Communication Complexity of Randomness Manipulation

Madhu Sudan
Harvard University

Joint work with Mitali Bafna (Harvard), Badih Ghazi (Google), and Noah Golowich (Harvard)
Randomness Processing Industry

- Dispersers, Extractors, Merges, Condensers, PRGs ...

![Diagram: X → E → E(X)]

- Interest not in the exact function. (E=RSA most boring.)
- Rather in its manipulation of distributions ...
  - Distribution of $X$ vs. that of $E(X)$
- Long history ... omitted. Key ingredients
  - single processor “$E$”
  - unknown source $X \in \mathcal{X}$
Distributed Randomness Processing

- **2-player setting:**

  \[ X_1, X_2, \ldots, X \rightarrow Alice \rightarrow U \]

  \[ Y_1, Y_2, \ldots, Y \rightarrow Bob \rightarrow V \]

  Infinite sequence \( \leq C \) bits, \( \leq r \) Rounds

  \( (X,Y) \sim P \) \quad \text{Goal: } (U,V) \approx_{\delta} Q

- Is \( \delta = 0 \) possible? If not minimize \( \delta \)! Etc.
- Alice + Bob can use **private randomness**
- Zero communication version: “Non-Interactive Simulation” (NIS).
A classical example

- **Zero communication**
  - \((X,Y) \sim \text{Unif}(((0,0), (0,1), (1,0)))\)
  - \((U,V) \sim \rho\)-correlated bits.
    - \(U,V \sim \text{Unif}([0,1])\), \(\Pr[U = V] = \frac{(1+\rho)}{2}\)

- **[Witsenhausen ‘70s]**:
  - Can’t achieve \(\rho = 1\). Not even \(\rho = .51\)
  - (Naïve strategy achieves \(\rho = \frac{1}{4}\))
  - Can achieve \(\rho = \frac{1}{3}\) by a general method.
  - Best \(\rho\)? Open!!
Intermediate Problem: Output \((A, B) \) w. \(A, B \sim N(0,1)\) with maximum correlation.

Definition: Max-Corr \( \rho(X, Y) \triangleq \max_{f,g} \mathbb{E}_{X,Y} [f(X)g(Y)] \)

Where \( f, g: \Omega \to \mathbb{R} \),
\[
\begin{align*}
\mathbb{E}_X[f(X)] &= \mathbb{E}_Y[g(Y)] = 0 \\
\mathbb{E}_X[f(X)^2] &= \mathbb{E}_Y[g(Y)^2] = 1
\end{align*}
\]

Thm 1: For \( (X, Y) \), can achieve \( \rho(U, V) \leq \rho(X, Y) \)

Thm 2: For Gaussian output \((A, B)\), can achieve \( \rho(A, B) = \rho(X, Y) \)

Thm 3: For binary \((U, V)\), can achieve
\[
1 - \frac{2 \cdot \cos^{-1}(\rho(X, Y))}{\pi} \leq \rho(U, V) \leq \rho(X, Y)
\]

---

This result was stated as Theorem 6.2 in [9]. However the proof given in [9] is incorrect. It hinges on a chain of inequalities of which the last actually holds in the direction opposite to the one asserted.
"Tensorization"

- **“Single-letter” Problems**: Given (1) structure $S$ (graph, distribution, game) (2) product operation $S^\otimes t$, (3) measure $M$, compute $\lim_{t\to\infty} M(S^\otimes t)$
  - Shannon capacity, Compression length, Channel capacity, parallel repetition value of 2-prover game, Direct sum complexity, NIS

- **“Tensorizing bound”**: 
  - Typically: $M(S^\otimes t) \geq M(S)$.
  - Find: $\mathcal{M}(\cdot)$ s.t. $M(S) \leq \mathcal{M}(S)$ and $\mathcal{M}(S^\otimes t) = \mathcal{M}(S)$

- **Max-Correlation “Tensorizes”**
  - If $(X,Y) \sim \mu$ & $(X^t,Y^t) \sim \mu^\otimes t$ then $\rho(X^t,Y^t) = \rho(X,Y)$
  - Proof idea: $\mu$ described by matrix $P$; $\rho$ related to its singular values, $\mu^\otimes t$ described by $P^\otimes t$
# Single-letter Embarrassment

<table>
<thead>
<tr>
<th>Problem</th>
<th>Base</th>
<th>Single-letter version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compressibility</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>Channel capacity</td>
<td>NP</td>
<td>P</td>
</tr>
<tr>
<td>Independent set/Shannon capacity</td>
<td>NP</td>
<td>[0, ∞]</td>
</tr>
<tr>
<td>Value of 2-prover game</td>
<td>NP</td>
<td>[NP, ∞]</td>
</tr>
<tr>
<td>Non-interactive-simulation (“correlation”)</td>
<td>NP</td>
<td>[0, CA]</td>
</tr>
<tr>
<td>Communication Complexity</td>
<td>[NP, Exp?]</td>
<td>[0, CA]</td>
</tr>
<tr>
<td>?</td>
<td>P</td>
<td>[NP, ∞]</td>
</tr>
</tbody>
</table>

**Glossary:** $0 \leq P \leq NP \leq EXP \leq \text{Computable} \leq CA (\text{Computably Approximable}) \leq ∞$
Complexity of NIS

“Computably approximable”

- [Ghazi,Kamath,S.], [De,Mossel,Neeman]², [Ghazi,Kamath,Raghavendra]
- There exists a finite time algorithm determining if we can get $\epsilon$-close to $(U,V)$ from i.i.d. samples of $(X,Y)$

Idea: “Invariance Priniciple” [Mossel]

- Either there is a method using few samples of $(X,Y)$ or many samples $(X_1,Y_1), ..., (X_n,Y_n)$ each with low influence on $(U,V)$.
- If latter, can replace $(X,Y)$ by finite # Gaussians.
- Gives computable upper bound on #samples needed to get close to optimal strategy.
Restrict to “Common Randomness Generation” ($k$-CRG): Output ($U = V$), with $U \sim \text{Unif}([0,1]^k)$

First considered ~70 [Ahlsedwee-Csizar]:

“Characterization”: Can communicate $R \cdot k$ + $o(k)$ bits, where $R = \min_{\Pi} \frac{\text{IC}_{\text{int}}(\Pi)}{\text{IC}_{\text{ext}}(\Pi)}$

Computable? Computably approximable?

For one-way communication: $R = 1 - \rho(X,Y)^2$ [Zhao-Chia] (see also [Guruswami-Radhakrishnan, Ghazi-Jayram, S.-Tyagi-Watanabe])

Oct. 31, 2018  Simons: Comm. Comp. Randomness Manipulation
This talk: Rounds in CRG

- Does increasing #rounds ease CRG? (For some \((X, Y)\)?) ... Is the following true?
  - Question: \(\forall \varepsilon > 0, r, \exists (X, Y) \text{ s.t. } \forall k\)
    - \(\exists (r, \varepsilon k)\)-protocol for \(k\)-CRG\((X, Y)\)
    - No \((r - 1, (1 - \varepsilon).k)\)-protocol for \(k\)-CRG\((X, Y)\)
- Still don’t know. Partial progress.
- Thm. [BGGS’19] : \(\forall n, k, r\) there exists \((X, Y)\) s.t.
  - \(\exists (r, O(r \log n))\)-protocol for \(k\)-CRG\((X, Y)\).
  - No \(\left(\frac{r}{2} - 3, \min \left\{ k, \frac{n}{\text{polylog } n} \right\} \right)\)-protocol for \(k\)-CRG\((X, Y)\)
Our Source:

\[ i_0 \sim [n] \]
\[ \pi_1, \ldots, \pi_n \sim S_n \]
\[ A_1, \ldots, A_n \sim \{0,1\}^k \]
\[ B_1, \ldots, B_n \sim \{0,1\}^k \]
\[ A_i = B_i \]

where \( i_j = \pi_j(i_{j-1}) \)

“Pointer Chasing Source” (PCS)
CRG from Pointer Chasing Source

- Easy direction easy: \((r, r \log n)\)-protocol for \(k\)-CRG
- Hardness?
  - Pointer chasing is hard: [Duris-Galil-Schnitger, Nisan-Wigderson,...].
  - CRG from PCS requires pointer chasing?
    - No! Can solve problem without pointers!
    - Small non-deterministic complexity!
    - Hardness needs to use hardness of disjointness? And of pointer-chasing?
    - How to combine modularly?
Our solution: Pointer Verification Problem

- **PVP:** Distinguish \((X, Y) \sim PV_{\text{YES}}\) from \((X, Y) \sim PV_{\text{NO}}\)
  - **PV_{\text{YES}}:** \(X = (\pi_1, \pi_3, \ldots, \pi_{r-1});\ Y = (i_0, i_r; \pi_2, \pi_4, \ldots, \pi_r)\)
  - **PV_{\text{NO}}:** \(X = (\pi_1, \pi_3, \ldots, \pi_{r-1});\ Y = (i_0, i'_r; \pi_2, \pi_4, \ldots, \pi_r)\)
- **YES instance satisfy:** \(i_j = \pi_j(i_{j-1})\)
- **NO instance:** \(i'_r\) random

**Claims:**

- Hardness of (Unique) Set Disjointness \(\Rightarrow\) Protocol for \(k\)-CRG(PCS) solves PVP.
- No \(\left(\frac{r}{2} - O(1), \frac{n}{\text{polylog } n}\right)\)-protocol for PVP.
**PVP ≤ CRG**

- Let $\mu_{XY}$ denote dist. of PCS source. $\mu_X, \mu_Y$ marginals
- CRG Hard $\iff \mu_{XY}$ indist. from $\mu_X \times \mu_Y$ (to $\left( \frac{r}{2}, C \right)$-protocols)
- Intermediate distribution: $\mu_{\text{mid}}$
  - $X = (\pi_{\text{odd}}, A_1, \ldots, A_n)$; $Y = (i_0, \pi_{\text{even}}, B_1, \ldots, B_n)$, with $A_j = B_j$ for random $j$.
  - (Correlation exists but pointers don't point to it.)
- Claims:
  - $\mu_{XY}$ indist. from $\mu_{\text{mid}}$ $\iff$ PVP is hard
  - $\mu_{\text{mid}}$ indist. from $\mu_X \times \mu_Y$ $\iff$ (Unique) disjointness hard.
Main idea: “Round elimination” a la NW ’93
- No black-box use 😞
- No simple variant 😞

Induction on #rounds
- Many invariants ... roughly
  - \( H(\pi_{\text{odd}}, \pi_{\text{even}} \mid \text{transcript}) \approx \text{Max} - O(C) \)
  - \( H(i_t \mid \text{transcript}) \approx \log n - o(1) \)
  - \( X \indep Y \mid \text{path, transcript} \)
Conclusions

- Many interesting problems in distributed randomness manipulation.
- Complexity of the “Single-letter characterizations”.
- Tight characterization of round complexity of CRG.
Thank You!