Communication with Imperfect Shared Randomness

(Joint work with Venkatesan Guruswami (CMU), Raghu Meka (UCLA) and Madhu Sudan (MSR))

Who? Clément Canonne (Columbia University)

When? January 12, 2015

Communication & Complexity

There is a world outside of n

Context

There is Alice, Bob, what they communicate *and what they don't have to*.

Communication & Complexity

There is a world outside of n

Context

There is Alice, Bob, what they communicate *and what they don't have to*.
The

$$f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}\ ,$$

they compute; the protocol

$$\Pi$$

they use; from which

$$D_x, D_y$$

their inputs come; what is blue and what red means.

Communication & Complexity

But context is not perfect...

Noise, misunderstandings, false assumptions

Context is almost never perfectly shared.

Communication & Complexity

But context is not perfect. . .

Noise, misunderstandings, false assumptions

Context is almost never perfectly shared.

- My periwinkle is your orchid.

Communication & Complexity

But context is not perfect. . .

Noise, misunderstandings, false assumptions

Context is almost never perfectly shared.

- My periwinkle is your orchid.
- the printer on the 5^{th} floor of Columbia is not *exactly* the model my laptop has a driver for.

Communication & Complexity

But context is not perfect...

Noise, misunderstandings, false assumptions

Context is almost never perfectly shared.

- My periwinkle is your orchid.
- the printer on the 5$^{\text{th}}$ floor of Columbia is not *exactly* the model my laptop has a driver for.
- what on Earth is a "French baguette" in New York?!

Communication & Complexity

But context is not perfect...

Noise, misunderstandings, false assumptions

Context is almost never perfectly shared.

- My periwinkle is your orchid.
- the printer on the 5th floor of Columbia is not *exactly* the model my laptop has a driver for.
- what on Earth is a "French baguette" in New York?!

This talk: Shared randomness in communication complexity as "context"

Equality testing

I have $x \in \{0,1\}^n$, you have $y \in \{0,1\}^n$, are they equal?

Complexity?

- Deterministic $\text{det}(\mathrm{EQ}) = \Theta(n)$
- Private randomness $\text{private}(\mathrm{EQ}) = \Theta(\log n)$
- Shared randomness $\text{psr}(\mathrm{EQ}) = O(1)$

(Recall Newman's Theorem:

$$\text{private}(P) \leq \text{psr}(P) + O(\log n).)$$

This work

Randomness *and* uncertainty

ISR (Imperfectly Shared Randomness)

What if the randomness ("context") was not perfectly in sync?

To compute $f(x, y)$:

- Alice: has access to $r \in \{\pm 1\}^*$, gets input $x \in \{0, 1\}^n$
- Bob: has access to $s \in \{\pm 1\}^*$, gets input $y \in \{0, 1\}^n$

w/ $r \sim_\rho s$: s is obtained by perturbing each bit of r independently

This work

Randomness *and* uncertainty

ISR (Imperfectly Shared Randomness)

What if the randomness ("context") was not perfectly in sync?

To compute $f(x, y)$:

- Alice: has access to $r \in \{\pm 1\}^*$, gets input $x \in \{0,1\}^n$
- Bob: has access to $s \in \{\pm 1\}^*$, gets input $y \in \{0,1\}^n$

w/ $r \sim_\rho s$: s is obtained by perturbing each bit of r independently

Studied (independently) by [BGI14] (different focus: "referee model"; more general correlations).

ISR: general relations

For every P with $x, y \in \{0,1\}^n$ and $0 \leq \rho \leq \rho' \leq 1$,

$$\begin{aligned} \mathsf{psr}(P) &\leq \mathsf{isr}_{\rho'}(P) \leq \mathsf{isr}_{\rho}(P) \\ &\leq \mathsf{private}(P) \leq \mathsf{psr}(P) + O(\log n). \end{aligned}$$

(also true for one-way: $\mathsf{psr}^{\mathrm{ow}}, \mathsf{isr}_{\rho}^{\mathrm{ow}}, \mathsf{private}^{\mathrm{ow}}$)

$\rightsquigarrow$ but for many problems, $\log n$ is already huge.

Rest of the talk

1. A first example: the COMPRESSION problem
2. General upperbound on ISR in terms of PSR
3. Strong lower bound: Alice, Bob, Charlie and Dana.

First result: Uncertain Compression

Compression with uncertain priors

Alice has P, gets $m \sim P$; Bob knows $Q \simeq P$, wants m.

Previous work

$P = Q$	$H(P)$ (Huffman coding)
$P \simeq_\Delta Q$	$H(P) + 2\Delta$ [JKKS11] (w/ shared randomness)
$P \simeq_\Delta Q$	$O(H(P) + \Delta + \log\log N)$ [HS14] (deterministic)

First result: Uncertain Compression

Compression with uncertain priors

Alice has P, gets $m \sim P$; Bob knows $Q \simeq P$, wants m.

Previous work

$P = Q$ — $H(P)$ (Huffman coding)

$P \simeq_\Delta Q$ — $H(P) + 2\Delta$ [JKKS11] (w/ shared randomness)

$P \simeq_\Delta Q$ — $O(H(P) + \Delta + \log\log N)$ [HS14] (deterministic)

This work

For all $\epsilon > 0$,
$\mathrm{isr}_\rho^{\mathrm{ow}}(\textsc{Compress}_\Delta) \leq \frac{1+\epsilon}{1-h(\frac{1-\rho}{2})}(H(P) + 2\Delta + O(1))$

"natural protocol"

General upperbound

It's inner products all the way down!

Theorem $\forall \rho > 0$, $\exists c < \infty$ *such that* $\forall k$, *we have*

$$\mathsf{PSR}(k) \subseteq \mathsf{ISR}_\rho^{\mathrm{ow}}(c^k)\,.$$

Proof. (Outline)

- Define $\textsc{GapInnerProduct}$, "complete" for $\mathsf{PSR}(k)$ (see strategies as $X_R, Y_R\{0,1\}^{2^k}$; use Newman's Theorem to bound $\#$ R's);
- Show there exists a (Gaussian-based) isr protocol for $\textsc{GapInnerProduct}$, with $O_\rho(4^k)$ bits of comm.

□

General upperbound

Can we do better?

For problems in $\mathsf{PSR}^{\mathrm{ow}}(k)$? $\mathsf{PSR}^{\mathrm{ow}}(k) \subseteq \mathsf{ISR}^{\mathrm{ow}}_{\rho}(c^{o(k)})$?

For ISR_{ρ}? $\mathsf{PSR}(\omega(k)) \subseteq \mathsf{ISR}_{\rho}(c^{k})$?

General upperbound

Can we do better?

For problems in $\mathrm{PSR}^{\mathrm{ow}}(k)$? $\mathrm{PSR}^{\mathrm{ow}}(k) \subseteq \mathrm{ISR}^{\mathrm{ow}}_{\rho}(c^{o(k)})$?

For ISR_{ρ}? $\mathrm{PSR}(\omega(k)) \subseteq \mathrm{ISR}_{\rho}(c^{k})$?

Answer No.

Strong converse: lower bound

It's as good as it gets.

Theorem $\forall k$, $\exists P = (P_n)_{n\in\mathbb{N}}$ *s.t.* $\mathrm{psr}^{\mathrm{ow}}(P) \leq k$, *yet* $\forall \rho < 1$

$$\mathrm{isr}_\rho(P) = 2^{\Omega_\rho(k)}\,.$$

Proof. (High-level)

- Define SparseGapInnerProduct, relaxation of GapInnerProduct.
- Show it has as $O(\log q)$-bit one-way psr protocol (Alice uses the shared randomness to send *one* coordinate to Bob)
- isr lower bound: argue that for any (fixed)* strategy of Alice and Bob using less that $\sqrt{q}$ bits, either (a) something impossible happens in the Boolean world, or (b) something impossible happens in the Gaussian world.

□

Strong converse: lower bound

Two-pronged impossibility: more on the prongs.

Case (a) The strategies $(f_r, g_s)_{r,s}$ have common high-influence variable: Charlie and Dana can use Alice and Bob's protocol to *distill randomness*.

Case (b) If no such variable: with a new Invariance Principle[1], Charlie and Dana can apply Alice and Bob's protocol "transferred to the Gaussian world" to solve (yet another) problem, *Gaussian Inner Product*.

[1] In the spirit of [Mos10]

Conclusions

Summary

- Dealing with more realistic situations: Alice, Bob, and what they do not know about each other;
- comes into play when n is **huge** (Newman's Theorem becomes loose);
- show general and tight relations and reductions in this model, with both upper and lower bounds.
- a new invariance theorem, and use in comm. complexity.

Conclusions

Summary

- Dealing with more realistic situations: Alice, Bob, and what they do not know about each other;
- comes into play when n is **huge** (Newman's Theorem becomes loose);
- show general and tight relations and reductions in this model, with both upper and lower bounds.
- a new invariance theorem, and use in comm. complexity.

What about...

- more general forms of correlations?
- cases where even randomness is expensive? (minimize its use)
- *one-sided error*?

Thank you.

(Questions?)

Mohammad Bavarian, Dmitry Gavinsky, and Tsuyoshi Ito.
On the role of shared randomness in simultaneous communication.
In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias, editors, *ICALP (1)*, volume 8572 of *Lecture Notes in Computer Science*, pages 150–162. Springer, 2014.

Andrej Bogdanov and Elchanan Mossel.
On extracting common random bits from correlated sources.
CoRR, abs/1007.2315, 2010.

Venkatesan Guruswami, Johan Håstad, Rajsekar Manokaran, Prasad Raghavendra, and Moses Charikar.
Beating the random ordering is hard: Every ordering CSP is approximation resistant.
SIAM J. Comput., 40(3):878–914, 2011.

Oded Goldreich, Brendan Juba, and Madhu Sudan.
A theory of goal-oriented communication.
Journal of the ACM, 59(2):8:1–8:65, May 2012.

Elad Haramaty and Madhu Sudan.
Deterministic compression with uncertain priors.
In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*, ITCS '14, pages 377–386, New York, NY, USA, 2014. ACM.

Brendan Juba, Adam Tauman Kalai, Sanjeev Khanna, and Madhu Sudan.
Compression without a common prior: an information-theoretic justification for ambiguity in language.
In Bernard Chazelle, editor, *ICS*, pages 79–86. Tsinghua University Press, 2011.

Elchanan Mossel.
Gaussian bounds for noise correlation of functions.
Geometric and Functional Analysis, 19(6):1713–1756, 2010.

Theorem (Our Invariance Principle)

Fix any two parameters $p_1, p_2 \in (-1, 1)$. For all $\varepsilon \in (0, 1]$, $\ell \in \mathbb{N}$, $\theta_0 \in [0, 1)$, and closed convex sets $K_1, K_2 \subseteq [0, 1]^\ell$ there exist $\tau > 0$ and mappings

$$T_1 \colon \{f \colon \{+1, -1\}^n \to K_1\} \to \{F \colon \mathbb{R}^n \to K_1\}$$
$$T_2 \colon \{g \colon \{+1, -1\}^n \to K_2\} \to \{G \colon \mathbb{R}^n \to K_2\}$$

such that for all $\theta \in [-\theta_0, \theta_0]$, if f, g satisfy

$$\max_{i \in [n]} \min \left(\max_{j \in [\ell]} \mathrm{Inf}_i(d) f_j, \max_{j \in [\ell]} \mathrm{Inf}_i(d) g_j \right) \leq \tau$$

then, for $F = T_1(f)$ and $G = T_2(g)$, we have where $N = N_{p_1, p_2, \theta}$ and $\mathcal{G}$ is the Gaussian distribution which matches the first and second-order moments of N.

Theorem (Invariance Theorem of [GHM+11])

Let (Ω, μ) be a finite prob. space with each prob. at least $\alpha \le 1/2$. Let $b = |\Omega|$ and $\mathcal{L} = \{\chi_0 = 1, \chi_1, \chi_2, \ldots, \chi_{b-1}\}$ be a basis for r.v.'s over Ω. Let $\Upsilon = \{\xi_0 = 1, \xi_1, \ldots, \xi_{b-1}\}$ be an ensemble of real-valued Gaussian r.v.'s with 1st and 2nd moments matching those of the χ_i's; and $h = (h_1, h_2, \ldots, h_t)\colon \Omega^n \to \mathbb{R}^t$ s.t.

$$\mathrm{Inf}_i(h_\ell) \le \tau, \qquad \mathrm{Var}(h_\ell) \le 1$$

for all $i \in [n]$ and $\ell \in [t]$. For $\eta \in (0,1)$, let H_ℓ ($\ell \in [t]$) be the multilinear polynomial associated with $T_{1-\eta} h_\ell$ w.r.t. $\mathcal{L}$. If $\Psi\colon \mathbb{R}^t \to \mathbb{R}$ is Λ-Lipschitz (w.r.t. the L_2-norm), then

$$\left|\mathbb{E}\Big[\Psi\big(H_1(\mathcal{L}^n), \cdots, H_t(\mathcal{L}^n)\big)\Big] - \mathbb{E}\Big[\Psi\big(H_1(\Upsilon^n), \cdots, H_t(\Upsilon^n)\big)\Big]\right| \le C(t) \cdot \Lambda \cdot \tau^{\frac{\eta}{18} \log \frac{1}{\alpha}} = o_\tau(1)$$

for some constant $C = C(t)$.

Strong converse: lower bound

Two-pronged impossibility, first prong.

Case (a)

The strategies $(f_r, g_s)_{r,s}$ have common high-influence variable *(recall the one-way psr protocol)*.
But then, two players Charlie and Dana can* leverage this strategies to win an *agreement distillation* game:

Definition (Agreement distillation)

Charlie and Dana have no inputs. Their goal is to output w_C and w_D satisfying:

$$\Pr[\, w_C = w_D \,] \geq \gamma;$$

$$H_\infty(w_C), H_\infty(w_D) \geq \kappa.$$

But this requires $\Omega(\kappa) - \log(1/\gamma)$ bits of communication (via [BM10, Theorem 1]).

Strong converse: lower bound

Two-pronged impossibility, second prong.

Case (b)

$f_r \colon \{0,1\}^n \to K_A \subset [0,1]^{2^k}$, $g_s \colon \{0,1\}^n \to K_B \subset [0,1]^{2^k}$ have no common high-influence variable.

We then show that this implies $k = 2^{\Omega(\sqrt{q})}$, by using an *Invariance Principle* (in the spirit of [Mos10]) to "go to the Gaussian world": if f, g are low-degree polynomials with no common influential variable, then

$$\mathbb{E}_{(x,y)\sim N^{\otimes n}}[\langle f(x), g(y)\rangle] \simeq \mathbb{E}_{(X,Y)\sim \mathcal{G}^{\otimes n}}[\langle F(X), G(Y)\rangle]$$

and Charlie and Dana can use this solve (yet another) problem, the Gaussian Inner Product ($\textsc{GaussianCorrelation}_\xi$).

But...

a reduction to $\textsc{Disjointness}$ shows that (even with psr) this requires $\Omega(1/\xi)$ bits of communication. $\square$