ($g = \vee$, $f = EQ^{\ell}$) there is a direct sum [Brody, Chakrabarti, Kondapally'12, Saglam, Tardos'13]