Direct Sum in Interactive Communication Models Using Information-theoretic Tools

COMS 6998 Communication Complexity Applications

Alexander Lindenbaum, Yunya Zhao

April 6, 2022

Review: Information Theory Preliminaries

- Entropy of a random variable X

$$\mathrm{H}(X) = \sum_x p(x) \cdot \log \frac{1}{p(x)} = \mathop{\mathbb{E}}_{p(x)} \left[\log \frac{1}{p(x)} \right]$$

- Conditional Entropy

$$\mathrm{H}(Y|X) = \mathop{\mathbb{E}}_{p(xy)} \left[\log \frac{1}{p(y|x)} \right] = \mathop{\mathbb{E}}_{p(x)} [\mathrm{H}(Y|X = x)]$$

- Chain Rule of Entropy

$$\mathrm{H}(XY) = \mathrm{H}(X) + \mathrm{H}(Y|X)$$

Review: Information Theory Preliminaries

- Mutual Information

$$I(A;B) = H(A) - H(A|B)$$

- Conditional Mutual Information

$$I(A;B|C) = H(A|C) - H(A|BC)$$

- Chain Rule of Mutual Information

$$I(AB;C) = I(A;C) + I(B;C|A)$$

- Chain Rule of Conditional Mutual Information

$$I(AB;C|D) = I(A;C|D) + I(B;C|AD)$$

Relating to Communication: Information Complexity

Analogous to Communication Cost and Communication Complexity:

Information Cost is related to the amount of information gained through the execution of a communication protocol π

Information Complexity is related to a function f (a problem) over all protocols that computes it.

Relating to Communication: Information Complexity

- **Transcript** of a protocol
  Given a protocol π, the **transcript** $\boldsymbol{\pi}(\mathbf{X}, \mathbf{Y})$ is the concatenation of the public randomness with all the messages that are sent during the execution of π on input X, Y

- **Internal Information Cost**
  (Distributional) Internal information cost $\mathsf{IC}^i_\mu(\pi)$ is how much each party learns about the other party's input during the execution of π

$$\mathsf{IC}^i_\mu(\pi) = I(X; \pi(X,Y)|Y) + I(Y; \pi(X,Y)|X)$$

Relating to Communication: Information Complexity

- **Internal Information Cost**
  (Distributional) Internal information cost $\mathrm{IC}_\mu^i(\pi)$ is how much each party learns about the other party's input during the execution of π

$$\mathrm{IC}_\mu^i(\pi) = I(X; \pi(X, Y)|Y) + I(Y; \pi(X, Y)|X)$$

- **External Information Cost**
  (Distributional) External information cost $\mathrm{IC}_\mu^{ext}(\pi)$ is how much information an outside observer learns about both parties' input just by looking at Alice and Bob chat

$$\mathrm{IC}_\mu^{ext}(\pi) = I(XY; \pi(X, Y))$$

Internal IC $\leq$ External IC

For protocol π and distribution μ, we have

$$\mathsf{IC}^{i}_{\mu}(\pi) \leq \mathsf{IC}^{ext}_{\mu}(\pi)$$

[Intuition] at each round, an independent observer is always going to learn more *new* info about XY than X and Y about each other, or more formally:

Internal IC $\leq$ External IC

For protocol π and distribution μ, we have

$$\mathsf{IC}^{i}_{\mu}(\pi) \leq \mathsf{IC}^{ext}_{\mu}(\pi)$$

Proof.
Let ω be any fixed prefix of the transcript of length $i-1$.

If it is the X player's turn to speak, the amount of info she learns about Y is zero

$$I(Y; \pi(X,Y)_i | X, \pi(X,Y)_{\leq i-1} = \omega) = 0$$

Similarly, if it is the Y player's turn to speak, the amount of info he learns about X is zero. So at each round, there has to be one player who learns nothing new.

Internal IC $\leq$ External IC

For protocol π and distribution μ, we have

$$\mathsf{IC}^{i}_{\mu}(\pi) \leq \mathsf{IC}^{ext}_{\mu}(\pi)$$

Proof.
On the other hand, an observer always learns something new at each round, and that amount is

$$I(XY; \pi(X,Y)_i | \pi(X,Y)_{\leq i-1} = \omega)$$

$$= I(X; \pi(X,Y)_i | \pi(X,Y)_{\leq i-1} = \omega) + I(Y; \pi(X,Y)_i | X\pi(X,Y)_{\leq i-1} = \omega)$$

$$\geq I(X; \pi(X,Y)_i | Y\pi(X,Y)_{\leq i-1} = \omega) + I(Y; \pi(X,Y)_i | X\pi(X,Y)_{\leq i-1} = \omega)$$

NOTE: if μ is a product distribution, $\mathsf{IC}^{i}_{\mu}(\pi) = \mathsf{IC}^{ext}_{\mu}(\pi)$ $\square$

Motivation: Direct Sum

The **direct sum question** is about the complexity of solving *several* copies of a given problem. In communication complexity, it can be phrased as follows:

given function

$$f : \{0,1\}^m \times \{0,1\}^m \longrightarrow \{0,1\}$$

define

$$f^n : (\{0,1\}^m)^n \times (\{0,1\}^m)^n \longrightarrow \{0,1\}^n$$

to be

$$f^n((x_1, ..., x_n), (y_1, ..., y_n)) = (f(x_1, y_1), ..., f(x_n, y_n))$$

What is the relationship between the communication costs of f and f^n?

Motivation: Direct Sum

Why direct sum?

Hardness Amplification

direct sum + lower bound on “primitive” problem = lower bound on “composite” problem

- Ex. Karchmer-Raz-Wigderson: $\mathsf{P} \neq \mathsf{NC}^1$ if circuit depth has strong direct sum (there are inherently sequential problems)

Very sensitive to models

Motivation: Direct Sum

The communication complexity for f^n is most n times the communication complexity of f.

$$D(f^n) \leq n \cdot D(f)$$

Is this the best we could do?
We don't know...

Motivation: Direct Sum

- **Strong Direct Sum Conjecture** "the naive is the optimal"

$$D_\rho^{\mu^n}(f^n) \stackrel{?}{=} \Omega(n) \cdot D_\rho^{\mu}(f)$$

One direction is trivial, need to prove the other direction

- Direct Sum Theorem for Simultaneous Communication (the equality function)[CSWY01]

$$C(\mathrm{EQ}_n^m) = \Omega(m\sqrt{n})$$

Why Information Complexity - Information Theoretical tools

- CSWY01 used information theoretic tools to arrive at direct sum.
- Information Complexity has a nice direct sum property

$$IC^n(f) \geq n \cdot IC(f)$$

- The above property bridges together direct sum of communication:

$$D^n(f) \geq IC^n(f) \geq n \cdot IC(f) \;\; ??? \;\; n \cdot D(f)$$

Notations

Given a function $f(x, y)$ and a distribution μ on inputs to f

- The communication complexity $D_\rho^\mu(f)$, maximum number of bits communicated by a protocol that computes f with error ρ
- $D_\rho^{\mu,n}(f)$, the communication involved in the best protocol that computes f on n **independent** pairs of input (x, y) drawn from μ, and getting the answer correct except an error ρ **on each coordinate**.
- Note that the above is different from $D_\rho^{\mu^n}(f^n)$, and

$$D_\rho^{\mu,n}(f) \leq D_\rho^{\mu^n}(f^n)$$

(Not direct sum but,) Information Equals Amortized Communication

- The **amortized communication complexity**

$$\lim_{n \to \infty} \frac{D_{\rho}^{\mu,n}(f)}{n}$$

- Information equals amortized communication complexity:

$$\lim_{n \to \infty} \frac{D_{\rho}^{\mu,n}(f)}{n} = IC_{\mu}^{i}(f)$$

Direct Sum Theorems

(Information Complexity Direct Sum) For every boolean function f, distribution μ,

$$IC_{\mu}^{n}(f) \geq n \cdot IC_{\mu}(f)$$

(Weak Direct Sum [BBCR10]) For every boolean function f, distribution μ, and any positive constant $\delta > 0$,

$$D_{\mu^n}(f^n, \epsilon) \geq \tilde{\Omega}(\sqrt{n} \cdot D_{\mu}(f, \epsilon + \delta))$$

Compression, IC

(Information Complexity Direct Sum) For every boolean function f, distribution μ,

$$IC^n_\mu(f) \geq n \cdot IC_\mu(f)$$

(Theorem 3.17 in [BR11]) For every μ, f, n, let π be a protocol realizing $D^{\mu,n}_\rho(f)$. Then there exists a protocol τ computing f with error ρ on inputs drawn from μ such that $CC(\tau) = CC(\pi)$, and $IC^i_\mu(\tau) \leq \frac{IC^i_{\mu^n}(\pi)}{n} \leq \frac{D^{\mu,n}_\rho(f)}{n}$ $(\leq \frac{D^{\mu^n}_\rho(f^n)}{n})$

Compression, CC

(Weak Direct Sum [BBCR10]) For every boolean function f, distribution μ, and any positive constant $\delta > 0$,

$$D_{\mu^n}(f^n, \epsilon) \geq \tilde{\Omega}(\sqrt{n} \cdot D_\mu(f, \epsilon + \delta))$$

(Interactive compression according to internal IC [BBCR10]) over any distribution μ on $X \times Y$, for every $\epsilon > 0$, π can be simulated with a protocol τ of length

$$O\Big(\sqrt{IC_\mu^i(\pi) \cdot CC(\pi)} \frac{\log(CC(\pi)/\epsilon)}{\epsilon}\Big),$$

and $\tau(X, Y) = \pi(X, Y)$ w.h.p.

Proving Information Complexity Direct Sum: Notations

Given a function $f(x, y)$ and a distribution μ on inputs to f

- The communication complexity $D_\rho^\mu(f)$, maximum number of bits communicated by a protocol that computes f with error ρ
- $D_\rho^{\mu,n}(f)$, the communication involved in the best protocol that computes f on n **independent** pairs of input (x, y) drawn from μ, and getting the answer correct except an error ρ **on each coordinate**.
- Note that the above is different from $D_\rho^{\mu^n}(f^n)$, and

$$D_\rho^{\mu,n}(f) \leq D_\rho^{\mu^n}(f^n)$$

Proving Information Complexity Direct Sum

(Theorem 3.17 in [BR11]) For every μ, f, n, let π be a protocol realizing $D_\rho^{\mu,n}(f)$. Then there exists a protocol τ computing f with error ρ on inputs drawn from μ such that $CC(\tau) = CC(\pi)$, and $IC_\mu^i(\tau) \leq \frac{IC_{\mu^n}^i(\pi)}{n} \leq \frac{D_\rho^{\mu,n}(f)}{n}$ $(\leq \frac{D_\rho^{\mu^n}(f^n)}{n})$

[Intuition] given a "more powerful" protocol, construct a new protocol that preserves the CC but saves IC by a factor of n.

Proving Information Complexity Direct Sum

(Theorem 3.17 in [BR11]) For every μ, f, n, let π be a protocol realizing $D_\rho^{\mu,n}(f)$. Then there exists a protocol τ computing f with error ρ on inputs drawn from μ such that $CC(\tau) = CC(\pi)$, and $IC_\mu^i(\tau) \leq \frac{IC_{\mu^n}^i(\pi)}{n} \leq \frac{D_\rho^{\mu,n}(f)}{n}$ $(\leq \frac{D_\rho^{\mu^n}(f^n)}{n})$

Proof.
First let us assume that π only uses private randomness (can easily extend to cover public randomness case). The new protocol $\tau(x,y)$ is defined as follows:

Proving Information Complexity Direct Sum

(Theorem 3.17 in [BR11]) For every μ, f, n, let π be a protocol realizing $D_\rho^{\mu,n}(f)$. Then there exists a protocol τ computing f with error ρ on inputs drawn from μ such that $CC(\tau) = CC(\pi)$, and $IC_\mu^i(\tau) \leq \frac{IC_{\mu^n}^i(\pi)}{n} \leq \frac{D_\rho^{\mu,n}(f)}{n}$ $(\leq \frac{D_\rho^{\mu^n}(f^n)}{n})$

Proof.

- the parties publicly sample J uniformly at random from [n]. J is understood as an index.
- The parties publicly sample $X_1, ..., X_{J-1}$ and $Y_{J+1}, ..., Y_n$.
- The first party privately samples $X_{J+1}, ..., X_n$ conditioned on the corresponding Y's; The second party does similar.
- The parties run the old protocol π on $X_1, ..., X_n, Y_1, ..., Y_n$ and output the result computed for the J'th coordinate. (i.e. viewing $X_J = x,\ Y_J = y$)

Proving Information Complexity Direct Sum

(Theorem 3.17 in [BR11]) For every μ, f, n, let π be a protocol realizing $D_\rho^{\mu,n}(f)$. Then there exists a protocol τ computing f with error ρ on inputs drawn from μ such that $CC(\tau) = CC(\pi)$, and $IC_\mu^i(\tau) \leq \frac{IC_{\mu^n}^i(\pi)}{n} \leq \frac{D_\rho^{\mu,n}(f)}{n}$ $(\leq \frac{D_\rho^{\mu^n}(f^n)}{n})$

Proof.

Analyze the protocol: observe CC and bounded error: $CC(\pi) = CC(\tau)$, and error is bounded by ρ.

Proving Information Complexity Direct Sum

(Theorem 3.17 in [BR11]) For every μ, f, n, let π be a protocol realizing $D_\rho^{\mu,n}(f)$. Then there exists a protocol τ computing f with error ρ on inputs drawn from μ such that $CC(\tau) = CC(\pi)$, and $IC_\mu^i(\tau) \leq \frac{IC^i_{\mu^n}(\pi)}{n} \leq \frac{D_\rho^{\mu,n}(f)}{n}$ $\left(\leq \frac{D_\rho^{\mu^n}(f^n)}{n}\right)$

Proof.
Analyze the protocol: bound $IC_\mu^i(\tau) = I(X;\tau|Y) + I(Y;\tau|X)$.
NOTE: X, Y are r.v. for τ's inputs (sampled according to μ).
Let's bound the first term:

$$
\begin{aligned}
I(X:\tau|Y) &\leq I(X:\tau Y_1\cdots Y_n|Y)\\
&= I(X;JX_1\cdots X_{J-1}Y_1\cdots Y_n\pi|Y)\\
&= I(X;JX_1\cdots X_{J-1}Y_1\cdots Y_n|Y) + I(X_J;\pi|JX_1\cdots X_{J-1}Y_1\cdots Y_n)\\
&= I(X_J;\pi|JX_1\cdots X_{J-1}Y_1\cdots Y_n)
\end{aligned}
$$

Proving Information Complexity Direct Sum

(Theorem 3.17 in [BR11]) For every μ, f, n, let π be a protocol realizing $D_\rho^{\mu,n}(f)$. Then there exists a protocol τ computing f with error ρ on inputs drawn from μ such that $CC(\tau) = CC(\pi)$, and $IC_\mu^i(\tau) \leq \frac{IC_{\mu^n}^i(\pi)}{n} \leq \frac{D_\rho^{\mu,n}(f)}{n}$ $(\leq \frac{D_\rho^{\mu^n}(f^n)}{n})$

Proof.
Expanding the expectation according to J, apply Chain Rule:

$$I(X;\tau|Y) \leq (1/n)\sum_{j=1}^{n} I(X_j;\pi|X_1\cdots X_{j-1}Y_1\cdots Y_n)$$
$$= I(X_1\cdots X_n;\pi|Y_1\cdots Y_n)/n$$

Similarly we can bound $I(Y;\tau|X) \leq I(Y_1\cdots Y_n;\pi|X_1\cdots X_n)/n$, and thus $\mathrm{IC}_\mu^i(\tau) \leq \mathrm{IC}_{\mu^n}^i(\pi)/n \leq \mathrm{CC}(\pi)/\mathrm{n}$ □

Interactive Compression

- Given a protocol π with low *information complexity*, can we get another protocol with lower *communication* and slightly more error?

- Yes, by simulating π while sending less bits. The cost is a small chance of error.
  [BBCR10]: over any distribution μ on $X \times Y$, for every $\epsilon > 0$, π can be simulated with a protocol τ of length

$$O\Big(\sqrt{IC_\mu^i(\pi) \cdot CC(\pi)}\frac{\log(CC(\pi)/\epsilon)}{\epsilon}\Big),$$

  and $\tau(X, Y) = \pi(X, Y)$ w.h.p.

- Sufficient to prove direct sum result, but stronger result exists (external IC).

Compression Proof Idea

- In τ, Alice and Bob privately guess π's transcript $M = m_1 m_2 \cdots m_C$ without communicating. Then communicate with few bits to correct their guesses.

- Alice will come up with $m_1^A, m_2^A, \ldots m_C^A$,
  Bob will come up with $m_1^B, m_2^B, \ldots m_C^B$.

- Once $m_i^A = m_i^B = m_i$ they can output $\pi(X, Y)$. How do Alice and Bob guess M?

- For each prefix $m_{<i}$ of bits sent in π, let

$$\gamma(m_{<i}) = p(M_i = 1|xym_{<i}).$$

  These numbers are how the messages in are distributed in $\pi(X, Y)$.

- How to sample from this distribution:
    - Use (public) randomness to get $\rho_1, \ldots, \rho_C \sim \mathrm{Unif}([0, 1])$.
    - set $m_1 = 1$ iff $\rho_1 < \gamma(m_{<1}) = p(M_1 = 1|xy)$,
    - set $m_2 = 1$ iff $\rho_2 < \gamma(m_{<2}) = p(M_2 = 1|xym_1)$,

      $\vdots$
    - set $m_C = 1$ iff $\rho_C < \gamma(m_{<C}) = p(M_C = 1|xym_{<C})$.

- If Alice, Bob sampled this way, they would have successfully simulated π.

- The problem: Alice does not have y, so does not know the value of $\gamma(m_{<i})$. Similarly, Bob is missing x...

- Key insight: if Alice communicates first in π, she knows $\gamma(m_{<1})$

$$\gamma(m_{<1}) = p(M_1 = 1|xy) = p(M_1 = 1|x)$$

since the first bit sent has no dependence on Bob's secret y.

- In general, if Alice speaks next in π and she knows $m_{<i}$, then she knows the value of

$$p(M_i = 1|xym_{<i}) = p(M_i = 1|xm_{<i})$$

Likewise, if Bob speaks next and knows $m_{<i}$, then he knows the value of

$$p(M_i = 1|xym_{<i}) = p(M_i = 1|ym_{<i})$$

and can sample correctly.

- Let

$$\gamma^A(m_{<i}) = p(M_i = 1|xm_{<i})$$
$$\gamma^B(m_{<i}) = p(M_i = 1|ym_{<i})$$

- In τ: Alice computes

$$m_1^A = 1 \iff \rho_1 < \gamma^A(m_{<1}),$$
$$m_2^A = 1 \iff \rho_2 < \gamma^A(m_{<2}),$$
$$\vdots$$
$$m_C^A = 1 \iff \rho_C < \gamma^A(m_{<C})$$

  Bob computes

$$m_1^B = 1 \iff \rho_1 < \gamma^B(m_{<1}),$$
$$\vdots$$
$$m_C^B = 1 \iff \rho_C < \gamma^B(m_{<C})$$

- Alice will sample M correctly, up until the first time $\gamma^A(m_{<i}) \neq \gamma(m_{<i})$ (when Bob speaks for the first time).
- Bob will sample M correctly, up until the first time $\gamma^B(m_{<i}) \neq \gamma(m_{<i})$ (when Alice speaks for the first time).
- Alice and Bob communicate to find the first i where $m_i^A \neq m_i^B$. Who is right?
    - If the ith bit is sent by Alice, m_i^A is sampled correctly.
    - If ith bit sent by Bob, m_i^B is sampled correctly.
- Whoever is wrong: correct their ith bit and recompute their guess. Repeat until $m^A = m^B$.

- How many bits must Alice and Bob communicate to find first i where m^A, m^B disagree?
- $O(\log C/\delta)$ bits using binary search + hashing, if probability of error is $\delta > 0$.
- By union bound, total error is at most $C\delta = \epsilon/2$. $O(\log(C/\epsilon))$ bits sent for each mistake i.

- Remains to bound the number of corrections Alice, Bob will have to make.
- Will see that $\mathbb{E}[\# \text{ mistakes made}] \leq \sqrt{I \cdot C}$

$$\implies \mathbb{E}[\text{length of } \tau] \leq O(\sqrt{IC} \cdot \log(C/\epsilon)).$$

- By Markov's inequality,

$$\Pr\left(|\tau| > \frac{2}{\epsilon} \cdot O(\sqrt{IC} \cdot \log(C/\epsilon))\right) \leq \epsilon/2.$$

With prob. $\geq 1 - \epsilon$, τ will simulate π correctly and have desired communication.

- What is the probability that Alice, Bob made the first mistake at i?
- Both have $m_{<i}$ sampled correctly, and ρ_i falls between $\gamma^A(m_{<i})$ and $\gamma^B(m_{<i})$.
- So probability of mistake at i is at most

$$\mathbb{E}_{xym}[|\gamma^A(m_{<i}) - \gamma^B(m_{<i})|]$$
$$\leq \mathbb{E}_{xym}[|p(m_i = 1|xm_{<i}) - p(m_i = 1|ym_{<i})|].$$

- Useful fact relating mutual information and independence: if A, B are random variables, then

$$\mathbb{E}_{b \sim B}[|p(a|b) - p(a)|] \leq \sqrt{I(A:B)}.$$

- If $I(A:B) = I(B:A)$ is small, then $p(a|b) \approx p(a)$ on average.

- Say Alice sends the ith bit in π. Fixing over $m_{<i}$,

$$\begin{aligned} &\mathbb{E}_{xym_{<i}}[|p(m_i = 1|xm_{<i}) - p(m_i = 1|ym_{<i})|] \\ &= \mathbb{E}_{xym_{<i}}[|p(m_i = 1|xym_{<i}) - p(m_i = 1|ym_{<i})|] \\ &\leq \sqrt{I(M_i : X|Ym_{<i})} \\ &= \sqrt{I(X : M_i|Ym_{<i})}. \end{aligned}$$

If Bob sends the ith bit, we get $\leq \sqrt{I(Y : M_i|Xm_{<i})}$

- An upper bound the expected number of corrections made:

$$\sum_{i=1}^{C} \sqrt{I(X : M_i|YM_{<i}) + I(Y : M_i|XM_{<i})}$$

$$\sum_{i=1}^{C} \sqrt{I(X : M_i | Y M_{<i}) + I(Y : M_i | X M_{<i})}$$

$$\leq \sqrt{C} \cdot \sqrt{\sum_{i=1}^{C} I(X : M_i | Y M_{<i}) + I(Y : M_i | X M_{<i})}$$

by Cauchy-Shwarz;

$$= \sqrt{C} \cdot \sqrt{I(X : M | Y) + I(Y : M | X)} = \sqrt{IC}.$$

□

Intuition for Compression

- If $IC^i_\mu(\pi)$ is small, then Alice doesn't need to know Bob's y to get a good idea for what M is. Same for Bob.
- Small $IC^i_\mu(\pi)$ means $m^A \approx m$ and $m \approx m^B$, as seen in proof.
- NOT guaranteed to give us lower communication. In fact, this is weak.
- Also in [BBCR10] can simulate π such that

$$CC(\tau) \leq O\Big(IC^o_\mu(\pi)\frac{\log(CC(\pi)/\epsilon)}{\epsilon^2}\Big).$$

Almost $CC(\tau) \leq O(IC(\pi))$!

Using Compression to Prove Direct Sum Lower Bound

- Let's show that

$$CC(T^n) = \tilde{\Omega}(\sqrt{n} \cdot CC(T)).$$

- Specifically, [BBCR10] for every $\epsilon > 0$,

$$R_\rho(f^n) \cdot \log(R_\rho(f^n)/\epsilon) \geq \Omega(R_{\rho+\epsilon}(f)\epsilon\sqrt{n}).$$

  Then apply min-max principle: $R_\rho(f) = \max_\mu D_\rho^\mu(f)$.

- Let π be any protocol for f^n on inputs drawn from μ^n with error prob. $\leq \rho$.

- Recall protocol for single copy f using randomness $R = (J, X_{<J}, Y_{>J})$, with

$$CC(\tau) \leq CC(\pi)$$
$$IC^i_\mu(\tau) \leq 2CC(\pi)/n.$$

- Compress τ with error ϵ to get a protocol for f with error $\rho + \epsilon$, communication

$$CC(\tau') \leq O\Big(\frac{CC(\pi)\log(CC(\pi)/\epsilon)}{\epsilon\sqrt{n}}\Big).$$

- τ' computes f: $CC(\tau') \geq R_{\rho+\epsilon}(f)$

- So for all π for f^n,

$$CC(\pi)\log(CC(\pi)/\epsilon) \geq \Omega(R_{\rho+\epsilon}(f)\epsilon\sqrt{n}).$$

□

Closing the Direct Sum Bound

- Is it possible to show that $CC(f^n) = \Theta(n \cdot CC(f))$?
- $CC(f^n) = O(n \cdot CC(f))$ is trivial.
- Lower bound: $CC(f^n) = \tilde{\Omega}(\sqrt{n} \cdot CC(f))$ (proved this).

Separation in IC and CC

- Answer: no. [GKR15] showed that there is a family of functions with information k and communication $2^{\Omega(k)}$.

- Amortized communication:

$$IC^i(T) = \lim_{n\to\infty} \frac{CC(T^n)}{n}$$

$CC(T) \geq 2^{\Omega(k)}$ but $CC(T^n) \approx nk$.

- Their T is played on a tree with $k \cdot 2^{100\cdot 4^k}$ layers, goal is to output a path from root to leaf satisfying Alice and Bob's inputs.

Rao and Sinha Easier Separation

- In [RS18], they show an exponential separation for the *k-ary pointer jumping function*:
  - Alice gets $X : [k]^{<n} \to [k]$ and $F : [k]^n \to [k]$.
  - Bob gets $Y : [k]^{<n} \to [k]$ and $G : [k]^n \to [k]$.
  - They have to find the unique $z \in [k]^n$ where for all $1 \le i < n$

$$X(z_{\le i}) + Y(z_{\le i}) = z_{r+1} \mod k,$$

and output $F(z) + G(z) \mod 2$.