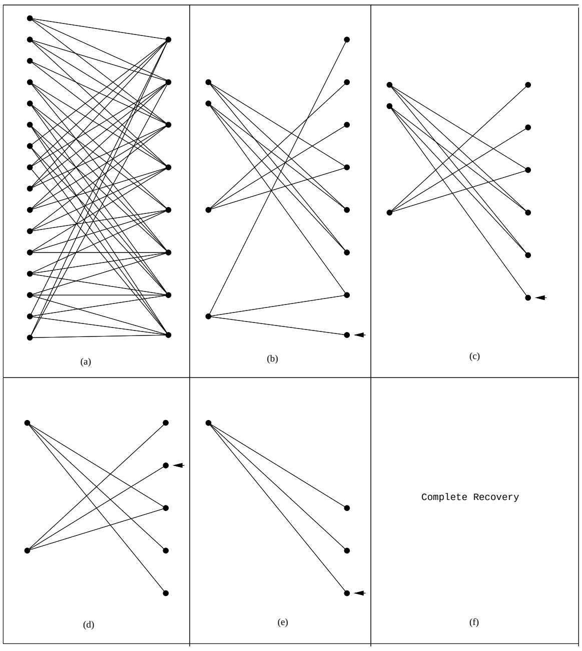
Low-Density Low-Complexity Codes



M. Amin Shokrollahi

Lucent Technologies
Bell Labs Innovations

Low-Complexity Codes

We will outline in this talk the design and analysis of error-correcting codes that can be encoded and decoded efficiently and protect against a fraction of errors that is almost as large as given by theoretical upper bounds.

Existence of such bounds and codes that asymptotically meet these bounds was proved in the landmark paper of C.E. Shannon in 1948.

Several codes can be proved to meet the asymptotic bounds. Almost none of them are equipped with efficient encoders and decoders.

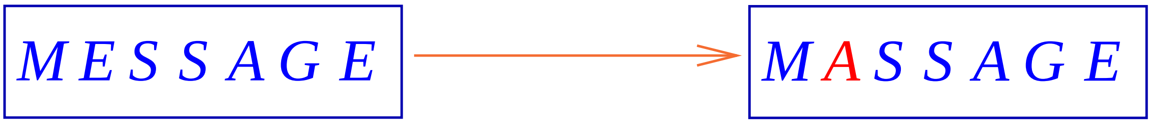
(Brief) History

Gallager	1963
Zyablov	1971
Zyablov-Pinsker	1976
Tanner	1981
Turbo Codes	1993
Berroux-Glavieux-Thitimajshima	
Sipser-Spielman, Spielman	1995
MacKay-Neal, MacKay	1995
Luby-Mitzenmacher-S-Spielman-Stemann	1997
Luby-Mitzenmacher-S-Spielman	1998
Richardson-S-Urbanke	1999

Codes

Codes are used when messages are to be transmitted over noisy channels.

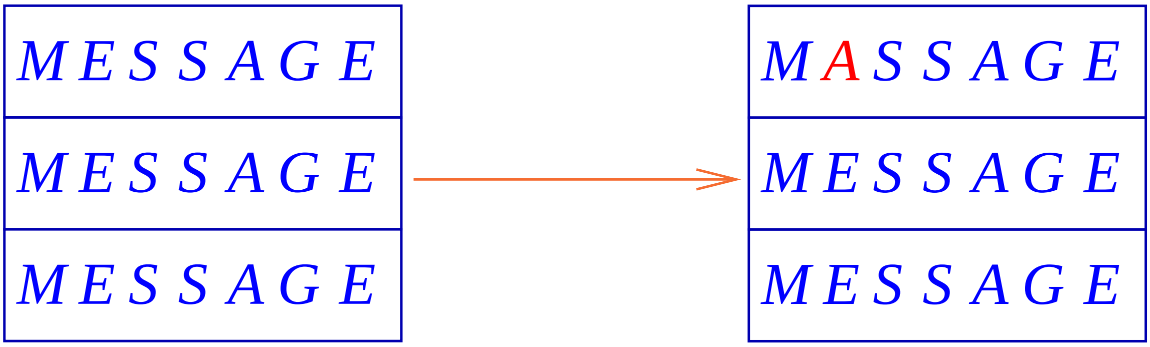
Without coding



MESSAGE

MASSAGE

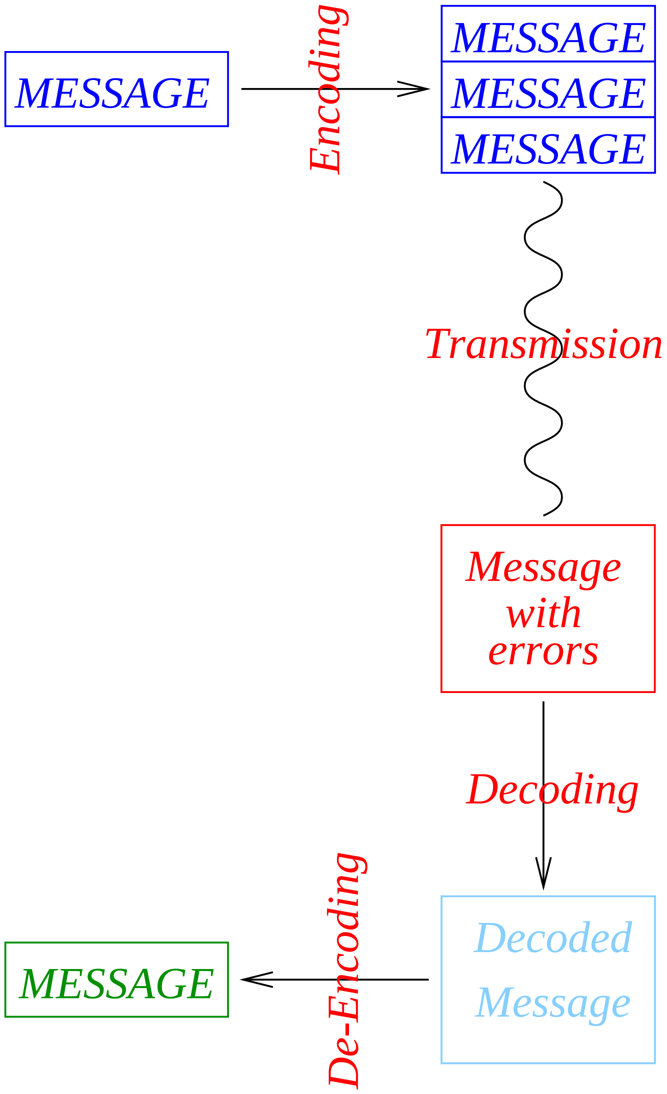
With coding



MESSAGE
MESSAGE
MESSAGE

MASSAGE
MESSAGE
MESSAGE

Encoding and Decoding



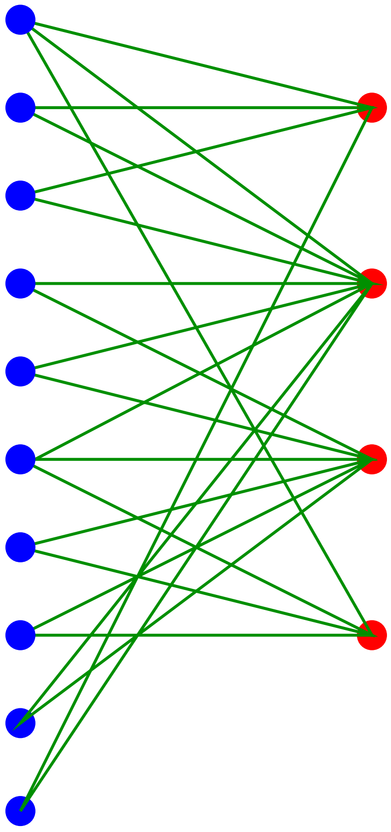
Encoding^{-1} $\neq$ Decoding!

Low Density Parity Check Codes

Were first introduced by R. Gallager in the early 1960's.

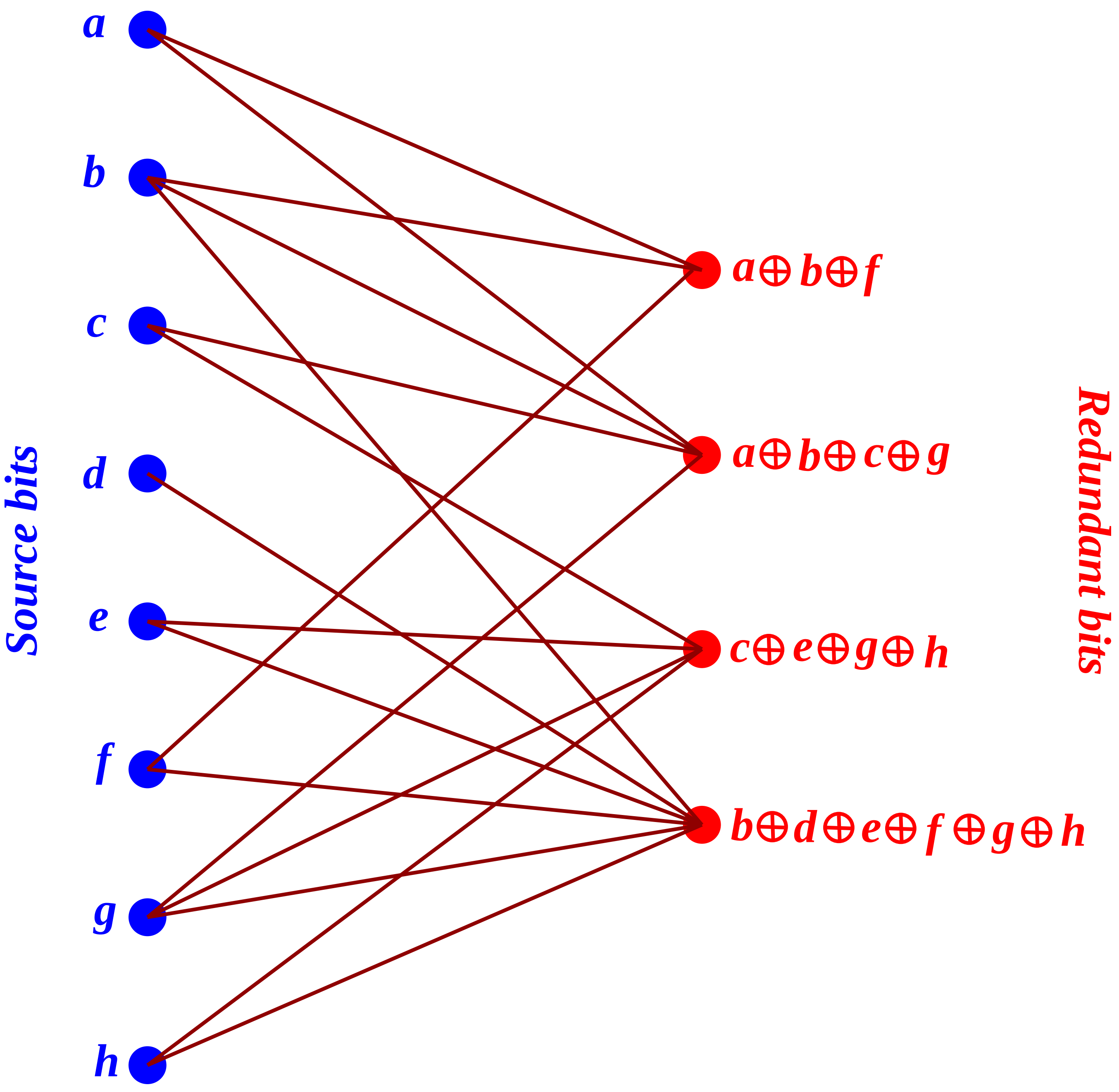
Codes are built from sparse bipartite graphs.

Encoding and Decoding are simple.



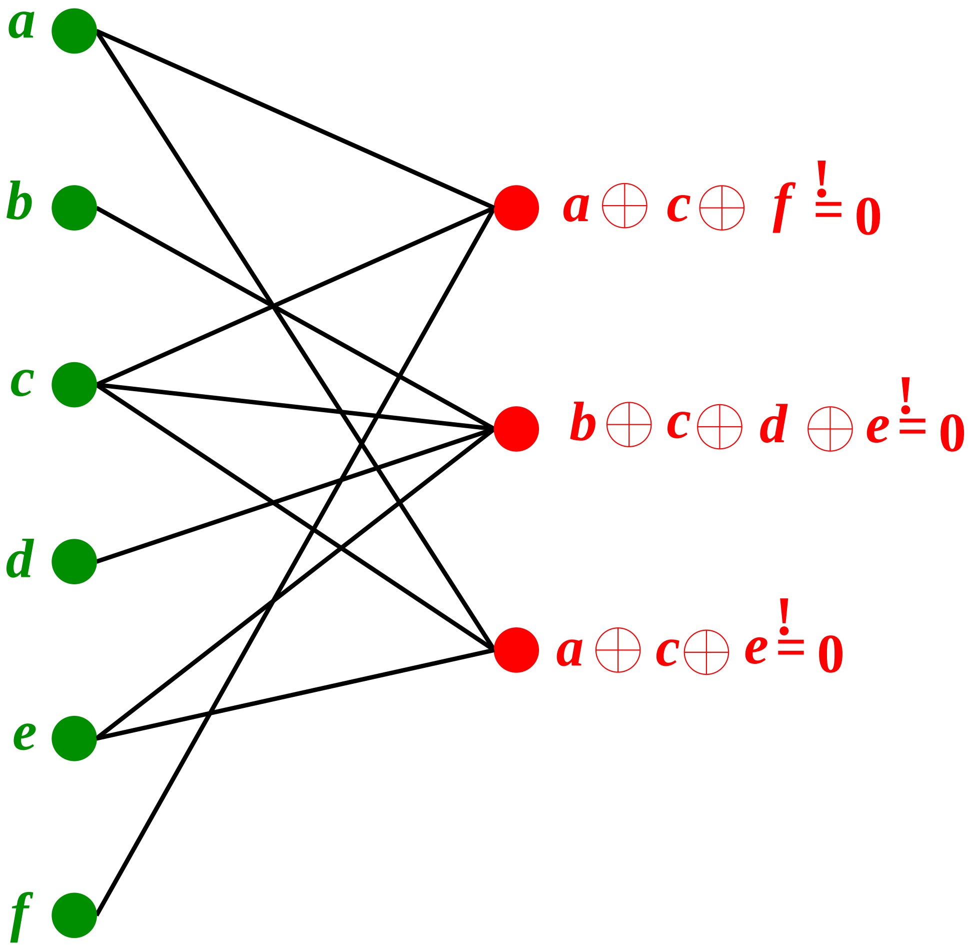
Encoding with Bipartite Graphs

Take a bipartite graph between k nodes on the left and $n-k$ nodes on the right. Label left nodes with the k packets to be encoded. Label right nodes with the redundant packets. Compute value of each right node as XOR of values of adjacent left nodes.


Source bits
a
b
c
d
e
f
g
h
$a \oplus b \oplus f$
$a \oplus b \oplus c \oplus g$
$c \oplus e \oplus g \oplus h$
$b \oplus d \oplus e \oplus f \oplus g \oplus h$
Redundant bits

Encoding time is proportional to number of edges in graph.

Original Gallager Codes



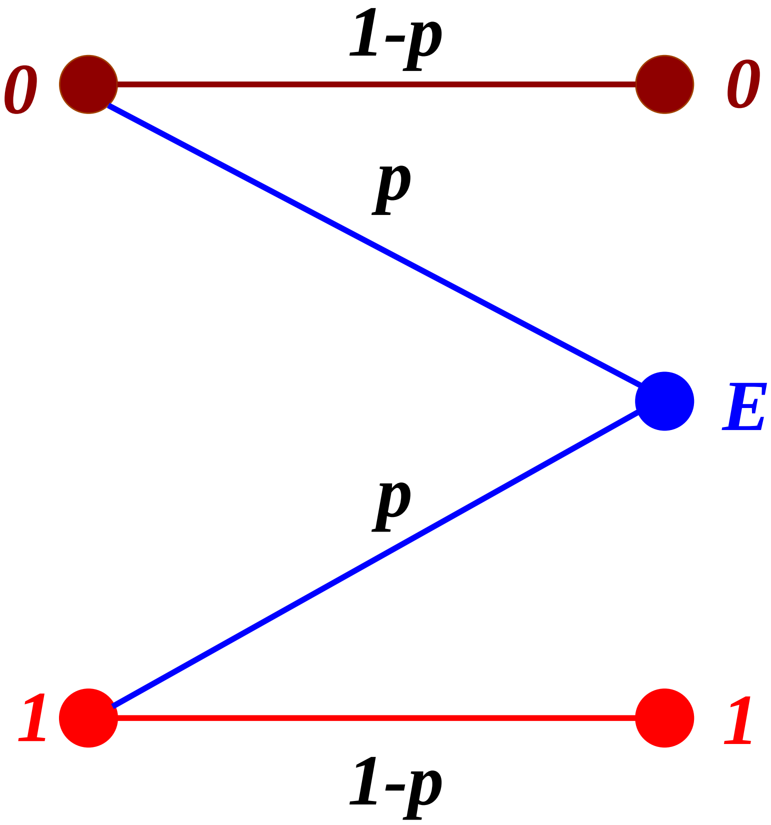
$a \oplus c \oplus f \overset{!}{=} 0$

$b \oplus c \oplus d \oplus e \overset{!}{=} 0$

$a \oplus c \oplus e \overset{!}{=} 0$

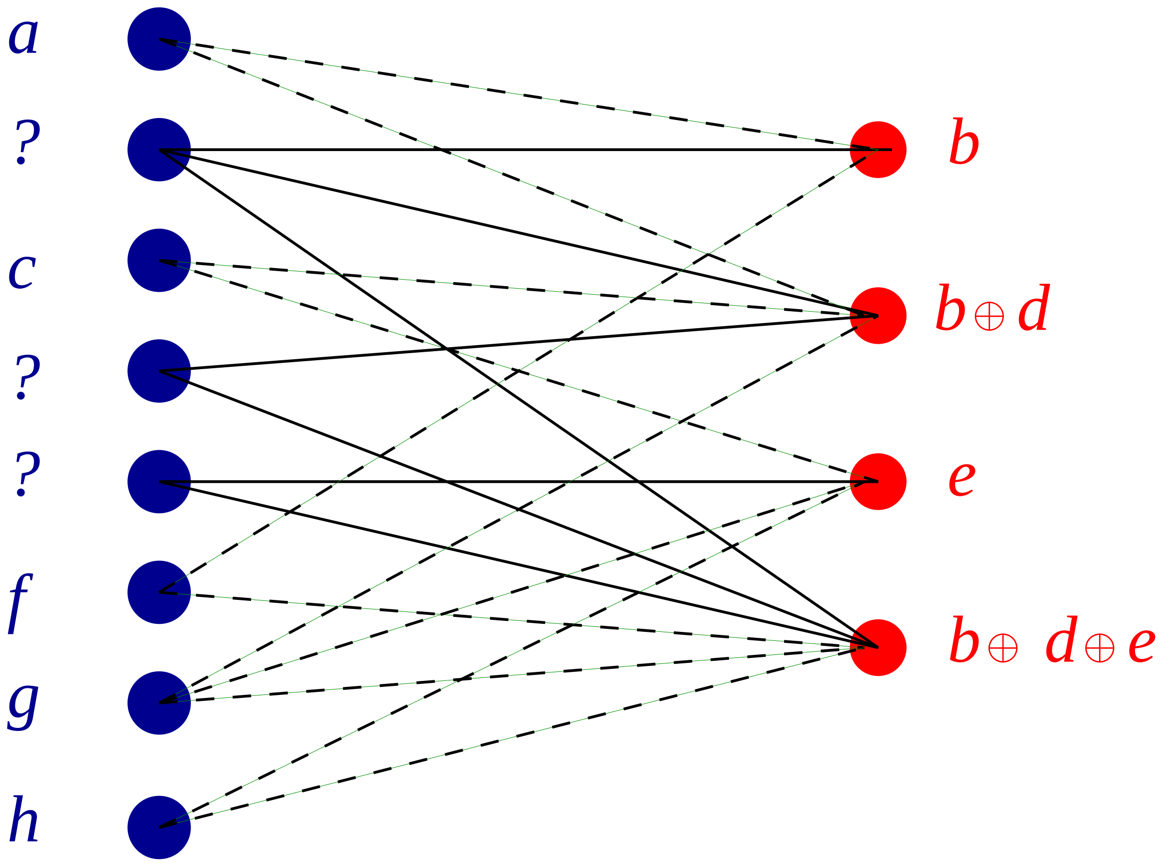
Decoding

We will adopt the model of an erasure channel.



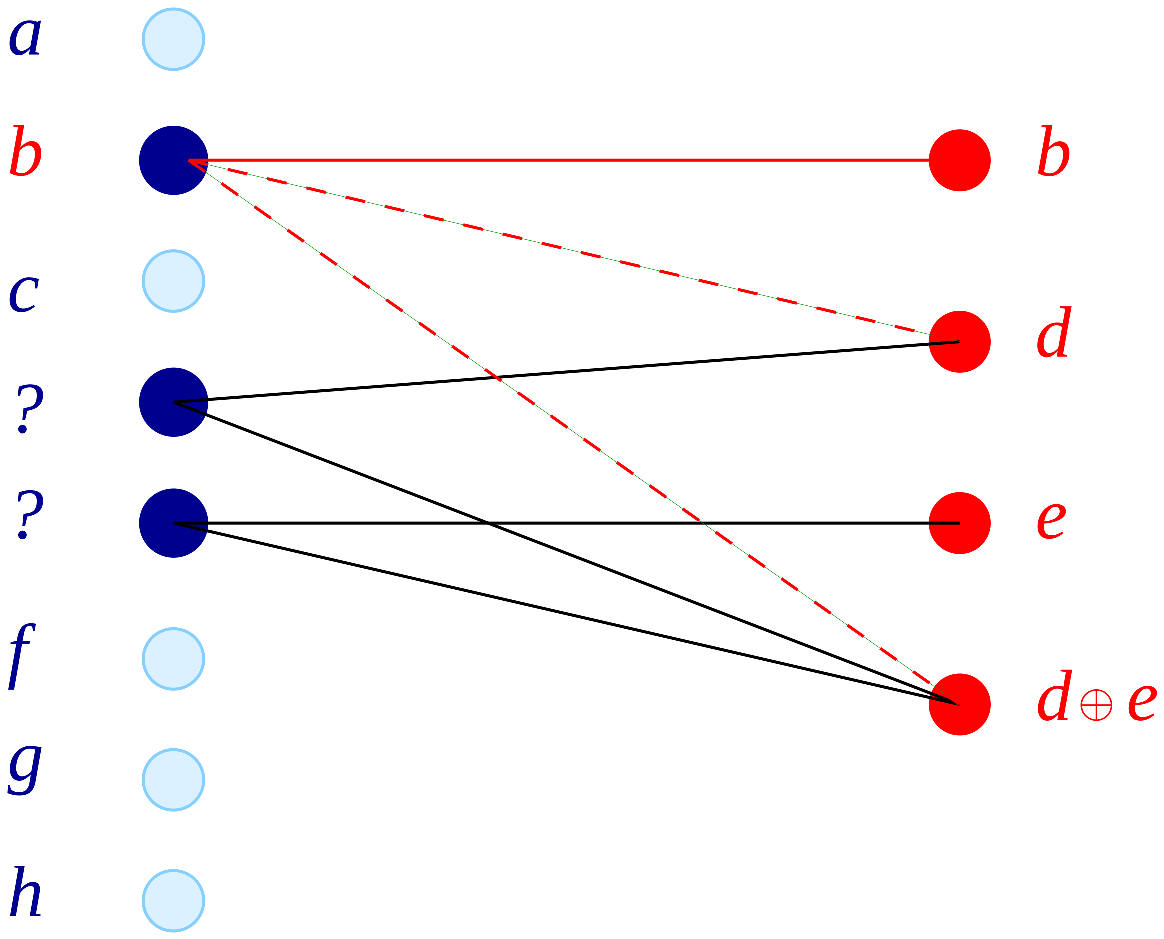
Decoding

Stage 1: Direct Recovery



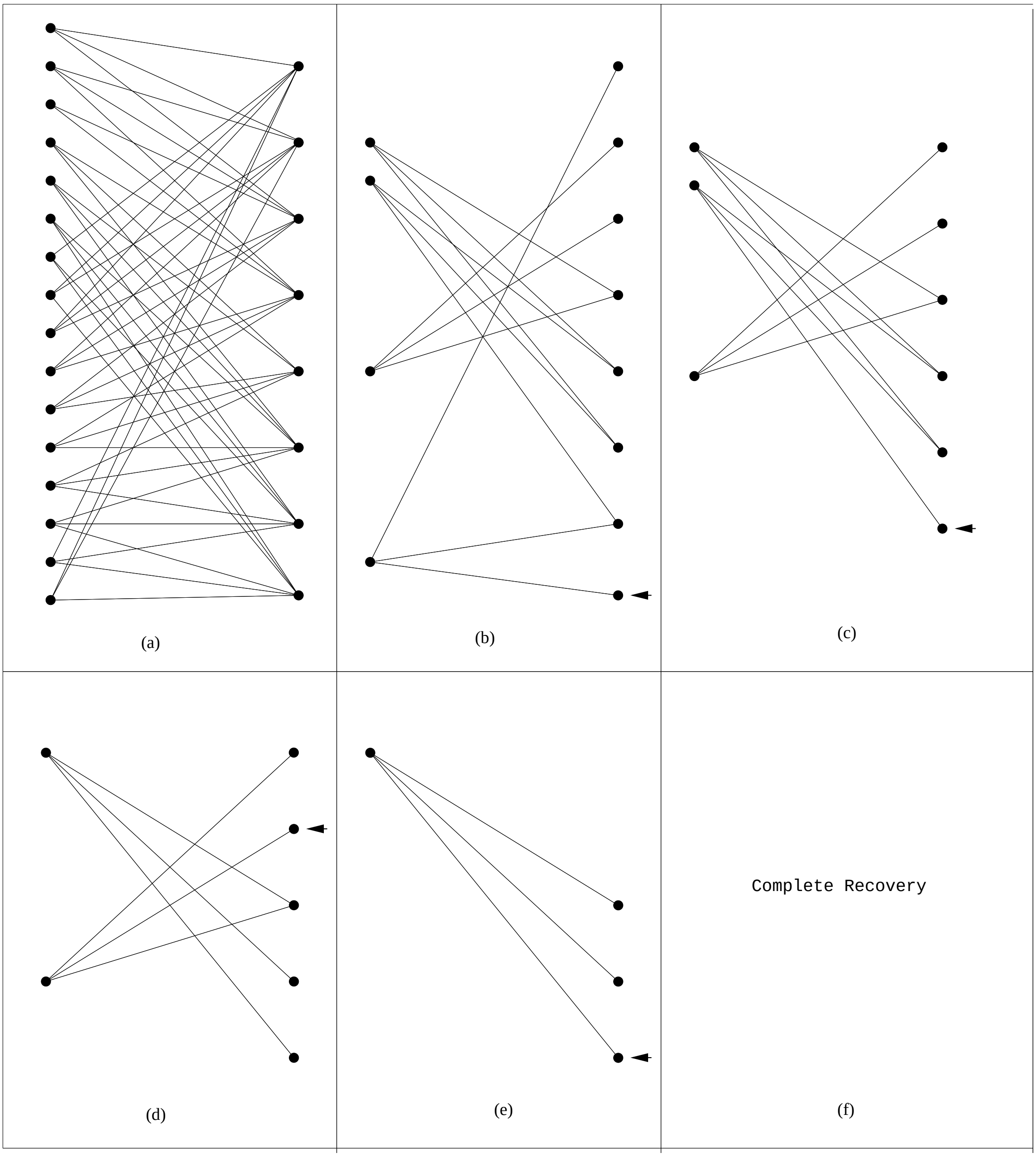
Decoding

Stage 2: Substitution Recovery



Decoding time is proportional to number of edges in graph.

Example



The (Inverse) Problem

We now have a fast encoding and decoding algorithm.

Want to design codes that perform good with respect to these algorithms.

How do we design the graphs?

Experiments

Choose regular graphs.

Experiments show that a $(3, 6)$-graph recovers from 42.9% erasures.

A $(4, 8)$-graph recovers from 38.3% erasures.

A $(5, 10)$-graph recovers from 34.1% erasures.

What are these numbers?

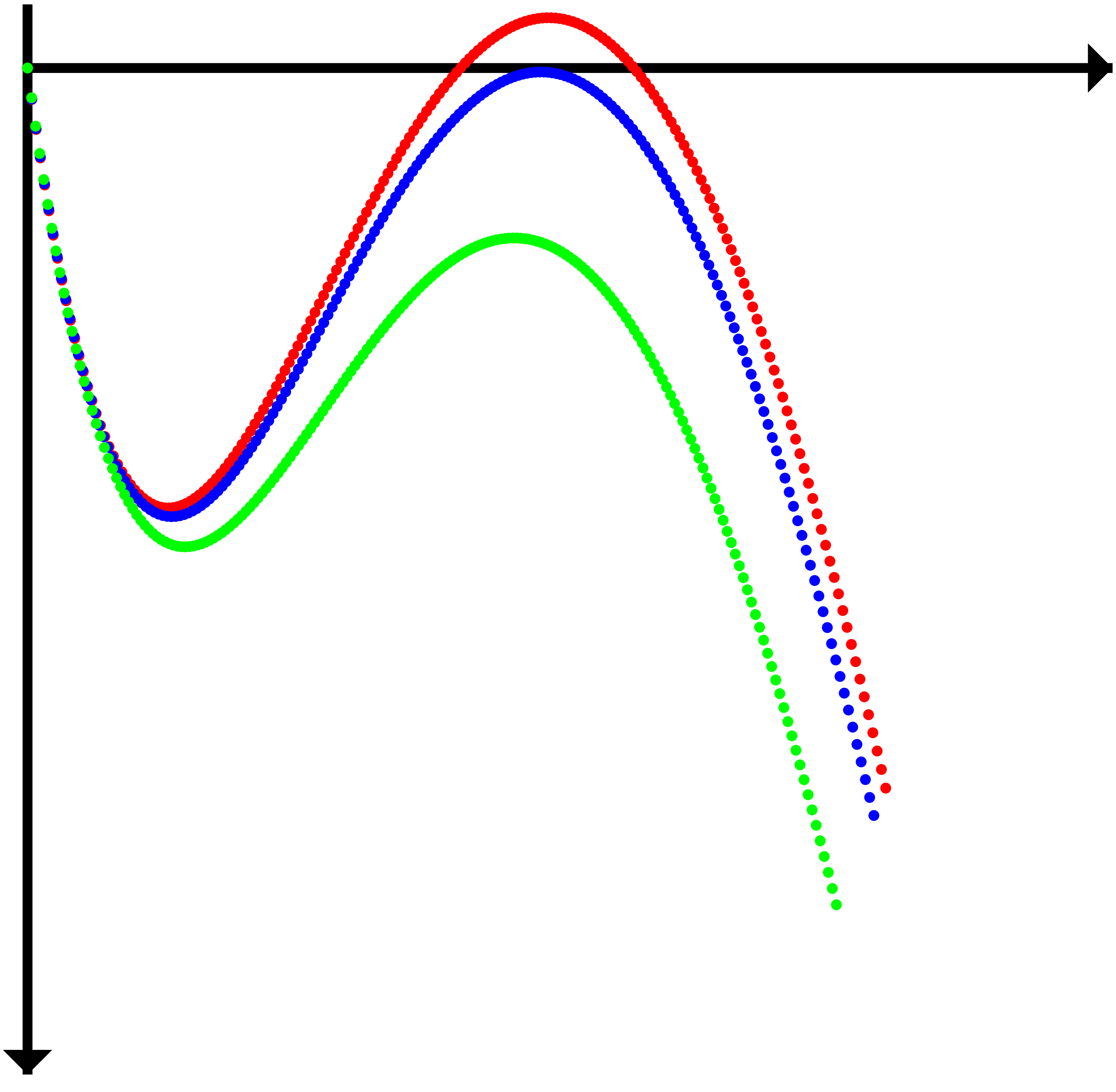
Revelation

Theorem. A random (k, d)-graph recovers from a p-fraction of erasures with high probability iff

$$p \cdot (1 - (1 - x)^{d-1})^{k-1} < x \qquad \text{for } x \in (0, p).$$

(Luby, Mitzenmacher, S, Spielman, Stemann)

Proof: uses martingales, tail inequalities, large deviation results.



General Case

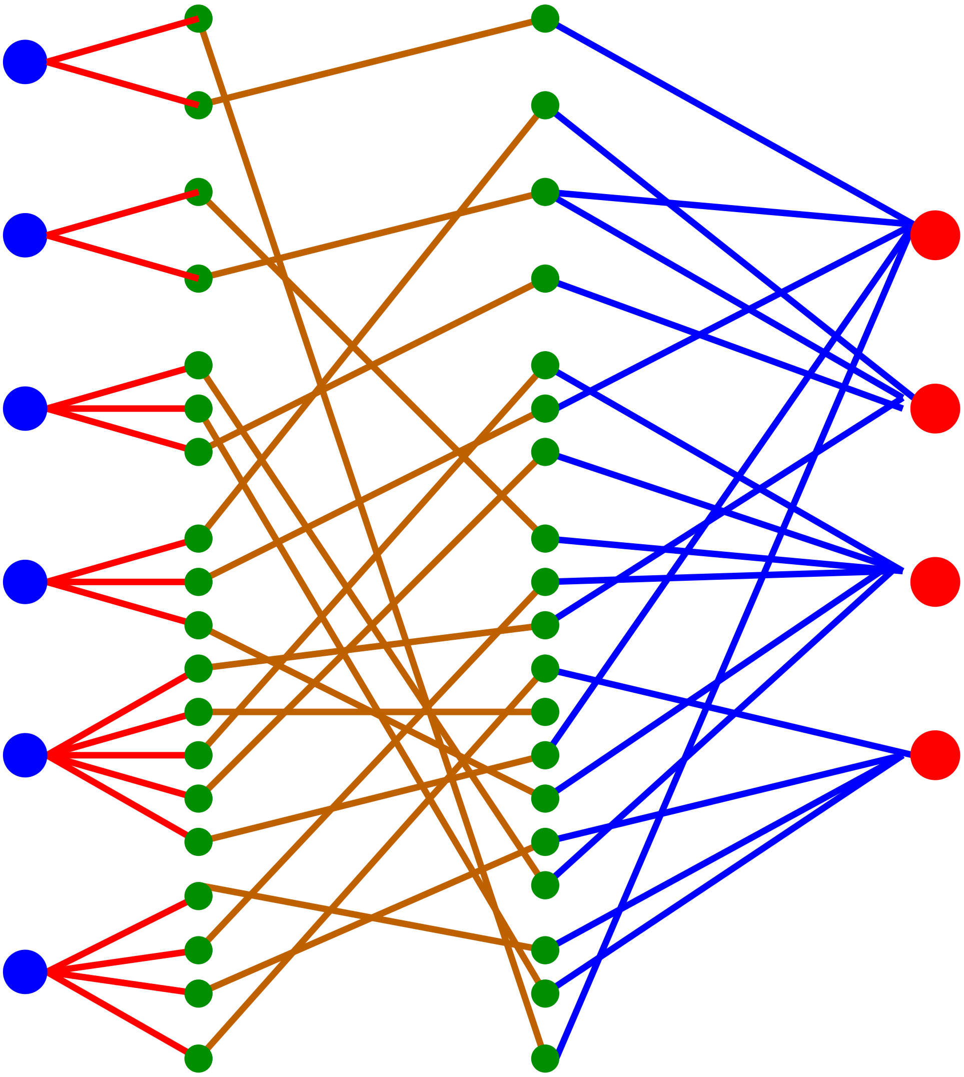
Let λ_i and ρ_i be the fraction of edges of degree i on the left and the right hand side, respectively.

Let $\lambda(x) := \sum_i \lambda_i x^{i-1}$ and $\rho(x) := \sum_i \rho_i x^{i-1}$.

Condition for successful decoding for erasure probability is then

$$p_0 \lambda\left(1 - \rho(1 - x)\right) < x$$

for all $x \in (0, p_0)$.



Design of Graphs: Linear Programming

Fix right hand side $\rho(x)$, and find best left hand side $\lambda(x)$ using the condition

$$p_0\lambda\left(1-\rho(1-x)\right) < x$$

on $(0,1)$ using linear programming.

Once best left hand side found, fix left hand side and use dual condition

$$\rho\left(1-p_0\lambda(1-x)\right) > x$$

on $(0,1)$ with linear programming to find best right hand side.

Iterate!

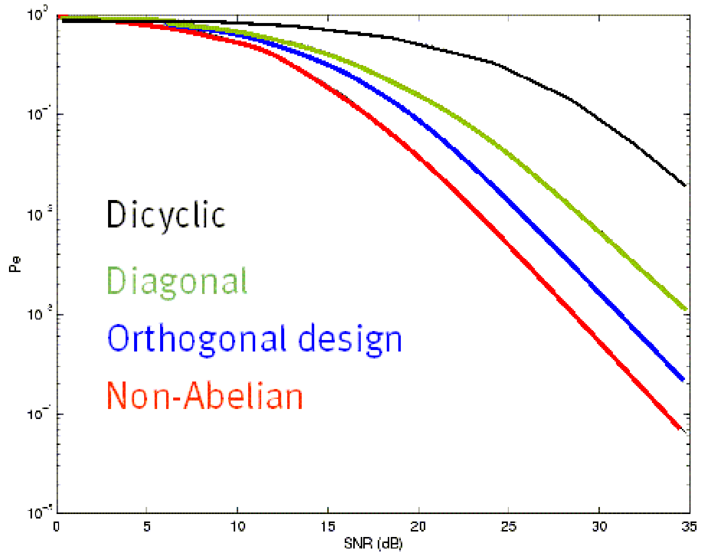
Asymptotically Optimal Codes

Using highly irregular graphs, one obtains, for any rate R sequences of codes that can get arbitrarily close to the capacity of the erasure channel.

Degree structure? Fix design parameter D.

$$\begin{aligned}\lambda(x) &:= \frac{1}{H(D)}\left(x+\frac{x^2}{2}+\cdots+\frac{x^D}{D}\right)\\ \rho(x) &:= \exp\left(\mu(x-1)\right)\end{aligned}$$

where $H(D)$ is harmonic sum $1+1/2+\cdots+1/D$ and $\mu = H(D)/\left(1-1/(D+1)\right)$.



10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
Pe
Dicyclic
Diagonal
Orthogonal design
Non-Abelian
0
5
10
15
20
25
30
35
SNR (dB)

Asymptotically Optimal Codes II

Another sequence (S):

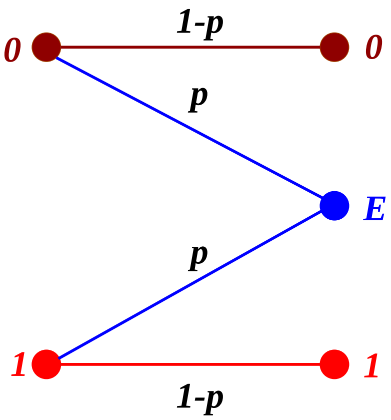
Right regular graphs, distribution of nodes on left hand side is connected with the power series expansion of

$$(1-x)^{1/m}.$$

hese are the only known examples of sequences of LDPC-codes that reach the channel capacity for any nontrivial channel.

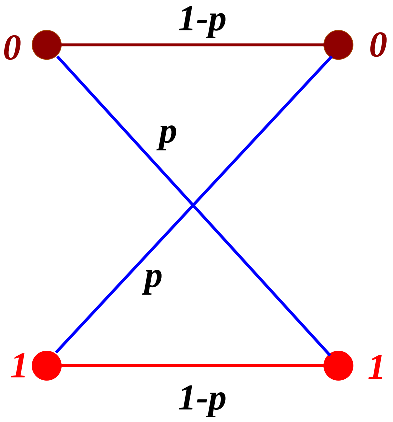
Other Channels

Erasure channel:



Capacity$=1-p$.

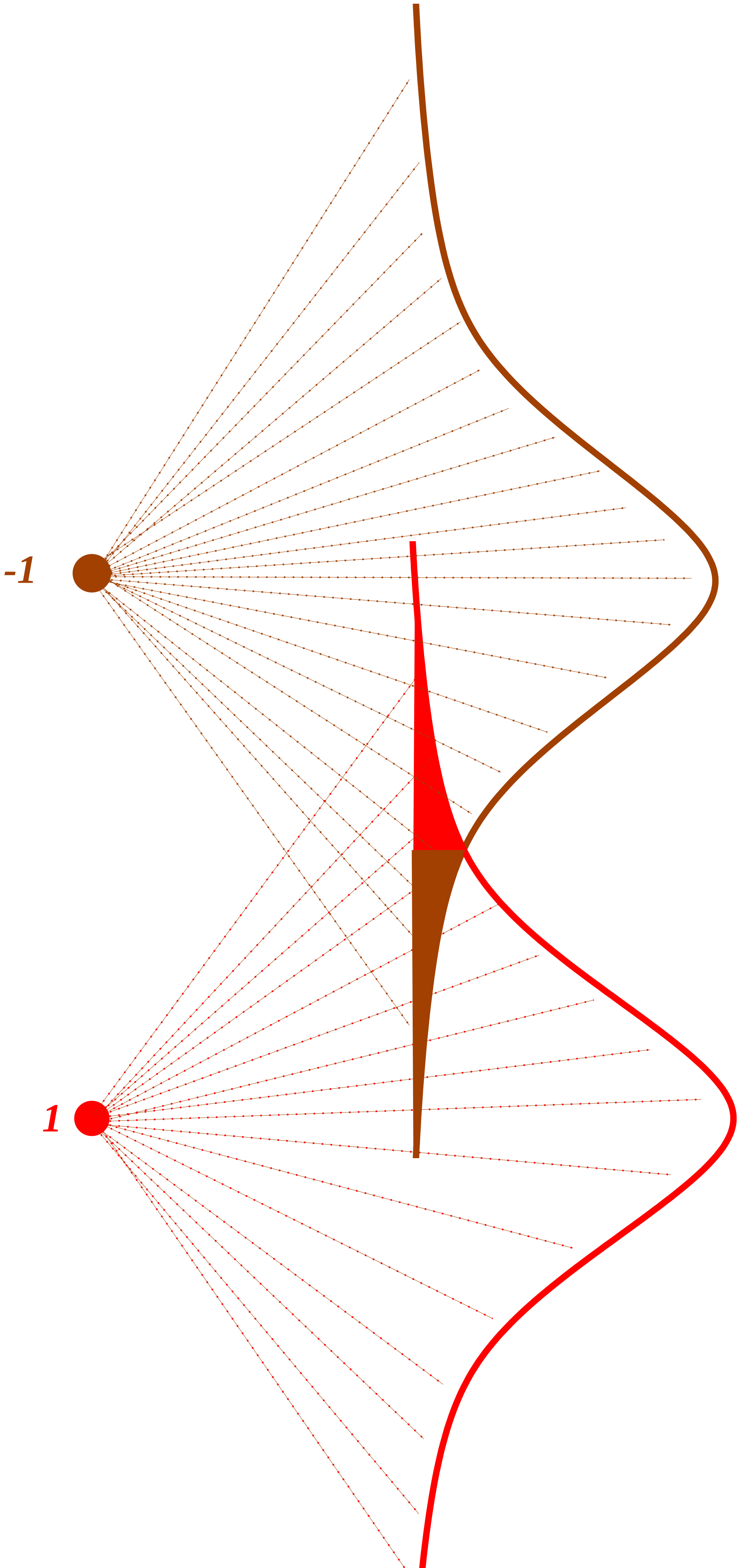
Binary symmetric channel:



Capacity$=1-p\log_2 p-(1-p)\log_2(1-p)$.

Other Channels

Additive White Gaussian Noise Channel:

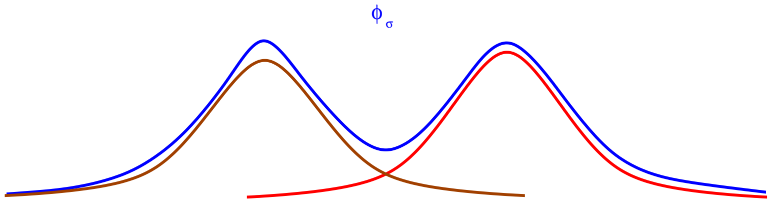

-1
1

Gaussian noise with variance σ^2.

Capacity=

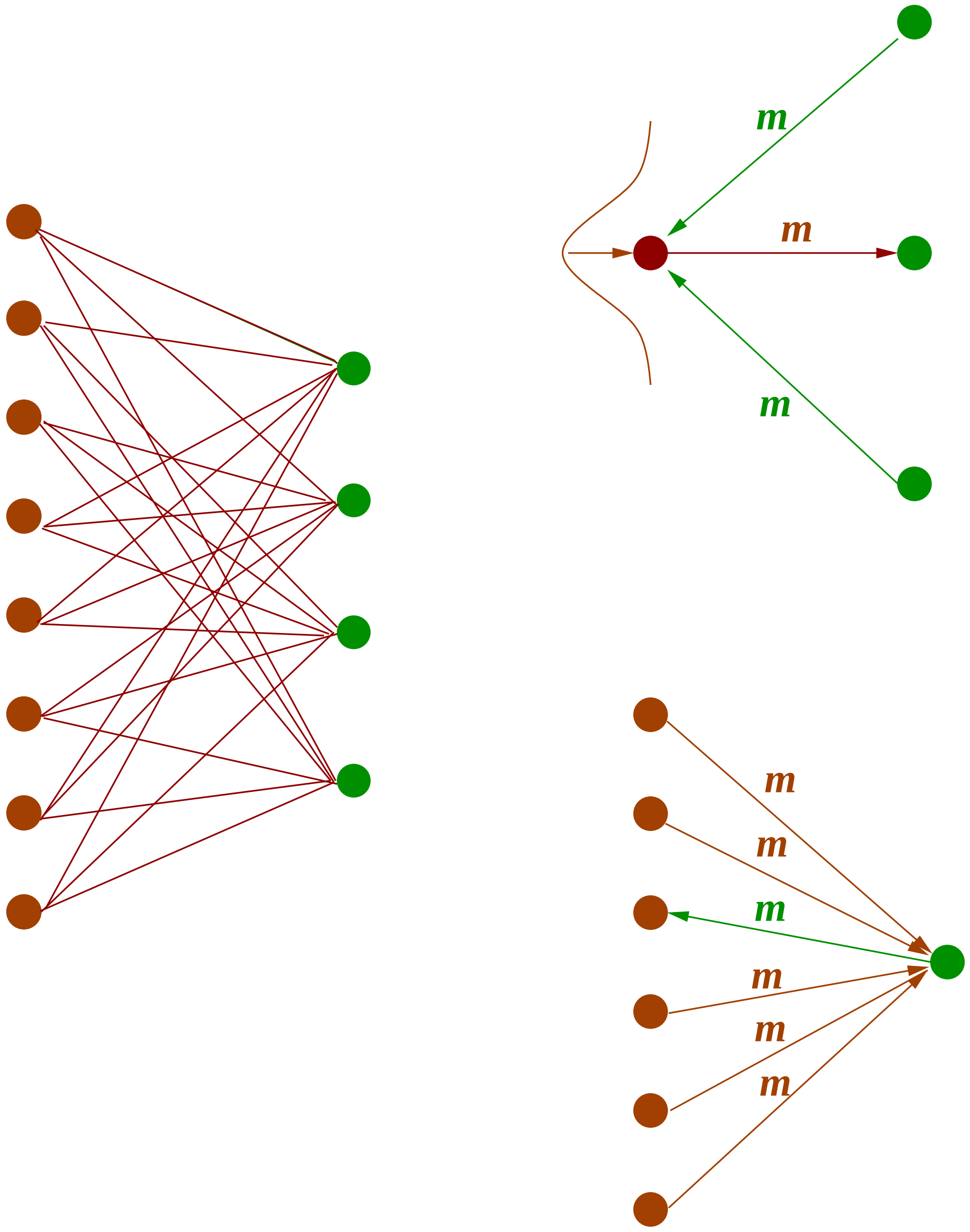
$$-\int_{-\infty}^{\infty} \phi_\sigma(x) \log_2 \phi_\sigma(x) \mathrm{d}x - \frac{1}{2}\log_2(2\pi \mathrm{e}\sigma^2),$$

where



Other Channels

Decoder? Belief propagation.



Other Channels

- Certain embodiments of the belief propagation can be rigorously analysed using the same methods as the ones used for erasure channels.

- Experiments show that some codes that are good for the erasure channel are also good for the binary symmetric or the AWGN-channel. But: can we prove it?

(Luby, Mitzenmacher, Shokrollahi, Spielman, 1998).

Analysis

Richardson and Urbanke (both Bell Labs) observed that the work of Luby, Mitzenmacher, S, Spielman can be generalized to analyse the full belief propagation algorithm (1998).

Based on this analysis, Richardson-S-Urbanke construct low-density codes that are closer to the Shannon capacity than other types of codes, such as Turbo codes.

Belief Propagation: Analysis

f_ℓ: density of the probability distribution of the messages passed from the check nodes to the message nodes at round ℓ of the algorithm.

P_0: density of the error distribution (in log-likelihood representation).

Consider (k, d) regular graph.

$$\Gamma\left(f_{\ell+1}\right) = \left(\Gamma\left(P_0 \otimes f_\ell^{\otimes(r-1)}\right)\right)^{\otimes(t-1)},$$

where Γ is a hyperbolic change of measure function,

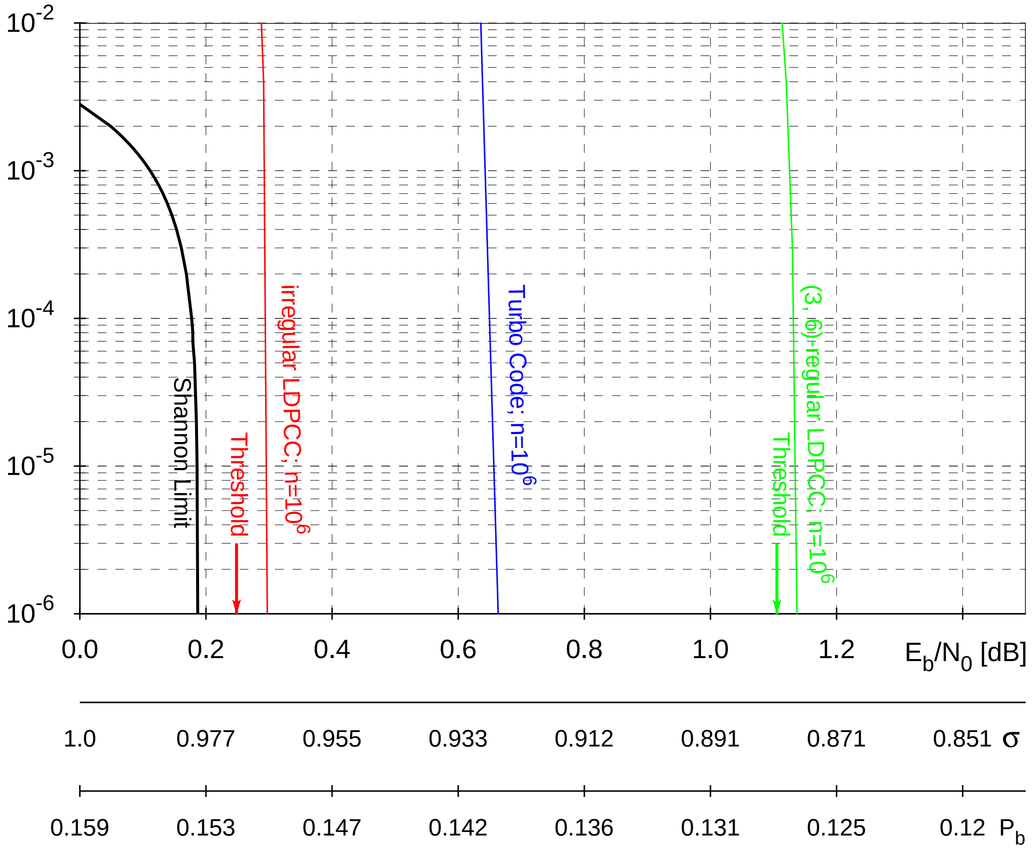
$$\Gamma(f)(y) := f(\ln \coth y/2)/\sinh(y),$$

and $\otimes$ denotes convolution.

We want f_ℓ to converge to a Delta function at ∞.

Gives rise to high-dimensional optimization algorithms.

Some Results



Race for Capacity

Explicit sequences of low-density codes which approach the Shannon-capacity when decoded with belief propagation!

Conjecture: They exist!!

Known only for the erasure channel (Luby et al., S).