Communication with imperfect shared randomness

Venkatesan Guruswami

Carnegie Mellon University

Banff workshop on Communication Complexity & Applications
August 26, 2014

Joint work with Clément Canonne, Raghu Meka & Madhu Sudan

Shared randomness

Shared randomness can yield big savings in comm. complexity.

Basic example: Equality testing

- Alice has $x \in \{0,1\}^n$ and Bob $y \in \{0,1\}^n$; Determine $x \stackrel{?}{=} y$.
- Deterministic communication complexity (CC) $\geqslant n$ bits
- CC with private randomness $= \Theta(\log n)$ (with some error)
- Shared randomness r (independent of x, y): $O(1)$ (one-way) CC

Shared randomness

Shared randomness can yield big savings in comm. complexity.

Basic example: Equality testing

- Alice has $x \in \{0,1\}^n$ and Bob $y \in \{0,1\}^n$; Determine $x \stackrel{?}{=} y$.
- Deterministic communication complexity (CC) $\geqslant n$ bits
- CC with private randomness $= \Theta(\log n)$ (with some error)
- Shared randomness r (independent of x, y): $O(1)$ (one-way) CC

Question studied in this work

Is *perfect* shared randomness (psr) required for comm. savings?

How is communication complexity of various problems affected with *imperfect* shared randomness (isr), when Alice and Bob have r and r' with some known correlation between r and r'?

Motivation

Imperfect shared randomnes is a natural model, pointing to new phenomena and raising interesting questions.

In many forms of natural communication, parties understand each other fairly well, but never perfectly.

- This imperfection in understanding creates an obstacle to many known solutions, and calls for new approaches/techniques
- Understand the inherent price one has to pay to solve communication problems with imperfect shared context.

Communication w/ Imperfect Shared Randomness

Focus on two-party setting.

Alice and Bob have randomness r and r' resp., consisting of i.i.d samples (r_i, r'_i) from some correlated distribution D.

- Note: D uniform on $\{00, 11\}$ is perfect shared randomness

[1] Our proposal was (amusingly) independent, but came later

Communication w/ Imperfect Shared Randomness

Focus on two-party setting.

Alice and Bob have randomness r and r' resp., consisting of i.i.d samples (r_i, r'_i) from some correlated distribution D.

- Note: D uniform on $\{00, 11\}$ is perfect shared randomness

Model introduced recently in [Bavarian-Gavinsky-Ito'14][1]

- Primary focus on simultaneous message model
- Technical focus on different kinds of correlation.
- Among basic results: Equality testing has an $O(1)$ CC protocol as long as D is not a product distribution.

[1] Our proposal was (amusingly) independent, but came later

Our Focus

- ρ-correlated randomness $r \sim_\rho r'$:
  - i'th bits $(r_i, r'_i) \sim \{1, -1\}^2$ with uniform marginals & $\mathbb{E}[r_i r'_i] = \rho$.
  - $\rho = 1$ is perfect shared randomness; $\rho = 0$ is private randomness.
- One-way communication model
  - Alice sends one message to Bob who announces the answer
- Main contributions:
  - Introduce some new communication problems of interest
  - Compare their comm. compl. with/without perfect shared randomness.

Basic observations

P = communication (promise) problem with inputs $x, y \in \{0,1\}^n$.

Various measures

Randomized comm. compl. (for error $1/10$ say):

- $\mathrm{isr}_\rho(P)$ = comm. compl. (CC) with ρ-correlated randomness
- $\mathrm{psr}(P)$ = CC with perfect shared randomness $(= \mathrm{isr}_1(P))$
- $\mathrm{private}(P)$ = CC with private randomness $(= \mathrm{isr}_0(P))$

$[\mathrm{isr}^{\mathrm{ow}}_\rho(P), \mathrm{psr}^{\mathrm{ow}}(P)$ for one-way model$]$

Basic observations

P = communication (promise) problem with inputs $x, y \in \{0,1\}^n$.

Various measures

Randomized comm. compl. (for error $1/10$ say):

- $\mathrm{isr}_\rho(P)$ = comm. compl. (CC) with ρ-correlated randomness
- $\mathrm{psr}(P)$ = CC with perfect shared randomness $(= \mathrm{isr}_1(P))$
- $\mathrm{private}(P)$ = CC with private randomness $(= \mathrm{isr}_0(P))$

$[\mathrm{isr}^{\mathrm{ow}}_\rho(P), \mathrm{psr}^{\mathrm{ow}}(P)$ for one-way model$]$

Observation

$\forall \rho \in [0,1]$, $\mathrm{psr}(P) \leqslant \mathrm{isr}_\rho(P) \leqslant \mathrm{private}(P) \leqslant \mathrm{psr}(P) + O(\log n)$

(holds for one-way model as well)

isr vs. psr

Observation

$\mathrm{psr}(P) \leqslant \mathrm{isr}_{\rho}(P) \leqslant \mathrm{psr}(P) + O(\log n)$ (holds for one-way model as well)

- Thus, CC with imperfectly shared randomness is never too far from CC with perfect shared randomness.

isr vs. psr

Observation

$\mathrm{psr}(P) \leqslant \mathrm{isr}_{\rho}(P) \leqslant \mathrm{psr}(P) + O(\log n)$ (holds for one-way model as well)

- Thus, CC with imperfectly shared randomness is never too far from CC with perfect shared randomness.
- But sometimes the $\log n$ overhead is undesirable (eg. when there is $O(1)$ complexity psr protocol)
- $\therefore$ interesting to study whether gap can be smaller in such cases

Main Results

Problems with $O(1)$ one-way CC in psr model also admit $O(1)$ bit one-way protocols with imperfect (ρ-correlated) shared randomness.

- But, with exponentially more bits
- And this exponential blow-up is necessary in general

Main Results

Problems with $O(1)$ one-way CC in psr model also admit $O(1)$ bit one-way protocols with imperfect (ρ-correlated) shared randomness.

- But, with exponentially more bits
- And this exponential blow-up is necessary in general

Theorem (Upper bound)

$\forall \rho > 0$, $\exists c_\rho < \infty$, *such that for any promise problem* P

$$\mathrm{psr}^{\mathrm{ow}}(P) = k \Longrightarrow \mathrm{isr}^{\mathrm{ow}}_{\rho}(P) \leqslant c_\rho \cdot 2^k.$$

Main Results

Problems with $O(1)$ one-way CC in psr model also admit $O(1)$ bit one-way protocols with imperfect (ρ-correlated) shared randomness.

- But, with exponentially more bits
- And this exponential blow-up is necessary in general

Theorem (Upper bound)

$\forall \rho > 0$, $\exists c_\rho < \infty$, *such that for any promise problem* P

$$\mathrm{psr}^{\mathrm{ow}}(P) = k \Longrightarrow \mathrm{isr}^{\mathrm{ow}}_{\rho}(P) \leqslant c_\rho \cdot 2^k.$$

Theorem (Lower bound, the main technical result)

$\forall k \in \mathbb{Z}^+$, $\exists$ *a promise problem* P *such that*

$$\mathrm{psr}^{\mathrm{ow}}(P) \leqslant k \quad \textit{but} \quad \forall \rho < 1, \quad \mathrm{isr}^{\mathrm{ow}}_{\rho}(P) \geqslant 2^{\Omega_\rho(k)}.$$

Other Results

1. Compression with *uncertain priors*
   - Alice has to send $m \sim P$ to Bob
   - Bob knows $P \Rightarrow$ can compress to $H(P)$ bits
   - What if Bob doesn't know P perfectly (knows Q "close" to P)?
   - $O(H(P) + \delta(P, Q))$ communication achievable in imperfect shared randomness model (extending psr protocol of [Juba-Kalai-Khanna-Sudan'11])

Other Results

1. Compression with *uncertain priors*
   - Alice has to send $m \sim P$ to Bob
   - Bob knows $P \Rightarrow$ can compress to $H(P)$ bits
   - What if Bob doesn't know P perfectly (knows Q "close" to P)?
   - $O(H(P) + \delta(P, Q))$ communication achievable in imperfect shared randomness model (extending psr protocol of [Juba-Kalai-Khanna-Sudan'11])

2. Agreement distillation problem (extracting common sample from ρ-correlated samples)
   - Communication complexity of extracting shared randomness with min-entropy m is $\Theta_\rho(m)$ for every $\rho \in (0, 1)$.
   - Used in previous $2^{\Omega_\rho(k)}$ isr model lower bound.

Rest of the Talk

Some ideas about:

1. simulating one-way psr protocols with imperfect (ρ-correlated) shared randomness
2. exponential separation between one-way CC in isr and psr models

Some open questions

A complete problem for one-way psr model

GapIndex$_q(n)$

Alice holds $x \in [q]^n$;
Bob holds $y \in (\{0,1\}^q)^n$ (equivalently, $y : [n] \times [q] \to \{0,1\}$)

Goal: Bob must distinguish between

- Yes instances: $\mathbb{E}_{i \sim [n]}[y(i, x_i)] > 2/3$
- No instances: $\mathbb{E}_{i \sim [n]}[y(i, x_i)] < 1/3$

A complete problem for one-way psr model

$\textsc{GapIndex}_q(n)$

Alice holds $x \in [q]^n$;
Bob holds $y \in (\{0,1\}^q)^n$ (equivalently, $y : [n] \times [q] \to \{0,1\}$)

Goal: Bob must distinguish between

- YES instances: $\mathbb{E}_{i \sim [n]}[y(i, x_i)] > 2/3$
- NO instances: $\mathbb{E}_{i \sim [n]}[y(i, x_i)] < 1/3$

Observation

In psr model, $\textsc{GapIndex}_q(n)$ has a $\lceil \lg q \rceil$ bit one-way protocol.

- (Alice & Bob use shared randomness to sample $i \sim [n]$;
  Alice sends $x_i \in [q]$; Bob accepts if $y(i, x_i) = 1$.)

Completeness for one-way psr

GapIndex$_q(n)$

Alice holds $x \in [q]^n$;
Bob holds $y : [n] \times [q] \to \{0, 1\}$)

Goal: Bob must distinguish between $\mathbb{E}_{i \sim [n]}[y(i, x_i)] > 2/3$ or $< 1/3$

Completeness for one-way psr

$\textsc{GapIndex}_q(n)$

Alice holds $x \in [q]^n$;
Bob holds $y : [n] \times [q] \to \{0,1\}$)

Goal: Bob must distinguish between $\mathbb{E}_{i\sim[n]}[y(i,x_i)] > 2/3$ or $< 1/3$

$\textsc{GapIndex}_{2^k}(\cdot)$ is *complete* for problems with k-bit one-way protocol in psr model.

- $x =$ Alice's k-bit message for each value of shared randomness
- $y =$ Bob's decision for each randomness and possible message from Bob.

CC of GapIndex in isr model

Theorem

Let $q \in \mathbb{Z}^+$ and $\rho > 0$. $\textsc{GapIndex}_q(n)$ admits $O_\rho(q)$ bit one-way protocol in ρ-correlated isr model. Formally,

$$\mathrm{isr}^{\mathrm{ow}}_\rho(\textsc{GapIndex}_q(n)) \leqslant O(q \cdot \rho^{-2}) \ .$$

Together with completeness of $\textsc{GapIndex}_{2^k}(\cdot)$ for k-bit one-way psr protocols, we get

$$\mathrm{psr}^{\mathrm{ow}}(P) = k \Longrightarrow \mathrm{isr}^{\mathrm{ow}}_\rho(P) \leqslant c_\rho \cdot 2^k \ .$$

isr protocol for GapIndex

Mapping to inner product

View $x \in \{0,1\}^{nq}$ by encoding $j \in [q]$ by basis vector $e_j \in \{0,1\}^q$.

- Normalize by $1/\sqrt{n}$ factor to get *unit vector* $u_x \in \mathbb{R}^{nq}$.

Let $u_y = y/\sqrt{n}$, and note $\|u_y\| \leqslant \sqrt{q}$.

- Now $\mathbb{E}_i[y(i, x_i)] = \langle u_x, u_y \rangle$.

Need a protocol to estimate $\langle u_x, u_y \rangle$ with error $< 1/6$.

Gaussians based intuition

Geometric setup

Alice holds $u_x \in \mathbb{R}^m$, $\|u_x\| = 1$; Bob holds $u_y \in \mathbb{R}^m$, $\|u_y\| \leqslant \sqrt{q}$.
Goal: Estimate $\langle u_x, u_y \rangle$ with error $< 1/6$.

View shared randomness as ρ-correlated Gaussian vectors g_1, g_2.

- Alice can compute $a = \langle g_1, u_x \rangle$, and Bob $b = \langle g_2, y_y \rangle$.
- When $g_1 = g_2$, $\mathbb{E}[a \cdot b] = \langle u_x, u_y \rangle$ (and $\mathrm{Var}[ab] \lesssim \|u_y\|^2$).

Gaussians based intuition

Geometric setup

Alice holds $u_x \in \mathbb{R}^m$, $\|u_x\| = 1$; Bob holds $u_y \in \mathbb{R}^m$, $\|u_y\| \leqslant \sqrt{q}$.
Goal: Estimate $\langle u_x, u_y \rangle$ with error $< 1/6$.

View shared randomness as ρ-correlated Gaussian vectors g_1, g_2.

- Alice can compute $a = \langle g_1, u_x \rangle$, and Bob $b = \langle g_2, y_y \rangle$.
- When $g_1 = g_2$, $\mathbb{E}[a \cdot b] = \langle u_x, u_y \rangle$ (and $\mathrm{Var}[ab] \lesssim \|u_y\|^2$).
- When $g_2 \sim_\rho g_1$, ab is still a good estimator with higher variance.
- Thus can estimate $\mathbb{E}[ab]$ from $O_\rho(\|u_y\|^2) \leqslant O_\rho(q)$ samples.

Gaussians based intuition

Geometric setup

Alice holds $u_x \in \mathbb{R}^m$, $\|u_x\| = 1$; Bob holds $u_y \in \mathbb{R}^m$, $\|u_y\| \leqslant \sqrt{q}$.
Goal: Estimate $\langle u_x, u_y \rangle$ with error $< 1/6$.

View shared randomness as ρ-correlated Gaussian vectors g_1, g_2.

- Alice can compute $a = \langle g_1, u_x \rangle$, and Bob $b = \langle g_2, y_y \rangle$.
- When $g_1 = g_2$, $\mathbb{E}[a \cdot b] = \langle u_x, u_y \rangle$ (and $\mathrm{Var}[ab] \lesssim \|u_y\|^2$).
- When $g_2 \sim_\rho g_1$, ab is still a good estimator with higher variance.
- Thus can estimate $\mathbb{E}[ab]$ from $O_\rho(\|u_y\|^2) \leqslant O_\rho(q)$ samples.

Carefuly analysis of *(a variant of this protocol)* shows it suffices for Alice to send $O(q/\rho^2)$ bits to Bob.

A variant: GapIP

For the protocol, only sparsity of x (and not its structure of one 1 in each q-sized block) is important. Motivates related problem:

$\textsc{GapIP}_q(n)$

Alice has $x \in \{0,1\}^n$ with $\|x\|^2 \in (1 \pm 0.01)\frac{n}{q}$;

Bob has $y \in \{1,-1\}^n$.

Goal: Distinguish between cases $\langle x, y \rangle > \frac{2}{3}\frac{n}{q}$ and $\langle x, y \rangle < \frac{1}{3}\frac{n}{q}$.

A variant: GapIP

For the protocol, only sparsity of x (and not its structure of one 1 in each q-sized block) is important. Motivates related problem:

$\mathrm{GapIP}_q(n)$

Alice has $x \in \{0,1\}^n$ with $\|x\|^2 \in (1 \pm 0.01)\frac{n}{q}$;
Bob has $y \in \{1,-1\}^n$.

Goal: Distinguish between cases $\langle x, y \rangle > \frac{2}{3}\frac{n}{q}$ and $\langle x, y \rangle < \frac{1}{3}\frac{n}{q}$.

Fact

$\mathrm{psr}^{\mathrm{ow}}(\mathrm{GapIP}_q(n)) \leqslant O(\log q)$ (for $n \to \infty$)

- Shared randomness gives sequence $i_1, i_2, \ldots$ of indices $\in [n]$
  Alice sends smallest ℓ such that $x_{i_\ell} \neq 0$; Bob outputs y_{i_ℓ}.

Tight lower bound in isr model

Theorem

$\forall \rho < 1$, $\exists \gamma > 0$ *such that* $\forall q \in \mathbb{Z}^+$ *and sufficiently large* n

$$\mathsf{isr}^{\mathsf{ow}}_{\rho}(\textsc{GapIP}_q(n)) \geqslant \gamma \cdot q \ .$$

Exponential gap between one-way CC in isr and psr models.

Lower Bound: High level intuition

Contrasting "Gaussian" protocol with the psr protocol:

- In the psr model, Alice uses randomness to pick one (or few) coordinates of x and sends some function of these bits to Bob
- In isr solution, Alice sends a very "non-junta" like function of x that is robust to perturbations of randomness (leads to CC $\Omega(q)$)

Lower Bound: High level intuition

Contrasting "Gaussian" protocol with the psr protocol:

- In the psr model, Alice uses randomness to pick one (or few) coordinates of x and sends some function of these bits to Bob
- In isr solution, Alice sends a very "non-junta" like function of x that is robust to perturbations of randomness (leads to CC $\Omega(q)$)

Guided by this difference, our proof has two parts:

1. Functions whose variables have "low influence" cannot be good strategies for Alice (unless they send $\Omega(q)$ bits)
2. If protocol uses functions with some influential variables, then Alice & Bob should agree on which variable to use.
   - *Agreement distillation* lower bound rules this out in isr model.

Abstracting strategies

Recall $\textsc{GapIP}_q(n)$:

- Alice has $x \in \{0,1\}^n$ with $\|x\|^2 \approx \frac{n}{q}$; Bob has $y \in \{1,-1\}^n$.
- Goal: Distinguish between cases $\langle x, y \rangle > \frac{2}{3}\frac{n}{q}$ and $\langle x, y \rangle < \frac{1}{3}\frac{n}{q}$.

Suppose Alice communicates one of ℓ possible messages:

- Alice's strategy given by functions $f_r : \{0,1\}^n \to \Delta(\ell)$
- Bob's strategy are functions $g_{r'} : \{1,-1\}^n \to [0,1]^\ell$.
- Acceptance probability on input (x, y) equals

$$\mathbb{E}_{r \sim_\rho r'}[\langle f_r(x), g_{r'}(y) \rangle] \ .$$

Low-influence protocols

Analyzed via a variant of the **invariance principle:**

- Strategies with low-influences for a "product-distributional" setting can be mapped to strategies behaving similarly for inputs that are *Gaussians with appropriate means and covariances.*

Low-influence protocols

Analyzed via a variant of the **invariance principle:**

- Strategies with low-influences for a "product-distributional" setting can be mapped to strategies behaving similarly for inputs that are *Gaussians with appropriate means and covariances.*

One-way protocol for $\mathrm{GAPIP}_q(n)$ implies one-way protocol of same communication complexity for following Gaussian problem:

- Alice is given $X \sim N(0,1)^n$, Bob $Y \sim N(0,1)^n$ with $\mathbb{E}[X_i Y_i] = \xi$
- Goal: tell apart $\xi = 1/\sqrt{q}$ (Yes instances) and $\xi = 0$ (No instances)

Low-influence protocols

Analyzed via a variant of the **invariance principle:**

- Strategies with low-influences for a "product-distributional" setting can be mapped to strategies behaving similarly for inputs that are *Gaussians with appropriate means and covariances.*

One-way protocol for $\mathrm{GAPIP}_q(n)$ implies one-way protocol of same communication complexity for following Gaussian problem:

- Alice is given $X \sim N(0,1)^n$, Bob $Y \sim N(0,1)^n$ with $\mathbb{E}[X_i Y_i] = \xi$
- Goal: tell apart $\xi = 1/\sqrt{q}$ (Yes instances) and $\xi = 0$ (No instances)

Solving Gaussian problem requires Alice to send $\Omega(q)$ bits ($\ell \geqslant 2^{\Omega(q)}$)

- Proof by reduction from indexing [Jayram-Kumar-Sivakumar] (described in David Woodruff's talk yesterday)

Influential case of proof

Invariance principle $\Rightarrow$ If $\ell = 2^{o(q)}$, then f_r and $g_{r'}$ *share* a coordinate with noticeable (low-degree) influence.

Influential case of proof

Invariance principle $\Rightarrow$ If $\ell = 2^{o(q)}$, then f_r and $g_{r'}$ *share* a coordinate with noticeable (low-degree) influence.

If Alice and Bob pick a random (low-degree) influential coordinate for their respective strategies, they will agree on a common coordinate $i \in [n]$ (with noticeable probability)

Influential case of proof

Invariance principle $\Rightarrow$ If $\ell = 2^{o(q)}$, then f_r and $g_{r'}$ *share* a coordinate with noticeable (low-degree) influence.

If Alice and Bob pick a random (low-degree) influential coordinate for their respective strategies, they will agree on a common coordinate $i \in [n]$ (with noticeable probability)

Can show: Varying r, r', the agreed upon i has high min-entropy

- This part uses distributions on Yes/No instances tailored to defeat specific strategies $\{f_r\}$, $\{g_{r'}\}$
- Unlike usual lower bounds, can't pick a single pair of distributions on Yes/No instances
  - Existence of psr protocol + Yao min-max principle $\Rightarrow$ *deterministic* strategy for *any* fixed distribution

Agreement Distillation

When $r \sim_\rho r'$, such randomness extraction requires $\gg q$ bits communication (for large n).

Agreement Distillation(m, η)

Two players Charlie and Dana, with access to correlated randomness $r \sim_\rho r'$.

Goal: Sample outputs w_C and w_D such that:

1. $H_\infty(w_C), H_\infty(w_D) \geqslant m$ and
2. $\Pr[w_C = w_D] \geqslant \eta$

Agreement Distillation

When $r \sim_\rho r'$, such randomness extraction requires $\gg q$ bits communication (for large n).

Agreement Distillation(m, η)

Two players Charlie and Dana, with access to correlated randomness $r \sim_\rho r'$.

Goal: Sample outputs w_C and w_D such that:

1. $H_\infty(w_C), H_\infty(w_D) \geqslant m$ and
2. $\Pr[w_C = w_D] \geqslant \eta$

Using Shannon capacity-achieving linear codes (for BSC_ρ), can solve above with $h\left(\frac{1-\rho}{2}\right) m + o(m)$ bits of communication (and $\eta = 1 - o_m(1)$).

Lower bound Agreement Distillation

Agreement Distillation(m, η):

Charlie and Dana, with access to correlated randomness $r \sim_\rho r'$.

Goal: Sample w_C, w_D s.t. $H_\infty(w_C), H_\infty(w_D) \geqslant m$ & $\Pr[w_C = w_D] \geqslant \eta$

Theorem

$\forall \rho < 1$, $\exists \delta > 0$ *such that Agreement Distillation*(m, η) *requires* $\delta m - \log(1/\eta)$ *bits of communication.*

Lower bound Agreement Distillation

Agreement Distillation(m, η):

Charlie and Dana, with access to correlated randomness $r \sim_\rho r'$.

Goal: Sample w_C, w_D s.t. $H_\infty(w_C), H_\infty(w_D) \geqslant m$ & $\Pr[w_C = w_D] \geqslant \eta$

Theorem

$\forall \rho < 1$, $\exists \delta > 0$ *such that Agreement Distillation*(m, η) *requires* $\delta m - \log(1/\eta)$ *bits of communication.*

Proof Idea

Consider zero-communication protocols.

- *Small set expansion* of the noisy hypercube implies agreement probability at most $2^{-m\left(\frac{1-\rho}{1+\rho}\right)}$.
- c bits sent by Alice can only improve success probability by factor 2^c.

Summary

Studied one-way communication in isr model when randomness of two parties are ρ-correlated.

One-way public randomness protocols sending k bits can be simulated by sending $\approx 2^k$ bits in isr model, and this is tight.

The isr protocols have to be "provably" different from the classical solutions in psr model.

New application domain for PCPs/inapproximability tools such as small set expansion, invariance principles, influence-based decoding.

Open questions

1. Other forms of correlations? Which ones are as powerful as ρ-correlated random bits?
2. Randomness reduction
   - Equality testing: [Bavarian-Gavinsky-Ito'14] use 2^n randomness, our Gaussian protocol uses $\mathrm{poly}(n)$ random bits. Could $O(\log n)$ random bits suffice?
   - How about for general problems?
3. All our protocols for isr model lead to two-sided error. Perhaps one-sided error in isr model cannot lead to any savings over private randomness?