Resource-Efficient Common Randomness and Secret-Key Schemes

Badih Ghazi (Google)

Joint work with T.S. Jayram (IBM Almaden), Madhu Sudan and Mitali Bafna (Harvard), Pritish Kamath (MIT), Noah Golowich (Harvard → Google → MIT), Prasad Raghavendra (UC Berkeley).

Randomness Processing Industry

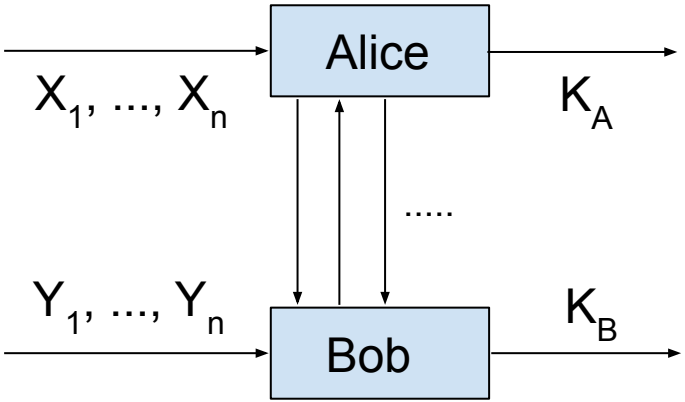
Dispersers, Extractors, Mergers, Condensers, PRGs …(long history omitted)

Key ingredients

Single processor

Unknown source

Distributed Randomness Processing



Objectives

Distribution of (K_A, K_B) is δ-close to target

Minimize δ, #n of samples, communication, # rounds, runtime

Examples of Correlated Sources

Alice gets input X

Bob gets input Y

Gaussian Source

$X \sim N(0,1)$

$Y \sim N(0,1)$

$E[XY] = \rho$

Binary Source

$X \sim U(\{-1, +1\})$

$Y \sim U(\{-1, +1\})$

$E[XY] = \rho$

Best Gaussian Correlation?

*Given i.i.d. samples from source P, largest ρ for which Alice and Bob can simulate a **Gaussian** source (without communication)?*

Maximal Correlation Coefficient:

$$\rho(P) = \sup_{\substack{f,g:\ \mathbb{E}[f(X)]=\mathbb{E}[g(Y)]=0 \\ \mathrm{Var}[X]=\mathrm{Var}[Y]=1}} \mathbb{E}[f(X)g(Y)]$$

[Witsenhausen, 1975]:

Best Gaussian correlation = *ρ(P)*

Computable in **polynomial time**! (SVD)

Best Binary Correlation?

*Given i.i.d. samples from source P, largest ρ for which Alice and Bob can simulate a **binary** source?*

[Witsenhausen, 1975]:

$$1 - \frac{2 \arccos \rho(P)}{\pi} \leq \rho^* \leq \rho(P)$$

Polynomial time quadratic approximation

Analogous to Goemans-Williamson rounding!

Best Binary Correlation?

$$1 - \frac{2 \arccos \rho(P)}{\pi} \leq \rho^* \leq \rho(P)$$

Binary source

Dictators are optimal! [Maximal Correlation]

Gaussian source

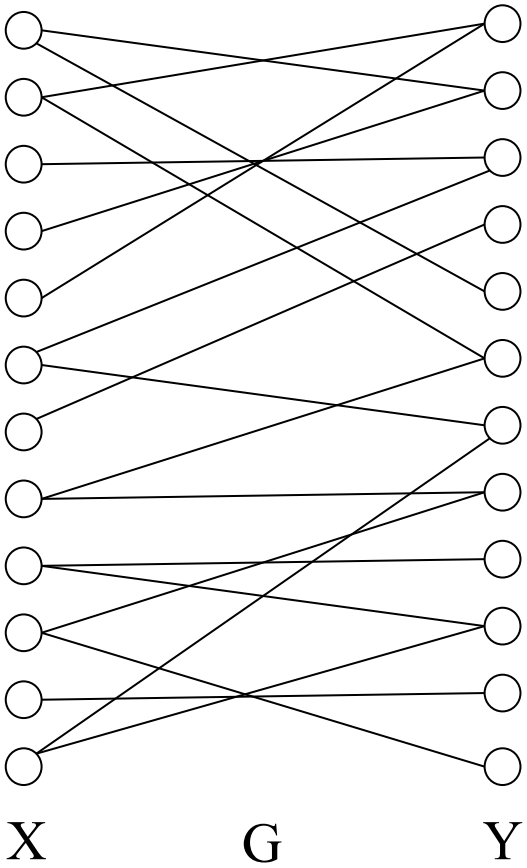
Halfspaces are optimal! [Borel's isoperimetric inequality, 1985]

Disjointness source

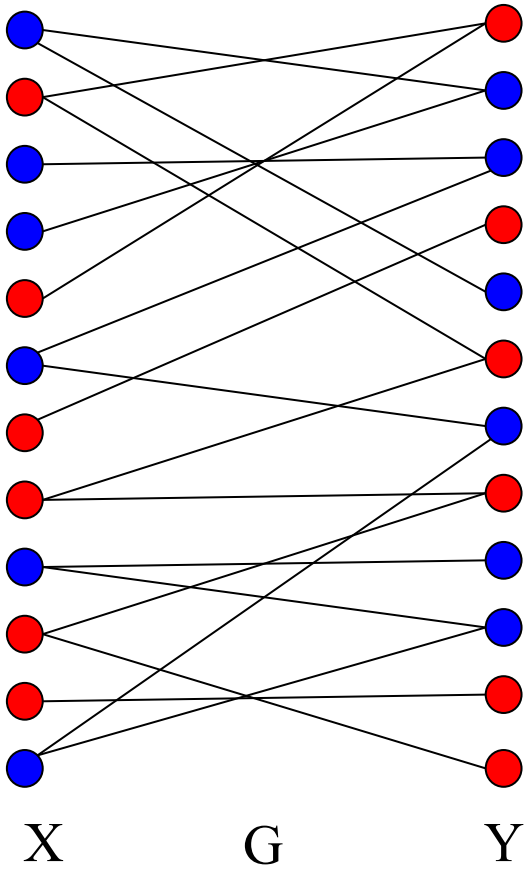
Uniform on {(0,0), (0,1), (1,0)}

Open in [1/3, 1/2]!

Exact Algorithm?


X
G
Y

Minimum Bipartite Bisection on Tensored Graphs



Minimize # edges cut
over all tensored graphs G^t

Equivalent to Best Binary Correlation!

Tensor-Power Problems

Problem	Base	Tensored
Compression	P	P
Channel Capacity	NP	P
Independent Set / Shannon Capacity	NP	?
Value of 2-prover game	NP	[NP, ∞]
Best Binary Correlation	NP	[0,CA]
Communication Complexity	[NP, Exp?]	[0, CA]
?	P	[NP, ∞]

Glossary: 0 ≤ P ≤ NP ≤ EXP ≤ Computable ≤ CA (Computable Approximately) ≤ ∞

Best Binary Correlation?

[G., Kamath, Sudan FOCS 2016]:

Computable Approximately
Doubly Exponential Time Algorithm

Ingredients:

Regularity Lemma
Invariance Principle [Mossel 2010]

Best Ternary Correlation?

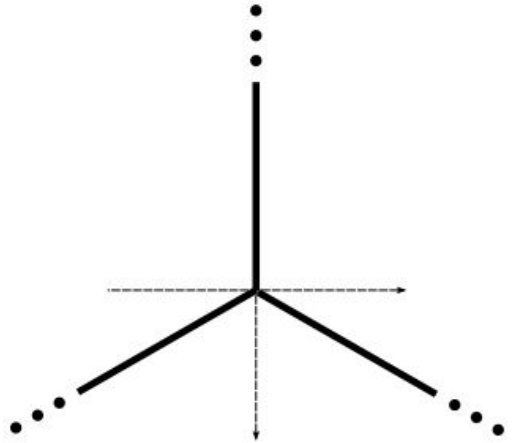
Gaussian source

Standard Simplex Conjecture

Peace Sign Partition

[Khot, Kindler, Mossel, O'Donnell 2007]

[Isaksson, Mossel 2012]



Simulating Arbitrary Given Source?

[De, Mossel, Neeman CCC 2017, SODA 2018]:

Approximately computable
Ackermann-type growth

[G., Kamath, Raghavendra CCC 2018]:

Doubly exponential
Dimension reduction for low-degree polynomials

Stronger goal:

Agreeing on k random bits using n samples from P

Common Randomness Generation

Objective:

Maximize k, Pr[agreement]

Minimize n, #rounds, communication

Equivalent to ***Secret Key Generation***

Key secure against eavesdropper

CRG: Zero Communication

Trivial Strategy:

Agreement probability 2^{-k}

[Bogdanov, Mossel IEEE Transactions on Information Theory 2011]:

Optimal tradeoff between agreement and entropy for **binary** source

$$\Pr[\text{agreement}] \approx 2^{-k\frac{(1-\rho)}{1+\rho}}$$

Ingredients:

Random binary linear error-correcting codes

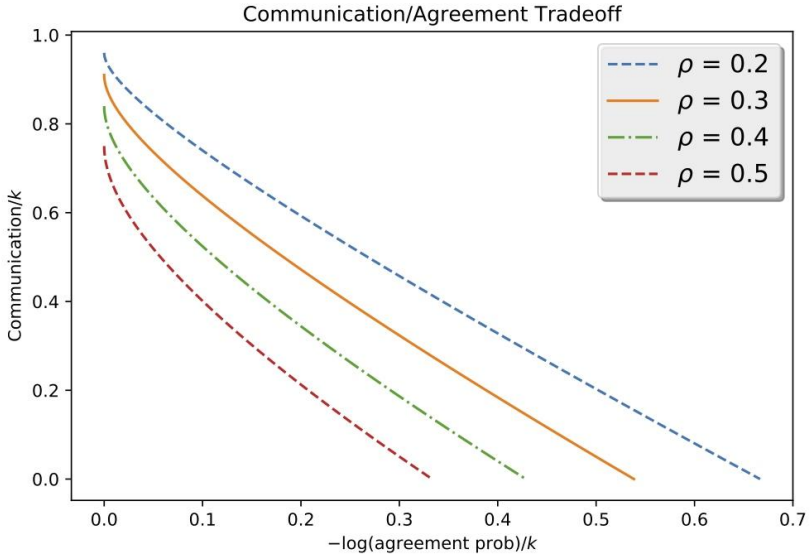
Hypercontractivity

CRG: One-Way Communication

[Guruswami, Radhakrishnan CCC 2016]:

Tight tradeoff for one-way communication

Similar ingredients



Explicit Schemes?

Sample-efficient?

Time-efficient?

CRG: Zero and One-Way Communication

[Jayram, G. SODA 2018]:

Explicit

Polynomial sample complexity

For **binary** and **Gaussian** sources

Ingredients:

Dual-BCH codes

Euclidean analogues

Computationally Efficient? Open!

Amortized CRG

*$\forall$ n large enough, agree on H*n bits of entropy with C*n communication*

[Ahlswede, Csiszar 1993]: characterization for one-way communication

Strong Data Processing Constant

$$s^*(X;Y) = \sup_{U:\ U-X-Y} \frac{I(U;Y)}{I(U;X)}$$

[Liu, Cuff, Verdu 2016]: multiple rounds

[Jayram, G., SODA 2018]: in terms of **Internal** and **External Information** Costs

Round Complexity

Do more rounds help?

For **binary** and **Gaussian** sources, question is open!

What about general sources?

[Tyagi 2013]: Separation between 1 and 2 rounds

[Bafna, G., Golowich, Sudan SODA 2019]: Round-communication tradeoffs for CRG & SKG

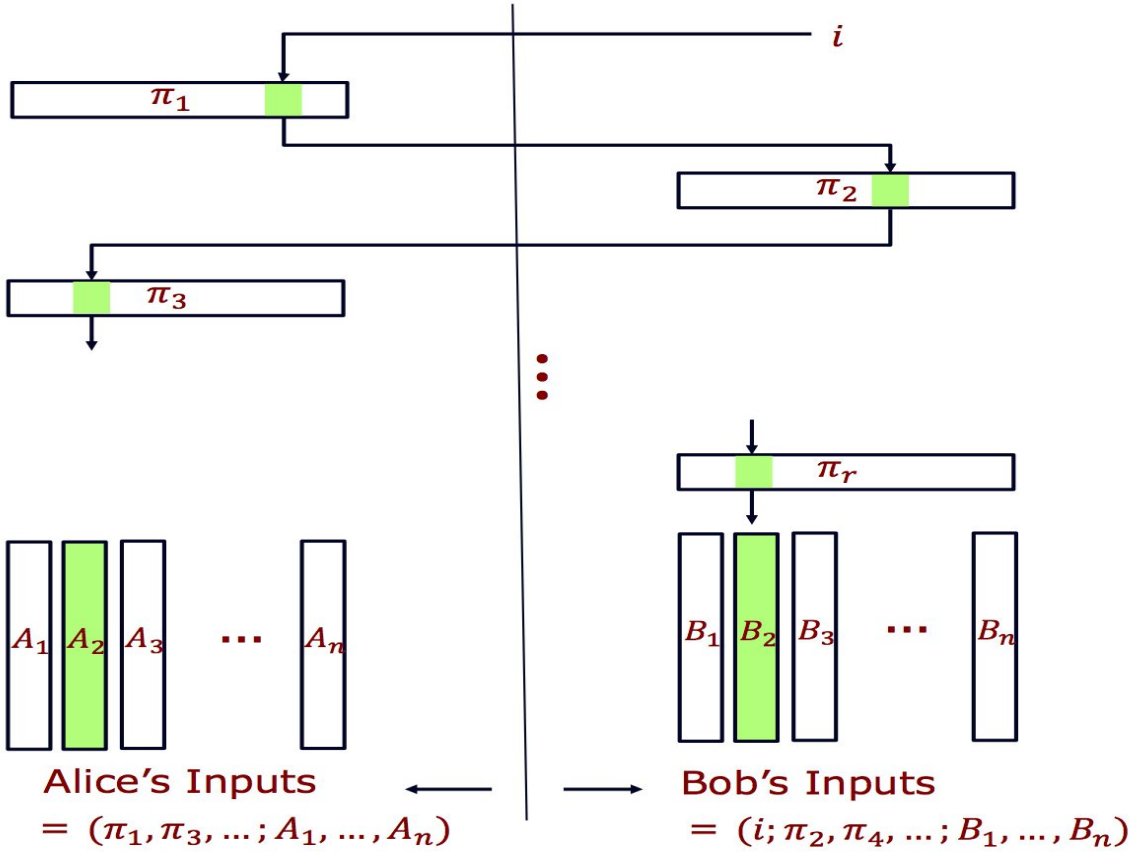
[Bafna, G., Golowich, Sudan SODA 2019]:

For every r and k, there is a source for which

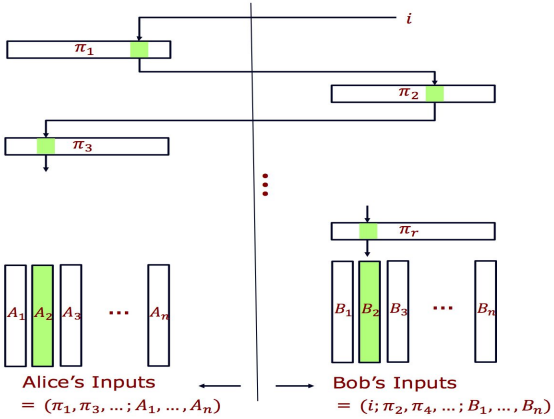
Agreeing on k random bits doable with r rounds and r*log(k) communication

Any protocol with r/2 rounds agreeing on k random bits has communication $\Omega(k)$

Pointer Chasing Source



Round-Communication Tradeoff



Upper Bound:

Immediate

Lower bound:

Reduce from Pointer Chasing [Nisan, Wigderson 1993]?

CRG problem can be solved without chasing pointers! (Equality Testing)

Pointer Verification Problem

Round elimination argument

Open Questions

Computational complexity of tensored graph problems?

The Houdre-Tetali conjecture

Time efficient common randomness generation?

Tight round-communication tradeoff for Pointer Chasing Source?

Connection to LSH?

Thank you!