1999 IMA Summer Program:
Codes, Systems and Graphical Models

yong

Schedule // Participants // Bibliographic Items Related to Week 1 // Bibliographic Items Related to Week 2 // Material from Talks

Partially supported by the National Security Agency

August 2-13, 1999

Organizers:

G. David Forney, Jr.
Massachusetts Institute of Technology
LUSE27@email.mot.com
forney@lids.mit.edu

Brian Marcus
IBM Almaden Research Center
marcus@almaden.ibm.com

Joachim Rosenthal
University of Notre Dame
rosen@nd.edu

Alexander Vardy
University of California, San Diego
vardy@ece.ucsd.edu

Note: The registration for this summer workshop has been closed due to an overwhelming response.

The invention of turbo codes and other capacity-approaching codes has led to an exciting cross-fertilization of ideas between researchers from different backgrounds.

The aim of the workshop is to bring together mathematicians, computer scientists, and electrical engineers in the area of coding theory, systems theory and symbolic dynamics so that the techniques from one area can be applied to problems in the other area. The two weeks of the workshop will be subdivided into two main focus areas:

Week 1:
Codes on Graphs and Iterative Decoding
Week 2:
Connections Among Coding Theory, System Theory and Symbolic Dynamics

Week 1

CODES ON GRAPHS AND ITERATIVE DECODING

Belief propagation in Bayesian networks has been extensively studied in artificial intelligence since the work of Pearl a decade ago, and turbo codes have recently become a subject of much research in coding theory. In the past year or two it has been recognized that the iterative decoding algorithm used for turbo codes and other capacity-approaching schemes are instances of belief propagation. This has led to an explosion of work devoted to understanding and exploiting this connection. A related problem is that of representing a given code by a graph, such as a Bayesian network. A central impetus of much of this work is to understand why iterative algorithms work so well empirically on graphs with cycles, where practically no theoretical results are known. Experts in the dynamics of algorithms have also begun to be drawn into this work. The major focus of week 1 of the IMA workshop will be to bring together researchers in these various fields to better understand these emerging connections. This will be a natural follow-on to a special session on this subject at the upcoming 1998 MTNS conference (Mathematical Theory of Networks and Systems, among the most mathematical of the systems theory conferences).

Topics for week 1 include: Codes defined on graphs, iterative decoding algorithms, factor graphs, turbo codes, connections with Bayesian networks.

Week 2

CONNECTIONS AMONG CODING THEORY, SYSTEM THEORY AND SYMBOLIC DYNAMICS

Coding Theory, System Theory and Symbolic Dynamics have much in common as evidenced by the following list of research topics that play a prominent role in each:

1. Construction of various types of finite- and finite-dimensional state representations of sequence spaces.
2. Investigation of fundamental structural properties of sequence spaces, such as observability and controllability.
3. Construction of input/output systems, i.e. mappings (or encoders) between sequence spaces.
4. Understanding the special role that algebraic structure (in particular, linearity and duality) plays in 1,2 and 3.

Yet these subjects have developed somewhat independently, and each has its own language and points of view. Until recently there has been very little communication among researchers in these subjects. A main purpose of week 2 of the IMA workshop is to further the communication among researchers and stimulate connections among these subjects. Week 2 will aim to continue a successful series of interdisciplinary meetings that has included an IEEE Information Theory Workshop on Coding, Systems and Symbolic Dynamics in 1993 (Mansfield, MA), a special invited session at the IEEE Conference on Decision and Control in 1995 (New Orleans), and two special sessions at the MTNS in 1998 (Padova).

Topics for week 2 include: Behavioral system theory, input/output mappings between spaces of sequences, state space representations, group codes, trellis codes, multi-dimensional systems and codes.

The organizers plan a number of invited tutorial lectures specifically for interspecialty communication. Leading workers in each field will also be invited to present surveys of current research, with less emphasis on solved problems than on open ones. Finally, there will be both invited and contributed papers presenting recent research results.

We expect the attendees to represent electrical engineering, mathematics and computer science departments in both academia and industry. As coding theory is the glue that holds the two weeks together, we expect that it will mostly be a subset of the coding theory participants who will attend both weeks.

WORKSHOP SCHEDULE

Week 1: August 2-6, 1999	Monday	Tuesday	Wednesday	Thursday	Friday
Week 2: August 9-13, 1999	Monday	Tuesday	Wednesday	Thursday	Friday

All talks are in Lecture Hall EE/CS 3-180 unless otherwise noted.

WEEK 1: CODES ON GRAPHS AND ITERATIVE DECODING
August 2-6, 1999

SCHEDULE for MONDAY, AUGUST 2

HISTORY AND TUTORIALS Day
G. David Forney, Jr. (chair)

8:30 am	Registration and Coffee	Reception Room EE/CS 3-176
9:10 am	**Willard Miller, Fred Dulles, and G. David Forney**	Introduction and Welcome
9:30 - 10:30 am	**R. Michael Tanner** University of California-Santa Cruz	Error-Correcting Codes and Graph-based Algorithms: Origins, Successes, the Current Quests
10:30 am	Coffee Break	Reception Room EE/CS 3-176
11:00 am - 12:00 pm	**Stephen B. Wicker** Cornell University	Markov Chains, Error Control, and the Advent of Turbo Coding
12:00 pm	Lunch	
2:00-3:00 pm	**Frank R. Kschischang** University of Toronto	Factor Graphs and the Sum-Product Algorithm
4:00 pm	IMA Tea	IMA East, 400 Lind Hall A variety of appetizers and beverages will be served.

SCHEDULE for TUESDAY, AUGUST 3

LOW DENSITY PARITY CHECK CODES DAY
R. Michael Tanner (chair)

9:15 am	Coffee	Reception Room EE/CS 3-176

9:30-10:30 am	**David J.C. MacKay** Cambridge University	Sparse Graph Codes
10:30 am	Coffee Break	Reception Room EE/CS 3-176
11:00 am - 12:00 pm	**Robert J. McEliece** California Institute of Technology	Some Simple Codes that Are Good in Both Theory and Practice
12:00 pm	Lunch	
2:00 - 3:00 pm	**Thomas J. Richardson** (Lucent Bell Labs) **Ruediger Urbanke** (Lucent Bell Labs)	Analysis and Design of Iterative Decoding Systems
3:00 pm	Coffee Break	Reception Room EE/CS 3-176

Contributed Talks and Informal Discussions

3:30 pm	**Amin Shokrollahi** Bell Labs	Capacity Achieving Low-density Erasure Codes
4:00 pm	**Gilles Zemor** ENST, Paris	Iterative Decoding of Cycle Codes of Graphs
4:30 pm	**Dakshi Agrawal** University of Illinois-Urbana Champaign	On the Phase Trajectories of the Turbo Decoding Algorithm

SCHEDULE for WEDNESDAY, AUGUST 4

INFERENCE DAY
Brendan J. Frey (chair)

9:15 am	Coffee	Reception Room EE/CS 3-176
9:30 - 10:30 am	**Tommi Jaakkola** Massachusetts Institute of Technology	Variational Methods for Inference
10:30 am	Coffee Break	Reception Room EE/CS 3-176
11:00 am - 12:00 pm	**Radford M. Neal** University of Toronto	Sparse Matrix Methods and Probabilistic Inference Algorithms
12:00 pm	Lunch	
2:00 - 3:00 pm	**Brendan J. Frey** University of Waterloo **Yair Weiss** University of California at Berkeley	The Sum-Product Algorithm in Gaussian Networks with Cycles
3:00 am	Coffee Break	Reception Room EE/CS 3-176

Contributed Talks and Informal Discussions

3:30 pm	**John B. Anderson** University of Lund	Properties of the Tailbiting BCJR Decoder
4:00 pm	**Amir Banihashemi** Carleton University	Tanner Graphs for Group Block Codes and Lattices: Construction and Complexity
4:30 pm	**Heeralal Janwa and Oscar Moreno** University of Puerto Rico	New Constructions of Ramanujan Graphs and Good Expander Graphs from Codes, Exponential Sums and Sequences

SCHEDULE for THURSDAY, AUGUST 5

Robert J. McEliece (chair)

9:15 am	Coffee	Reception Room EE/CS 3-176
9:30 - 10:30 am	**Randall E. Bryant** Carnegie Mellon University	Symbolic Boolean Manipulation with Ordered Binary Decision Diagrams
10:30 am	Coffee Break	Reception Room EE/CS 3-176
11:00 am - 12:00 pm	**John Lafferty** Carnegie Mellon University	Trellises, Decision Diagrams, and Factor Graphs
12:00 pm	Lunch	
2:00 - 3:00 pm	**James L. Massey** ETH Zurich and Lund University	Linear Systems over Fields and Rings, Linear Complexity, and Fourier Transforms
3:00 am	Coffee Break	Reception Room EE/CS 3-176
6:00 pm	**Workshop Dinner**	**Bona Vietnamese Restaurant** Located near the IMA and the Day's Inn at 802 Washington Avenue, the south side of Washington very near the

intersection of Washington
and Oak St.
Phone: 612-331-5011

SCHEDULE for FRIDAY, AUGUST 6

CODING THEORY DAY Alexander Vardy (chair)

8:45 am	Coffee	Reception Room EE/CS 3-176
9:00 - 10:00 am	**G. David Forney, Jr.** Massachusetts Institute of Technology	Codes and Systems on Graphs: Generalized State Realizations
10:00 am	Coffee Break	Reception Room EE/CS 3-176
10:15 - 11:15 am	**Ralf Koetter** University of Illinois at Urbana-Champaign	Factor Graphs, Trellis Formations, and Generalized State Realizations
11:15 am	Coffee Break	Reception Room EE/CS 3-176
11:30 am	**Hans-Andrea Loeliger** Endora Tech AG, Switzerland	Decoding and Equalization: Iterative Algorithms and Analog Networks

Week 1: August 2-6, 1999 Monday Tuesday Wednesday Thursday Friday
Week 2: August 9-13, 1999 Monday Tuesday Wednesday Thursday Friday

All talks are in Lecture Hall EE/CS 3-180 unless otherwise noted.

WEEK 2: CONNECTIONS AMONG CODING THEORY, SYSTEM THEORY AND SYMBOLIC DYNAMICS
August 9-13, 1999

SCHEDULE for MONDAY, AUGUST 9

8:30 am	Registration and Coffee	Reception Room EE/CS 3-176
9:10 am	**Willard Miller, Fred Dulles, Joachim Rosenthal, and Brian Marcus**	Introduction and Welcome

Automata and Systems
Jorn Justesen (Chair)

9:30 am	**Roger W. Brockett** Harvard University	Dynamical Systems and their Associated Automata
10:30 am	Coffee Break	Reception Room EE/CS 3-176
11:00 am - 12:00 pm	**Dominique Perrin** Université de Marne-la-Vallée	Symbolic Dynamics and Automata

Algebra and Geometry Applied to Systems
Ethan Coven (Chair)

1:30 pm	**Paul A. Fuhrmann** Ben Gurion University	A Polynomial Module Approach to Linear Systems Theory
2:30 pm	**Clyde Martin** Texas Tech University	Linear Systems as Vector Bundles on Spheres
3:30 pm	Coffee Break	Reception Room EE/CS 3-176
4:00 pm	**M.S. Ravi** Eastern Carolina University	An Algebraic Geometric Point of View to Linear Systems Theory
5:00 pm	IMA Tea	IMA East, 400 Lind Hall A variety of appetizers and beverages will be served.

SCHEDULE for TUESDAY, AUGUST 10

8:45 am	Coffee	Reception Room EE/CS 3-176

Convolutional Codes
Karl Petersen (Chair)

9:00 am	**Rolf Johannesson** University of Lund	Woven Convolutional Codes: Encoder Properties and Error Exponents
10:00 am	**Roxana Smarandache** University of Notre Dame	Construction of Convolutional Codes with Large Free Distance
11:00 am	Coffee Break	Reception Room EE/CS 3-176
11:30 am	**Fabio Fagnani** Politecnico di Torino Joint talk with **Sandro Zampieri** Universita di Padova	On Convolutional Codes over Rings

Contributed Talks
Joachim Rosenthal (Chair)

All talks will be 25 minutes long, including questions.

2:00 pm	**Thomas Mittelholzer** IBM Zurich Research Laboratory	Duals over Artinian Rings and the MacWilliams Identity
2:30 pm	**Sergio R. Lopez-Permouth** Ohio University	Finite Fields, Permutations and Trellis
3:00 pm	Coffee Break	Reception Room EE/CS 3-176
3:30 pm	**Danrun Huang** St. Cloud State	Period Three, Chaos, and the Golden Mean Shift
4:00 pm	**Dharmendra S. Modha** IBM Almaden Research Center	Art of Constructing Low-complexity Encoders/Decoders for Constrained Block Codes
4:30 pm	**Natasha Jonoska** University of South Florida	On Encoding in DNA Words

SCHEDULE for WEDNESDAY, AUGUST 11

8:45 am	Coffee	Reception Room EE/CS 3-176

Multidimensional Systems
Jon Hall (Chair)

9:00 am	**Klaus Schmidt** University of Vienna	Multi-dimensional Symbolic Dynamical Systems
10:00 am	**Paul H. Siegel** University of California-San Diego	Capacity of Constrained Systems in One and Two Dimensions
11:00 am	Coffee Break	Reception Room EE/CS 3-176
11:30 am	**Paul A. Weiner** Saint Mary's University of Minnesota	Multidimensional Convolutional Codes

Systems Theory
Roy Adler (Chair)

2:00 pm	**Jan C. Willems** University of Groningen	Systems, States and their Representations
3:00 pm	Coffee Break	Reception Room EE/CS 3-176
3:30 pm	**Sanjoy Mitter** MIT	Path Space View of Probabilistic Systems

SCHEDULE for THURSDAY, AUGUST 12

8:45 am	Coffee	Reception Room EE/CS 3-176

Symbolic Dynamics and Applications
Uwe Helmke (Chair)

9:00 am	**M. Michael Boyle** University of Maryland	Applications of Symbolic Dynamics to the Structure Theory of Nonnegative Matrices
10:00 am	**Natasha Jonoska** University of South Florida	Multiplicities of SFT Covers
11:30 am	**Selim Tuncel** University of Washington	Codings of Markov Chains and Weighted Graphs

Contributed Talks
Brian Marcus (Chair)

2:00 pm	**Marie-Pierre Béal** Université de Marne-la-Vallée	A Finite State Version of the Kraft-McMillan Theorem
2:30 pm	**Olivier Carton** Université de Marne-la-Vallée	Asynchronous Sliding Block Maps
3:00 pm	Coffee Break	Reception Room EE/CS 3-176
3:30 pm	**Christiane Frougny** LIAFA	Deterministic Synchronization of Bounded Delay 2-tape Finite Automata
4:00-4:30 pm	**Michael E. O'Sullivan** University College Cork	The Key Equation for One-point Codes
4:30 - 5:00 pm	**Fernando Guzmán** Binghamton University	Ambiguity in Codes
6:00 pm	**Workshop Dinner**	**Campus club** Located on the 4th floor of Coffman Student Union and serves a wide-ranging buffet. Coffman Union is located on the opposite side of Washington Avenue from the IMA and slightly to the west.

SCHEDULE for FRIDAY, AUGUST 13

8:45 am	Coffee	Reception Room EE/CS 3-176

Decoding and Interpolation
Zhe-Xian Wan (Chair)

9:00 am	**Margreet Kuijper** University of Melbourne	Algorithms for Decoding and Interpolation
10:00 am	**Patrick Fitzpatrick** National University of Ireland, Cork	Realization and Interpolation via Gröbner Bases
11:00 am	Coffee Break	Reception Room EE/CS 3-176
11:30 am	**Brian M. Allen** University of Notre Dame	Linear Systems Decoding of Convolutional Codes

Informal Contributed Talks

Lind Hall 409 with an option to switch to Lecture Hall EE/CS 3-180 contingent on the participants size

2:00 pm	**Karl Petersen** University of North Carolina	Good Measures for Bad Codes between SFT's
2:30 pm	**Ethan Coven** Wesleyan University	The Symbolic Dynamics of Tiling the Integers
3:00 pm	Coffee Break	IMA East Lind Hall room 400
3:15 pm	**Kimberly Johnson** University of North Carolina	Automata, and Pumping Lemmas for Beta-shifts
3:45 pm	**Paul Trow** University of Memphis	Mappings between Group Shifts

Week 1: August 2-6, 1999 Monday Tuesday Wednesday Thursday Friday
Week 2: August 9-13, 1999 Monday Tuesday Wednesday Thursday Friday

Back to top of page

IMA

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

Gallager Codes - Recent Results
August 3, 1999

A talk that was presented by David J.C. MacKay, an IMA Visitor from the Cambridge University

Gallager codes – recent results

David MacKay
Dept. of Physics
Univ. of Cambridge

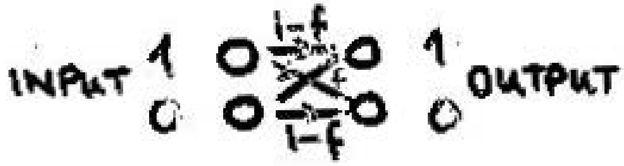
with
Radford Neal
Matthew Davey

- How good regular Gallager codes are.
- Enhancing Gallager codes.
- Open questions.
- Gallager codes for high rate, short block applications.

INFORMATION THEORY + CODING THEORY

<u>Aim:</u> virtually error-free communication over noisy channels

eg: binary symmetric channel



eg $f = 0.1$

<u>Method:</u>

Adds redundancy

Adds noise **n**

Good news: Can reduce error <u>probability</u>

Bad news: Redundancy means <u>rate</u> R is reduced

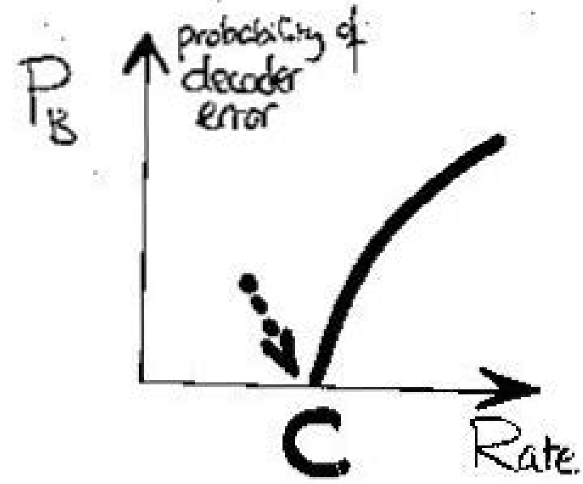
Code families

VERY GOOD CODES

can achieve

$P_B \to 0$

as $R \to C$

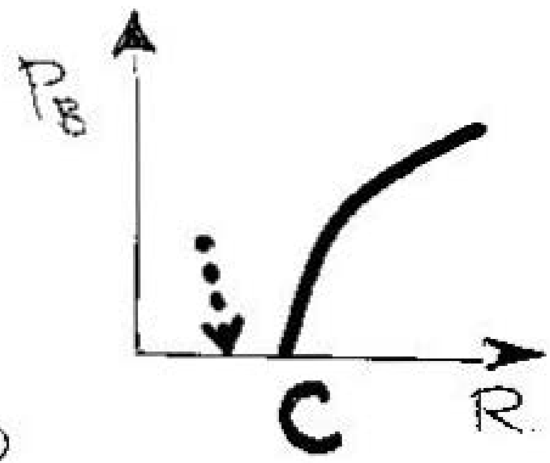


GOOD CODES

can achieve

$P_B \to 0$

at some $R > 0$

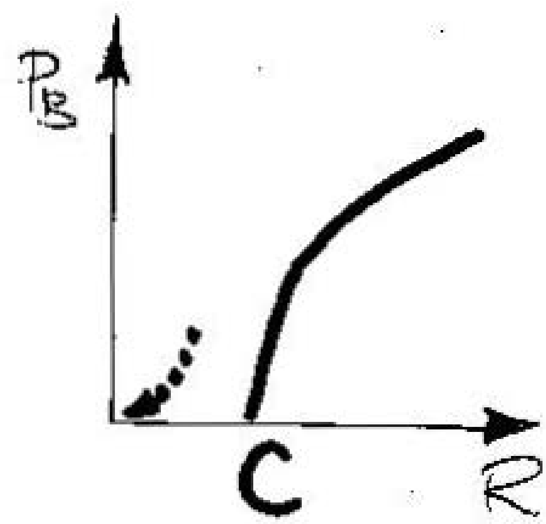


BAD CODES

only achieve

$P_B \to 0$

as $R \to 0$



Codes have <u>Good distance</u>

if $\frac{d}{N} \to$ const. as $N \to \infty$

Codes have <u>Very Good distance</u>

if $\frac{d}{N} \to \frac{d_{GV}}{N}$

$H_2\left(\frac{d_{GV}}{N}\right) = 1 - R$

Codes have <u>Bad distance</u>

if $\frac{d}{N} \to 0$

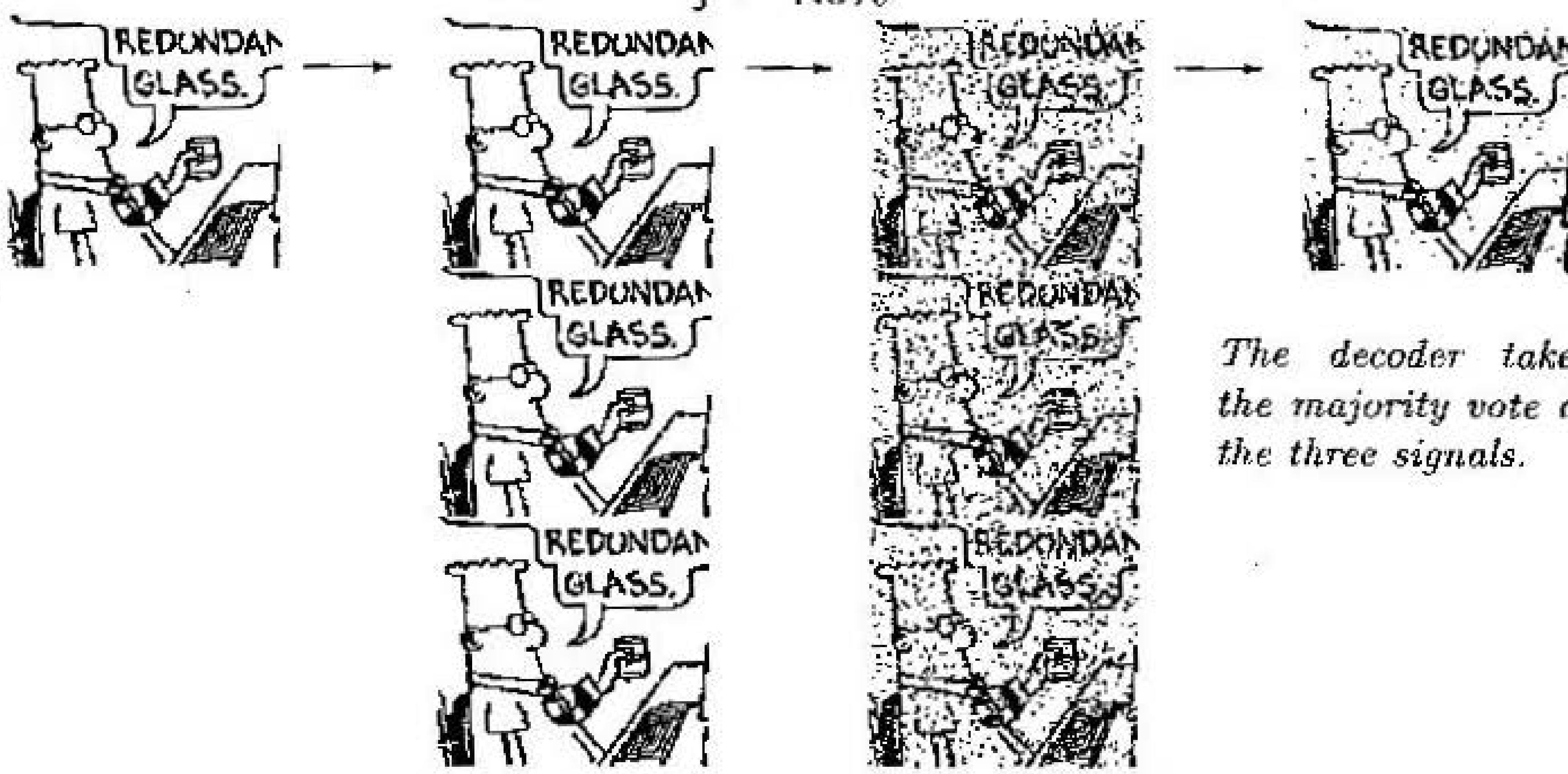
& <u>very bad distance</u>

if $d =$ const.

Repetition Code "R3"

s ENCODER t CHANNEL r DECODER $\hat{\mathbf{s}}$

$f = 7.5\%$



The decoder takes the majority vote of the three signals.

Good news: only 1.6% of decoded bits are in error
Bad news: rate of communication reduced to 1/3

Hamming (7,4) code

Encodes **K = 4** source bits
into **N = 7** transmitted bits.

PARITY CHECK MATRIX

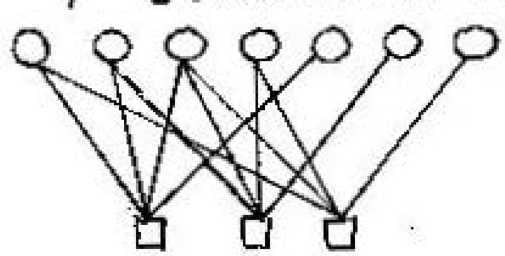
$$\underline{\underline{H}} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \quad \updownarrow M=3 \text{ constraints}$$

← K = 4 → source bits

M = 3 parity bits

GRAPH

7 transmitted bits



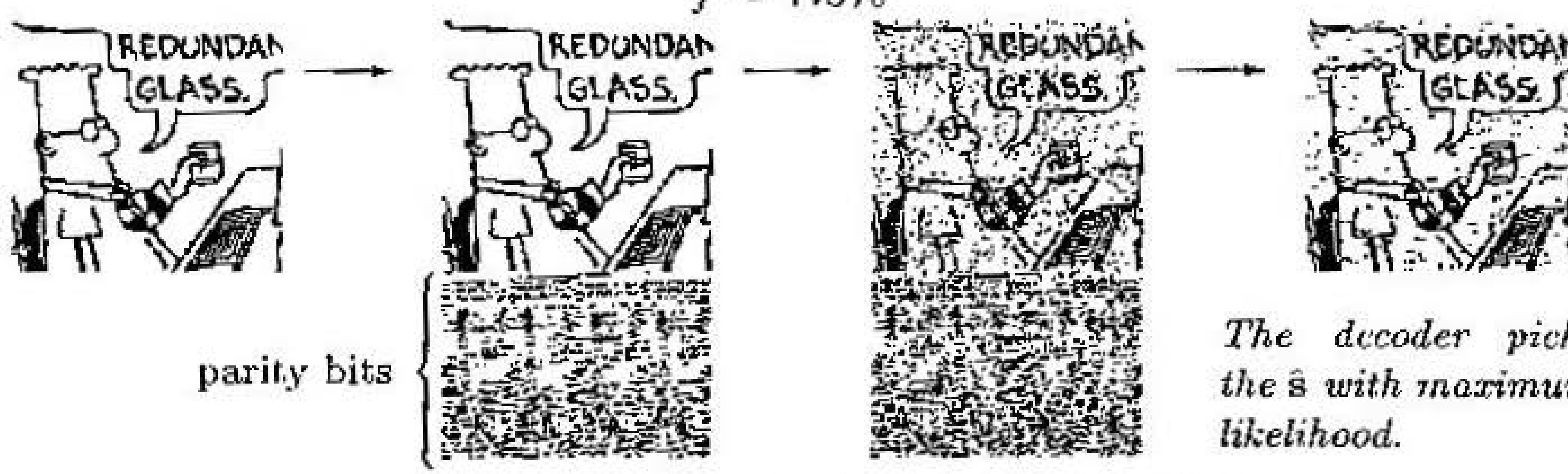
3 constraints

THE Hamming (7,4) code in actio

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

s ENCODER t CHANNEL r DECODER $\hat{\mathbf{s}}$

$f = 7.5\%$



The decoder picks the $\hat{\mathbf{s}}$ *with maximum likelihood.*

4% of decoded bits are in error

rate of communication is 4/7

Decoding the Hamming code

received = transmitted + noise

$$\underline{r} = \underline{t} + \underline{n} \quad \text{mod } 2.$$

$$\underline{\underline{H}}\,\underline{r} = \underbrace{\underline{\underline{H}}\,\underline{t}} + \underline{\underline{H}}\,\underline{n} \quad \text{mod } 2.$$

$\underline{\underline{H}}\,\underline{t} = 0$ for all valid transmissions

The "syndrome" of the received signal,

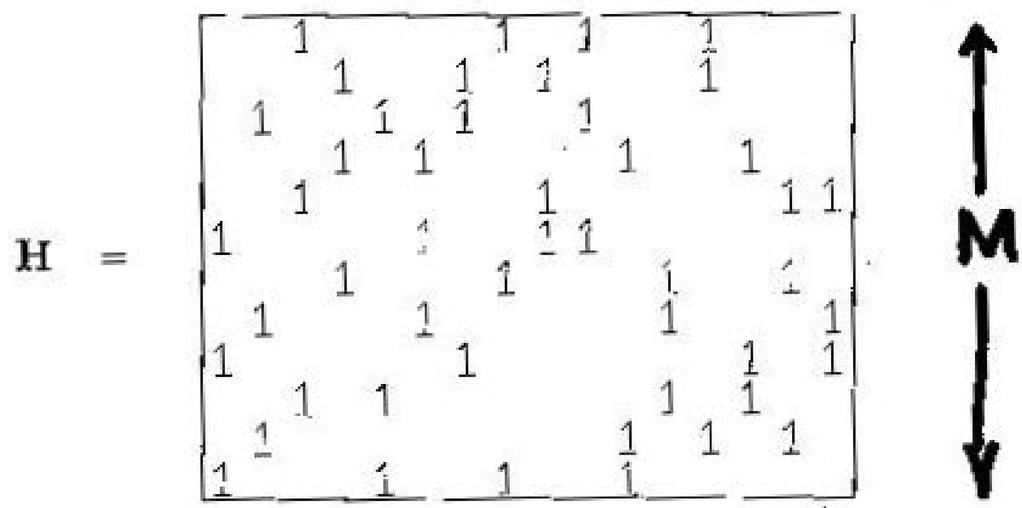
$$\underline{z} \equiv \underline{\underline{H}}\,\underline{r} \quad \textbf{mod 2}$$

$\Rightarrow$ want to solve for the sparsest $\underline{n}$ such that $\underline{\underline{H}}\,\underline{n} = \underline{z}$ mod 2

Low Density Parity Check Codes

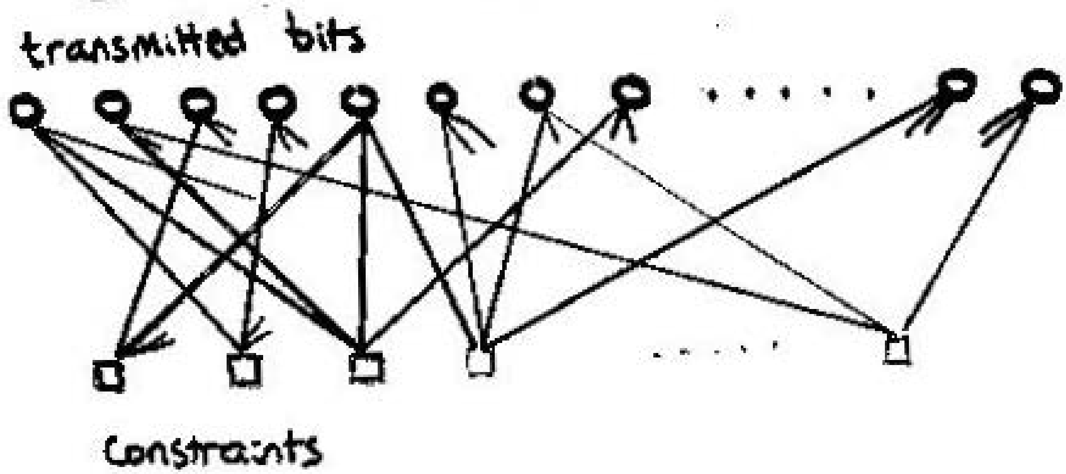
Gallager 1962

- Parity check matrix $\underline{\underline{H}}$ is very sparse



eg, every column has weight 3

- The graph is very sparse



THEORY

Gallager codes with fixed $t \geq 3$ are GOOD, and have GOOD DISTANCE.

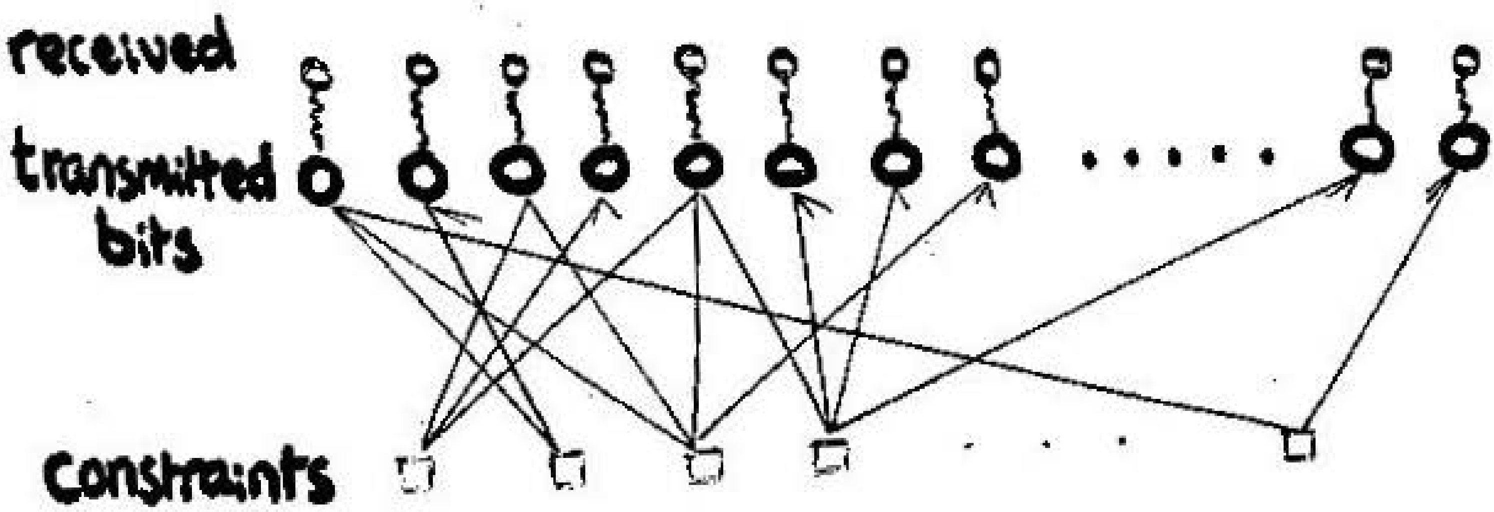
Gallager 1962 15
MacKay 1999 27

As t increases*, Gallager codes become VERY GOOD*, and have VERY GOOD DISTANCE.

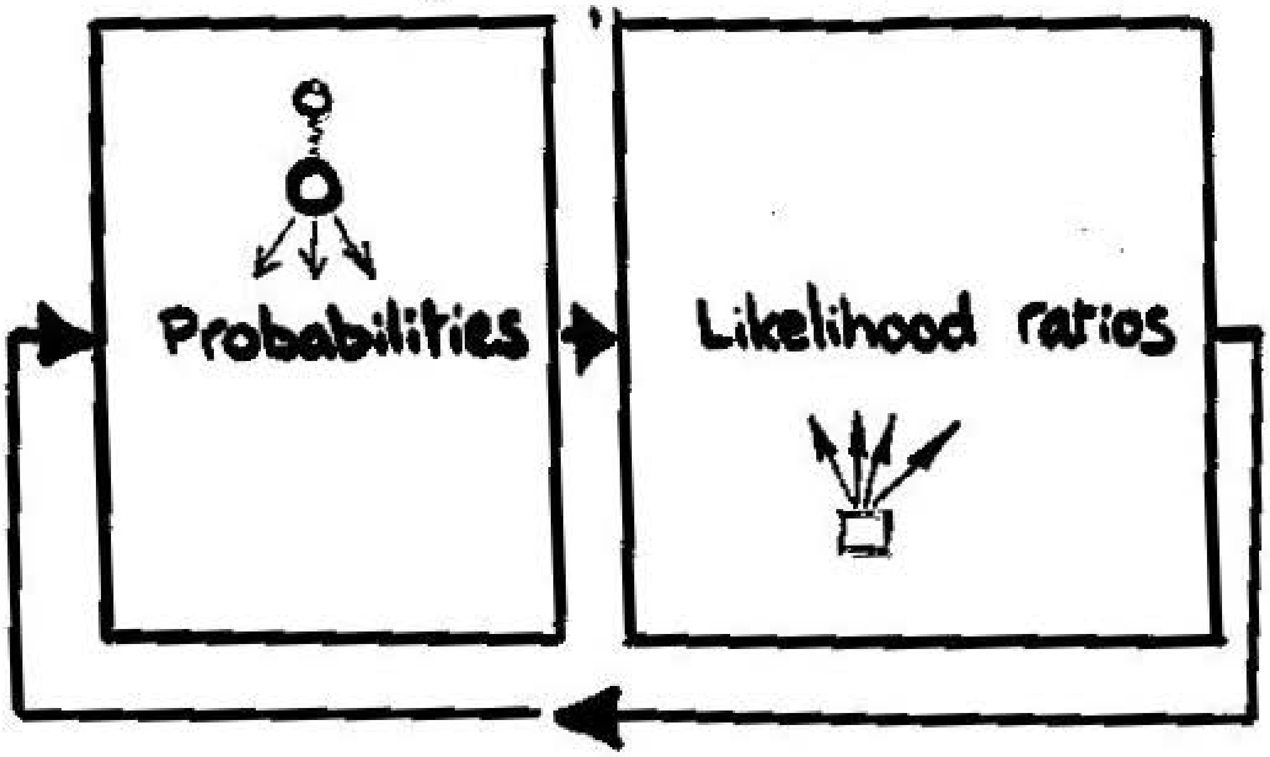
* $\left(\frac{t}{M} \text{ still vanishing}\right)$

* for a wide range of channels with and without memory. MacKay 1999

How to Solve the Decoding Problem (DIFFICULT)



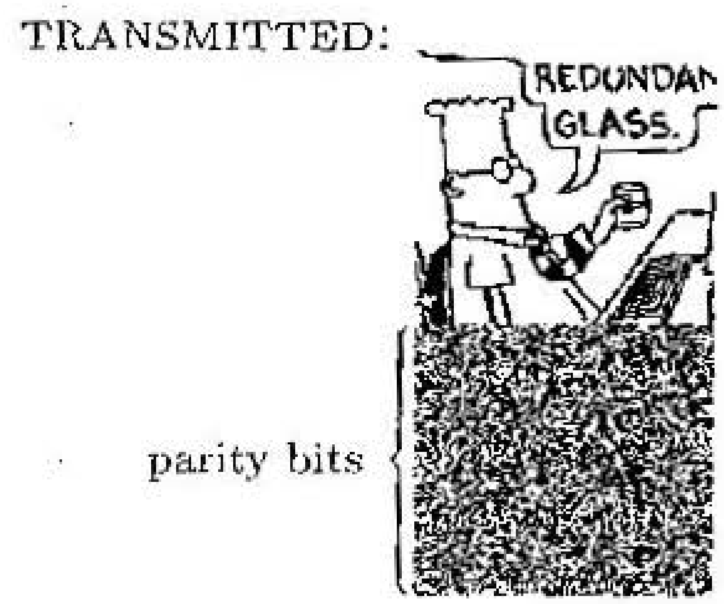
- sum-product algorithm



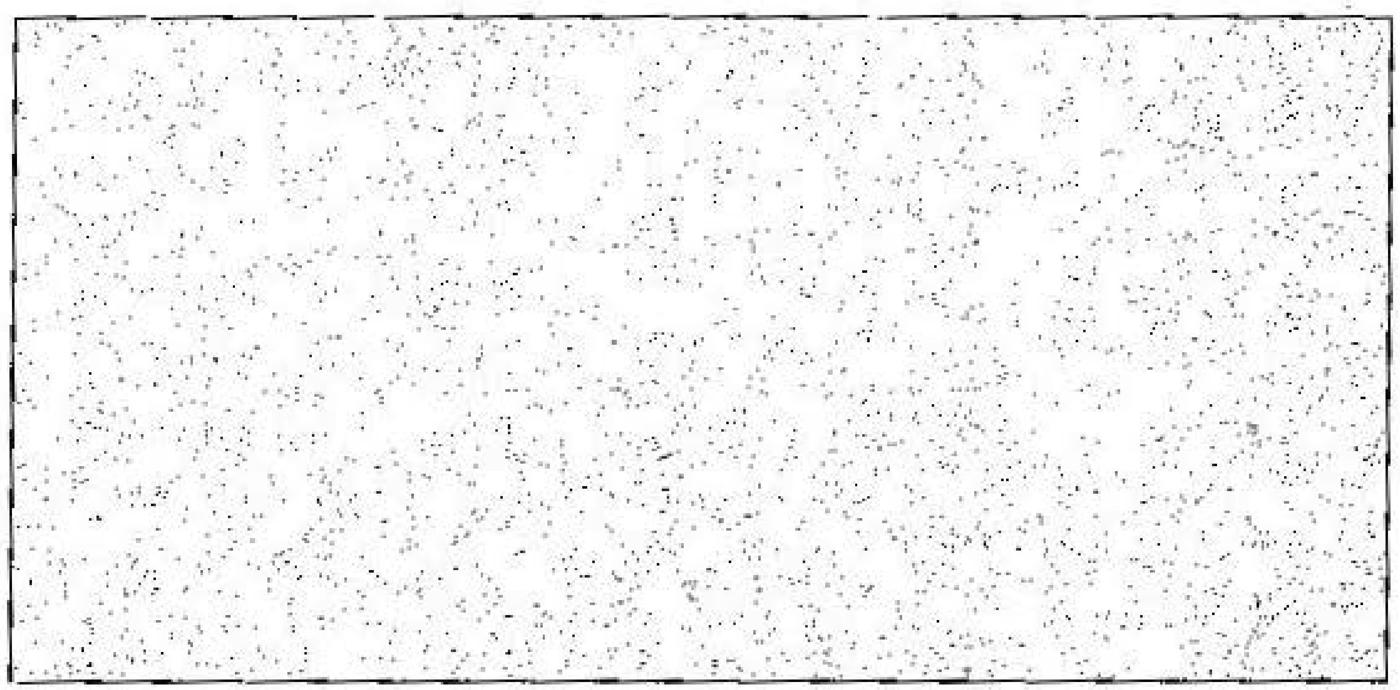
cycle 

THE ENCODER

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.



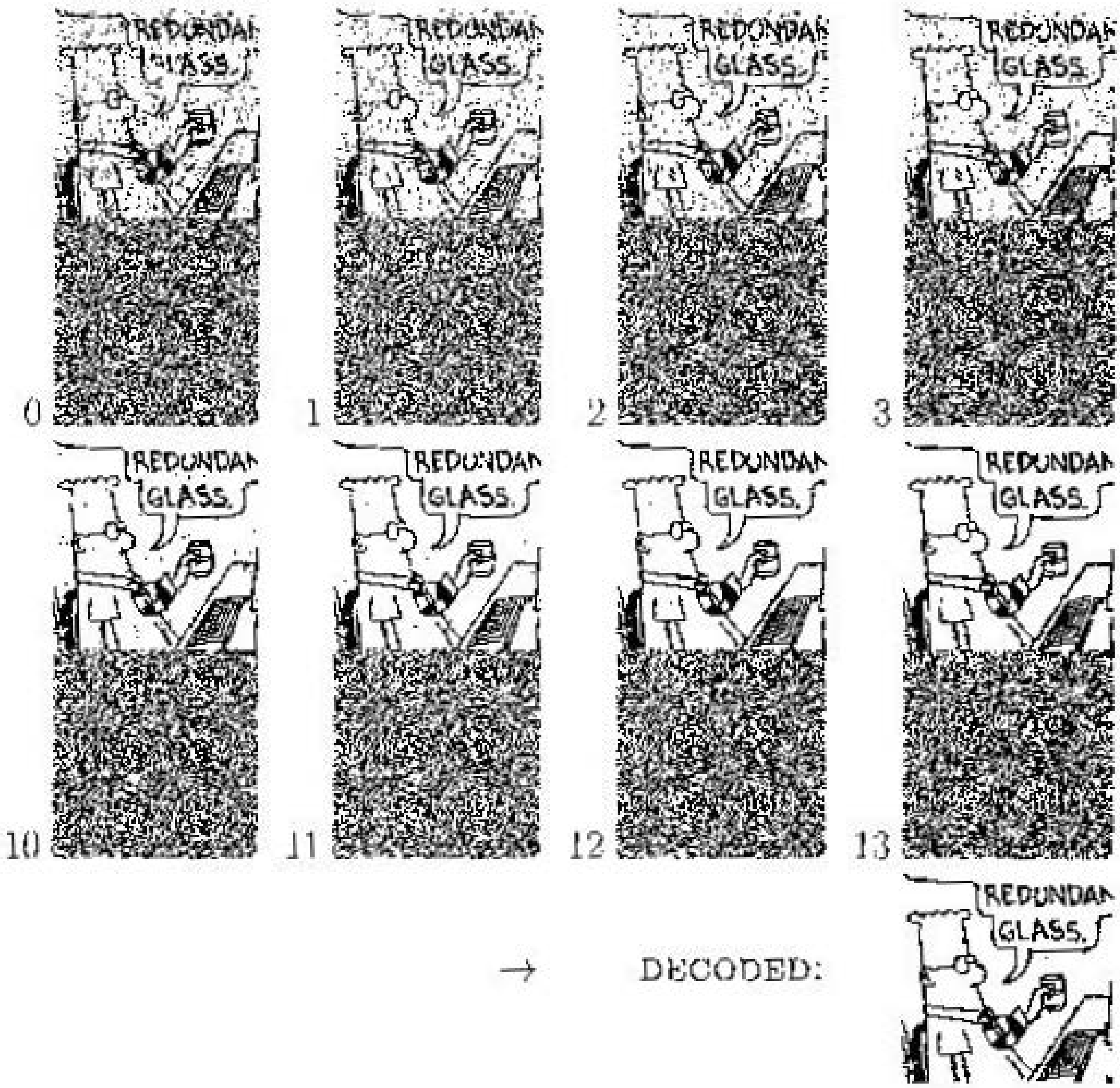
$\mathbf{H}$ =



Iterative decoding

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

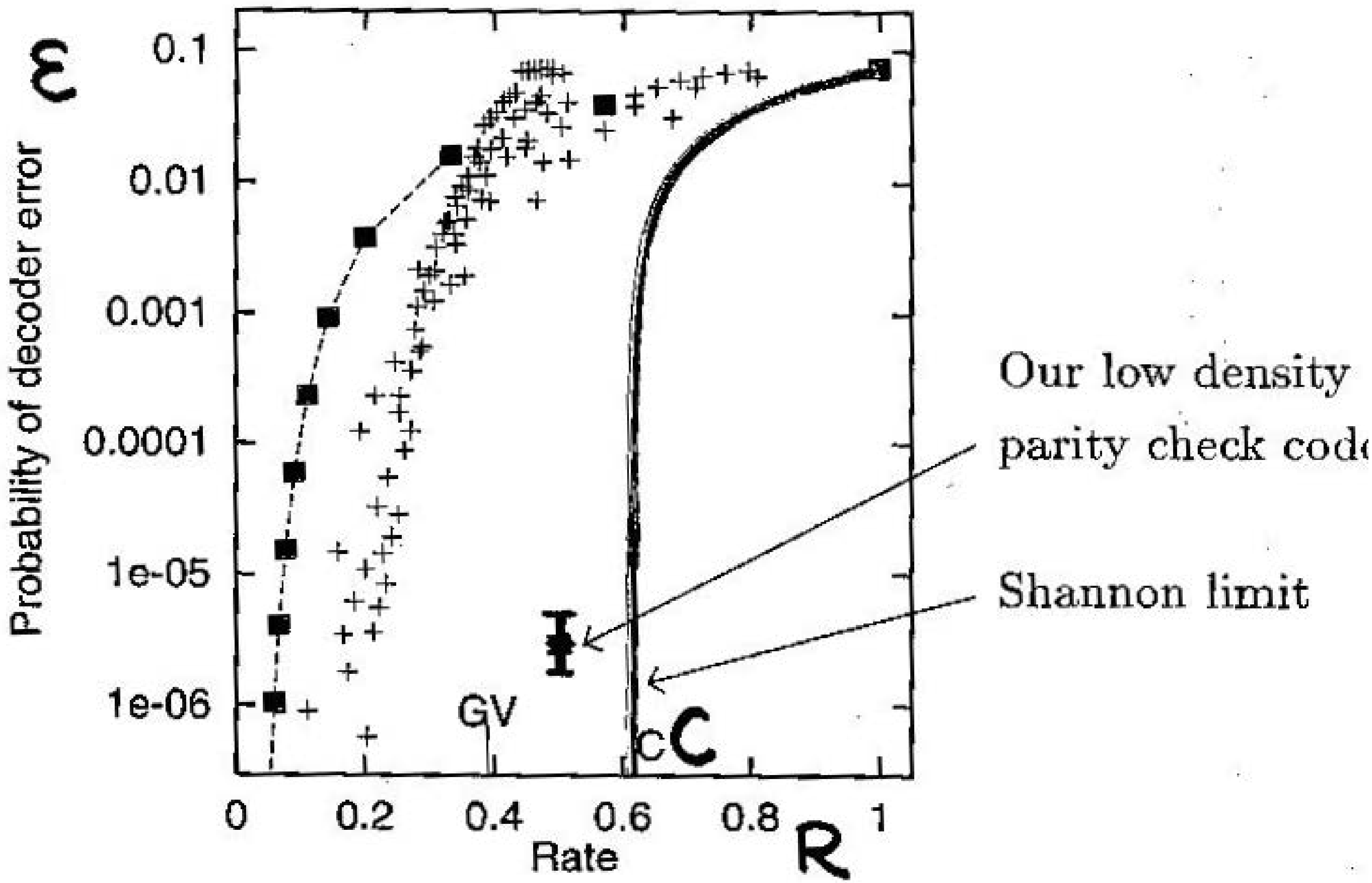
RECEIVED:



→ DECODED:

This final decoding is error free.

In the case of an unusually noisy transmission, the decoding algorithm fails to find a valid decoding. For this code and a channel with $f = 7.5\%$, such failures happen about once in every 100,000 transmissions.

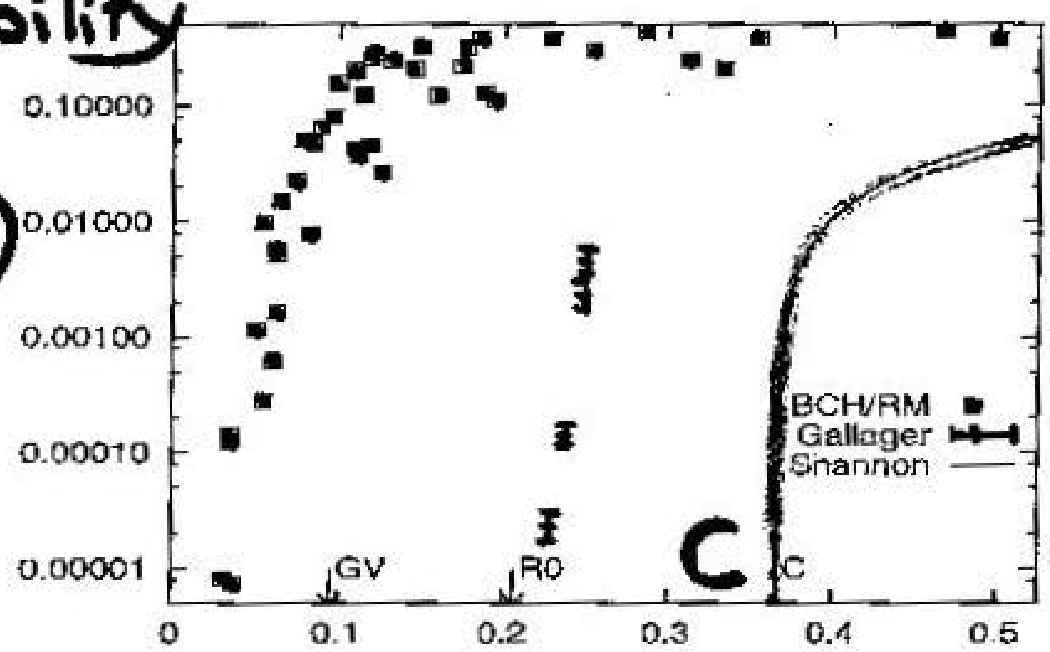


ε
Probability of decoder error
0.1
0.01
0.001
0.0001
1e-05
1e-06
GV
C
C
0
0.2
0.4
0.6
0.8
1
Rate
R
Our low density
parity check cod
Shannon limit

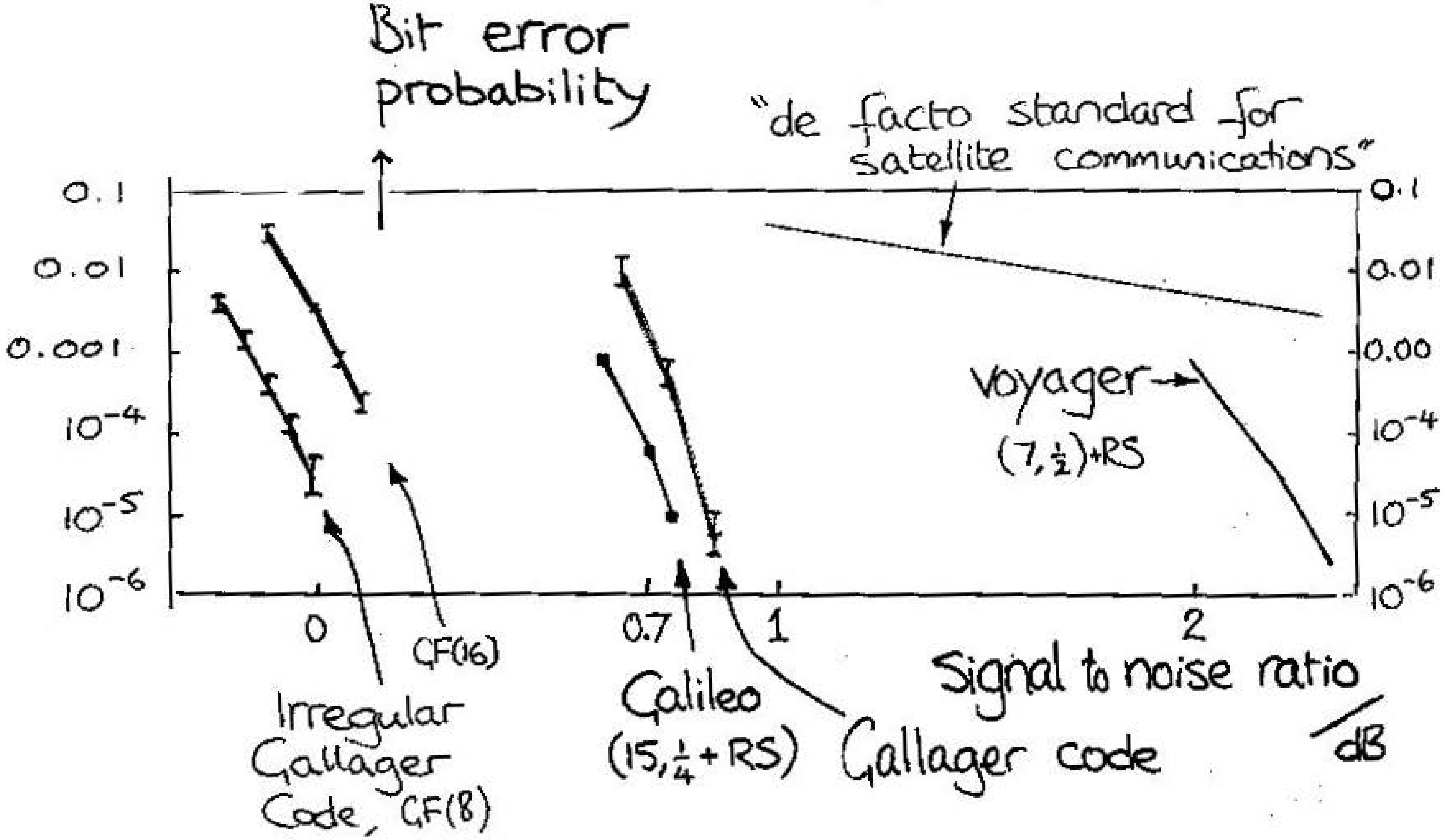
Block lengths $N \simeq 13{,}000$

Decoder error probability (log scale)

B.S.C. $f_n = 0.16$



R



N=24,000 bits N=24,448 bits N=40,000

ERRORS MADE BY Gallager CODES

Each iteration the best guess is checked — if $H\hat{\underline{n}} = \underline{z}$, halt.

Two potential failure modes:

1) Undetected errors
   — when decoder halts in a nearby codeword.

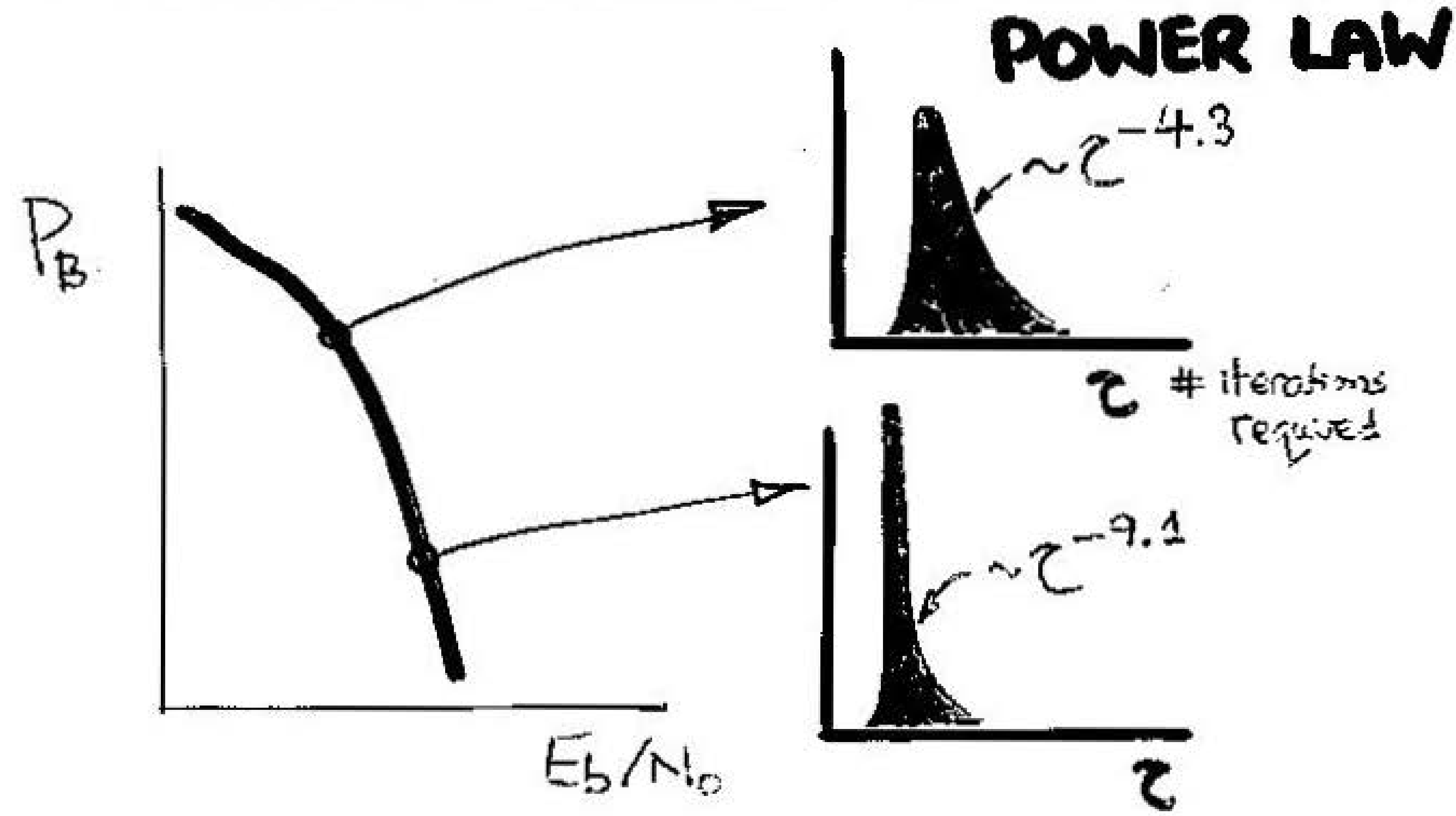
2) Detected errors
   — decoder reaches some max. number of iterations.

ALL* ERRORS ARE DETECTED ERRORS.

Gallager codes do not have low-weight codewords.

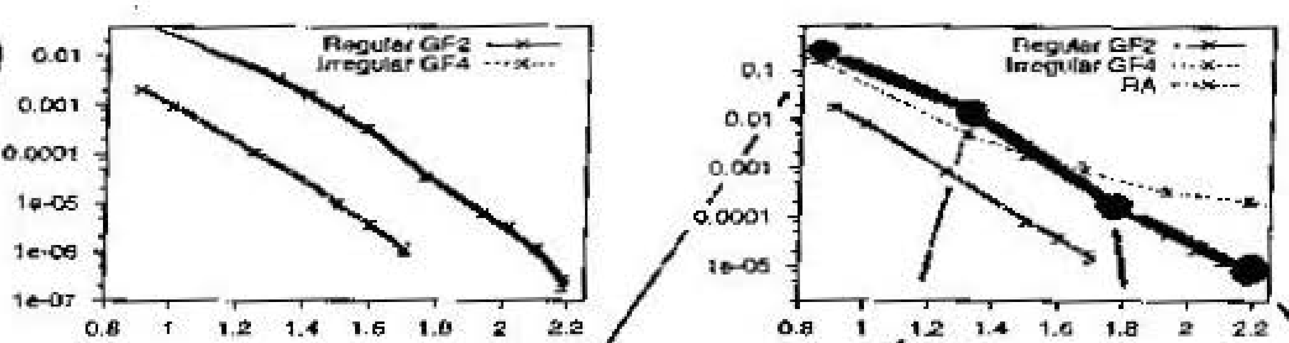
* in 10^8 blocks of experiments, for all codes with $N > 400$, $R \in (\frac{1}{4}, \frac{2}{3})$

Decoding times of sparse graph codes



DECODING TIMES HAVE A POWER-LAW DISTRIBUTION

N=1920
R=$\frac{1}{3}$

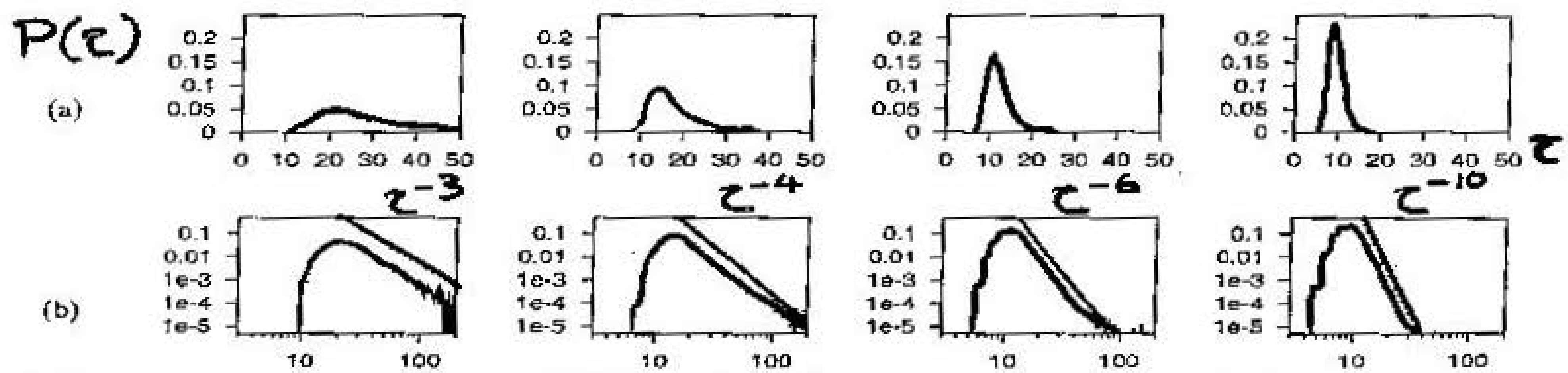


Profile

Columns		Rows	
2	449	5	635
3	389	6	5
4	72		
5	1		
11	2		
13	4		
19	39		

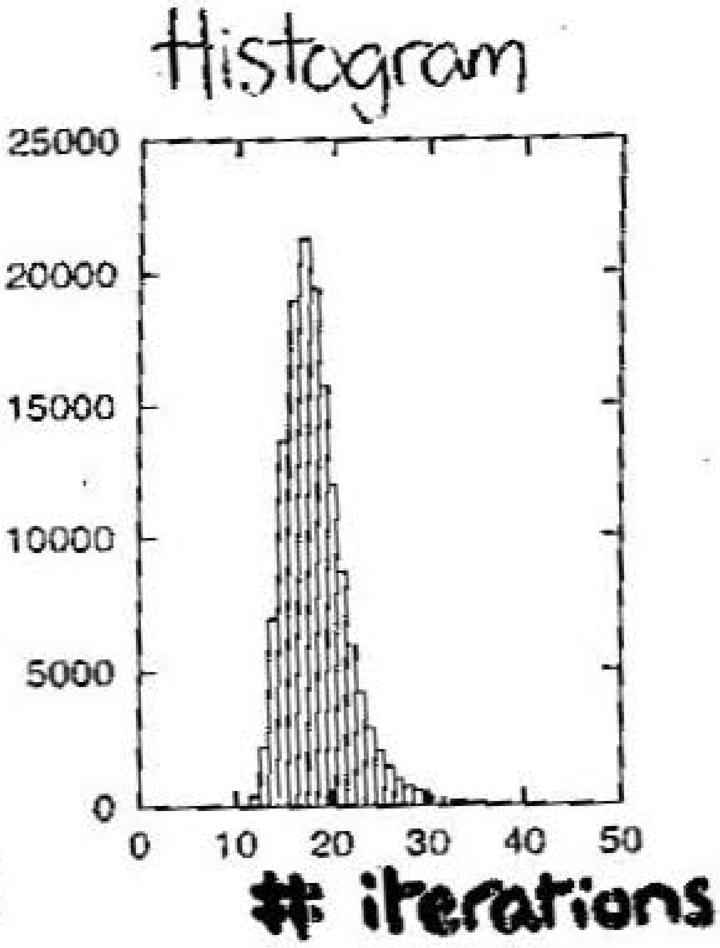
(c)

Figure 1. (a,b) Performance of Gallager codes with $N = 1920$, $R = 1/3$, as a function of E_b/N_0. In (b) we also show the performance of a repeat–accumulate code with $N = 3000$. (c) The profile of the irregular code over GF(4).

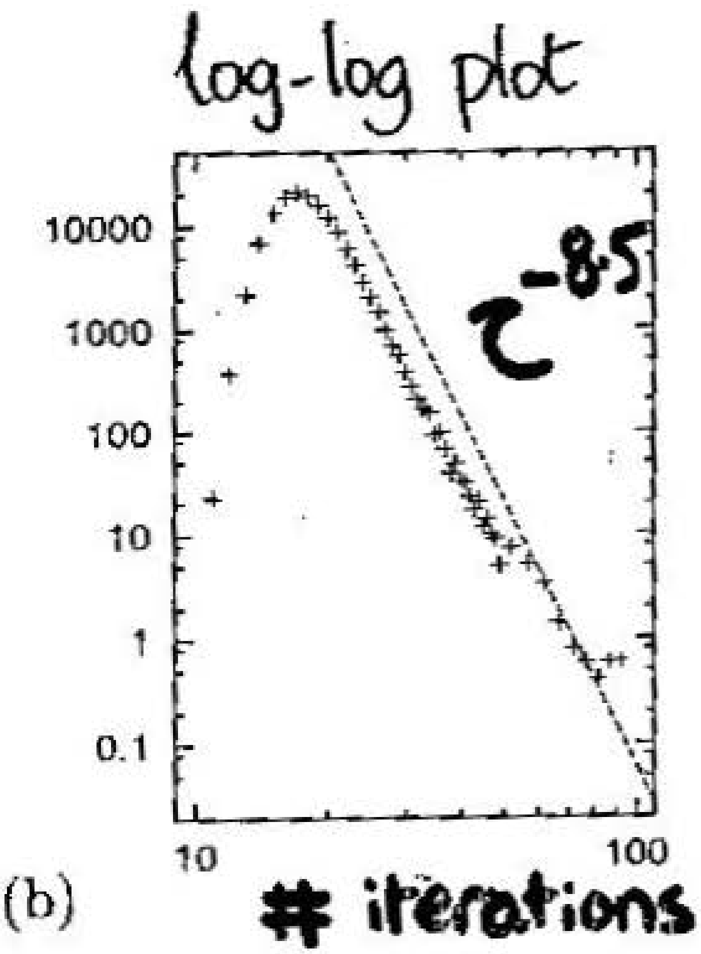


x/σ	0.90	0.95	1.00	1.05
(E_b/N_0)/dB	0.846	1.315	1.761	2.185
P_w	0.19	0.014	1.7×10^{-4}	9.0×10^{-6}
Power, p	3	4	6	10

Figure 2. Histograms showing the frequency distribution of decoding times for the binary Gallager code from figure 1: (a) linear plot; (b) log–log plot. The graphs show the number of iterations taken to reach a valid decoding; the value of P_w gives the frequency with which no valid decoding was reached after 1000 iterations. The power p which gives a good fit of the power law distribution $P(\tau) \propto \tau^{-p}$ (for large τ) is also shown.



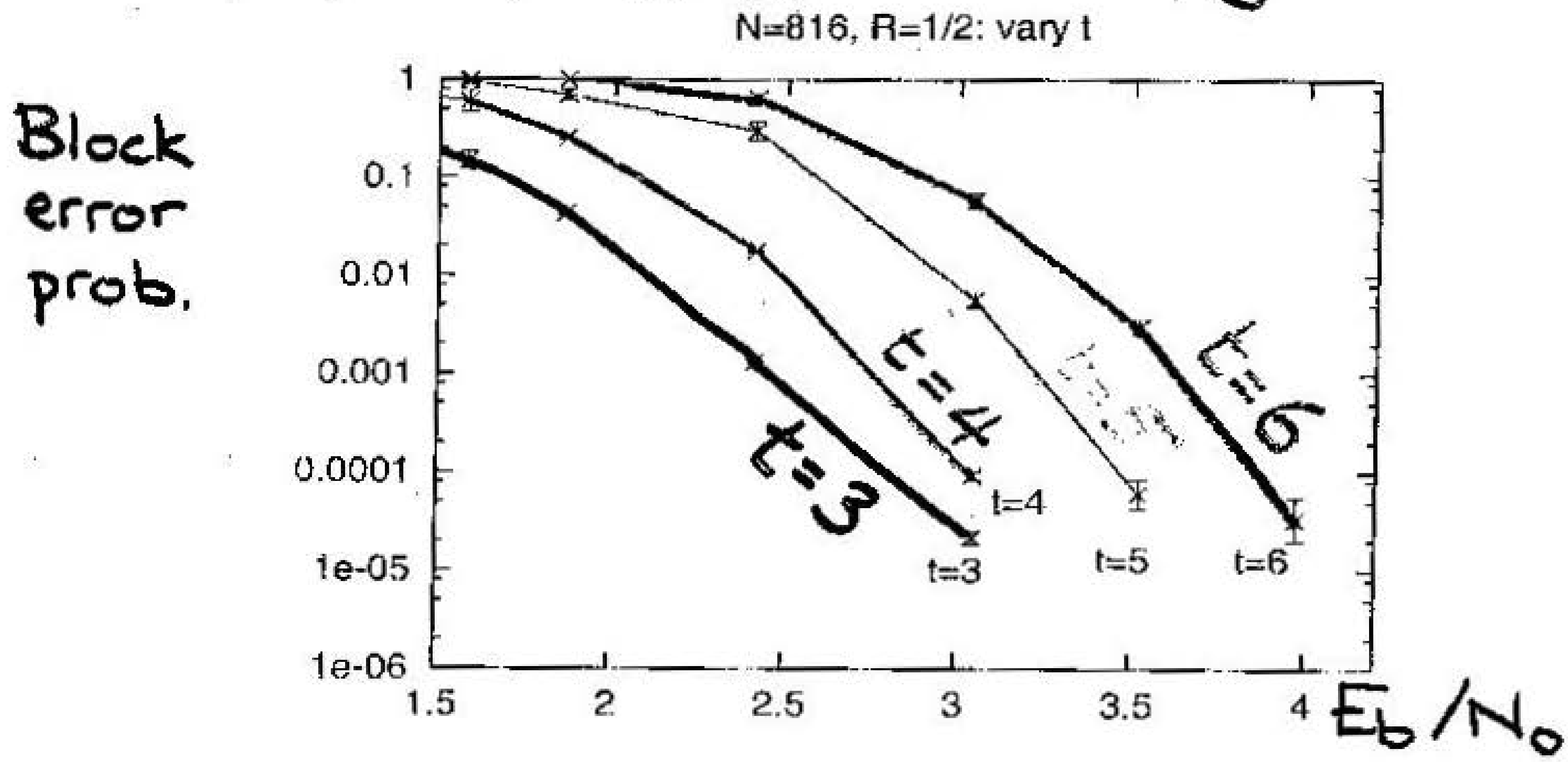
(a)



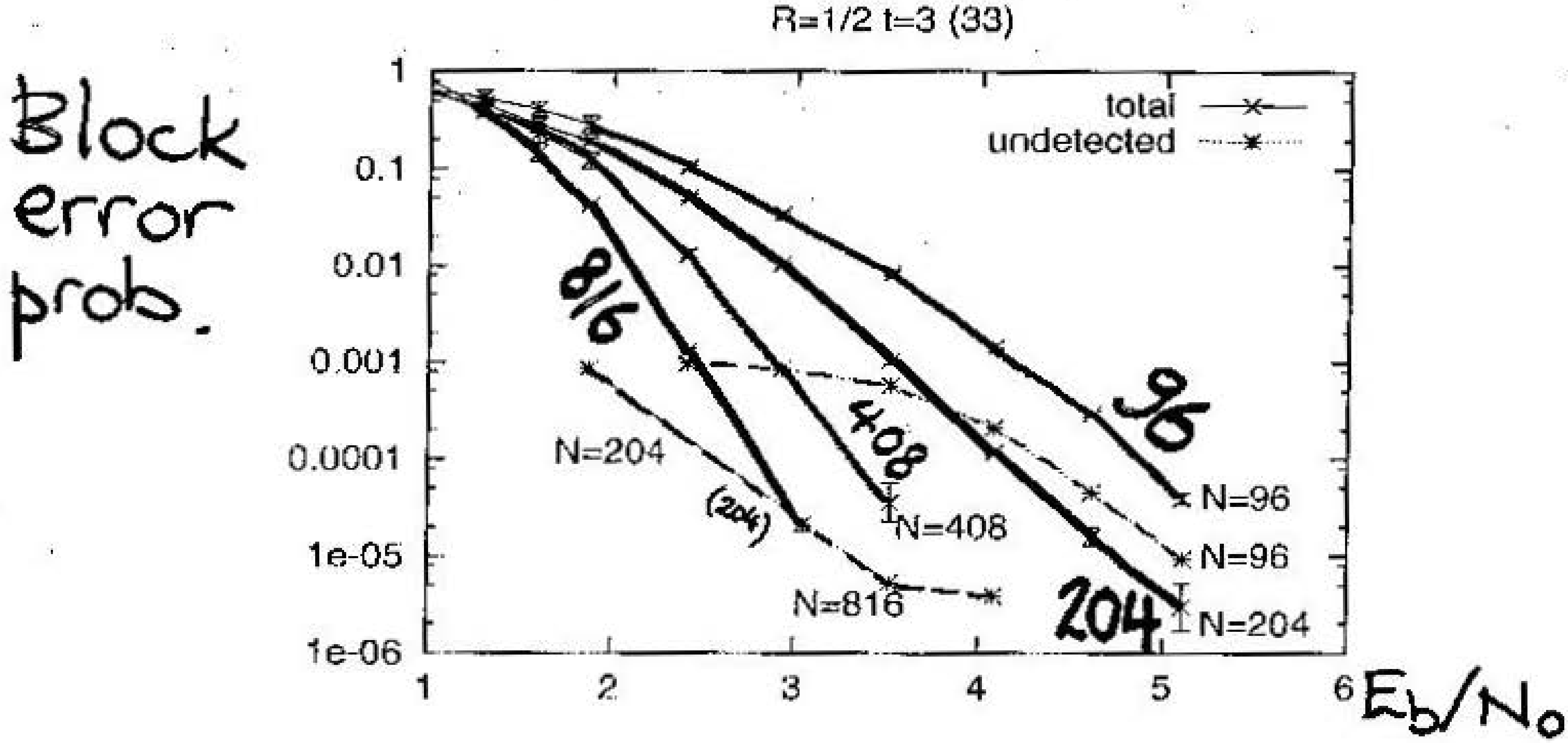
(b)

irregular Gallager code, $N = 9972$, $R = 1/2$, $E_b/N_0 = 1.0$, $P_{block} \approx 10^{-4}$

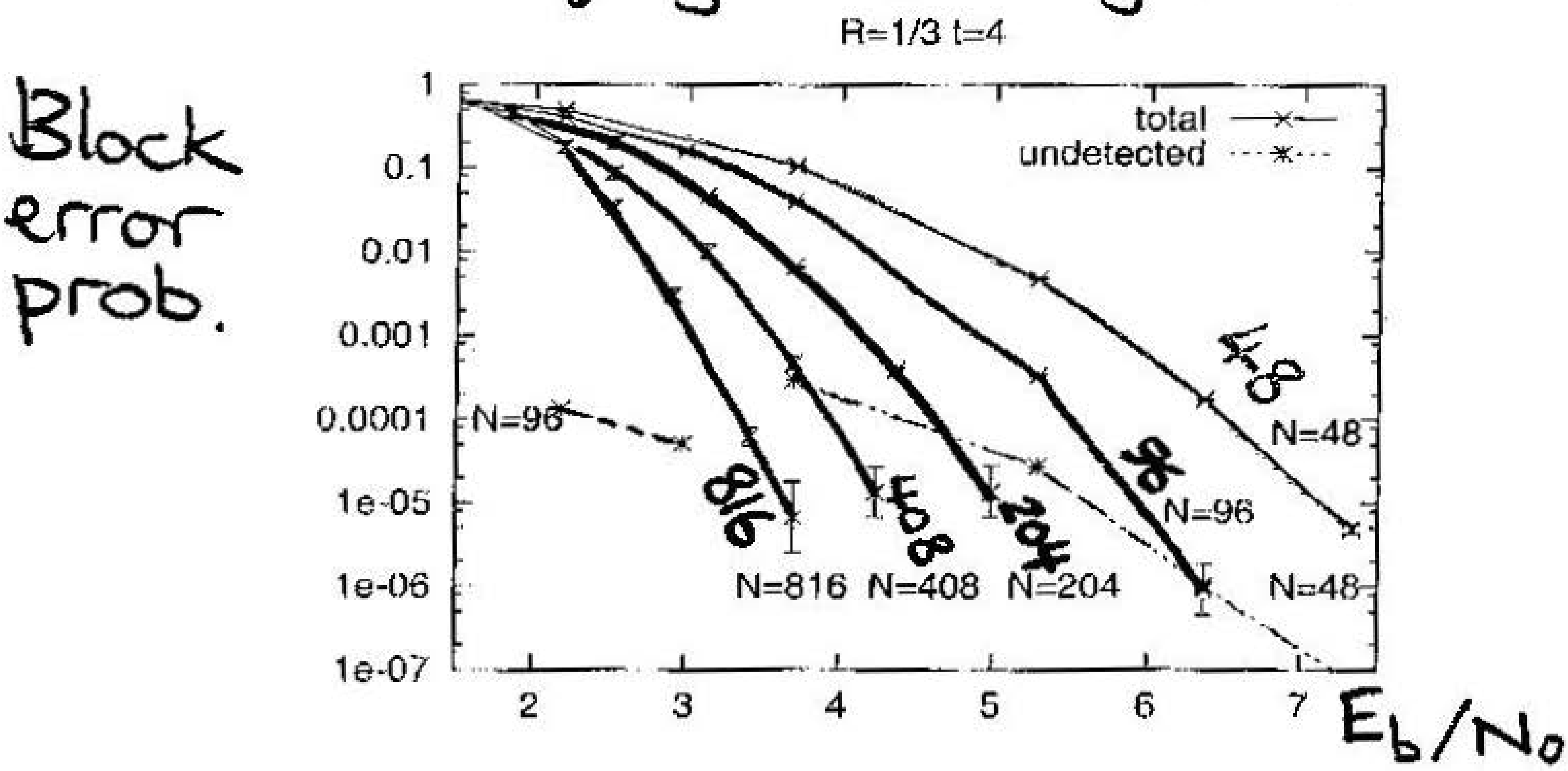
Varying column weight t



varying block length N



varying block length N



CRITERIA FOR COMPARING CODES

- E_b / N_0 for given P(block error)
  OR
  Distance from Shannon limit

- Whether block errors are undetected errors.

- Encoding complexity (Time, Space)

- Decoding complexity

- Ease of construction

"Gallager codes better than Turbo codes"

What does better mean?

SITUATION AT ~~PRESENT~~ 1995

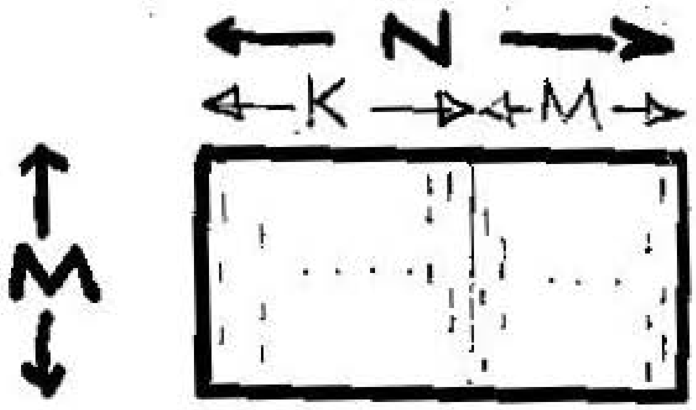
Definition	Standard Turbo Code	Standard Gallager Code
Closest to Shannon limit E_b/N_0 for given total error rate	T	
Smallest undetected error rate		G
Largest minimum distance		G
Faster decoding		G
Faster encoding	T	
More flexible choice of N,R		G
Can prove Shannon-esque properties		G
Patent-free		G

TOTAL = DETECTED ERROR RATE + UNDETECTED ERROR RATE

ENCODING

$\underline{s} \in \{0,1\}^K \rightarrow \underline{t}$

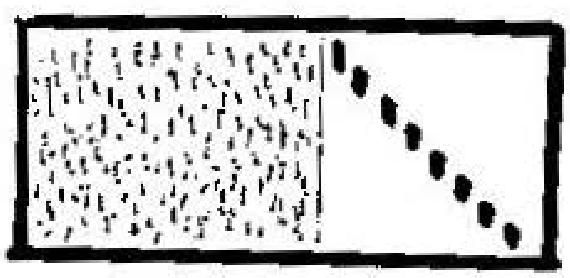
$Ht = 0$



$= \underline{\underline{H}}$

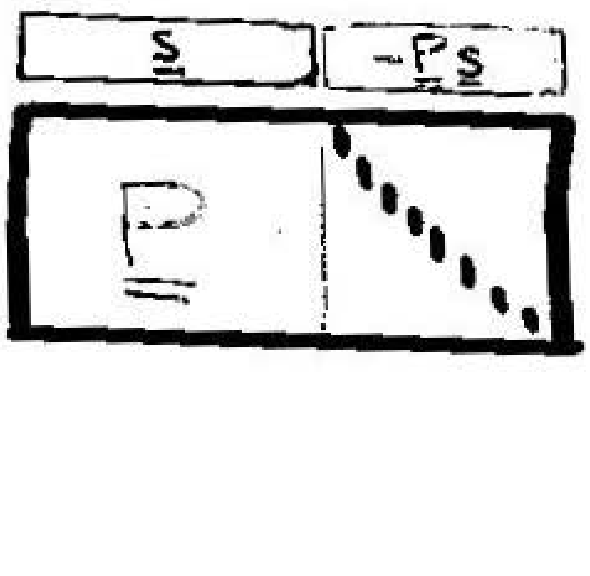
$\Downarrow$ Gaussian Elimination

$O(N^3)$

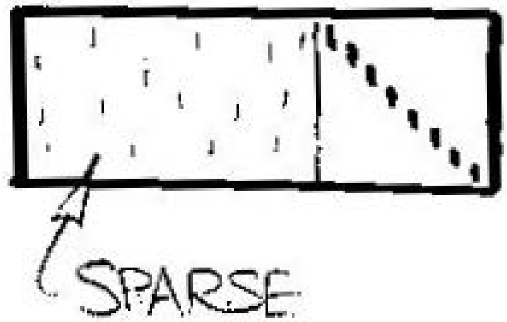


$= \underline{\underline{H}}'$

$\left[\; \underline{\underline{H}}\,\underline{t} = 0 \iff \underline{\underline{H}}'\underline{t} = 0 \; \right]$



$O(N^2)$
if $\underline{\underline{P}}$ dense

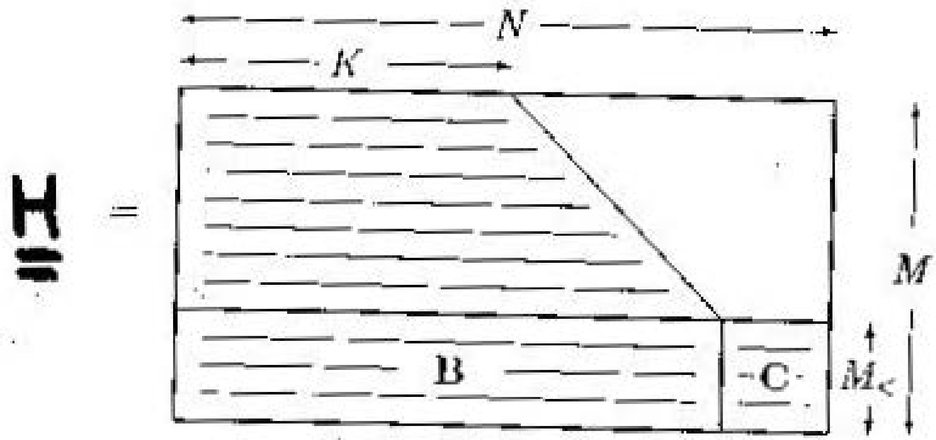
If $H =$ 

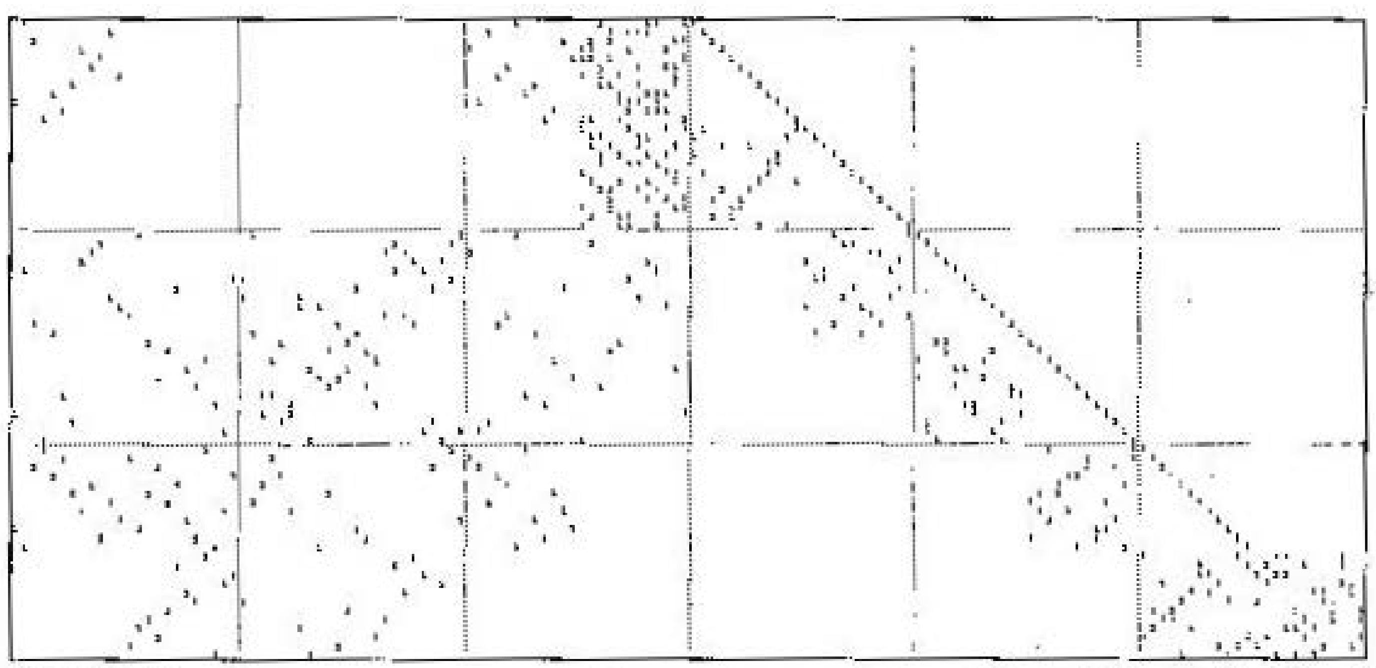
then encoding is fast – $O(N)$

But such codes are <u>bad</u>.

(Many low-weight codewords)

FAST-ENCODING GALLAGER CODES

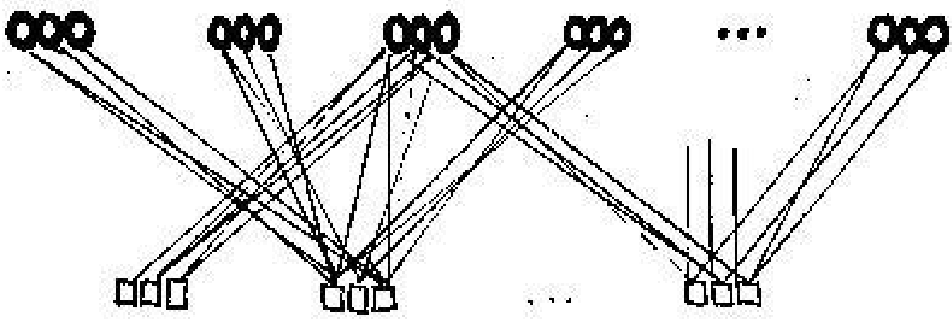




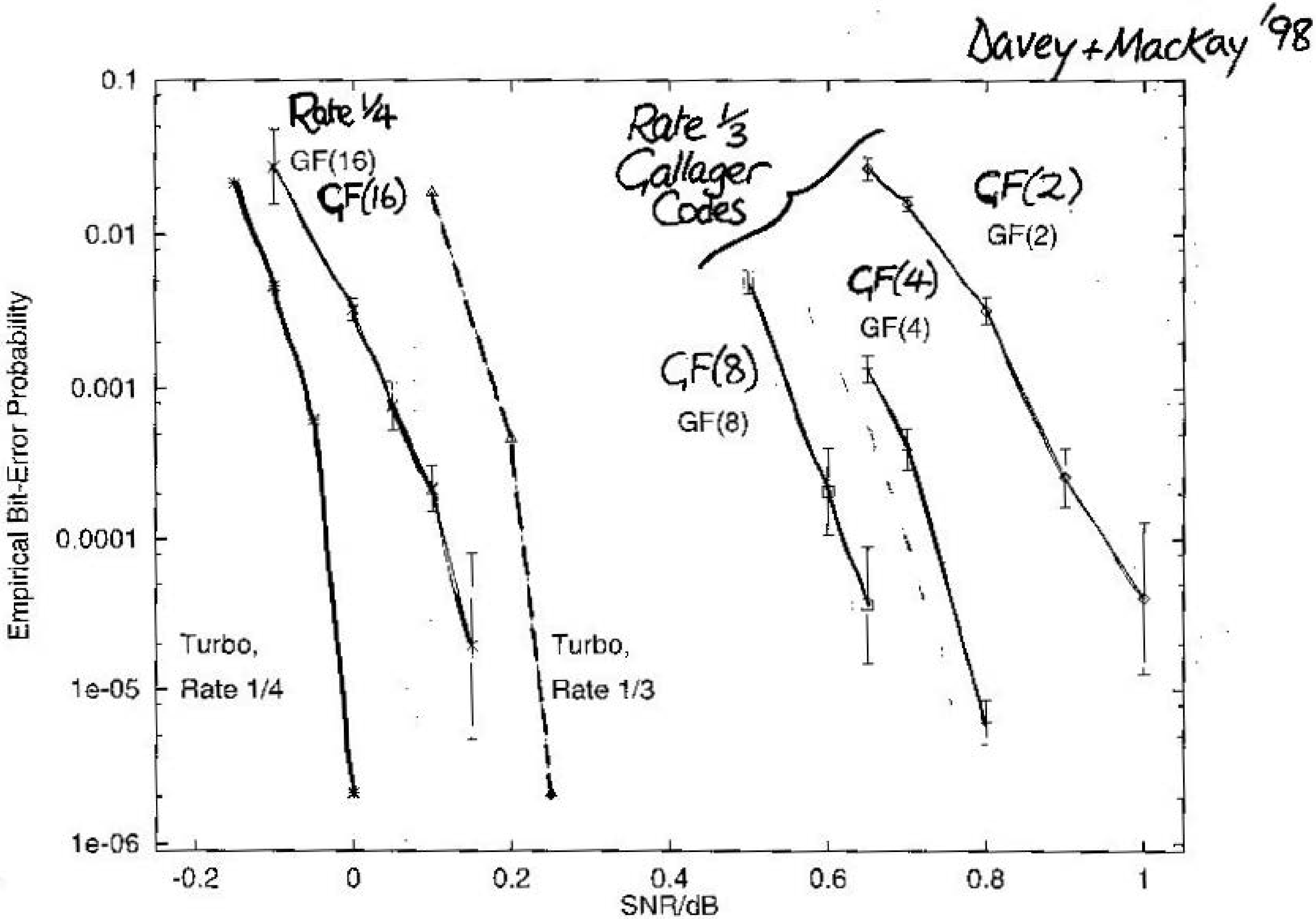
FAST-ENCODING GALLAGER CODE (Super-Poisson)

IMPROVING GALLAGER CODES I

- Clump bits together and track correlations during decoding



(Davey and MacKay)

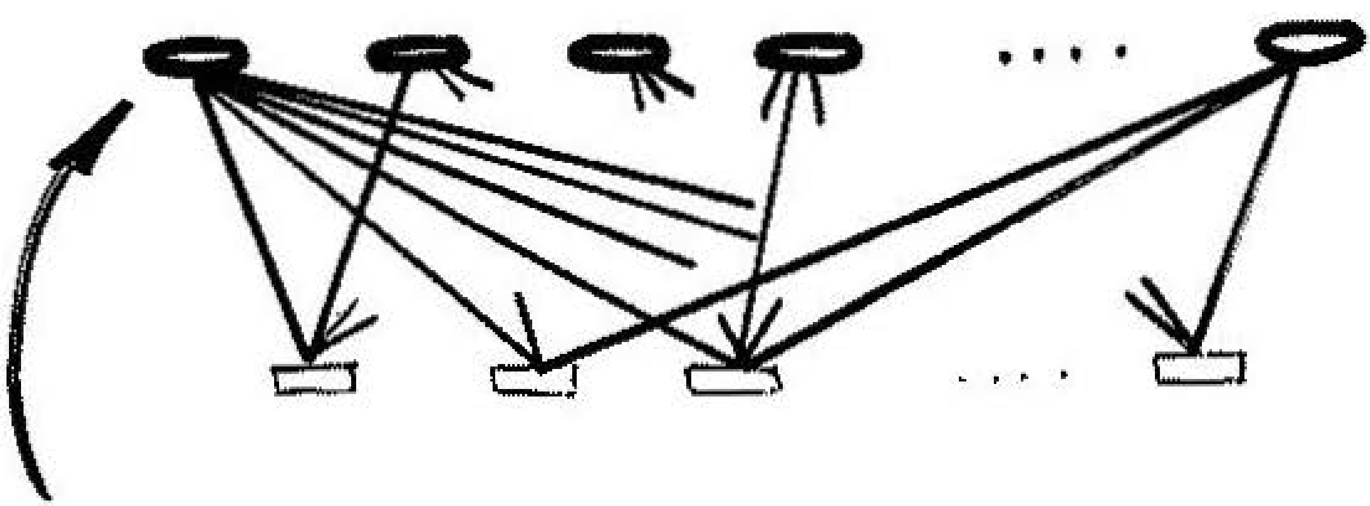


Davey + MacKay '98
Rate 1/4
Rate 1/3 Gallager Codes
GF(16)
GF(16)
GF(2)
GF(2)
GF(4)
GF(4)
GF(8)
GF(8)
Turbo, Rate 1/4
Turbo, Rate 1/3
Empirical Bit-Error Probability
SNR/dB
0.1
0.01
0.001
0.0001
1e-05
1e-06
-0.2
0
0.2
0.4
0.6
0.8
1

IMPROVING GALLAGER CODES II

- Make graph <u>irregular</u>

  (Luby, Mitzenmacher, Shokrollahi, & Spielman ;
  Davey and MacKay)



a privileged
bit participates
in many constraints

1999: Richardson, Shokrollahi & Urbanke

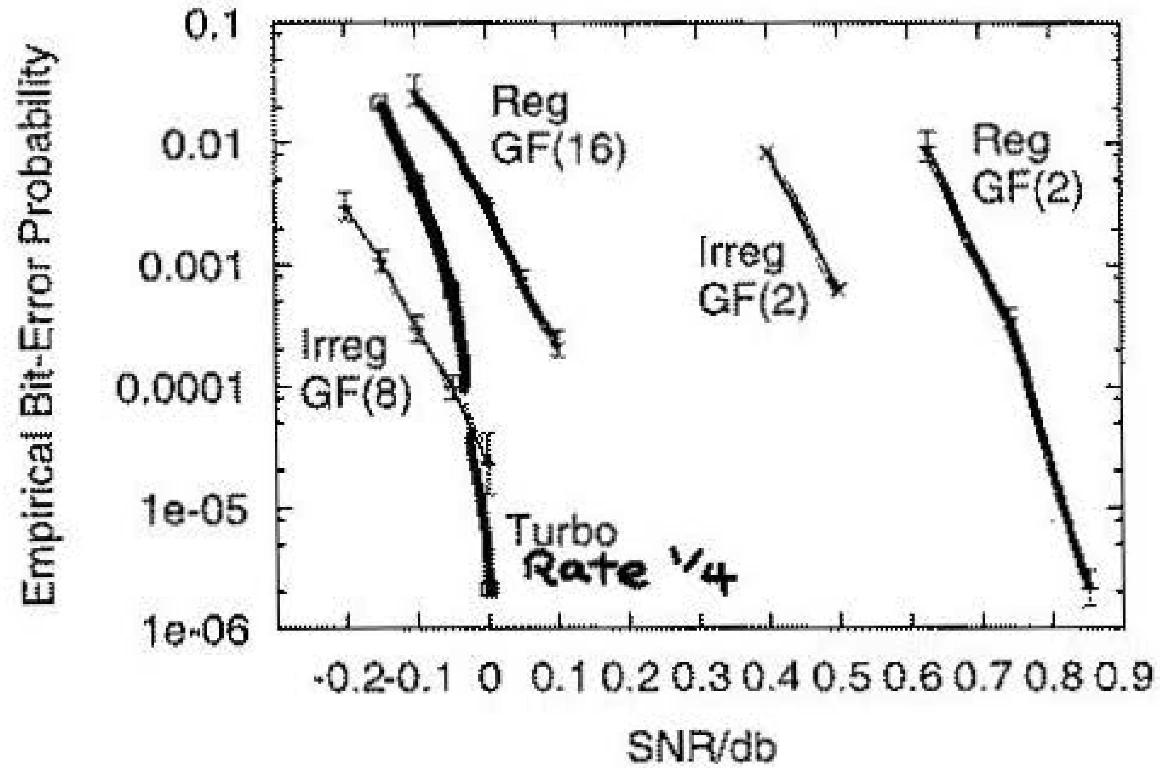


Figure 1: Empirical results for Gaussian Channel, Rate 1/4 Left–Right : Irregular LDPC, $GF(8)$ blocklength 24000 bits; JPL Turbo, blocklength 65536 bits; Regular LDPC, $GF(16)$, blocklength 24448 bits; Irregular LDPC , $GF(2)$, blocklength 64000 bits; Regular LDPC, $GF(2)$, blocklength 40000 bits. (Reproduced from [1].)

Col. Weight	2	3	7	9	11	15	19	23	29	43
fraction	0.700	0.191	0.037	0.024	0.013	0.017	0.008	0.005	0.003	0.002

Table 4.1: Parameters of good irregular code for $GF(8)$, rate 0.25

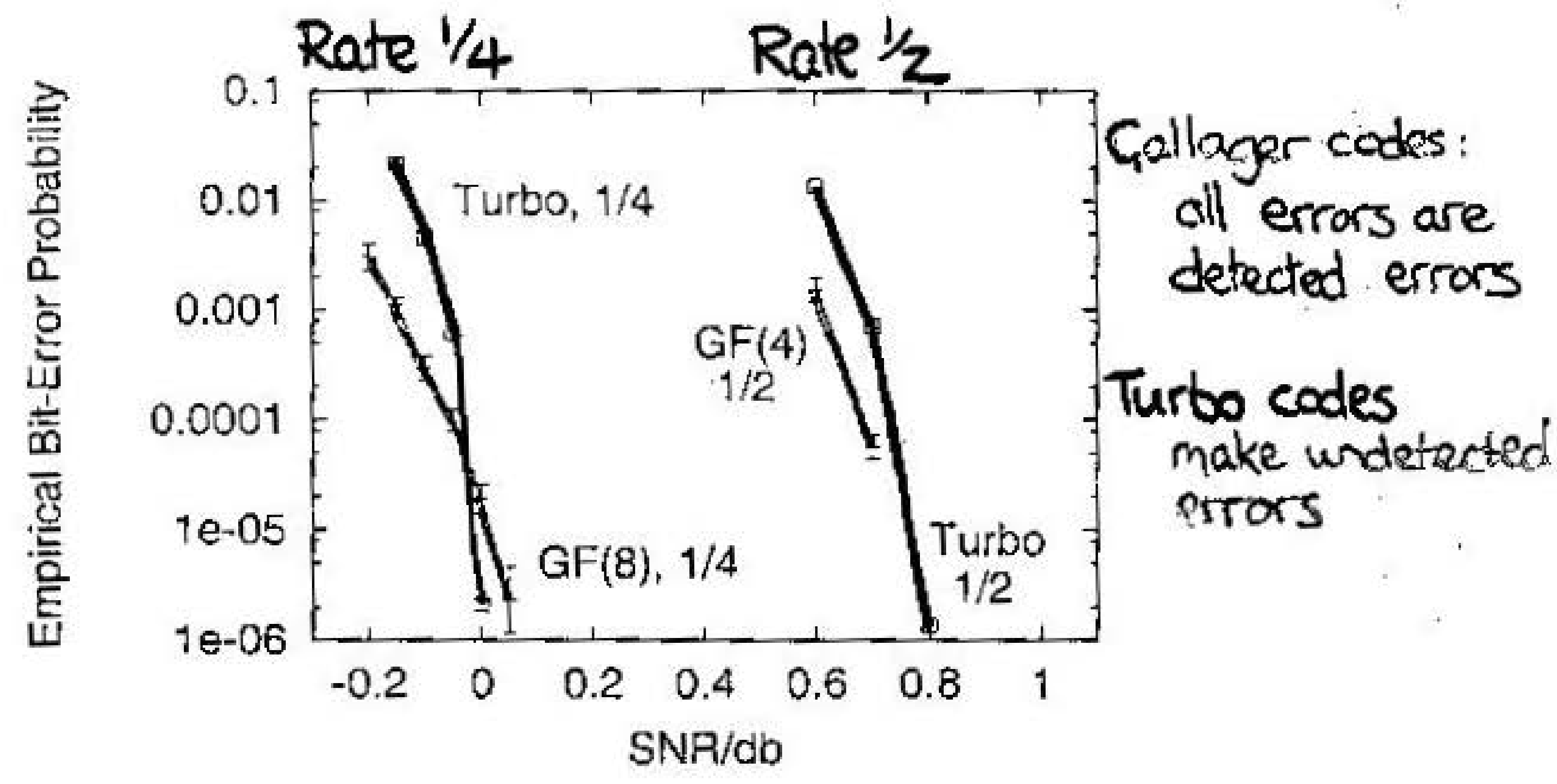
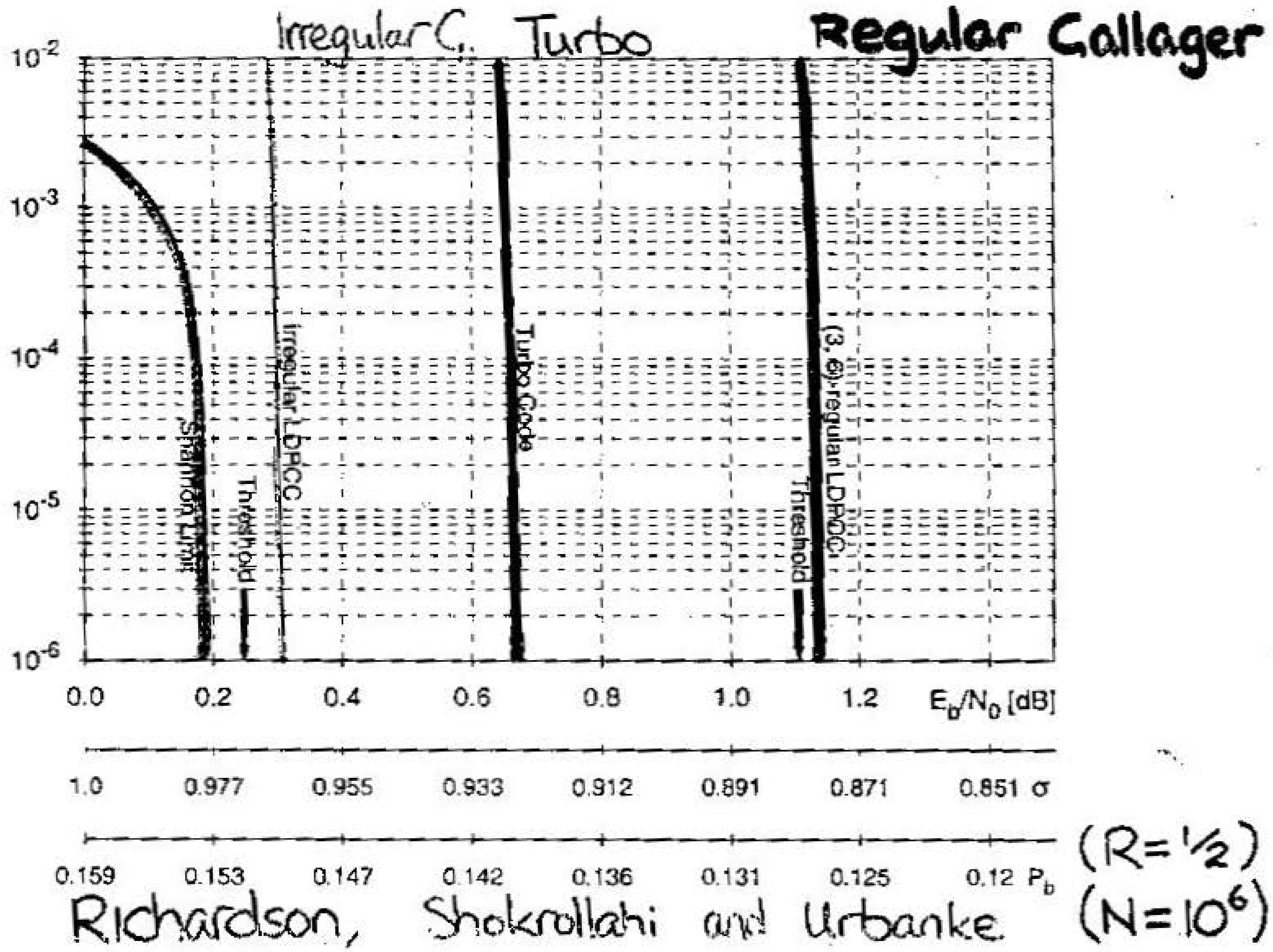
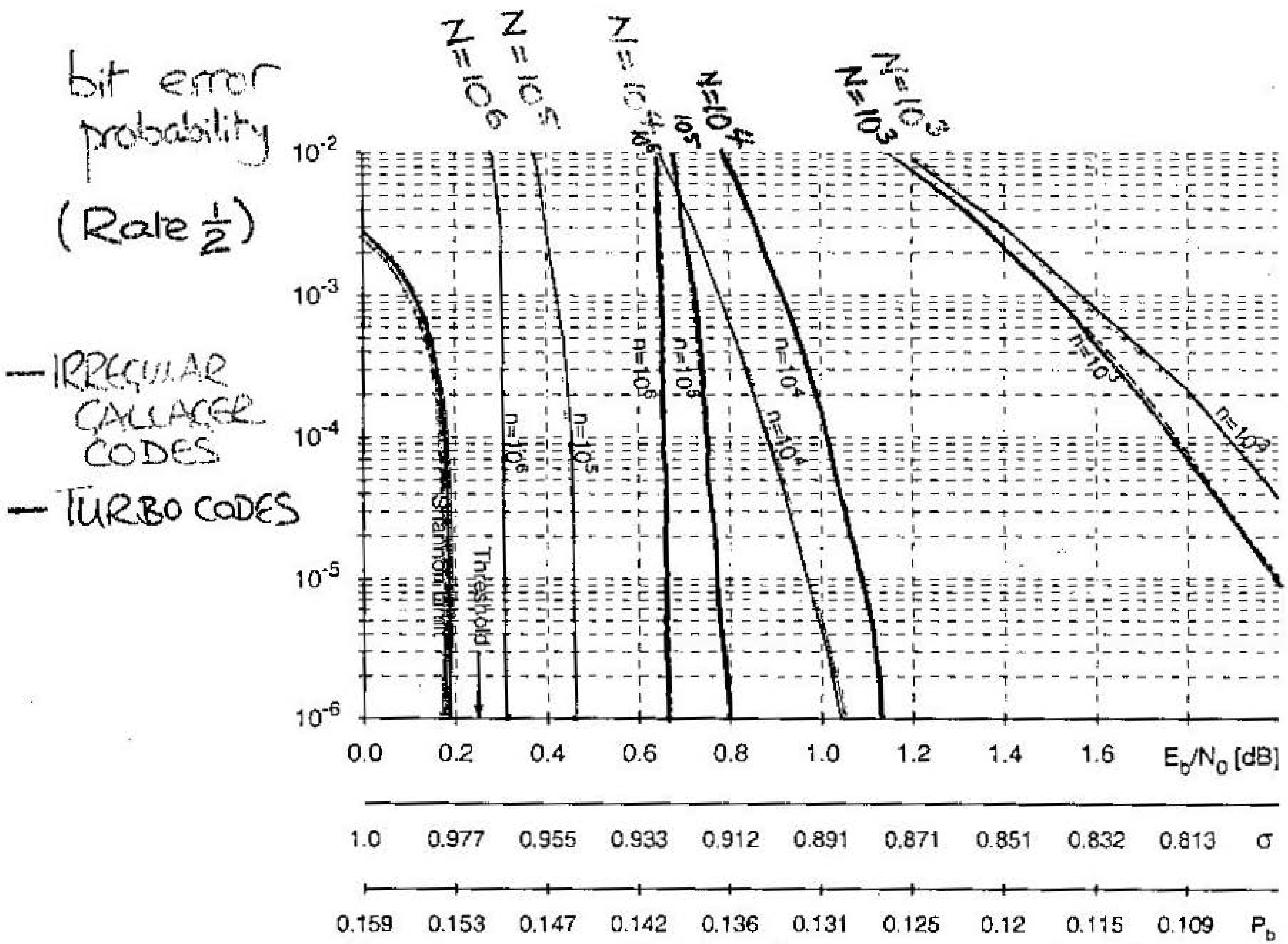


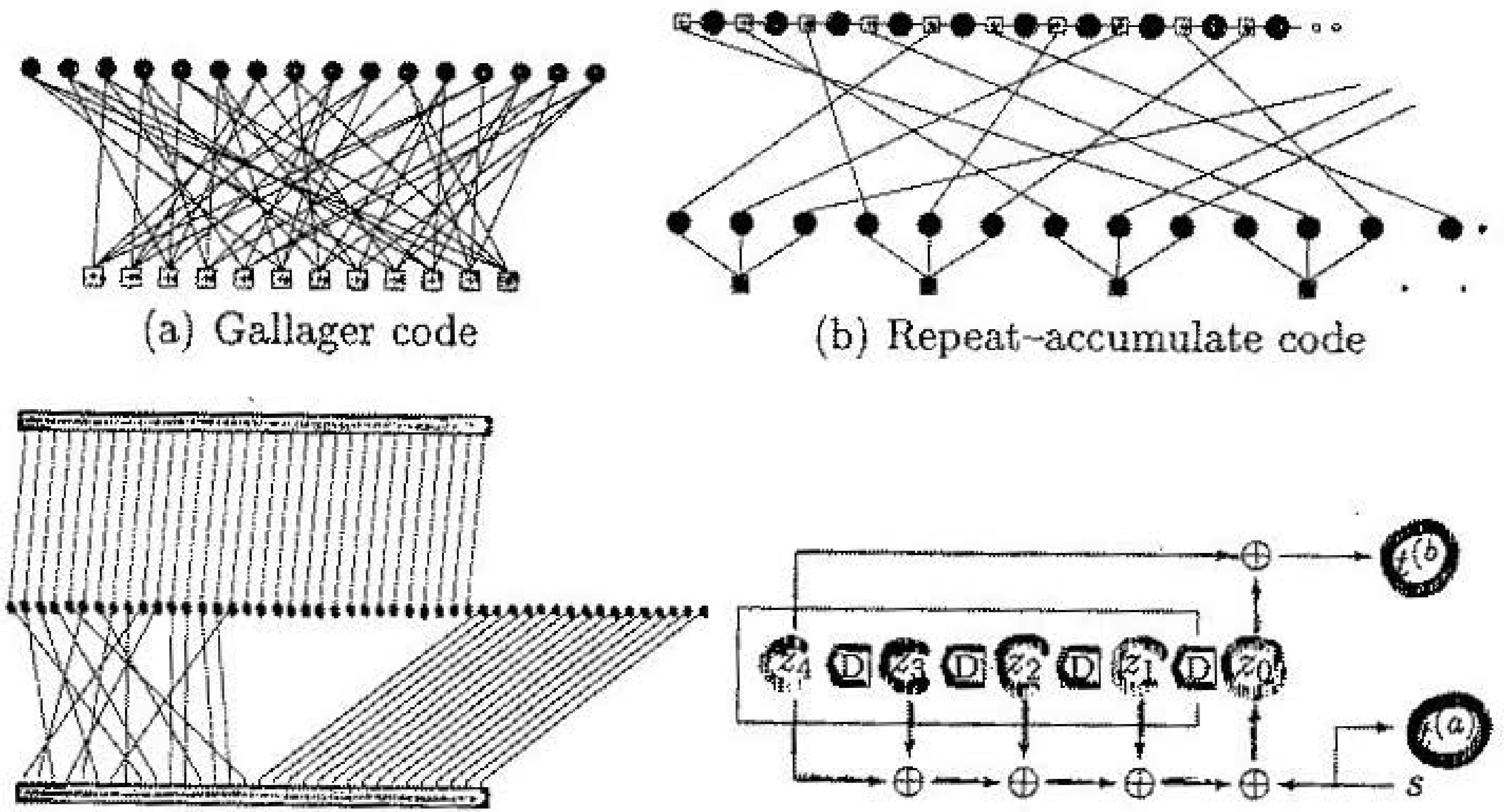
Figure 4-2: Current state of the art in error correction. Left to right: Rate 1/4 irregular LDPC over $GF(8)$ blocklength 24000 bits; Rate 1/4 JPL turbo code [8] blocklength 65536; Rate 1/2



Richardson, Shokrollahi and Urbanke



Richardson, Shokrollahi and Urbanke



(a) Gallager code

(b) Repeat–accumulate code

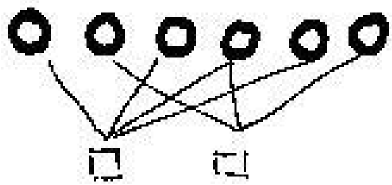
(c1) Turbo code

(c2) $(21/37)_8$ recursive convolutional code

Figure 1. Graphs of three sparse graph codes.

Where to go from Regular Gallager Codes

- Clump bits, & clump checks
- Make irregular



- Change type of constraints
- Add state-variables
- Algebraic constructions

DIFFERENCE-SET CYCLIC CODES

N	7	21	73	273	1057	4161
M	4	**10**	**28**	**82**	**244**	**730**
K	3	11	45	191	813	3431
d	4	**6**	**10**	**18**	34	66
k	3	**5**	**9**	17	33	65

THE FUTURE OF GALLAGER CODES ?

Gallager codes that satisfy
more than M low weight constraints.

eg Difference-set cyclic codes

- $N = 273$ $\quad M = 82$

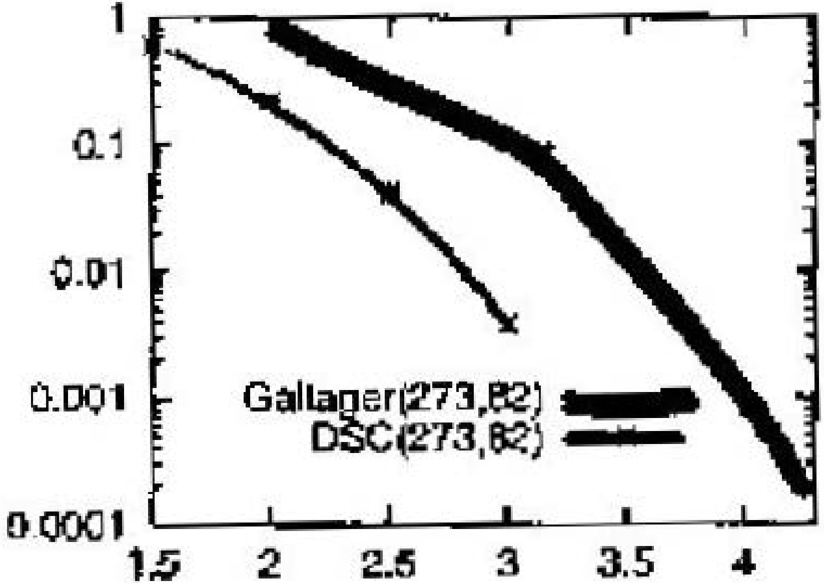
  has $H = \underbrace{\left[\begin{array}{c} \cdots \\ \text{weight} \\ 17 \end{array}\right]}_{\leftarrow 273 \rightarrow} \updownarrow 273$

- $N = 73$ $\quad M = 28$

  has 73×73 checks of weight 9.

implemented on a chip by
Tanner et al 1980s

The Tanner Challenge:
Find all such codes.

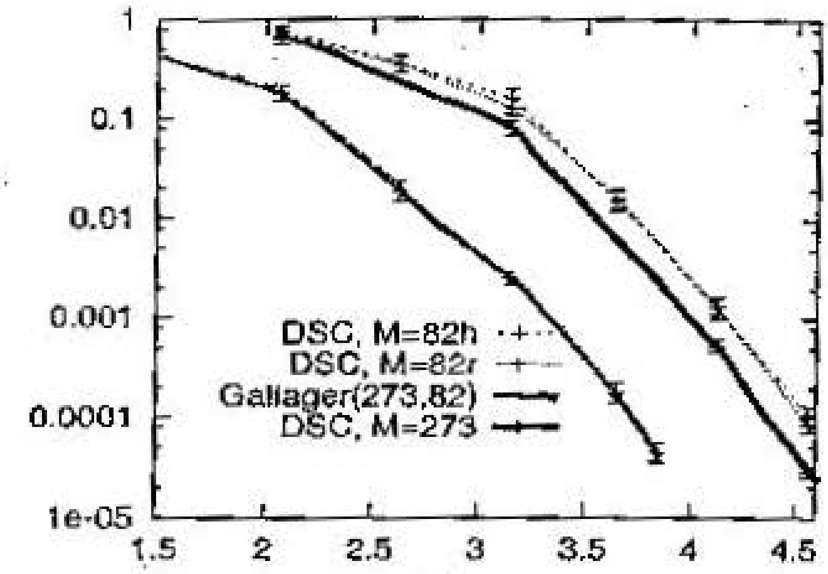


1
0.1
0.01
0.001
0.0001
1.5
2
2.5
3
3.5
4
Gallager(273,82)
DSC(273,82)

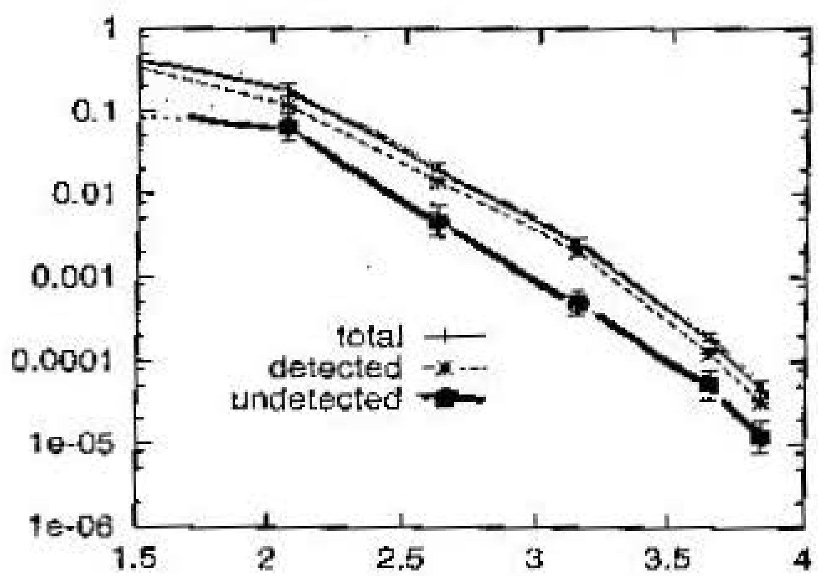
DIFFERENCE–SET CYCLIC CODES

N	7	21	73	273	1057	4161
M	4	**10**	**28**	**82**	**244**	**730**
K	3	11	45	191	813	3431
d	4	**6**	**10**	**18**	34	66
k	3	**5**	**9**	17	33	65

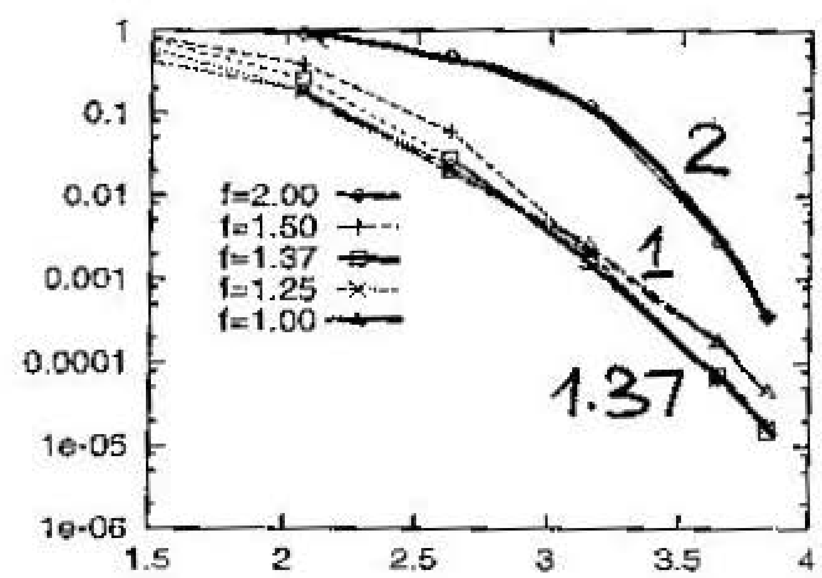
(a)



(b)



(c)



(d)

Figure 7. Difference-set cyclic codes — low-density parity-check codes satisfying many redundant constraints — outperform equivalent Gallager codes. (a) The table shows the N, M, K, distance d, and row weight k of some difference-set cyclic codes, highlighting the codes that have large d/N, small k, and large N/M. All DSC codes satisfy N constraints of weight k. In the comparison the Gallager code had $(j, k) = (4, 13)$, and rate identical to the DSC code. Vertical axis: block error probability; horizontal axis: E_b/N_0/dB.

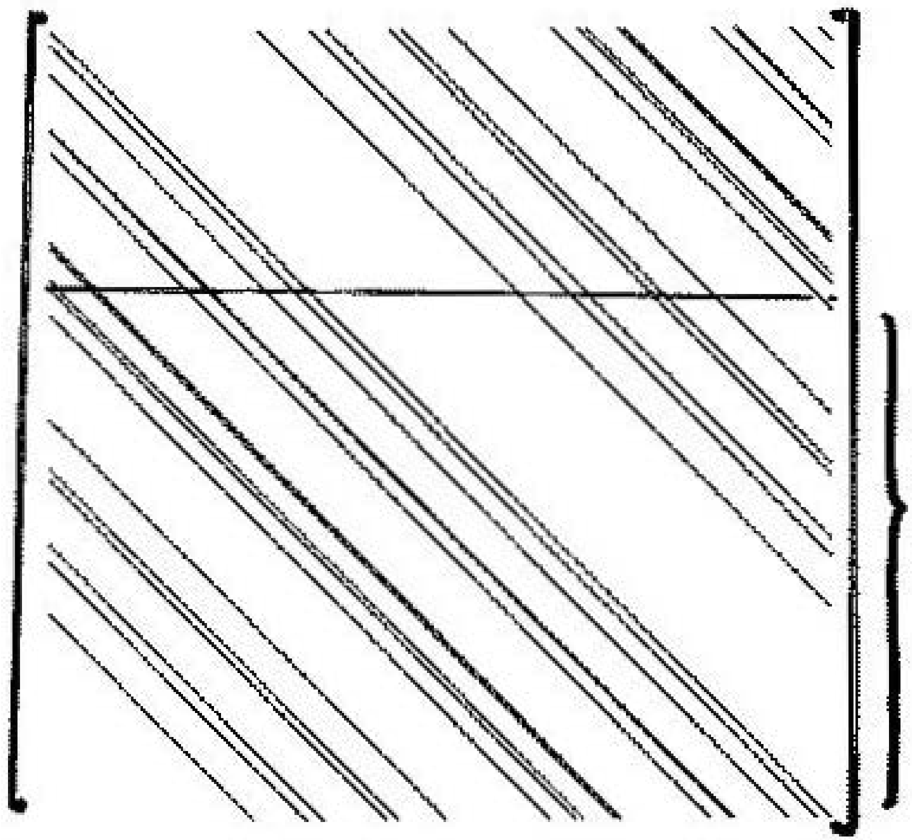
(c) The DSC code makes some undetected errors.

(d) The error rate of the DSC code can be slightly reduced by using a 'fudge factor' of 1.25 or 1.37 during the sum-product decoding.

Difference-set cyclic code

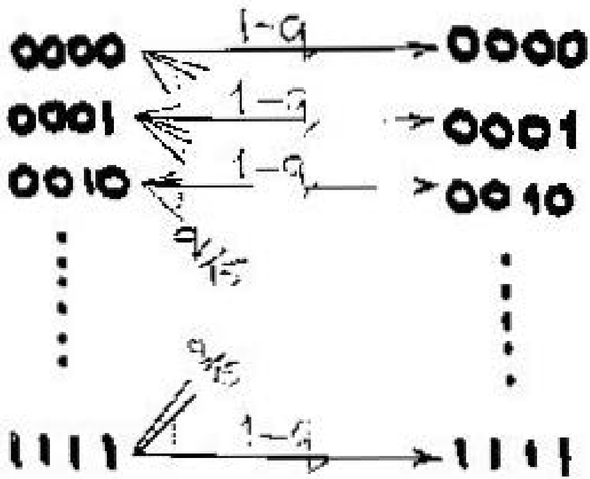
$N = 273 \quad M = 82$

$\underline{\underline{H}} =$



redundant checks

16-ary symmetric channel



$R = 8/9$

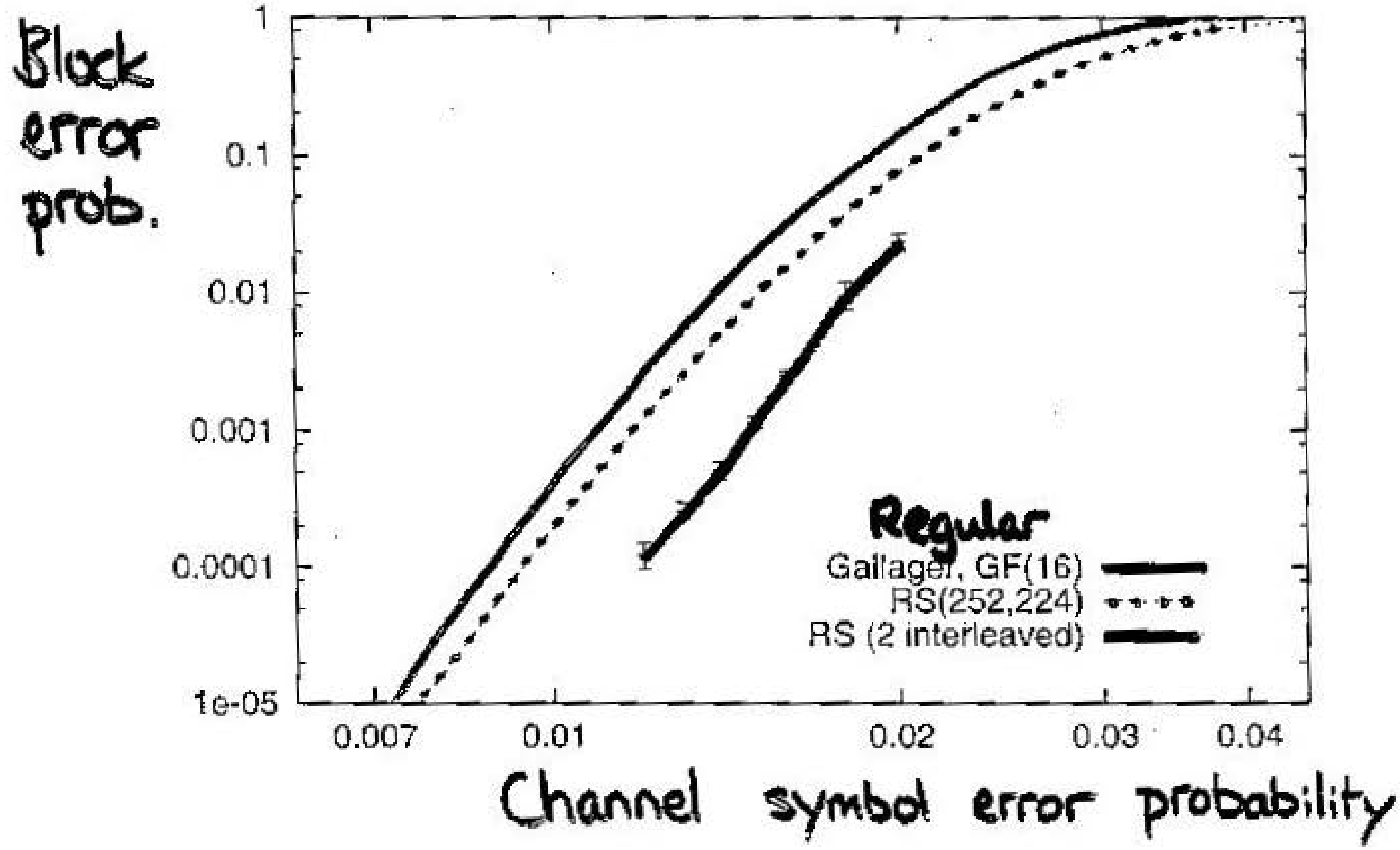
$N = 999$ (ie 3996 bits)

16-ary Symmetric Channel

N = 999 x (4bits)

R = 8/9 = 0.89

[redacted]



N = 999 . 4 = 3996 bits
M = 111.4
R = 8/9

RS. N = 252 8
= 2016 bits
Interleave 2 → 4032.

Rate $\frac{1}{2}$ $N \gtrsim 10{,}000$

Gallager codes beat

Turbo codes

How far can we push Gallager codes?

Rate 0.9 $N \simeq 4096$

Gallager codes beat

Reed-Solomon codes

(on the channels I've looked at)

Conclusions

For many applications, Gallager codes are best.