united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

SMR.1347 - 5

WORKSHOP ON STATISTICAL PHYSICS AND CAPACITY-APPROACHING CODES

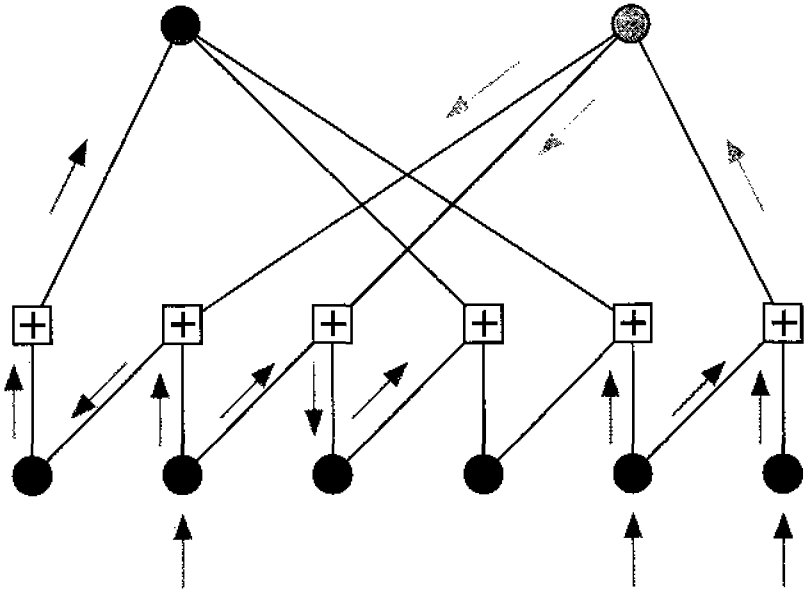
APPROACHING THE SHANNON LIMIT: A CODING THEORIST'S PROGRESS REPORT

R.J. McELIECE
California Institute of Technology
Department of Electrical Engineering
91125 Pasadena, California, U.S.A.

Please note: These are preliminary notes intended for internal distribution only.

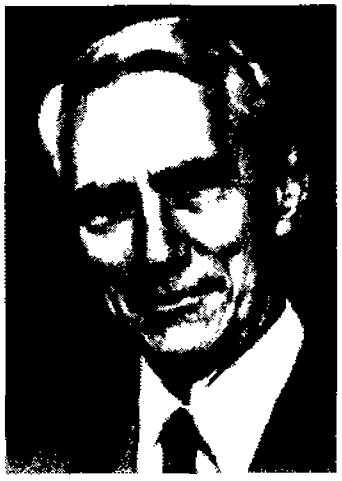
Approaching the Shannon Limit: A Coding Theorist's Progress Report

Robert J. McEliece
California Institute of Technology
Pasadena, California, USA



The Abdus Salam International Centre for Theoretical Physics
Grignano, Italy, May 21, 2001

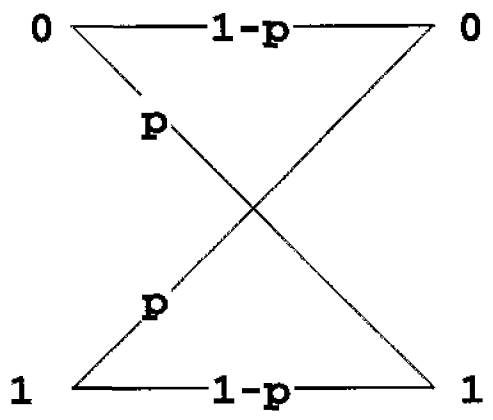
(A Tale of Two Claudes)



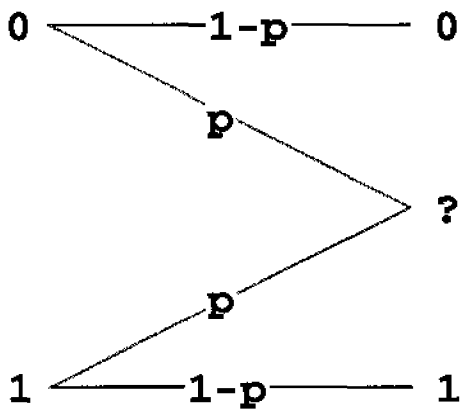
Shannon

Berrou

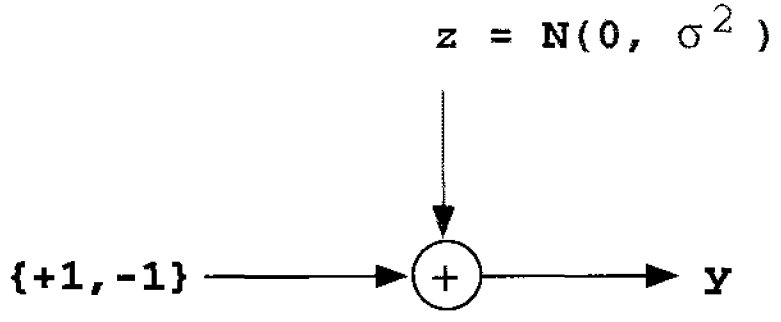
Three Common Channel Models



Binary Symmetric



Binary Erasure



Binary Input, Additive Gaussian Noise

Shannon's Channel Coding Theorem

Theorem: **For any (discrete-input memoryless) channel, there exists a number C, the channel capacity, such that for any desired data rate $R < C$ and any desired error probability $\pi > 0$, it is possible to design an encoder-decoder pair that permits the transmission of data over the channel at rate R and decoded error probability $< \pi$.**

How Hard is it to Achieve Channel Capacity?

- Channel capacity $= C$.
- Desired code rate $= C(1-\epsilon)$. For example, $\epsilon = .01$ means $R = .99C$.
- Desired decoder error probability $= \pi$. For example, $\pi = 10^{-7}$.
- $\chi_E(\epsilon, \pi) =$ the *encoding* complexity, operations *per information bit.*
- $\chi_D(\epsilon, \pi) =$ the *decoding* complexity, operations *per information bit.*

We are interested in the behavior of $\chi_E(\epsilon, \pi)$, *and especially* $\chi_D(\epsilon, \pi)$, *for fixed* π, *as* $\epsilon \to 0$.

The Classical Results.

Theorem A. *On a discrete memoryless channel of capacity C, for any fixed $\pi > 0$, for the Shannon-Gallager ensemble of rate $R = C(1-\epsilon)$, as $\epsilon \to 0$,*

$$\overline{\chi}_E(\epsilon, \pi) = O(1/\epsilon^2)$$
$$\overline{\chi}_D(\epsilon, \pi) = 2^{O(1/\epsilon^2)}.$$

The Classical Results.

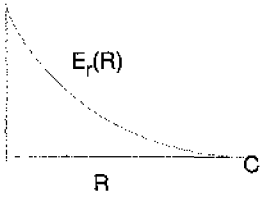
Theorem A. *On a discrete memoryless channel of capacity C, for any fixed $\pi > 0$, for the Shannon-Gallager ensemble of rate $R = C(1-\epsilon)$, as $\epsilon \to 0$,*

$$\overline{\chi}_E(\epsilon, \pi) = O(1/\epsilon^2)$$

$$\overline{\chi}_D(\epsilon, \pi) = 2^{O(1/\epsilon^2)}.$$

Proof: Use linear codes with (per-bit) encoding complexity $O(n)$, and ML decoding with complexity $2^{O(n)}$. To estimate n, use the random coding exponent:

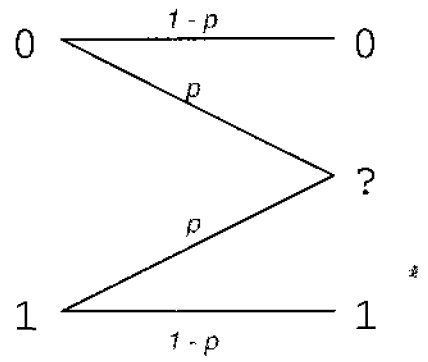
$$\overline{\pi} \leq e^{-nE_r(R)}$$



and the fact that

$$E_r(C(1-\epsilon)) \approx K\epsilon^2 \qquad \text{as } \epsilon \to 0. \qquad \blacksquare$$

An Improvement for the Binary Erasure Channel.



Theorem B. *For the binary erasure channel,* $\overline{\chi}_D$ *can be improved to*

$$\overline{\chi}_D(\epsilon, \pi) = O(1/\epsilon^4),$$

for fixed π*, as* $\epsilon \to 0$.

Proof: Decode with (per-bit) complexity $O(n^2)$ by solving linear equations for the erased positions. ∎

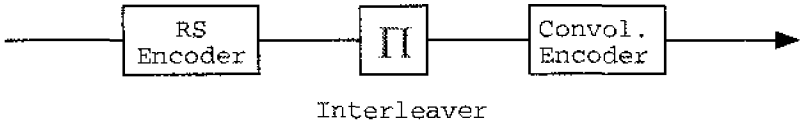
The Central Problem of Channel Coding:

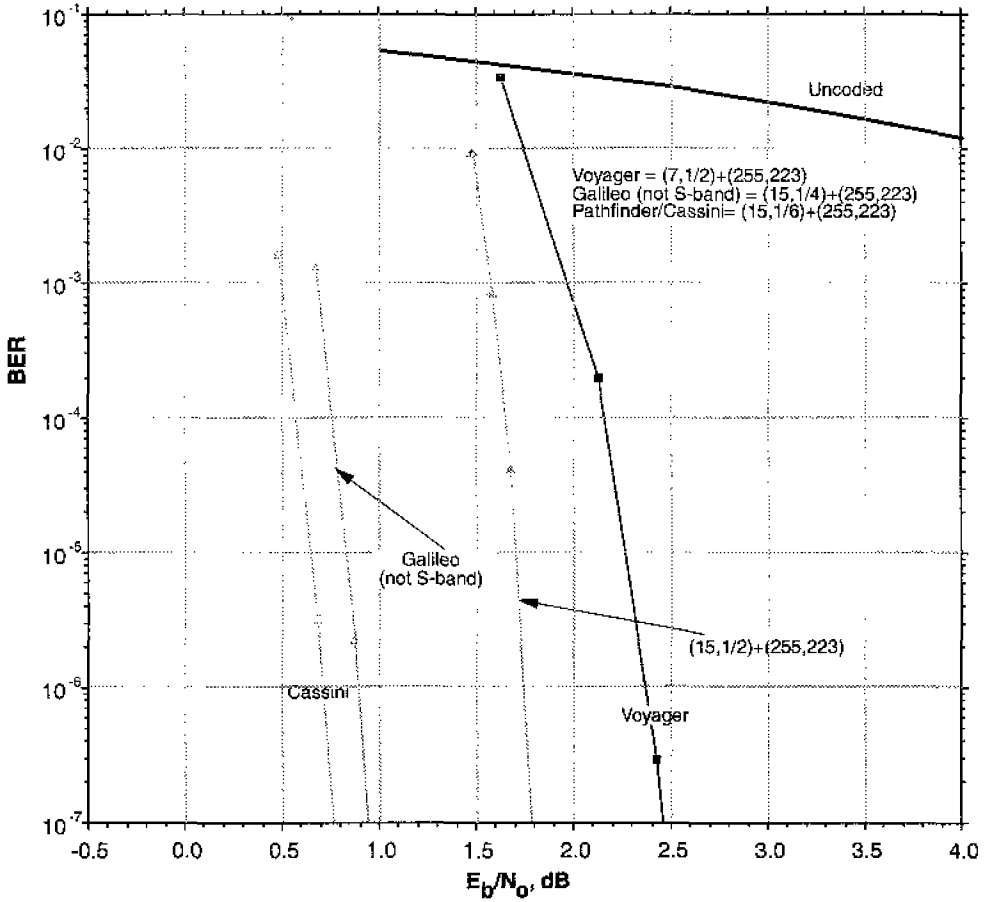
To find near Shannon-limit codes with practical encoding and decoding algorithms.

Pre-1993 Highlights:

- Algebraic Block Codes: BCH, Reed-Solomon, Algebraic-geometry: (Deep theory, leading to easy encoding and decoding, but suited primarily to storage applications, and do not approach capacity.)
- Convolutional Codes with Viterbi decoding: (System-theoretic approach, superior to algebraic codes in most transmission applications, but still far from Shannon limit.)
- Concatenated Codes (Forney): Hybrid approach, combining algebraic block codes with convolutional codes.

Pre-1993 State of the Art on the AWGN Channel





1993: And Then Came...

NEAR SHANNON LIMIT ERROR - CORRECTING CODING AND DECODING : TURBO-CODES (1)

Claude Berrou, Alain Glavieux and Punya Thitimajshima

Claude Berrou, Integrated Circuits for Telecommunication Laboratory

Alain Glavieux and Punya Thitimajshima, Digital Communication Laboratory

Ecole Nationale Supérieure des Télécommunications de Bretagne, France

(1) Patents N° 9105279 (France), N° 92460011.7 (Europe), N° 07/870,483 (USA)

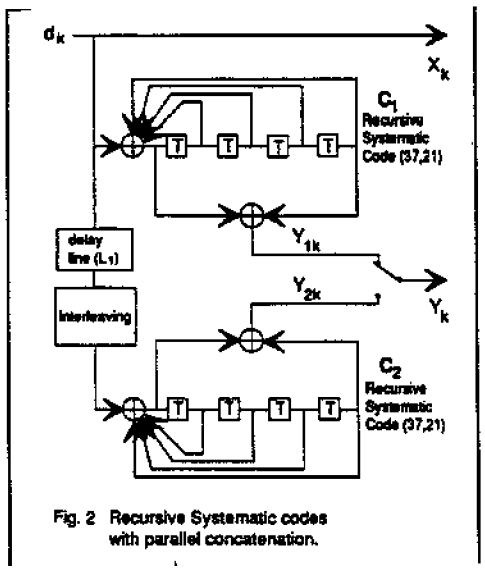


Fig. 2 Recursive Systematic codes with parallel concatenation.

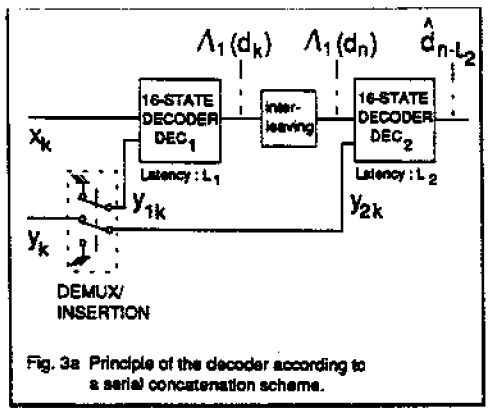
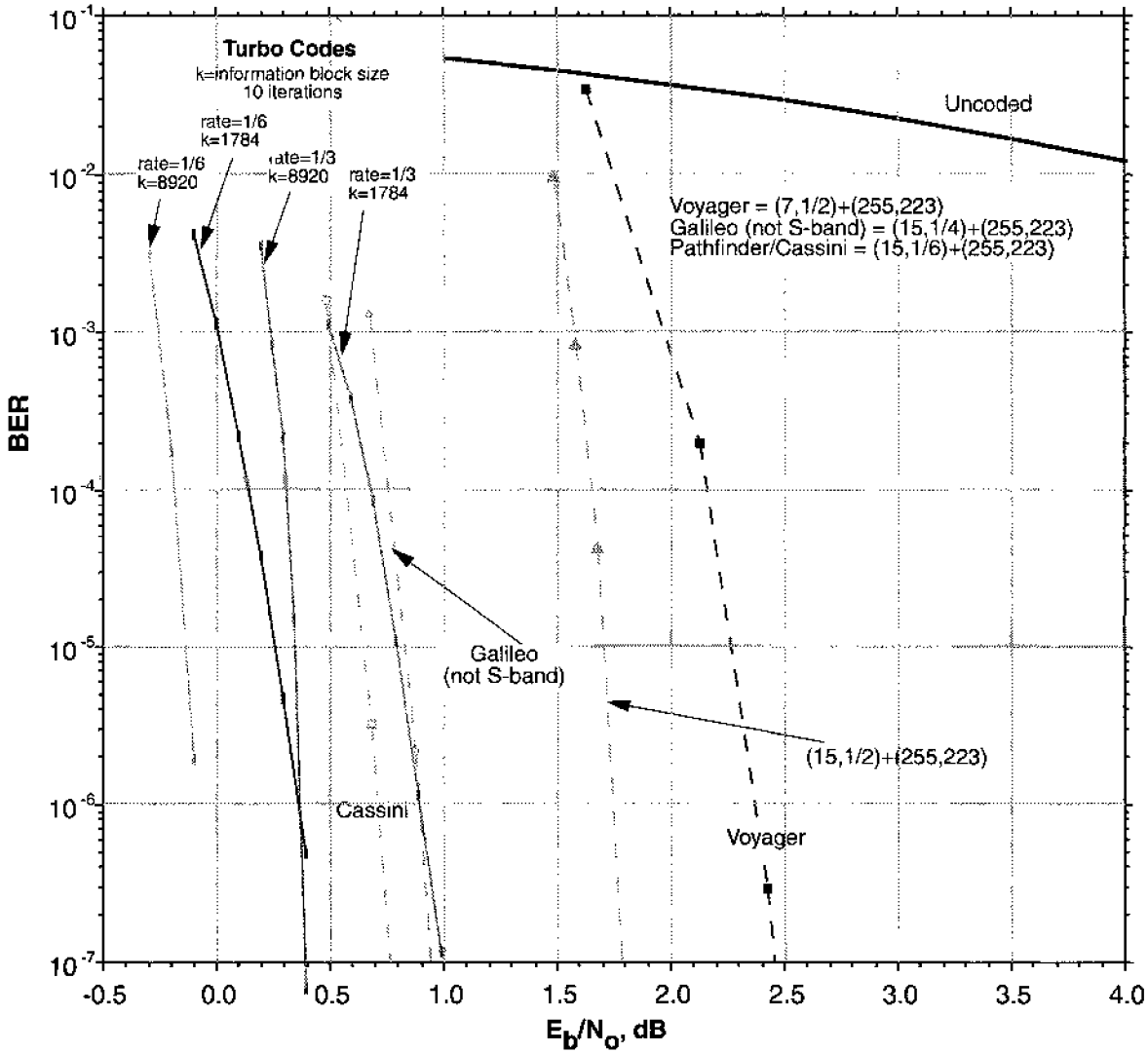


Fig. 3a Principle of the decoder according to a serial concatenation scheme.

The Turbo-Era State of the Art on the AWGN Channel



1997: Another Landmark Paper (For the BEC)

Practical Loss-Resilient Codes

Michael G. Luby* Michael Mitzenmacher† M. Amin Shokrollahi‡

Daniel A. Spielman§ Volker Stemann¶

Abstract

We present randomized constructions of linear-time encodable and decodable codes that can transmit over lossy channels at rates extremely close to capacity. The encoding and decoding algorithms for these codes have fast and simple software implementations. Partial implementations of our algorithms are faster by orders of magnitude than the best software implementations of any previous algorithm for this problem. We expect these codes will be extremely useful for applications such as real-time audio and video transmission over the Internet, where lossy channels are common and fast decoding is a requirement.

Despite the simplicity of the algorithms, their design and analysis are mathematically intricate. The design requires the ficients determined by the graph structure. Based on these polynomials, we design a graph structure that guarantees successful decoding with high probability.

1 Introduction

Studies show that the Internet exhibits packet loss, and the measurements in [10] show that the situation has become worse over the past few years. A standard solution to this problem is to request retransmission of data that is not received. When some of this retransmission is lost, another request is made, and so on. In some applications, this introduces technical difficulties. For real-time transmission this solution can lead to unacceptable delays caused by several rounds of communication between sender and receiver. For

1963: The Grandaddy of Them All:

LOW-DENSITY PARITY-CHECK CODES

PUBLISHED 1963 BY THE M.I.T. PRESS, CAMBRIDGE, MASSACHUSETTS

ROBERT G. GALLAGER

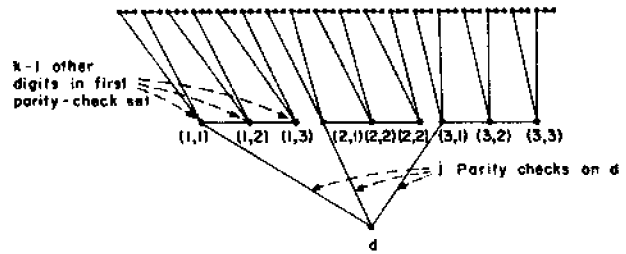


Figure 4.1. Parity-check set tree.

Assume now that both digit d and several of the digits in the first tier are transmission errors. Then on the first decoding attempt, the error-free digits in the second tier and their parity-check constraints will allow correction of the errors in the first tier. This in turn will allow correction of digit d on the second decoding attempt. Thus digits and parity-check equations can aid in decoding a digit seemingly unconnected with them. The probabilistic decoding scheme to be described next utilizes these extra digits and extra parity-check equations more systematically.

4.2 Probabilistic Decoding

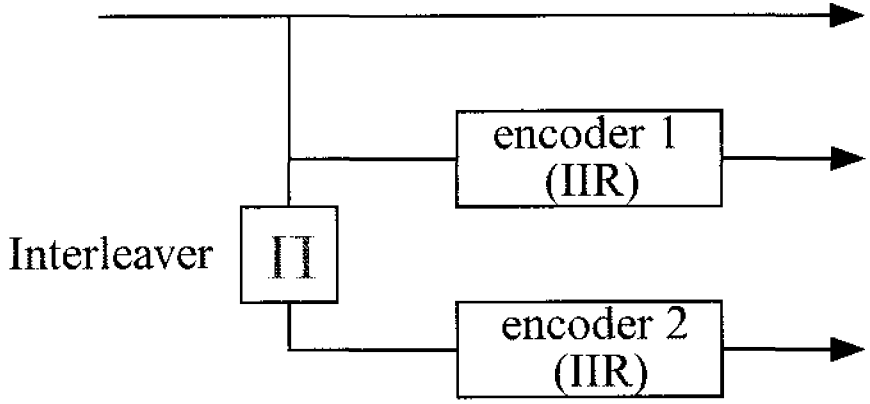
Assume that the code words from an (n, j, k) code are used with equal probability on an arbitrary binary-input channel. For any digit d, using the notation of Figure 4.1, an iteration process will be derived that on the m^{th} iteration computes the probability that the transmitted digit in position d is a 1 conditional on the received symbols out to and including the m^{th} tier. For the first iteration, we can consider digit d and the digits in the first tier to form a subcode in which all sets of these digits that satisfy the j parity-check equations in the tree have equal probability of transmission.*

The Secret is Iterative
Message-passing (Turbo) Decoding

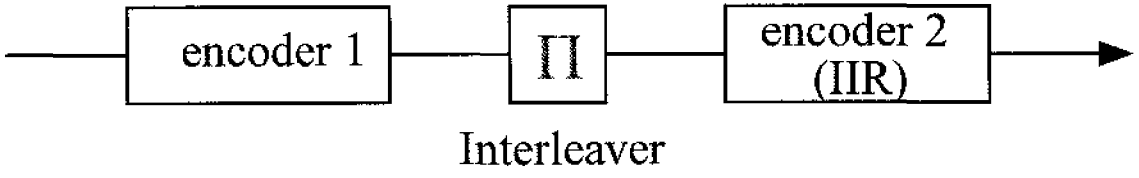
A Low-Complexity Approximation To "Exact" Decoding.

Codes that can be Decoded in the "Turbo-Style"

- Classical turbo codes:

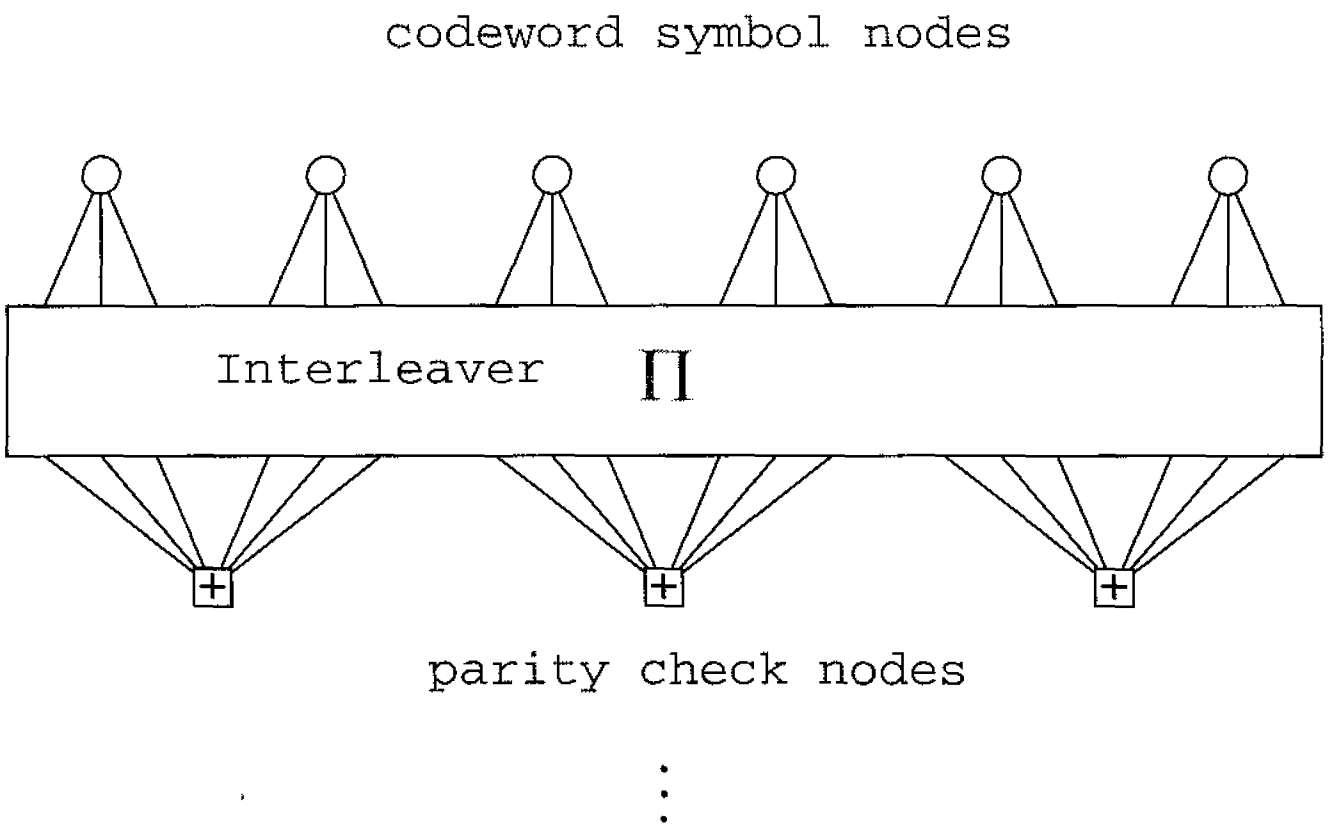


- "Serial" turbo codes:



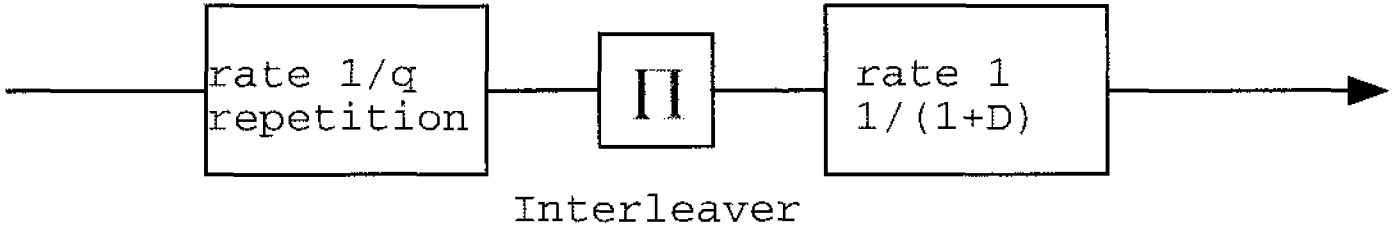
Codes that can be Decoded in the "Turbo-Style"

• Gallager codes (Low-Density Parity-Check), regular and irregular:



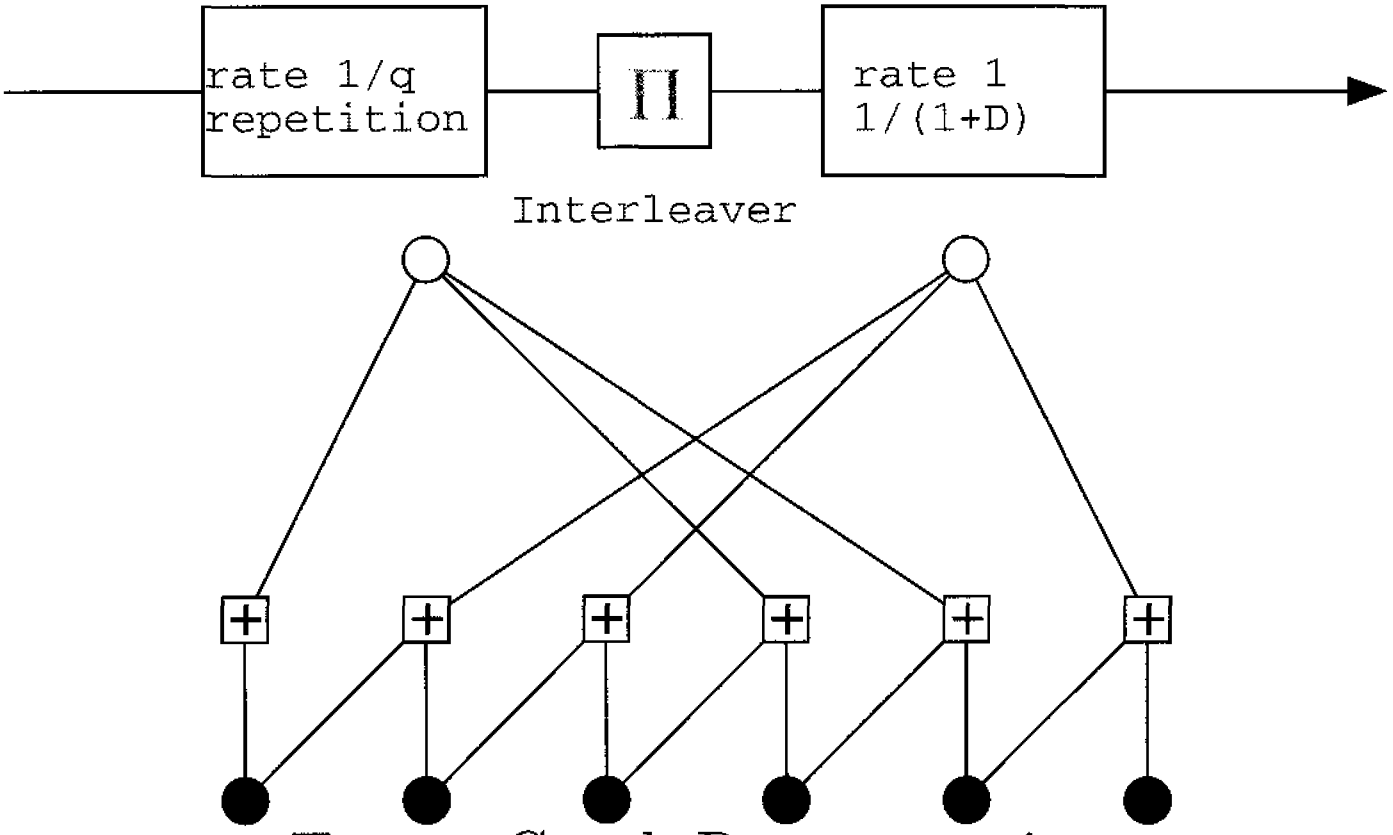
Repeat-Accumulate (RA) Codes

(nonsystematic)



Repeat-Accumulate (RA) Codes

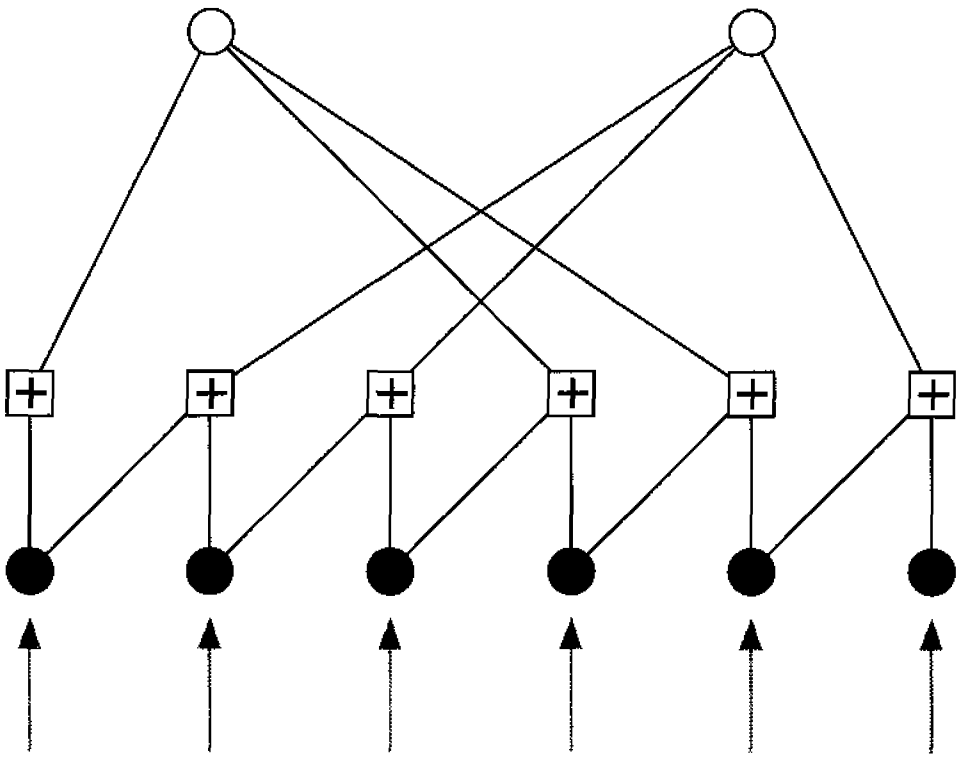
(nonsystematic)



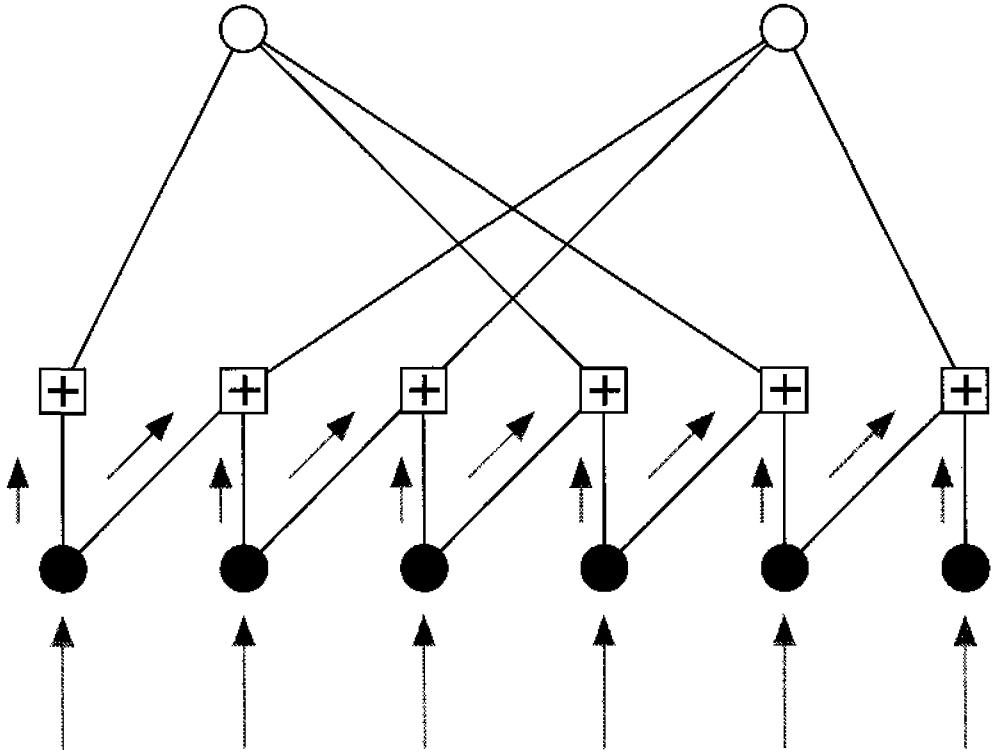
Tanner Graph Representation

$(k = 2,\ q = 3)$

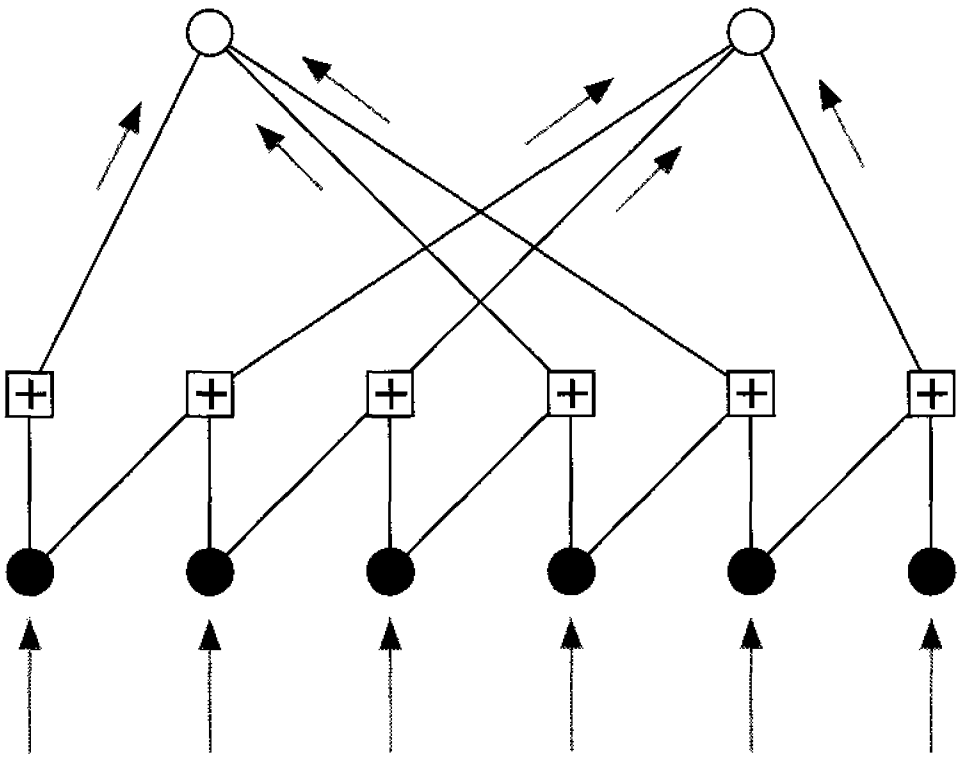
Decoding an RA Code Using Message Passing



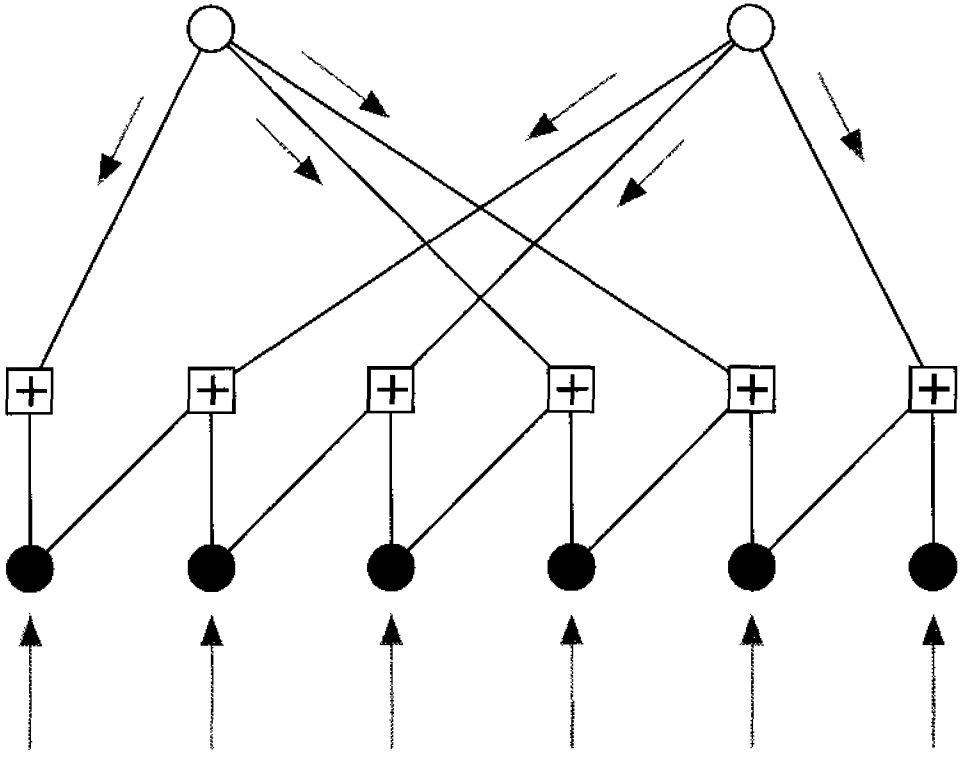
Decoding an RA Code Using Message Passing



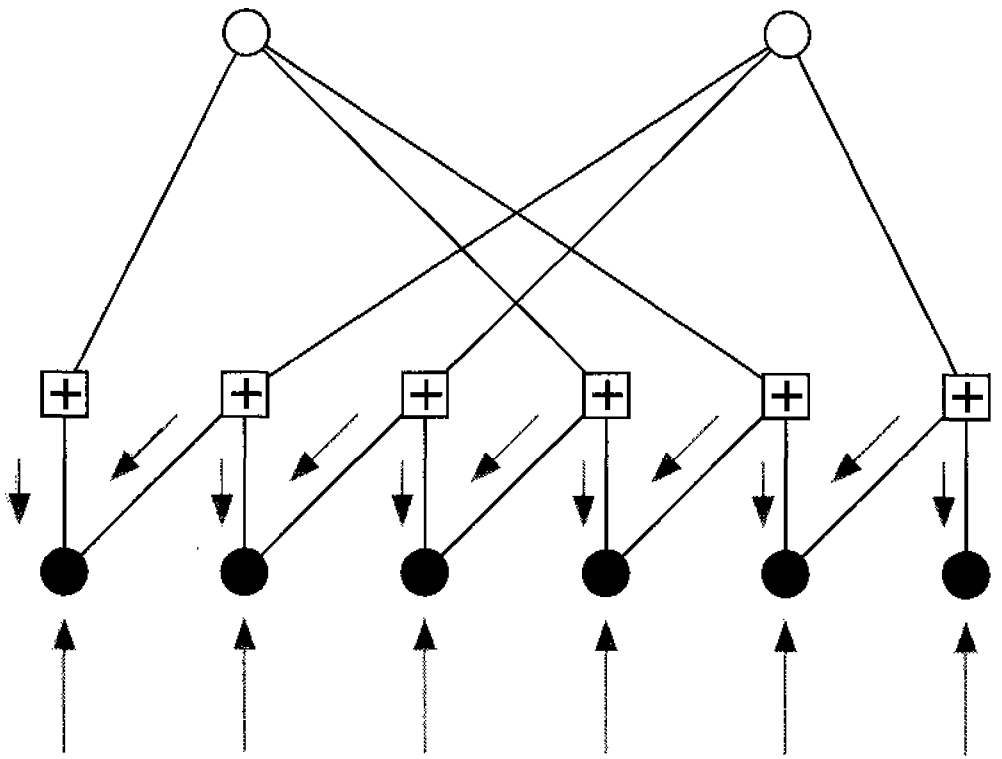
Decoding an RA Code Using Message Passing



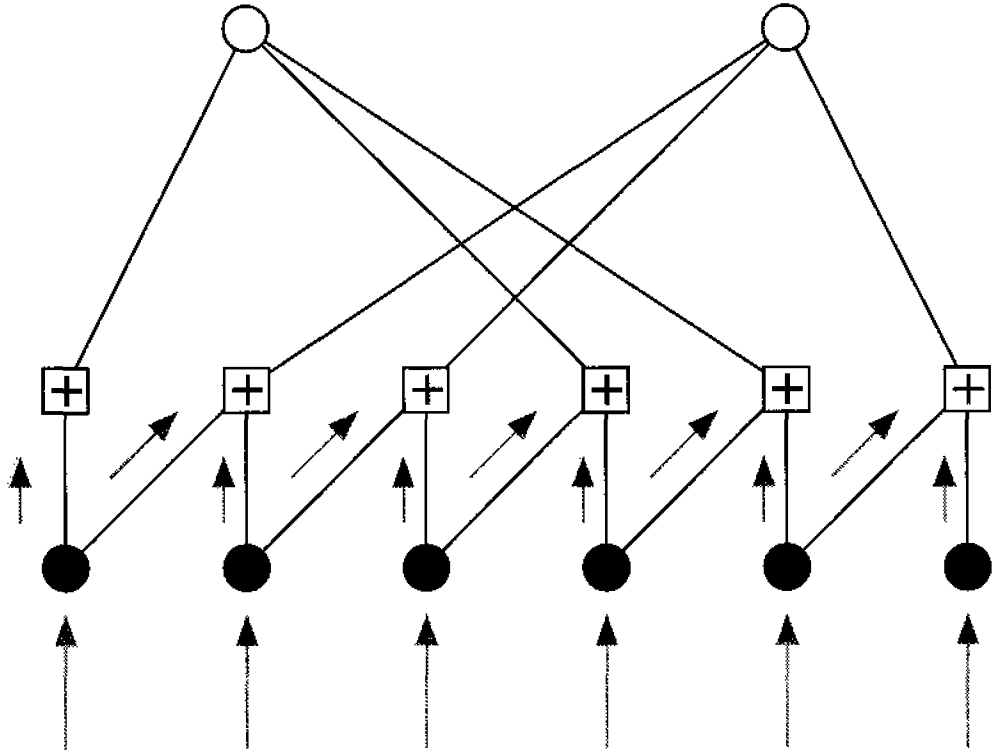
Decoding an RA Code Using Message Passing



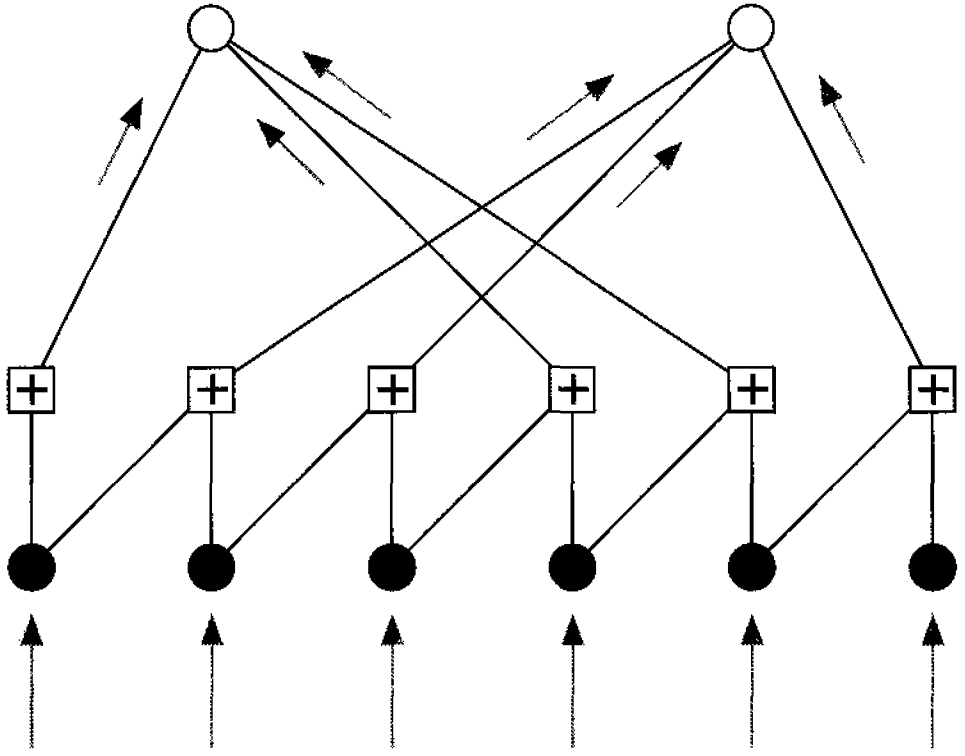
Decoding an RA Code Using Message Passing



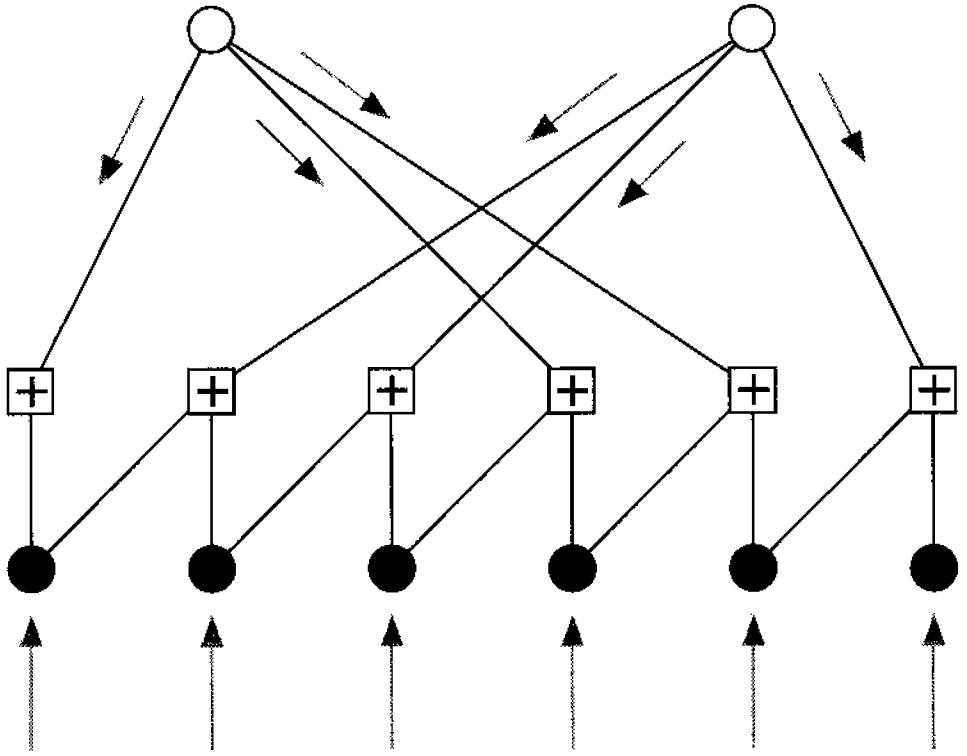
Decoding an RA Code Using Message Passing



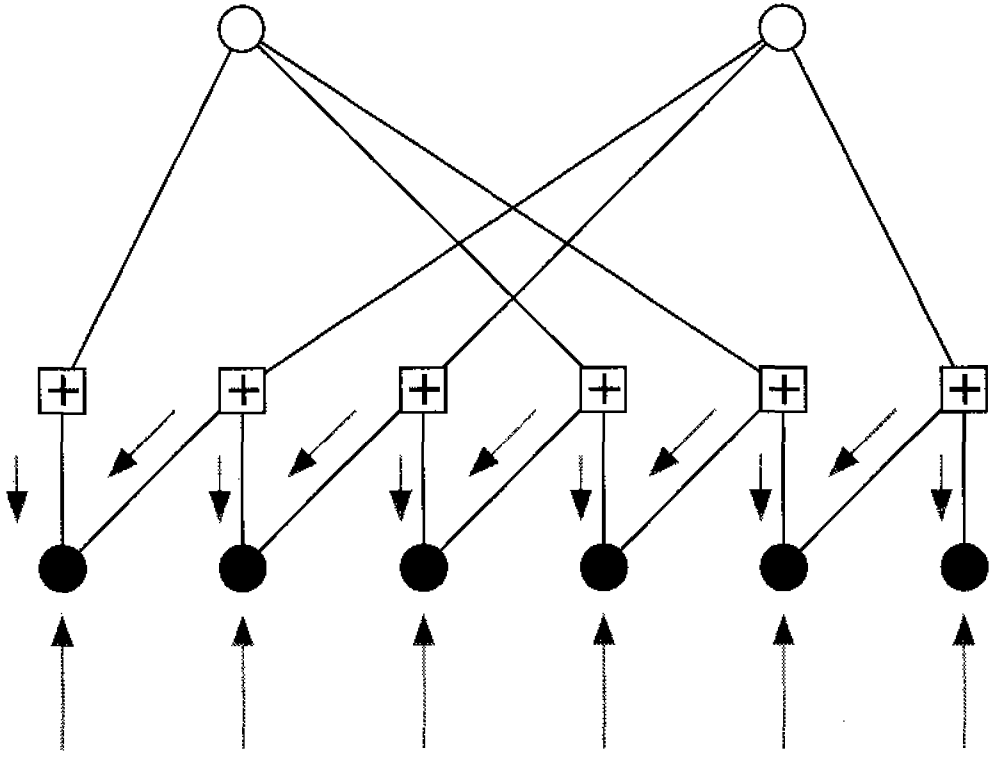
Decoding an RA Code Using Message Passing



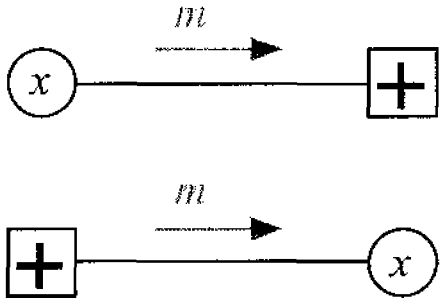
Decoding an RA Code Using Message Passing



Decoding an RA Code Using Message Passing

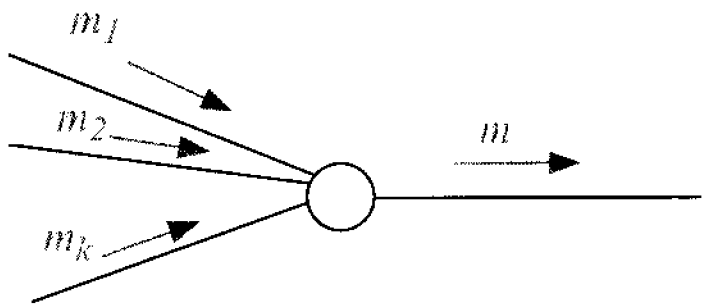


What are the Messages?



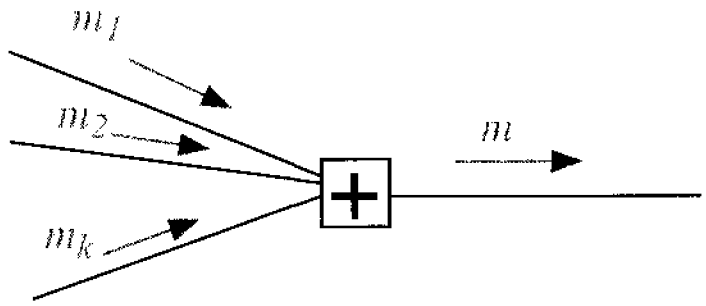
$$m = \log \frac{p(x=0)}{p(x=1)}.$$

How Messages are Updated At Variable Nodes



$$m = m_1 + m_2 + \cdots + m_k.$$

How Messages are Updated At Check Nodes

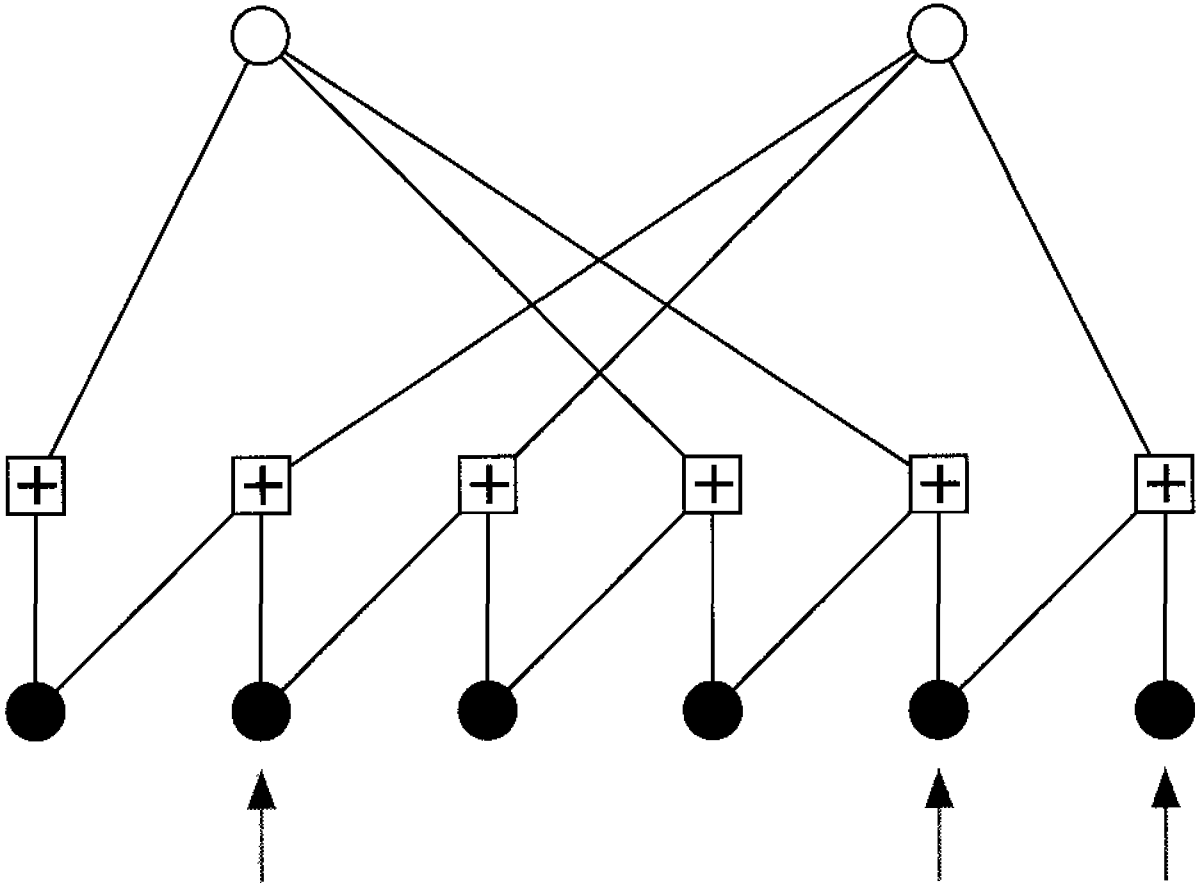


$$m = m_1 \boxplus m_2 \boxplus \cdots \boxplus m_k$$

$$\tanh(\frac{m}{2}) = \tanh(\frac{m_1}{2}) \tanh(\frac{m_2}{2}) \cdots \tanh(\frac{m_k}{2})$$

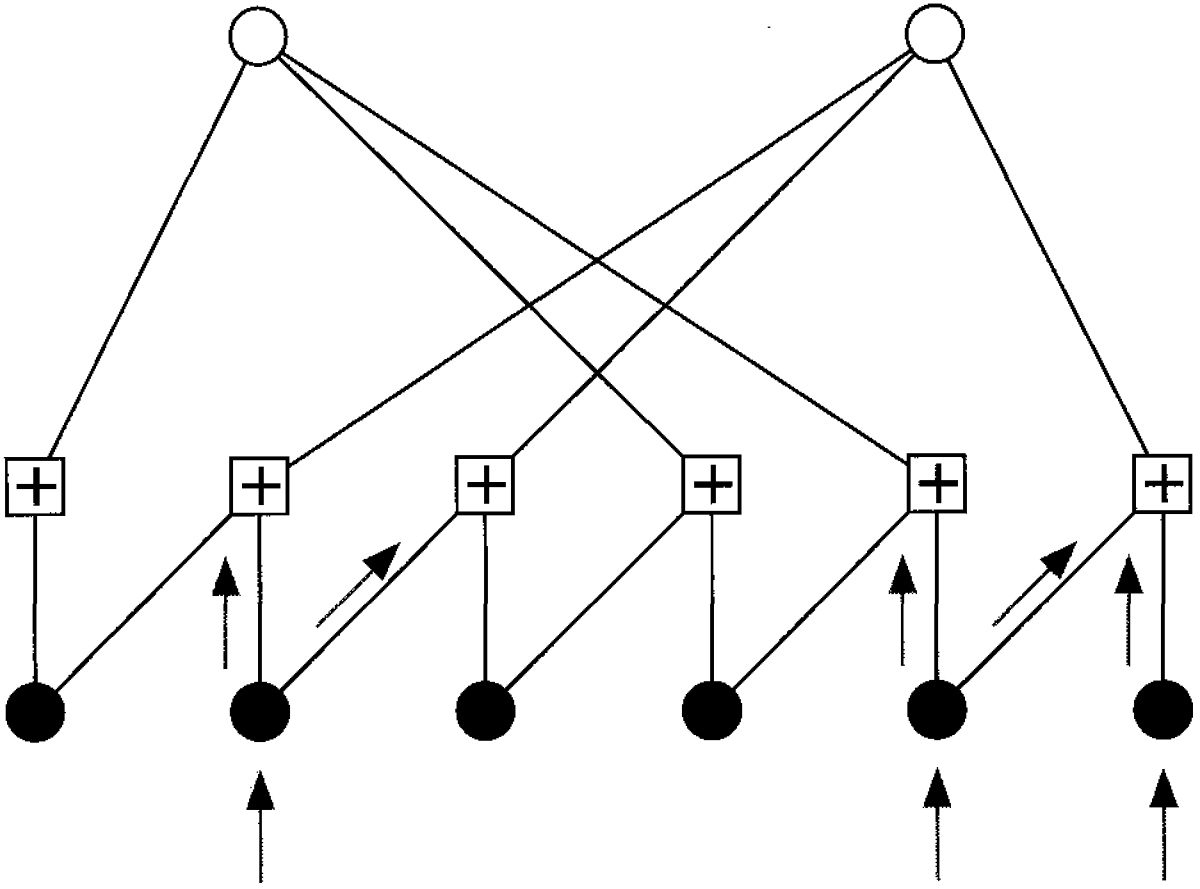
Decoding an RA Code on the BEC Using Message Passing

$\longrightarrow = 1, \dashrightarrow = 0$



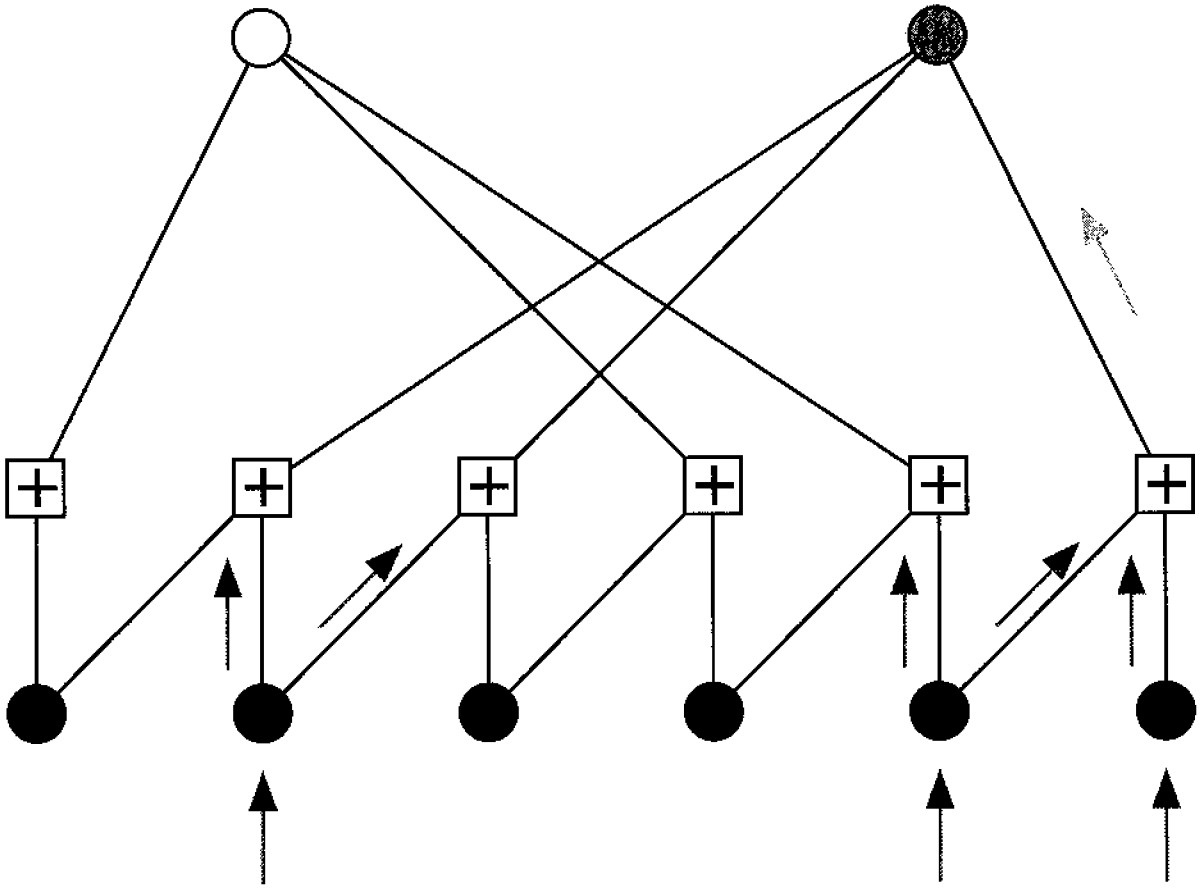
Decoding an RA Code on the BEC Using Message Passing

$\longrightarrow = 1, \longrightarrow = 0$



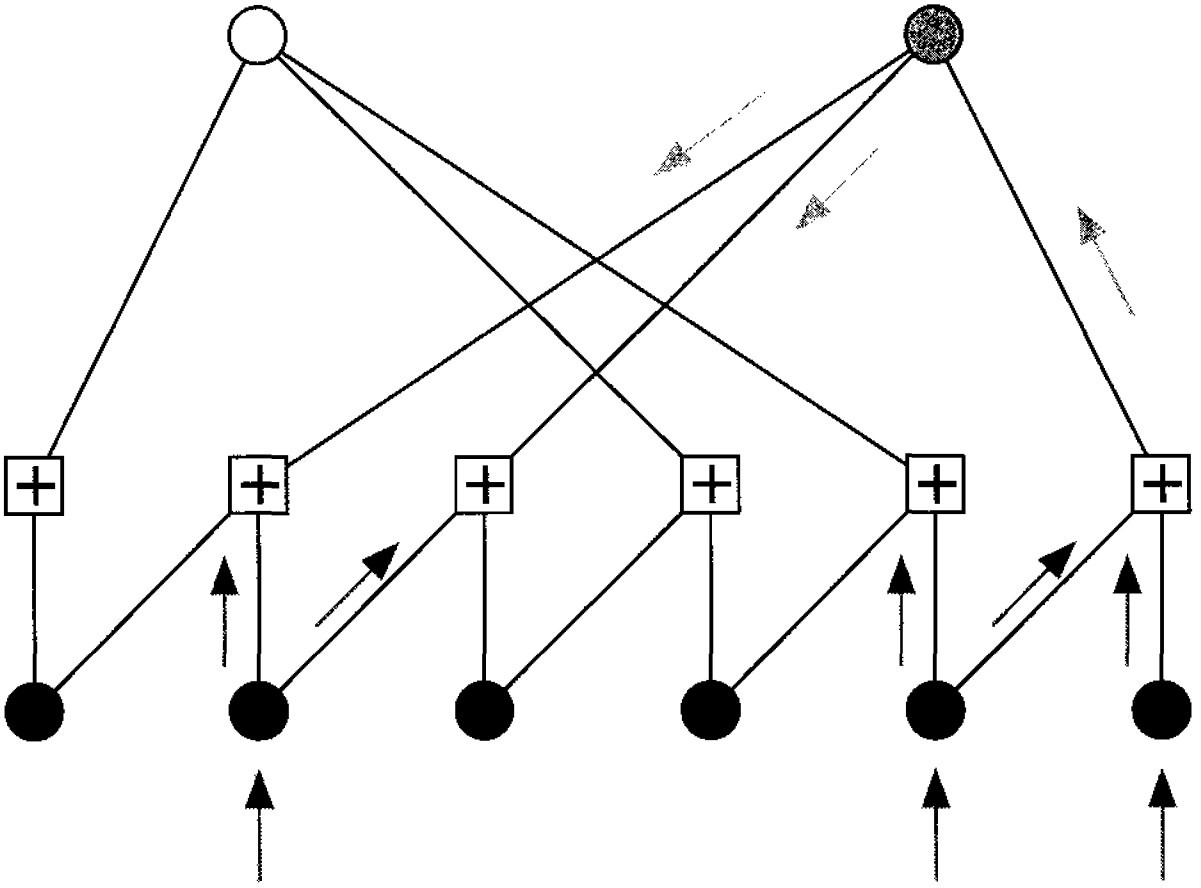
Decoding an RA Code on the BEC Using Message Passing

$\longrightarrow = 1, \quad \cdots\!\!\rightarrow = 0$



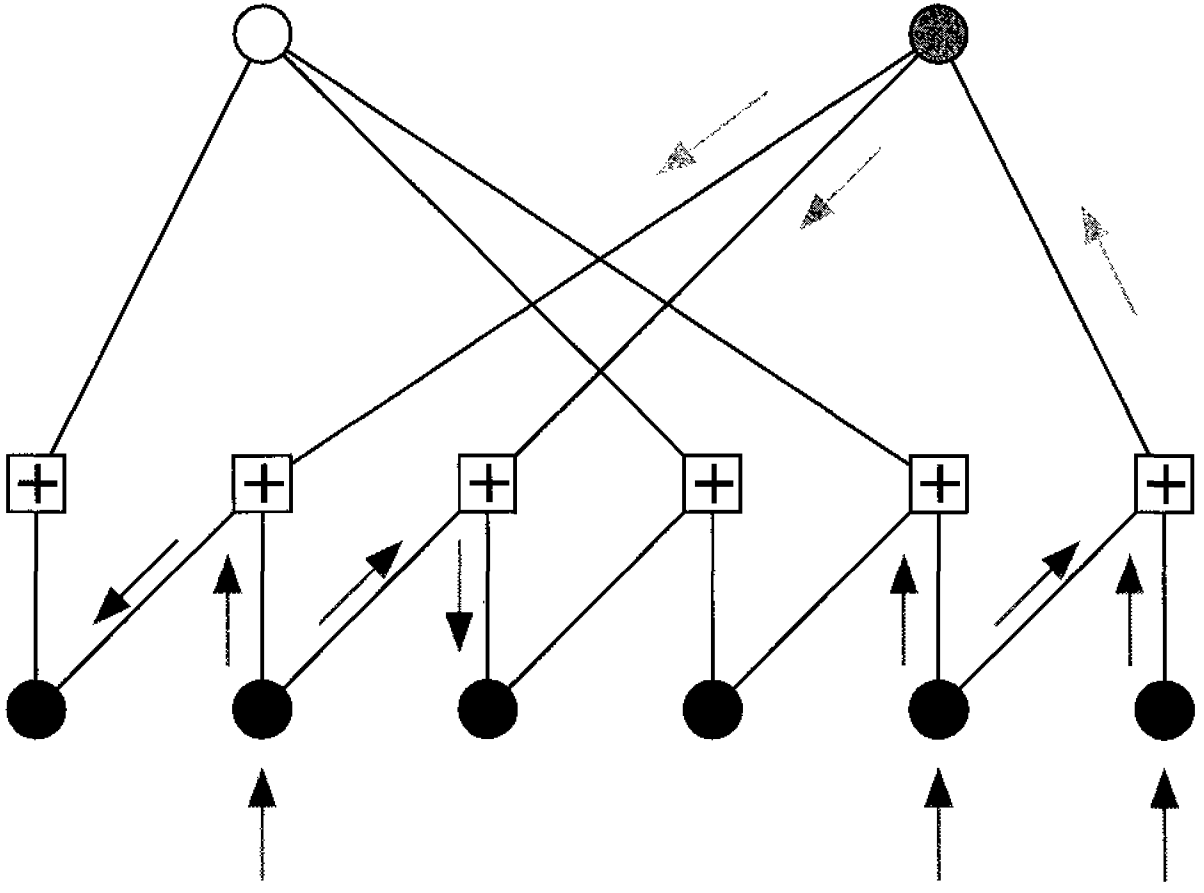
Decoding an RA Code on the BEC Using Message Passing

$\longrightarrow = 1$, $\dashrightarrow = 0$



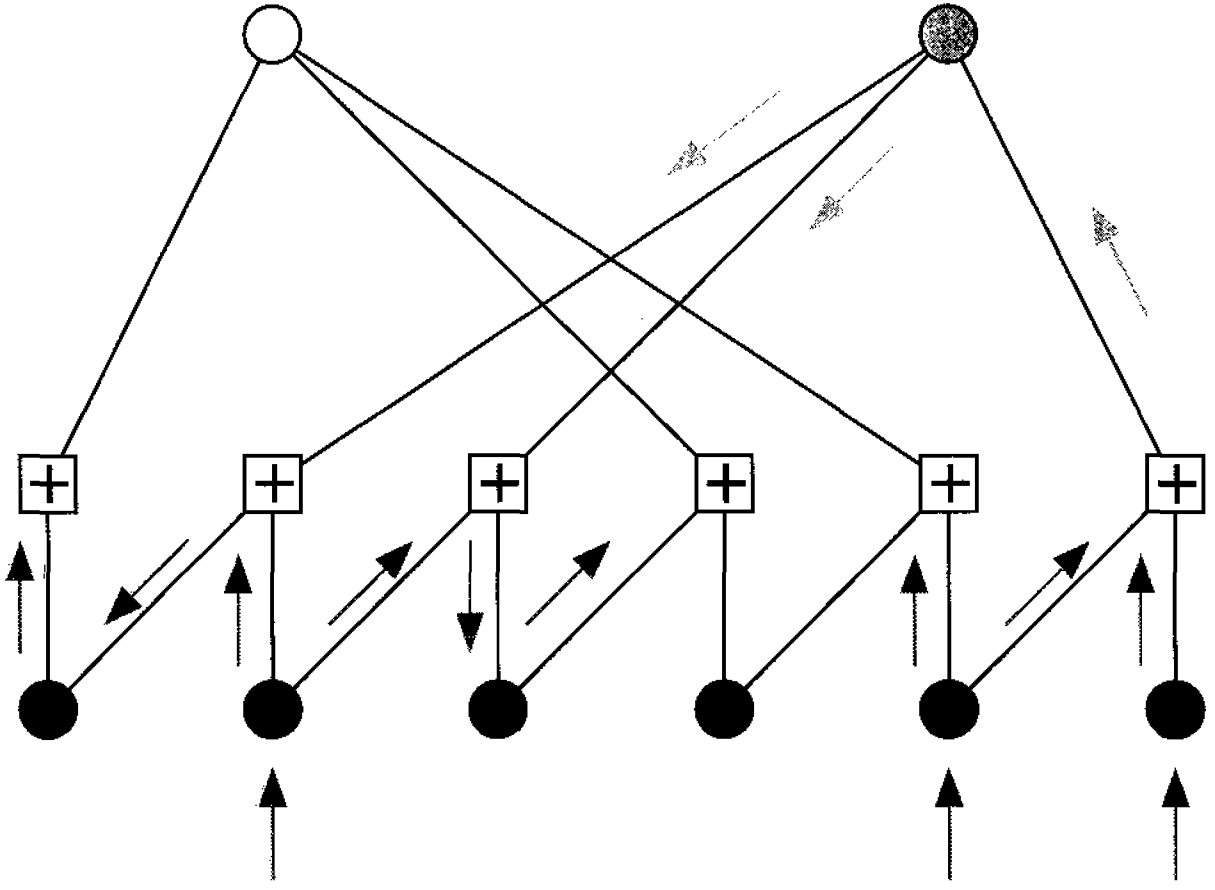
Decoding an RA Code on the BEC Using Message Passing

$\longrightarrow = 1, \longrightarrow = 0$



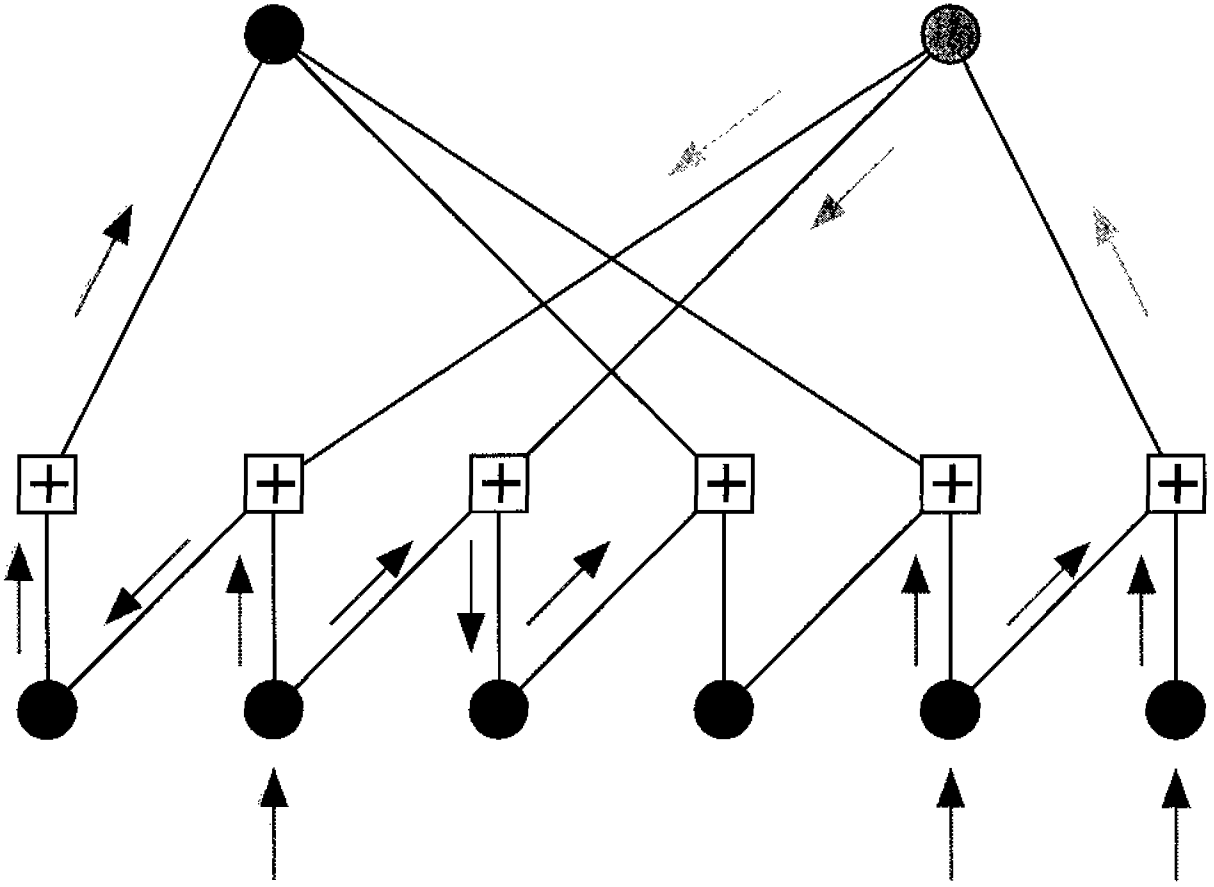
Decoding an RA Code on the BEC Using Message Passing

$\longrightarrow = 1$, $\dashrightarrow = 0$



Decoding an RA Code on the BEC Using Message Passing

$\longrightarrow = 1$, $\longrightarrow = 0$



What is the Complexity of Iterative Message-Passing Decoding?

- Complexity *per iteration*:

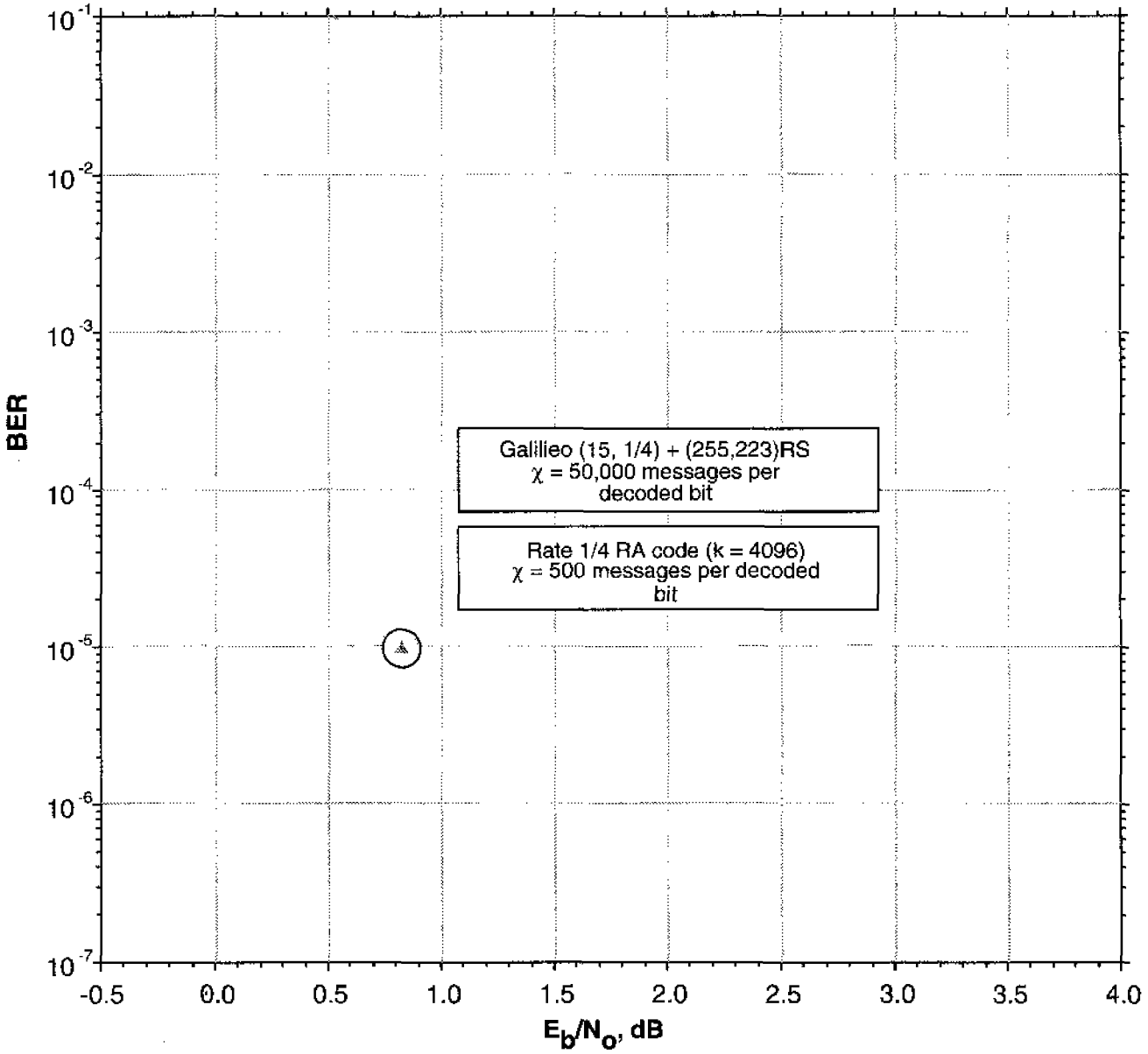
$$\chi_{IT} = 2\frac{E}{k},$$

where E is the number of edges in the Tanner graph, and k is the number of information bits (χ_{IT} is an ensemble invariant).

- $N(\epsilon, \pi)$ = Number of iterations needed to achieve error probability π.

$$\boxed{\chi_D(\epsilon, \pi) = \chi_{IT} \cdot N(\epsilon, \pi).}$$

One Interesting Point



Theory is Available!

The Capacity of Low-Density Parity Check Codes under Message-Passing Decoding

Tom Richardson, Rüdiger Urbanke
Bell Labs, Lucent Technologies
Murray Hill, NJ 07974

November 7, 1998

Design of Provably Good Low-Density Parity Check Codes

Tom Richardson, Amin Shokrollahi and Rüdiger Urbanke
Bell Labs, Lucent Technologies
Murray Hill, NJ 07974

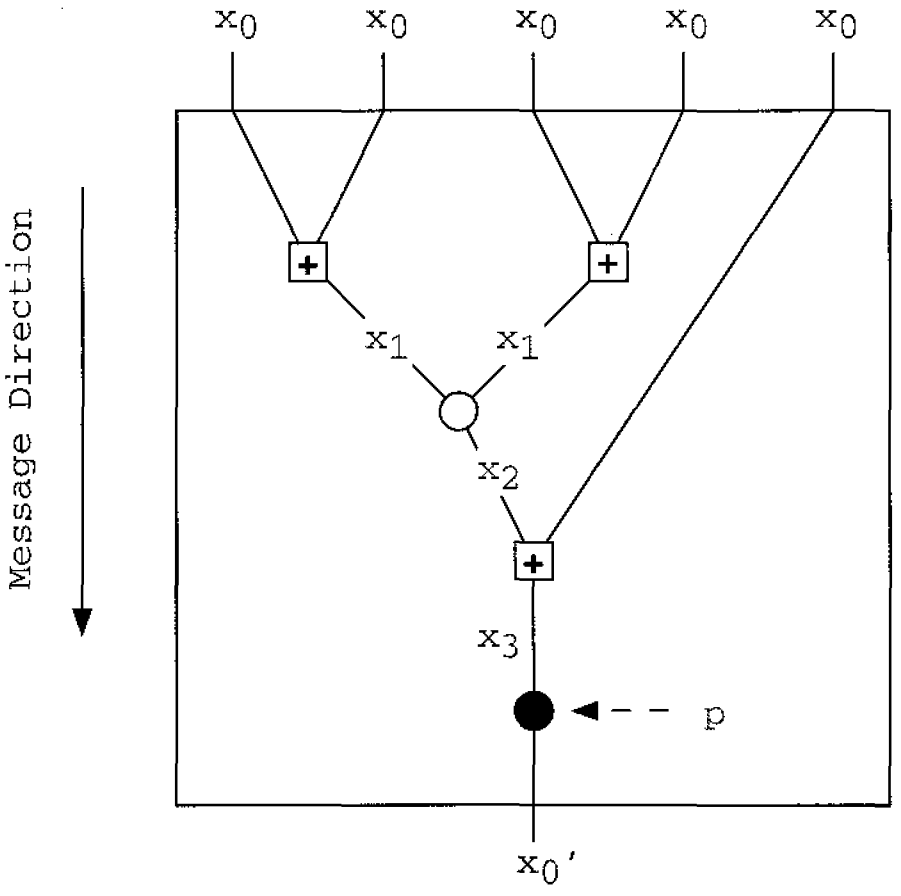
April 5, 1999

Analysis and Design of RA Codes
(The Fine Print)

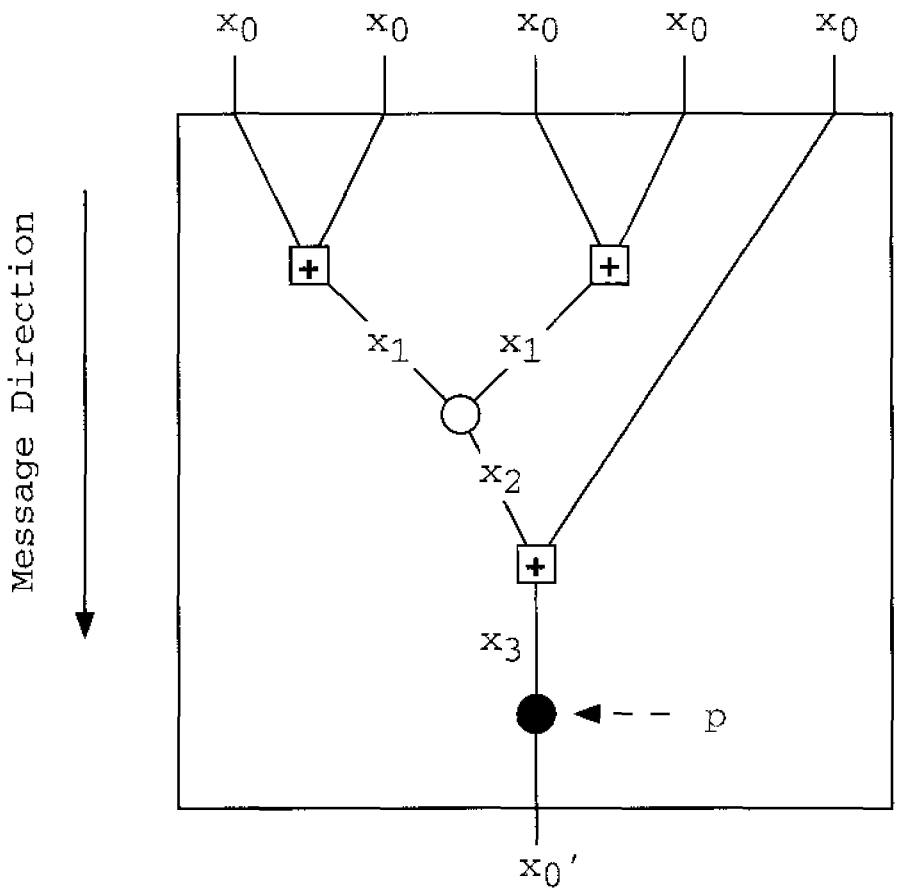
• The ensemble of RA codes satisfies the **RU condition:** For any fixed L, the probability that the depth-L neighborhood of a randomly selected edge contains a cycle goes to zero as $k \to \infty$.

• Therefore L-fold *density evolution* gives the limiting value ($k \to \infty$) of the ensemble bit error probability after L iterations. This limiting value will depend on the "noise parameter" of the channel. The largest noise parameter for which the limiting bit error probability is zero is called the *ensemble noise threshold.*

The Computation Tree for a $q = 3$ RA Code.

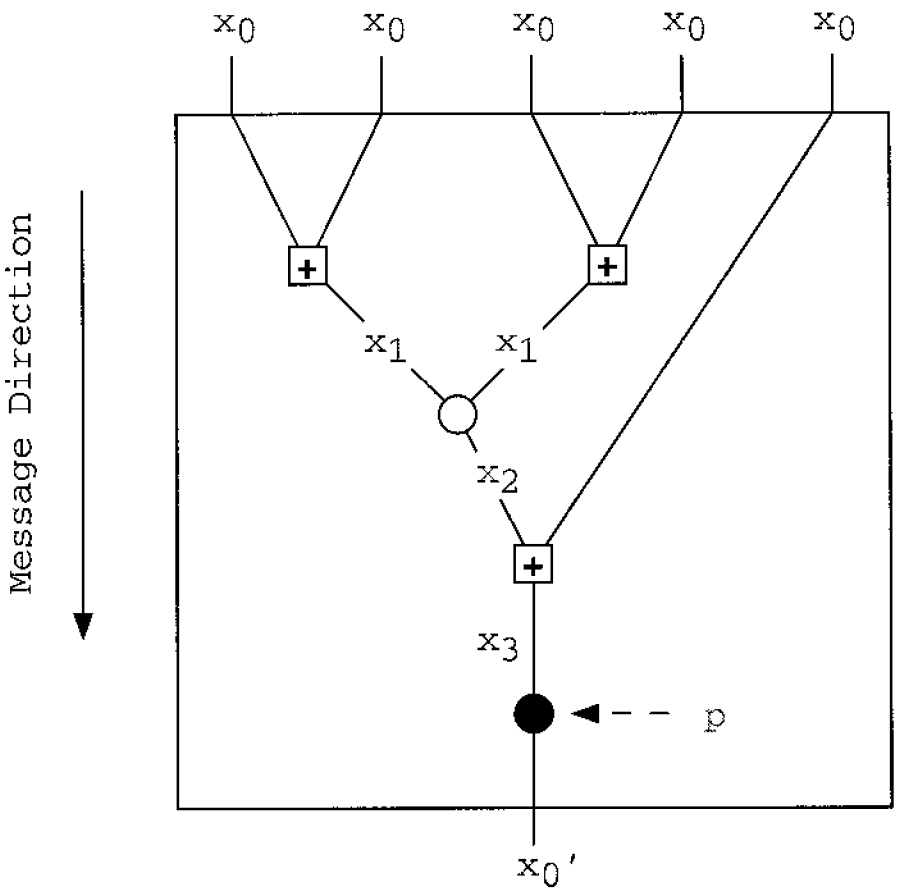


Density Evolution for a $q = 3$ RA Code.



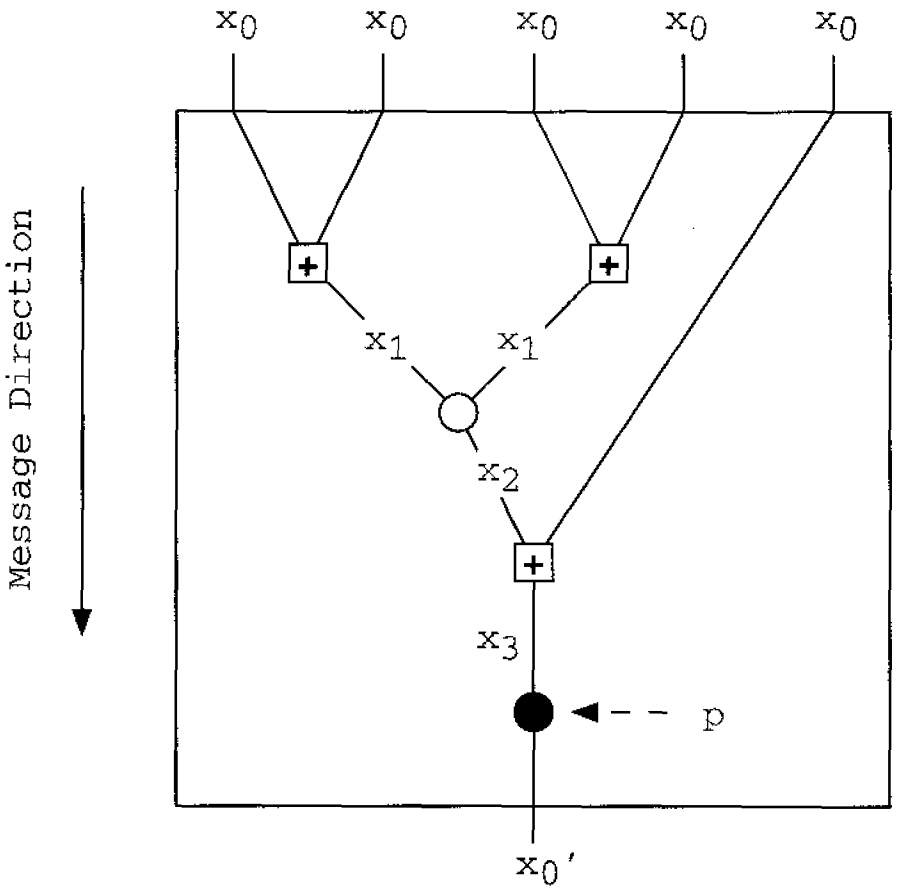
$$x_1 = 1 - (1 - x_0)^2.$$

Density Evolution for a $q = 3$ RA Code.



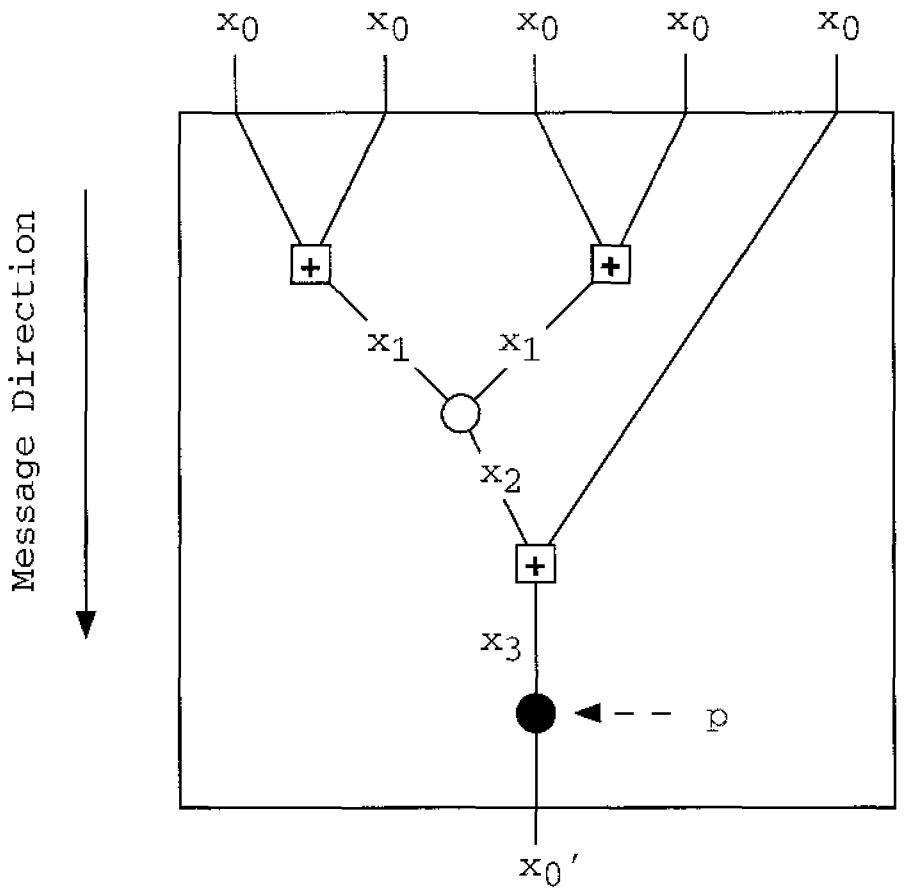
$$x_2 = x_1^{q-1}.$$

Density Evolution for a $q = 3$ RA Code.



$$x_3 = 1 - (1 - x_0)(1 - x_2).$$

Density Evolution for a $q = 3$ RA Code.



$$x_0' = x_3 p.$$

Summary:

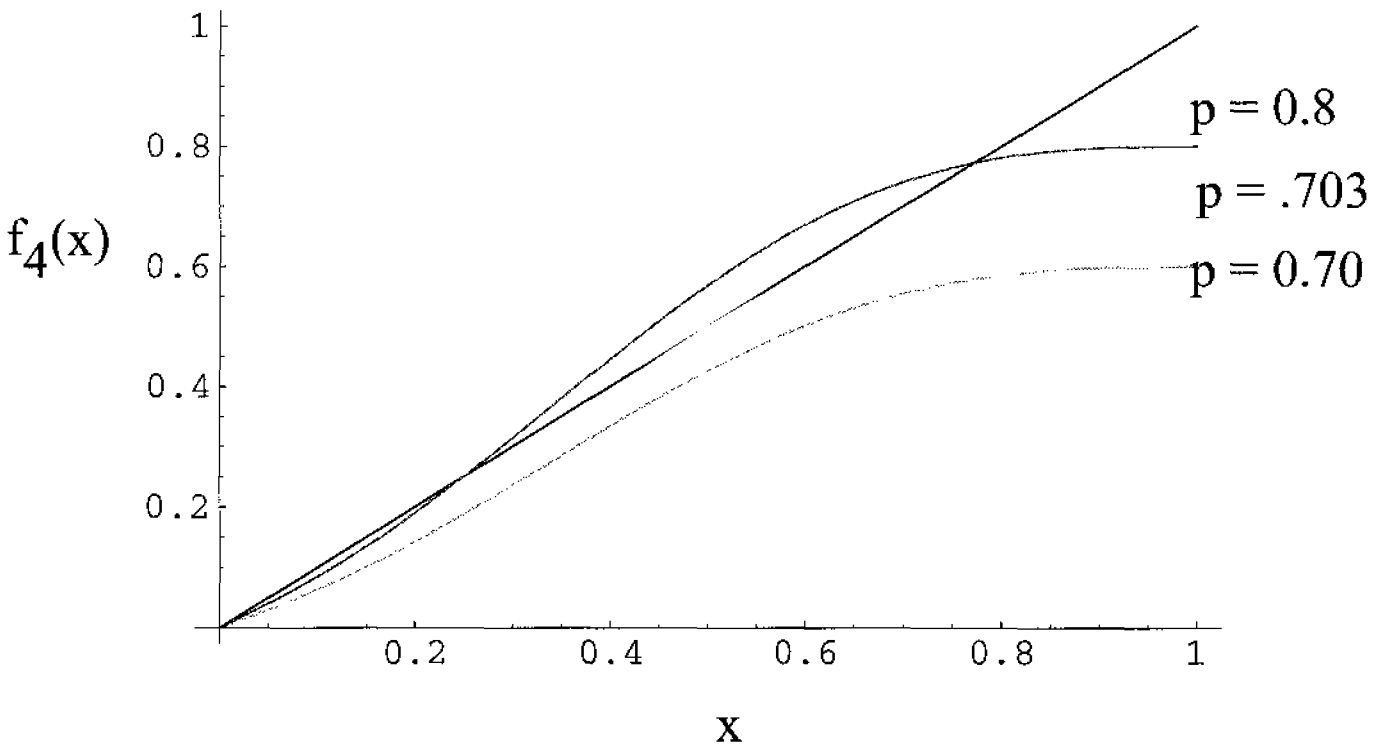
If $x_i^{(L)}$ denotes the value of $x_i t$ on the Lth iteration, then

$$x_0^{(L+1)} = f_q(x_0^{(L)}),$$

where

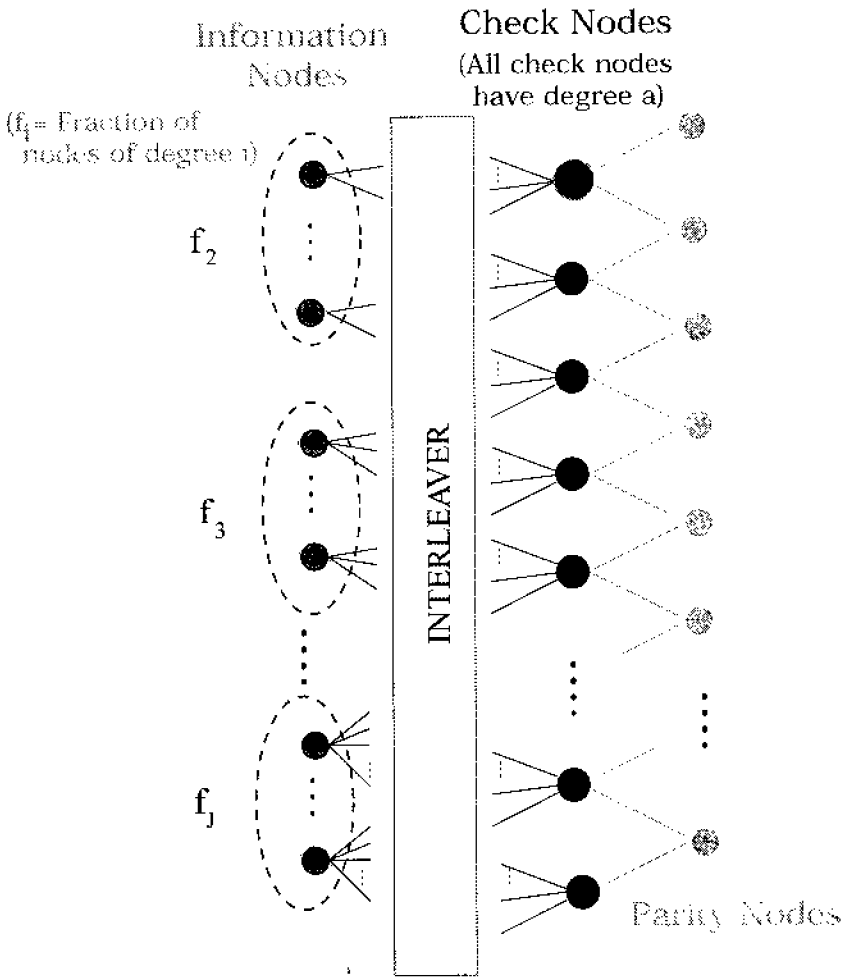
$$f_q(x) = p(1 - (1 - x)(1 - (1 - (1 - x)^2)^{q-1})).$$

The Erasure threshold for $R = 1/4$ RA codes is $p = 0.703$



"Irregular" Repeat-Accumulate Codes

(systematic)

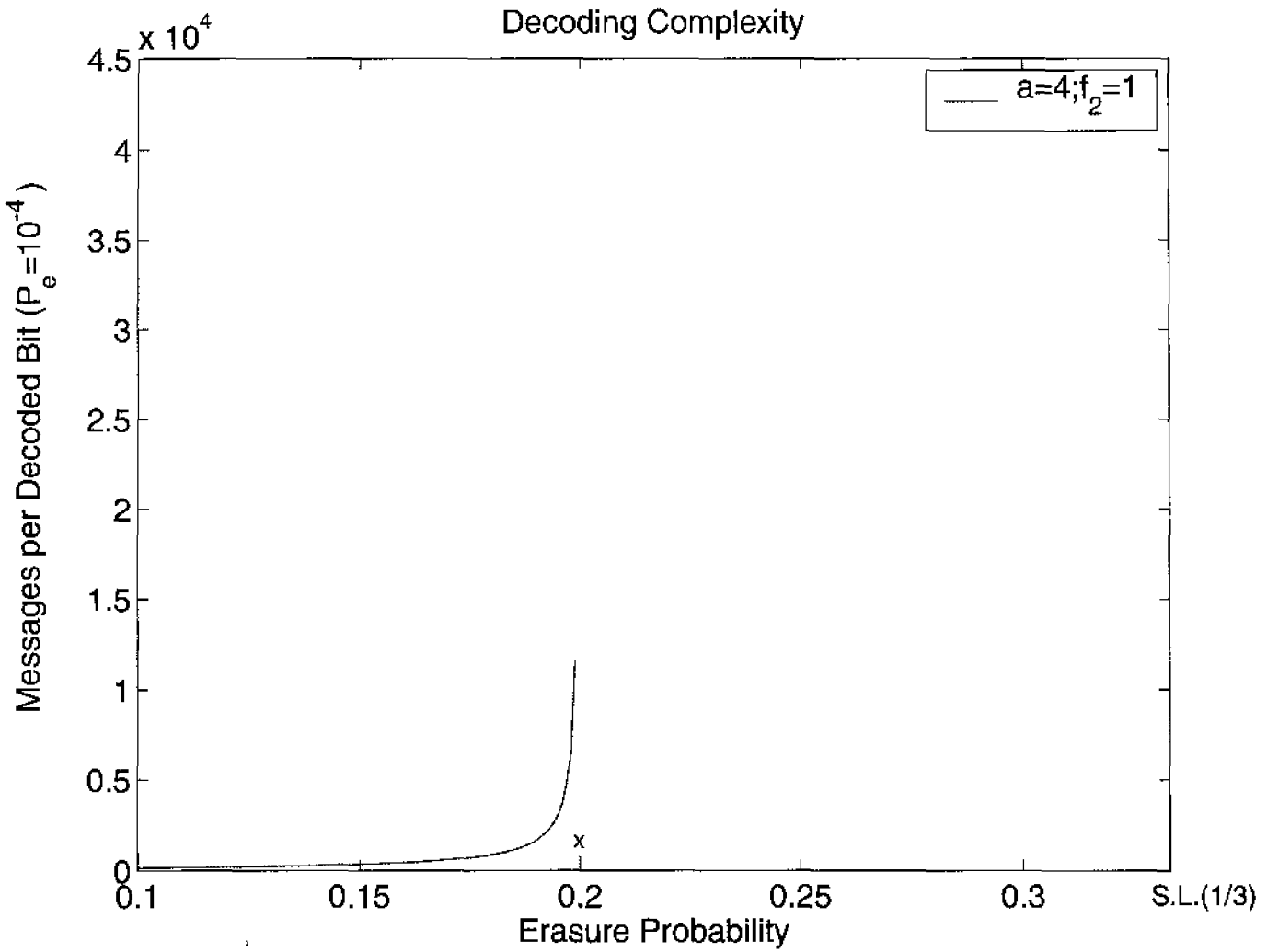


$$\overline{q} = \sum_{q \geq 2} q f_q$$

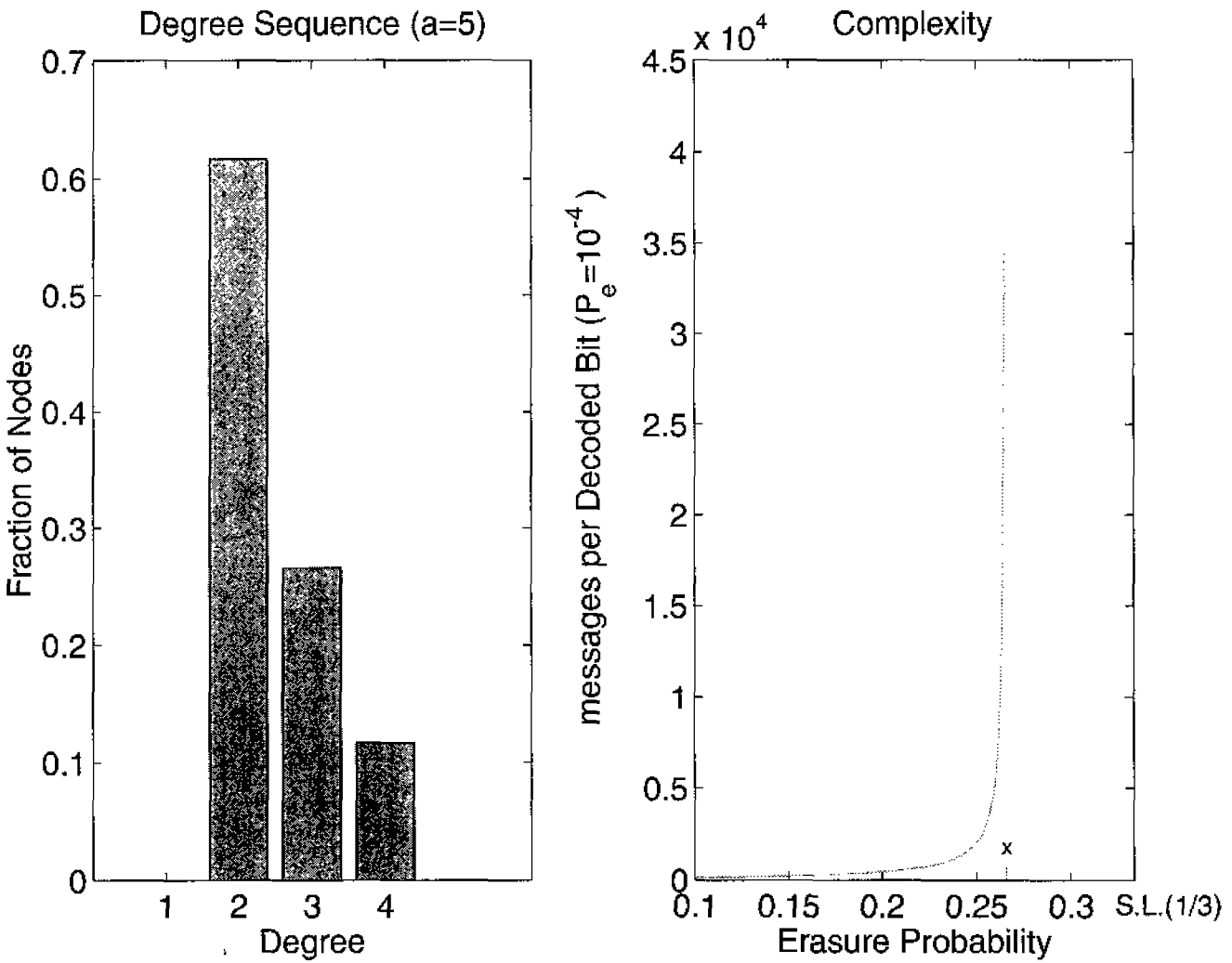
$$R = (1 + \frac{\overline{q}}{a})^{-1}$$

$$\chi_{IT} = 2(\frac{1}{R} - 1)(a + 2).$$

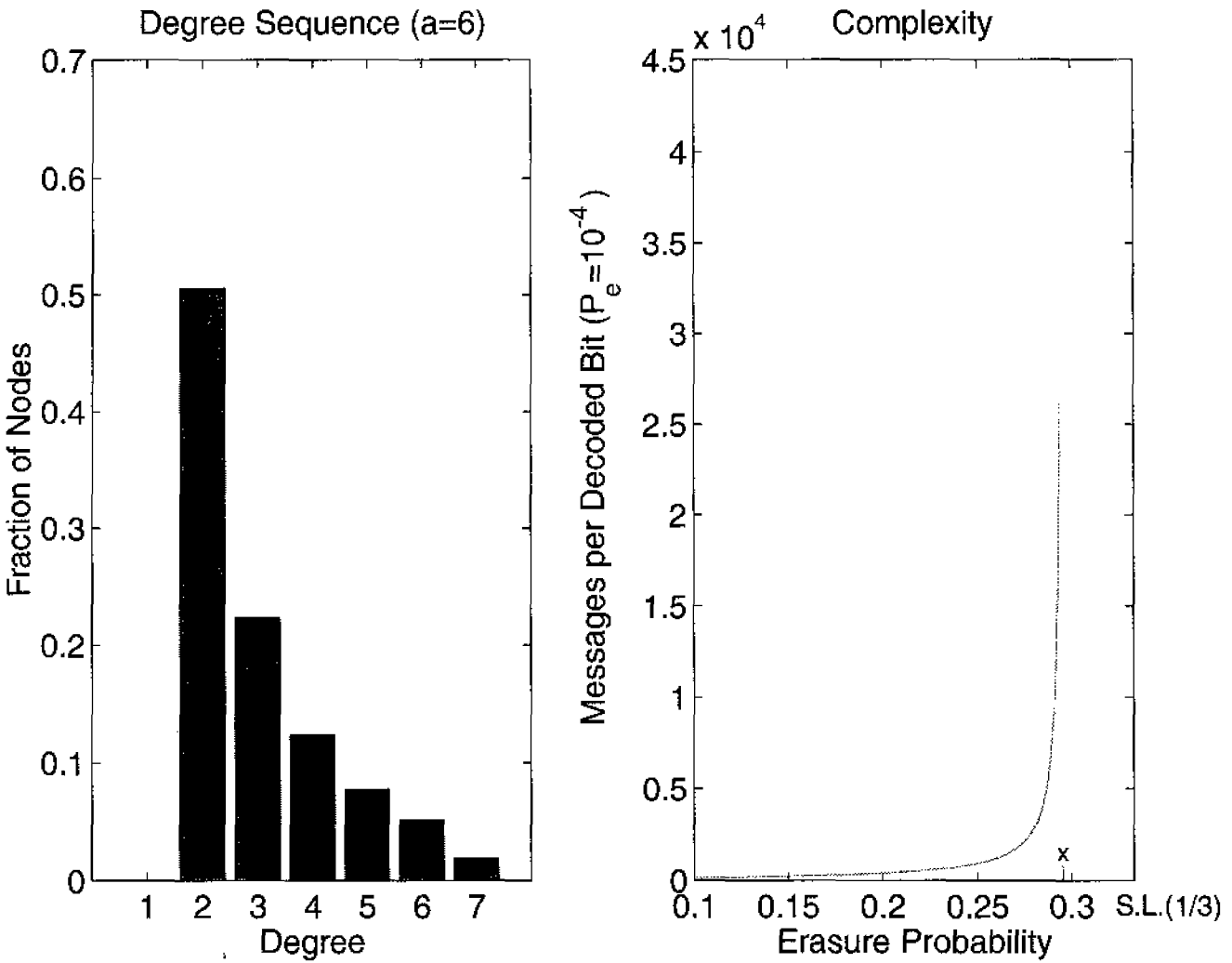
$R = 2/3$ IRA Codes for the Binary Erasure Channel



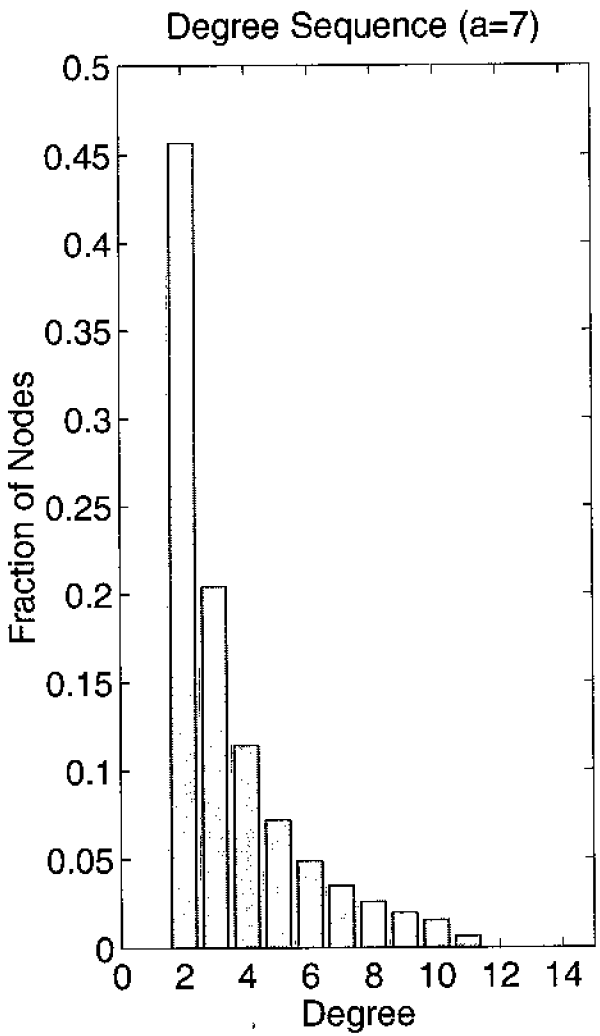
$R = 2/3$ IRA Codes for the Binary Erasure Channel

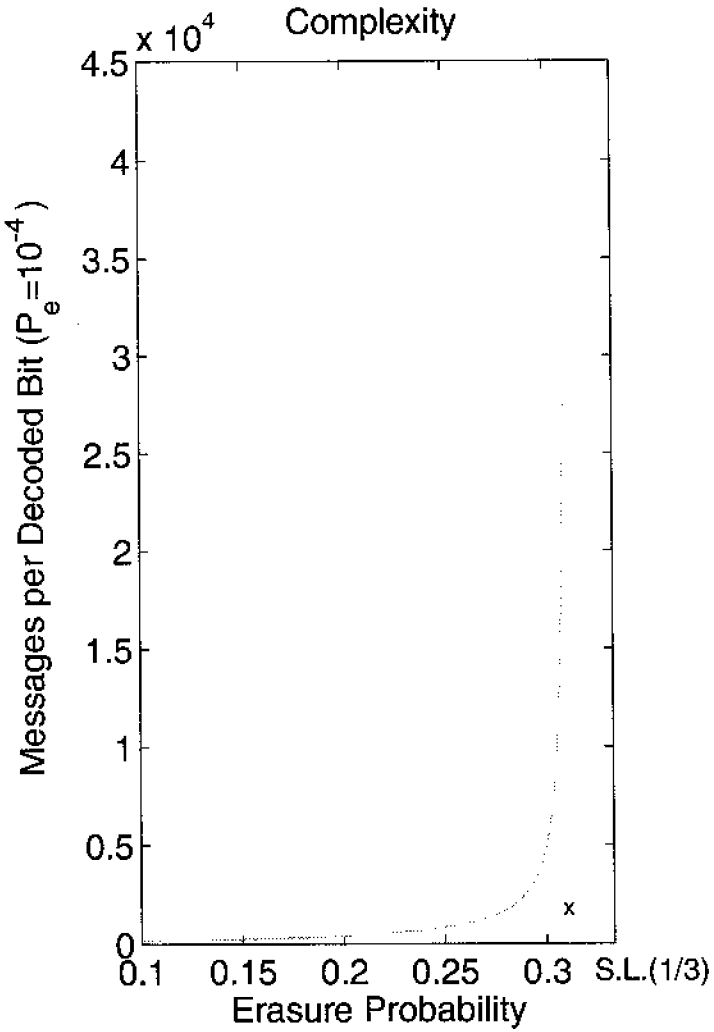


$R = 2/3$ IRA Codes for the Binary Erasure Channel

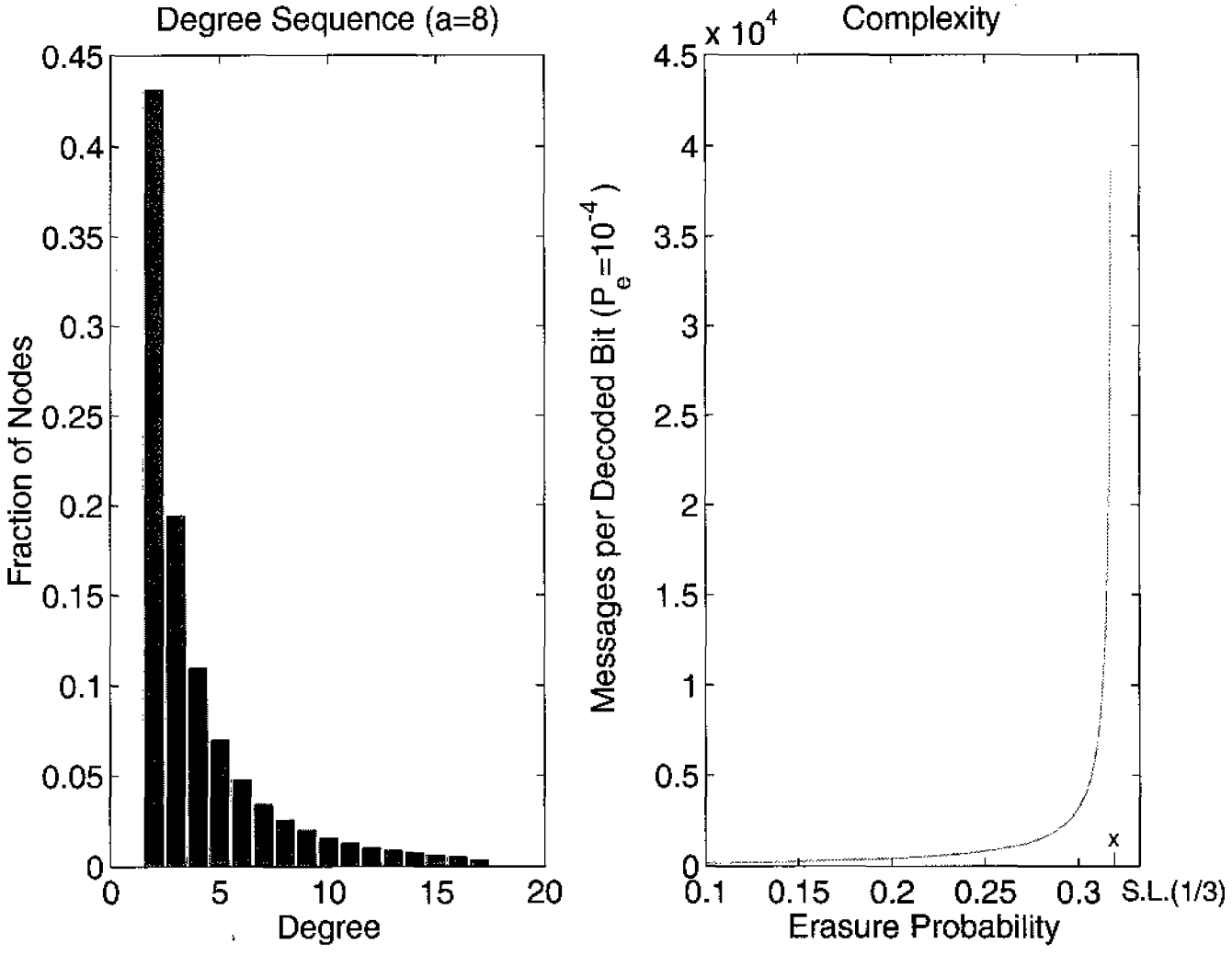


$R = 2/3$ IRA Codes for the Binary Erasure Channel

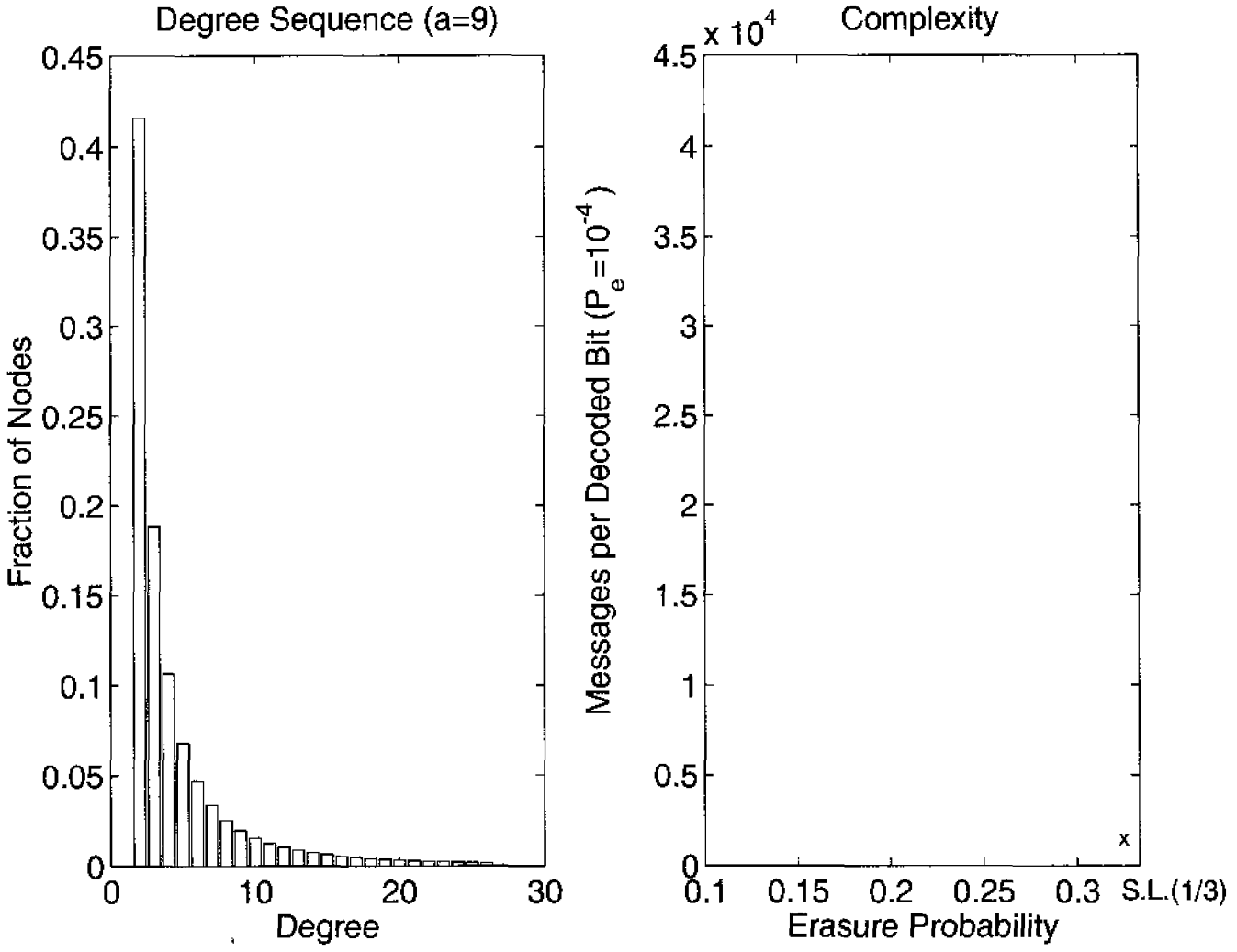




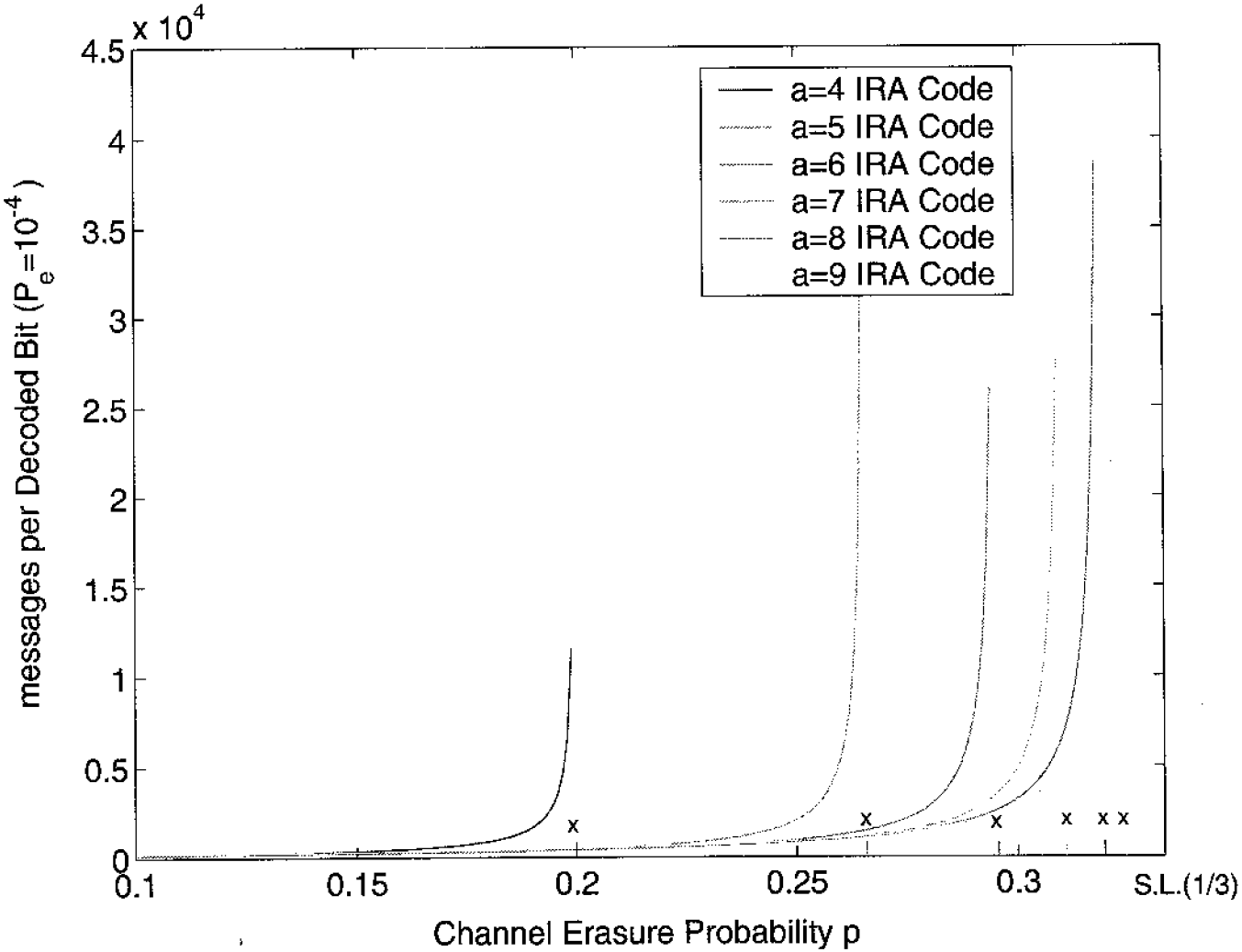
$R = 2/3$ IRA Codes for the Binary Erasure Channel



$R = 2/3$ IRA Codes for the Binary Erasure Channel



$R = 2/3$ IRA Codes for the Binary Erasure Channel



A New Result

Theorem C. *For the binary erasure channel, for IRA codes,*

$$\overline{\chi}_E(\epsilon, \pi) = O(\frac{1}{\epsilon})$$

$$\overline{\chi}_D(\epsilon, \pi) = O(\frac{1}{\epsilon^2} \log \frac{1}{\epsilon})$$

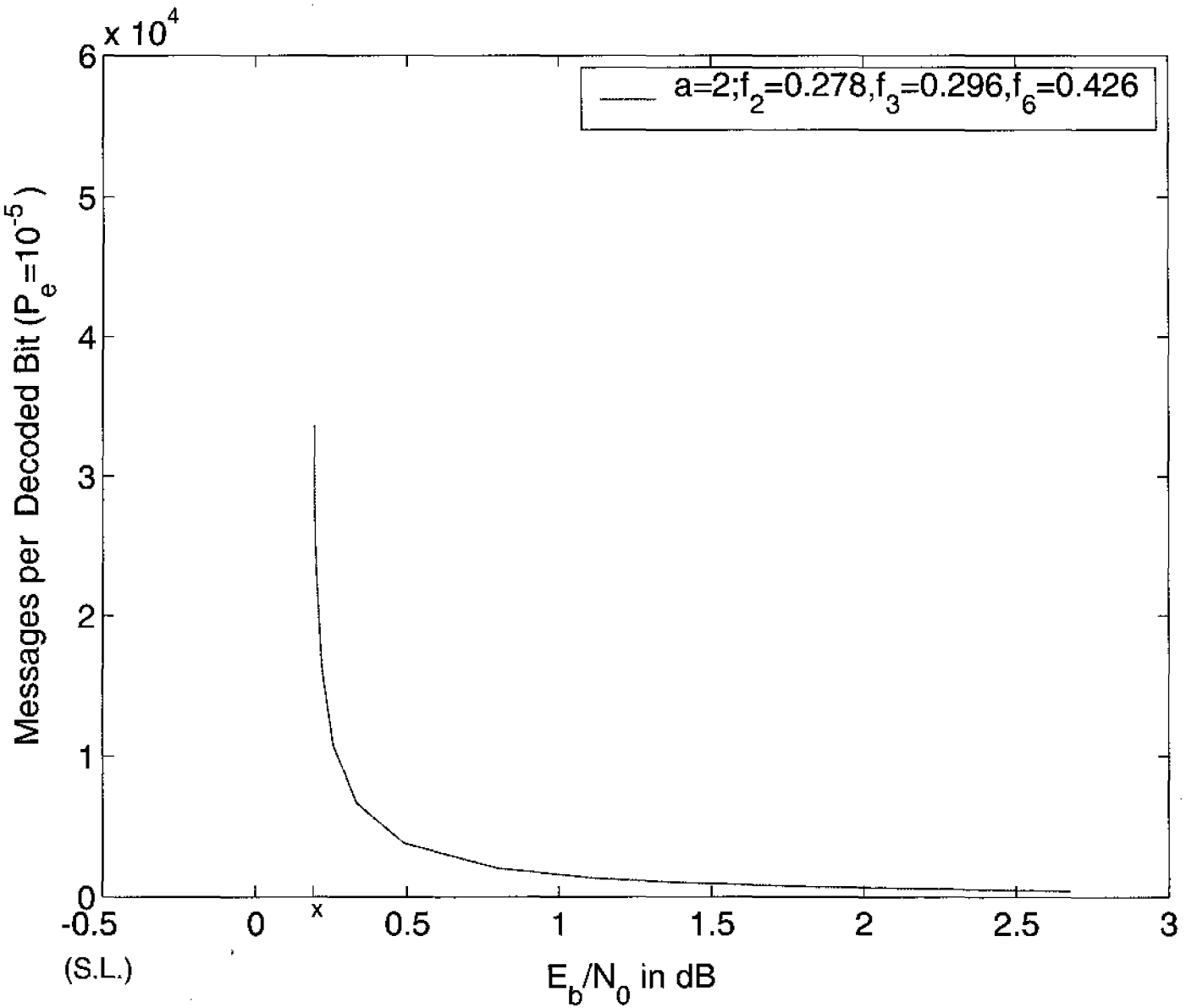
Conjecture D. *For the binary erasure channel, for IRA codes,*

$$\overline{\chi}_E(\epsilon, \pi) = O(\log \frac{1}{\epsilon})$$

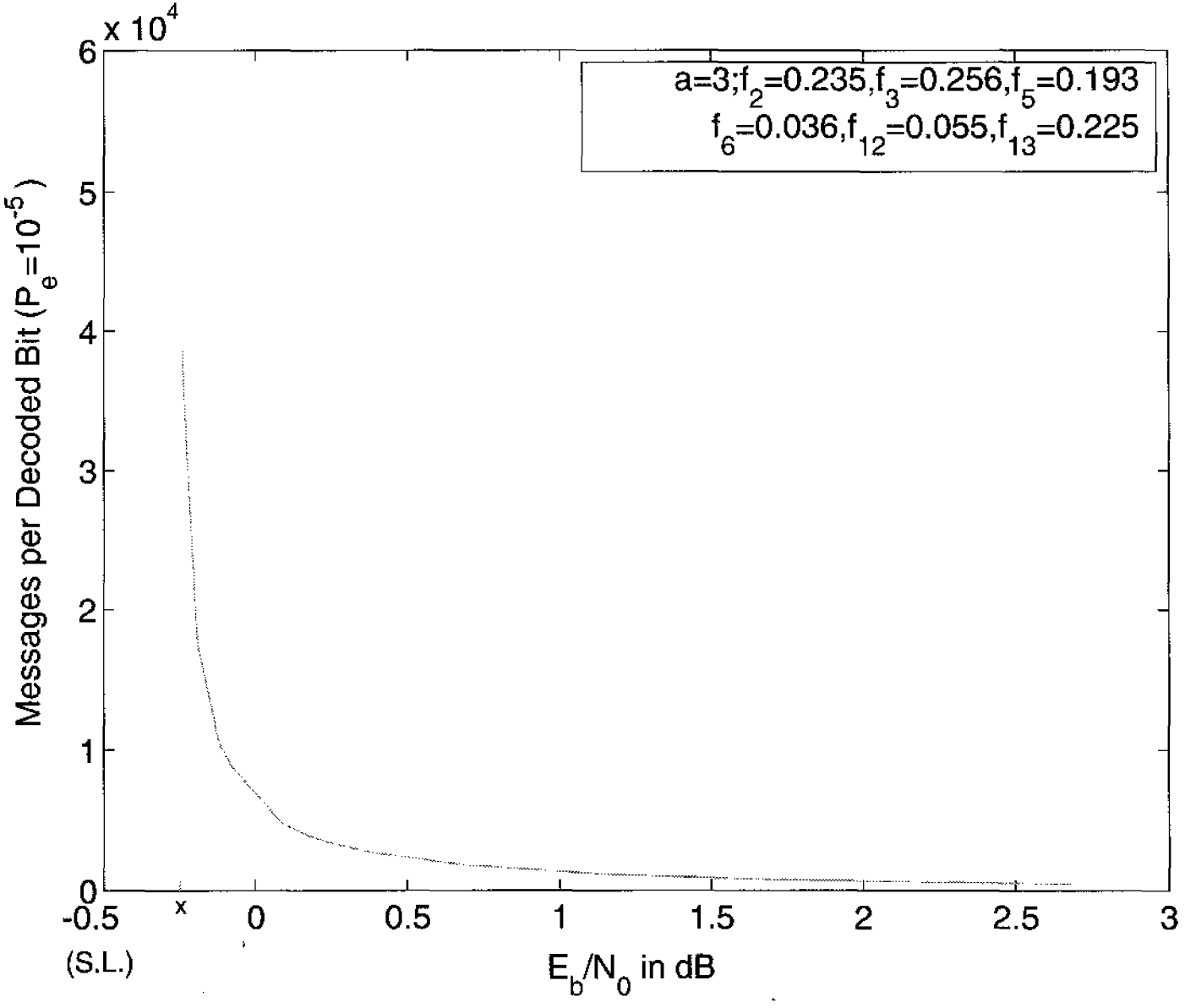
$$\overline{\chi}_D(\epsilon, \pi) = O(\frac{1}{\epsilon} \log \frac{1}{\epsilon})$$

Note: We can prove that $\overline{\chi}_D(\epsilon, \pi) = O(\frac{1}{\epsilon} \log \frac{1}{\epsilon})$ for irregular LDPC codes.

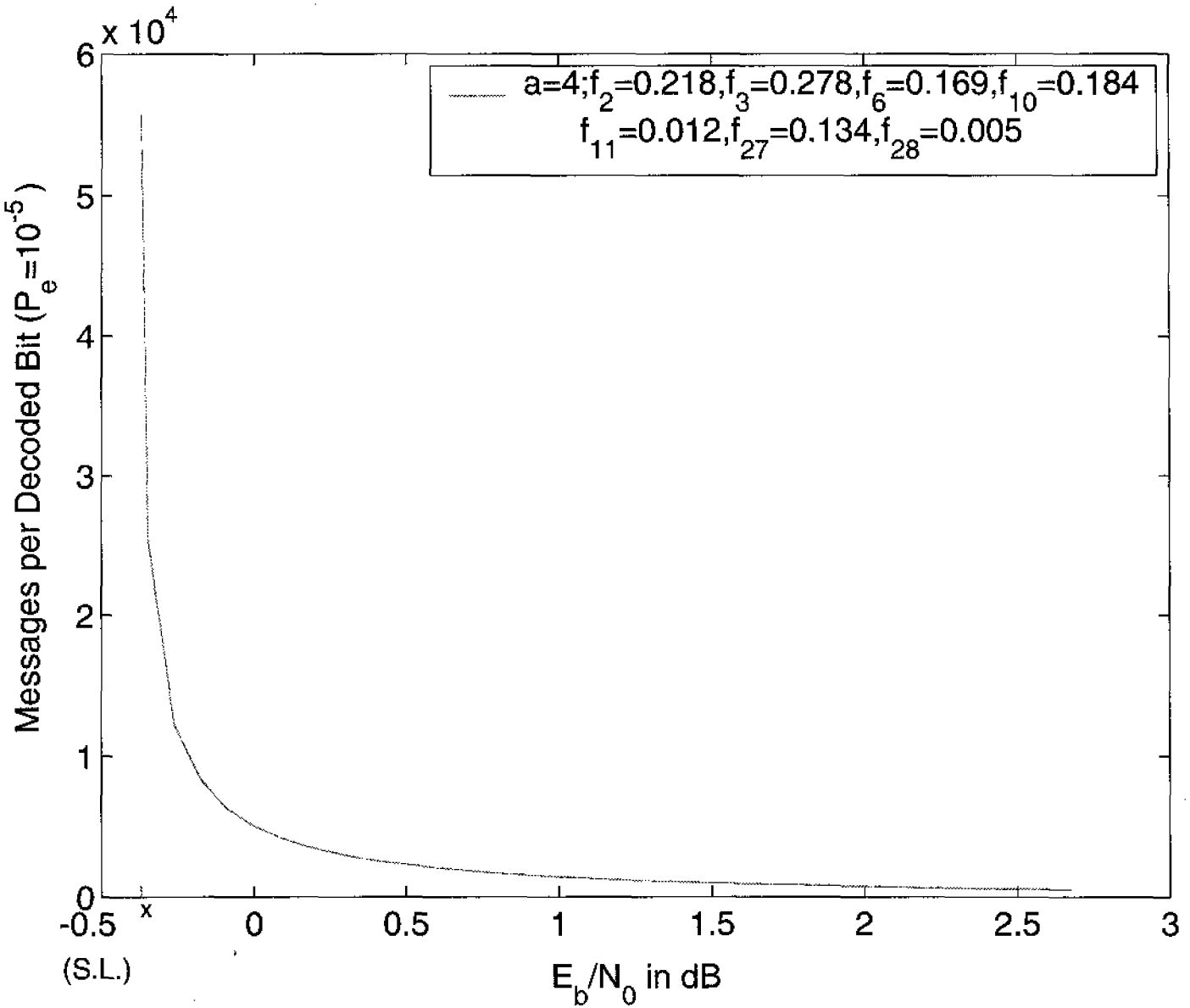
$R = 1/3$ IRA Codes for the AWGN Channel



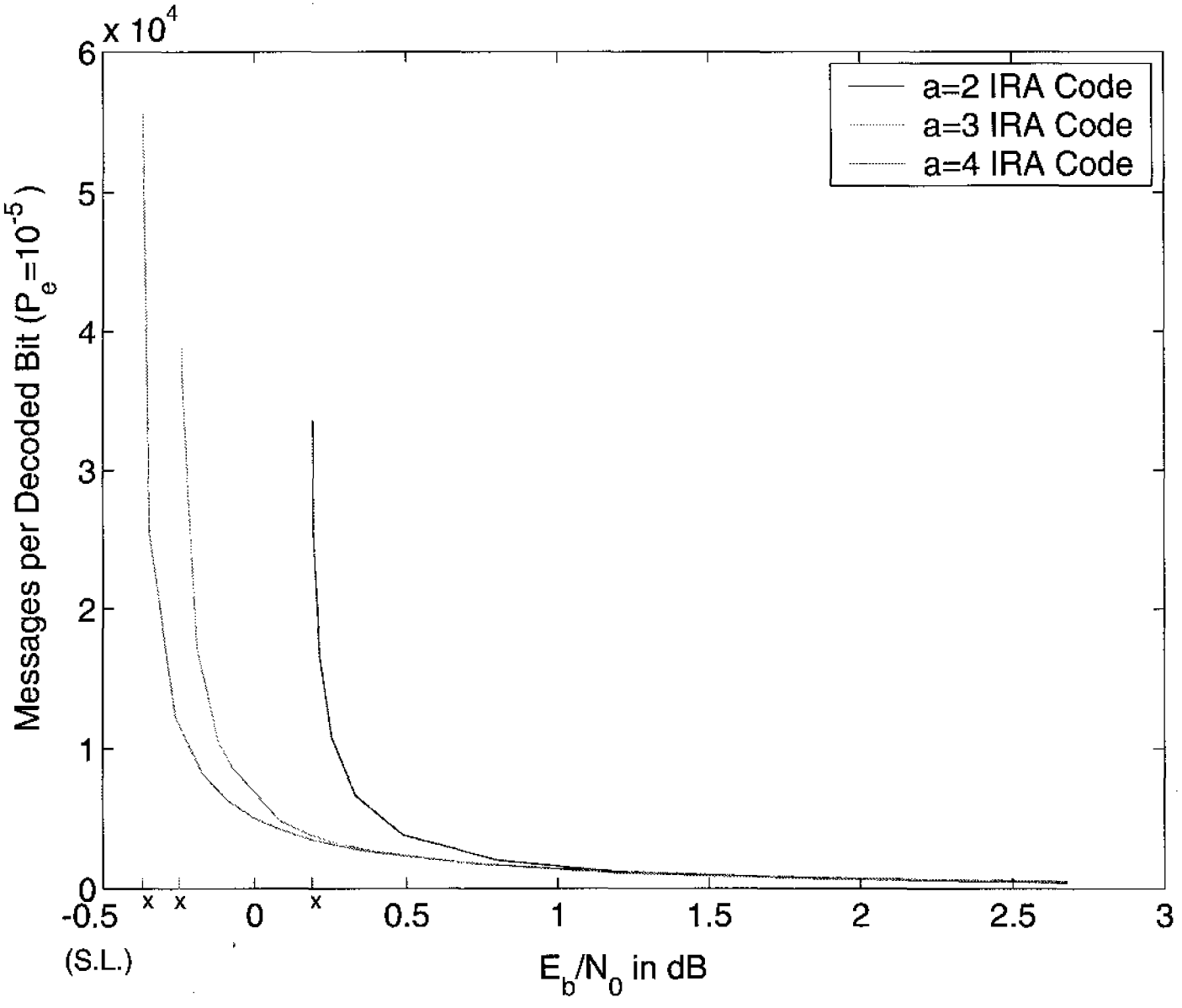
$R = 1/3$ IRA Codes for the AWGN Channel



$R = 1/3$ IRA Codes for the AWGN Channel



$R = 1/3$ IRA Codes for the AWGN Channel



A Conjecture

Conjecture E. *For any discrete memoryless channel, there exists a sequence of ensembles plus matched iterative decoding algorithms, such that for any fixed π, as $\epsilon \to 0$,*

$$\overline{\chi}_E(\epsilon, \pi) = O(\log \frac{1}{\epsilon})$$

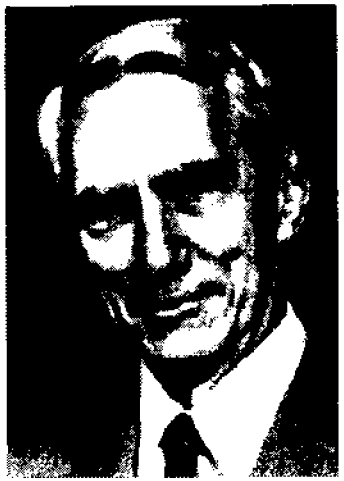
$$\overline{\chi}_D(\epsilon, \pi) = O(\frac{1}{\epsilon} \log \frac{1}{\epsilon})$$

Variations on the Theme

- *Irregular* Turbo Codes (Frey and MacKay)
- *Asymmetric* Turbo codes (Costello and Massey)
- *Mixture* Inner and/or outer codes (Divsalar and Dolinar)
- *Doped* Turbo codes (ten Brink)
- *Irregular* LDPC codes (Richardson, Shokrollahi and Urbanke)
- *Finite Geometry* LDPC codes (Fossorier and Lin)
- *Concatenated Tree* codes (Ping and Wu)

$\vdots$

How We May Appear to Future Generations



Claude Shannon (1918—2001 A.D.). *Generally regarded as the father of the Information Age, he formulated the notion of channel capacity in 1948 A.D. Within several decades, mathematicians, engineers, and physicists had devised practical ways to communicate reliably at data rates within 1% of the Shannon limit ...*

Encyclopedia Galactica, 166th ed.