Deterministic Randomness Extraction from Generalized and Distributed Santha-Vazirani Sources

Omid Etesami, IPM

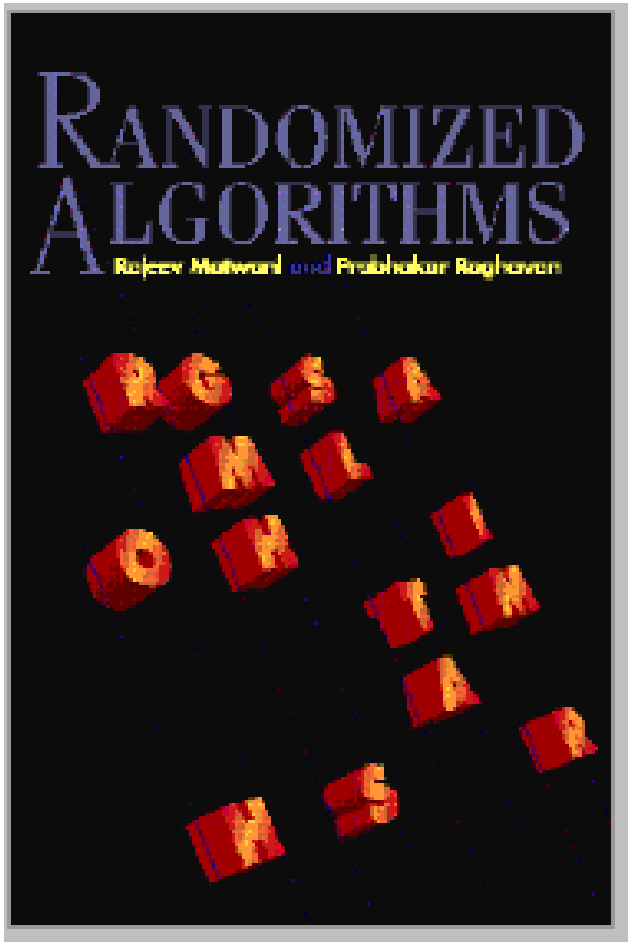
(Joint work with S. Beigi, A. Gohari)

Randomized algorithms

Sometimes simpler and more efficient than deterministic algorithms

Examples:

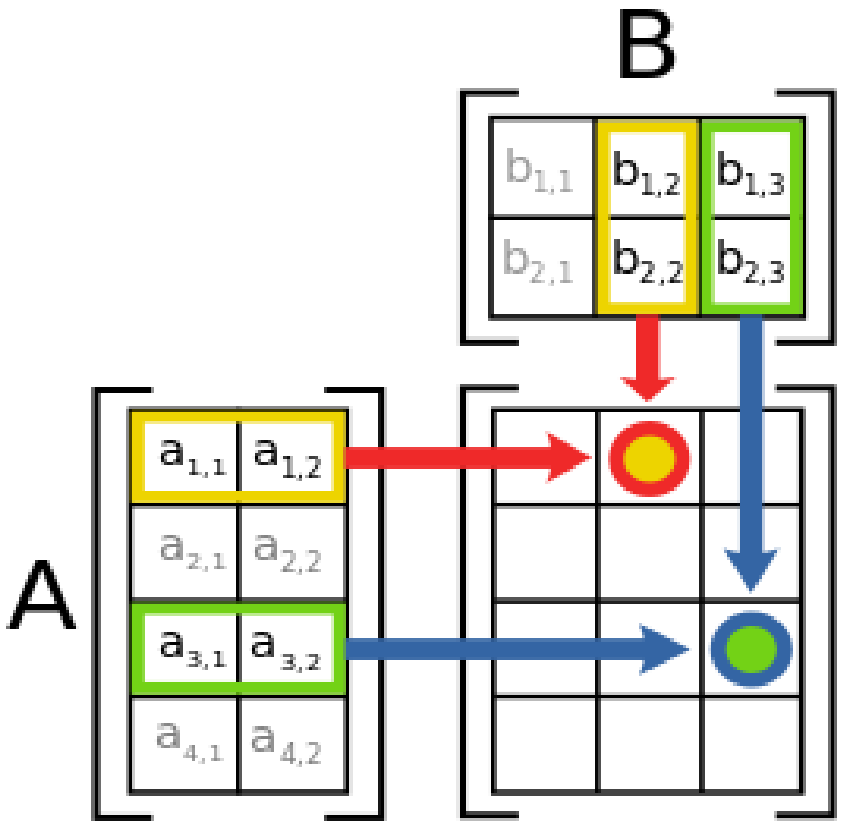
- Primality testing
- Quicksort
- Verifying matrix multiplication
- Min Cut
- Communication complexity
- Interactive proofs
- Cryptography
- Distributed algorithms
- …



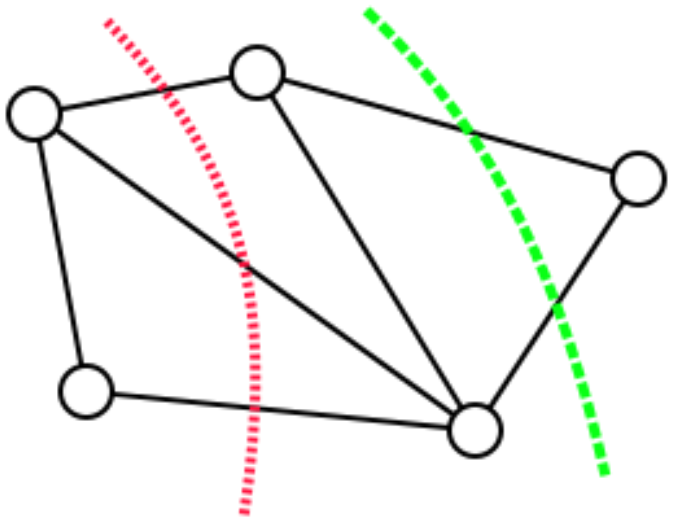
Verifying matrix multiplication

Given matrices A, B, C, is A B = C?

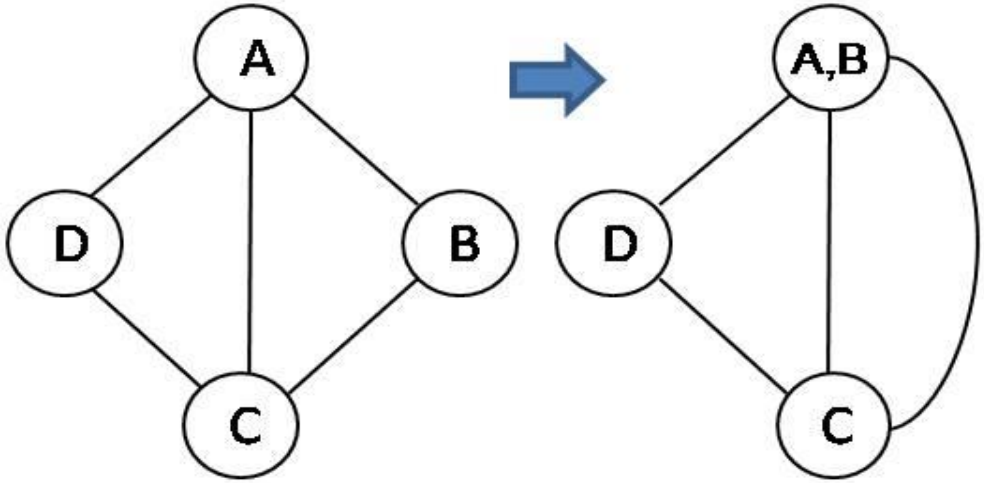
- Choose random x
- Let y := Bx
- Is Ay = Cx?



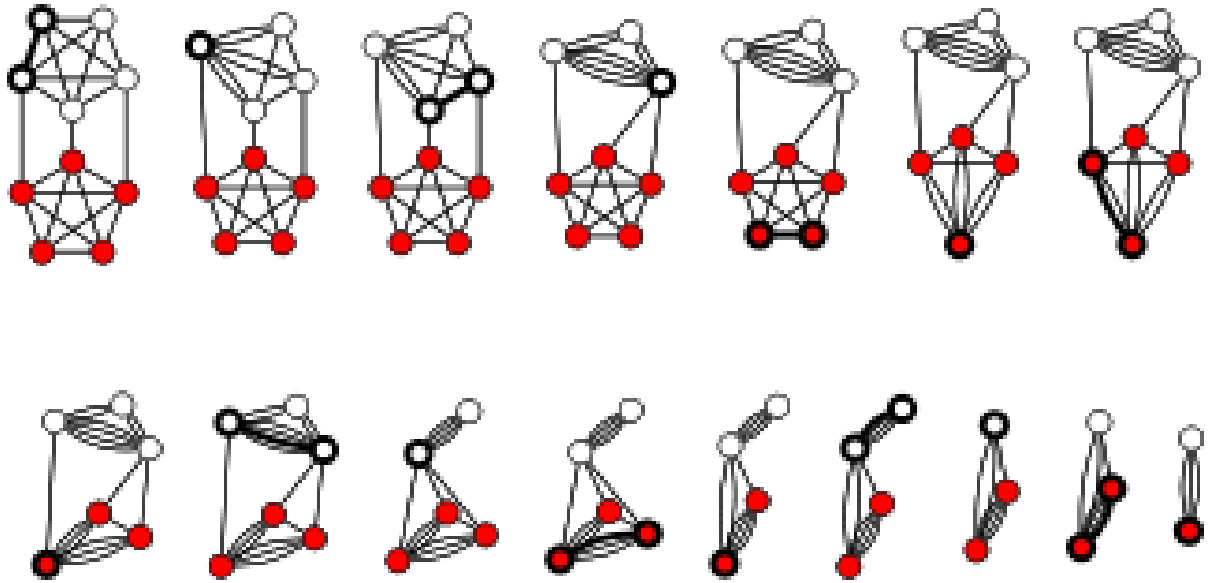
Min Cut



- Continue choosing an edge and contracting it



Run of Karger's algorithm



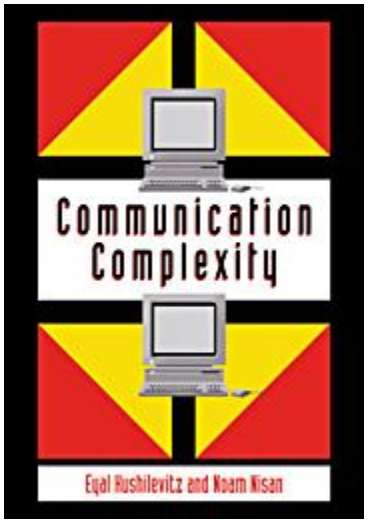
Communication complexity

Alice has X **Bob has Y**

EQ	000	001	010	011	100	101	110	111
000	1	0	0	0	0	0	0	0
001	0	1	0	0	0	0	0	0
010	0	0	1	0	0	0	0	0
011	0	0	0	1	0	0	0	0
100	0	0	0	0	1	0	0	0
101	0	0	0	0	0	1	0	0
110	0	0	0	0	0	0	1	0
111	0	0	0	0	0	0	0	1

Communication complexity of Equality

Deterministic algorithm requires communication at least length(x)



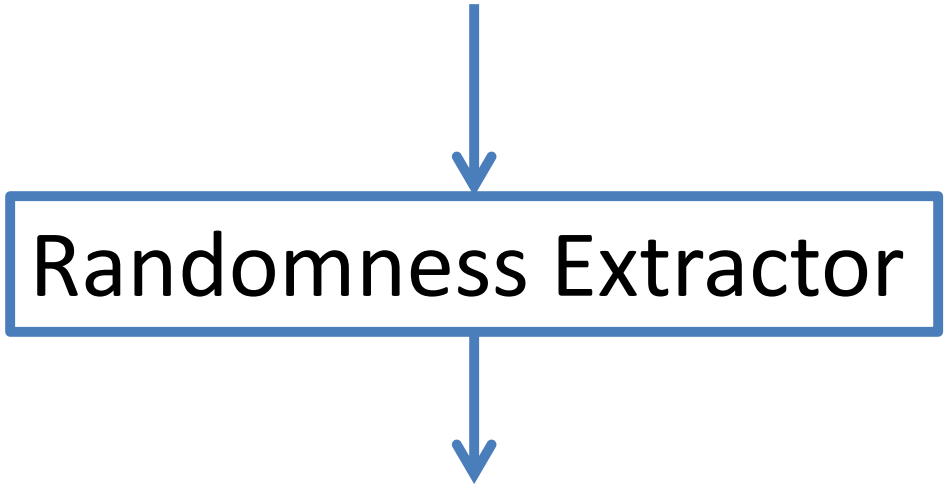
Randomized algorithm:

- Alice and Bob choose random string r
- Alice sends inner product of r, x mod 2
- Bob checks if this equals inner product of r, y

- Requires common randomness

Randomness

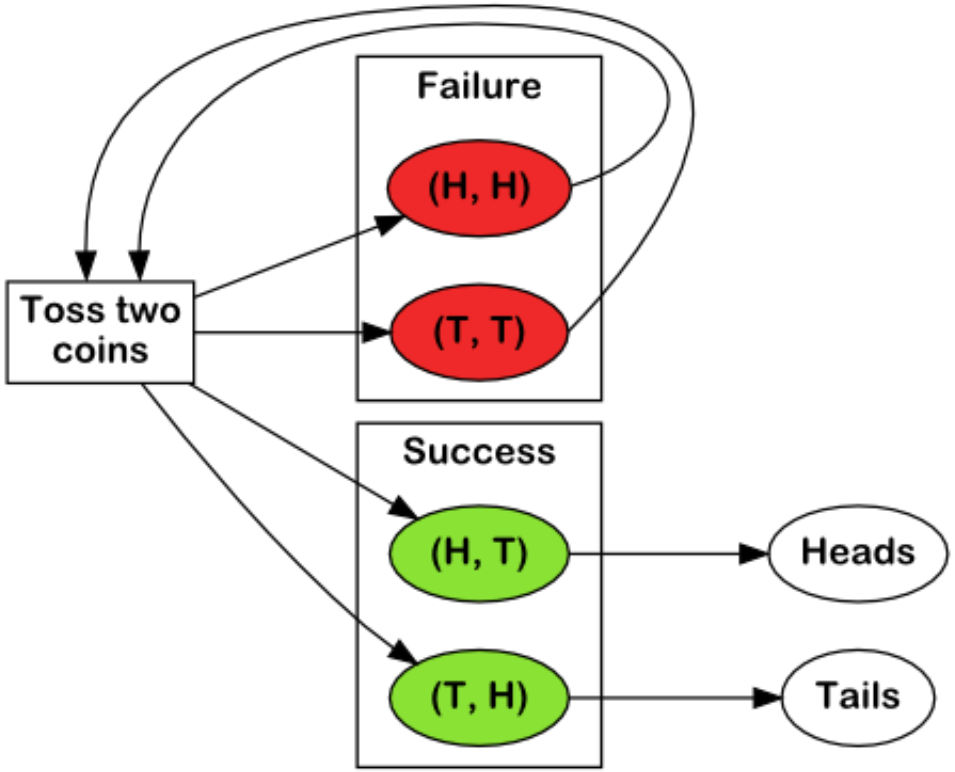
real-world imperfect source of randomness



perfect unbiased and independent bits

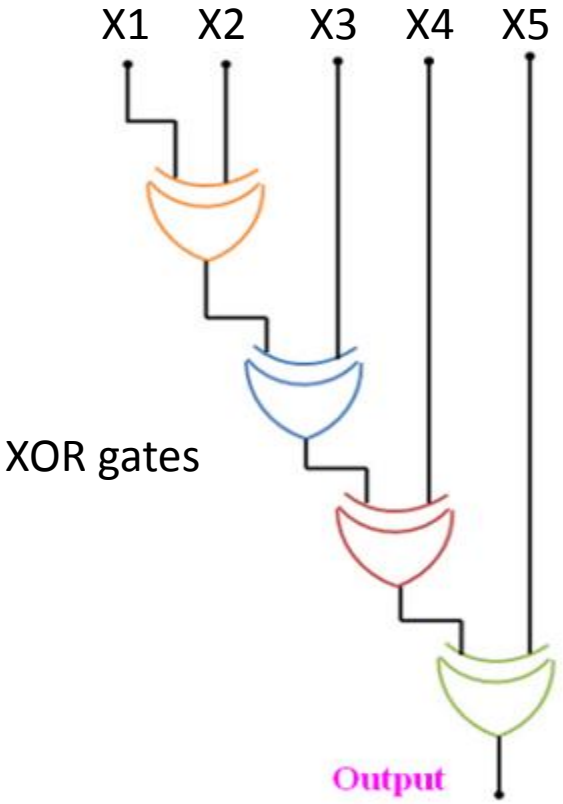
von Neumann extractor

Source: Independent trials of the same biased coin



Parity extractor

Source: Independent trials of different biased coins



Santha-Vazirani (SV) source

- Two coins

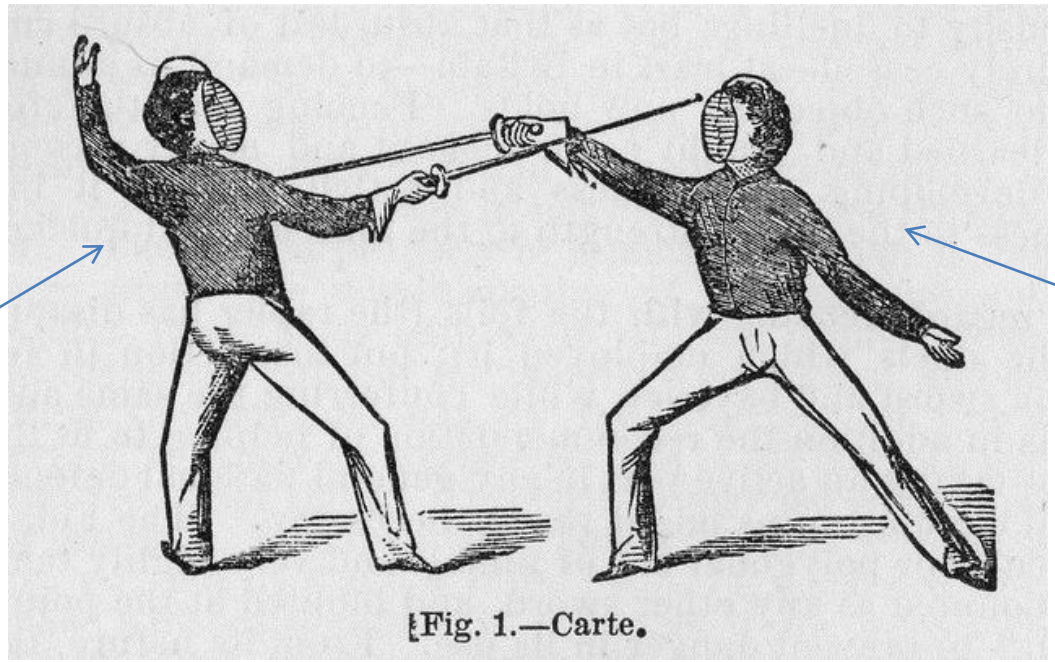
Pr[heads] = 1/3

Pr[heads] = 2/3

- Adversary each time, depending on previous outcomes, chooses one of the coins to toss

Deterministic Extractor for SV sources

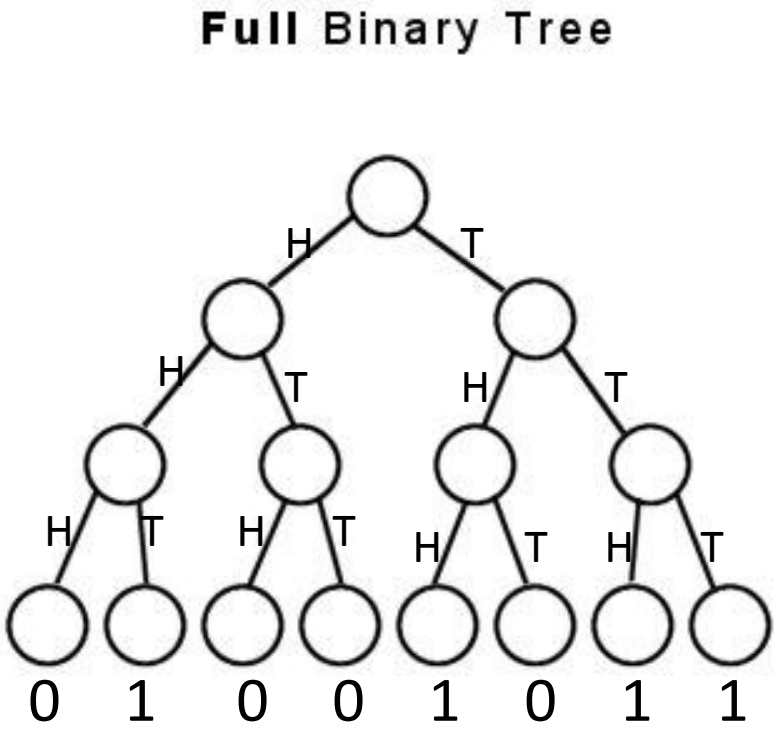
For every deterministic way of extracting one random bit, adversary has a strategy such that the extracted bit is 1 with probability ≥ 2/3 or ≤ 1/3.



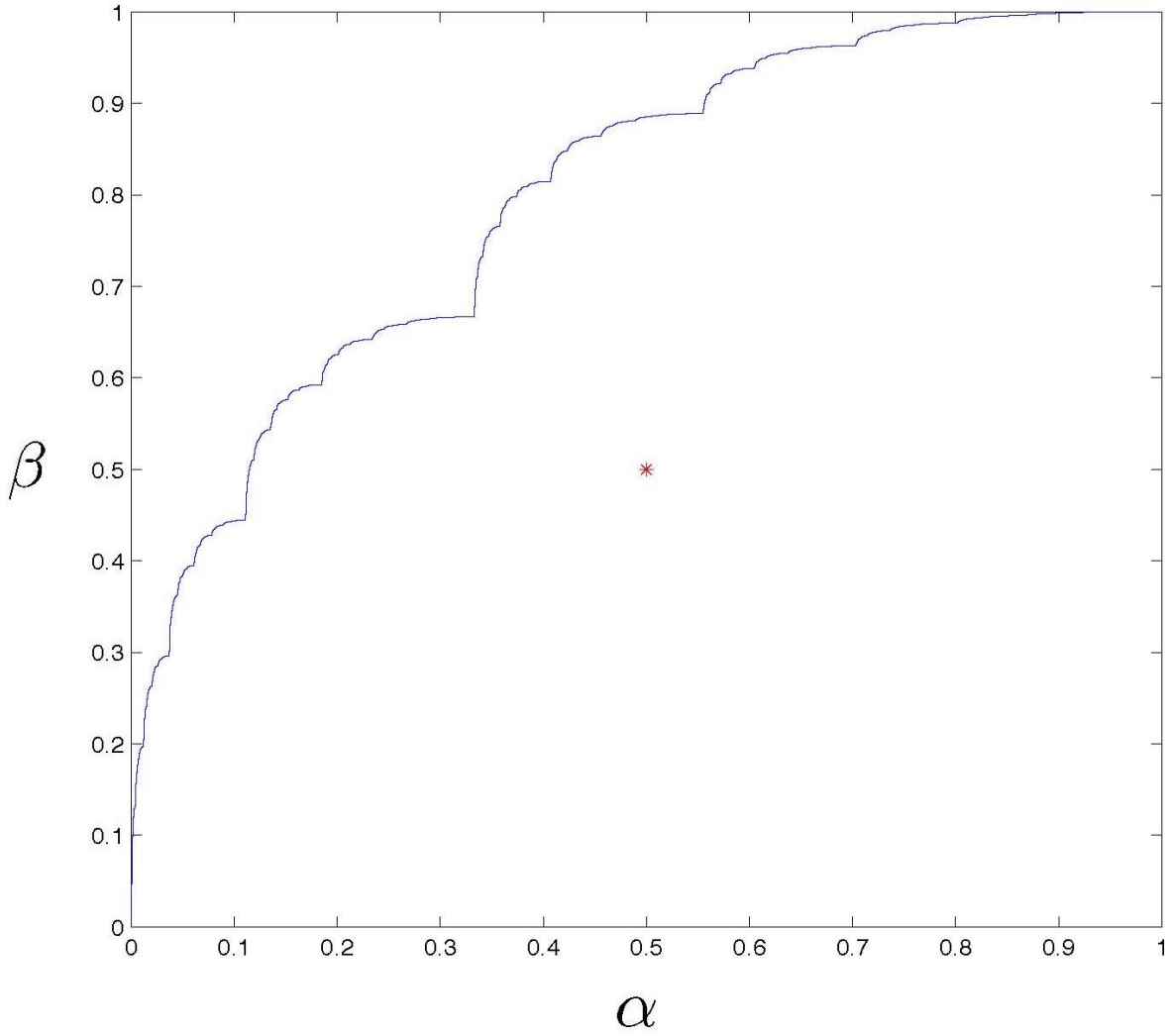
Deterministic extractor

can be thought as a tree with labeled leaves.

We argue recursively on subtrees.



Let α, β be min and max probability
given extractor outputs 1.
(α, β) lies above curve by induction.



SV sources have high min-entropy

The maximum probability of seeing any binary sequence is at most $(2/3)^n$.

Seeded extractor (e.g. Zuckerman)

source of high min-entropy

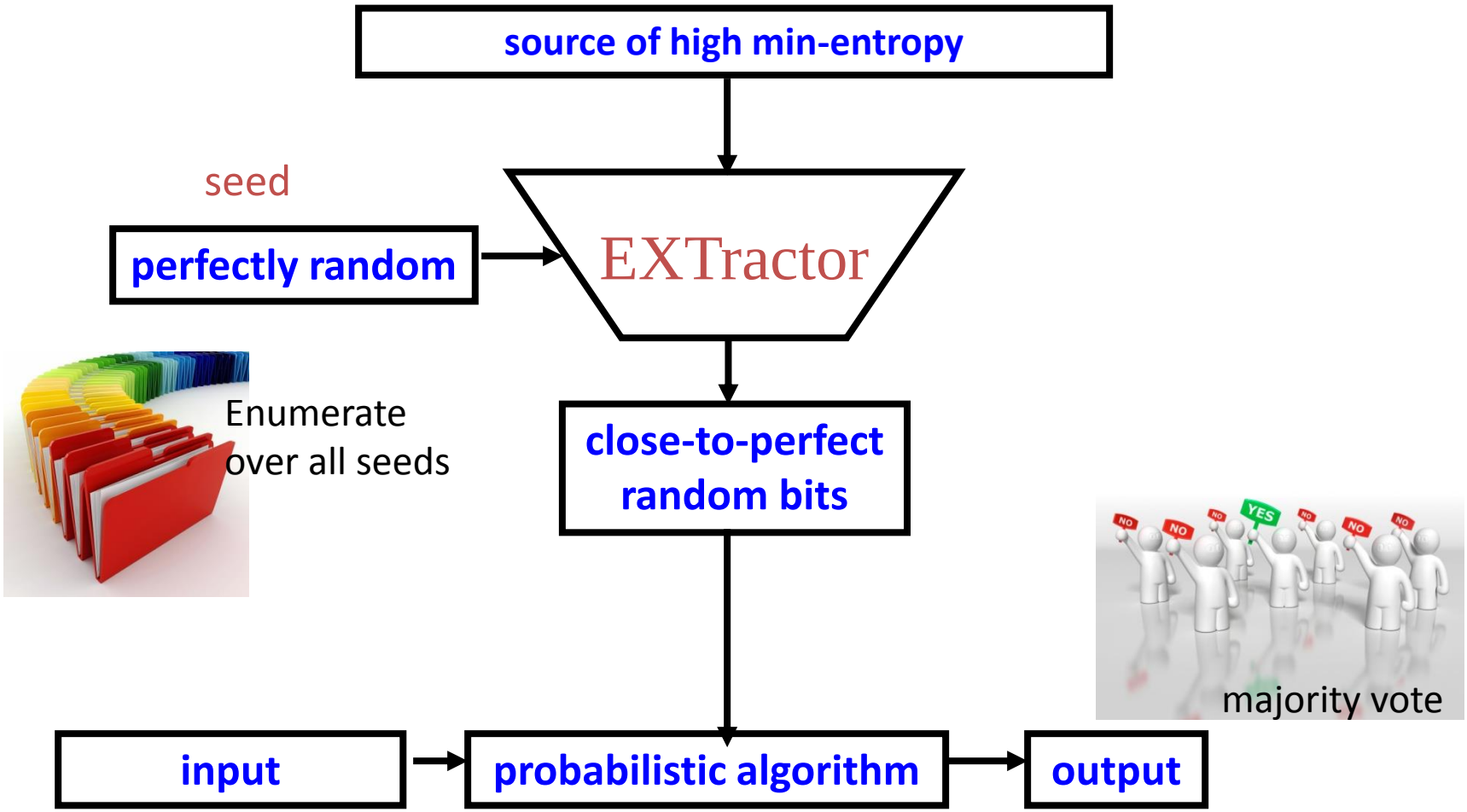
(short logarithmic) "seed"

perfectly random

EXTractor

close-to-perfect random bits

Simulating probabilistic algorithms



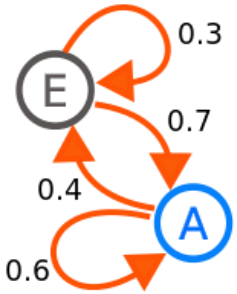
Need for efficient deterministic extractors

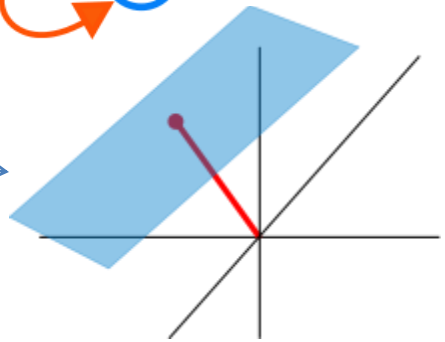
- Enumeration impossible in one-shot scenarios like cryptography or interactive proofs

- Extraction should be done efficiently (but the probabilistic method often gives implicit constructions)

Deterministic extractors

- i.i.d. with unknown bias [von Neumann]
- Markov chains [Blum]
- Affine sources [Bourgain, Gabizon-Raz]
- Polynomial sources [Dvir-Gabizon-Wigderson]
- Independent blocks [Bourgain]
- …





Deterministic extraction

for non-binary SV sources, unlike the binary case, is sometimes possible.

Generalized (non-binary) SV sources

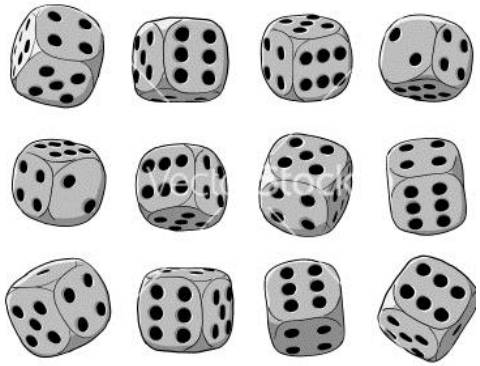
- The adversary has a set of dice. Different dice have different probabilities for a given face value.

- Adversary chooses each time which die to throw depending on past outcomes

Necessary and sufficient condition for extracting 1 bit

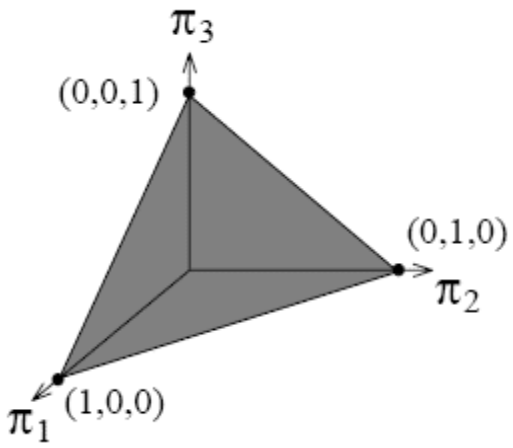
- Assume 6-sided dice that are non-degenerate
  (each face has non-zero probability)
- Extraction always possible for 5 dice

- Extraction often impossible for 6 or more dice

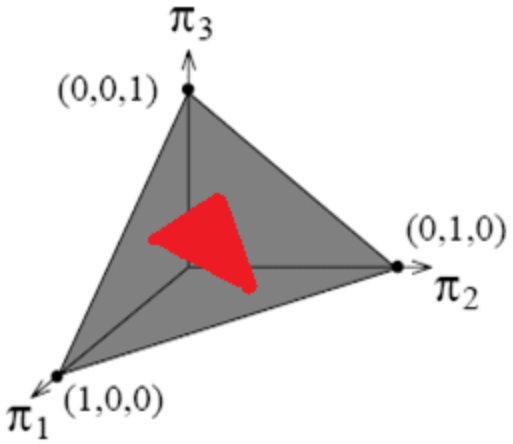


Probability Simplex

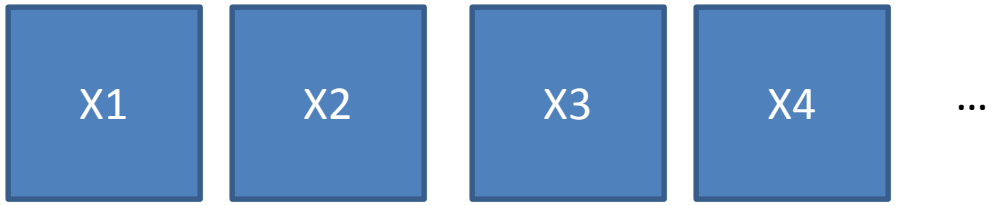
- Non-negative points in 6-dimensional Euclidean space whose coordinates sum to 1
- Each dice is one such point



- Extraction impossible iff the convex hull of the dice has interior in the simplex



Generalization of block sources

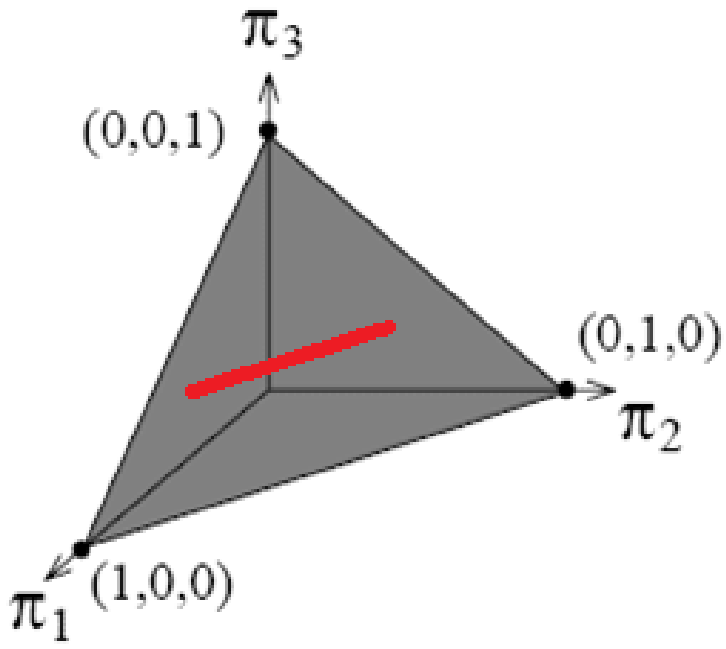


- Each block has some min-entropy conditioned on previous blocks

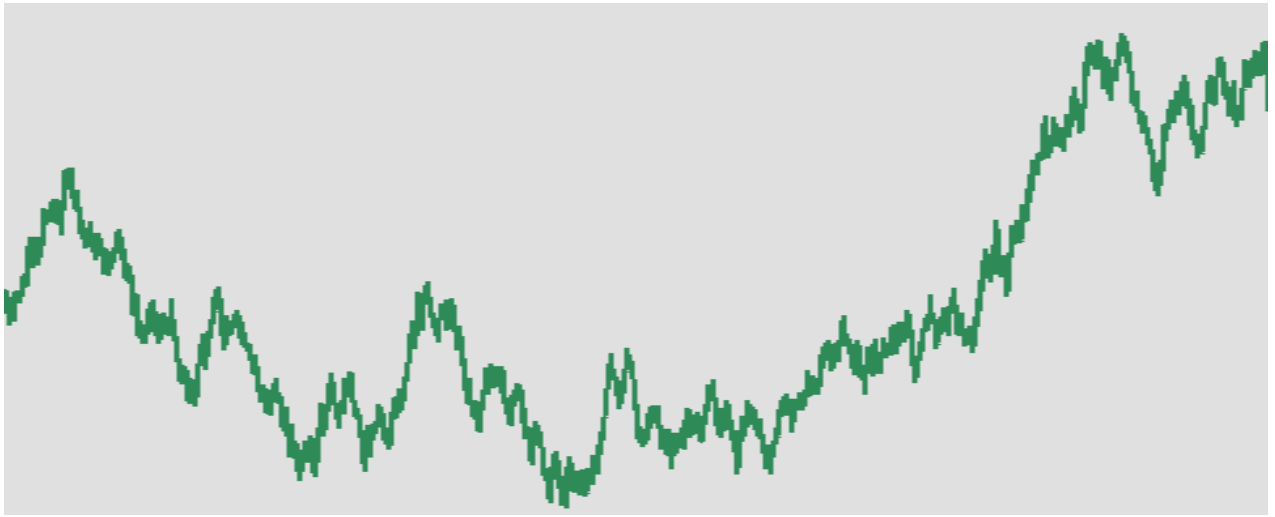
- Deterministic extraction from block sources is impossible

Proof of possibility of extraction

When the convex hull has no interior, there exist nonzero weights $w_1, w_2, w_3, w_4, w_5, w_6$ such that for die i we have $\sum_j p_i(j) w_j = 0.$



Construct a martingale



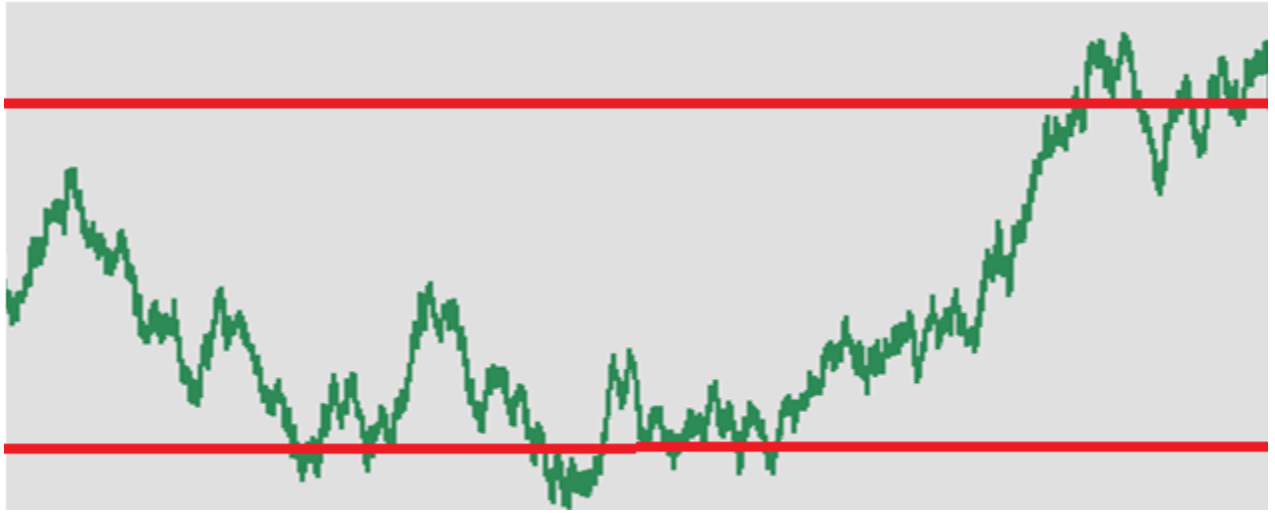
If $x_1, x_2, \ldots$ is the sequence of outcome of dice, then

$$y_i = \sum_{t=1}^{i-1} w_{x_t}$$

is a martingale.

Extract 1 bit

depending on whether the martingale first hits an upper limit or a lower limit.



With high probability, we hit one of the limits.

The bit is unbiased by the martingale stopping theorem

Common Randomness

used to synchronize actions in distributed algorithms.

Generate common randomness from i.i.d. sequence

(A_1, B_1)

(A_2, B_2)

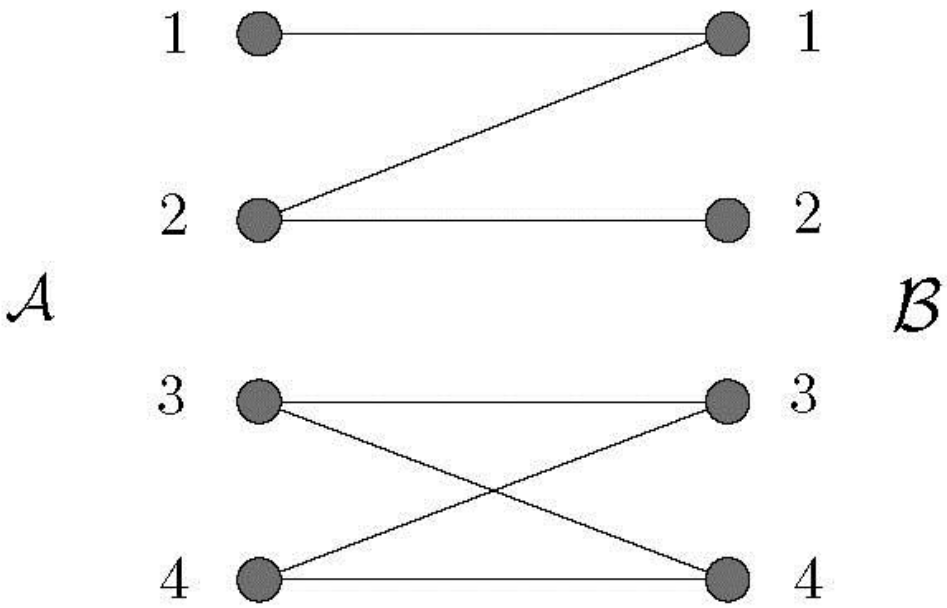
(A_3, B_3)

$\vdots$

=

[Gacs, Korner], [Witsenhausen]

In i.i.d. case, common randomness possible iff common data exists, i.e. common randomness reduces to extracting (ordinary) randomness.

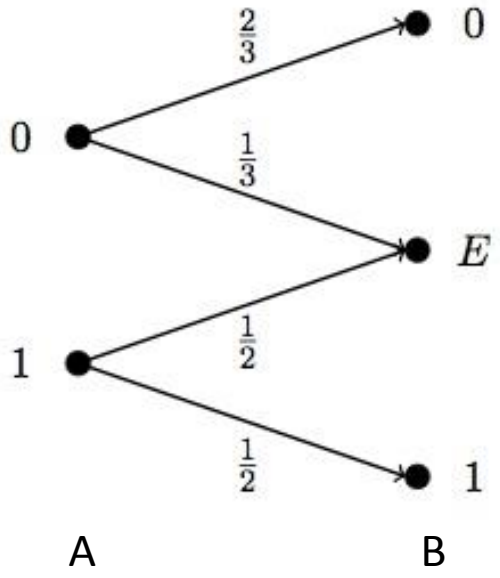


A \ B	1	2	3	4
1	0.1	0	0	0
2	0.1	0.2	0	0
3	0	0	0.1	0.1
4	0	0	0.2	0.2

Maximal correlation

Proof uses *tensorization* of *maximal correlation.*

E.g. maximal correlation of i.i.d. copies of



is still small.

Generate common randomness from distributed SV sources

The joint distribution of each pair is chosen by adversary out of a set of distributions, and may depend on previous pairs

(A_1, B_1)

(A_2, B_2)

(A_3, B_3)

$\vdots$

Common randomness possible only through extraction from common data

E.g. the superposition of the following three-component distributions still has two connected components and the common data is not enough

		B			
		1	2	3	4
A	1	0.1	0	0	0
	2	0	0.2	0	0
	3	0	0	0.1	0.1
	4	0	0	0.3	0.2

		B			
		1	2	3	4
A	1	0.2	0	0	0
	2	0.1	0.1	0	0
	3	0	0	0.3	0
	4	0	0	0	0.3

Proof of impossibility

Rather technical

We construct a continuous function that not only captures min and max probability of being equal to 1, but also the probability of agreement. The function has two terms:

- One term similar to non-distributed case
- One term inspired by maximal correlation


Thank
you!