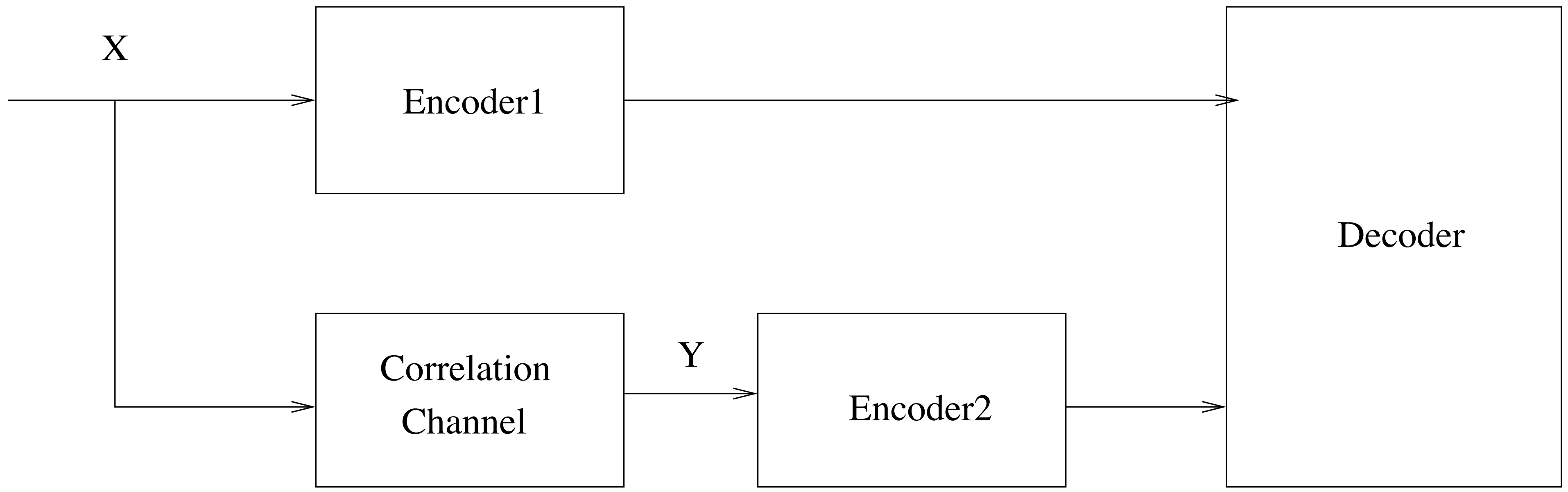
Fountain Codes for the Slepian-Wolf Problem



Amin Shokrollahi
EPFL

Joint work with B. Ndzana Ndzana (EPFL)

Allerton Conference, September 2006

Synopsis

- Fountain codes
- Applications of fountain codes in joint source-channel coding
- The Slepian-Wolf problem
- Fountain codes for the Slepian-Wolf problem
- Simulations
- Conclusions

Fountain codes

Allerton Conference, September 2006

Fountain Codes

Fountain codes are a class of codes designed for solving various data transmission problems, at the same time.

Fountain codes with fast encoding and decoding algorithms, and (arbitrarily) small overhead are particularly interesting for solving these problems.

Fountain codes were stipulated by Byers et al in 1998, and their applications discussed. A construction was, however, not given.

First construction of efficient Fountain codes was given by Luby (1998, published 2002).

(Binary) Fountain Codes

Fix distribution $\mathcal{D}$ on $(\mathbb{F}_2^k)^*$, where k is the number of input symbols.

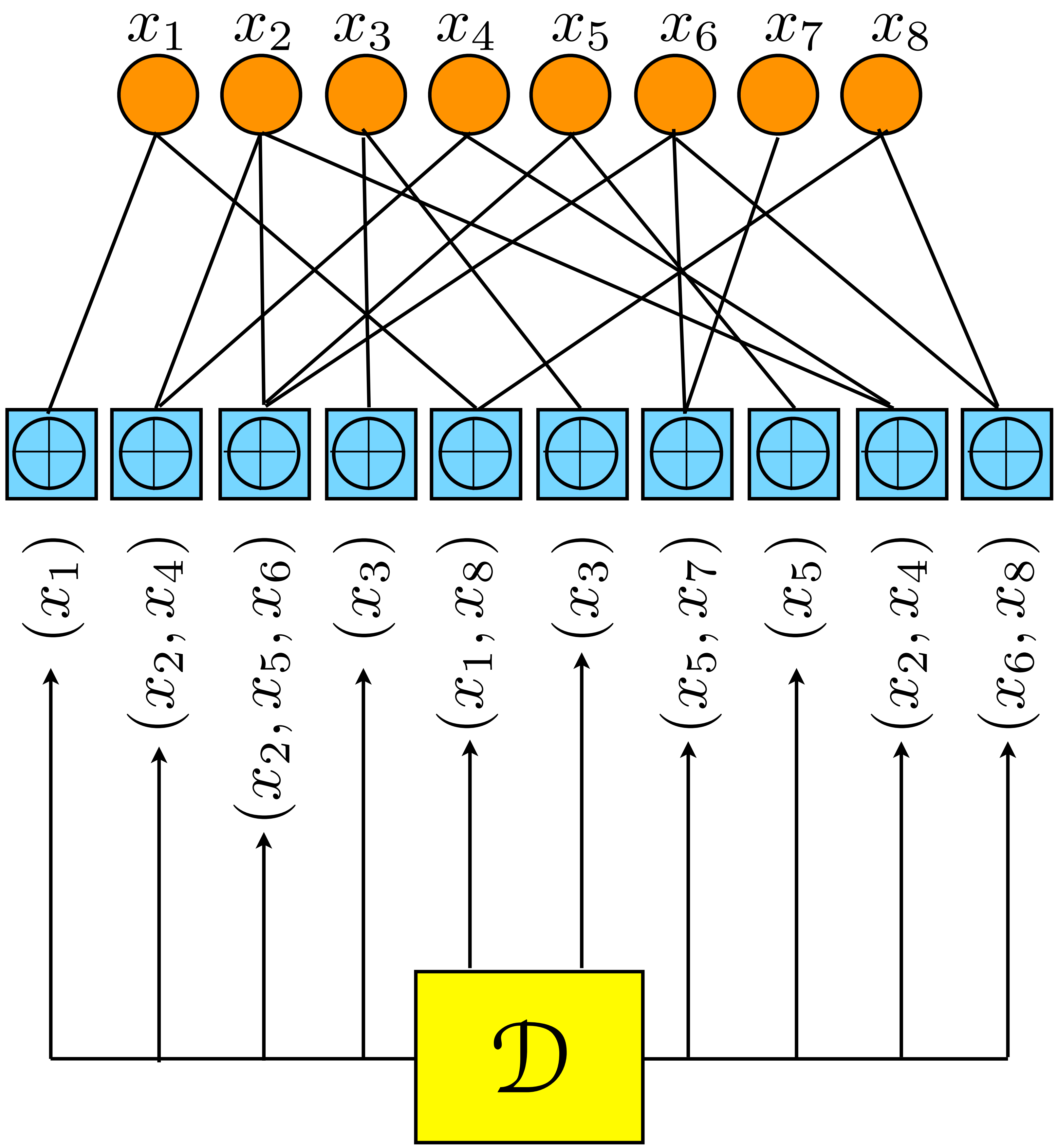
A fountain code with parameters $(\mathcal{D}, k)$ is a vector in $\left((\mathbb{F}_2^k)^*\right)^{\mathbb{N}}$ sampled from $\mathcal{D}^{\mathbb{N}}$.

Distribution $\mathcal{D}$ can be identified with a distribution $\mathfrak{D}$ on $\mathbb{F}_2^k$.

Operation:
For each output symbol sample independently from $\mathfrak{D}$ and add symbols corresponding to the sampled output.

Symbols are understood to be binary vectors, and additions are understood to be over $\mathbb{F}_2$.

(Binary) Fountain Codes



Example: LT Codes

Invented by Michael Luby in 1998.

First class of universal and almost efficient Fountain Codes.

Output distribution has a very simple form.

Encoding and decoding are very simple.

LT Codes

Fix distribution $(\Omega_1, \Omega_2, \ldots, \Omega_k)$ on $\{1, \ldots, k\}$

Distribution $\mathcal{D}$ is given by

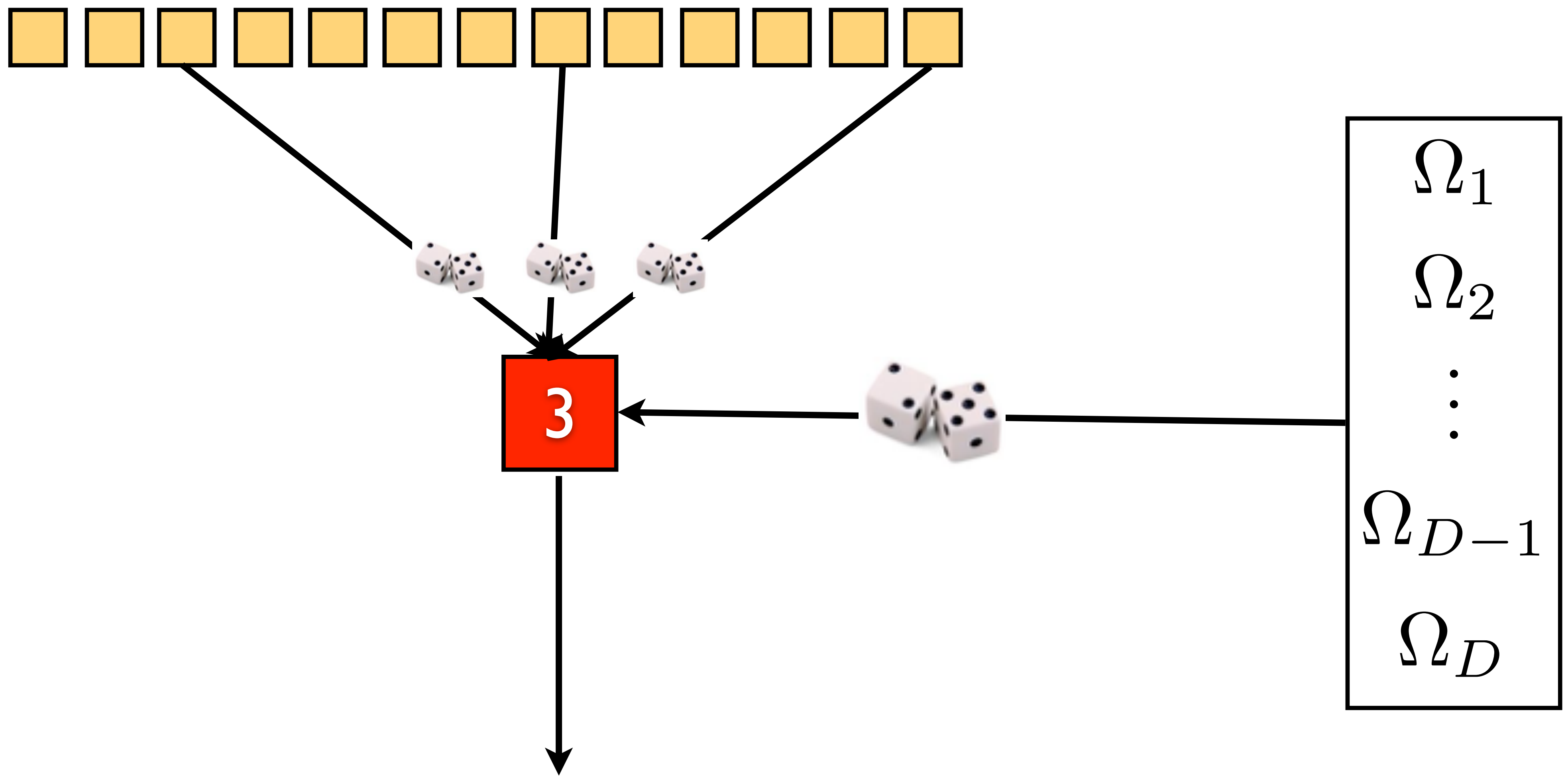
$$\Pr_{\mathcal{D}}(x) = \frac{\Omega_w}{\binom{k}{w}}$$

where w is the Hamming weight of x.

Parameters of the code are $(k, \Omega(x))$

$$\Omega(x) = \sum_{w=1}^{k} \Omega_w x^w$$

LT Coding Process



Example: Raptor Codes

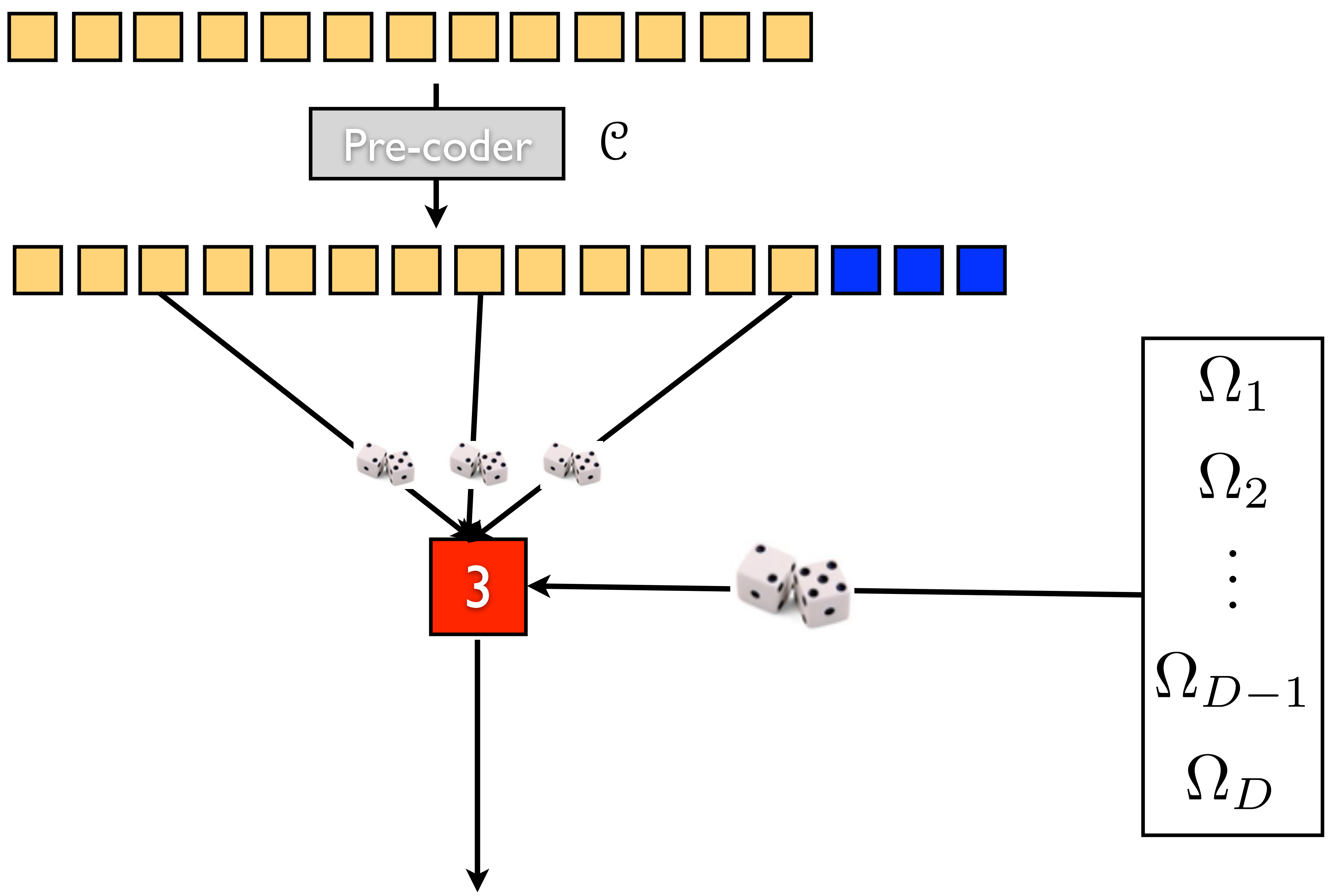
Invented by A. S. in late 2000.

First class of universal fountain codes with linear time encoding/decoding.

Extends LT-codes by using a pre-code.

Encoding and decoding are very simple.

Raptor Codes



Parameters: $(k, \mathcal{C}, \Omega(x))$

Raptor codes are fountain codes

Fountain codes & Compression

Fountain Codes and Compression

Fountain codes were first used for lossless compression in a series of two papers by Caire, Shamai, S, and Verdu.

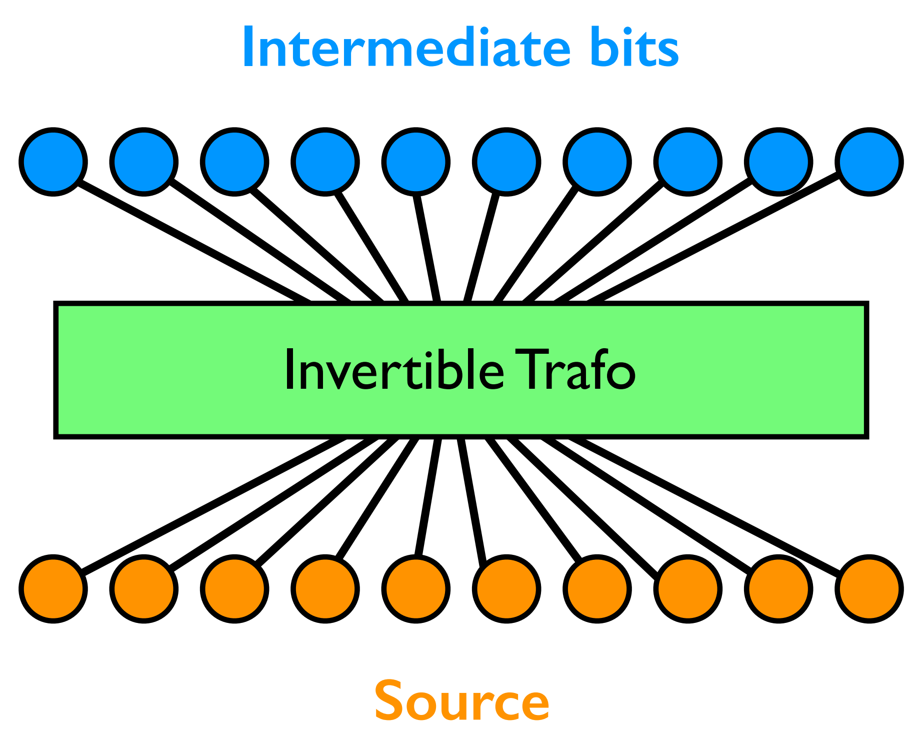
The procedure was modeled after the LDPC compression algorithms of Caire, Shamai, and Verdu, with a twist.

But what makes them interesting?

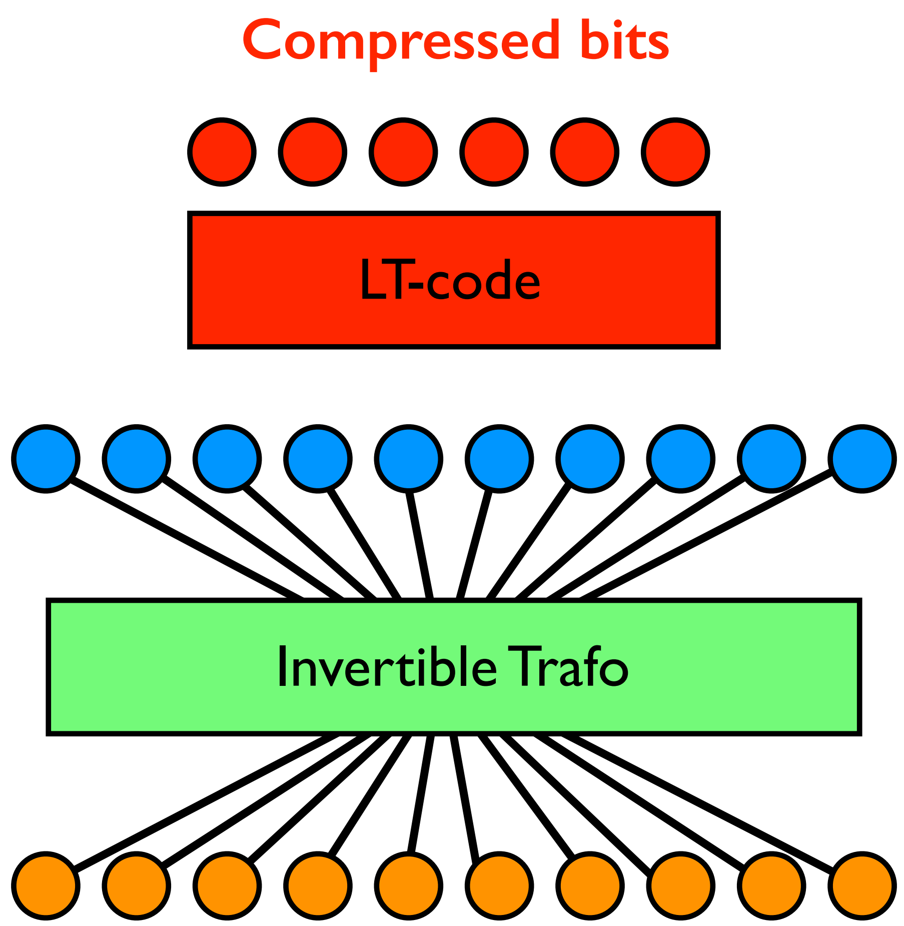
The “rateless” feature allows for two interesting applications:

1. It can be used when the source statistics is a-priori unknown, and
2. It can be used as a joint source-channel coding scheme (without the use of puncturing).

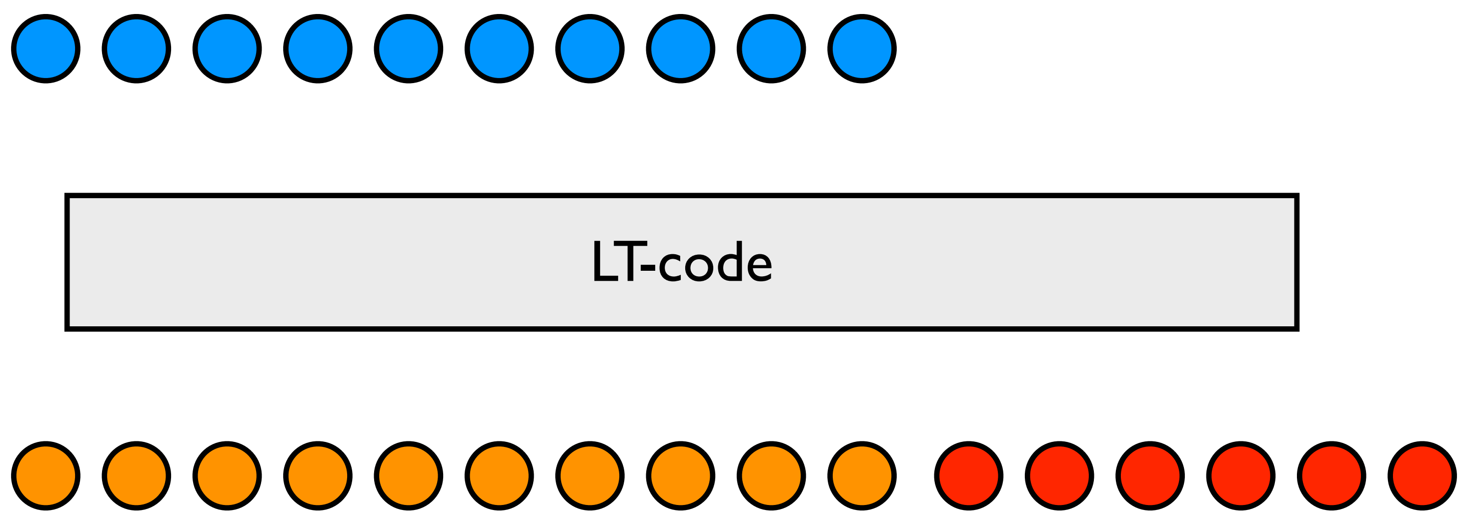
How Does it Work?



How Does it Work?

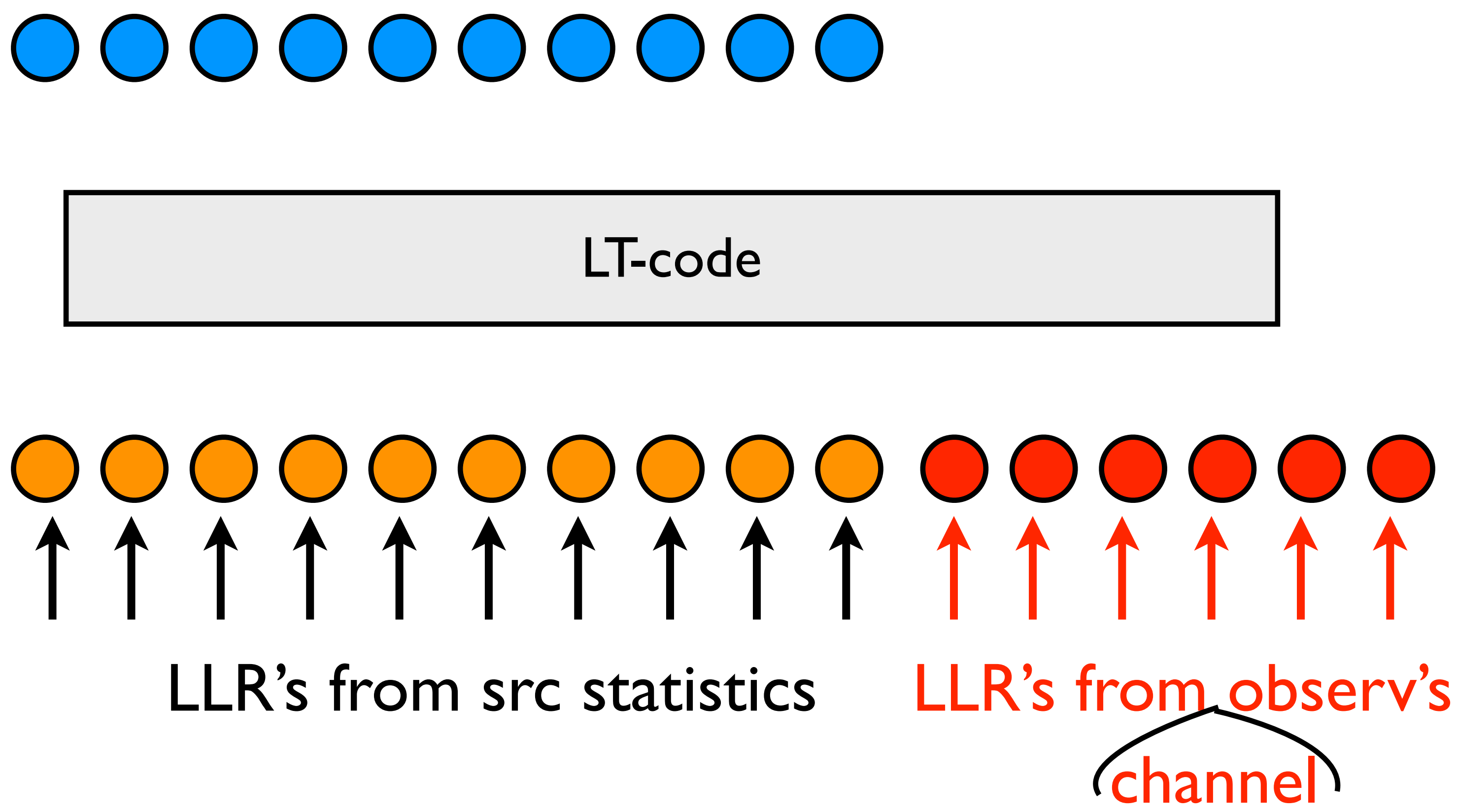


How Does it Work?



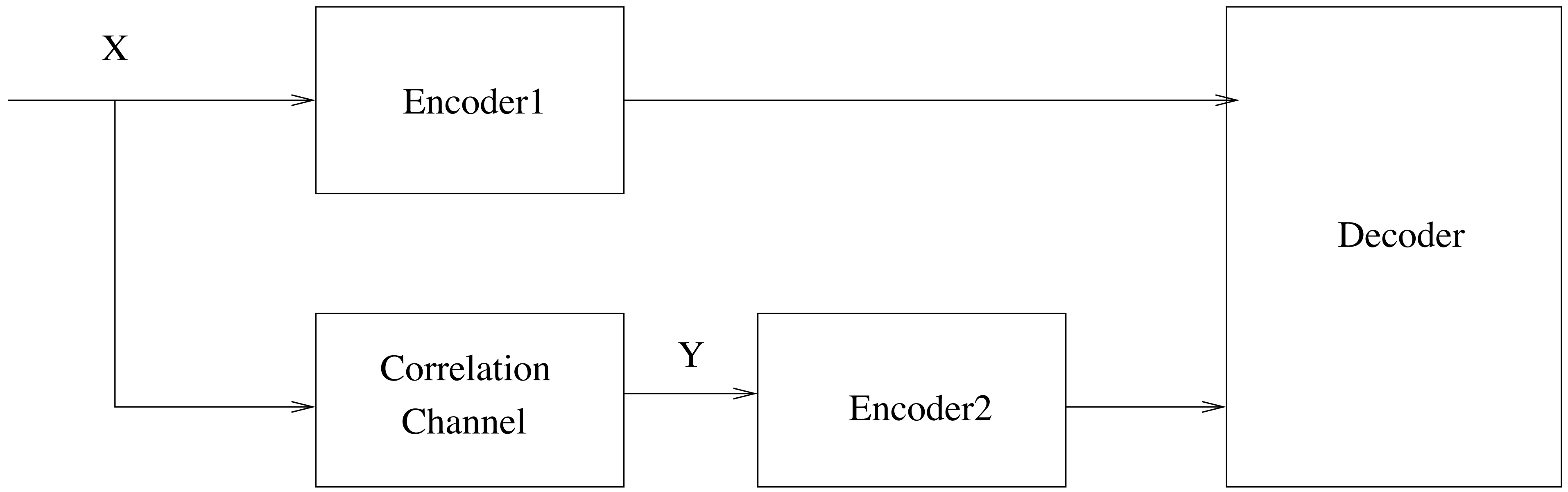
Combination with Channel Coding

BP-decoding with doping



Slepian-Wolf Problem

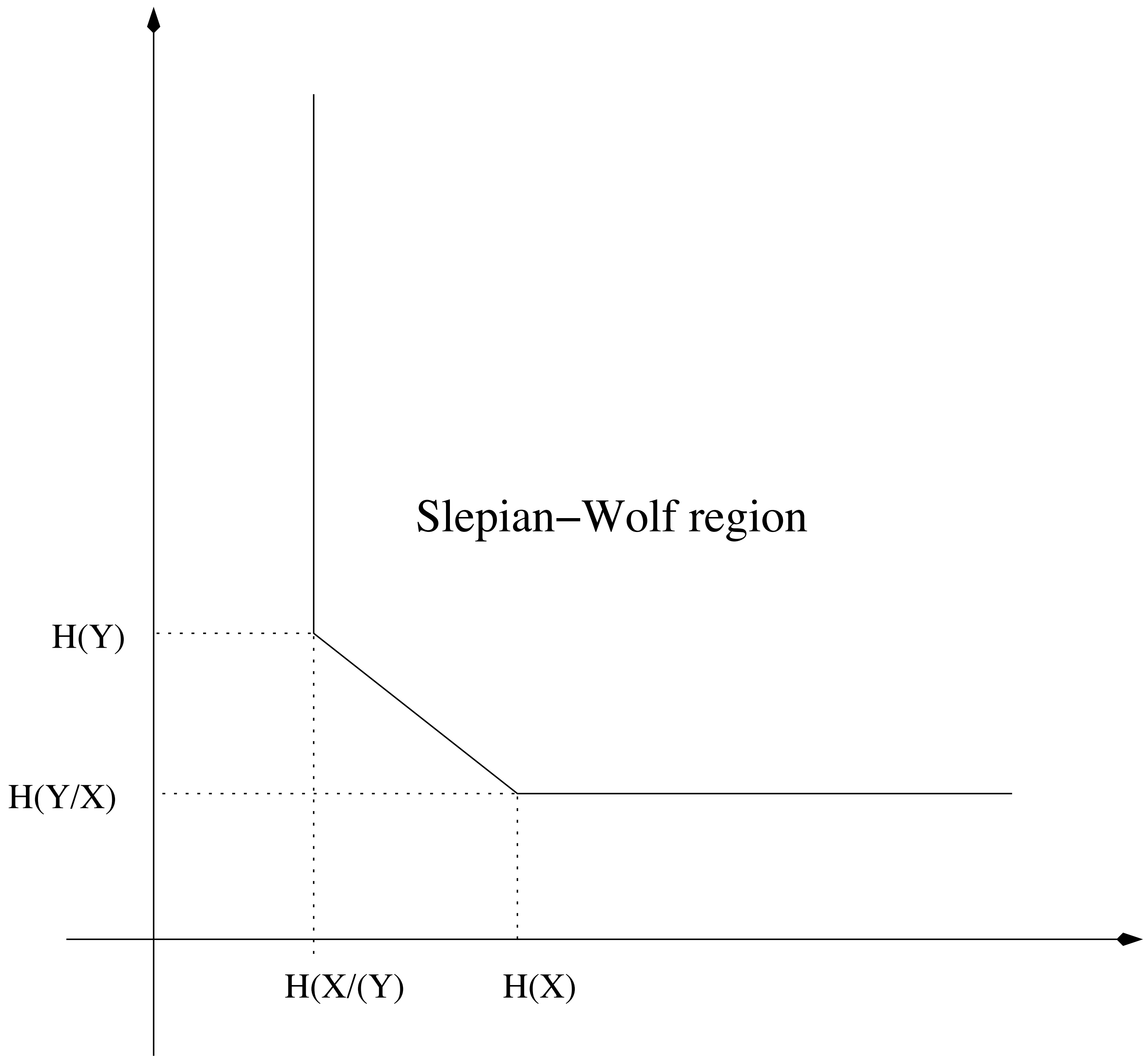
Problem Definition



1. Want to communicate with X and Y with $H(X,Y)$ bits.

2. Sources X and Y do not communicate with one another.

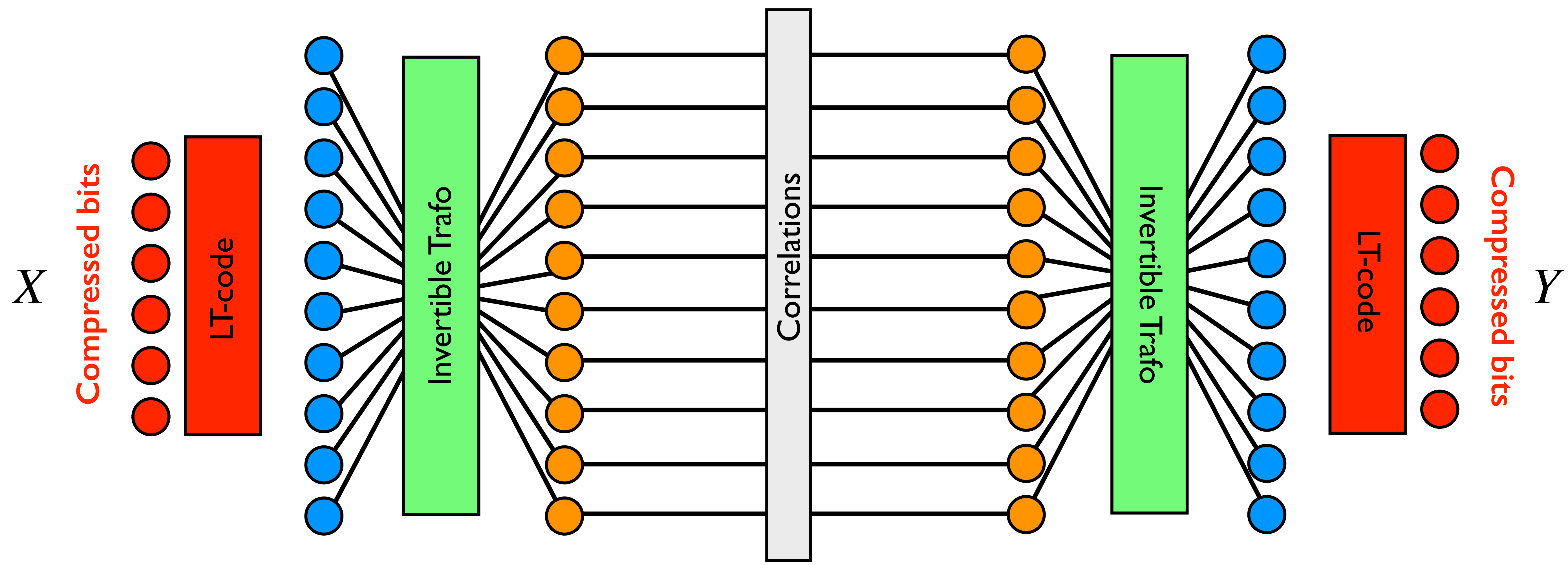
Achievable Region



Fountain Slepian-Wolf

The Scheme - Part I

General Slepian-Wolf with 2 users

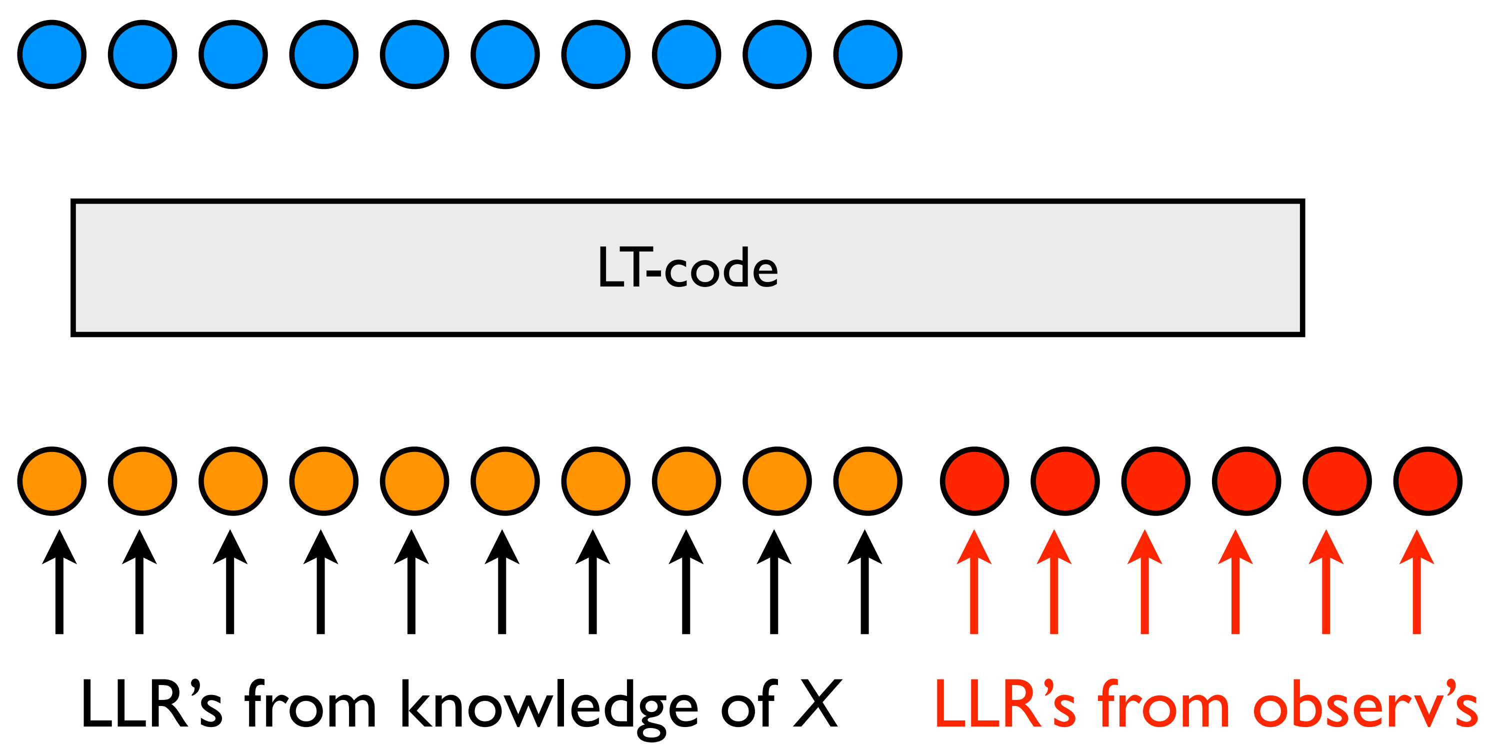


BP-decoding on the entire graph

Idea related to paper by Eckford-Yu

The Scheme - Part II

Slepian-Wolf with side information



BP-decoding on the entire graph

The Scheme - Part III

To make the scheme work, we need to add a few features.

1. Blind iterative doping (BLID): Y sends bits, and if decoding is not possible, a small fraction of the intermediate bits need to be sent as well.

2. Use of a repeat-accumulate mechanism: Y does not send the compressed bits directly, but uses an accumulator to do so (details in the paper). Leads to better performance.

3. Incremental transmission of the compressed bits, since encoder is unaware of the correlation.

Simulations

Simulations

1. Correlation between the sources is a BSC(p) with p unknown to the encoders.

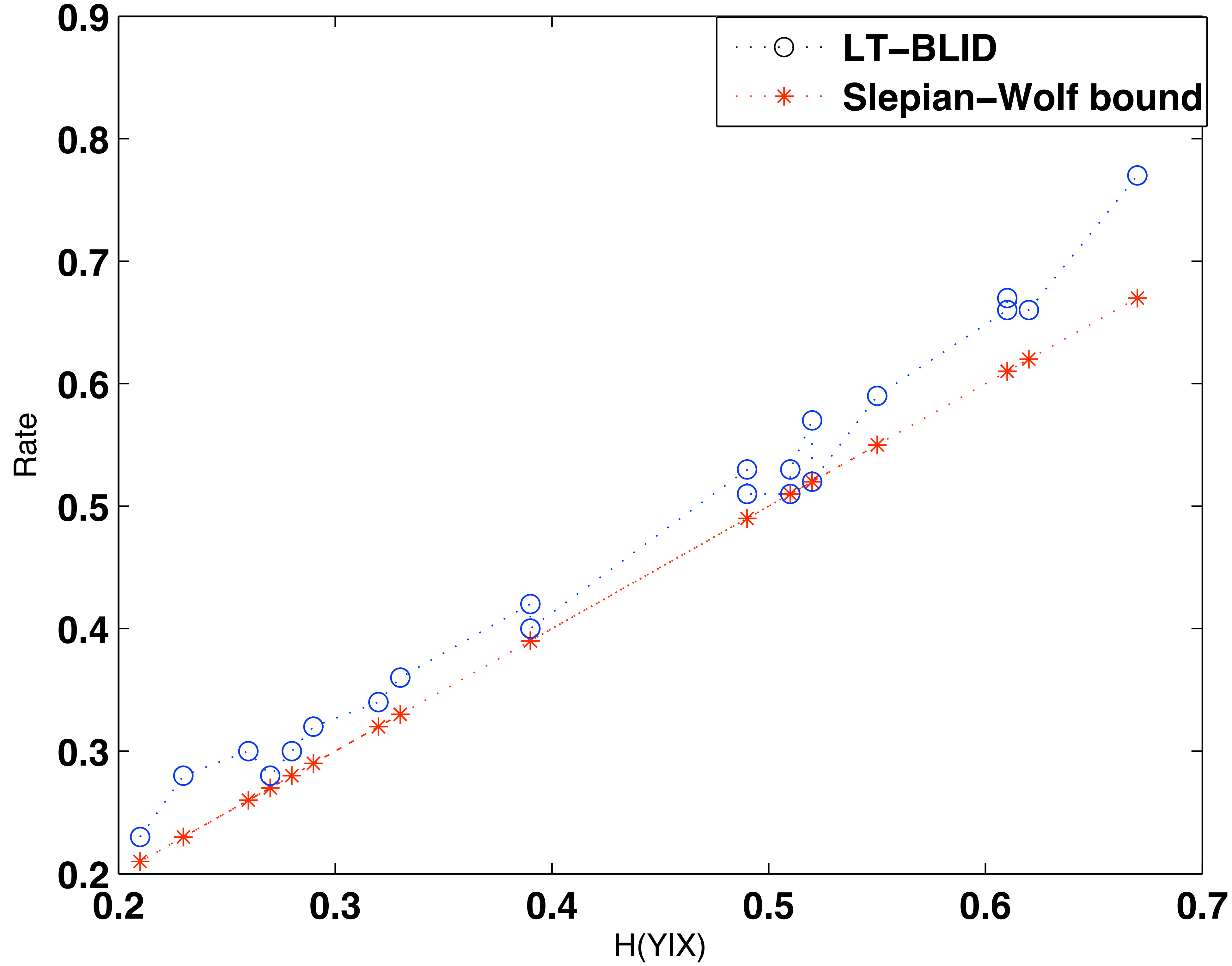
2. $\Pr[x_i = y_i] = 1 - p$

3. $H(Y \mid X)$ is between 0.2 and 0.7

4. Source X is completely available at the decoder (side information)

5. Length n of the sequences is 396

Simulations



Conclusions

Conclusions and Open Questions

1. Presented a fountain code scheme for the Slepian-Wolf.

2. Method can be used in conjunction with transmission over a noisy channel

3. Method can be used when the correlations are more complicated than the BSC

4. Question: how about the non-corner cases of the Slepian-Wolf region?

5. How about having more than two users?

6. Code design for the general case?