Aspects of 1-Way Communication of Index

- Alice has $x \in \{0,1\}^n$
- Bob has $i \in [n]$
- Alice sends a (randomized) message $M$ to Bob
- $I(M ; X \mid R) = \sum_i I(M ; X_i \mid X_{<i}, R)$
  \[ \geq \sum_i I(M; X_i \mid R) \]
  \[ = n - \sum_i H(X_i \mid M, R) \]
- **Fano**: $H(X_i \mid M, R) \leq H(\delta)$ if Bob can guess $X_i$ with probability $> 1 - \delta$
- $CC_\delta(\text{Index}) \geq I(M ; X \mid R) \geq n(1-H(\delta))$

*The same lower bound applies if the protocol is only correct on average over $x$ and $i$ drawn independently from a uniform distribution*
Distributional Communication Complexity

- $(X, Y) \sim \mu$

- \textit{\(\mu\)-distributional complexity} \(D_\mu(f)\): the minimum communication cost of a protocol which outputs \(f(X,Y)\) with probability 2/3 for \((X, Y) \sim \mu\)
  - Yao’s minimax principle: \(R(f) = \max_{\mu} D_\mu(f)\)

- 1-way communication: Alice sends a single message \(M(X)\) to Bob
Indexing is Universal for Product Distributions [Kremer, Nisan, Ron]

• Communication matrix $A_f$ of a Boolean function $f: X \times Y \rightarrow \{0,1\}$ has $(x,y)$-th entry equal to $f(x,y)$

• $\max_{\text{product } \mu} D_\mu(f) = \Theta(\text{VC} - \text{dimension})$ of $A_f$

• Implies a reduction from Index is optimal for product distributions

\[
\begin{bmatrix}
0 & 0 & 1 \\
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 1 & 0
\end{bmatrix}
\]
Indexing with Low Error

- Index Problem with 1/3 error probability and 0 error probability both have $\Omega(n)$ communication

- Sometimes, want lower bounds in terms of error probability

- Indexing on Large Alphabets:
  - Alice has $x \in \{0,1\}^{n/\delta}$ with $\text{wt}(x) = n$, Bob has $i \in [n/\delta]$
  - Bob wants to decide if $x_i = 1$ with error probability $\delta$
  - [Jayram, W] 1-way communication is $\Omega(n \log(1/\delta))$
  - Can be used to get an $\Omega(\log(\frac{1}{\delta}))$ bound for norm estimation
  - We’ve seen an $\Omega(\log n + \epsilon^{-2} + \log(\frac{1}{\delta}))$ lower bound for norm estimation
  - There is an $\Omega(\epsilon^{-2} \log \frac{1}{\delta} \log n)$ bit lower bound
Beyond Product Distributions

Although $R(f) = \max_{\mu} D_{\mu}(f)$, it may be that

$$\max_{\mu} D_{\mu}(f) \gg \max_{\text{product } \mu} D_{\mu}(f),$$

so one often can’t get good lower bounds by looking at product distributions…
Non-Product Distributions

• Needed for stronger lower bounds

• Example: approximate $|x|_\infty$ up to a multiplicative factor of $B$ in a stream
  – Lower bounds for $p$-norms

  $$\text{Gap}_\infty(x, y)$$
  Problem

  \[ x \in \{0, \ldots, B\}^n \quad y \in \{0, \ldots, B\}^n \]

  • Promise: $|x - y|_\infty \leq 1$ or $|x - y|_\infty \geq B$

  • Hard distribution non-product

  • $\Omega(n/B^2)$ lower bound [Saks, Sun] [Bar-Yossef, Jayram, Kumar, Sivakumar]
Direct Sums

- $\text{Gap}_\infty(x,y)$ doesn’t have a hard product distribution, but has a hard distribution $\mu = \lambda^n$ in which the coordinate pairs $(x_1, y_1), \ldots, (x_n, y_n)$ are independent
  
  - w.pr. $1-1/n$, $(x_i, y_i)$ random subject to $|x_i - y_i| \leq 1$
  
  - w.pr. $1/n$, $(x_i, y_i)$ random subject to $|x_i - y_i| \geq B$

- **Direct Sum**: solving $\text{Gap}_\infty(x,y)$ requires solving $n$ single-coordinate sub-problems $g$

- In $g$, Alice and Bob have $J, K \in \{0, \ldots, B\}$, and want to decide if $|J-K| \leq 1$ or $|J-K| \geq B$
Direct Sum Theorem

• Let $M$ be the message from Alice to Bob

• For $(X, Y) \sim \mu$, $I(M ; X, Y) = H(X,Y) - H(X,Y \mid M)$ is the information cost of the protocol

• [BJKS]: why not measure $I(M ; X, Y)$ when $(X,Y)$ satisfy $|X - Y|_\infty \leq 1$?
  – Is $I(M ; X, Y)$ large?
  – Let us go back to protocols correct on each $X, Y$ w.h.p.

• Define $\mu = \lambda^n$, where $(X_i, Y_i) \sim \lambda$ is random subject to $|X_i - Y_i| \leq 1$

• $IC(g) = \inf_\psi I(\psi ; J, K)$, where $\psi$ ranges over all $2/3$-correct 1-way protocols for $g$, and $J,K \sim \lambda$

  Is $I(M ; X, Y) = \Omega(n) \cdot IC(g)$?
The Embedding Step

- $I(M; X, Y)$

- We need to show $I(M; X_i, Y_i)$ for each $i$

Suppose Alice and Bob could fill in the remaining coordinates $j$ of $X, Y$ so that $(X_j, Y_j) \sim \lambda$

Then we get a correct protocol for $g$!
Conditional Information Cost

- \((X_j, Y_j) \sim \lambda\) is not a product distribution

- [BJKS] Define \(D = ((P_1, V_1), \ldots, (P_n, V_n))\):
  - \(P_j\) uniform on \{Alice, Bob\}
  - \(V_j\) uniform on \{1, \ldots, B\} if \(P_j = \text{Alice}\)
  - \(V_j\) uniform on \{0, \ldots, B-1\} if \(P_j = \text{Bob}\)
  - If \(P_j = \text{Alice}\), then \(Y_j = V_j\) and \(X_j\) is uniform on \(\{V_j-1, V_j\}\)
  - If \(P_j = \text{Bob}\), then \(X_j = V_j\) and \(Y_j\) is uniform on \(\{V_j, V_j+1\}\)

  \(X\) and \(Y\) are independent conditioned on \(D\)!

- \(I(M ; X, Y | D) = \Omega(n) \cdot IC(g | (P,V))\) holds now!

- \(IC(g | (P,V)) = \inf_{\psi} I(\psi ; J, K | (P,V))\), where \(\psi\) ranges over all 2/3-correct protocols for \(g\), and \(J,K \sim \lambda\)
Primitive Problem

- Need to lower bound $IC(g \mid (P,V))$

- For fixed $P = Alice$ and $V = v$, this is $I(\psi ; K)$ where $K$ is uniform over $v-1$, $v$

- From previous lecture: $I(\psi ; K) \geq D_{JS}(\psi_{v-1,v}, \psi_{v,v})$

- $IC(g \mid (P,V)) \geq E_v [D_{JS}(\psi_{v-1,v}, \psi_{v,v}) + D_{JS}(\psi_{v,v}, \psi_{v,v+1})] / 2$

  *Forget about distributions, let’s move to unit vectors!*
Hellinger Distance

- For distribution $\mu$, let $\sqrt{\mu}$ be the vector with coordinate $i$ equal to $\mu_i^{1/2}$

- $D_{JS}(\psi_{v-1,v}, \psi_{v,v}) \geq h(\psi_{v-1,v}, \psi_{v,v})^2$

(*) $IC(g \mid (P,V)) \geq E_v [h(\psi_{v-1,v}, \psi_{v,v})^2 + h(\psi_{v,v}, \psi_{v,v+1})^2] / 2$

- Properties
  - (Correctness) $h(\psi_{0,0}, \psi_{0,0})^2 = \Omega(1)$
  - (1-way Protocol) $\psi_{a,b}(m, out) = p_a(m) \cdot q_{b,m}(out)$
  - (Pythagorean) $h^2(\psi_{a,b}, \psi_{c,d}) \geq \frac{1}{2} (h^2(\psi_{a,b}, \psi_{a,d}) + h^2(\psi_{c,b}, \psi_{c,d}))$
Pythagorean Property

$$\frac{1}{2} (1 - h^2(\psi_{a,b}, \psi_{a,d}) + 1 - h^2(\psi_{c,b}, \psi_{c,d}))$$

$$= \frac{1}{2} \sum_{m,b} \left( \sqrt{p_a(m)} \cdot \sqrt{q_{b,m}(\text{out})} \sqrt{p_a(m)} \sqrt{q_{d,m}(\text{out})} + \sqrt{p_c(m)} \sqrt{q_{b,m}(\text{out})} \sqrt{p_c(m)} \sqrt{q_{d,m}(\text{out})} \right)$$

$$= \sum_{m,b} \frac{p_a(m) + p_c(m)}{2} (\sqrt{q_{b,m}(\text{out})} \sqrt{q_{d,m}(\text{out})})$$

$$\geq \sum_{m,b} \sqrt{p_a(m)} \sqrt{p_c(m)} \sqrt{q_{b,m}(\text{out})} \sqrt{q_{d,m}(\text{out})}$$

$$= 1 - h^2(\psi_{a,b}, \psi_{c,d})$$
Lower Bounding the Primitive Problem

- \text{IC}(f | (P, V))
  \geq E_v \left[ h(\psi_{v,v}, \psi_{v,v+1})^2 + h(\psi_{v,v}, \psi_{v+1,v})^2 \right] / 2
  \geq 1/(2B) \sum_v |\sqrt{\psi_{v-1,v}} - \sqrt{\psi_{v,v}}|^2 + |\sqrt{\psi_{v,v}} - \sqrt{\psi_{v,v+1}}|^2
  \geq 1/(2B^2) \left( \sum_v |\sqrt{\psi_{v-1,v}} - \sqrt{\psi_{v,v}} + |\sqrt{\psi_{v,v}} - \sqrt{\psi_{v,v+1}}|)^2 \right) (1/2)
  \geq 1/(2B^2) \left( \sum_v |\sqrt{\psi_{v,v}} - \sqrt{\psi_{v+1,v+1}}| \right)^2 (1/2)
  \geq 1/(2B^2) |\sqrt{\psi_{0,0}} - \sqrt{\psi_{B,B}}|^2 (1/2)
  \geq 1/(4B^2) \left( |\sqrt{\psi_{0,0}} - \sqrt{\psi_{0,B}}|^2 + |\sqrt{\psi_{B,0}} - \sqrt{\psi_{B,B}}|^2 \right)^2 \right) (1/2)
  = \Omega(1/B^2)
Direct Sum Wrapup

• $\Omega(n/B^2)$ bound for $\text{Gap}_\infty(x,y)$

• Similar argument gives $\Omega(n)$ bound for disjointness [BJKS]

• [Molinaro, Yaroslavtsev, W] Sometimes can “beat” a direct sum: solving all $n$ copies simultaneously with constant probability as hard as solving each copy with probability $1-1/n$
  – E.g., 1-way communication complexity of Equality

• Direct sums are nice, but often a problem can’t be split into simpler smaller problems, e.g., no known embedding step in gap-Hamming