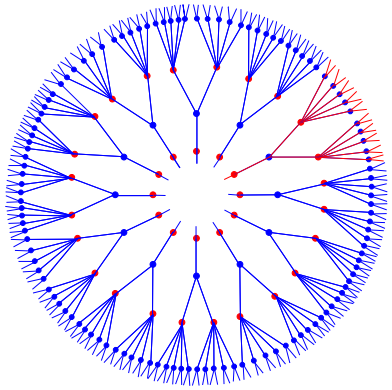
LDPC Codes for Erasure Correction



Amin Shokrollahi

digitalfountain

Content

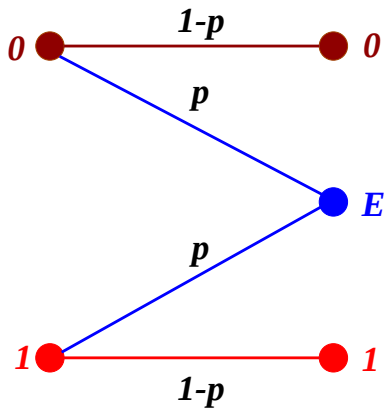
- Codes
- Erasure Correction
- LDPC codes
- Examples
- Tornado codes

Codes

A binary linear code of dimension k and block length n is a k-dimensional subspace of $\mathbb{F}_2^n$.

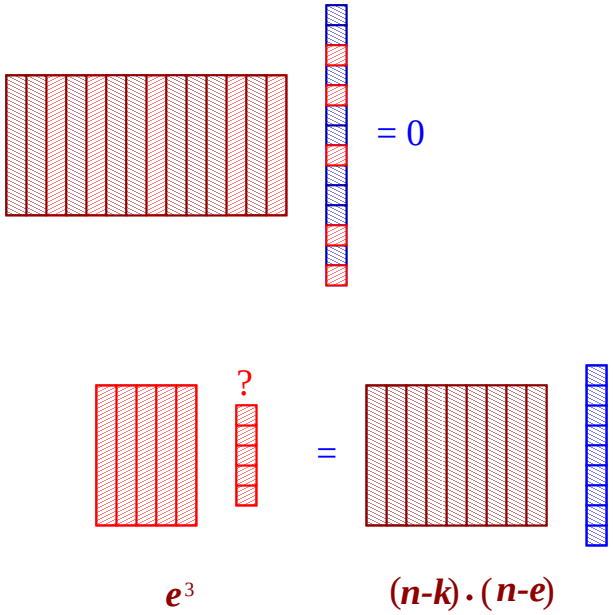
The ratio k/n is called the rate of the code.

Want to use codes to transmit information reliably over the binary erasure channel with erasure probability p:



Correcting erasures

Erasure correction can be done by solving system of equations using the check matrix:



Erasure correction is possible iff system of equations solvable. This decoder is called maximum likelihood decoder.

Capacity

Capacity of binary erasure channel with probability p is $1 - p$.

Shannon proves existence of codes whose rate approaches $1 - p$, and which can correct erasures over the channel with high probability.

Such codes are called capacity achieving.

How can we find those codes?

Can we obtain faster algorithms?

Computational Efficiency

Let a rate R be given.

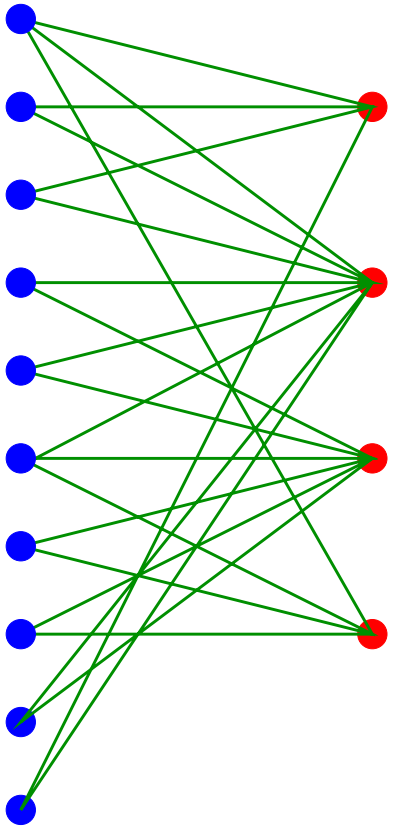
It can be shown that random codes of rate $1 - p - \epsilon$ can decode over the binary erasure channel of probability p with high probability.

But: encoding complexity is $O(n^2)$, and decoding complexity is $O(n^3)$. (Using maximum likelihood decoding.)

Want faster algorithms, ideally $O(n)$.

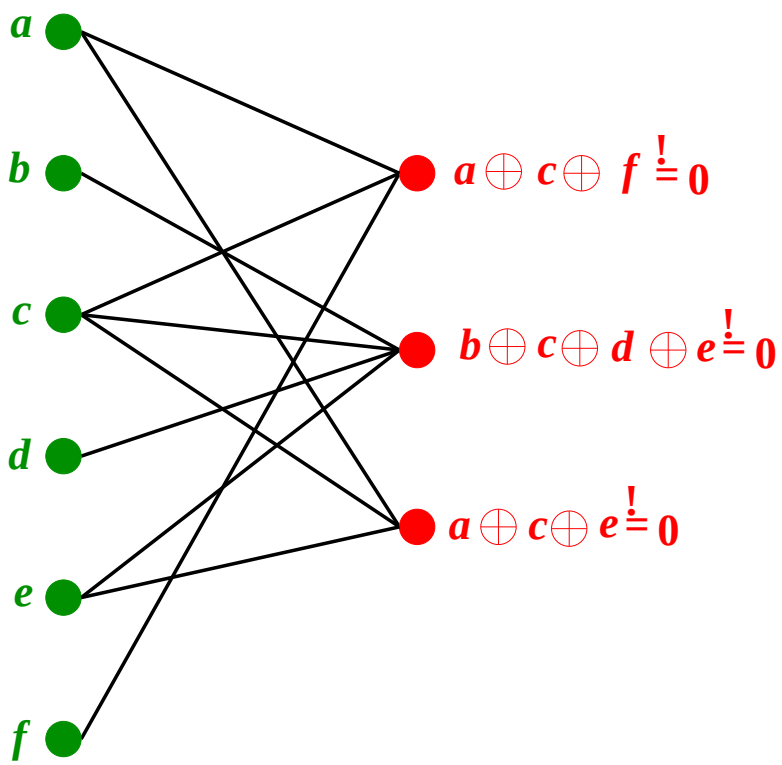
LDPC Codes

Constructed from sparse bipartite graphs.



Left nodes are called message nodes, right nodes are called check nodes.

Construction



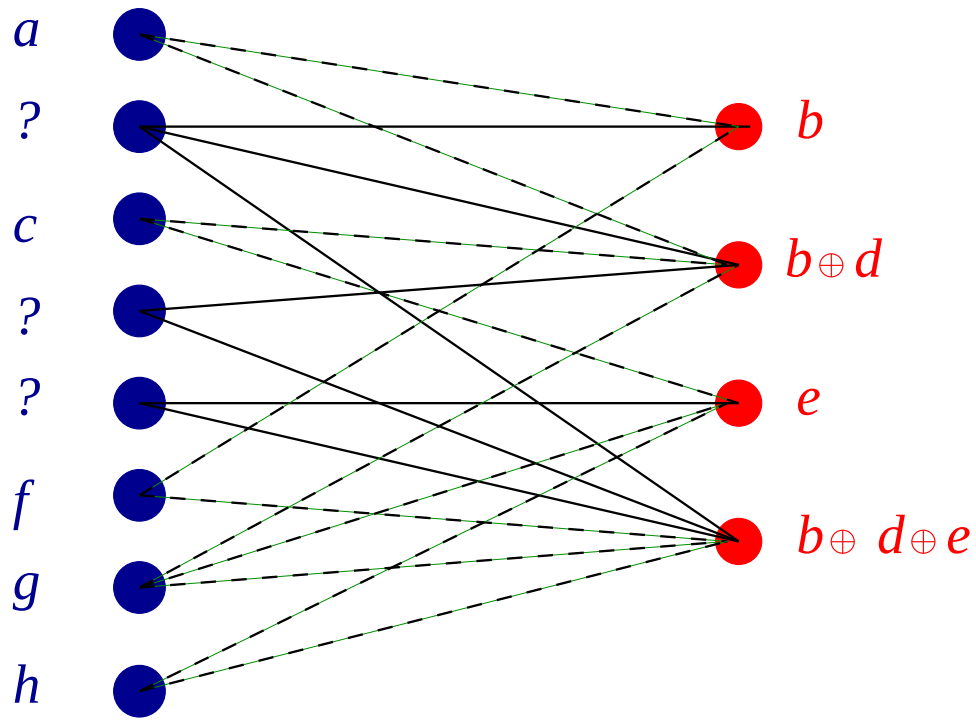
Every binary linear code has such a representation, but not every code can be represented by a sparse graph.

Encoding/decoding times?

Decoding

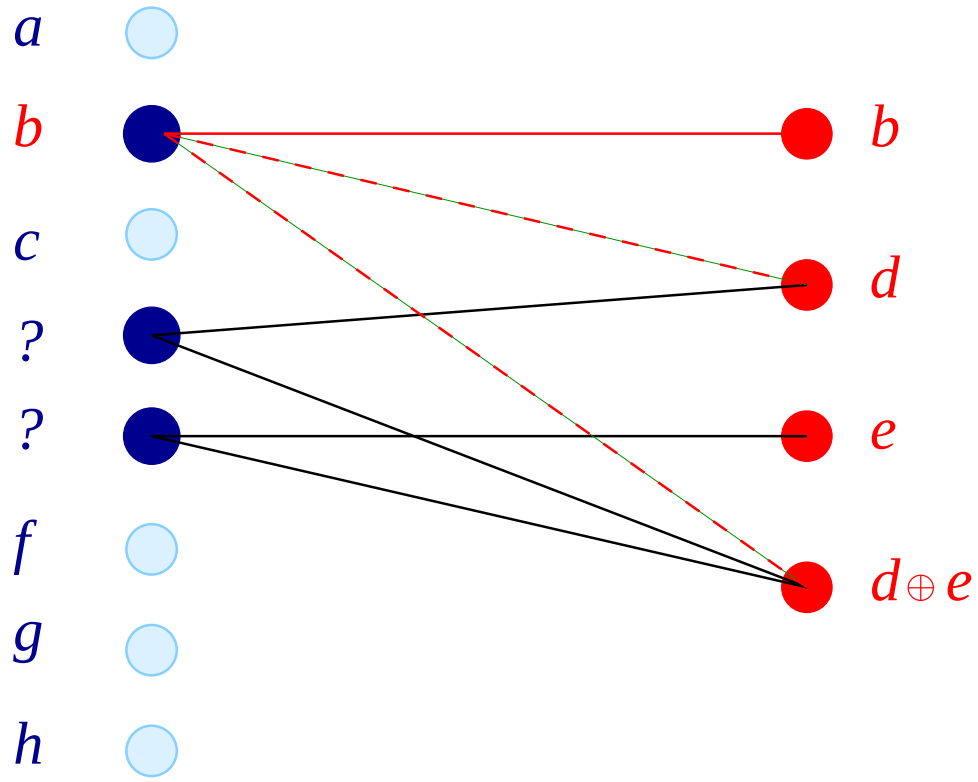
Luby-Mitzenmacher-Shokrollahi-Spielman-Stemann, 1997:

Phase 1: Direct recovery



Decoding

Phase 2: Substitution



Decoding Time

An addition is necessary for each edge in the graph.

If decoding is possible, then it can be done in time linear in n, the block length of the code.

Is decoding possible? (Depends on number of erasures.)

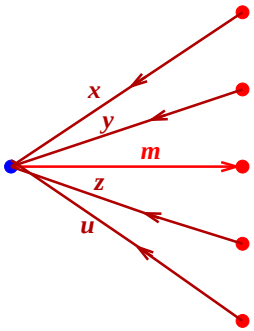
Inverse problem: We have a fast decoding algorithm. We want to design the graph in such a way that many erasures are correctable with this algorithm.

Message Passing Algorithm

Phrase decoding as algorithm in which messages are passed from message nodes to check nodes and back.

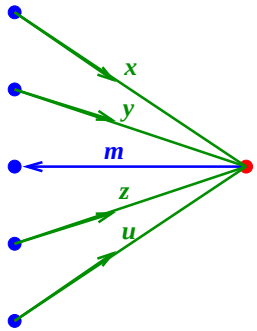
Messages are 0 or 1.

First round: erased message nodes send message 0, non-erased message nodes send message 1.



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

MESSAGE



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

CHECK

Heuristic Analysis

Let p_i denote the probability that a message node sends message 0 to a check node at round i, and let q_i be the probability that check node sends message 0 to a message node at round i.

What is p_{i+1} in terms of p_i and q_i?

If degree of message node is d, and incoming messages are independent random variables, then $p_{i+1} = q_i^{d-1}$.

If degree of check node is d and incoming messages are independent random variables, then $q_i = 1 - (1 - p_i)^{d-1}$.

General Recursion

λ_d is probability that edge in graph is connected to message node of degree d.

ρ_d is probability that edge in graph is connected to check node of degree d.

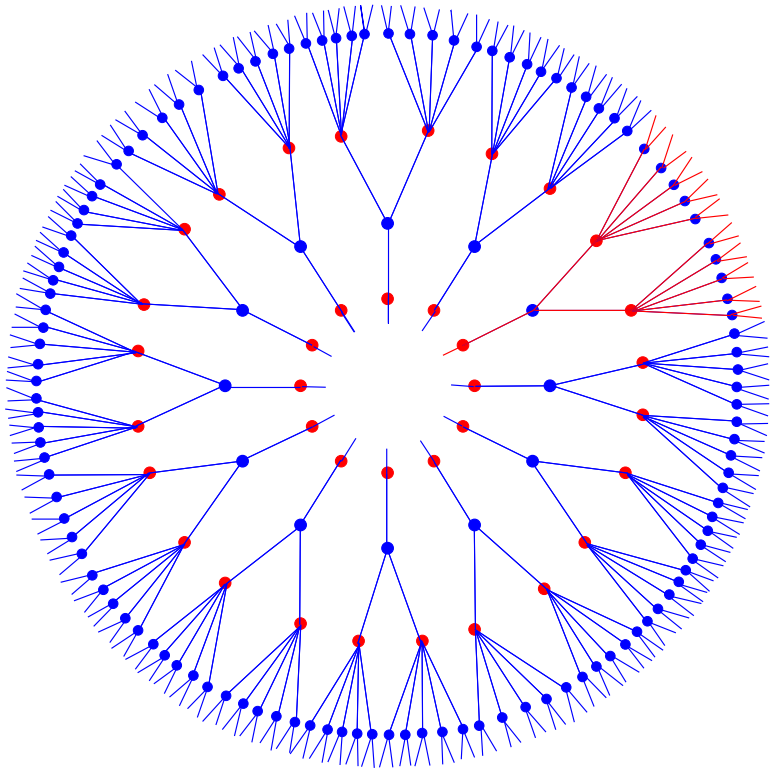
Then $p_{i+1} = p \cdot \lambda(1 - \rho(1 - x))$ if incoming messages are independent.

Condition for successful decoding:

$$p \cdot \lambda(1 - \rho(1 - x)) < x \qquad \text{for } x \in (0, p) \ .$$

Exact Proof

Expand neighborhood of message node



If neighborhood is tree up to depth 2ℓ, then ℓ rounds of algorithm can be accounted for by heuristic analysis.

Exact Proof

Fix iteration number ℓ. If the number of message (and check) nodes is large enough, and minimum degree of a message node is larger than 2, and the graph is chosen randomly, then the neighborhood of depth 2ℓ of most message nodes is a tree.

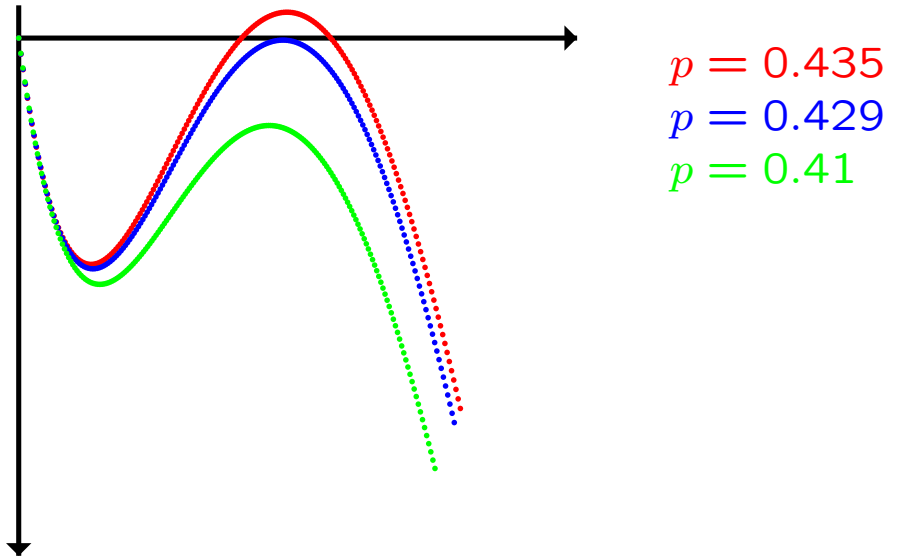
This (and some martingale arguments) shows that if above condition is satisfied, then decoder can reduce fraction of erasures to below any ϵn.

Rest of analysis is done via expander arguments.

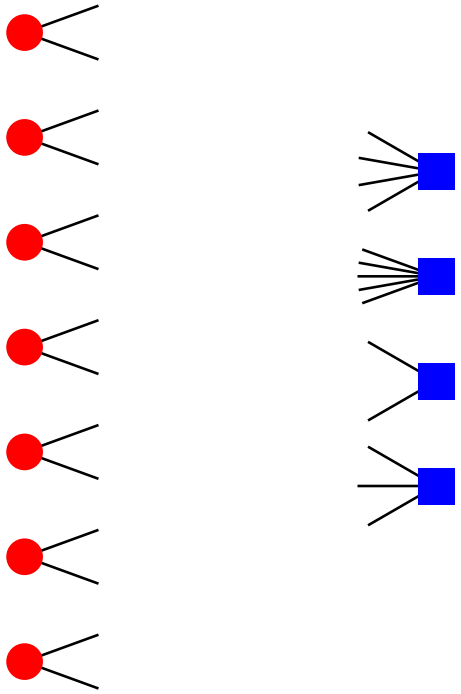
First Example: $(3,6)$-Graph

$\lambda(x) = x^2$, $\rho(x) = x^5$, $p(1-(1-x)^5)^2 < x$?

Consider $f(x) = p(1-(1-x)^5)^2 - x$, and test whether $f(x) < 0$ on $(0,p)$:

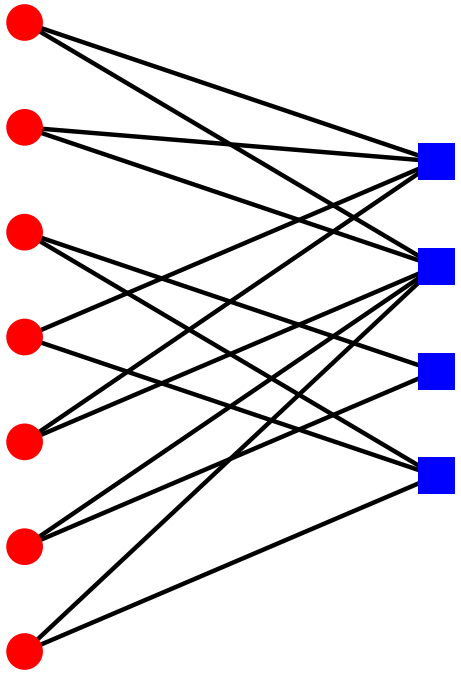


Example



Message nodes are of degree 2, choose their neighbors randomly.

Example



Message nodes are of degree 2, choose their neighbors randomly.

Erasure Correction

$\lambda(x) = x$.

$\rho(x) = \mathrm{e}^{\alpha(x-1)}$ asymptotically, where $\alpha = 2n/r$, n is number of message nodes and r is number of check nodes. (Poisson distribution.)

Condition:

$$p \cdot (1 - \mathrm{e}^{-\alpha x}) < x \qquad \text{for } x \in (0, p)$$

is equivalent to

$$1 - x - \mathrm{e}^{-\beta x} < 0 \qquad \text{for } x \in (0, 1), \beta = p\alpha.$$

This is true only if $\beta < 1$, i.e., $p < r/(2n)$.

Connection to Random Graphs

Consider graph on check nodes in which edges are formed by the erased message nodes.

This is a random (multi-)graph on the check nodes with average degree β.

Well-known theorem: a random graph with average degree β has a giant 2-core iff $1 - x - \mathrm{e}^{-\beta x} = 0$ has a solution x with $0 < x < 1$. This solution is the size of the giant 2-core. (2-core of a graph is maximal subgraph in which minimal degree is 2.)

A 2-core in the graph leads to failure of the decoder.

We have recovered this theorem using coding theory!

Encoding

Can be done in linear time for these graphs.

Will describe the algorithm on the blackboard.

Achieving Capacity

Choose parameter D, and set

$$\lambda(x) = \frac{1}{H(D)} \sum_{d=1}^{D} \frac{x^d}{d},$$

$$\rho(x) = \mathrm{e}^{\alpha(x-1)},$$

where α is chosen such that rate of code is R.

Then for $p = (1 - R)(1 - 1/(D + 1))$ we have $p\lambda(1 - \rho(1 - x)) < x$ on $(0, p)$, hence these codes come arbitrarily close to channel capacity with a linear time algorithm!

These codes are called Tornado codes. (Luby, Mitzenmacher, Shokrollahi, Spielman, 1997).