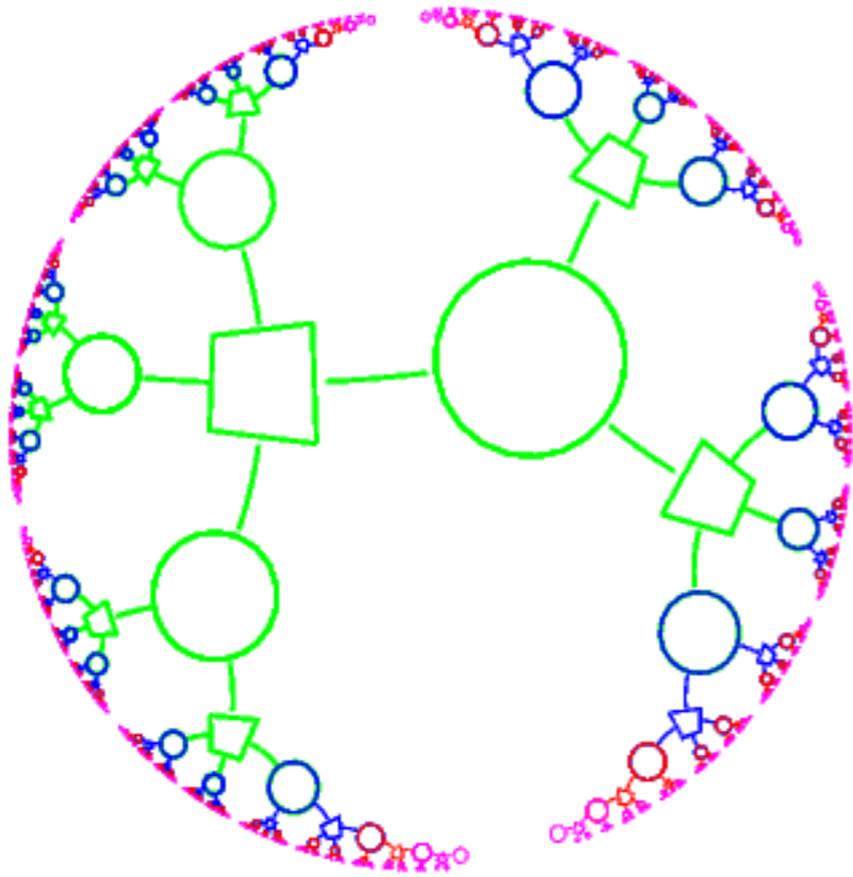
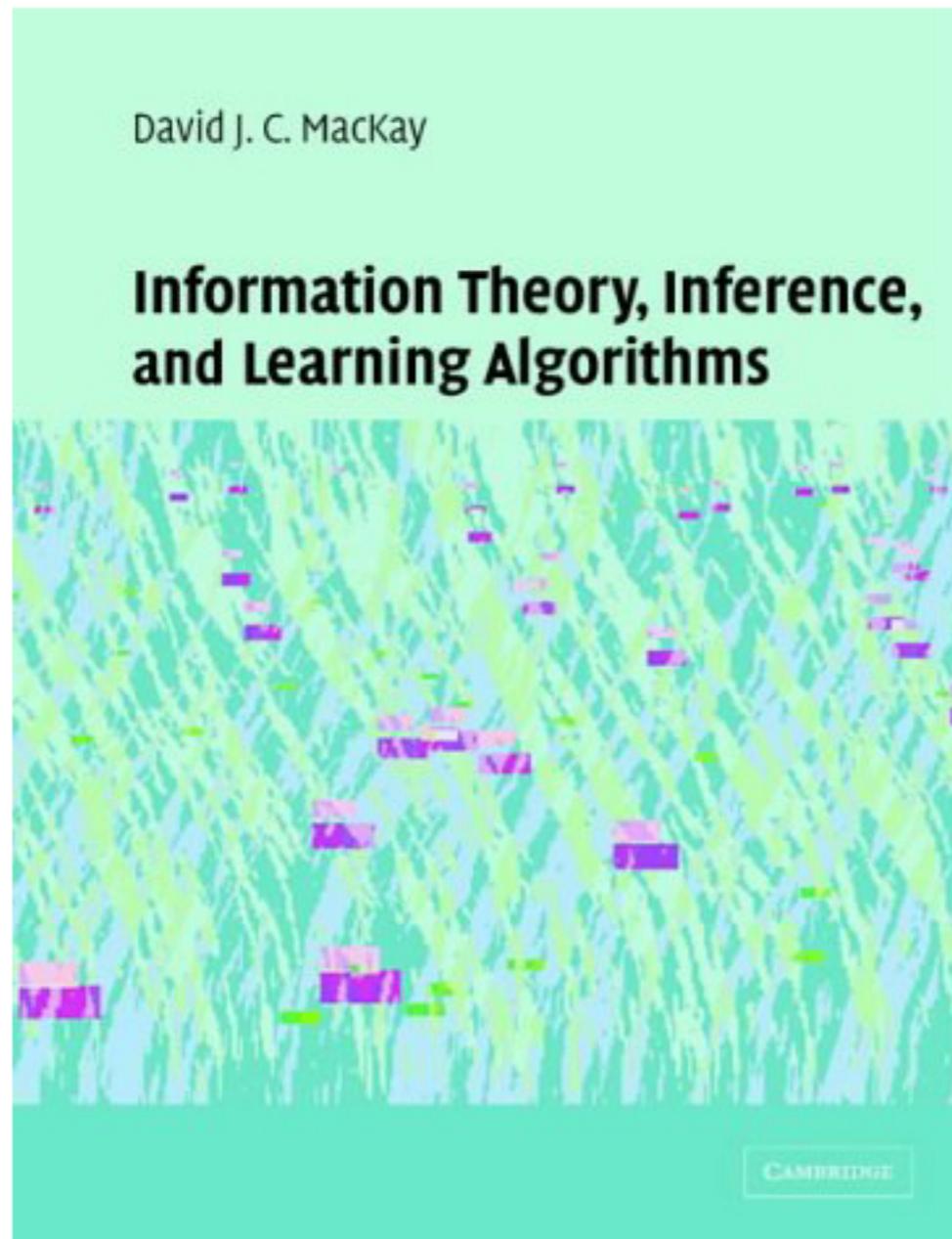


Error-correcting codes



Lecture notes - Chapters 1 & 47

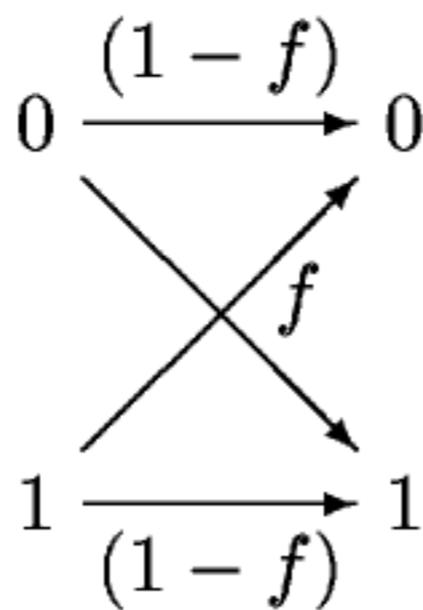


- Cambridge University Press
- 640 pages, 35 pounds
- 28 pounds at CUP bookshop
- Also available **free online**

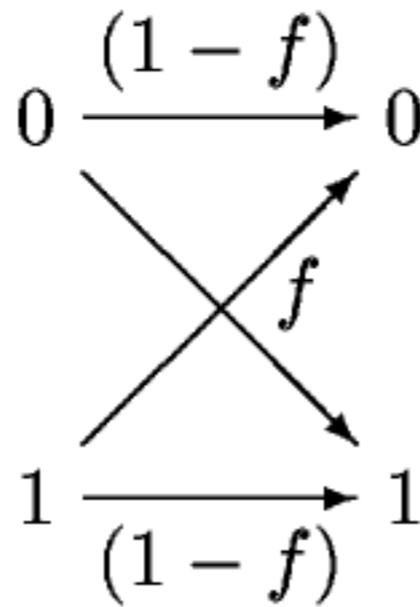
www.inference.phy.cam.ac.uk/mackay/itila/

Purpose: reliable communication over unreliable channels

eg, Binary symmetric channel



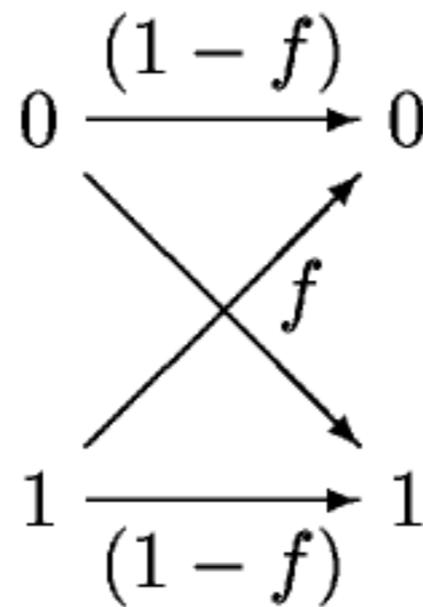
$$f = 0.1$$



Q: A file of $N = 10\,000$ bits is stored on this disc drive (with $f = 0.1$), then read.

Roughly how many bits are flipped?

\pm



Q: To make a successful business selling 1 Gigabyte disc drives, how small does the flip probability f need to be?

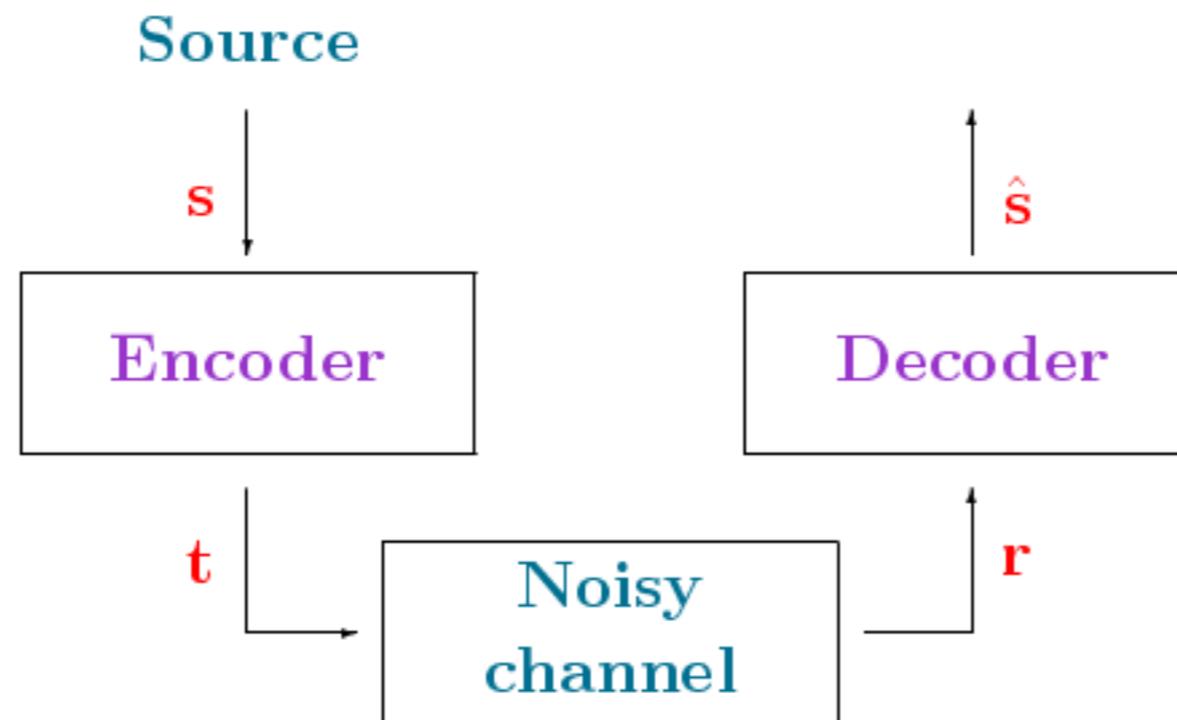
Solutions:

● Physical solution



● System solution

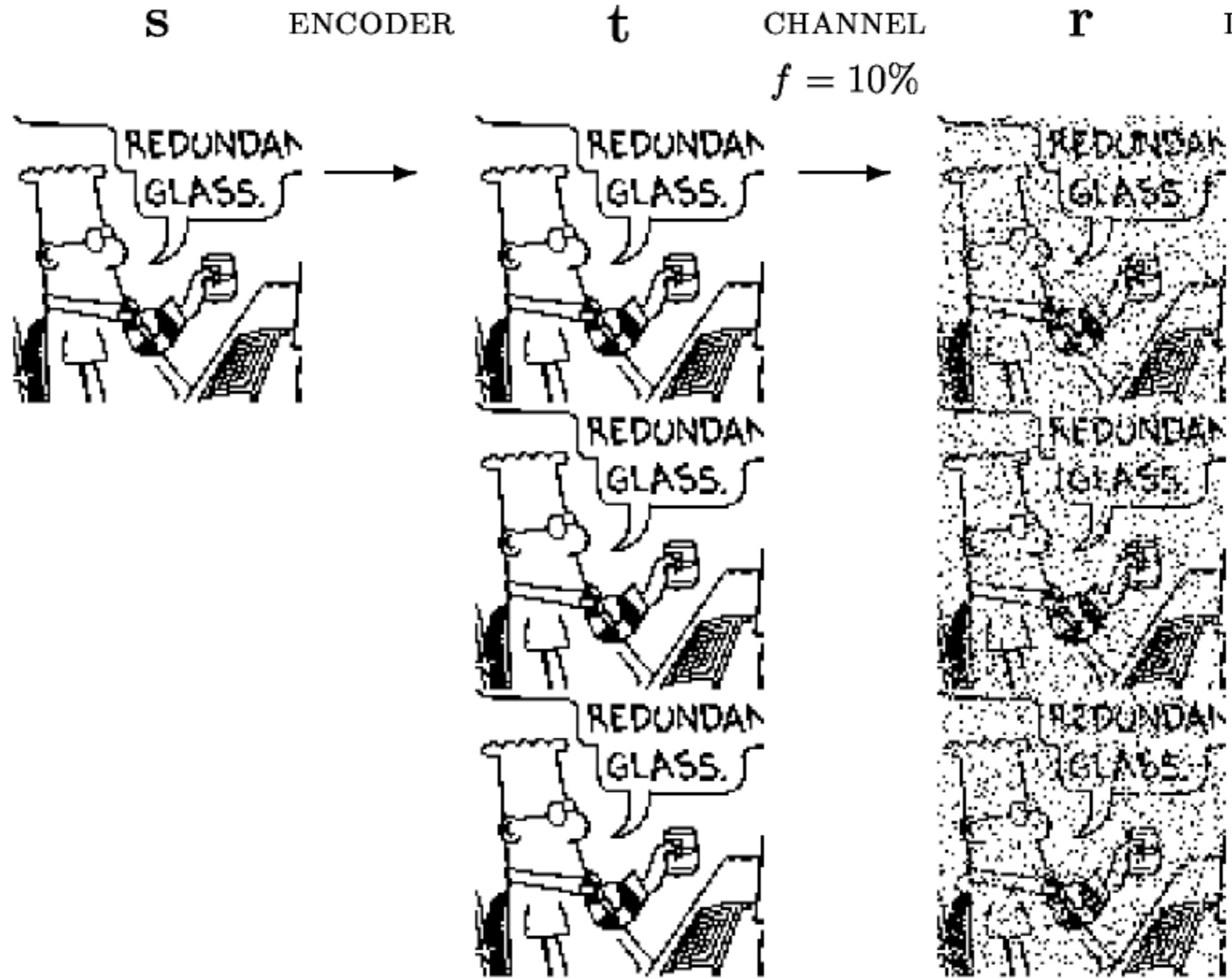
System solution



adds **redundancy**

does inference

Repetition code 'R3'



Repetition code 'R3'

S

ENCODER

t

CHANNEL

r

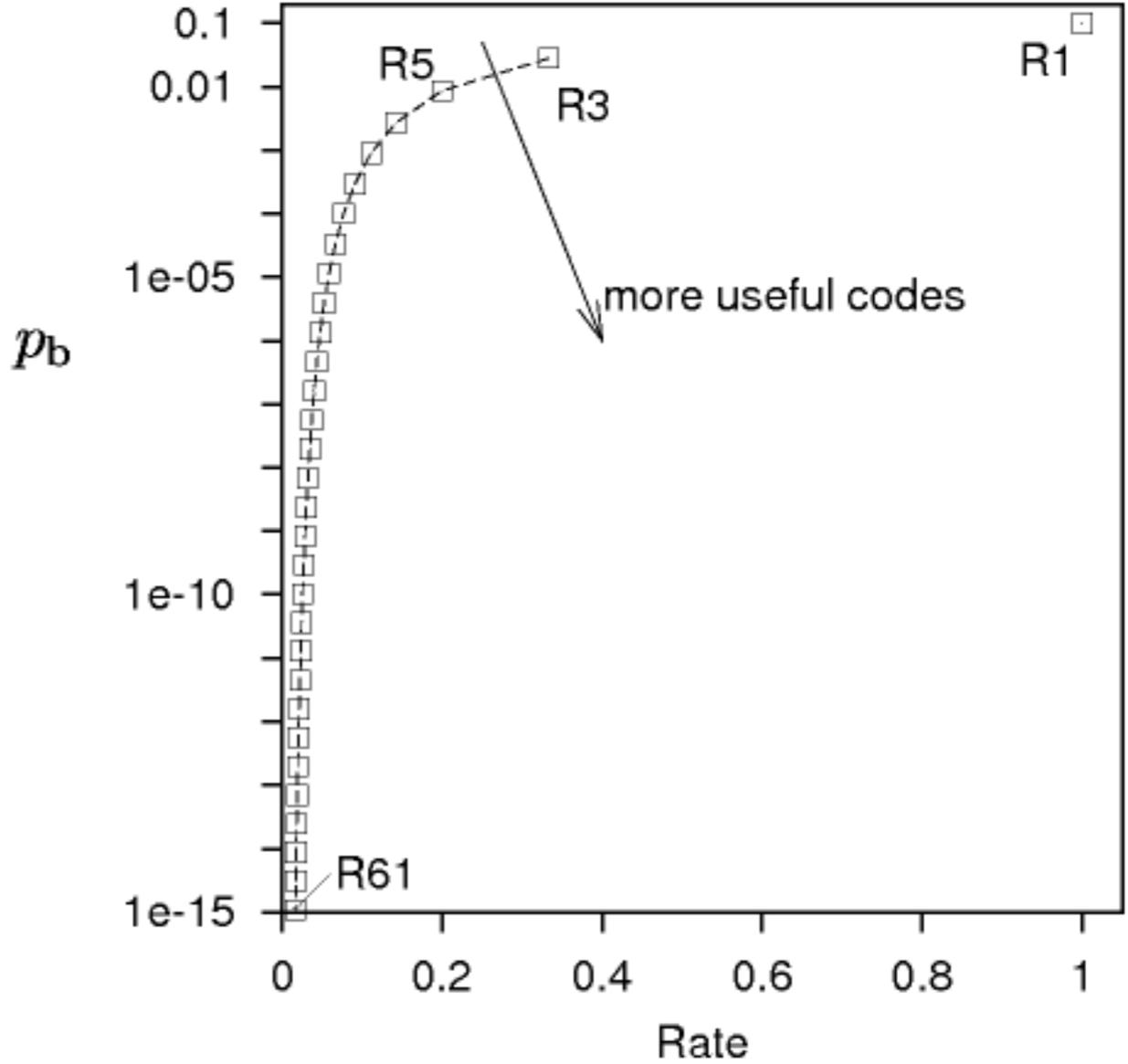
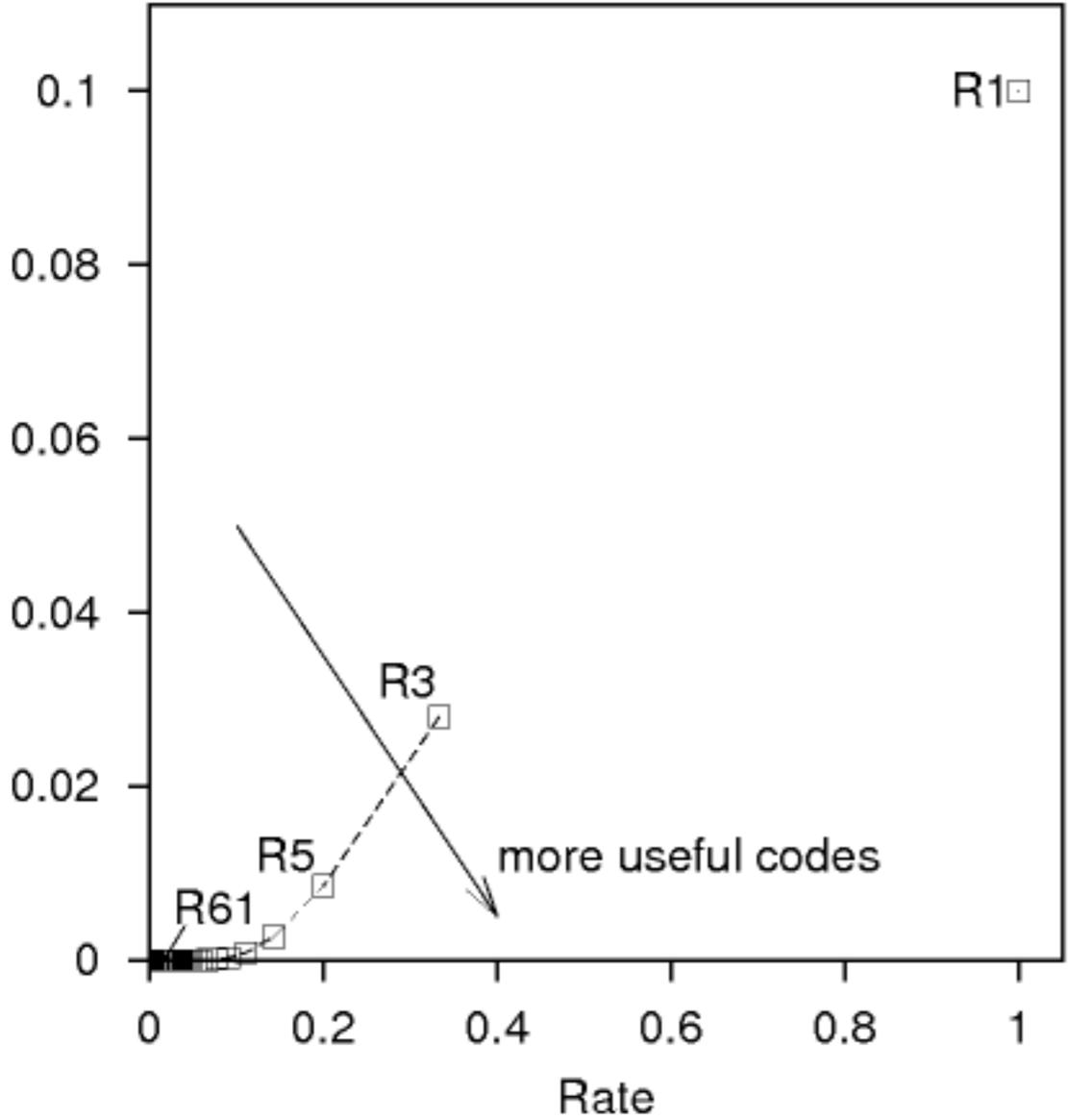
DECODER

\hat{S}

$f = 10\%$



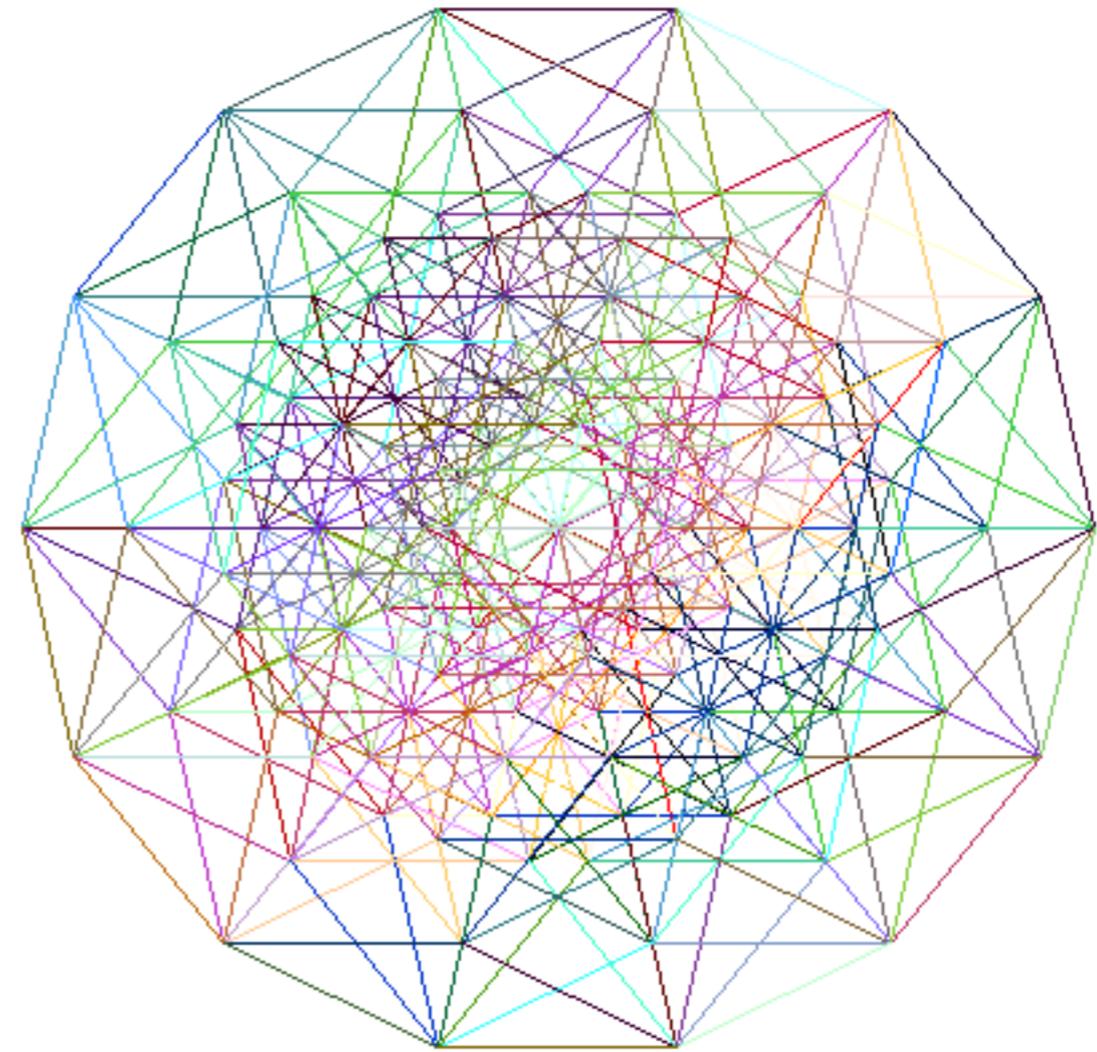
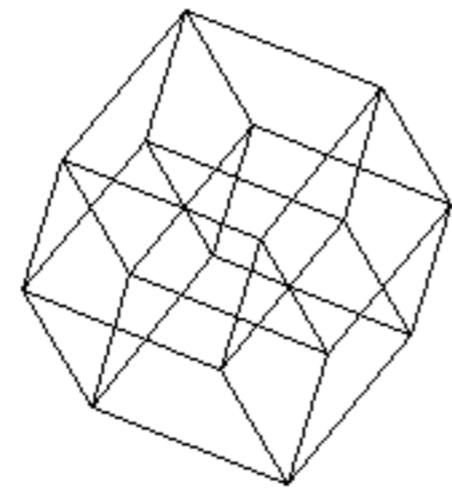
Performance of repetition codes



Block codes

(7,4) Hamming Code

s	t	s	t	s	t	s	t
0000	0000 000	0100	0100 110	1000	1000 101	1100	1100 011
0001	0001 011	0101	0101 101	1001	1001 110	1101	1101 000
0010	0010 111	0110	0110 001	1010	1010 010	1110	1110 100
0011	0011 100	0111	0111 010	1011	1011 001	1111	1111 111



(7,4) Hamming Code

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \uparrow \\ M = 3 \\ \downarrow \end{matrix}$$

$\leftarrow N = 7 \rightarrow$

Valid transmissions \mathbf{t} satisfy

$$\mathbf{H} \mathbf{t} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \pmod{2}$$

(7,4) Hamming Code

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{array}{l} \uparrow \\ M = 3 \\ \downarrow \end{array}$$

$\leftarrow N = 7 \rightarrow$

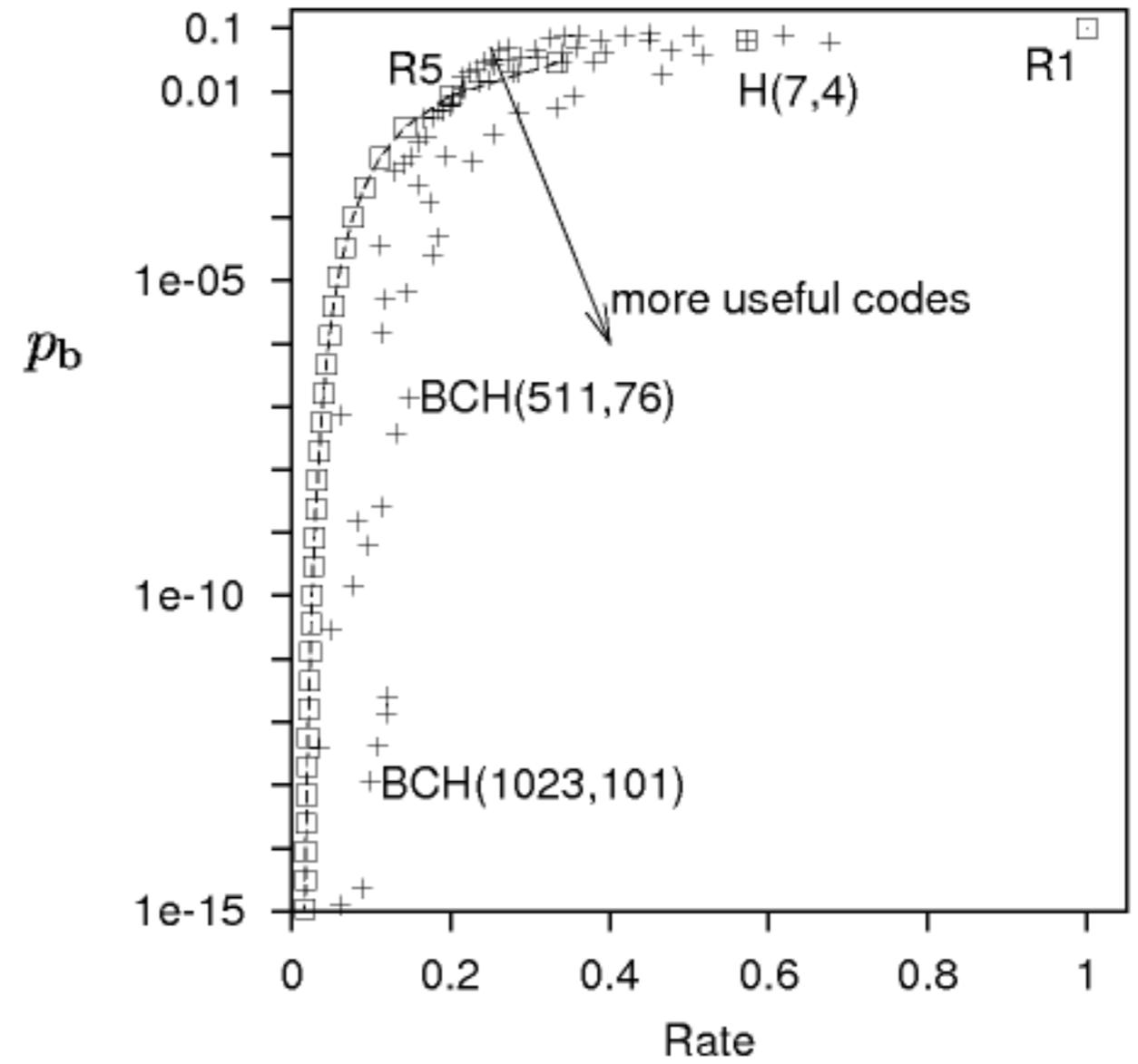
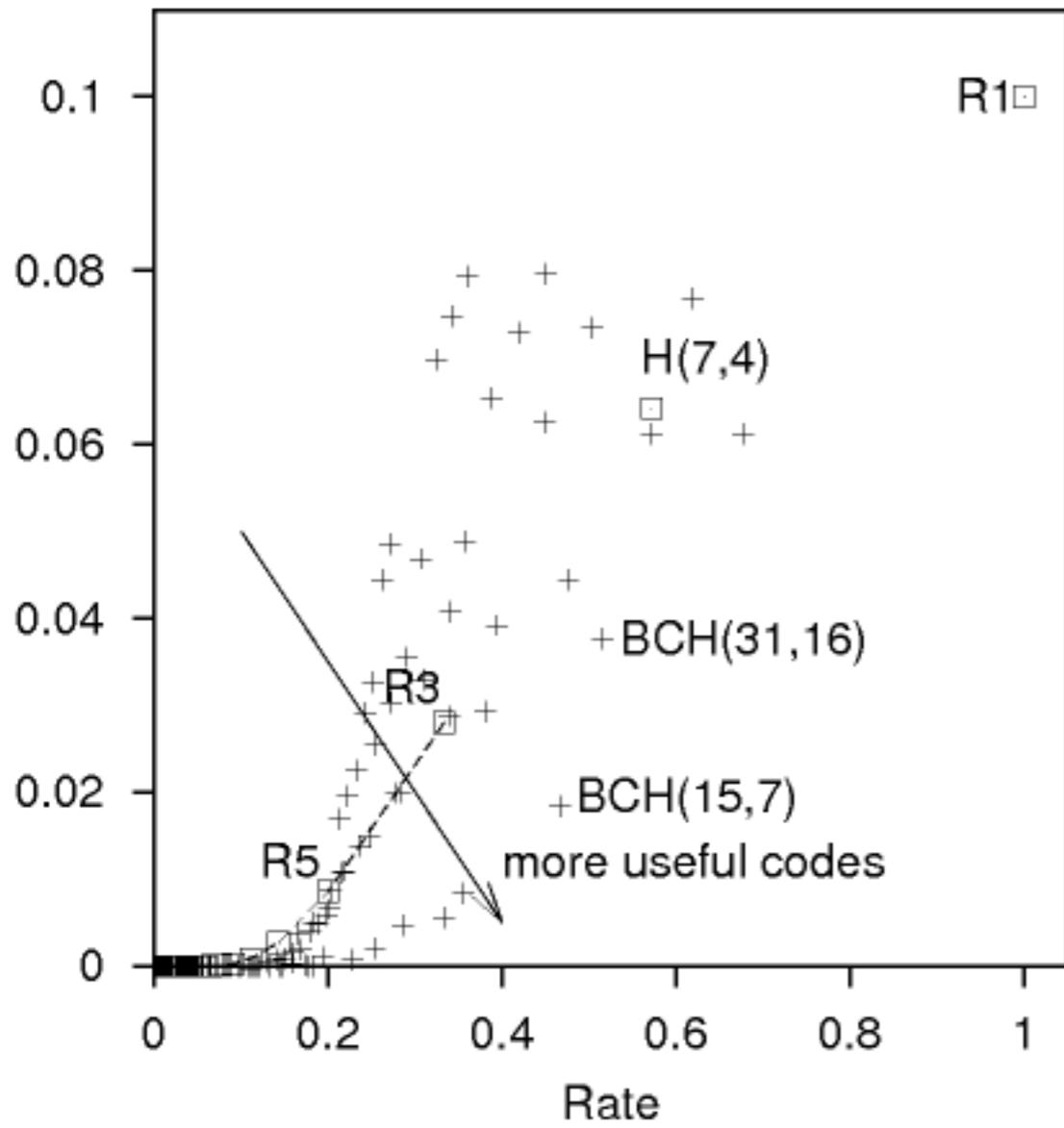
Valid transmissions \mathbf{t} satisfy

$$\mathbf{H} \mathbf{t} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ mod } 2$$

Received signal $\mathbf{r} = \mathbf{t} + \mathbf{n}$

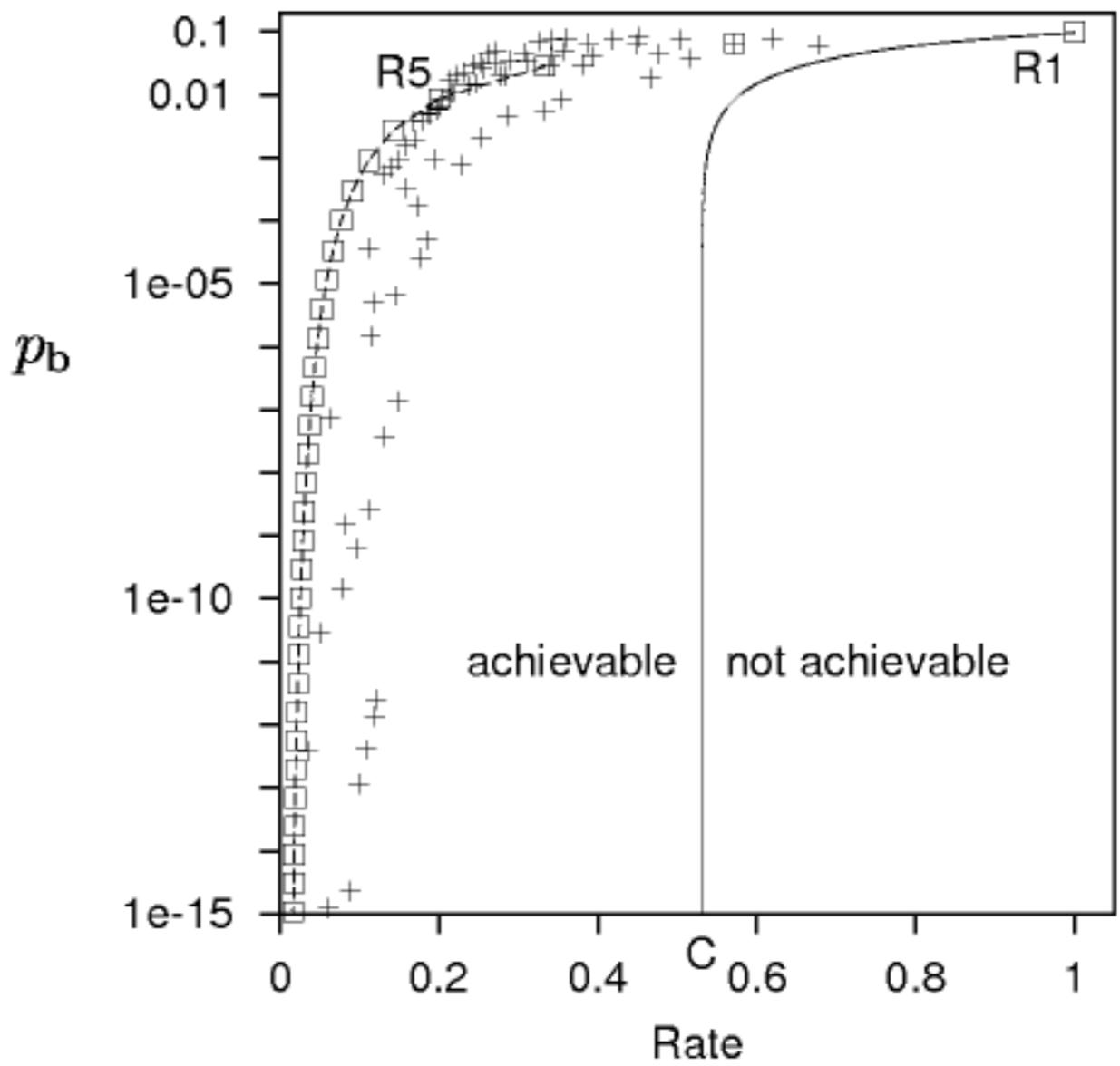
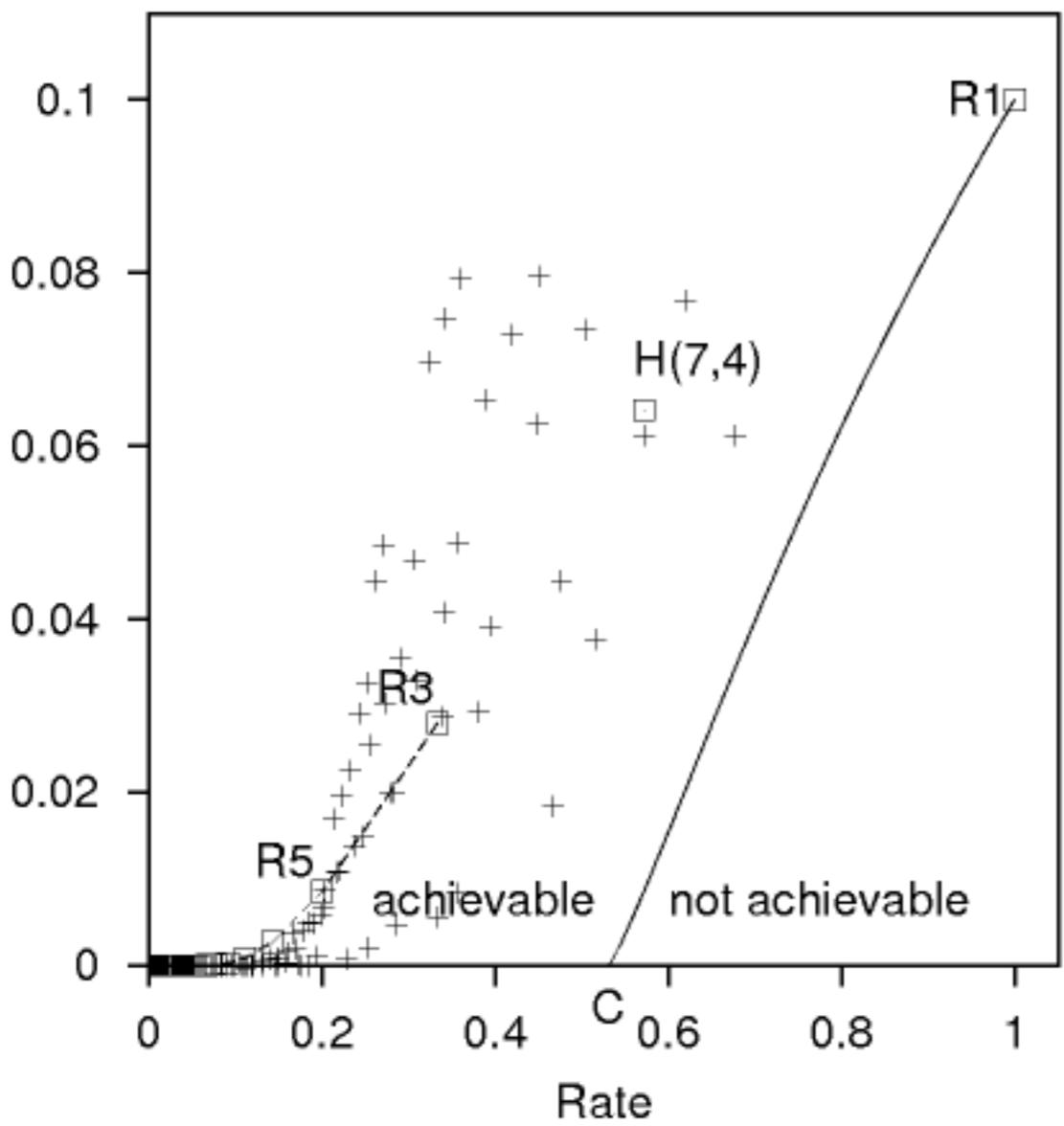
Syndrome $\mathbf{z} = \mathbf{H} \mathbf{r} = \mathbf{H} \mathbf{n}$.

Syndrome decoder $\mathbf{z} \longrightarrow \hat{\mathbf{n}}$.



What's achievable?

Shannon's noisy-channel coding theorem



$$C_{\text{BSC}}(f) = 1 - H_2(f)$$

$$H_2(f) = f \log_2 \frac{1}{f} + (1 - f) \log_2 \frac{1}{1 - f}$$

How to prove good codes exist

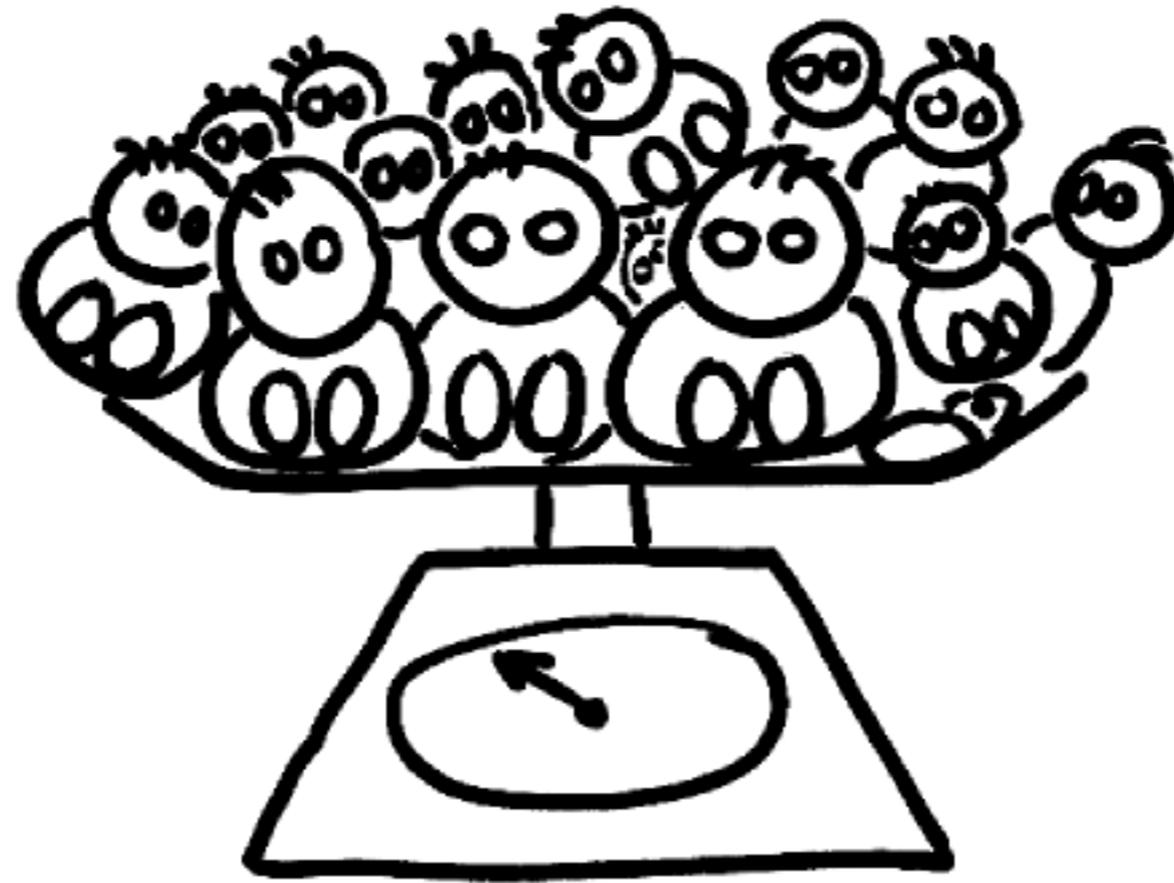
● Constructive proof

Given required $R < C$, and $\epsilon > 0$,

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 0 & & & \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & \dots & \dots & \dots \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & & & \\ \vdots & & & \vdots & & \vdots & & \ddots & & \\ \vdots & & & \vdots & & \vdots & & & \ddots & \\ \vdots & & & \vdots & & \vdots & & & & \ddots \end{bmatrix}$$

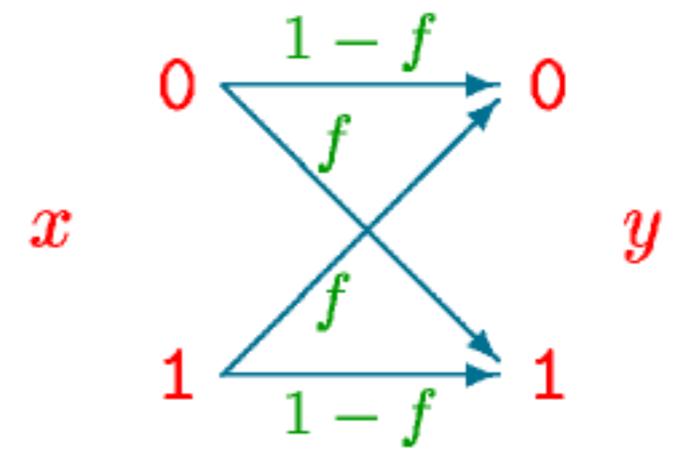
● Non-constructive proof

Shannon's way of proving malnutrition

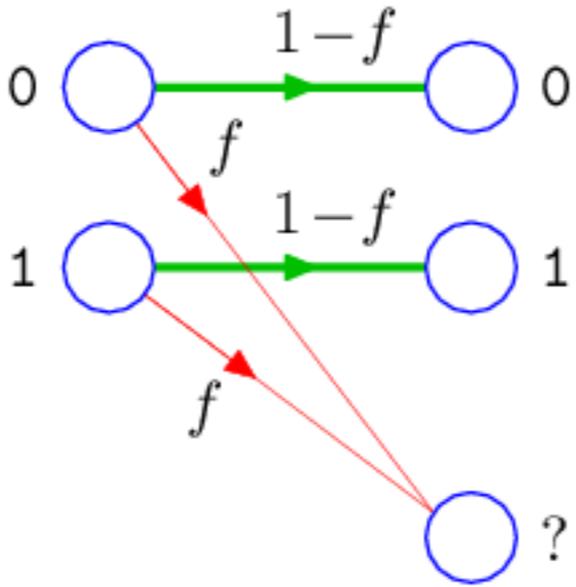


If **average** weight
of all babies is $< \epsilon$,
there must be **(at least!)**
one baby with weight $< \epsilon$.

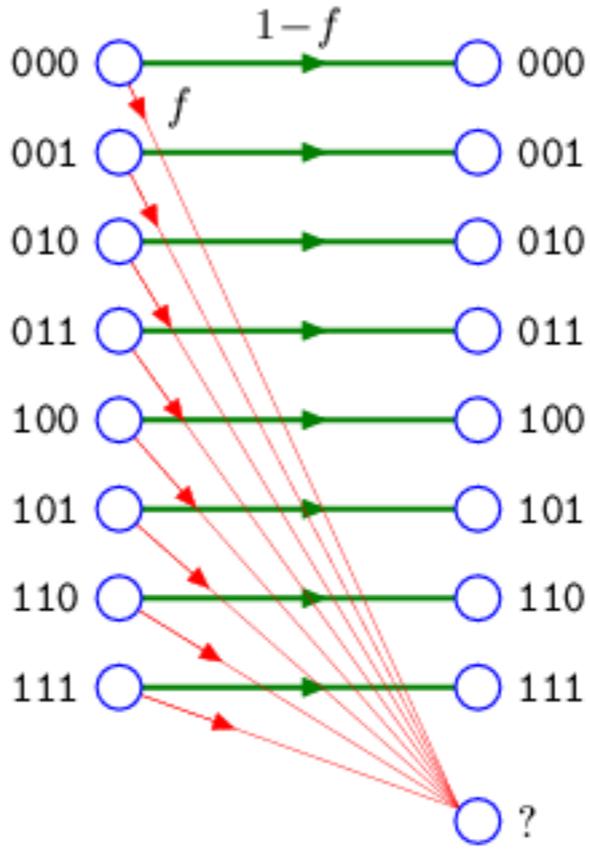
Shannon proved his noisy-channel theorem for any discrete memoryless channel



Channels with erasures



Binary erasure channel



8-ary erasure channel

Alice

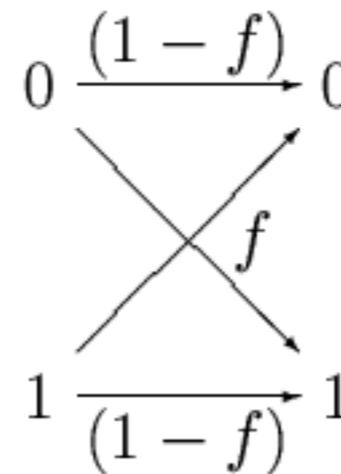


Packet-erasing channel

Bob



There are other noisy channels...

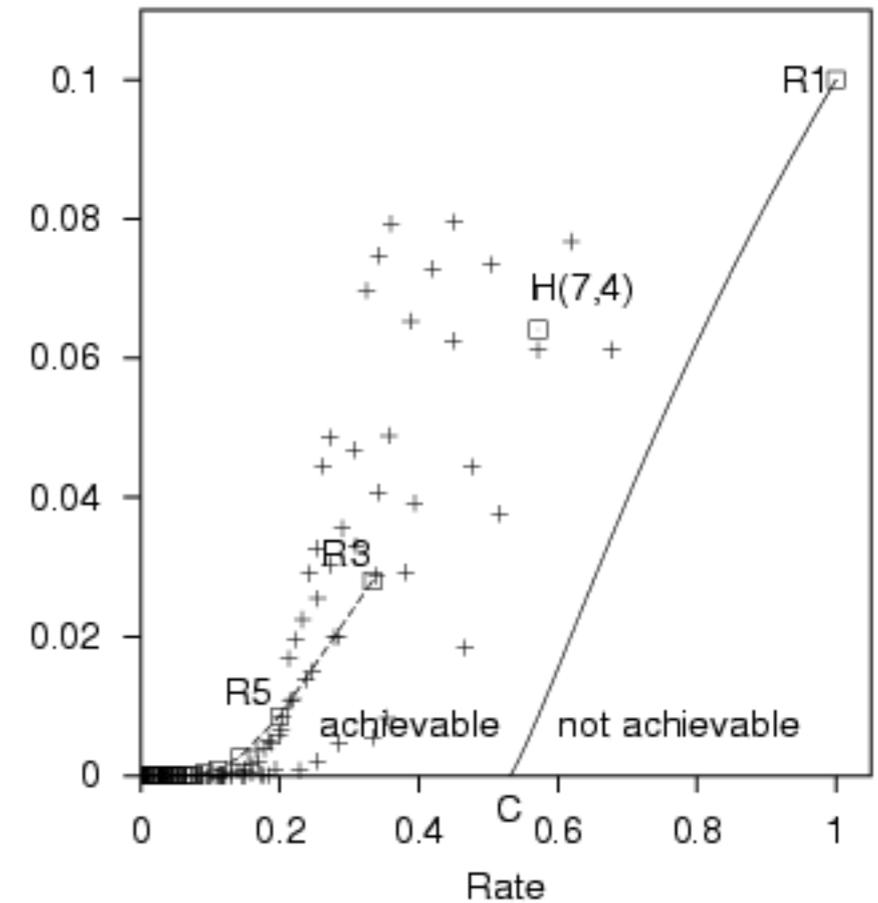
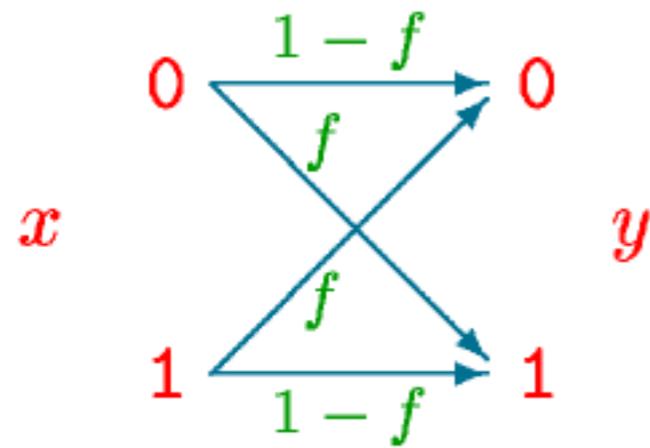


Insertions
and
Deletions
→



synchronization errors

Shannon's noisy channel coding theorem

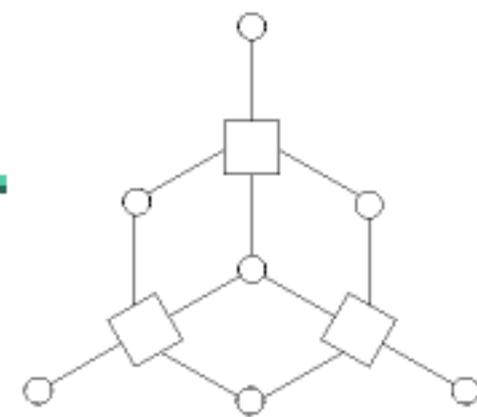


For any channel:
Reliable (virtually error-free) communication is possible
at rates up to C

Information theory Shannon, 1948

Coding theory Hamming, 1948; Reed-Solomon; Forney (Convolutional & concatenated codes)

Idea



● Decoding problems, such as

$$P_1(\mathbf{x}) = \frac{1}{Z_1} e^{\beta [x_1 x_2 x_3 x_5 + x_2 x_3 x_4 x_6 + x_1 x_3 x_4 x_7] + \sum_{n=1}^N b_n x_n}$$

look a bit like Boltzmann machines + Hopfield networks.... so

● Solve the decoding problem

$$\max_{\mathbf{x}} P_1(\mathbf{x})$$

using variational methods?

Electronics Letters, March 1995

Free energy minimisation algorithm for decoding and cryptanalysis

D.J.C. MacKay

Indexing terms: Decoding, Cryptography

An algorithm is derived for inferring a binary vector \mathbf{s} given noisy observations of \mathbf{As} modulo 2, where \mathbf{A} is a binary matrix. The binary vector is replaced by a vector of probabilities, optimised by free energy minimisation. Experiments on the inference of the state of a linear feedback shift register indicate that this algorithm supersedes the Meier and Staffelbach polynomial algorithm.

Decoding error-correcting codes

- For which codes are **approximate message-passing methods**
effective?

Electronics Letters, August 1996

Near Shannon limit performance of low density parity check codes

D.J.C. MacKay and R.M. Neal

[Low Density Parity Check Codes: Gallager 1962]

Indexing terms: Probabilistic decoding, Error correction codes

The authors report the empirical performance of Gallager's low density parity check codes on Gaussian channels. They show that performance substantially better than that of standard convolutional and concatenated codes can be achieved; indeed the performance is almost as close to the Shannon limit as that of turbo codes.

(7,4) Hamming Code - recap

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \uparrow \\ M \\ \downarrow \end{matrix}$$

$\leftarrow N=7 \rightarrow$

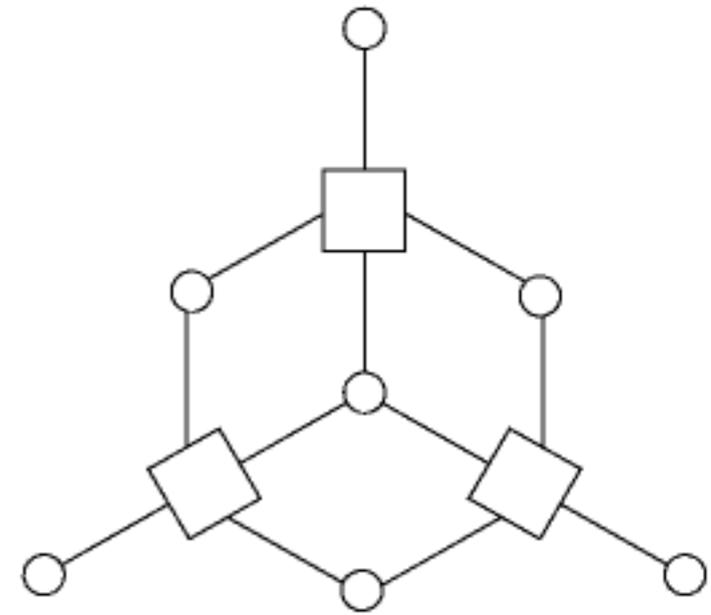
Valid transmissions \mathbf{t} satisfy

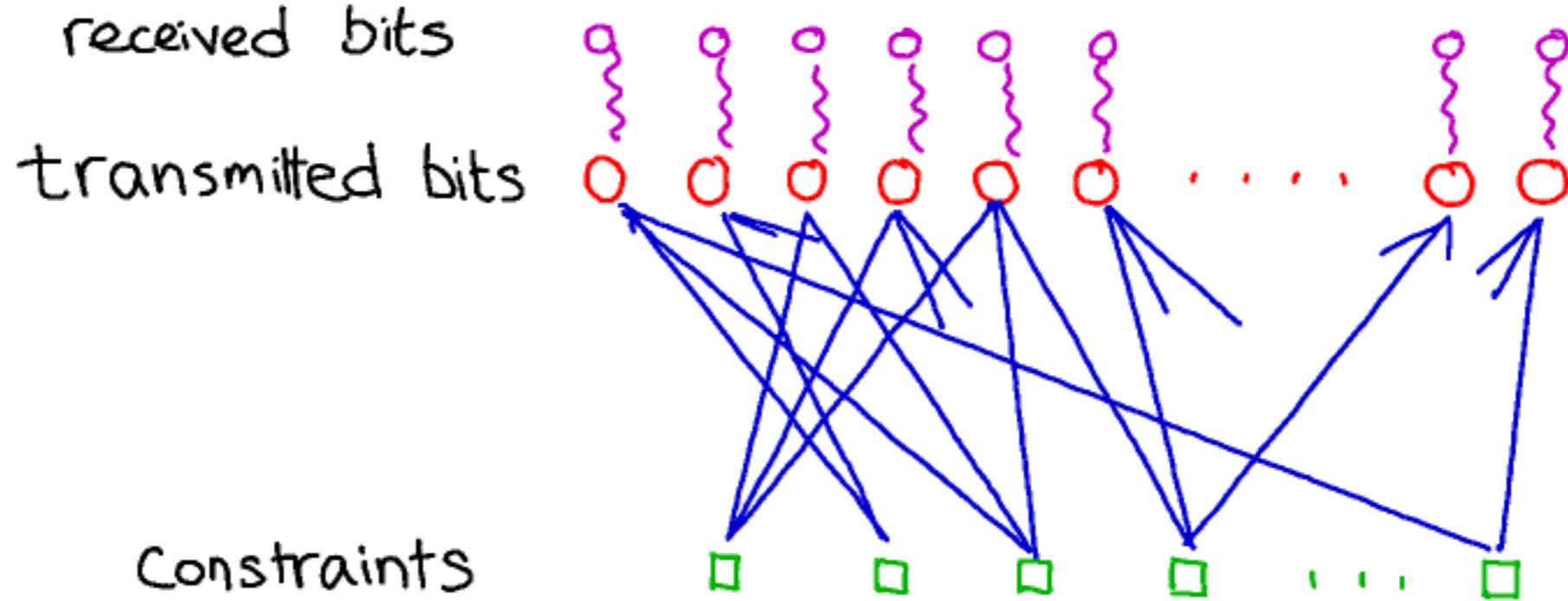
$$\mathbf{H} \mathbf{t} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ mod } 2$$

Received signal $\mathbf{r} = \mathbf{t} + \mathbf{n}$

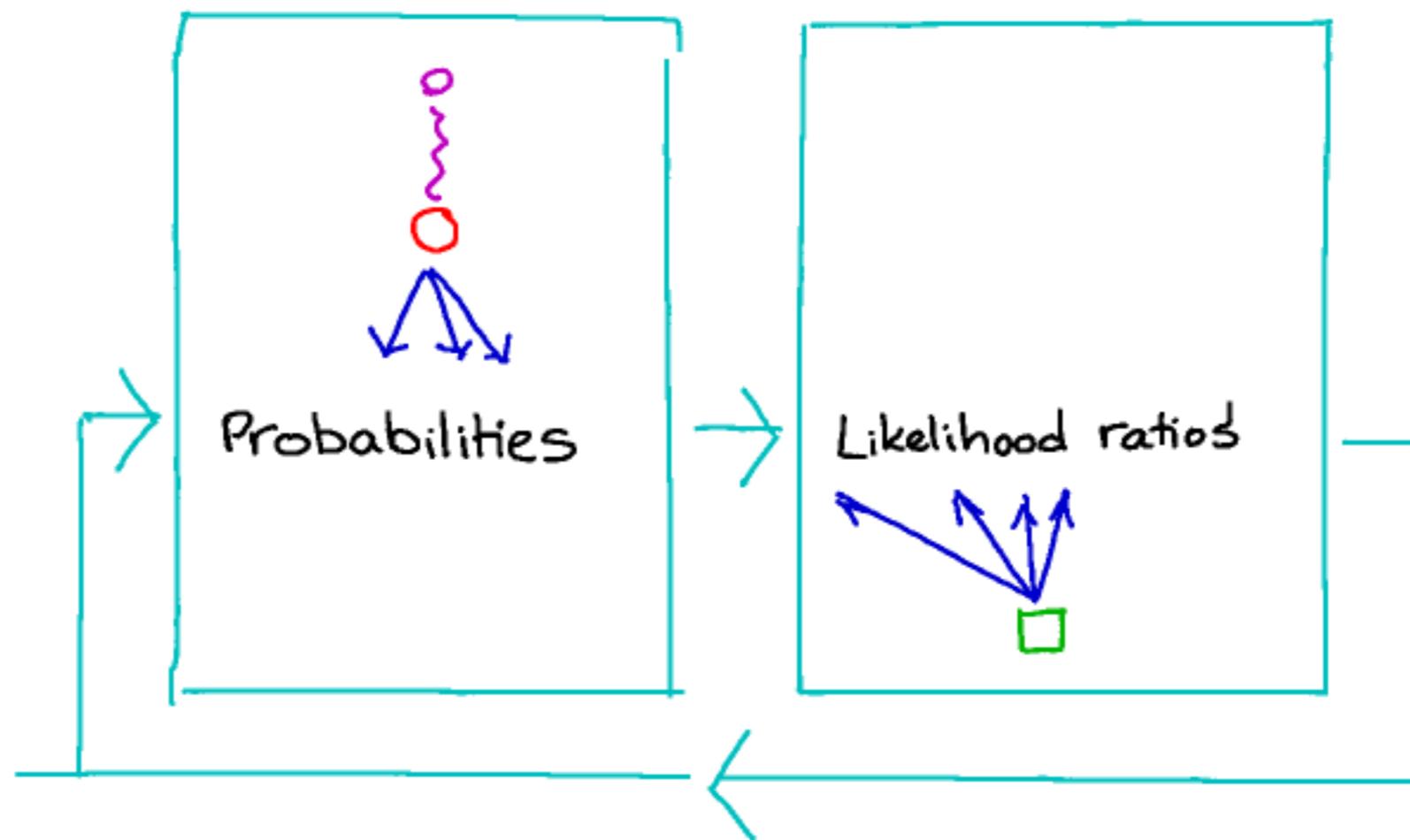
Syndrome $\mathbf{z} = \mathbf{H} \mathbf{r} = \mathbf{H} \mathbf{n}$.

Syndrome decoder $\mathbf{z} \rightarrow \hat{\mathbf{n}}$.





Decoding by the sum-product algorithm



Low Density Parity Check Code

We demonstrate a large code that encodes $K = 10000$ source bits into $N = 20000$ transmitted bits.

Each parity bit depends on about 5000 source bits.

The encoder is derived from a very sparse 10000×20000 matrix \mathbf{H} with three 1s per column.

TRANSMITTED:

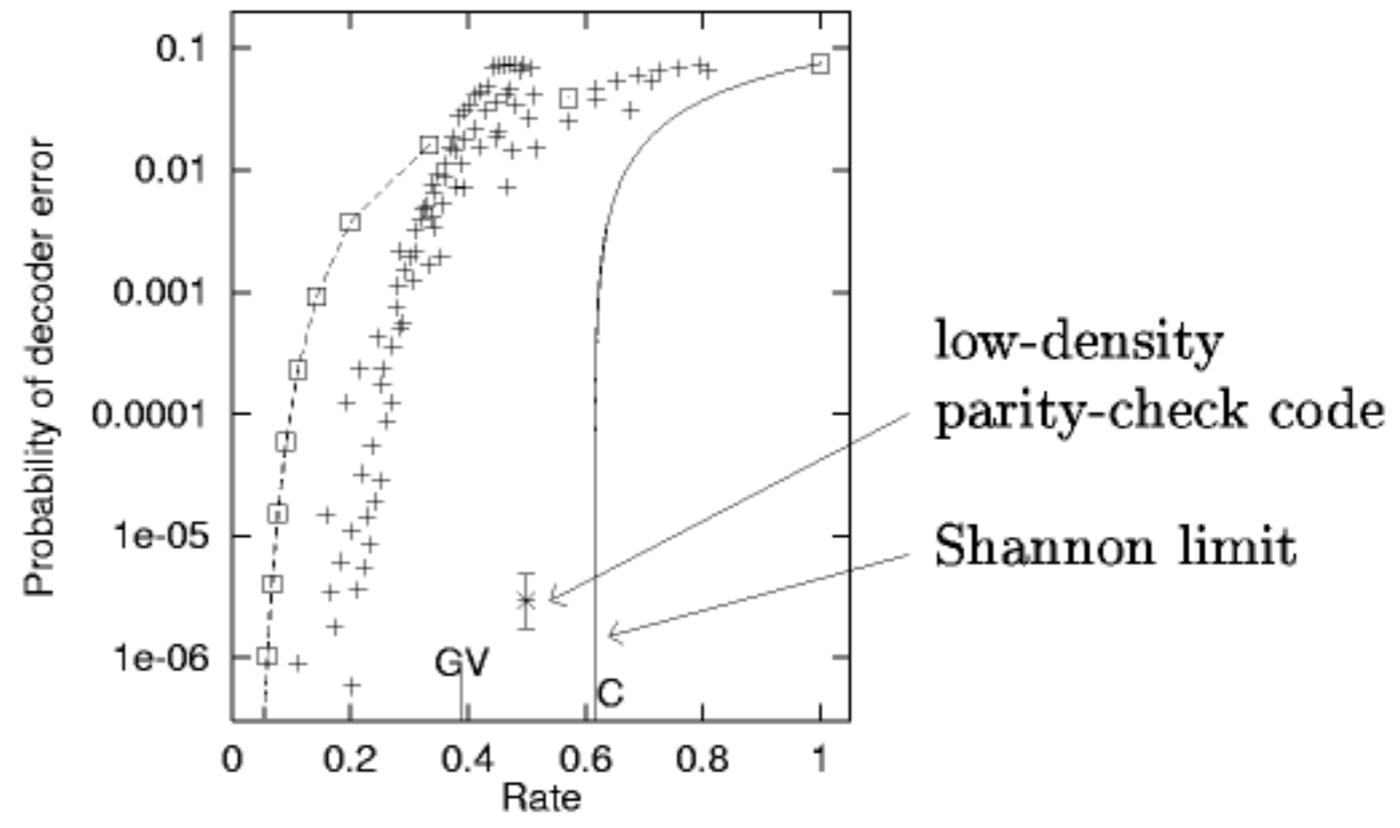


$\mathbf{H} =$



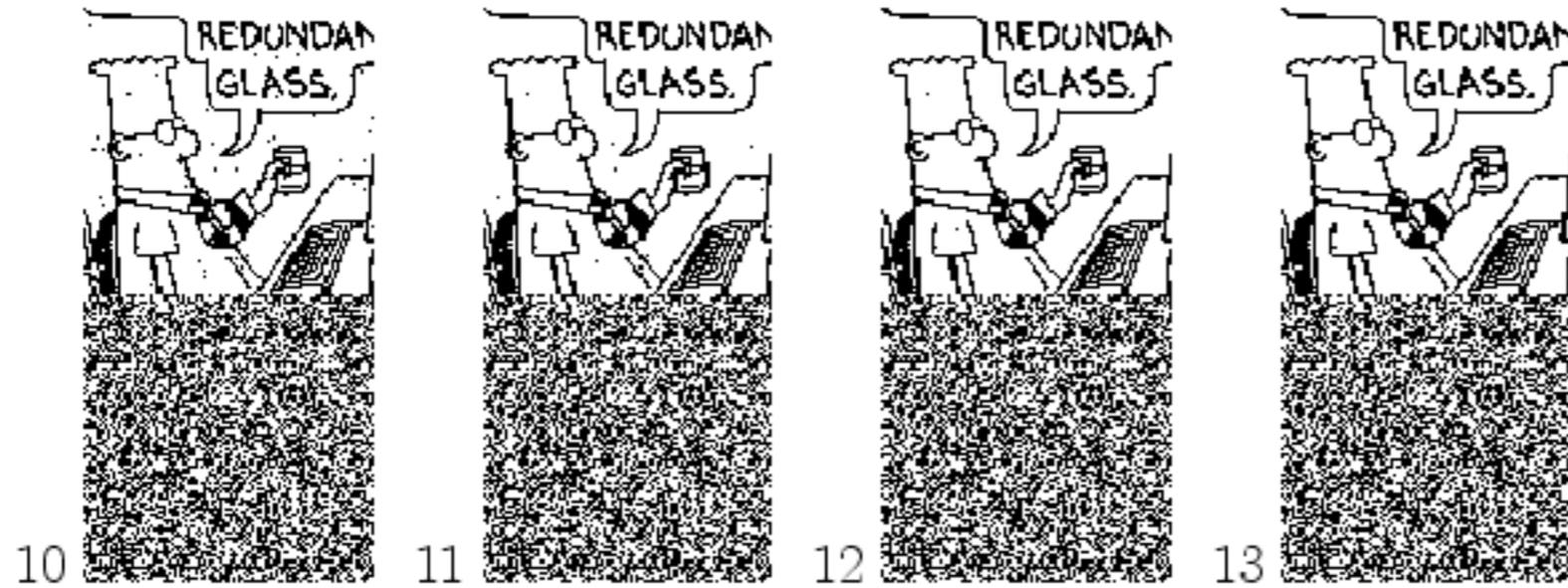
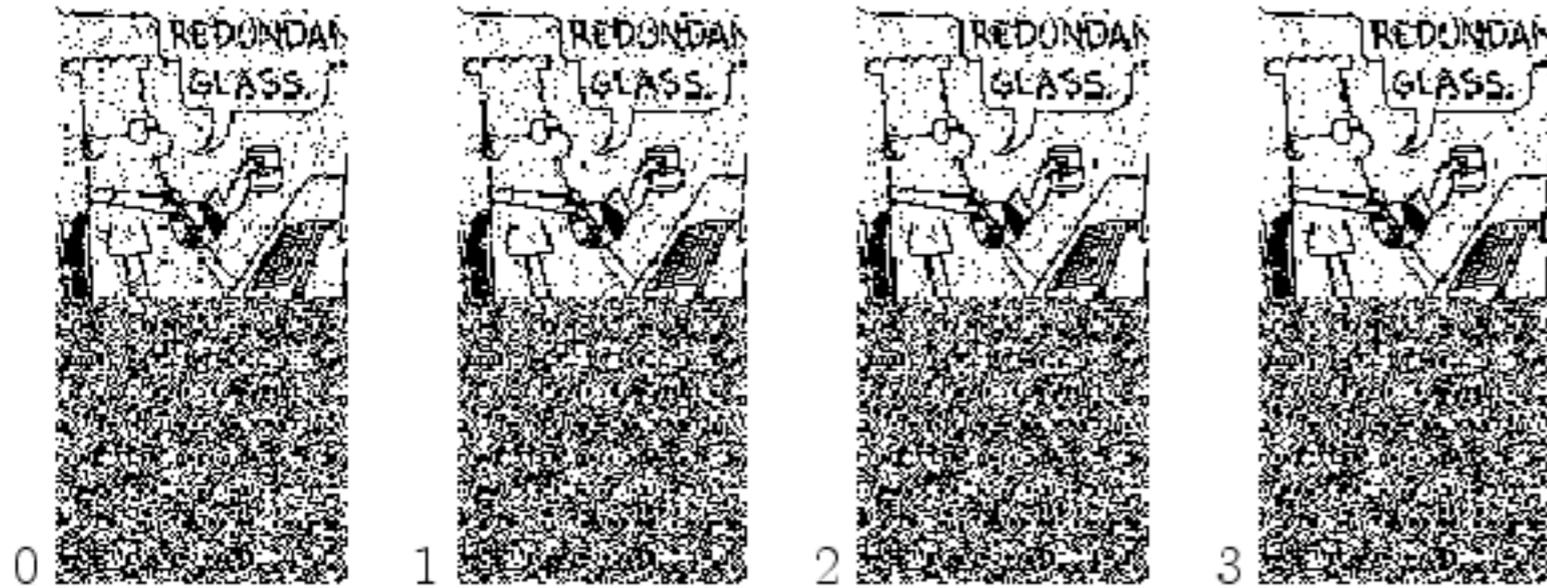
Low Density Parity Check Code (f = 7.5%)

Iterative probabilistic decoding



After the transmission is sent over a channel with noise level $f = 7.5\%$:

RECEIVED:

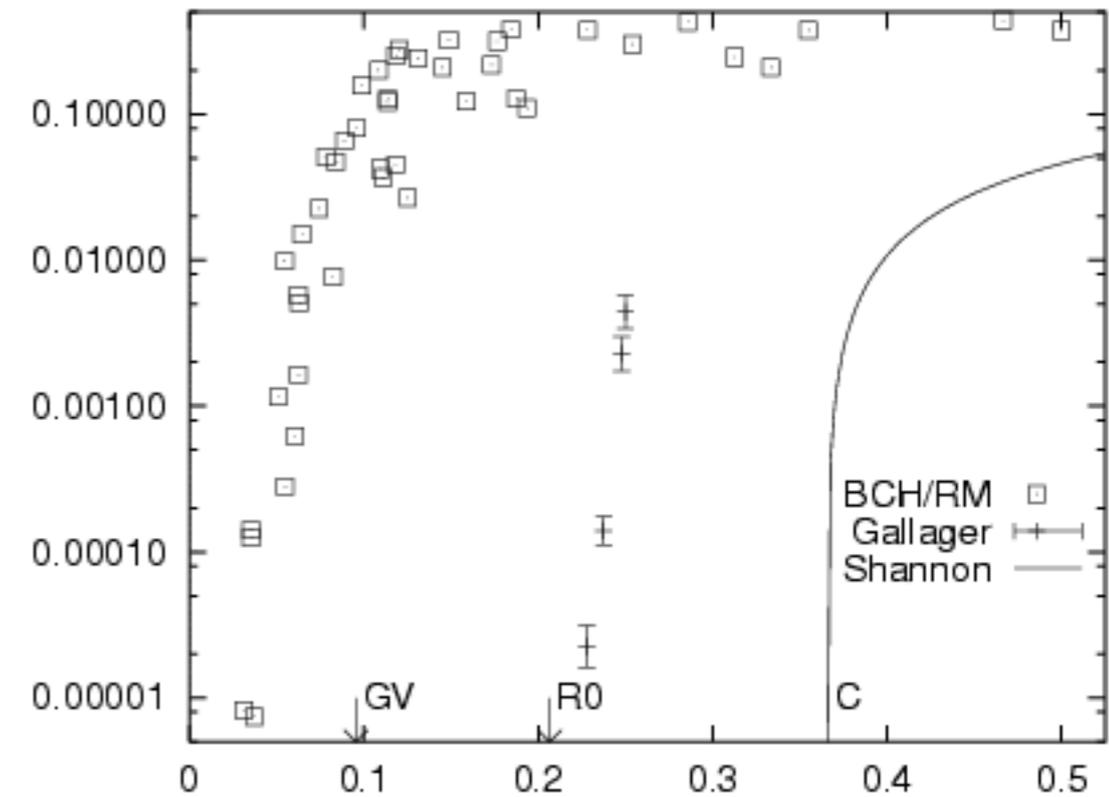
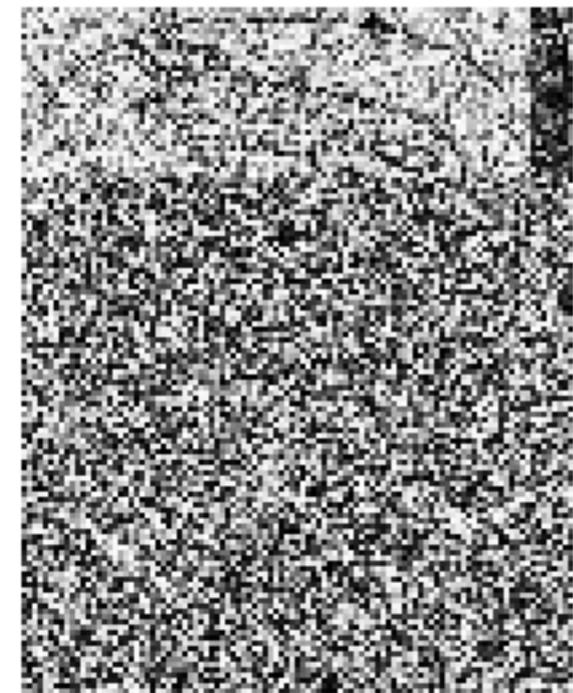


DECODED:

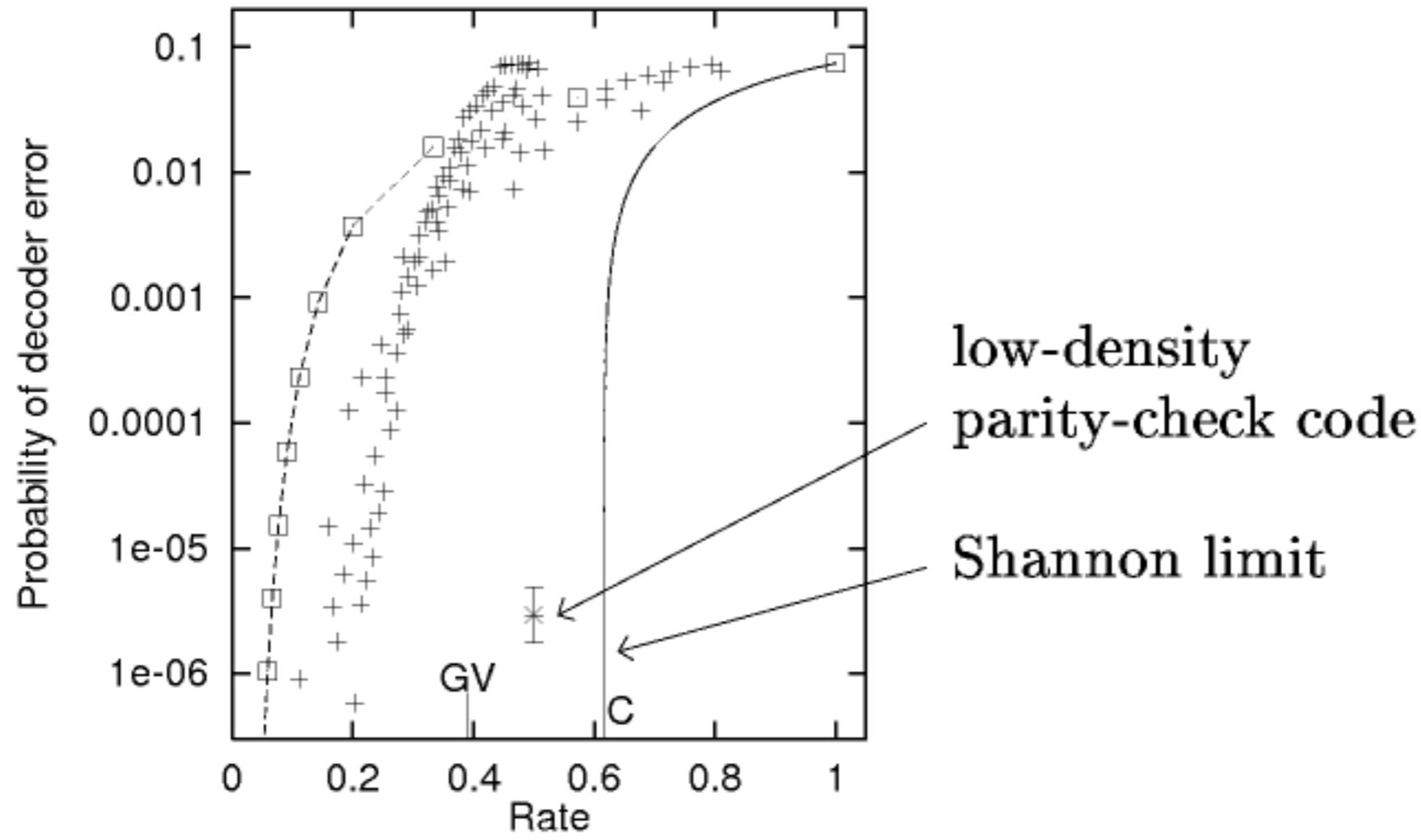


Low Density Parity Check Code

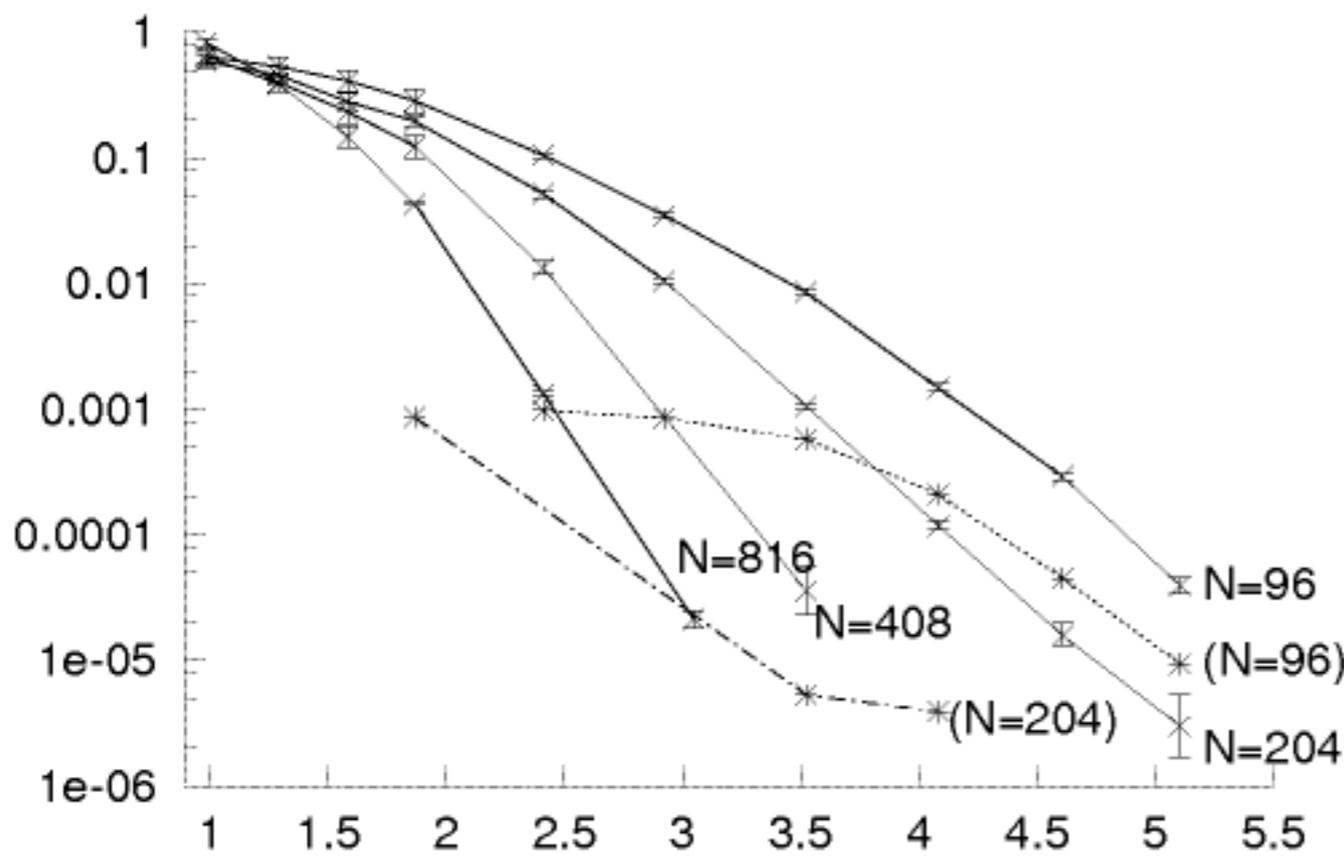
Iterative probabilistic decoding



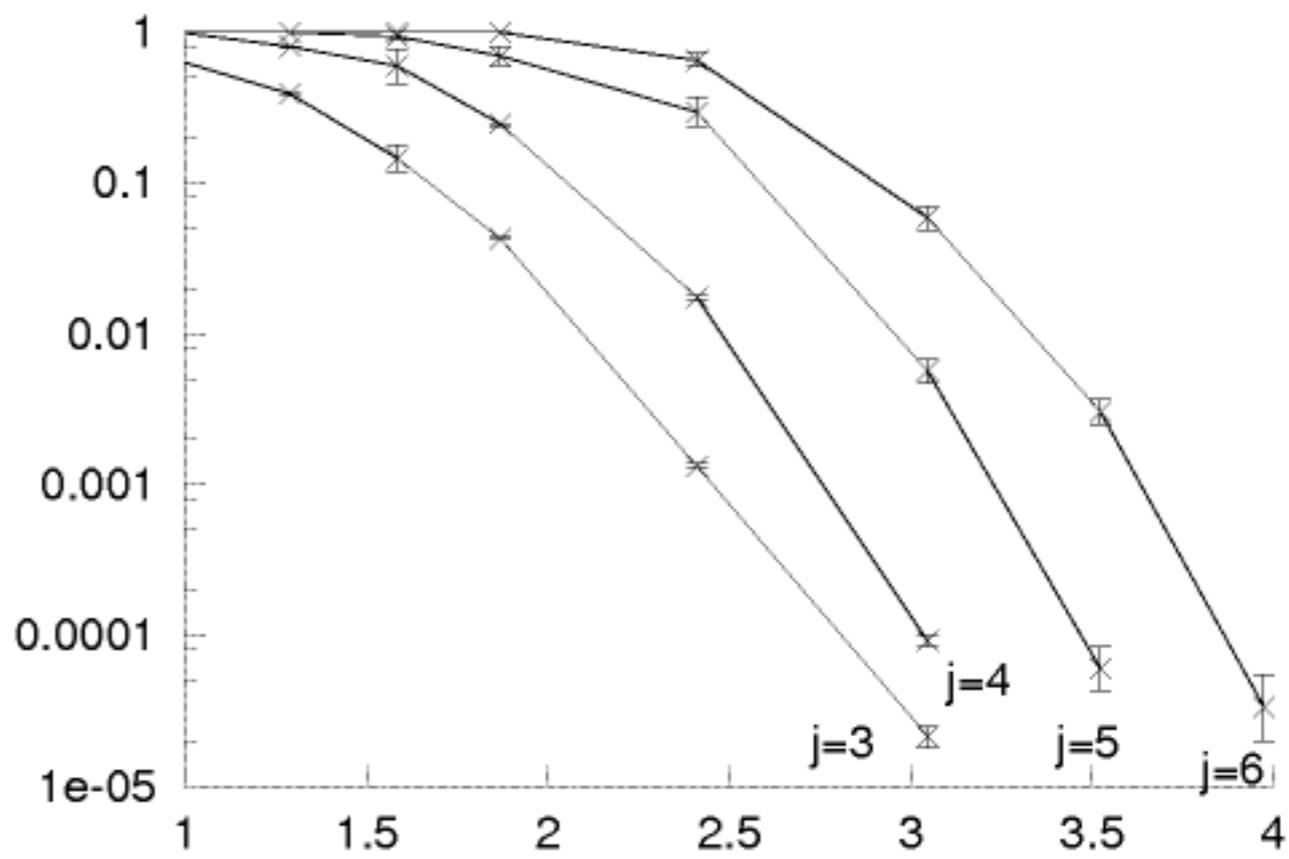
BSC: $f=7.5\%$



Dependence on blocklength and column weight



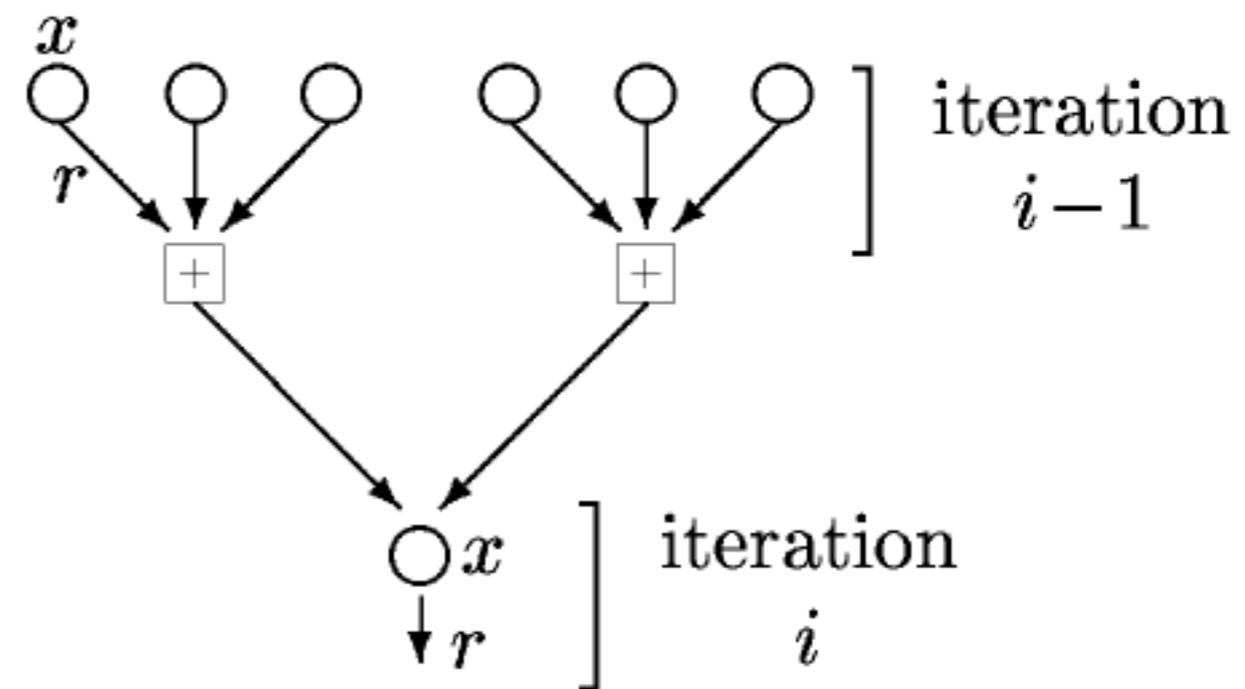
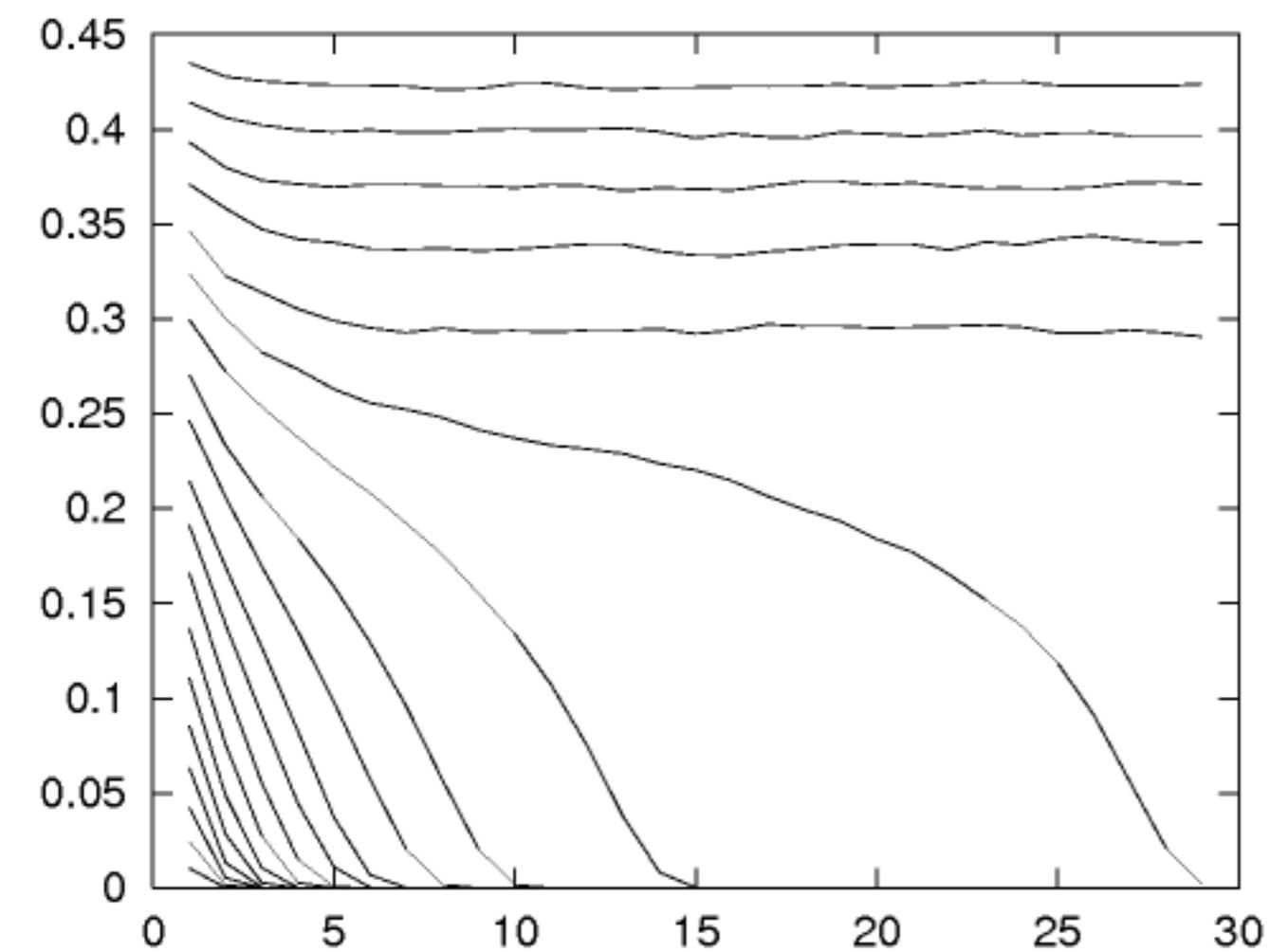
(a)



(b)

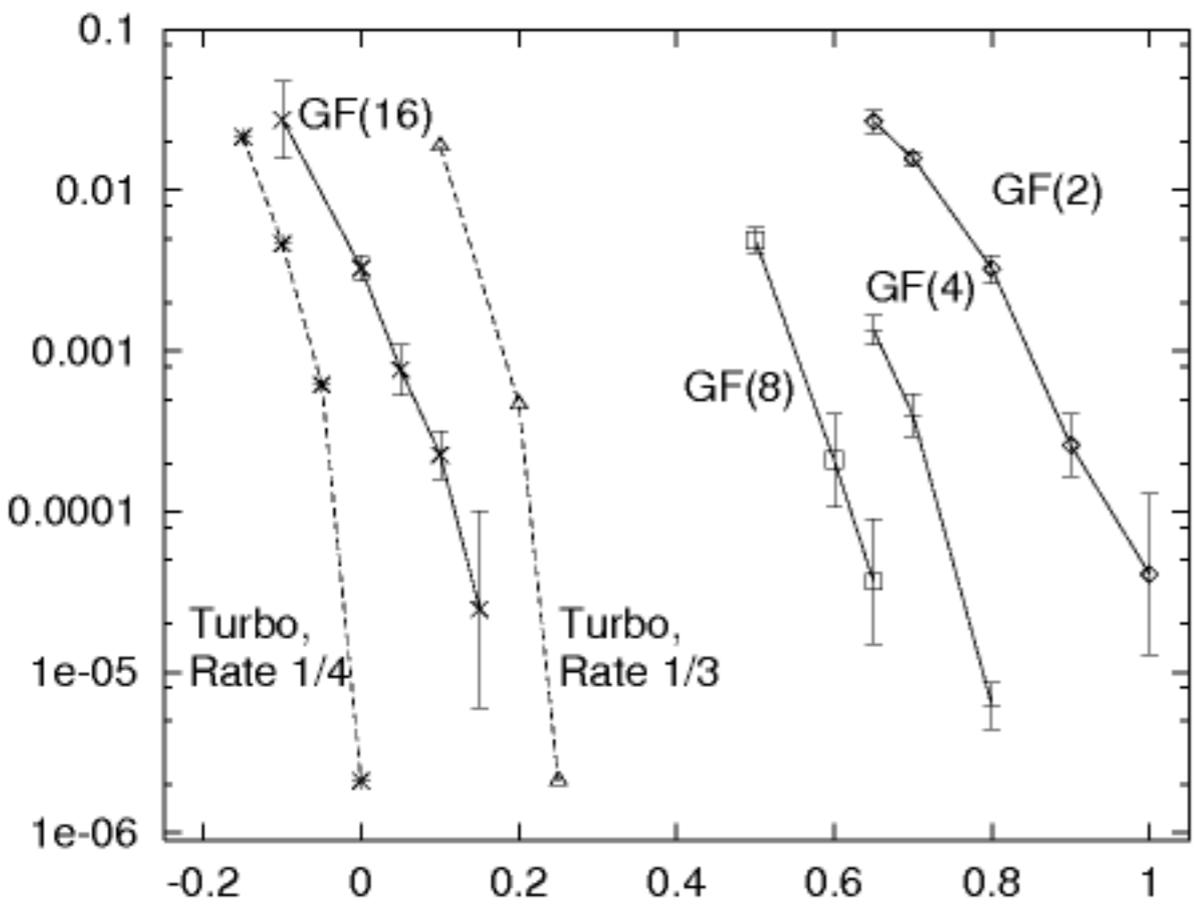
'Density evolution' on infinite graphs

Entropy versus iteration number

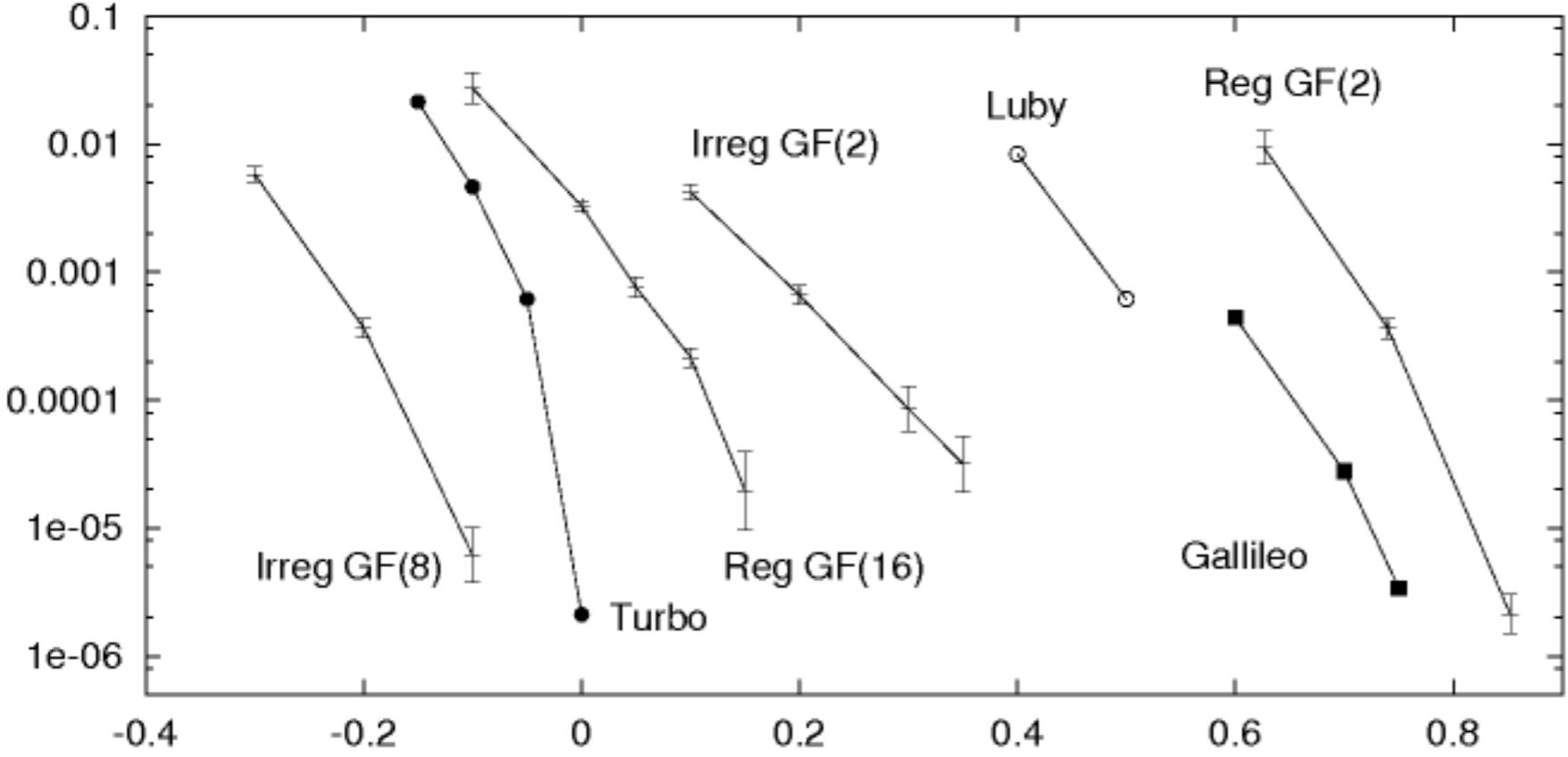


Identifies the **threshold** for sum-product decoding

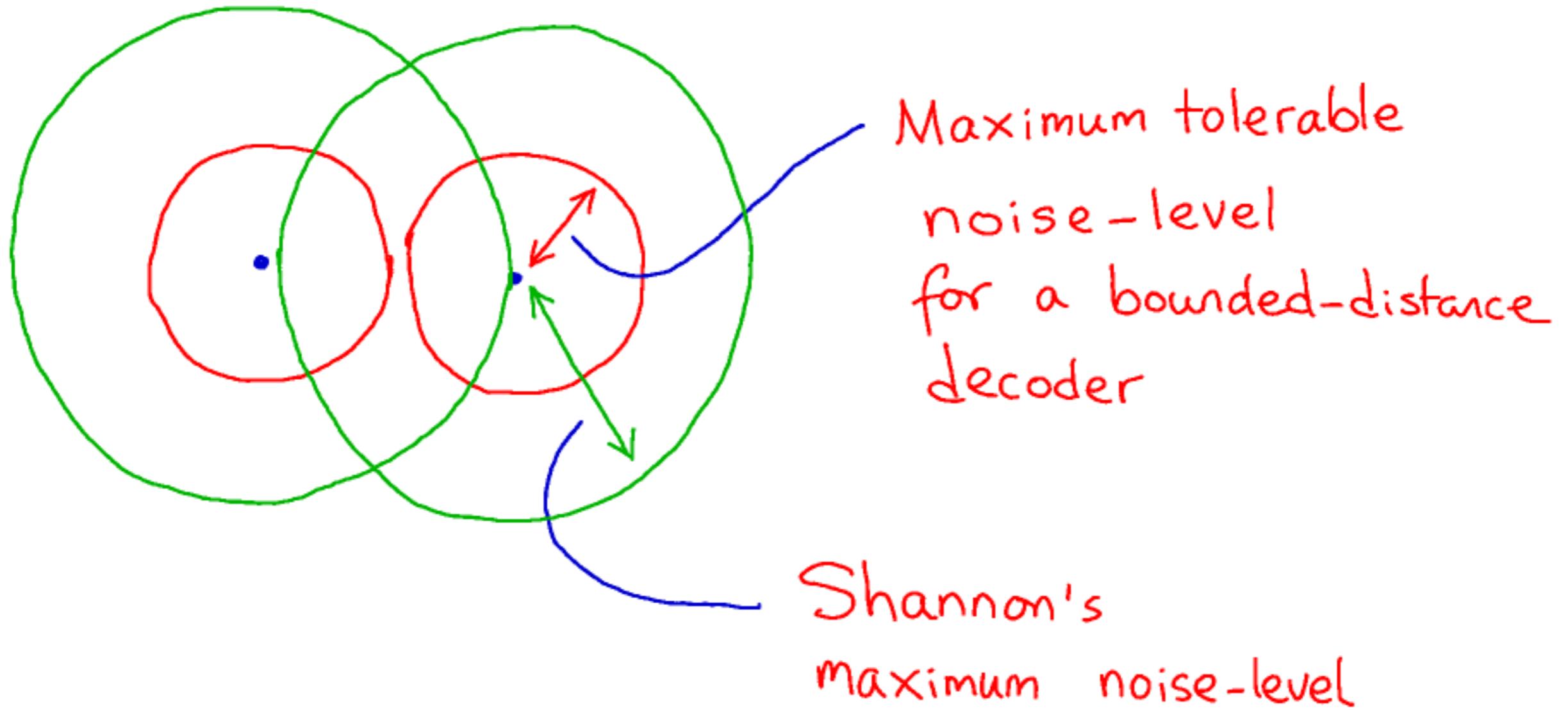
Beyond simple LDPC codes - Gaussian channel results



Beyond simple LDPC codes - Gaussian channel results

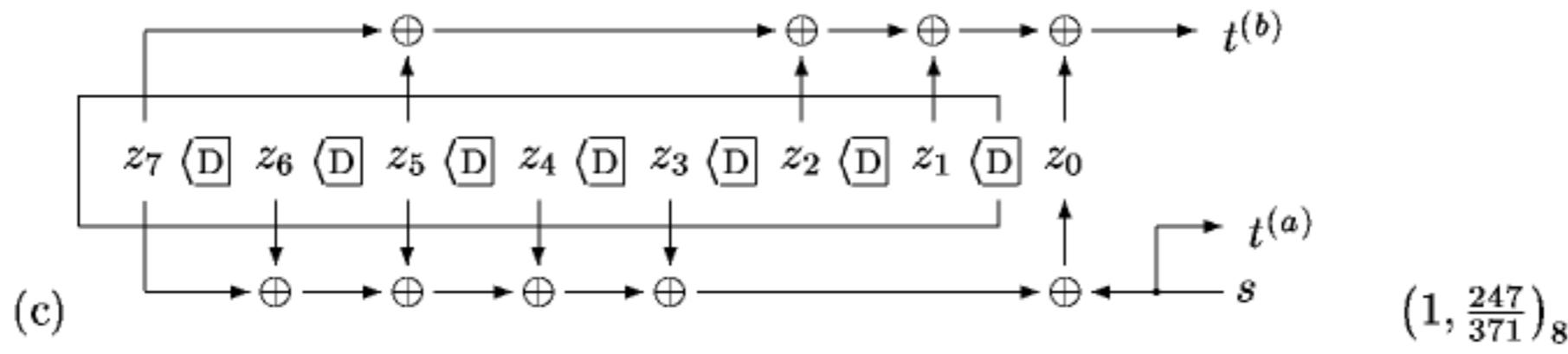
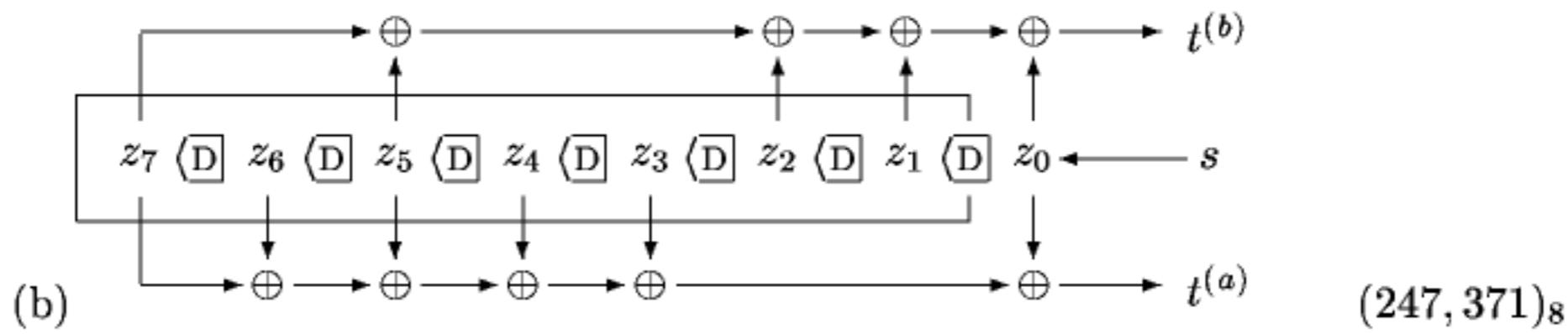
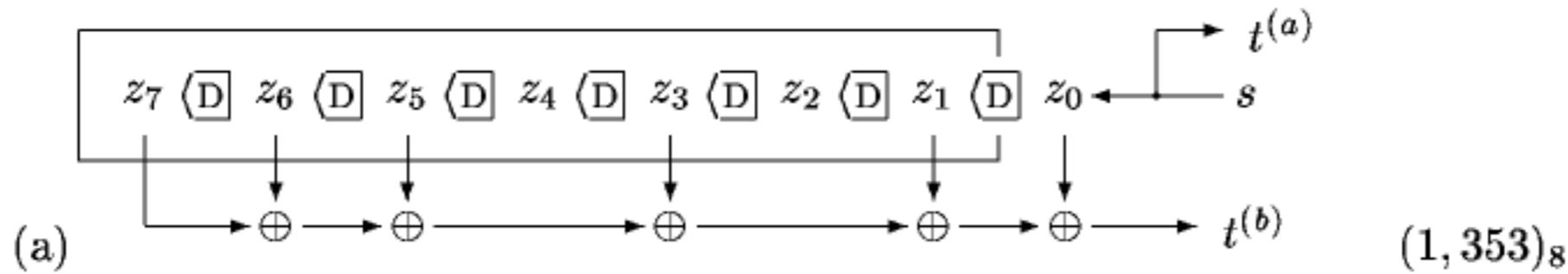


Two codewords

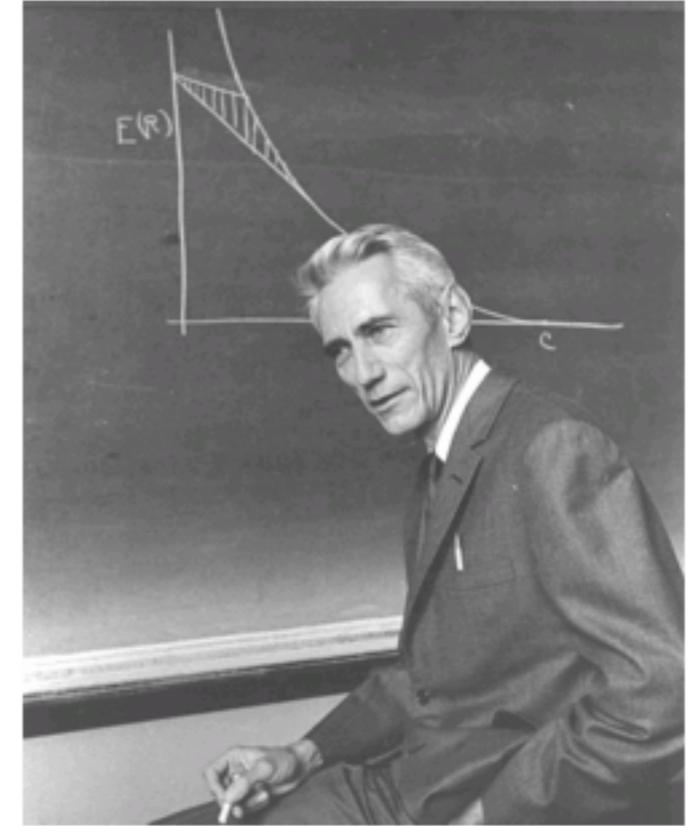


Convolutional codes

Octal name



Feedback



`Feedback? Pah! Who needs feedback?
Just use a random code!'

Sphere packing view

Count inputs and outputs \rightarrow get a bound on what's achievable.

Given a transmission of length N ,
the output will probably have Nf bits flipped,
so it will be in a typical set of size

$$\binom{N}{Nf} \simeq 2^{NH_2(f)}$$

So if we have 2^K alternative inputs, and almost all these typical outputs are distinct, we must have

TOTAL NUMBER OF TYPICAL OUTPUTS

$$\overbrace{2^K \times 2^{NH_2(f)}}$$

\leq

TOTAL SIZE OF OUTPUT SPACE

$$\overbrace{2^N}$$

i.e.,

$$K + NH_2(f) \leq N$$

i.e.,

$$\frac{K}{N} \leq 1 - H_2(f)$$