* Achieving Compression \( \sim \sqrt{IC \cdot IC \cdot \text{poly}lg} \)

[Inevitably we need to eliminate lots of rounds in communication]

First, assume we wish to simulate the behavior of \( T \) with no communication. What's the best we can achieve?

\[
\begin{align*}
\text{Alice} & \rightarrow \\
\text{Bob} & \rightarrow \\
\text{Alice} & \rightarrow \\
\text{Bob} & \rightarrow
\end{align*}
\]

wlog, Alice & Bob alternate, each round transmits one bit and the tree is full.

* Node \( v \) at depth \( i \) is uniquely represented by a bit-string of length \( i \) corresponding to the transcript leading to \( v \).

* Suppose Alice owns a node \( v \)

Then, if \( T \) reaches \( v \), Alice chooses the right child w.p. \( P_A^v \).  
(1)
Bob doesn't know $p^V$, but he can produce a "best-guess":

$$p^v_B = \text{Prob}(\text{Alice chooses 1} \mid Y = y, \Pi \text{ reaches node } v)$$

(i.e., what Bob knows)

In general, $p^v_A$ and $p^v_B$ may be very different, and if this is so, Bob learns substantially from Alice's message, and the bit sent by Alice isn't "redundant".

Without any communication, Alice and Bob can compute $p^v_A$ and $p^v_B$ for all nodes, and accordingly, sample a root-leaf path.

**Problem:**

1) $p^v_A$ and $p^v_B$ different, means the distribution with the paths would be different.

2) Even if prob. estimates are close, Alice and Bob must coordinate to sample "similar" paths.
For "problem 2", Alice and Bob can use public randomness for correlated sampling.

**Correlated Sampling:** Alice and Bob are given parameters $p, q \in [0,1]$, and want to sample

- Alice: $a$
- Bob: $b$

s.t. $\Pr(a = 1) = p$, $\Pr(b = 1) = q$,

- $|p - q|$ small $\Rightarrow \Pr(a \neq b)$ small.

**Solution:**

1) Using public randomness, sample $\rho \in [0,1]$ u.a.r.

2) Alice outputs $a = 1$ iff $\rho \leq p$.

3) Bob outputs $b = 1$ iff $\rho \leq q$.

Observe $\Pr(a \neq b) = |p - q|$.

Alice and Bob use correlated sampling at each node to create a root-leaf path (using their estimates).
N.B.: At each node, one either Alice or Bob samples "correctly".

* If all estimates are the same, Alice & Bob both sample the correct path with no communication (i.e., perfectly simulate Π).

Good luck, Alice & Bob!

After sampling root-leaf paths, Alice and Bob find the earliest point where the paths differ.

* Problem: Find the first disagreement.

Alice and Bob are given $k$-bit strings $x$ and $y$.

Goal: Output the smallest $i$ where $x_i \neq y_i$, or 0 if $x = y$. (error ≤ $\varepsilon$)

* Binary search and hashing-based equality tests, using public coin: $CC = O\left( \log k \log \left( \frac{\log k}{\varepsilon} \right) \right)$.

* More careful idea leads to $CC = O\left( \log \frac{k}{\varepsilon} \right)$.

⇒ Each inconsistency can be found using $O\left( \frac{\log CC(\Pi)}{\varepsilon} \right)$ bits of communication.
* After fixing the first inconsistency, Alice and Bob resample the rest of the path using the updated information and repeat the disagreement test.

* The result follows if we show that, at any node \( v \),

\[
\Pr(\text{disagreement, Alice and Bob guess differently}) \leq \sqrt{\sum_{\pi} \frac{IC(\pi)}{|\pi|}}. \quad (\text{idea: Information theoretic inequalities + Cauchy Schwarz})
\]

**Def.** \( E_i = 1 \) iff Alice and Bob disagree at level \( i \). 

**NB.** * \( E_i \) happens at node \( v \) exactly when

\[
\min\left\{ p_v^A, p_v^B \right\} \leq p \leq \max\left\{ p_v^A, p_v^B \right\}.
\]

(Fix public randomness for now)

\[
\Rightarrow \quad \mathbb{E}\left( E_i \right) = \mathbb{E}_{x y v_{i-1} v_i} \left| \frac{p_v A - p_v B}{v_{i-1}} \right|
\]

\( (v = \text{correct path of } \pi) \)

\[
\leq \mathbb{E}_{x y v_{i-1} v_i} \Delta_{TV}(V_i | X V < i, V_i | Y V < i)
\]

\[
\leq \mathbb{E}_{x y v_{i-1} v_i} \Delta_{TV}(V_i | X V < i, V_i | X Y V < i)
\]

\[
+ \Delta_{TV}(V_i | Y V < i, V_i | X Y V < i)
\]
\[
\begin{align*}
\text{(Note: } \Delta_{TV}(p,q) \leq \sqrt{D(p\|q)}) \quad \Rightarrow \quad \sqrt{\mathbb{E} \sum_{xy} \mathbb{D}(V_i | xy, V < i \parallel V_i | x, V < i)} + 2 \\
\quad \leq \sqrt{\mathbb{E} \sum_{xy} \mathbb{D}(V_i | xy, V < i \parallel V_i | y, V < i)} \\
\text{Jensen} \quad \leq \sqrt{\mathbb{E} \sum_{xy} \mathbb{D}(m \parallel m)} \\
\quad \leq \sqrt{I(V_i; Y_i | X, V < i) + I(V_i; X_i | Y, V < i)} \\
\Rightarrow \quad \mathbb{E}(\# \text{ disagreements}) \leq \sum_i \sqrt{I(m) + I(m)} \\
\text{Cauchy-Schwarz} \quad \leq \sqrt{|\Pi| \sum_i (I(m) + I(m))} \\
\text{Chain Rule} \quad = \sqrt{|\Pi| (I(V_i; Y_i | X) + I(V_i; X_i | Y))} \\
= \sqrt{|\Pi| \cdot IC(\Pi)} \\
[\text{need to take } \mathbb{E} \text{ over public coin, to be precise.}]
\end{align*}
\]