Imperfectly Shared Randomness in Communication

Madhu Sudan
Microsoft Research

Joint work with Clément Canonne (Columbia), Venkatesan Guruswami (CMU) and Raghu Meka (UCLA).
Communication Complexity

The model (with shared randomness)

\[ f : (x, y) \mapsto \Sigma \]

\[ R = $$$ \]

\[ CC(f) = \# \text{ bits exchanged by best protocol} \]

\[ f(x, y) \text{ w.p. } 2/3 \]
Communication Complexity: Motivation

- **Lower bounds:**
  - Circuit complexity, Streaming, Data Structures, extended formulations ...

- **Upper bounds?**
  - What is the right model for Communication (e.g., this talk)? - Shannon’48 or Yao’79?
    - If you wish to reproduce this talk ...
      - Shannon ‘48
    - If goal is for you to learn something, or if we expect to use interaction ...
      - Yao ’79!!
Natural (Contextual) communication

- Communication among humans:
  - Large context.
  - (Small) uncertainty about context.
  - Short communications.
- Can we use CC to study such communication?
  - What are example problems?
  - What are reliability mechanisms?
  - How do you leverage small uncertainty about large context?
- What are examples of problems with small communication complexity?
Aside: Easy CC Problems

- Equality testing:
  \[ EQ(x, y) = 1 \iff x = y; \quad O \]

- Hamming distance:
  \[ H_k(x, y) = 1 \iff \Delta(x, y) \leq k; \quad \]

- Small set intersection:
  \[ \cap_k (x, y) = 1 \iff \text{wt}(x), \text{wt}(y); \quad \]
  \[ CC(\cap_k) = O(k) \quad \text{[Håstad Wigderson]} \]

- Gap (Real) Inner Product:
  \[ x, y \in \mathbb{R}^n; |x|_2, |y|_2 = 1; \]
  \[ GIP_c(x, y) = 1 \iff \langle x, y \rangle > c; \quad \]

Thanks to Badih Ghazi and Pritish Kamath
Uncertainty in Communication

- Overarching question: Are there communication mechanisms that can overcome uncertainty?

- What is uncertainty? Some possible models
  - Bob wishes to compute f. Alice only has “approximate” knowledge of f.
  - Alice & Bob’s inputs are strongly correlated.

- This talk: Alice, Bob don’t share randomness perfectly; only approximately.
Rest of this talk

- Model: Imperfectly Shared Randomness

- Positive results: Coping with imperfectly shared randomness.

- Negative results: Analyzing weakness of imperfectly shared randomness.
Model: Imperfectly Shared Randomness

- Alice ← $r$; and Bob ← $s$ where $(r, s_i) = \text{i.i.d. sequence of correlated pairs } (r_i, s_i)_i; \newline r_i, s_i \in \{-1, +1\}; \mathbb{E}[r_i] = \mathbb{E}[s_i] = 0; \mathbb{E}[r_i s_i] = \rho \geq 0 .$

- Notation:
  - $isr_\rho(f) = \text{cc of } f \text{ with } \rho\text{-correlated bits.}$
  - $cc(f): \text{Perfectly Shared Randomness cc. } = isr_1(f)$
  - $priv(f): \text{cc with PRIVate randomness } = isr_0(f)$

- Starting point: for Boolean functions $f$
  - $cc(f) \leq isr_\rho(f) \leq priv(f) \leq cc(f) + \log n$
  - What if $cc(f) \ll \log n? \ E.g. \ cc(f) = O(1)$

$03/04/2015 \hspace{10em} \text{TCS+: ISR in Communication}$
Results

- Model first studied by [Bavarian,Gavinsky,Ito’14] ("Independently and earlier").
  - Their focus: Simultaneous Communication; general models of correlation.
  - They show $isr(Equality) = O(1)$ (among other things)

- Our Results:
  - Generally: $cc(f) \leq k \Rightarrow isr(f) \leq 2^k$
  - Converse: $\exists f$ with $cc(f) \leq k \land isr(f) \geq 2^k$
Equality Testing (our proof)

- Key idea: Think inner products.
  - Encode \( x \mapsto X = E(x); y \mapsto Y = E(y); X,Y \in \{-1, +1\}^N \)
    - \( x = y \Rightarrow \langle X, Y \rangle = N \)
    - \( x \neq y \Rightarrow \langle X, Y \rangle \leq N/2 \)

- Estimating inner products:
  - Building on sketching protocols ...
  - Alice: Picks Gaussians \( G_1, \ldots, G_t \in \mathbb{R}^N \),
  - Sends \( i \in [t] \) maximizing \( \langle G_i, X \rangle \) to Bob.
  - Bob: Accepts iff \( \langle G'_i, Y \rangle \geq 0 \)
  - Analysis: \( O_\rho(1) \) bits suffice if \( G \approx_\rho G' \)

Gaussian Protocol
General One-Way Communication

- Idea: All communication $\leq$ Inner Products
- (For now: Assume one-way-cc$(f) \leq k$)
  - For each random string $R$
    - Alice’s message $= i_R \in [2^k]$
    - Bob’s output $= f_R(i_R)$ where $f_R : [2^k] \rightarrow \{0,1\}$
    - W.p. $\geq \frac{2}{3}$ over $R$, $f_R(i_R)$ is the right answer.
General One-Way Communication

- For each random string $R$
  - Alice’s message $= i_R \in [2^k]$
  - Bob’s output $= f_R(i_R)$ where $f_R : [2^k] \to \{0,1\}$
  - W.p. $\geq \frac{2}{3}$, $f_R(i_R)$ is the right answer.

- Vector representation:
  - $i_R \mapsto x_R \in \{0,1\}^{2^k}$ (unit coordinate vector)
  - $f_R \mapsto y_R \in \{0,1\}^{2^k}$ (truth table of $f_R$).
  - $f_R(i_R) = \langle x_R, y_R \rangle$; Acc. Prob. $\propto \langle X, Y \rangle; X = (x_R)_R; Y = (y_R)_R$
  - Gaussian protocol estimates inner products of unit vectors to within $\pm \epsilon$ with $O_{\rho} \left( \frac{1}{\epsilon^2} \right)$ communication.
Two-way communication

- Still decided by inner products.

- Simple lemma:
  - \( \exists K_A^k, K_B^k \subseteq \mathbb{R}^{2^k} \) convex, that describe private coin k-bit comm. strategies for Alice, Bob s.t. accept prob. of \( \pi_A \in K_A^k, \pi_B \in K_B^k \) equals \( \langle \pi_A, \pi_B \rangle \)

- Putting things together:

Theorem: \( cc(f) \leq k \Rightarrow isr(f) \leq O_\rho(2^k) \)
Main Technical Result: Matching lower bound

Theorem: There exists a (promise) problem $f$ s.t.
- $cc(f) \leq k$
- $isr_\rho(f) \geq \exp(k)$

The Problem:
- Gap Sparse Inner Product (G-Sparse-IP).
- Alice gets sparse $x \in \{0,1\}^n$; $wt(x) \approx 2^{-k} \cdot n$
- Bob gets $y \in \{0,1\}^n$
- Promise: $\langle x, y \rangle \geq (.9)2^{-k} \cdot n$ or $\langle x, y \rangle \leq (.6)2^{-k} \cdot n$.
- Decide which.
**Protocol for G-Sparse-IP**

- Note: Gaussian protocol takes $O(2^k)$ bits.
  - Need to get exponentially better.
- Idea: $x_i \neq 0 \Rightarrow y_i$ correlated with answer.
- Use (perfectly) shared randomness to find random index $i$ s.t. $x_i \neq 0$.
- Shared randomness: $i_1, i_2, i_3, \ldots$ uniform over $[n]$.
- Alice → Bob: smallest index $j$ s.t. $x_{ij} \neq 0$.
- Bob: Accept if $y_{ij} = 1$
- Expect $j \approx 2^k$; $cc \leq k$. 

---

G-Sparse-IP:

$x, y \in \{0, 1\}^n; \text{wt}(x) \approx 2^{-k} \cdot n$

Decide $\langle x, y \rangle \geq (.9) 2^{-k} \cdot n$

or $\langle x, y \rangle \leq (.6) 2^{-k} \cdot n$?
Towards a lower bound: Ruling out a natural approach

- Natural approach:
  - Alice and Bob use (many) correlated bits to agree perfectly on few random bits?
  - For G-Sparse-IP need $O(2^k \log n)$ random bits.

- Agreement Distillation Problem:
  - Alice & Bob exchange $t$ bits; generate $k$ random bits, with agreement probability $\gamma$.
  - Lower bound [Bogdanov, Mossel]:
    
    $$t \geq k - O\left(\log \frac{1}{\gamma}\right)$$

Towards Lower Bound

- Explaining two natural protocols:
  - Gaussian Inner Product Protocol:
    - Ignore sparsity and just estimate inner product.
    - Uses $\sim 2^{2k}$ bits. Need to prove it can’t be improved!

G-Sparse-IP:
$x, y \in \{0, 1\}^n; \text{wt}(x) \approx 2^{-k} \cdot n$
Decide $\langle x, y \rangle \geq (.9) 2^{-k} \cdot n$
or $\langle x, y \rangle \leq (.6) 2^{-k} \cdot n$?
Optimality of Gaussian Protocol

- **Problem:**
  - \((x, y) \leftarrow \mu^n: \mu = \mu_{YES} \text{ or } \mu_{NO} \text{ supported on } \mathbb{R} \times \mathbb{R}
    - \mu_{YES}: \epsilon\text{-correlated Gaussians}
    - \mu_{NO}: \text{uncorrelated Gaussians}

- **Lemma:** Separating \(\mu^n_{YES} \text{ vs. } \mu^n_{NO}\) requires \(\Omega(\epsilon^{-1})\) bits of communication.

- **Proof:** Reduction from Disjointness

- **Conclusion:** Can’t ignore sparsity!
Towards Lower Bound

- Explaining two natural protocols:
  - Gaussian Inner Product Protocol:
    - Ignore sparsity and just estimate inner product.
    - Uses $\sim 2^{2k}$ bits. Need to prove it can’t be improved!
  - Protocol with perfectly shared randomness:
    - Alice & Bob agree on coordinates to focus on:
      
      $$(i_1, i_2, ..., i_{2k}, ...)$$;
    - Either $i_1$ has high entropy (over choice of $r, s$)
      - Violates agreement distillation bound
    - Or has low-entropy:
      - Fix distributions of $x, y$ s.t. $x_{i_1} \perp y_{i_1}$

---

**G-Sparse-IP:**

$x, y \in \{0, 1\}^n; \text{wt}(x) \approx 2^{-k} \cdot n$

Decide $\langle x, y \rangle \geq (.9) 2^{-k} \cdot n$

or $\langle x, y \rangle \leq (.6) 2^{-k} \cdot n$?
Aside: Distributional lower bounds

- **Challenge:**
  - Usual CC lower bounds are distributional.
  - \( cc(G\text{-Sparse-IP}) \leq k, \quad \forall \text{ inputs}. \)
  - \( \Rightarrow cc(G\text{-Sparse-IP}) \leq k, \quad \forall \text{ distributions}. \)
  - \( \Rightarrow det\text{-}cc(G\text{-Sparse-IP}) \leq k, \quad \forall \text{ distributions}. \)
- So usual approach can’t work ...
  - Need to fix strategy first and then “identify” a hard distribution for the strategy ...

\[ G\text{-Sparse-IP}: \]
\[ x, y \in \{0, 1\}^n; wt(x) \approx 2^{-k} \cdot n \]

Decide \( \langle x, y \rangle \geq (.9) 2^{-k} \cdot n \)

or \( \langle x, y \rangle \leq (.6) 2^{-k} \cdot n? \)
Towards lower bound

- Summary so far:
  - Symmetric strategy $\Rightarrow 2^k$ bits of comm.
  - Strategy asymmetric; $x_1, y_1 \ldots x_k, y_k$ have high influence $\Rightarrow$ fix the distribution so these coordinates do not influence answer.
  - Strategy asymmetric; with random coordinate having high influence $\Rightarrow$ violates agreement lower bound.

- Are these exhaustive? How to prove this?
  - Invariance Principle!!
    - [Mossel, O’Donnell, Oleskiewisz], [Mossel] ...
ISR lower bound for GSIP.

- One-way setting (for now)
- Strategies: Alice $f_r(x) \in [K]$; Bob $g_s(y) \in \{0,1\}^K$;
- Distributions:
  - If $x_i, y_i$ have high influence on $(f_r, g_s)$ w.h.p. over $(r,s)$ then set $x_i = y_i = 0$. [i is BAD]
  - Else $y_i$ correlated with $x_i$ in YES case, and independent in NO case.
- Analysis:
  - $i \in BAD$ influential in both $f_r, g_s \Rightarrow$ No help.
  - $i \notin BAD$ influential ... $\Rightarrow$ violates agreement lower bound.
  - No common influential variable $\Rightarrow x, y$ can be replaced by Gaussians $\Rightarrow 2^k$ bits needed.

G-Sparse-IP: $x, y \in \{0,1\}^n; \text{wt}(x) \approx 2^{-k} \cdot n$
Decide $\langle x, y \rangle \geq (.) 2^{-k} \cdot n$ or $\langle x, y \rangle \leq (.) 2^{-k} \cdot n?$

03/04/2015 TCS+: ISR in Communication
Invariance Principle + Challenges

- Informal Invariance Principle: $f, g$ low-degree polynomials with no common influential variable
  \[ \Rightarrow \operatorname{Exp}_{x,y}[f(x)g(y)] \approx \operatorname{Exp}_{x,y}[f(X)g(Y)] \]
  (caveat $f \approx f; g \approx g$)
  - where $x, y$ Boolean $n$-wise product dist.
  - and $X, Y$ Gaussian $n$-wise product dist

- Challenges [+ Solutions]:
  - Our functions not low-degree [Smoothening]
  - Our functions not real-valued
    - $g: \{0,1\}^n \to \{0,1\}^\ell$: [Truncate range to $[0,1]^\ell$]
    - $f: \{0,1\}^n \to [\ell]$: [???, [work with $\Delta(\ell)$]]
Invariance Principle + Challenges

- Informal Invariance Principle: \( f, g \) low-degree polynomials with no common influential variable
  \[ \Rightarrow \text{Exp}_{x,y}[f(x)g(y)] \approx \text{Exp}_{X,Y}[f(X)g(Y)] \] (caveat \( f \approx f; g \approx g \))

- Challenges
  - Our functions not low-degree [Smoothening]
  - Our functions not real-valued [Truncate]
  - Quantity of interest is not \( f(x) \cdot g(y) \) ...
    - [Can express quantity of interest as inner product.]
  - ... (lots of grunge work ...)
  - Get a relevant invariance principle (next)
Invariance Principle for CC

Theorem: For every convex $K_1, K_2 \subseteq [-1,1]^\ell$

$\exists$ transformations $T_1, T_2$ s.t.
if $f: \{0,1\}^n \to K_1$ and $g: \{0,1\}^n \to K_2$

have no common influential variable, then

$F = T_1 f: \mathbb{R}^n \to K_1$ and $G = T_2 g: \mathbb{R}^n \to K_2$ satisfy

$\text{Exp}_{x,y}[(f(x), g(y))] \approx \text{Exp}_{x,y}[(F(X), G(Y))]$

- Main differences: $f, g$ vector-valued.
- Functions are transformed: $f \mapsto F; g \mapsto G$
- Range preserved exactly ($K_1 = \Delta(\ell); K_2 = [0,1]^\ell$!)
  - So $F, G$ are still communication strategies!
Summarizing

- $k$ bits of comm. with perfect sharing
  $\rightarrow 2^k$ bits with imperfect sharing.
- This is tight
- Invariance principle for communication
  - Agreement distillation
  - Low-influence strategies

G-Sparse-IP:
$x, y \in \{0, 1\}^n; wt(x) \approx 2^{-k} \cdot n$

Decide $\langle x, y \rangle \geq (0.9) 2^{-k} \cdot n$
  or $\langle x, y \rangle \leq (0.6) 2^{-k} \cdot n$?
Conclusions

- Imperfect agreement of context important.
  - Dealing with new layer of uncertainty.
  - Notion of scale (context LARGE)

- Many open directions+questions:
  - Imperfectly shared randomness:
    - One-sided error?
    - Does interaction ever help?
    - How much randomness?
    - More general forms of correlation?
Thank You!