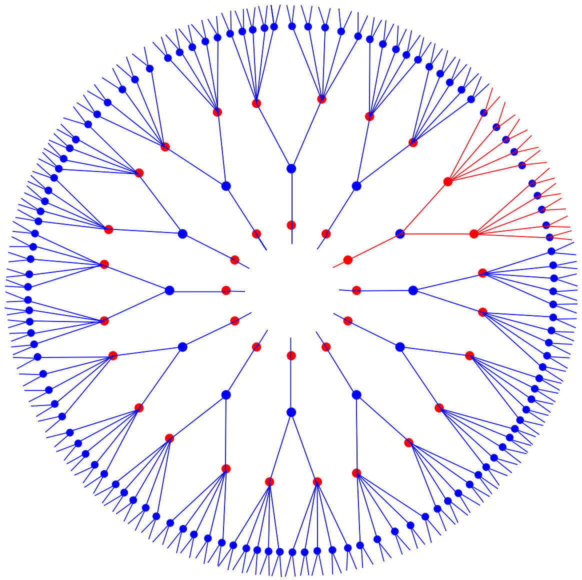
An Introduction to Low-Density Parity-Check Codes



Amin Shokrollahi

digitalfountain

Outline

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.

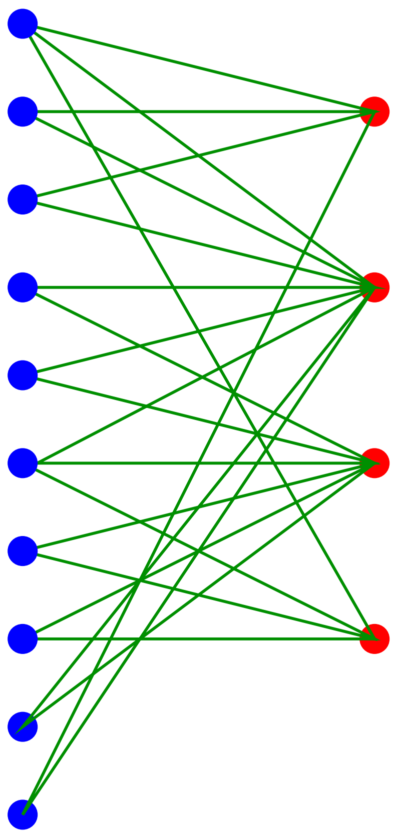
Several codes can be proved to meet the asymptotic bounds. Almost none of them are equipped with efficient encoders and decoders.

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-Urbanke	1999
Richardson-Shokrollahi-Urbanke	1999

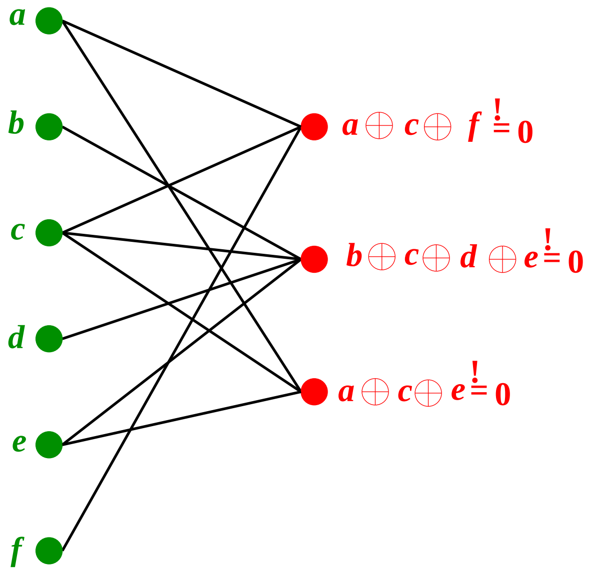
Code Construction

Codes are constructed from sparse bipartite graphs.



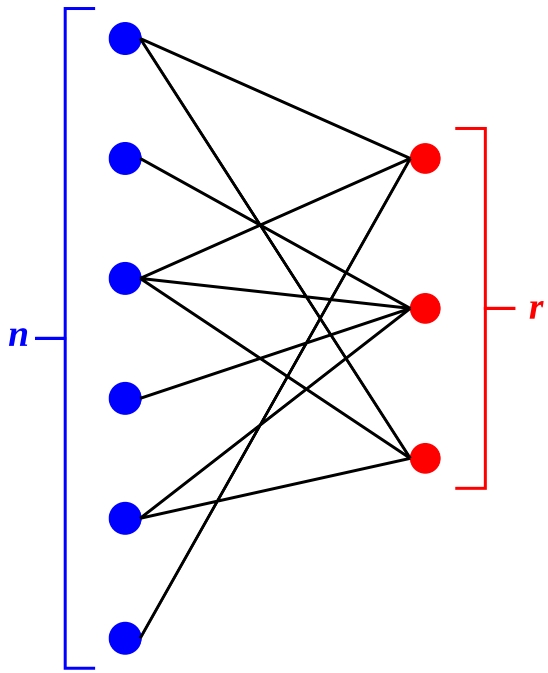
Code Construction

Any binary linear code has a graphical representation.



Not any code can be represented by a sparse graph.

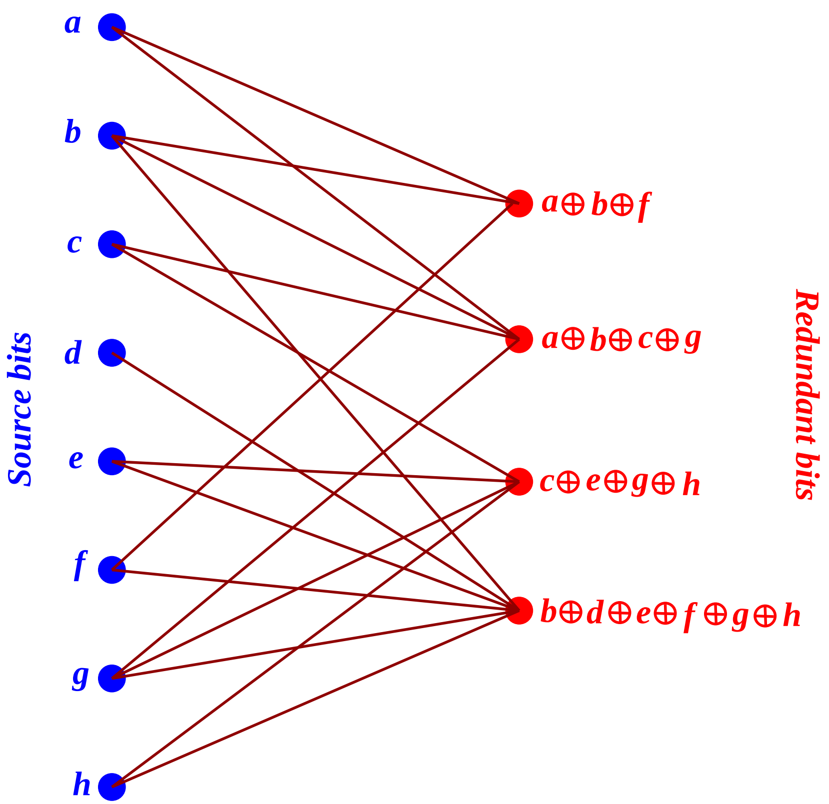
Parameters



$$\textit{Rate} \geq \frac{n-r}{n}$$

$$\textit{Rate} \geq 1 - \frac{\textit{average left degree}}{\textit{average right degree}}$$

Dual Construction



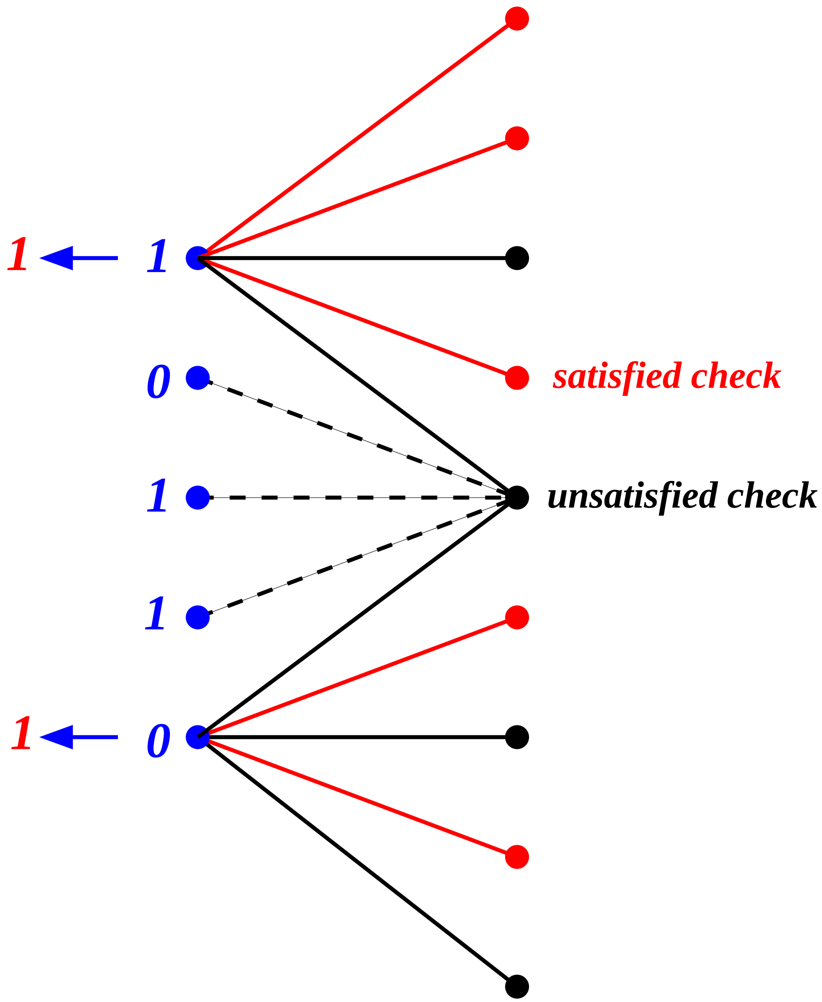
Encoding time is proportional to number of edges.

Algorithmic Issues

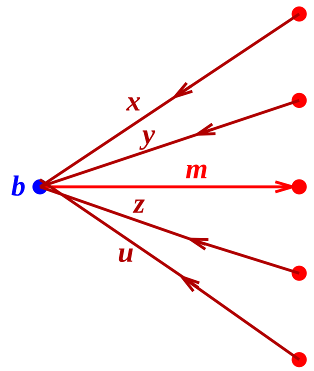
- Encoding?
  - Is linear time for the dual construction
  - Is quadratic time (after preprocessing) for the Gallager construction. More later!

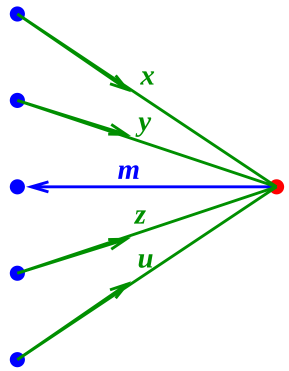
- Decoding?
  - Depends on the channel,
  - Depends on the fraction of errors.

Decoding on a BSC: Flipping



Decoding on a BSC: Gallager Algorithm A (Message passing)



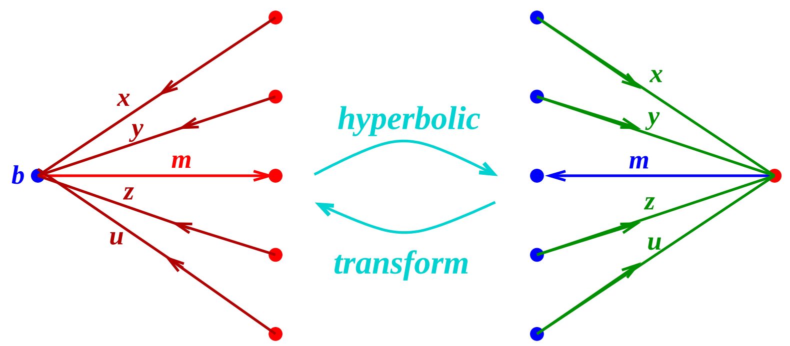


$$m = \begin{cases} x & \text{if } x=y=z=u \\ b & \text{else} \end{cases}$$

$$m = x \oplus y \oplus z \oplus u$$

MESSAGE **CHECK**

Decoding on a BSC: Belief Propagation



$m = x + y + z + u + b$

$m = x * y * z * u$

$(a,b) * (c,d) := (a+c,\ b+d \bmod 2)$

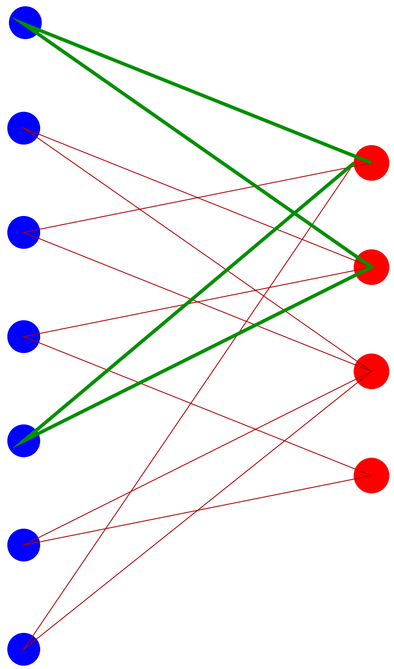
MESSAGE

CHECK

Messages in log-likelihood ratios.

Optimality of Belief Propagation

Belief propagation is bit-optimal if graph has no loops.



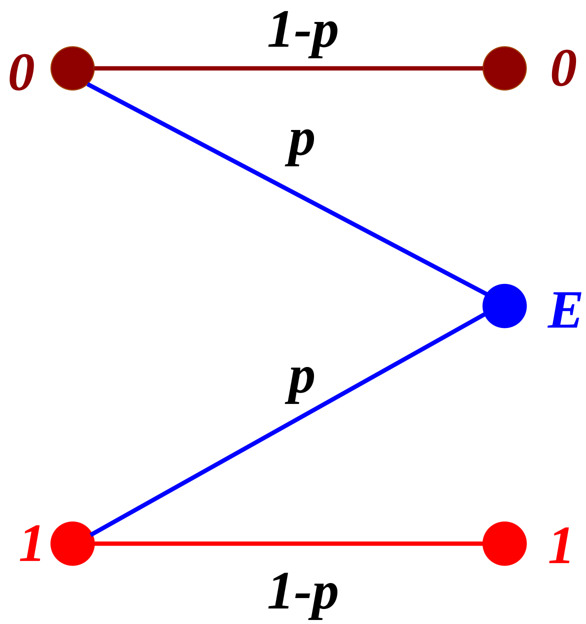
Maximizes the probability

$$P(c_m = b \mid y) = \sum_{c \in C} P(c \mid y).$$

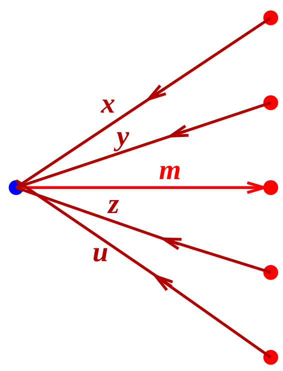
Performance on a (3,6)-graph

Shannon limit:	11%	
Flipping algorithm:	1%?	
Gallager A:	4%	
Gallager B:	4%	(6.27%)
Erasure decoder:	7%	
Belief propagation:	8.7%	(10.8%)

The Binary Erasure Channel (BEC)

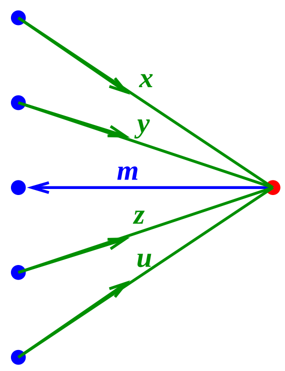


Decoding on a BEC:
Luby-Mitzenmacher-Shokrollahi-Spielman-Stemann



$$m = \begin{cases} 1 & \text{if} \quad x \vee y \vee z \vee u = 1 \\ 0 & \text{else} \end{cases}$$

MESSAGE

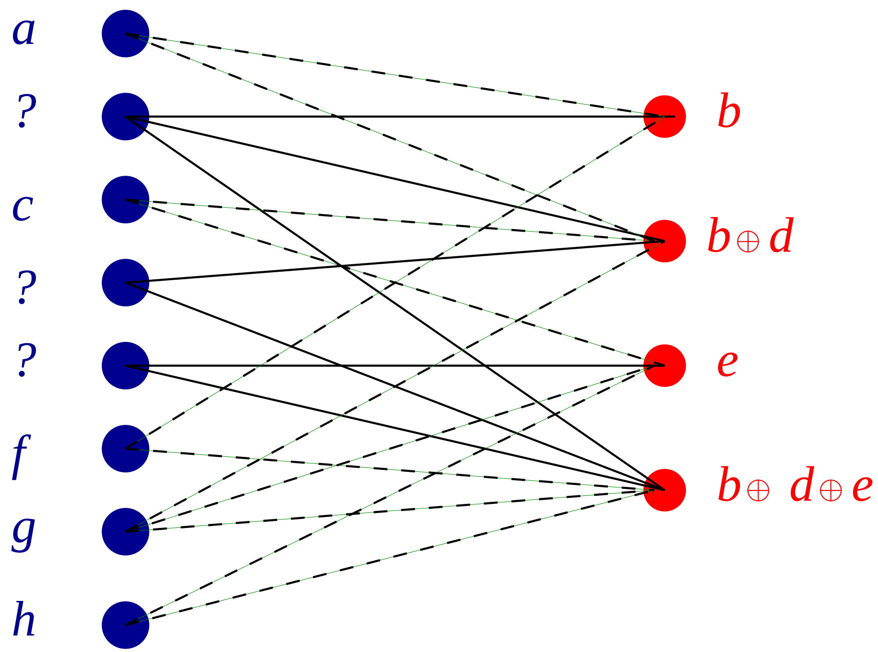


$$m = \begin{cases} 1 & \text{if} \quad x = y = z = u = 1 \\ 0 & \text{else} \end{cases}$$

CHECK

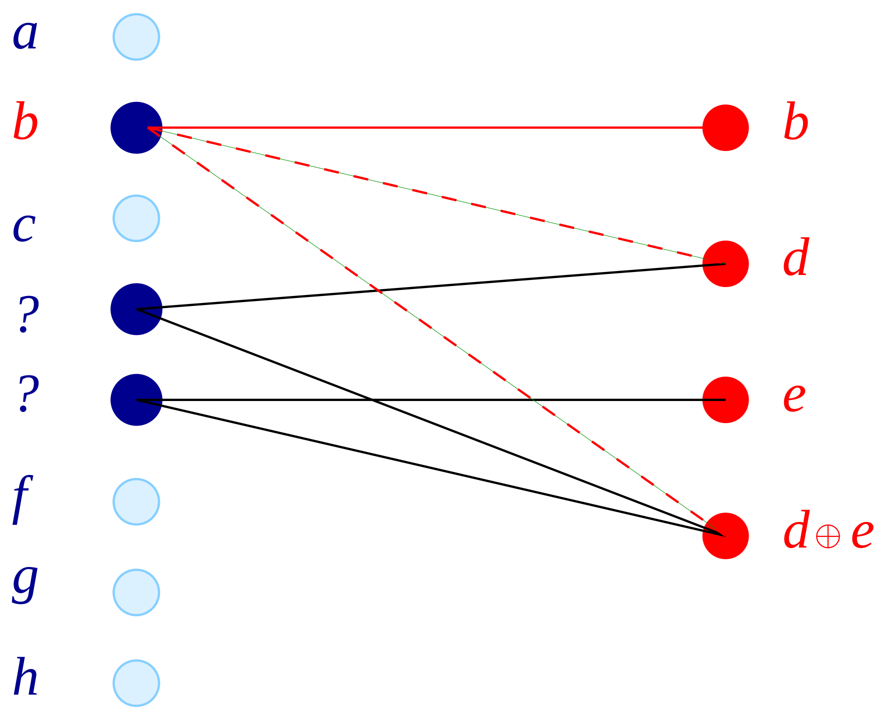
Decoding on a BEC

Phase 1: Direct recovery

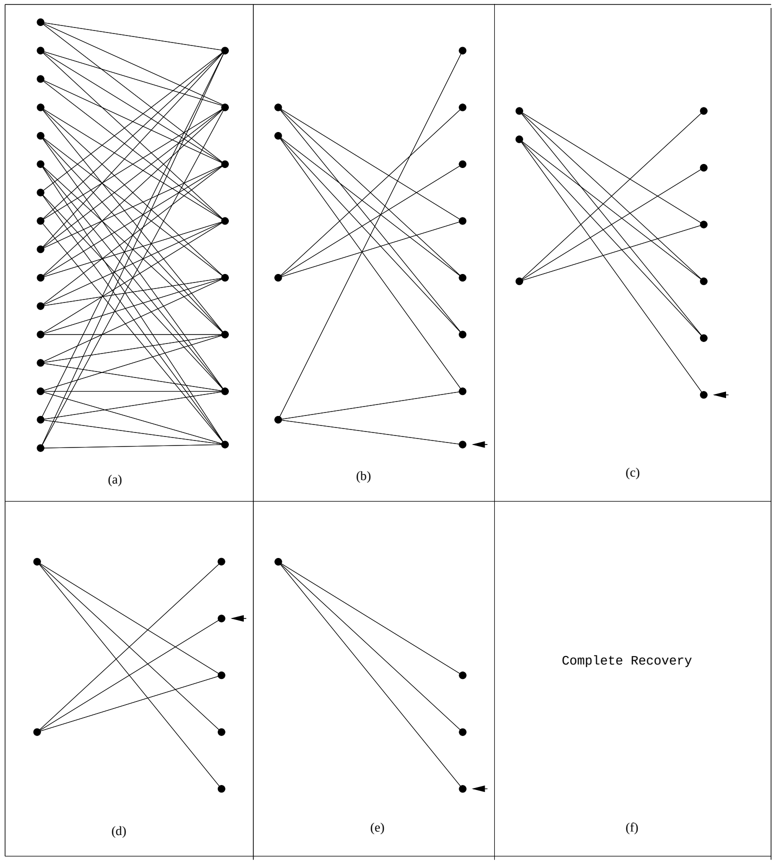


Decoding on a BEC

Phase 2: Substitution



Example



The (inverse) problem

Have: fast decoding algorithms.

Want: design codes that can correct many errors using these algorithms.

Focus on the BEC in the following.

Experiments

Choose regular graphs.

An (d, k)-regular graph has rate at least $1 - d/k$. Can correct at most an d/k-fraction of erasures.

Choose a random (d, k)-graph.

p_0 := maximum fraction of erasures the algorithm can correct.

d	k	d/k	p_0
3	6	0.5	0.429
4	8	0.5	0.383
5	10	0.5	0.341
3	9	0.33	0.282
4	12	0.33	0.2572

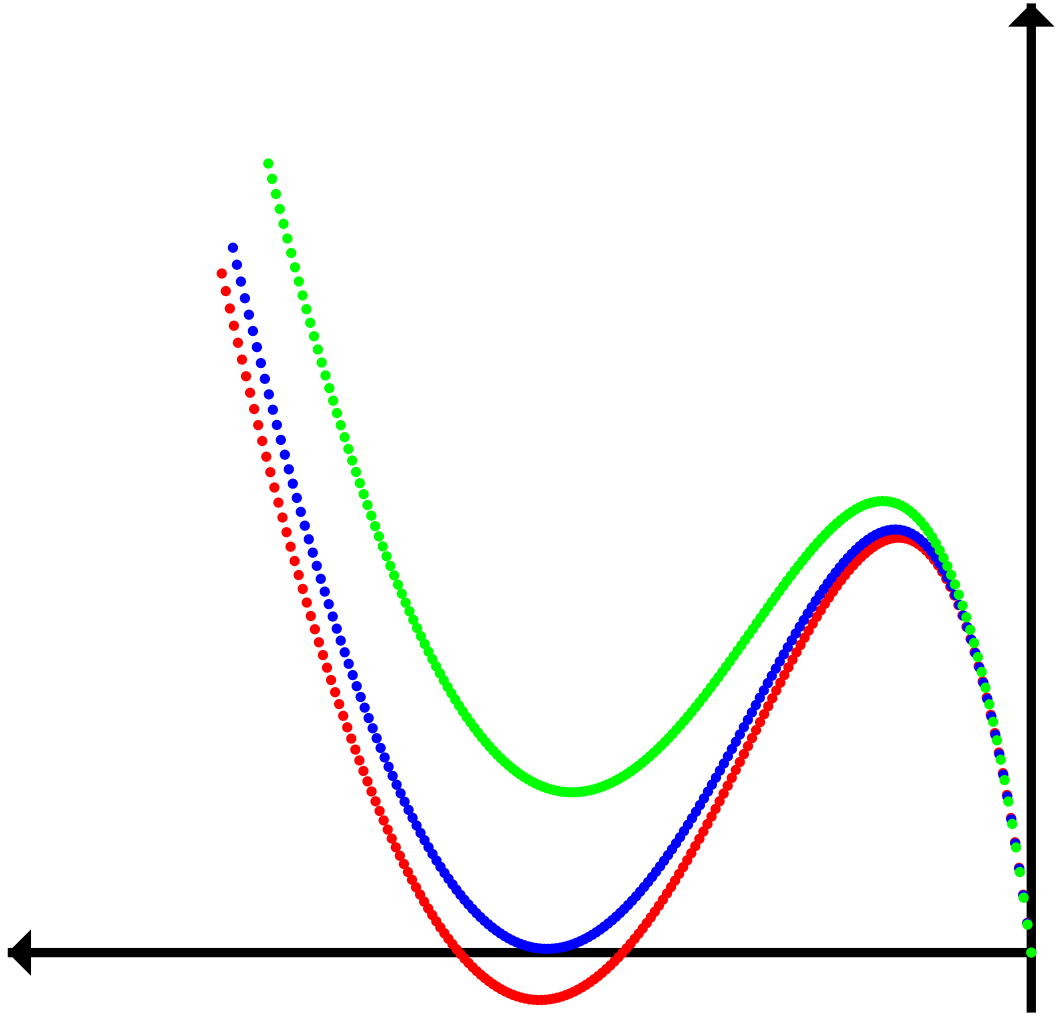
What are these numbers?

A Theorem

Luby, Mitzenmacher, Shokrollahi, Spielman, Stemann, 1997:

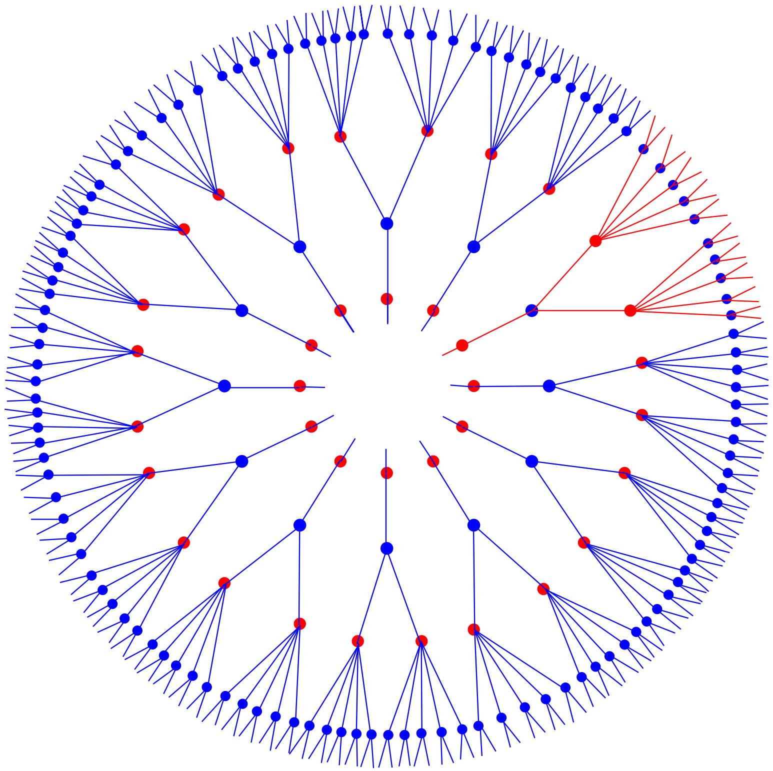
A randomly chosen (d, k)-graph can correct a p_0-fraction of erasures with high probability if and only if

$$p_0 \cdot (1 - (1 - x)^{k-1})^{d-1} < x \qquad \text{for } x \in (0, p_0).$$



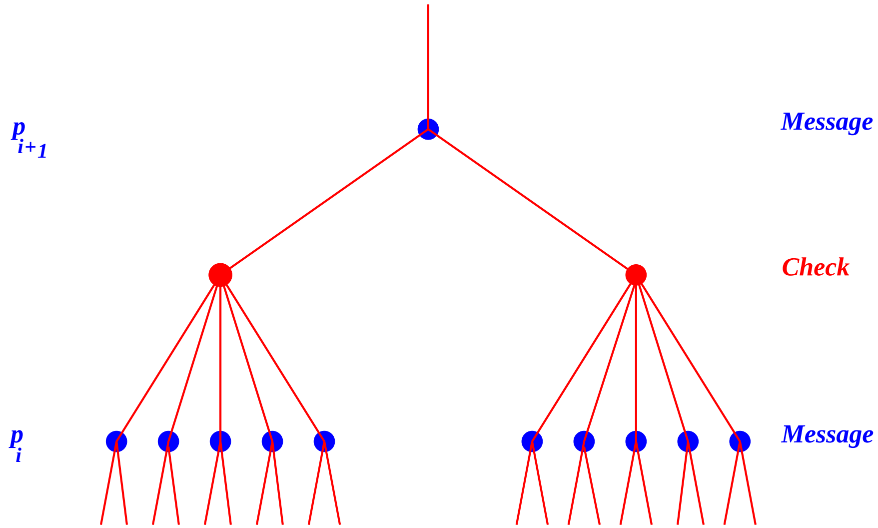
Analysis: $(3, 6)$-graphs

Expand neighborhoods of message nodes.



Analysis: $(3, 6)$-graphs

p_i probability that message node is still erased after ith iteration.



$$p_{i+1} = p_0(1 - (1 - p_i)^5)^2.$$

Successful Decoding

Condition:

$$p_0(1-(1-p_i)^5)^2 < p_i$$

Analysis: $(3, 6)$-graphs

Making arguments exact:

- Neighborhood is tree-like: high probability, standard argument.

- Above argument works for expected fraction of erasures at ℓth round.

  Real value is sharply concentrated around expected value p_ℓ: Edge exposure martingale, Azuma's inequality.

The 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 p_0 is then

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

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

Belief propagation

Richardson-Urbanke, 1999:

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 (d, k) regular graph.

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

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.

Achieving capacity

Want to design codes that can recover from a fraction of $1-R$ of erasures (asymptotically).

Want to have λ and ρ so that

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

for all $x \in (0, p_0)$, and p_0 arbitrarily close to

$$1 - R = \frac{\int_0^1 \rho(x) \mathrm{d}x}{\int_0^1 \lambda(x) \mathrm{d}x}.$$

Tornado codes

Extremely irregular graphs provide for any rate R sequences of codes which come arbitrarily close to the capacity of the erasure channel!

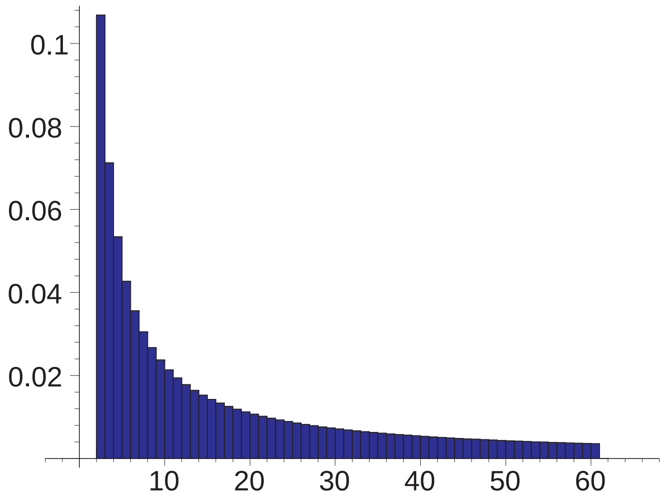
Degree structure?

Choose design parameter D.

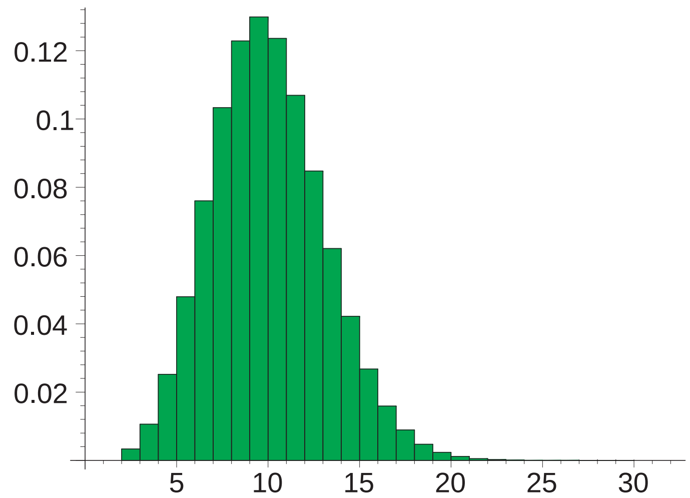
$$
\begin{aligned}
\lambda(x) &:= \frac{1}{H(D)}\left(x+\frac{x^2}{2}+\cdots+\frac{x^D}{D}\right) \\
\rho(x) &:= \exp\left(\mu(x-1)\right),
\end{aligned}
$$

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

Tornado Codes: Degree Distribution



Heavy tail



Poisson

Right regular codes

Shokrollahi, 1999:

Graphs that are regular on the right.

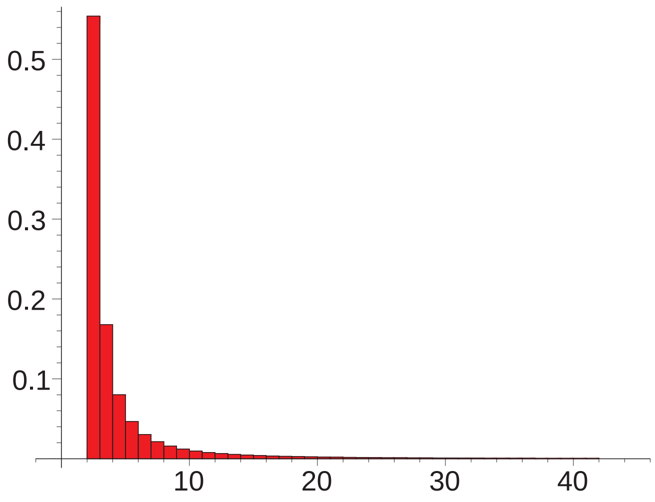
Degrees on the left are related to the Taylor expansion of

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

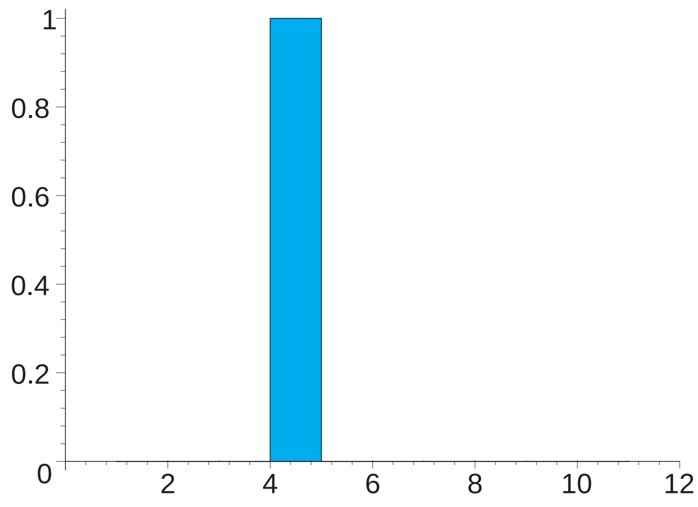
Methodology for constructing capacity-achieving sequences by Oswald-Shokrollahi, 2000.

Also show that the right regular sequence is the best in a certain sense.

Right Regular Codes: Degree Distribution



Left



Right

Other channels?

f density function.

$\lambda(f) := \sum_i \lambda_i f^{\otimes(i-1)}$.

$\rho(f) := \sum_i \rho_i f^{\otimes(i-1)}$.

$$\Gamma\left(f_{\ell+1}\right) = \rho\left(\Gamma\left(P_0 \otimes \lambda(f_\ell)\right)\right).$$

Want P_0 such that $f_\ell \to \Delta_\infty$.

Conditions on the density functions

Richardson-Shokrollahi-Urbanke, 1999:

- Consistency: if the channel is *"symmetric"*, then the density functions f_ℓ satisfy $f(x) = f(-x)\mathrm{e}^x$.

- Fixed point theorem: If $P_{\mathrm{err}}(f_i) = P_{\mathrm{err}}(f_j)$ for $i < j$, then $f_i = f_j$ is a fixed point of the iteration.

Conditions on the density functions

- Stability: let

$$r := -\lim_{n\to\infty} \frac{1}{n} \log P_{\text{err}}(P_0^{\otimes n}).$$

Then for $\lambda_2 \rho'(1) > e^r$ we have $P_{\text{err}}(f_\ell) > \epsilon$ for some fixed ϵ and all ℓ.

If $\lambda_2 \rho'(1) < e^r$, then the fixed point Δ_∞ is stable.

$$P_{\text{err}}(f) := \int_{-\infty}^{0} f(x)\mathrm{d}x$$

is the error probability.

Stability

- Erasure channel with erasure probability p_0:

$$\lambda_2 \rho'(1) \leq \frac{1}{p_0}.$$

- BSC channel: with probability p:

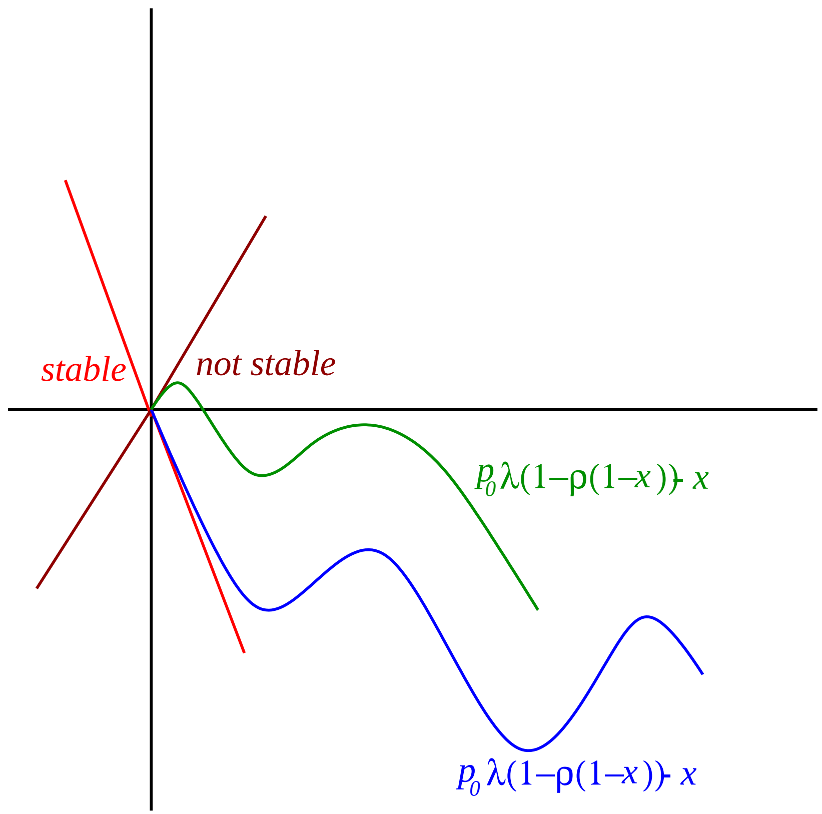
$$\lambda_2 \rho'(1) \leq \frac{1}{2\sqrt{p(1-p)}}.$$

- AWGN channel: with variance σ^2:

$$\lambda_2 \rho'(1) \leq e^{-\frac{1}{2\sigma^2}}.$$

Stability for the Erasure Channel

Shokrollahi, 1999:



Flatness: Higher Stability Conditions

Shokrollahi, 2000:

$(\lambda_m(x), \rho_m(x))$ capacity achieving sequence of degree distributions.

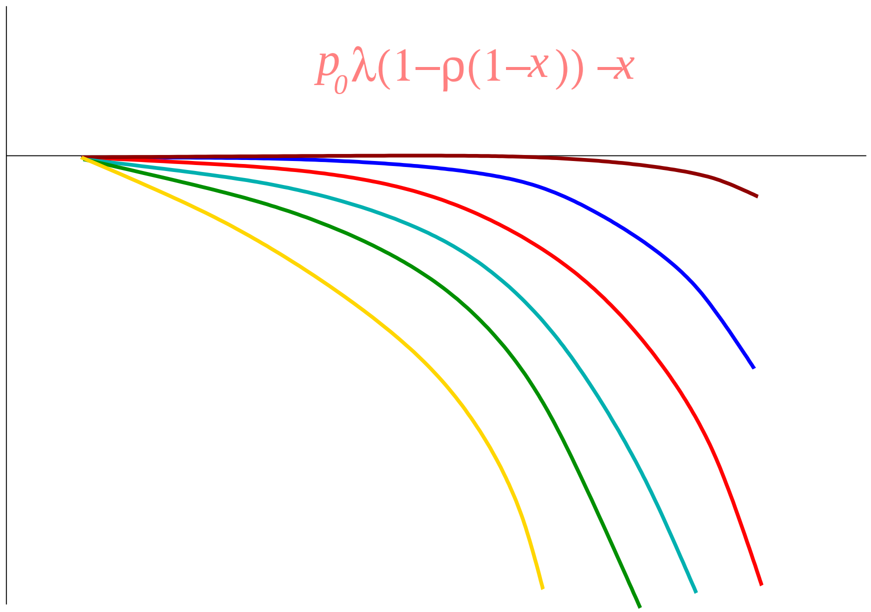
Then:

$$(1-R)\lambda_m(1-\rho_m(1-x)) - x$$

converges uniformly to the zero-function on the interval $[0, 1-R]$.

No equivalent known for other channels.

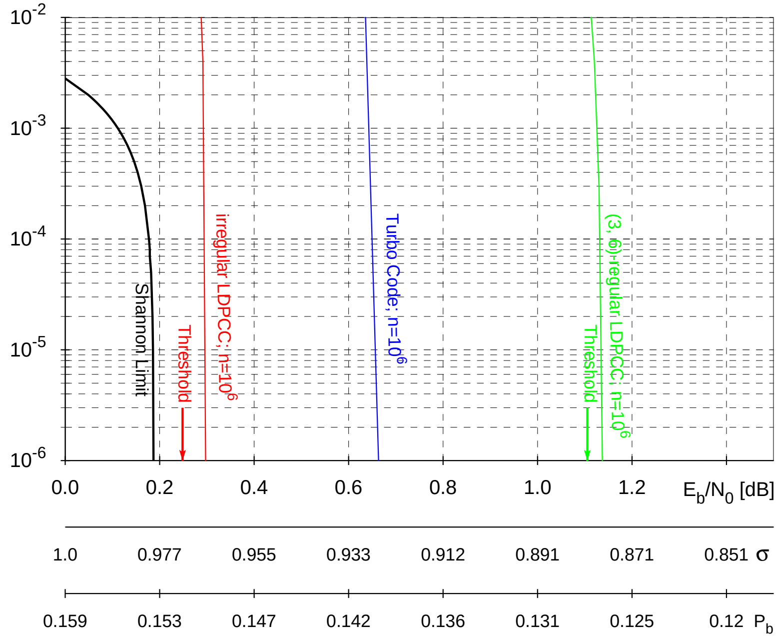
Flatness: Higher Stability Conditions



Capacity achieving

No sequences of c.a. degree distributions for channels other than the erasure channel known.

Conjecture: They exist!



Open problems

Asymptotic theory

1. Classification of capacity achieving sequences for the erasure channel.
2. Capacity achieving sequences for other channels.
3. Exponentially small error probabilities for the decoder (instead of polynomially small).

Explicit constructions

1. Constructions using finite geometries.
2. Construction using Reed-Solomon-Codes.
3. Algebraic constructions.

Short codes

Graphs with loops.

Algorithmic issues

1. Design and analysis of new decoding algorithms.
2. Design of new encoders.

Randomness

Use of randomness in other areas: random convolutional codes?.