Télécom Paris (ENST)
Institut Eurécom

THESE

Présentée pour Obtenir le Grade de Docteur
de l'Ecole Nationale Supérieure
des Télécommunications

Spécialité: Communication et Electronique

Souad Guemghar

Techniques de Codage Avancées et Applications au CDMA

Président	P. Solé, I3S (Sophia Antipolis, France)
Rapporteurs	E. Biglieri, Politecnico de Torino (Turin, Italie)
	J. Boutros, ENST (Paris, France)
Examinateurs	A. Glavieux, ENST Bretagne (Brest, France)
	A. Roumy, IRISA (Rennes, France)
Directeur de thèse	G. Caire, Institut Eurécom (Sophia Antipolis, France)

29 Janvier 2004

Télécom Paris (ENST)
Institut Eurécom

THESIS

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
from Ecole Nationale Supérieure
des Télécommunications

Specializing: Communication and Electronics

Souad Guemghar

Advanced Coding Techniques and Applications to CDMA

President	P. Solé, I3S (Sophia Antipolis, France)
Readers	E. Biglieri, Politecnico de Torino (Torino, Italy)
	J. Boutros, ENST (Paris, France)
Examiners	A. Glavieux, ENST Bretagne (Brest, France)
	A. Roumy, IRISA (Rennes, France)
Thesis supervisor	G. Caire, Institut Eurécom (Sophia Antipolis, France)

January 29th 2004

A mes parents, ma soeur et mon mari

Remerciements

Mon travail de thèse est maintenant presque arrivé à sa fin et c'est le moment d'exprimer ma gratitude et mes remerciements envers certaines personnes.

Tout d'abord, je remercie mon directeur de thèse Giuseppe Caire de m'avoir donné l'occasion de faire cette thèse qui constitue une expérience extrêmement enrichissante sur plusieurs plans. Il a souvent eu de brillantes idées de recherche, ce qui m'a fait explorer plusieurs domaines différents.

Je remercie l'ensemble des membres de mon jury qui m'ont fait l'honneur de siéger à ma soutenance.

En particulier, je tiens à remercier les deux rapporteurs de ma thèse, Professeur Ezio Biglieri et Docteur Joseph Boutros. Leur lecture attentive et leurs suggestions ont contribué à l'amélioration de la qualité de ce rapport.

Je remercie aussi Professeur Alain Glavieux et Professeur Patrick Solé d'avoir siégé en tant qu'examinateur et président, respectivement, à ma soutenance publique et d'avoir apporté une contribution critique à ce travail.

Ma collaboration avec Aline Roumy a été d'une grande importance. Elle m'a beaucoup apporté dans ma recherche. J'ai beacoup apprécié les échanges d'idées qu'on a eus, ainsi que sa modestie et son amabilité.

Je veux aussi remercier Professeur Sergio Verdú pour son apport à ce travail qui a grandement contribué à sa qualité scientifique.

Ensuite, je veux remercier la région PACA qui m'a financièrement permis de réaliser ce travail de thèse en m'octroyant une bourse de recherche, avec le soutien d'Infineon Technologies Sophia.

Que dire d'Eurécom? Merci au personnel qui a été efficace et solidaire à plusieurs étapes de ma thèse. Merci aux doctorants. Venant des quatre coins du monde, ils apportent un réel enrichissement et une si grande ouverture d'esprit que j'aimerais bien que le monde soit à leur image.

Je veux remercier mes amis. Ils m'ont soutenue à des moments difficiles et m'ont apporté beaucoup de joie de vivre. Je pense en particulier à (par

ordre alphabétique) Carine, Farouk, Kader, Maxime, Mari et Navid. Je veux aussi remercier Daniela qui m'a apporté son soutien à des moments difficiles.

Ma soeur, qui va bientôt connaître les joies de finir sa thèse, m'a toujours soutenue. Elle a toujours cru en moi et m'a donné le courage et la volonté d'aller au bout de moi-même.

Jan, la rencontre majeure de ma vie, merci pour ton amour, ta confiance, ton soutien et ta présence.

Je dédie ce travail à mes parents. J'arrive au point où je ne sais plus comment exprimer ma gratitude éternelle à mes parents chéris. Ils ont consenti à beaucoup de sacrifices pour que ma soeur et moi ayons une bonne éducation, et ils croient très fort en nous.

Souad Guemghar-Exner
Sophia Antipolis, le 18 février 2004

Abstract

In this work, we propose low-complexity coding/decoding schemes to approach the capacity of binary-input symmetric-output channels and code division multiple access channels.

In the first part of this thesis, we consider systematic random-like irregular repeat accumulate code ensembles of infinite block length, assuming transmission over a binary-input symmetric-output channel. The code ensemble is described by a Tanner graph, and is decoded by the message-passing belief propagation algorithm. Density evolution describes the evolution of message distributions that are passed on the Tanner graph of the code ensemble. Applying density evolution under the belief propagation decoder results in a dynamic system on the set of symmetric distributions. We formulate a general framework to approximate the exact density evolution with a one-dimensional dynamic system on the ensemble of real numbers. Based on this general framework, we propose four low-complexity methods to design irregular repeat accumulate code ensembles. These optimization methods are based on Gaussian approximation, reciprocal (dual) channel approximation and extrinsic mutual information transfer function, among other recently-developed tools. These methods allow us to design irregular repeat accumulate codes, of various rates, with vanishing bit error rate guaranteed by a local stability condition of the fixed-point of the exact density evolution recursions. Using the exact density evolution, the thresholds of the designed codes are evaluated, and are found to be very close to the Shannon limits of the binary input additive white Gaussian noise channel and the binary symmetric channel. For the binary-input additive white Gaussian noise channel, we investigate the performance of finite length irregular repeat accumulate codes, whose graph is conditioned so that either the girth or the minimum stopping set size is maximized. We compare the performances of the resulting IRA codes to those of random ensembles under maximum likelihood

decoding, and to the performances of the best low density parity check codes of comparable graph conditioning.

In the second part of this thesis, we develop a low-complexity coding/decoding scheme, to approach the capacity of the Gaussian multiple access channel, using random-spreading code division multiple access in the large system limit. Our approach is based on the use of quaternary phase shift keying modulation, capacity-achieving binary error-correcting codes, linear minimum mean square error filtering and successive decoding. We optimize the power profile (respectively rate profile) in the case of equal-rate (respectively equal-power) users. In the equal-rate setting, it is found that the achievable spectral efficiency, when using low-rate binary error correcting codes, is very close to the optimum. Through simulations, we show that the system optimization carried out in the large-system limit and for infinite block length can be used to dimension finite-size practical systems with no error propagation throughout the successive decoding.

Résumé

Ce travail propose des schémas de codage et de décodage à complexité réduite, afin d'approcher la capacité des canaux à entrée binaire et sortie symétrique, ainsi que des canaux d'accès multiple par répartition de codes.

Dans la première partie de cette thèse, nous nous attelons à étudier l'ensemble aléatoire de codes irréguliers dits "répétition-accumulation", de longueur infinie, transmis sur un canal à entrée binaire et sortie symétrique, et décodés par l'algorithme somme-produit. En utilisant la technique de l'évolution de densités, on écrit un système récursif qui décrit l'évolution des densités des messages qui sont propagés sur le graphe de Tanner qui représente l'ensemble de codes. Ensuite, on formule un cadre général dans lequel l'évolution des densités est approximée par un système dynamique dont les variables appartiennent à l'ensemble des nombres réels. A partir de ce cadre, on propose quatre méthodes de complexité réduite pour optimiser des codes répétition-accumulation. Ces méthodes sont basées sur l'approximation Gaussienne, l'approximation réciproque (duale), et la fonction de transfert de l'information mutuelle extrinsèque. Ces méthodes permettent de construire des codes de différents rendements, et dont les taux d'erreur tendent vers zéro, pour peu que la condition de stabilité locale soit satisfaite. Les seuils de décodage, évalués par la technique d'évolution de densités exacte, sont très proches de la limite de Shannon du canal Gaussien à entrée binaire et du canal binaire symétrique. Pour le canal Gaussien à entrée binaire, nous nous intéressons à la performance de ces codes dans le cas de la longueur finie, avec un graphe de Tanner conditionné pour maximiser les tailles des cycles les plus courts ou de certains ensembles dits bloquants ou stopping sets. La performance de ces codes est comparée à celle de l'ensemble aléatoire décodé au maximum de vraisemblance, ainsi qu'à celle des meilleurs codes de Gallager de même rendement et niveau de conditionnement.

La deuxième partie de cette thèse développe un schéma de codage/décodage

à complexité réduite afin d'approcher la capacité du canal Gaussien à accès multiple. On considère l'accès multiple par répartition de codes aléatoires dans la limite d'un système de taille infinie. Notre approche est basée sur l'utilisation d'une modulation à déplacement de phase quadrivalente, de codes binaires correcteurs d'erreurs atteignant la capacité du canal, de filtres à erreur quadratique moyenne minimale et d'un décodage successif. On optimise le profil des puissances (respectivement des rendements) en supposant que les utilisateurs du système à accès multiple ont tous le même rendement (respectivement la même puissance). Dans le cas où tous les utilisateurs ont le même rendement, l'efficacité spectrale du système optimisé est très proche de l'efficacité spectrale optimale. Au travers de simulations numériques, il est montré que la méthode d'optimisation permet de passer du système à taille infinie à un système pratique de taille finie, dont le décodage successif ne propage pas d'erreur de décodage.

Contents

List of Figures

List of Tables

Acronyms

Here are the main acronyms used in this document. The meaning of an acronym is usually indicated once, when it first occurs in the text. The english acronyms are also used for the french summary.

ACE	Approximate Cycle EMD
APP	*A Posteriori* Probability
AWGN(C)	Additive White Gaussian Noise (Channel)
BCJR	Bahl, Cocke, Jelinek and Raviv (algorithm)
BEC	Binary Erasure Channel
BER	Bit Error Rate
BIAWGNC	Binar-Input Additive White Gaussian Noise Channel
BP	Belief Propagation
BPSK	Binary Phase Shift Keying
BSC	Binary Symmetric Channel
CDMA	Code-Division Multiple Access
DE	Density Evolution
EMD	Extrinsic Message Degree
EXIT	Extrinsic Mutual Information Transfer (function)
GA	Gaussian Approximation
GMAC	Gaussian Multiple Access Channel
IC	Interference Cancellation
i.i.d.	independent and identically distributed
IOWE	Input-Output Weight Enumerator
IRA	Irregular Repeat Accumulate (code)
ISI	Inter Symbol Interference
LDPC	Low Density Parity Check (code)
MAC	Multiple Access Channel

MAI	Multiple Access Interference
MAP	Maximum *A Posteriori*
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error (filter or decoder)
pdf	probability density function
PEG	Progressive Edge Growth (algorithm)
QPSK	Quaternary Phase Shift Keying
RA	Repeat Accumulate (code)
SINR	Signal to Noise Plus Interference Ratio
SISO	Soft-Input Soft-Output (decoder)
SNR	Signal-to-Noise Ratio
SSMAX	Stopping Set Maximization (algorithm)
TSB	Tangential Sphere Bound
WER	Word Error Rate

Notations

Here is a list of the main notations and symbols used in this document. We have tried to keep consistent notations throughout the document, but some symbols have different definitions depending on when they occur in the text.

General Notations

$\mathbb{C}$	The set of complex numbers
E_b	Energy per information bit
E_b/N_0	Signal to noise ratio per information bit
E_s/N_0	Signal to noise ratio per symbol
N_0	One sided noise power spectral density of the AWGN channel
R	The rate of the considered code
$\mathbb{R}$	The set of real numbers
σ^2	Real Gaussian noise variance
x	Scalar variable
$\mathbf{x}$	Vector variable
X	Scalar random variable
$\mathbf{X}$	Matrix variable

Part 1: Irregular Repeat Accumulate Codes

a	Grouping factor
$A_{w,h}$	Input output weight enumerator (or IOWE)
d_{min}	The minimum Hamming distance of the considered code
$\mathcal{E}_{sym}$	Symmetric distribution of a BEC
$\mathcal{F}_{sym}$	Set of symmetric distributions
$I(X;Y)$	Mutual information between random variables X and Y

$\mathfrak{I}$	Binary-input symmetric-output capacity functional
$J(\mu)$	Capacity functional for $\mathcal{N}_{sym}(\mu)$
k	Information block length
ℓ	Iteration number
λ_i	Fraction of edges connected to a degree-i information bitnode
m	Repetition block length
n	Output code block length
N	Repetition frame length
$\mathcal{N}_{sym}$	Symmetric Gaussian distribution $\mathcal{N}(\mu, 2\mu)$
P_b	Bit error probability (or BER)
P_w	Word error probability (or WER)
Pe	Probability of error
P_ℓ	Average distribution of messages from an information bitnode to a checknode
$\tilde{P}_\ell$	Average distribution of messages from a parity bitnode to a checknode
Q_ℓ	Average distribution of messages from checknode to an information bitnode
$\tilde{Q}_\ell$	Average distribution of messages from a checknode to a parity bitnode
r	Chernoff bound exponent

Part 2: Coded CDMA under Successive Decoding

β	Total channel load
β_j	Class load of class j
C	Optimal spectral efficiency of random CDMA
C^{mmse}	Spectral efficiency of random CDMA with MMSE filtering
C_{qpsk}	Capacity of QPSK-input AWGN channel
C_{qpsk}	Spectral efficiency of QPSK-input random CDMA
C^*	Single-user AWGN channel capacity
η	Large-system multiuser efficiency of the MMSE receiver
γ_j	SNR of users in class j
L	Total number of classes
K	Total number of users
K_j	Number of users in class j

$\kappa(X)$	Kurtosis of the distribution of X
N	Spreading factor (or gain)
R_j	Code rate of users in class j
ρ	Spectral efficiency of CDMA system
$\mathbf{S}$	Matrix of spreading sequences

Chapter 1

Introduction

1.1 Advanced Coding Techniques

Claude Shannon proved [1] the existence of codes that allow reliable transmission, provided that the information rate in bits per channel use is less than the channel *capacity*. A randomly generated code with large block size has a high probability to be a good code, i.e. to closely approach the Shannon limit. However, its decoding complexity is exponential in the block size, and is thus prohibitive in practice. Hence, the central challenge of coding theory consists of designing codes that perform as close as possible to the Shannon limit, and are still decodable with a reasonable complexity. An important step in this direction was the introduction of concatenated codes by David Forney [2], which consist of a powerful *outer* block code and an *inner* convolutional code. The inner decoder makes use of the Viterbi algorithm, and is followed by an outer decoder based on hard decisions and algebraic decoding.

In 1993, Berrou, Glavieux and Thitimajshima introduced a novel coding structure, named *Turbo Codes* [3], which consists of the parallel concatenation of two convolutional codes through an interleaver. This structure admits an iterative decoding scheme, based on the recursive estimation of *a posteriori* probabilities (APP) using the BCJR algorithm [4] and exchanging *extrinsic*

information between the constituent decoders. The performance of turbo codes approaches the Shannon capacity of the additive white Gaussian noise channel within a fraction of a dB.

The introduction of turbo codes constitutes a major breakthrough in coding theory as it gave rise to a large amount of work in the field of random-like codes, leading to the introduction of "turbo-like" code families. In particular, we note the re-discovery of the low density parity check (LDPC) codes, originally proposed in [5], the introduction of irregular LDPC codes [6, 7] and the introduction of the Repeat-Accumulate (RA) codes [8]. In many relevant settings, the iterative decoding of these codes achieves performances that are very close to the Shannon limit. In [6, 7], irregular LDPC codes were shown to asymptotically achieve the capacity of the binary erasure channel (BEC) under iterative message-passing decoding. Although the BEC is the only channel for which such a result currently exists, irregular LDPC codes have been designed for other binary-input channels, e.g., the binary symmetric channel (BSC), the binary input additive white Gaussian noise channel (BI-AWGNC) [9], and the binary-input inter-symbol interference (ISI) channel [10, 11, 12, 13], and have been shown to achieve very good performances.

The *Tanner* (*bipartite*) *graph* [14] is a powerful formalism that provides a description of these turbo-like codes and their iterative decoding. The codes are decoded iteratively by performing local computations at *nodes*, and passing the resulting information along *edges* in the graph. The complexity of the decoder depends on the complexity of the local node computation, the complexity of the information passing on the edges, and finally on the number of iterations of the decoder. We will mainly be interested in one type of iterative decoders, namely message passing decoders, for which messages represent estimates of the transmitted bits. Moreover, we concentrate on the *belief propagation* (BP) decoder which assumes local independence of *incoming* messages, and applies probability rules at the computation nodes.

The introduction of irregular LDPC codes motivated other turbo-like coding schemes such as irregular RA codes (IRA), for which the achievability of the BEC capacity has been shown [15], and irregular turbo codes [16, 17]. IRA codes are in fact special subclasses of both irregular LDPC and irregular turbo codes. IRA codes are an appealing choice because they have a low encoding/decoding complexity and their performance is quite competitive with that of turbo codes and LDPC codes.

An IRA code is characterized by $\{f_i \geq 0, i = 2, 3, \ldots : \sum_{i=2}^{\infty} f_i = 1\}$ referred to as the *repetition profile*, and a *grouping factor* a. A fraction f_i of

information bits is repeated i times, for $i = 2, 3, \ldots$. The resulting sequence is interleaved and input to a recursive finite-state machine, the *accumulator*, which outputs one bit for every a input symbols. $\{f_i\}$ and a are considered as degrees of freedom in the optimization of the IRA code ensemble.

The recursive finite-state machine is the simplest one which gives full freedom to choose any rational number between 0 and 1 as the coding rate. In this work, we restrict our study to IRA codes that use a two-state convolutional code, obeying the same simple recursion as in [15], although it might be expected that better codes can be obtained by including the finite-state machine as a degree of freedom in the overall ensemble optimization.

First attempts to optimize irregular LDPC codes ([18] for the BEC and [19] for other channels) were based on the density evolution (DE) technique, which computes the expected performance of a random-like code ensemble in the limit of infinite code block length. In order to reduce the computational burden of ensemble optimization based on the DE, faster techniques have been proposed, based on the approximation of the DE by a one-dimensional dynamical system (recursion). These techniques are exact only for the BEC, for which DE is one-dimensional. The most popular techniques proposed so far are based on the Gaussian approximation (GA) of messages exchanged in the message passing decoder. GA in addition to the *symmetry condition* of message densities allows the Gaussian density of messages to be expressed by a single parameter. Techniques differ in the parameter to be tracked and in the mapping functions defining the dynamic system that describes the evolution of probability distributions on the Tanner graph associated to the code ensemble [20, 21, 22, 23, 24, 25, 26].

In this thesis, we tackle the problem of optimizing IRA code ensembles on binary-input symmetric-output channels. We propose four design methods based on the approximation of DE by the evolution of a one-dimensional parameter, namely the mutual information between the transmitted bits and the *log likelihood ratio* messages propagated on the graph. These methods differ in the way message densities and BP computations are approximated. The first method relies on Gaussian approximation and *reciprocal channel approximation*. This allows to write the recursions of the approximate DE in closed-form. This is equally the case for method 2, except that here it is assumed that messages are outputs of a virtual BEC whose capacity is the same as the computed mutual information. On the other hand, methods 3 and 4 do not have closed-form expressions of the approximate DE recursions. Indeed, there we make use of Monte Carlo simulation in order to track the

mutual information evolution on the bipartite graph. These methods are formulated such that they can be used to design IRA code ensembles on any binary-input symmetric-output channel.

If the code block length is finite, i.e., the bipartite graph has a finite length, the assumption of local independence of messages does not generally hold. Indeed, randomly-constructed turbo-like codes of finite length may have poor performances under the BP decoder, and the finite-length gap from channel capacity may not be as good as predicted by the infinite length DE analysis. This gives rise to the interleaver design issue which has a fundamental role in the finite-length performance of codes on graphs. Interleaver design criteria are mainly based on heuristic arguments: the elimination of short cycles in order to limit the propagation of unreliable messages, the maximization of the *minimum distance* in order to eliminate low-weight codewords responsible for poor performance for medium to high signal to noise ratios (SNR), the maximization of *stopping sets* [27] responsible for decoding errors of LDPC codes on the BEC.

In this thesis, we analyze the performance of finite-length IRA codes on the BIAWGNC, whose interleavers are subject to one of the following constraints: (a) the size of the smallest cycle is larger than a certain value, which is a design parameter, (b) the size of the smallest stopping set is larger than a certain value, which is a design parameter. We also determine the average random regular RA code performance on the BIAWGNC under maximum likelihood (ML) decoding, which is then compared to the performance with graph-conditioning.

1.2 Coded CDMA with Successive Decoding

In the multiple access channel (MAC), several transmitters (users) share the same transmission medium (physical channel). The output of the channel is the noisy superposition of the transmitted signals. In the present work, we consider the Gaussian multiple access channel (GMAC) in which the noise is Gaussian. Multiuser information theory [28] teaches us that the capacity of a multiple access channel, i.e. the total number of users that can be transmitted reliably on the channel, is generally maximized by transmitting mutually interfering signals. The main figure of merit is the *spectral efficiency*, defined as the total data rate per unit bandwidth (bits per second per Hertz, or bits per channel use) .

In the present work, we investigate low complexity practical coding/decoding schemes to approach the maximum spectral efficiency of the GMAC. Reaching the optimal performance requires suitable coding strategies, and a decoding scheme that can decode the stream of transmitted bits arbitrarily reliably from the noisy superposition of transmitted signals. This is done by the *successive decoding* technique which decodes a given user treating all other users as noise, then subtracts the contribution of the decoded user from the signal, and repeats this process until all users have been successfully decoded.

The coding strategy that we adopt in the present work is code division multiple access (CDMA) also known as spread spectrum multiple access [29, 30]. In CDMA, each user is assigned a different *spreading sequence*, which is a unit-norm vector in an N-dimensional signal space. The elements of a spreading sequence are called *chips*. The *spreading factor* N is the number of chips per transmitted symbol. We assume that the spreading is random, i.e. spreading sequences are assigned randomly and chips are chosen equally likely and independently. The use of random spreading is justified by the following facts:

- Random spreading accurately models practical CDMA systems which use long pseudo-noise sequences that span many symbol periods [30].
- The spectral efficiency averaged with respect to the choice of signatures is a lower bound to the optimum spectral efficiency achievable with an optimum choice of deterministic spreading sequences.

The spectral efficiency of synchronous CDMA systems with Gaussian noise and random spreading has been found in the large-system limit, where the number of users and the spreading factor are infinite, while their ratio, the *channel load*, is finite [31, 32]. The maximum spectral efficiency is achieved by Gaussian inputs. However, practical systems make use of discrete small-size input alphabets. Some recent works use the asymptotic analysis of large CDMA systems with random spreading sequences and various receivers to design practical CDMA systems [33].

In this thesis, we investigate the maximum spectral efficiency achievable by random synchronous CDMA, in the large system limit and infinite code block length, in the following low-complexity setting: quaternary phase shift keying (QPSK) modulation, binary capacity-achieving error-control codes, and successive decoding . For given codes, we maximize the spectral efficiency

of the CDMA system assuming successive decoding, for the cases of equal rate and equal power users. In both cases, the maximization of the spectral efficiency can be formulated as a linear program and admits a simple closed-form solution that can be readily interpreted in terms of power and rate control. We provide examples of the proposed optimization methods based on off-the-shelf LDPC codes and show that the use of low-rate binary codes, in the equal rate setting, is almost optimal in approaching the GMAC capacity. We also investigate by simulation the performance of practical systems with finite code block length, finite number of users and spreading factor. Our low-complexity decoding approach is based on a twofold iterative decoding scheme:

- Binary LDPC codes are decoded iteratively at the single user level.
- The overall stripping decoder is iterated more than once in order to make it more robust to error propagation due to non-zero bit error rate after single-user decoding.

1.3 Thesis Outline

In Chapter 2, we present the random IRA code ensemble with infinite code block length limit. Section 2.1 introduces the systematic IRA encoder, Section 2.2 presents binary-input symmetric-output channels, and Section 2.3 introduces the message passing belief propagation decoder related to the systematic IRA code ensemble. Section 2.4, the density evolution technique is summarized and applied to the IRA code ensemble. The DE recursions are established and a local stability condition of the fixed-point corresponding to vanishing error is derived.

In Chapter 3, we optimize the repetition profiles of random IRA code ensembles for the collection of binary-input symmetric-output channels. The optimization is formalized as a linear program. In Section 3.1, we formalize the approximate DE of the IRA code ensemble in a general framework, by introducing two mappings from the set of symmetric distributions (respectively the set of real numbers) into the set of real numbers (respectively the set of symmetric distributions). Section 3.2 presents the extrinsic mutual information transfer function which describes the evolution of mutual information on a bipartite graph under message passing decoding. We also introduce the binary-input symmetric-output capacity functional, whose proper-

ties are summarized in Section 3.3. In Section 3.4, we propose four ensemble optimization methods, which stem from our general framework. Section 3.5 shows some properties the DE approximation methods introduced in the previous section. In particular, we analyze the local stability conditions of the new fixed-points and derive some properties of the optimized rates of methods 2 and 4. In Section 3.6, we compare the code optimization methods by evaluating the iterative decoding thresholds of the optimized IRA code, using the exact DE evolution, over the BIAWGNC and the BSC.

In Chapter 4, we are concerned with the performance of finite length regular and irregular Repeat and Accumulate codes when used on the BIAWGNC. Section 4.2 describes how to construct IRA codes whose bipartite graph is free of cycles or stopping sets up to a certain length. Section 4.3 compares the theoretical *girth* (size of the smallest cycle) of the bipartite graph of short-length regular RA codes to that obtained by girth-conditioning. Then, Section 4.4 shows how to determine the average random regular RA code performance under ML decoding. In Section 4.5 gives simulation results on the performances of finite length regular and irregular RA codes of rate 1/2.

Chapter 5 states the theoretical limits on the spectral efficiency achievable by the successive decoder on the power-constrained CDMA channel, in the large system limit. Section 5.1 presents the basic synchronous CDMA AWGN model where users are grouped into a finite number of classes such that users in a given class have the same rate and received SNR. Section 5.2 introduces the Gaussian multiple access channel and the successive decoder. In Section 5.3, we state existing results on the optimum spectral efficiency of the power-constrained CDMA channel in the large system limit. Our choice for the QPSK input constellation is justified on the basis of complexity and asymptotic optimality, as shown in Section 5.4.

In Chapter 6, we tackle the problem of approaching the optimum spectral efficiency of CDMA with QPSK input modulation, binary-input capacity-approaching binary error codes and a low-complexity successive decoding algorithm. In Section 6.1.1, we consider the optimization of the received power profile of the different classes. Conversely, in Section 6.1.2, we optimize the code rate profile assuming the same received SNR for all users. Section 6.2 presents numerical examples of both system design settings when the binary user codes are optimum irregular LDPC codes found in [9]. Simulation results for finite block length and finite number of users validate the large-system infinite block length assumption made in the proposed optimization methods.

In Chapter 7, we summarize the contributions of the thesis and propose

some directions for future research.

Part I

Irregular Repeat Accumulate Codes

Chapter 2

Irregular Repeat Accumulate Codes and Decoding

This chapter introduces the systematic irregular repeat accumulate encoder and its related decoder: the belief propagation message-passing algorithm. In the infinite block length limit, and for binary-input symmetric-output channels, the density evolution technique is used to analyze the decoder, leading to a two-dimensional dynamical system on the space of symmetric distributions. A local stability condition around the fixed-point of the system is derived.

2.1 Encoding of IRA Codes

Fig. 2.1 shows the block-diagram of a systematic IRA encoder. A block of information bits $\mathbf{b} = (b_1, \ldots, b_k) \in \mathbb{F}_2^k$ is encoded by an (irregular) repetition code of rate k/N. Each bit b_j is repeated r_j times, where $(r_1, \ldots, r_k)$ is a sequence of integers such that $2 \leq r_j \leq d$ and $\sum_{j=1}^{k} r_j = N$ (d is the maximum repetition factor). The block of repeated symbols is interleaved, and the resulting block $\mathbf{x}_1 = (x_{1,1}, \ldots, x_{1,N}) \in \mathbb{F}_2^N$ is encoded by an *accumulator*,

defined by the recursion

$$x_{2,j+1} = x_{2,j} + \sum_{i=0}^{a-1} x_{1,aj+i}, \quad j = 0, \ldots, m-1 \tag{2.1}$$

with initial condition $x_{2,0} = 0$, where $\mathbf{x}_2 = (x_{2,1}, \ldots, x_{2,m}) \in \mathbb{F}_2^m$ is the accumulator output block corresponding to the input $\mathbf{x}_1$, $a \geq 1$ is a given integer (referred to as *grouping factor*), and we assume that $m = N/a$ is an integer. Finally, the codeword corresponding to the information block $\mathbf{b}$ is given by $\mathbf{x} = (\mathbf{b}, \mathbf{x}_2)$ and the output block length is $n = k + m$.

The transmission channel is memoryless, binary-input and symmetric-output, i.e., its transition probability $p_{Y|X}(y|x)$ satisfies

$$p_{Y|X}(y|0) = p_{Y|X}(-y|1) \tag{2.2}$$

where $y \mapsto -y$ indicates a *reflection* of the output alphabet[1]. Binary-input symmetric-output channels are presented in the next section.

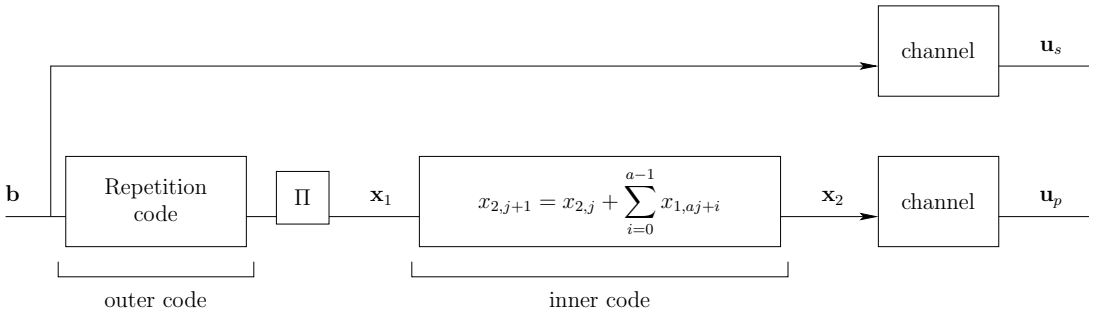


Figure 2.1: Systematic IRA encoder.

IRA codes are best represented by their Tanner graph [14] (see Fig. 2.2). In general, the Tanner graph of a linear code is a bipartite graph whose node set is partitioned into two subsets: the *bitnodes*, corresponding to the coded symbols, and the *checknodes*, corresponding to the parity-check equations that codewords must satisfy. The graph has an edge between bitnode α and checknode β if the symbol corresponding to α participates in the parity-check equation corresponding to β.

[1]If the output alphabet is the real line, then $-y$ coincides with ordinary reflection with respect to the origin. Generalizations to other alphabets are immediate.

Since the IRA encoder is systematic (see Fig. 2.1), it is useful to further classify the bitnodes into two subclasses: the information bitnodes $\{v_j, j = 1, \cdots, k\}$, corresponding to information bits, and the parity bitnodes $\{p_j, j = 1, \cdots, k\}$, corresponding to the symbols output by the accumulator. Those information bits that are repeated i times are represented by bitnodes with degree i, as they participate in i parity-check equations. Each checknode $\{c_j, j = 1, \cdots, m\}$ is connected to a information bitnodes and to two parity bitnodes and represents one of the equations (for a particular j) (2.1). The connections between checknodes and information bitnodes are determined by the interleaver and are highly randomized. On the contrary, the connections between checknodes and parity bitnodes are arranged in a regular zig-zag pattern since, according to (2.1), every pair of consecutive parity bits are involved in one parity-check equation.

A random IRA code ensemble with parameters $(\{\lambda_i\}, a)$ and information block length k is formed by all graphs of the form of Fig. 2.2 with k information bitnodes, grouping factor a and $\lambda_i N$ edges connected to information bitnodes of degree i, for $i = 2, \ldots, d$. The sequence of non-negative coefficients $\{\lambda_i\}$ such that $\sum_{i=2}^{d} \lambda_i = 1$ is referred to as the *degree distribution* of the ensemble. The probability distribution over the code ensemble is induced by the uniform probability over all interleavers (permutations) of N elements.

The information bitnode average degree is given by $\bar{d} \stackrel{\Delta}{=} 1/(\sum_{i=2}^{d} \lambda_i/i)$. The number of edges connecting information bitnodes to checknodes is $N = k/(\sum_{i=2}^{d} \lambda_i/i)$. The number of parity bitnodes is $m = k/(a \sum_{i=2}^{d} \lambda_i/i)$. Finally, the code rate is given by

$$R = \frac{k}{k+m} = \frac{a \sum_{i=2}^{d} \lambda_i/i}{1 + a \sum_{i=2}^{d} \lambda_i/i} = \frac{a}{a + \bar{d}} \tag{2.3}$$

Under the constraints $0 \leq \lambda_i \leq 1$ and $\sum_{i \geq 2} \lambda_i = 1$, we get $\bar{d} \geq 2$. Therefore the highest rate with parameter a set to 1 is 1/3. This motivates the use of $a \geq 2$ in order to get higher rates.

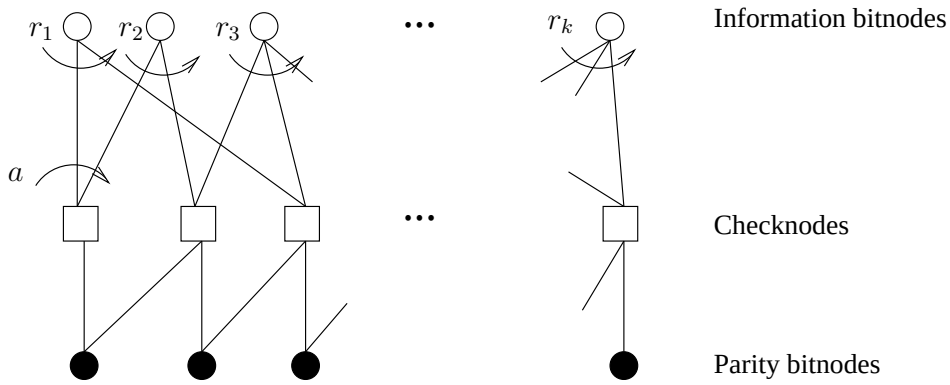


Figure 2.2: Tanner graph of an IRA code.

2.2 Binary-Input Symmetric-Output Channels

Let X and Y be the input and output random variables of a channel with transition probability $p_{Y|X}(y|x)$. Let $\mathcal{X}$ be the (discrete) input alphabet and $\mathcal{Y}$ be the (discrete or continuous) output alphabet. Throughout this thesis, we will mainly deal with memoryless binary-input symmetric-output channels.

- The channel is *memoryless* if $p_{Y|X}(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p_{Y|X}(y_i|x_i)$, where $\mathbf{x}$ and $\mathbf{y}$ are of length n.

- The channel is *binary-input* if the input alphabet $\mathcal{X}$ has cardinality 2. Typically, $\mathcal{X}$ is either $\{0,1\}$ or $\{+1,-1\}$ as is the case for binary phase shift keying (BPSK). We let $\mathcal{X} = \{0,1\}$ from this moment on, unless otherwise stated.

- The channel is said to be *output-symmetric* if $p_{Y|X}(y|0) = p_{Y|X}(-y|1)$.

The channel is characterized by transition probabilities $p_{Y|X}(y|0)$ and $p_{Y|X}(y|1)$, for $x \in \{0,1\}$ and $y \in \mathcal{Y}$. If the output alphabet $\mathcal{Y}$ is discrete, then the transition probability is equal to the probability of the event $(Y = y|X = x)$

$$Pr(Y = y|X = x) = p_{Y|X}(y|x)$$

If the output alphabet $\mathcal{Y}$ is continuous (a subset of $\mathbb{R}$), the transition probability corresponds to the conditional probability density function (pdf) asso-

ciated to the channel

$$Pr(Y \in S|X = x) = \int_S p_{Y|X}(y|x)dy$$

where $S \subseteq \mathcal{Y} \subseteq \mathbb{R}$. The channel is further characterized by a noise-related parameter r given by

$$r \stackrel{\Delta}{=} -\log \int_{\mathcal{Y}} \sqrt{p_{Y|X}(y|0)p_{Y|X}(y|1)}dy \tag{2.4}$$

which is the exponent of the Bhattacharyya and Chernoff bounds [34]. For binary-input symmetric-output channels, the parameter r has an alternative expression shown in Section 2.4.

Here are examples of the parameter r for some useful binary-input symmetric-output channels.

Example 2.1 A BEC with erasure probability p has $r = -\log p$. ◊

Example 2.2 A BSC with crossover probability p has $r = -\log(2\sqrt{p(1-p)})$. ◊

Example 2.3 A BIAWGNC with Gaussian noise distribution $\mathcal{N}(0, \sigma^2)$ and

$$\begin{aligned} p_{Y|X}(y|0) &= \tfrac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-1)^2/2\sigma^2} \\ p_{Y|X}(y|1) &= \tfrac{1}{\sqrt{2\pi\sigma^2}} e^{-(y+1)^2/2\sigma^2} \end{aligned}$$

has $r = \frac{1}{2\sigma^2}$. ◊

2.3 Belief Propagation Decoding of IRA Codes

As maximum likelihood decoding is exponentially difficult for some graph codes, including IRA codes, for increasing block length, we consider BP message-passing decoding [35, 36, 37], which has a linear complexity in block length per decoder iteration. If a graph is cycle-free, then the message-passing BP algorithm computes exact marginal posterior probabilities.

In message-passing decoding algorithms, the graph nodes receive messages from their neighbors, compute new messages and forward them to their neighbors. The message output by a node u along an edge e is *extrinsic*

as it does not depend on the incoming message along the same edge e. The algorithm is defined by the code Tanner graph, by the set on which messages take on values, by the node computation rules and by the node activation scheduling method.

In BP decoding, messages take on values in the extended real line $\mathbb{R} \cup \{-\infty, \infty\}$. The BP decoder is initialized by setting all messages output by the checknodes to zero. Each bitnode α is associated with the *channel observation* message (log-likelihood ratio)

$$u_\alpha = \log \frac{p_{Y|X}(y_\alpha|x_\alpha = 0)}{p_{Y|X}(y_\alpha|x_\alpha = 1)} \tag{2.5}$$

where y_α is the channel output corresponding to the transmission of the code symbol x_α.

The BP node computation rules are given as follows. For a given node we identify an adjacent edge as *outgoing* and all other adjacent edges as *incoming*. Consider a bitnode α of degree i and let $m_1, \ldots, m_{i-1}$ denote the messages received from the $i-1$ incoming edges and u_α the associated channel observation message. The message $m_{o,\alpha}$ passed along the outgoing edge is given by

$$m_{o,\alpha} = u_\alpha + \sum_{j=1}^{i-1} m_j \tag{2.6}$$

Consider a checknode β of degree i and let $m_1, \ldots, m_{i-1}$ denote the messages received from the $i-1$ incoming edges. Then, the message $m_{o,\beta}$ passed along the outgoing edge is given by the following "tanh rule" [38]

$$\tanh \frac{m_{o,\beta}}{2} = \prod_{j=1}^{i-1} \tanh \frac{m_j}{2} \tag{2.7}$$

Taking the logarithm on each side of (2.7), we convert the product into a sum, and get

$$m_{o,\beta} = \gamma^{-1} \left(\gamma(m_1) + \cdots + \gamma(m_{i-1}) \right), \tag{2.8}$$

where the mapping $\gamma : \mathbb{R} \to \mathbb{F}_2 \times \mathbb{R}_+$ is defined by [19]

$$\gamma(z) = \left(\text{sign}(z), -\log \tanh \frac{|z|}{2} \right) \tag{2.9}$$

and where the sign function is defined as [19]

$$\text{sign}(z) = \begin{cases} 0 & \text{if } z > 0 \\ 0 & \text{with prob. } 1/2 \text{ if } z = 0 \\ 1 & \text{with prob. } 1/2 \text{ if } z = 0 \\ 1 & \text{if } z < 0 \end{cases}$$

Since the code Tanner graph has cycles, different scheduling methods yield in general non-equivalent BP algorithms. In this work we shall consider the following "classical" scheduling strategies:

- LDPC-like scheduling [15]. In this case, all bitnodes and all checknodes are activated alternately and in parallel. Every time a node is activated, it sends outgoing messages to all its neighbors. A decoding iteration (or "round" [39]) consists of the activation of all bitnodes and all checknodes.

- Turbo-like scheduling. Following [40], a good decoding scheduling consists of isolating large trellis-like subgraphs (or, more generally, normal realizations in Forney's terminology) and applying locally the forward-backward BCJR algorithm [4] (that implements efficiently the BP algorithm on normal cycle-free graphs), as done for Turbo codes [3]. A decoding iteration consists of activating all the information bitnodes in parallel (according to (2.6)) and of running the BCJR algorithm over the entire accumulator trellis. In particular, the checknodes do not send messages to the information bitnodes until the BCJR iteration is completed.

These two scheduling methods arise from the fact that IRA codes are subclasses of both LDPC codes and irregular turbo codes. Notice that for both of the above scheduling methods one decoder iteration corresponds to the activation of all information bitnodes in the graph exactly once.

2.4 Density Evolution and Stability

The bit error rate (BER) performance of BP decoding averaged over the IRA code ensemble can be analyzed, for any finite number ℓ of iterations and in the limit of $k \to \infty$, by the DE technique [19]. For a given bitnode α and iteration ℓ, the message sent over an outgoing edge (say edge e) is

a random variable that depends on the transmitted codeword, the channel noise and the interleaver (uniformly distributed over the set of permutations of N elements). The DE method finds the distribution of this random variable averaged over the channel noise and the interleaver, assuming that the block length goes to infinity. Under such an assumption, the probability that an oriented neighborhood of depth 2ℓ of the edge e contains cycles vanishes. Therefore, DE can be computed under the cycle-free condition, implying that the input messages at any node in the BP algorithm are statistically independent. For binary-input symmetric-output channels, the average message distributions do not depend on the transmitted codeword [39], so the transmission of the all-zero codeword can be assumed. The usefulness of the DE method stems from the *Concentration Theorem* [39, 18] which guarantees that, with high probability, the BER after ℓ iterations of the BP decoder applied to a randomly selected code in the ensemble and to a randomly generated channel noise sequence is close to the BER computed by DE with high probability, for sufficiently large block length.

Next, we formulate the DE for IRA codes and we study the stability condition of the fixed-point corresponding to zero BER. As in [19, section III-B], we introduce the space of *distributions* whose elements are non-negative non-decreasing right-continuous functions with range in $[0, 1]$ and domain the extended real line.

It can be shown that, for a binary-input symmetric-output channel, the distributions of messages at any iteration of the DE satisfy the symmetry condition [19]

$$\int h(x)dF(x) = \int e^{-x}h(-x)dF(x) \tag{2.10}$$

for any function h for which the integral exists. If F has density f, (2.10) is equivalent to

$$f(x) = e^{x}f(-x) \tag{2.11}$$

With some abuse of terminology, distributions satisfying (2.10) are said to be *symmetric*. The space of symmetric distributions will be denoted by $\mathcal{F}_{\text{sym}}$.

The bit error probability operator $\text{Pe} : \mathcal{F}_{\text{sym}} \to [0, 1/2]$ is defined by

$$\text{Pe}(F) = \frac{1}{2}(F^{-}(0) + F(0))$$

where $F^{-}(z)$ is the left-continuous version of $F(z)$. We introduce the "delta at zero" distribution, denoted by Δ_0, for which $\text{Pe}(\Delta_0) = 1/2$, and the "delta at infinity" distribution, denoted by Δ_∞, for which $\text{Pe}(\Delta_\infty) = 0$.

The symmetry property (2.10) implies that a sequence of symmetric distributions $\{F^{(\ell)}\}_{\ell=0}^{\infty}$ converges to Δ_∞ if and only if $\lim_{\ell\to\infty} \mathrm{Pe}(F^{(\ell)}) = 0$, where convergence of distributions is in the sense given in [19, Sect. III-F].

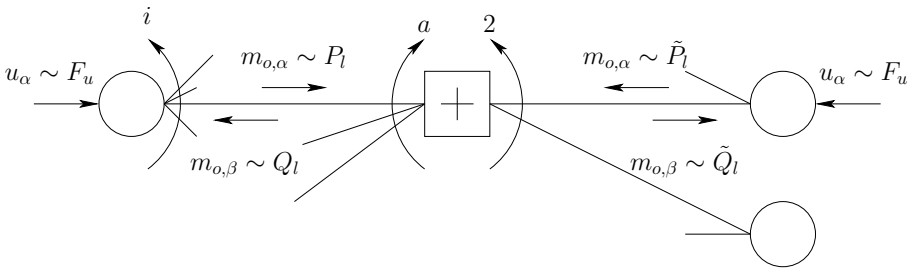


Figure 2.3: Message flow on the graph of a systematic IRA code

The DE for IRA code ensembles is given by the following proposition.

Proposition 2.4 *Let P_ℓ [resp., $\widetilde{P}_\ell$] denote the average distribution of messages passed from an information bitnode [resp., parity bitnode] to a checknode, at iteration ℓ . Let Q_ℓ [resp., $\widetilde{Q}_\ell$] denote the average distribution of messages passed from a checknode to an information bitnode [resp., parity bitnode], at iteration ℓ (see Fig. 2.3).*

Under the cycle-free condition, $P_\ell, \widetilde{P}_\ell, Q_\ell, \widetilde{Q}_\ell$ satisfy the following recursion:

$$P_\ell = F_u \otimes \lambda(Q_\ell) \tag{2.12}$$

$$\widetilde{P}_\ell = F_u \otimes \widetilde{Q}_\ell \tag{2.13}$$

$$Q_\ell = \Gamma^{-1}\left(\Gamma(\widetilde{P}_{\ell-1})^{\otimes 2} \otimes \Gamma(P_{\ell-1})^{\otimes(a-1)}\right) \tag{2.14}$$

$$\widetilde{Q}_\ell = \Gamma^{-1}\left(\Gamma(\widetilde{P}_{\ell-1}) \otimes \Gamma(P_{\ell-1})^{\otimes a}\right) \tag{2.15}$$

for $\ell = 1, 2, \ldots$, with initial condition $P_0 = \widetilde{P}_0 = \Delta_0$, where F_u denotes the distribution of the channel observation messages (2.5), $\otimes$ denotes convolution of distributions, defined by

$$(F \otimes G)(z) = \int F(z-t)dG(t) \tag{2.16}$$

$\otimes m$ denotes m-fold convolution, $\lambda(F) \stackrel{\Delta}{=} \sum_{i=2}^{d} \lambda_i F^{\otimes(i-1)}$, $\Gamma(F_x)$ is the distribution of $y = \gamma(x)$ (defined on $\mathbb{F}_2 \times \mathbb{R}_+$), when $x \sim F_x$, and Γ^{-1} denotes the

inverse mapping of Γ*, i.e.,* $\Gamma^{-1}(G_y)$ *is the distribution of* $x = \gamma^{-1}(y)$ *when* $y \sim G_y$.

Proof: See Appendix 2.A.

The DE recursion (2.12 – 2.15) is a two-dimensional non-linear dynamical system with state-space $\mathcal{F}^2_{\text{sym}}$ (i.e., the state trajectories of (2.12 – 2.15) are sequences of pairs of symmetric distributions $(P_\ell, \widetilde{P}_\ell)$). For this system, the BER at iteration ℓ is given by $\text{Pe}(P_\ell)$. A property of DE, given in Proposition 2.4, is that $\text{Pe}(P_\ell)$ is a non-negative and non-increasing function of ℓ [19, 39]. Hence, $\lim_{\ell\to\infty} \text{Pe}(P_\ell)$ exists.

It is easy to see that $(\Delta_\infty, \Delta_\infty)$ is a fixed-point of (2.12 – 2.15). The local stability of this fixed-point is given by the following result:

Theorem 2.5 *The fixed-point* $(\Delta_\infty, \Delta_\infty)$ *for the DE is locally stable if and only if*

$$\lambda_2 < \frac{e^r(e^r - 1)}{a + 1 + e^r(a - 1)} \tag{2.17}$$

where $r = -\log(\int e^{-z/2} dF_u(z))$.

Proof: See Appendix 2.B.

Here necessity and sufficiency are used in the sense of [19]. By following steps analogous to [19], it can be shown that if (2.17) holds, then there exists $\xi > 0$ such that if for some $\ell \in \mathbb{N}$, $\text{Pe}(RP_\ell(P_0, \widetilde{P}_0) + (1 - R)\widetilde{P}_\ell(P_0, \widetilde{P}_0)) < \xi$ then $\text{Pe}(RP_\ell + (1 - R)\widetilde{P}_\ell)$ converges to zero as ℓ tends to infinity. On the contrary, if λ_2 is strictly larger than the RHS of (2.17), then there exists $\xi > 0$ such that for all $\ell \in \mathbb{N}$ $\text{Pe}(RP_\ell(P_0, \widetilde{P}_0) + (1 - R)\widetilde{P}_\ell(P_0, \widetilde{P}_0)) > \xi$.

Consider a family of channels $\mathcal{C}(\nu) = \{p^\nu_{Y|X} : \nu \in \mathbb{R}_+\}$, where the channel parameter ν is, for example, an indicator of the noise level in the channel. Following [39], we say that $\mathcal{C}(\nu)$ is monotone with respect to the IRA code ensemble $(\{\lambda_i\}, a)$ under BP-decoding if, for any finite ℓ

$$\nu \leq \nu' \Leftrightarrow \text{Pe}(P_\ell) \leq \text{Pe}(P'_\ell)$$

where P_ℓ and P'_ℓ are the message distributions at iteration ℓ of DE applied to channels $p^\nu_{Y|X}$ and $p^{\nu'}_{Y|X}$, respectively.

Let $\text{BER}(\nu) = \lim_{\ell\to\infty} \text{Pe}(P_\ell)$, where $\{P_\ell\}$ is the trajectory of DE applied to the channel $p^\nu_{Y|X}$. The *threshold* $\nu^\star$ of the ensemble $(\{\lambda_i\}, a)$ over the

monotone family $\mathcal{C}(\nu)$ is the "worst" channel parameter for which the limiting BER is zero, i.e.,

$$\nu^\star = \sup\{\nu \geq 0 \ : \ \mathrm{BER}(\nu) = 0\} \tag{2.18}$$

We are interested in determining the thresholds of the ensemble $(\{\lambda_i\}, a)$ over different binary-input symmetric-output channels. Therefore, we optimize the IRA ensemble parameters $\{\lambda\}$ and a so as to maximize the threshold. Thus, for every value of ν, the optimal IRA ensemble parameters a and $\{\lambda_i\}$ maximize R subject to vanishing $\mathrm{BER}(\nu) = 0$, i.e., are solution of the optimization problem

$$\begin{cases} \text{maximize} & a \displaystyle\sum_{i=2}^{d} \lambda_i / i \\ \text{subject to} & \displaystyle\sum_{i=2}^{d} \lambda_i = 1, \quad \lambda_i \geq 0 \ \forall \ i \\ \text{and to} & \mathrm{BER}(\nu) = 0 \end{cases} \tag{2.19}$$

the solution of which can be found by some numerical techniques, as in [19]. However, the constraint $\mathrm{BER}(\nu) = 0$ is given directly in terms of the fixed-point of the DE recursion, making the optimization computationally very intensive. In the next chapter, we present low-complexity IRA code design methods, over a broad class of binary-input symmetric-output channels, and with performances close to the Shannon limit. These methods are based on replacing the infinite-dimensional message distribution with a one-dimensional quantity, rendering the optimization more tractable.

2.5 Conclusion

This chapter has presented the systematic IRA code ensemble encoder and its associated belief propagation decoder, in the limit of large code block length. This assumption allows to consider a cycle-free graph and enables to use DE to evaluate the densities of messages passed on the graph under message passing decoding. The threshold of the code ensemble can then be evaluated by iteratively calculating the message densities. We have derived a general stability condition for IRA codes under exact DE, which guarantees a vanishing BER if the error probability, at a given iteration, is small enough.

APPENDIX

2.A Proof of Proposition 2.4

The derivation of DE for IRA codes is analogous to the derivation of DE in [19] for irregular LDPC codes.

We are interested in the evolution of the distributions of messages given by (2.6) and (2.8), under the independence assumption (stated in Section 2.4).

(2.6) is the sum of independent random variables (log likelihood ratio messages). It can easily be shown that, if z_1 and z_2 are independent random variables, with distributions F_{z_1} and F_{z_2}, respectively, then the distribution of $z_1 + z_2$ is $F_{z_1} \otimes F_{z_2}$, the convolution of F_{z_1} and F_{z_2} defined in (2.16).

(2.8) involves the mapping γ and its inverse γ^{-1}. Given a random variable x with distribution $F_x(z)$, the distribution of $\gamma(x)$ is given by:

$$\Gamma(F_x)(s,z) = \chi_{\{s=0\}}\Gamma_0(F_x)(z) + \chi_{\{s=1\}}\Gamma_1(F_x)(z) \tag{2.20}$$

where

$$\Gamma_0(F_x)(z) = 1 - F_x^-(-\log\tanh\frac{z}{2}),$$

$$\Gamma_1(F_x)(z) = F_x(\log\tanh\frac{z}{2}),$$

and where $\chi_{\mathcal{A}}$ denotes the indicator function of the event $\mathcal{A}$.

By (2.8), the message passed from a check node to an information bitnode is the image under γ^{-1} of a sum of independent random variables (images of messages under γ), and therefore the distribution of their sum is the convolution of their distributions. The outgoing message from a checknode to a parity bitnode is the result of the operation (2.8) on a incoming messages from information bitnodes (with distribution $P_{\ell-1}$) and one incoming message from a parity bitnode (with distribution $\tilde{P}_{\ell-1}$). Therefore, the distribution of the message from a checknode to a parity bitnode is $Q_\ell = \Gamma^{-1}\left(\Gamma(\tilde{P}_{\ell-1}) \otimes \Gamma(P_{\ell-1})^{\otimes a}\right)$. Likewise, the outgoing message from a checknode to an information bitnode is the result of the operation (2.8) on $a-1$ incoming messages from information bitnodes and two incoming messages from a parity bitnode. Therefore, the distribution of the message from a checknode to a parity bitnode is $\tilde{Q}_\ell = \Gamma^{-1}\left(\Gamma(\tilde{P}_{\ell-1})^{\otimes 2} \otimes \Gamma(P_{\ell-1})^{\otimes(a-1)}\right)$.

A randomly chosen edge among the set of edges connected to information bitnodes, has probability λ_i to be connected to an information bitnode of

degree i. Then, by (2.6), the distribution of the output message of this degree-i information bitnode is $\lambda_i F_u \otimes Q_\ell^{\otimes(i-1)}$. Summing up over all instances of information bitnode degrees, we find that the distribution of the message from an information bitnode to a checknode is $P_\ell = F_u \otimes \sum_{i=2}^{d} \lambda_i Q_\ell^{\otimes(i-1)}$ which can be written as $P_\ell = F_u \otimes \lambda(Q_\ell)$. The distribution of the message from a parity bitnode to a checknode is simply the convolution of the channel observation message distribution and the distribution of the message from a checknode to a parity bitnode, i.e., $\tilde{P}_\ell = F_u \otimes \tilde{Q}_\ell$. □

2.B Proof of Theorem 2.5

We follow in the footsteps of [19] and analyze the local stability of the zero-BER fixed-point by using a small perturbation approach.

The mapping Γ applied to Δ_0 and Δ_∞ yields

$$\begin{aligned} \Gamma(\Delta_0)(s,z) &= \tfrac{1}{2}\chi_{\{s=0\}}\Delta_\infty(z) + \tfrac{1}{2}\chi_{\{s=1\}}\Delta_\infty(z) \\ \Gamma(\Delta_\infty)(s,z) &= \chi_{\{s=0\}}\Delta_0(z). \end{aligned} \tag{2.21}$$

Given $G(s,z) = \chi_{\{s=0\}}G_0(z) + \chi_{\{s=1\}}G_1(z)$, applying Γ^{-1} yields

$$\Gamma^{-1}(G)(z) = \chi_{\{z>0\}}(1 - G_0(-\log\tanh\frac{z}{2})) + \chi_{\{z<0\}}G_1(-\log\tanh\frac{-z}{2}) \tag{2.22}$$

For the sake of brevity, we introduce the short-hand notation

$$G(s,z) = \chi_{\{s=0\}}G_0(z) + \chi_{\{s=1\}}G_1(z) = \chi_0 G_0 + \chi_1 G_1$$

The m-fold convolution of $G(s,z)$ by itself is given by

$$\begin{aligned} &(\chi_0 G_0(z) + \chi_1 G_1(z))^{\otimes m} = \\ &\chi_0 \left(\sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \tbinom{m}{2j} G_0^{\otimes(m-2j)} \otimes G_1^{\otimes 2j} \right) + \chi_1 \left(\sum_{j=0}^{\lfloor \frac{m-1}{2} \rfloor} \tbinom{m}{2j+1} G_0^{\otimes(m-2j-1)} \otimes G_1^{\otimes 2j+1} \right) \end{aligned} \tag{2.23}$$

where $\lfloor . \rfloor$ stands for the integer part.

In order to study the local stability of the fixed-point $(\Delta_\infty, \Delta_\infty)$, we initialize the DE recursion at the point

$$\begin{cases} P_0 &= (1-2\epsilon)\Delta_\infty + 2\epsilon\Delta_0 \\ \widetilde{P}_0 &= (1-2\delta)\Delta_\infty + 2\delta\Delta_0 \end{cases}$$

for some small $\epsilon, \delta > 0$, and we apply one iteration of the DE recursion (2.12 – 2.15). The step-by-step derivation is as follows. From (2.21) we have

$$\begin{cases} \Gamma(P_0) &= \chi_0\left((1-2\epsilon)\Delta_0 + \epsilon\Delta_\infty\right) + \chi_1\left(\epsilon\Delta_\infty\right) \\ \Gamma(\widetilde{P}_0) &= \chi_0\left((1-2\delta)\Delta_0 + \delta\Delta_\infty\right) + \chi_1\left(\delta\Delta_\infty\right) \end{cases}$$

By applying (2.23) we obtain

$$\begin{cases} \Gamma(P_0)^{\otimes n} &= \chi_0\left((1-2n\epsilon)\Delta_0 + n\epsilon\Delta_\infty\right) + \chi_1\left(n\epsilon\Delta_\infty\right) + O(\epsilon^2) \\ \Gamma(\widetilde{P}_0)^{\otimes 2} &= \chi_0\left((1-4\delta)\Delta_0 + 2\delta\Delta_\infty\right) + \chi_1\left(2\delta\Delta_\infty\right) + O(\delta^2) \end{cases}$$

By applying Γ^{-1} we get

$$\begin{cases} Q_1 &= \Gamma^{-1}\left(\Gamma(P_0)^{\otimes(a-1)} \otimes \Gamma(\widetilde{P}_0)^{\otimes 2}\right) \\ \widetilde{Q}_1 &= \Gamma^{-1}\left(\Gamma(P_0)^{\otimes a} \otimes \Gamma(\widetilde{P}_0)\right) \end{cases}$$

and

$$\begin{cases} Q_1 &= (1-2(a-1)\epsilon - 4\delta)\Delta_\infty + (2(a-1)\epsilon + 4\delta)\Delta_0 + O(\epsilon^2, \delta^2) \\ \widetilde{Q}_1 &= (1-2a\epsilon - 2\delta)\Delta_\infty + (2a\epsilon + 2\delta)\Delta_0 + O(\epsilon^2, \delta^2) \end{cases}$$

Hence, by noticing that

$$\begin{aligned} Q_1^{\otimes n} &= \sum_{j=0}^{n} \binom{n}{j} (1-2(a-1)\epsilon - 4\delta)^{n-j} (2(a-1)\epsilon + 4\delta)^j \Delta_\infty^{\otimes n-j} \otimes \Delta_0^{\otimes j} + O(\epsilon^2, \delta^2) \\ &= \begin{cases} \Delta_\infty + O(\epsilon^2, \delta^2), & \text{for } n \geq 2 \\ (1-2(a-1)\epsilon - 4\delta)\Delta_\infty + (2(a-1)\epsilon + 4\delta)\Delta_0 + O(\epsilon^2, \delta^2), & \text{for } n = 1 \end{cases} \end{aligned}$$

we have

$$\lambda(Q_1) = (1-2(a-1)\lambda_2\epsilon - 4\lambda_2\delta)\Delta_\infty + (2(a-1)\lambda_2\epsilon + 4\lambda_2\delta)\Delta_0 + O(\epsilon^2, \delta^2).$$

Finally, by using the fact that $P_1 = F_u \otimes \lambda(Q_1)$ and that $\widetilde{P}_1 = F_u \otimes \widetilde{Q}_1$, the message distributions after one DE iteration are given by

$$\begin{bmatrix} P_1 \\ \widetilde{P}_1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} 2\epsilon \\ 2\delta \end{bmatrix} F_u + \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \mathbf{A} \begin{bmatrix} 2\epsilon \\ 2\delta \end{bmatrix} \right) \Delta_\infty + \begin{bmatrix} O(\epsilon^2) \\ O(\delta^2) \end{bmatrix}$$

where

$$\mathbf{A} = \begin{bmatrix} (a-1)\lambda_2 & 2\lambda_2 \\ a & 1 \end{bmatrix} \tag{2.24}$$

After ℓ iterations we obtain

$$\begin{bmatrix} P_\ell \\ \widetilde{P}_\ell \end{bmatrix} = \mathbf{A}^\ell \begin{bmatrix} 2\epsilon \\ 2\delta \end{bmatrix} F_u^{\otimes \ell} + \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \mathbf{A}^\ell \begin{bmatrix} 2\epsilon \\ 2\delta \end{bmatrix} \right) \Delta_\infty + \begin{bmatrix} O(\epsilon^2) \\ O(\delta^2) \end{bmatrix} \tag{2.25}$$

Then, by applying $\mathrm{Pe}(\cdot)$ to P_ℓ in (2.25) we obtain that $\lim_{\ell\to\infty} \mathrm{Pe}(P_\ell) = 0$ (implying that $\lim_{\ell\to\infty} P_\ell = \Delta_\infty$) if the eigenvalues of the matrix $\mathbf{A}e^{-r}$ are inside the unit circle. r is given by the large deviation theory as [19]

$$\begin{aligned} r &= -\lim_{\ell\to\infty} \frac{1}{\ell} \log \mathrm{Pe}(F_u^{\otimes\ell}) \\ &= -\log\left(\inf_{s>0} \int e^{-sz} dF_u(z) \right) \\ &= -\log\left(\int e^{-z/2} dF_u(z) \right) \end{aligned} \tag{2.26}$$

where the last equality follows from the fact that $F_u(z) \in \mathcal{F}_{\mathrm{sym}}$. Note that r coincides with the Chernoff bound exponent given by (2.4), which can easily be seen using the symmetry of F_u.

The stability condition is obtained by explicitly computing the largest eigenvalue in magnitude. We obtain

$$\frac{1}{2}\left(1 + \lambda_2(a-1) + \sqrt{1 + (2+6a)\lambda_2 + (a-1)^2\lambda_2^2} \right) < e^r. \tag{2.27}$$

Since the LHS of (2.27) is increasing, condition (2.27) is indeed an upper bound on λ_2, given explicitly by (2.17). □

Chapter 3

Design of Irregular Repeat Accumulate Code Ensembles

In this chapter, we address the optimization of the IRA code ensemble for the class of binary-input symmetric-output channels. We propose and compare four low-complexity ensemble optimization methods which lead to optimization problems that are solvable by linear programming. Our approach to IRA code design is based on the following tools: the EXtrinsic mutual Information Transfer function and its analytical properties [20, 41, 42], reciprocal channel (duality) approximation [41, 43], and the non-strict convexity of mutual information. We design code ensembles for various rates on the BIAWGNC and the BSC and evaluate the thresholds of the codes thus designed using the true density evolution.

3.1 IRA Ensemble Optimization

A variety of methods have been developed in order to simplify the code ensemble optimization [15, 22, 43, 44, 45, 46]. They consist of replacing the DE with a dynamical system defined over the reals (rather than over the space of distributions), whose trajectories and fixed-points are related in some way to the trajectories and fixed-point of the DE. The only other

work that has proposed a method to design IRA codes is [15, 47] where the design focuses on the choice of the grouping factor and the repetition profile. The method used in [15] to choose the repetition profile was based on the infinite-block length Gaussian Approximation of message passing decoding proposed in [22].

Essentially, all proposed approximated DE methods can be formalized as follows. Let $\Phi : \mathcal{F}_{\text{sym}} \to \mathbb{R}$ and $\Psi : \mathbb{R} \to \mathcal{F}_{\text{sym}}$ be mappings of the set of symmetric distributions to the real numbers and vice versa. Then, a dynamical system with state-space $\mathbb{R}^2$ can be derived from (2.12 – 2.15) as

$$\begin{aligned} x_\ell &= \Phi\left(F_u \otimes \lambda\left(\mathsf{Q}_\ell\right)\right) && (3.1)\\ \widetilde{x}_\ell &= \Phi\left(F_u \otimes \widetilde{\mathsf{Q}}_\ell\right) && (3.2)\\ \mathsf{Q}_\ell &= \Gamma^{-1}\left(\Gamma\left(\Psi(\widetilde{x}_{\ell-1})\right)^{\otimes 2} \otimes \Gamma\left(\Psi(x_{\ell-1})\right)^{\otimes(a-1)}\right) && (3.3)\\ \widetilde{\mathsf{Q}}_\ell &= \Gamma^{-1}\left(\Gamma\left(\Psi(\widetilde{x}_{\ell-1})\right) \otimes \Gamma\left(\Psi(x_{\ell-1})\right)^{\otimes a}\right) && (3.4) \end{aligned}$$

for $\ell = 1, 2, \ldots$, with initial condition $x_0 = \widetilde{x}_0 = \Phi(\Delta_0)$, and where $(x_\ell, \widetilde{x}_\ell)$ are the system state variables.

By eliminating the intermediate distributions Q_ℓ and $\widetilde{\mathsf{Q}}_\ell$, we can put (3.1 – 3.4) in the form

$$\begin{aligned} x_\ell &= \phi(x_{\ell-1}, \widetilde{x}_{\ell-1}) \\ \widetilde{x}_\ell &= \widetilde{\phi}(x_{\ell-1}, \widetilde{x}_{\ell-1}) \end{aligned} \qquad (3.5)$$

For all DE approximations considered in this work, the mappings Φ and Ψ and the functions ϕ and $\widetilde{\phi}$ satisfy the following desirable properties:

1. $\Phi(\Delta_0) = 0$, $\Phi(\Delta_\infty) = 1$.

2. $\Psi(0) = \Delta_0$, $\Psi(1) = \Delta_\infty$.

3. ϕ and $\widetilde{\phi}$ are defined on $[0, 1] \times [0, 1]$ and have range in $[0, 1]$.

4. $\phi(0, 0) > 0$ and $\widetilde{\phi}(0, 0) > 0$.

5. $\phi(1, 1) = \widetilde{\phi}(1, 1) = 1$, i.e., $(1, 1)$ is a fixed-point of the recursion (3.5). Moreover, this fixed-point corresponds to the zero-BER fixed-point $(\Delta_\infty, \Delta_\infty)$ of the exact DE.

6. If $F_u \neq \Delta_0$, the function $\widetilde{\phi}(x,\widetilde{x}) - \widetilde{x}$ is strictly decreasing in $\widetilde{x}$ for all $x \in [0,1]$. Therefore, the equation

$$\widetilde{x} = \widetilde{\phi}(x,\widetilde{x})$$

has a unique solution in $[0,1]$ for all $x \in [0,1]$. This solution will be denoted by $\widetilde{x}(x)$.

It follows that all fixed-points of (3.5) must satisfy

$$x = \phi(x,\widetilde{x}(x)) \tag{3.6}$$

and that in order to avoid fixed-points other than $(1,1)$, (3.6) must not have solutions in the interval $[0,1)$, i.e., it must satisfy

$$x < \phi(x,\widetilde{x}(x)), \quad \forall \ x \in [0,1) \tag{3.7}$$

Notice that, in general, (3.7) is neither a necessary nor a sufficient condition for the uniqueness of the zero-BER fixed-point of the exact DE. However, if the choice of the DE approximation is good, this provides a heuristic for the code ensemble optimization.

By replacing the constraint $\mathrm{BER}(\nu) = 0$ by (3.7) in (2.19), we obtain the *approximated* IRA ensemble optimization method as

$$\begin{cases} \text{maximize} & a \displaystyle\sum_{i=2}^{d} \lambda_i / i \\ \text{subject to} & \displaystyle\sum_{i=2}^{d} \lambda_i = 1, \quad \lambda_i \geq 0 \ \forall \ i \\ \text{and to} & x < \phi(x,\widetilde{x}(x)), \quad \forall \ x \in [0,1) \end{cases} \tag{3.8}$$

Approximations of the DE recursion differ essentially in the choice of Φ and Ψ, and in the way the *intermediate* distributions Q_ℓ and $\widetilde{\mathsf{Q}}_\ell$ and the channel message distribution F_u are approximated.

3.2 EXIT Functions

Several recent works show that DE can be accurately described in terms of the evolution of the mutual information between the variables associated with the bitnodes and their messages (see [20, 48, 21, 49, 42, 50, 26]).

The key idea in order to approximate DE by mutual information evolution is to describe each computation node in BP-decoding by a *mutual information transfer function*. For historical reasons, this function is usually referred to as the EXtrinsic mutual Information Transfer (EXIT) function.

EXIT functions are generally defined as follows. Consider the model of Fig. 3.1, where the box represents a generalized computation node of the BP algorithm (i.e., it might contain a subgraph formed by several nodes and edges, and might depend on some other random variables such as channel observations, not shown in Fig. 3.1). Let $m_1, \ldots, m_{i-1}$ denote the input messages, assumed independent and identically distributed (i.i.d.) $\sim F_{\text{in}}$, and let $m_o \sim F_{\text{out}}$ denote the output message. Let X_j denote the binary code symbol associated with message m_j, for $j = 1, \ldots, i-1$, and let X denote the binary code symbol associated with message m_o. Since $F_{\text{in}}, F_{\text{out}} \in \mathcal{F}_{\text{sym}}$, we can think of m_j and m_o as the outputs of binary-input symmetric-output channels with inputs X_j and X and transition probabilities

$$P(m_j \leq z | X_j = 0) = F_{\text{in}}(z) \tag{3.9}$$

$$P(m_o \leq z | X = 0) = F_{\text{out}}(z), \tag{3.10}$$

respectively.

Channel (3.9) models the *a priori* information that the node receives about the symbols X_j's, and the channel (3.10) models the *extrinsic information* [3] that the node generates about the symbol X.

We define the binary-input symmetric-output capacity functional $\mathfrak{I} : \mathcal{F}_{\text{sym}} \to [0, 1]$, such that

$$\mathfrak{I}(F) \triangleq 1 - \int_{-\infty}^{\infty} \log_2 \left(1 + e^{-z}\right) dF(z) \tag{3.11}$$

Namely, $\mathfrak{I}$ maps any symmetric distribution F into the capacity [1] of the binary-input symmetric-output channel with transition probability $p_{Y|X}(y|0) = F(y)$.

Then, we let

$$I_A = I(X_j; m_j) = \mathfrak{I}(F_{\text{in}})$$

$$I_E = I(X; m_o) = \mathfrak{I}(F_{\text{out}})$$

[1] Recall that the capacity of a binary-input symmetric-output memoryless channel is achieved by uniform i.i.d. inputs.

denote the capacities of the channels (3.9) and (3.10), respectively. The EXIT function of the node of Fig. 3.1 is the set of pairs (I_A, I_E), for all $I_A \in [0,1]$ and for some (arbitrary) choice of the input distribution F_{in} such that $\mathcal{I}(F_{\text{in}}) = I_A$. Notice that the EXIT function of a node is not uniquely defined, since it depends on the choice of F_{in}. In general, different choices yield different transfer functions.

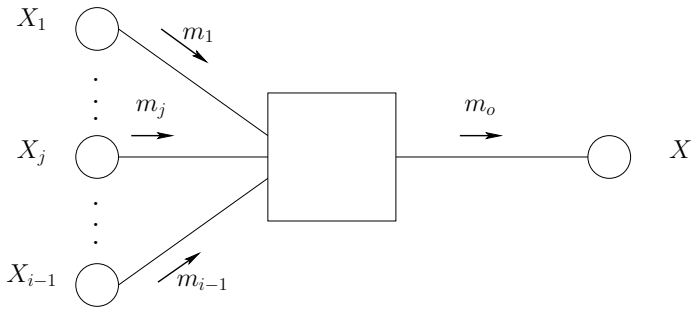


Figure 3.1: EXIT model

The approximations of the DE considered in this work are based on EXIT functions, and track the evolution of the mutual information between the messages output by the bitnodes and the associated code symbols.

3.3 Some Properties of Binary-Input Symmetric-Output Channels

3.3.1 Property 1

Consider a binary-input symmetric-output channel with probability distribution function $p_{Y|X}(y|0) = G(y)$, where G is not necessarily symmetric, in the sense of (2.10). Let G have a density $g(y)$. The channel capacity can be written as

$$C = 1 - \int_{-\infty}^{\infty} \log_2 \left(1 + \frac{g(-z)}{g(z)}\right) g(z) dz \tag{3.12}$$

By concatenating the transformation $y \mapsto u = \log \frac{p_{Y|X}(y|0)}{p_{Y|X}(y|1)}$ to the channel output, we obtain a new binary-input symmetric-output channel with $p'_{U|X}(u|0) = F(u)$ such that $F \in \mathcal{F}_{\text{sym}}$. Moreover, since U is a sufficient

statistic for Y, the original channel has the same capacity as the new channel, given by $C = \mathcal{I}(F)$. Therefore, by defining appropriately the channel output, the capacity of any binary-input symmetric-output channel can always be put in the form (3.11).

3.3.2 Property 2

The following is an interesting property of symmetric distributions.

Proposition 3.1 *Let X be binary with $P[X = 0] = p$ and $P[X = 1] = 1-p$. Let S be independent of X and take M (finite) values with $P[S = i] = q_i$. Conditioned on $S = i$, Y is a continuous random variable with conditional density function*

$$f^{(i)}_{Y|X=1}(y) = e^{-y} f^{(i)}_{Y|X=0}(y)$$

Then

$$I(X;Y|S) = I(X;Y)$$

Proof: See Appendix 3.A.

The following corollary restricts the above result to binary-input symmetric-output channels.

Corollary 3.2 *The mutual information functional is not strictly convex on the set of binary-input symmetric-output channels with transition probability $p_{Y|X}(y|0) \in \mathcal{F}_{\text{sym}}$.*

Proof: See Appendix 3.A.

3.3.3 Property 3

The third property gives a lower bound on the capacity of binary-input symmetric-output channels as a function of the Chernoff bound exponent.

Lemma 3.3 *The capacity of a binary-input symmetric-output channel, with transition probability $p_{Y|X}(y|0) = F(y)$, is upper-bounded by*

$$C(F) \geq \frac{1}{2}\left(r - r^2\right) \log_2 e \tag{3.13}$$

where $r = -\log(\int e^{-z/2} dF(z))$.

Proof: See Appendix. 3.E

3.4 DE Approximation Methods

3.4.1 Method 1

The first approximation of the DE considered in this work assumes that the distributions at any iteration are Gaussian. A Gaussian distribution satisfies the symmetry condition (2.11) if and only if its variance is equal to twice the absolute value of its mean. We introduce the short-hand notation $\mathcal{N}_{\text{sym}}(\mu)$ to denote the symmetric Gaussian distribution (or density, depending on the context) with mean μ, i.e., $\mathcal{N}_{\text{sym}}(\mu) \triangleq \mathcal{N}(\mu, 2|\mu|)$.

For a distribution $F \in \mathcal{F}_{\text{sym}}$, we let the mapping Φ be equal to $\mathcal{I}$ defined in (3.11), and for all $x \in [0,1]$ we define the mapping

$$\Psi : x \mapsto \mathcal{N}_{\text{sym}}(J^{-1}(x)) \tag{3.14}$$

where

$$J(\mu) \triangleq \mathcal{I}(\mathcal{N}_{\text{sym}}(\mu)) = 1 - \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-z^2} \log_2\left(1 + e^{-2\sqrt{\mu}z - \mu}\right) dz \tag{3.15}$$

Namely, Ψ maps $x \in [0,1]$ into the symmetric Gaussian distribution $\mathcal{N}_{\text{sym}}(\mu)$ such that the BIAWGNC with transition probability $p_{Y|X}(y|0) = \mathcal{N}_{\text{sym}}(\mu)$ has capacity x. Note that $J(\mu)$ is the capacity of a BIAWGNC with signal to noise ratio $E_s/N_0 = |\mu|/4$.

The first key approximation in Method 1 is

$$\begin{aligned} \mathsf{Q}_\ell &\approx \mathcal{N}_{\text{sym}}(\mu_\ell) \\ \widetilde{\mathsf{Q}}_\ell &\approx \mathcal{N}_{\text{sym}}(\widetilde{\mu}_\ell) \end{aligned} \tag{3.16}$$

for some $\mu_\ell, \widetilde{\mu}_\ell \geq 0$.

In order to compute μ_ℓ and $\widetilde{\mu}_\ell$ we make use of the reciprocal channel approximation [43] also called *approximate* duality property of EXIT functions in [41]. This states that the EXIT function of a checknode is accurately approximated by the EXIT function of a bitnode with the same degree after the change of variables $I_A \mapsto 1 - I_A$ and $I_E \mapsto 1 - I_E$ (see Fig. 3.2). Using approximate duality, we replace the checknode by a bitnode and change $(x_{\ell-1}, \widetilde{x}_{\ell-1})$ into $(1 - x_{\ell-1}, 1 - \widetilde{x}_{\ell-1})$. Since the output message from a bitnode is the sum of the input messages (see eq. (2.6)), and since the input distributions

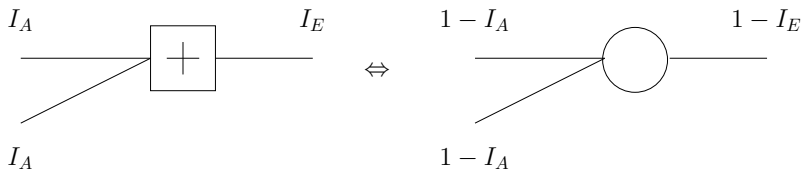


Figure 3.2: Reciprocal (dual) channel approximation

$\Psi(1 - x_{\ell-1})$ and $\Psi(1 - \widetilde{x}_{\ell-1})$) are Gaussian, then the output distribution is Gaussian too, with mean

$$(a-1)J^{-1}(1 - x_{\ell-1}) + 2J^{-1}(1 - \widetilde{x}_{\ell-1})$$

for messages sent to information bitnodes and

$$aJ^{-1}(1 - x_{\ell-1}) + J^{-1}(1 - \widetilde{x}_{\ell-1})$$

for messages sent to parity bitnodes. Finally, μ_ℓ and $\widetilde{\mu}_\ell$ are given by

$$\begin{cases} \mu_\ell &= J^{-1}\left(1 - J\left((a-1)J^{-1}(1 - x_{\ell-1}) + 2J^{-1}(1 - \widetilde{x}_{\ell-1})\right)\right) \\ \widetilde{\mu}_\ell &= J^{-1}\left(1 - J\left(aJ^{-1}(1 - x_{\ell-1}) + J^{-1}(1 - \widetilde{x}_{\ell-1})\right)\right) \end{cases} \tag{3.17}$$

The second key approximation in Method 1 is to replace F_u with a discrete symmetric distribution such that

$$F_u \approx \sum_{j=1}^{D} p_j \Delta_{v_j} \tag{3.18}$$

for some integer $D \geq 2$, $v_j \in \mathbb{R}$ and $p_j \in \mathbb{R}_+$ such that $\sum_{j=1}^{D} p_j = 1$.

With this assumption, from the definition (3.11) of the operator $\mathfrak{I}$, from Corollary 3.2 and since [19]: a) the convolution of symmetric distributions is symmetric, and b) the convex combination of symmetric distributions is symmetric, it is immediate to write (3.1) and (3.2) as

$$\begin{cases} x_\ell &= 1 - \sum_{i=2}^{d} \sum_{j=1}^{D} \lambda_i p_j \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-z^2} \log_2\left(1 + e^{-2\sqrt{(i-1)\mu_\ell} z - (i-1)\mu_\ell - v_j}\right) dz \\ \widetilde{x}_\ell &= 1 - \sum_{j=1}^{D} p_j \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-z^2} \log_2\left(1 + e^{-2\sqrt{\widetilde{\mu}_\ell} z - \widetilde{\mu}_\ell - v_j}\right) dz \end{cases} \tag{3.19}$$

The desired DE approximation in the form (3.5) is obtained (implicitly) by combining (3.17) and (3.19). Notice that (3.19) is linear in the repetition profile and the optimization problem (3.8) can be solved as linear programming.

Example 3.4 Discrete-output channels. In general, when the channel output is discrete then the approximation (3.18) holds exactly. For example, for the BSC with transition probability p we have

$$F_u = p\Delta_{-\log\frac{1-p}{p}} + (1-p)\Delta_{\log\frac{1-p}{p}}$$

$\Diamond$

Example 3.5 The BIAWGNC defined by $y = (-1)^x + z$, where $z \sim \mathcal{N}(0, \sigma^2)$, is a channel such that

$$F_u = \mathcal{N}_{\text{sym}}(2/\sigma^2) \tag{3.20}$$

In this case, since convolving symmetric Gaussian distributions yields a symmetric Gaussian distribution whose mean is the sum of the means, the discretization approximation (3.18) is not necessary and we have

$$\begin{array}{rcl} F_u \otimes \lambda(\mathsf{Q}_\ell) & = & \displaystyle\sum_{i=2}^{d} \lambda_i \mathcal{N}_{\text{sym}}(2/\sigma^2 + (i-1)\mu_\ell) \\ F_u \otimes \widetilde{\mathsf{Q}}_\ell & = & \mathcal{N}_{\text{sym}}(2/\sigma^2 + \widetilde{\mu}_\ell) \end{array} \tag{3.21}$$

By applying the operator $\mathfrak{I}$, using (3.15) and Corollary 3.2, we obtain the DE approximation for the BIAWGNC as

$$\begin{cases} x_\ell & = & \displaystyle\sum_{i=2}^{d} \lambda_i J\Big(\frac{2}{\sigma^2} + (i-1)J^{-1}\big(1 - J((a-1)J^{-1}(1-x_{\ell-1}) + \\ & & \qquad 2J^{-1}(1-\widetilde{x}_{\ell-1}))\big)\Big) \\ \widetilde{x}_\ell & = & J\left(\frac{2}{\sigma^2} + J^{-1}\left(1 - J\left(aJ^{-1}(1-x_{\ell-1}) + J^{-1}(1-\widetilde{x}_{\ell-1})\right)\right)\right) \end{cases} \tag{3.22}$$

$\Diamond$

3.4.2 Method 2

The second approximation of the DE assumes that the distributions of messages at any iteration consist of two mass points, one at zero and the other at $+\infty$. For such distributions, we introduce the short-hand notation $\mathcal{E}_{\text{sym}}(\epsilon) \stackrel{\Delta}{=} \epsilon\Delta_0 + (1-\epsilon)\Delta_\infty$.

We let the mapping Φ be equal to $\mathcal{I}$ defined in (3.11) and the mapping Ψ be

$$\Psi : x \mapsto \mathcal{E}_{\text{sym}}(1-x) \tag{3.23}$$

for all $x \in [0,1]$. Namely, Ψ maps $x \in [0,1]$ into the symmetric distribution $\mathcal{E}_{\text{sym}}(x)$ of a BEC with erasure probability $1-x$.

With these mappings, (3.3 – 3.4) can be put in the form

$$\begin{array}{rl} \mathsf{Q}_\ell & = \mathcal{E}_{\text{sym}}(1 - x_{\ell-1}^{a-1}\widetilde{x}_{\ell-1}^2) \\ \widetilde{\mathsf{Q}}_\ell & = \mathcal{E}_{\text{sym}}(1 - x_{\ell-1}^{a}\widetilde{x}_{\ell-1}) \end{array} \tag{3.24}$$

where we used the fact that, as it can be seen from the definitions of Γ and Γ^{-1} in (2.20 – 2.22),

$$\Gamma^{-1}\left(\Gamma(\mathcal{E}_{\text{sym}}(\epsilon_1)) \otimes \Gamma(\mathcal{E}_{\text{sym}}(\epsilon_2))\right) = \mathcal{E}_{\text{sym}}(1-(1-\epsilon_1)(1-\epsilon_2)) \tag{3.25}$$

Notice that, while Q_ℓ and $\widetilde{\mathsf{Q}}_\ell$ in Method 1 were *assumed* to be symmetric Gaussian (see (3.16)), here (3.24) holds exactly.

As a consequence of these mappings, the communication channel of the parity bits, with distribution F_u, is replaced by a BEC with erasure probability $\epsilon = 1 - \mathcal{I}(F_u)$. Furthermore, for any $F \in \mathcal{F}_{\text{sym}}$, we have

$$\mathcal{I}(F \otimes \mathcal{E}_{\text{sym}}(\epsilon)) = 1 - (1-\mathcal{I}(F))\epsilon \tag{3.26}$$

From this result, it is immediate to obtain the approximated DE recursion as

$$\left\{ \begin{array}{rl} x_\ell & = 1 - (1-\mathcal{I}(F_u)) \displaystyle\sum_{i=2}^{d} \lambda_i \left(1 - x_{\ell-1}^{a-1}\widetilde{x}_{\ell-1}^2\right)^{i-1} \\ \widetilde{x}_\ell & = 1 - (1-\mathcal{I}(F_u))\left(1 - x_{\ell-1}^{a}\widetilde{x}_{\ell-1}\right) \end{array} \right. \tag{3.27}$$

Notice that (3.27) is the standard (exact) DE for the IRA ensemble $(\{\lambda_i\}, a)$ over a BEC (see [15]) with the same capacity of the actual binary-input symmetric-output channel, given by $\mathcal{I}(F_u)$. We point out here that this

method, consisting of replacing the actual channel with a BEC with equal capacity and optimizing the code ensemble for the BEC (physical channel), was proposed in [43] for the optimization of LDPC ensembles. Interestingly, this method follows as a special case of our general approach for DE approximation, for a particular choice of the mappings Φ and Ψ.

Assume that we are at a fixed-point $(x, \widetilde{x})$ of the decoding algorithm and we solve for x. From (3.27), $\widetilde{x}$ is a function of x given by

$$\widetilde{x} = \frac{\mathcal{I}(F_u)}{1 - (1 - \mathcal{I}(F_u))x^a} \tag{3.28}$$

Then from (3.27), we obtain the fixed-point equation corresponding to (3.6) in closed form as

$$x = 1 - (1 - \mathcal{I}(F_u)) \sum_{i=2}^{d} \lambda_i \left(1 - \frac{x^{a-1}\mathcal{I}(F_u)^2}{(1 - (1 - \mathcal{I}(F_u))x^a)^2}\right)^{i-1} \tag{3.29}$$

3.4.3 Methods 3 and 4

Methods 1 and 2 yield (almost) closed-form DE approximations at the price of some approximations of the message distributions and, above all, of the checknode output distributions Q_ℓ and $\widetilde{\mathsf{Q}}_\ell$.

In much of the current literature on random-like code ensemble optimization, the EXIT function of a decoding block is obtained by Monte Carlo simulation, by generating i.i.d. input messages, estimating the distribution of the output messages and computing a one-dimensional quantity [20, 21, 22, 23, 24, 25, 26]. Following this approach, we shall consider the IRA decoder with Turbo-like scheduling (see Fig. 3.3) and obtain the EXIT functions of the inner and outer decoders.

The inner (accumulator) and outer (repetition) decoders are characterized by an EXIT function as defined in Section 3.2, for some guess of the symmetric distribution F_{in}. In general, the EXIT function of the decoders can be obtained as follows:

1. Let the channel observation messages be i.i.d., with distribution F_u.
2. Assume the decoder input messages are i.i.d., with distribution F_{in}.
3. Obtain either in closed form or by Monte Carlo simulation the corresponding marginal distribution F_{out} of the decoder output messages.

4. Let $I_A = \mathcal{I}(F_{\text{in}})$, $I_E = \mathcal{I}(F_{\text{out}})$ be a point on the EXIT function curve.

Our Methods 3 and 4 consist of applying the above approach under the assumptions $F_{\text{in}} = \mathcal{N}_{\text{sym}}(J^{-1}(I_A))$ and $F_{\text{in}} = \mathcal{E}_{\text{sym}}(1 - I_A)$, respectively.

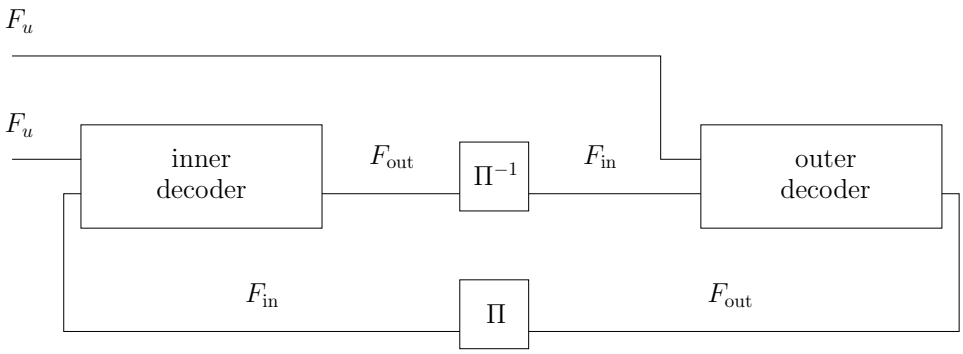


Figure 3.3: Turbo-like IRA decoder

Example 3.6 Method 3 with $a > 1$. The complexity of the BCJR decoder becomes prohibitive as a increases. Using the reciprocal channel approximation, the determination of the EXIT function of the accumulator is easier. The accumulator can be seen as the serial concatenation of a single parity check code and a 2-state convolutional code of rate 1 (Fig. 3.4). Let $I_a = \mathcal{I}(F_a)$ and $I_e = \mathcal{I}(F_e)$, where F_a (resp. F_e) is the distribution of the input (resp. output) message of the convolutional code. The EXIT function of the convolutional decoder, denoted by $I_e = f(I_a)$, is obtained by Monte Carlo simulation (see Appendix 3.B). Using the reciprocal channel approximation, the (I_A, I_E) point on the EXIT function curve of the accumulator is then given by

$$I_E = 1 - J\left((a-1)J^{-1}(1-I_A) + J^{-1}\left(1 - f\left(J(aJ^{-1}(I_A))\right)\right)\right) \tag{3.30}$$

$\Diamond$

Let the resulting EXIT functions of the inner and outer decoders be denoted by $I_E = g(I_A)$ and by $I_E = h(I_A)$, respectively, and let x denote the mutual information between the messages at the output of the outer decoder (repetition code) and the corresponding symbols (information bitnodes).

The resulting approximated DE is given by

$$x_\ell = h(g(x_{\ell-1})) \tag{3.31}$$

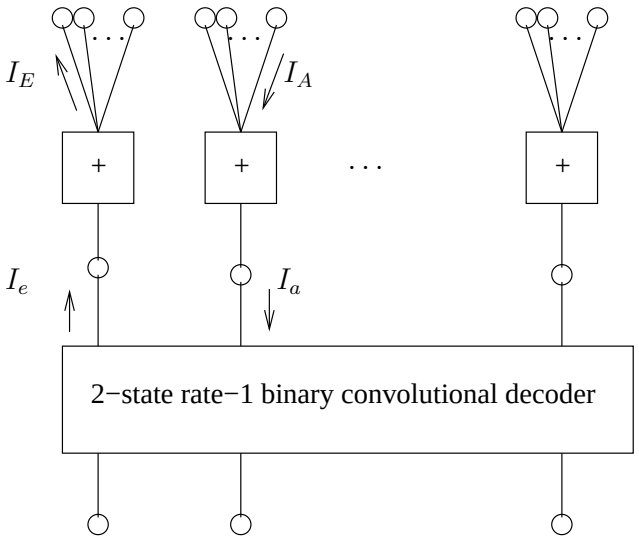


Figure 3.4: Accumulator as the serial concatenation of a single parity check code and a 2-state convolutional code

The corresponding fixed-point equation is given by $x = h(g(x))$, and the condition for the uniqueness of the fixed point at $x = 1$, corresponding to (3.7), is $x < h(g(x))$ for all $x \in [0, 1)$. The resulting IRA optimization methods are obtained by using this condition in (3.8).

It must be noted that $h(0) > 0$, because the IRA codes are systematic. If the IRA codes were not systematic, the iterative decoding would never start as $h(0)$ would be 0, unless $a = 1$ in which case the decoder starts but the rate range is limited (see Section 2.1). This is why we have considered systematic IRA codes in all our analysis [15].

While for the inner decoder (accumulator) we are forced to resort to Monte Carlo simulation, it is interesting to notice that, due to the simplicity of the repetition code, for both Methods 3 and 4 the EXIT function of the outer decoder ($I_E = h(I_A)$) can be obtained in closed form. Note that $h(I_A)$ can be written as

$$h(I_A) = \sum_{i=2}^{d} \lambda_i h_i(I_A) \tag{3.32}$$

where $I_E = h_i(I_A)$ is the EXIT function of the repetition "sub-decoder" with repetition degree i.

For Method 3, by discretizing the channel observation distribution as in

(3.18), we have

$$h(I_A) = 1 - \sum_{i=2}^{d} \sum_{j=1}^{D} \lambda_i p_j \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-z^2} \log_2 \left(1 + e^{-2\sqrt{(i-1)J^{-1}(I_A)}z - (i-1)J^{-1}(I_A) - v_j}\right) dz \tag{3.33}$$

For Method 4 we have

$$h(I_A) = 1 - (1 - \mathcal{I}(F_u)) \sum_{i=2}^{d} \lambda_i (1 - I_A)^{i-1} \tag{3.34}$$

3.5 Properties of the Approximated DE

In this section we show some properties of the approximated DE derived in Section 3.1.

3.5.1 Stability condition.

Consider the DE approximation of Method 1. As said in Section 3.4.1, $(x, \widetilde{x}) = (1, 1)$ is a fixed point of the system (3.17–3.19). We have the following result:

Theorem 3.7 *The fixed-point at* $(1, 1)$ *of the system (3.17 – 3.19) is stable if and only if the fixed-point* $(\Delta_\infty, \Delta_\infty)$ *of the exact DE (2.12 – 2.15) is stable.*

Proof: See Appendix 3.C.

For other DE approximations, stability does not generally imply stability of the corresponding exact DE. Consider the DE approximation of Method 2. $(1, 1)$ is a fixed-point of the system (3.27). We have the following result:

Proposition 3.8 *The local stability condition of the approximated DE with Method 2 is less stringent than that of the exact DE.*

Proof: See Appendix 3.D.

It is interesting to notice that the looseness of the stability condition was already mentioned in [43] when designing LDPC codes over the BIAWGNC with the erasure-channel approximation, as in our Method 2.

If an approximated DE has a less stringent stability condition, then the exact stability condition must be added to the ensemble optimization as an

explicit additional constraint. It should be noticed that the DE approximations used in [15, 22, 43] require the additional stability constraint. For example, the codes presented in [15] for the BIAWGNC and for which $\lambda_2 > 0$ are not stable. Therefore, the BER after an arbitrary large number of decoder iterations is not vanishing.

3.5.2 Fixed-Points, Coding Rate and Channel Capacity.

An interesting property of optimization Methods 2 and 4 is that the optimized ensemble for a given channel with channel observation distribution F_u and capacity $C = \mathcal{I}(F_u)$ has coding rate not larger than C. In fact, as a corollary of a general result of [42] (see Appendix 3.F), we have that

Theorem 3.9 *The DE approximations of Methods 2 and 4 have a unique fixed point* $(1, 1)$ *only if the IRA ensemble coding rate* R *satisfies* $R < C = \mathcal{I}(F_u)$.

Proof: See Appendix 3.F.

We show in Section 3.6.1 through some examples that this property does not hold in general for other code ensemble optimization methods, for which the ensemble rate R might result to be larger than the (nominal) capacity $\mathcal{I}(F_u)$. This means that the threshold $\nu^\star$, evaluated by exact DE, is worse than the channel parameter ν used for the ensemble design.

3.6 Numerical Results

3.6.1 Design Example for Rate $1/2$ Codes

In this subsection we present the result of optimization for codes of rate $1/2$ and give examples for the BSC with cross-over probability p and the BIAWGNC with symbol SNR

$$\mathsf{SNR} \triangleq \frac{E_s}{N_0} = \frac{1}{2\sigma^2}$$

and energy per bit divided by noise power spectral density N_0

$$\frac{E_b}{N_0} \triangleq \frac{1}{R}\frac{E_s}{N_0} = \frac{1}{R}\frac{1}{2\sigma^2}$$

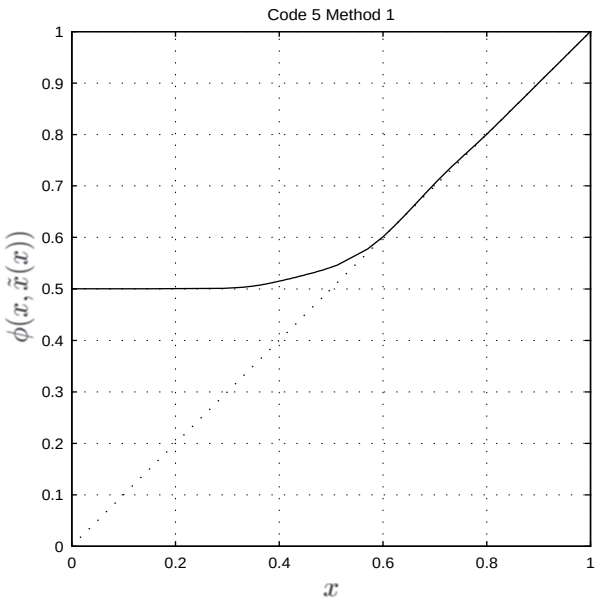


Figure 3.5: Fixed-point equation for BIAWGNC with Method 1, IRA code rate 1/2

In Fig. 3.5, the curve (solid line) shows $\phi(x, \widetilde{x}(x))$ as a function of $x \in [0, 1]$ for method 1, for the BIAWGNC and code rate 1/2. The solutions of the fixed-point equation (3.6) correspond to the intersection of this curve with the main diagonal (dotted line). Fig.3.6 shows the inverse of the EXIT function of the accumulator $g(x)$ (dashed line) and the EXIT functions of repetition "sub-decoders" $h_i(x)$ with repetition degrees $i = 2, 3, \cdots, 100$ (solid lines), on the BIAWGNC for method 3. The weighted combination (with weights λ_i) of the EXIT functions $h_i(x)$ yields the EXIT function of the repetition decoder $h(x)$. The fixed-point equations and EXIT functions are very similar for the four methods.

Tables 3.1 and 3.2 give the degree sequences, the grouping factors and the information bitnode average degrees for the four methods, for codes of rate 1/2 over the BIAWGNC and the BSC, respectively. We compute the true iterative decoding thresholds (by using the exact DE) for all the ensembles (denoted by SNR(DE) and p(DE) in the tables) and report also the gap of these thresholds with respect to the Shannon limit (denoted by SNR_{gap}(DE) and p_{gap}(DE) in the tables). Then, we compare it to the threshold of the approximated DE (denoted by SNR_{gap}(approx.) and p_{gap}(approx.) in the

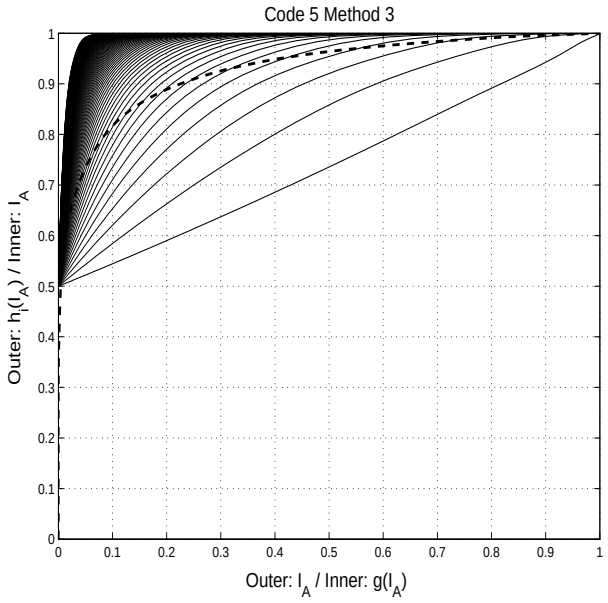


Figure 3.6: EXIT functions for BIAWGNC with Method 3, IRA code rate 1/2

tables). We observe that the rates of the codes designed using methods 2 or 4 are below capacity, which is consistent with Theorem 3.9. On the contrary, the codes designed using methods 1 or 3 have rate possibly larger than the capacity corresponding to the channel parameter used for the design. It can easily be checked that all the designed codes are stable.

3.6.2 Thresholds of IRA Ensembles

In this section we present results for codes designed according to the four methods, for rates from 0.1 to 0.9, and we compare the methods on the basis of the true thresholds obtained by DE. We present the code rate, the grouping factor, the average repetition factor and the gap to Shannon limit, for both BSC and BIAWGNC.

Tables 3.3 and 3.4 show the performance of IRA codes on the BIAWGNC. Tables 3.5 and 3.6 show the performance of IRA codes on the BSC. For all rates, and for both channels, IRA codes designed assuming Gaussian a priori (methods 1 and 3) perform much better than those designed assuming BEC a priori (methods 2 and 4). Nevertheless, method 4 yields better codes than

	Method 1		Method 2		Method 3		Method 4	
	i	λ_i	i	λ_i	i	λ_i	i	λ_i
	2	0.04227	2	0.05554	2	0.05266	2	0.05554
	3	0.16242	3	0.16330	3	0.11786	3	0.14480
	7	0.06529	8	0.06133	5	0.05906	7	0.18991
	8	0.06489	9	0.19357	6	0.06517	8	0.00996
	9	0.06207	25	0.14460	8	0.03615	19	0.03721
	10	0.01273	26	0.08842	9	0.11288	20	0.25894
	11	0.13072	100	0.29323	13	0.06068	100	0.30366
	14	0.04027			14	0.04650		
	25	0.00013			22	0.08606		
	26	0.05410			23	0.01610		
	36	0.13031			34	0.11019		
	37	0.13071			35	0.11919		
	100	0.10402			100	0.11751		
Rate	0.50183		0.49697		0.50154		0.49465	
a	8		8		8		8	
d	7.94153		8.09755		7.95087		8.17305	
SNR(DE)	-2.739		-2.457		-2.727		-2.588	
SNR_{gap}(DE)	0.059		0.406		0.075		0.306	
SNR_{gap}(approx.)	-0.025		0.040		-0.021		0.071	

Table 3.1: IRA codes of rate 1/2, designed with methods 1, 2, 3 and 4, for the BIAWGNC, with threshold evaluated with exact DE

method 2, especially at low rates. This is due to the fact that, in method 2, the communication channel is replaced with a BEC with the same capacity, while this is not the case in method 4. This difference in performance decreases as the rate increases.

Fig. 3.7 compares the performance of IRA ensembles with the best known LDPC ensembles [9] on the BIAWGNC. As expected, the performance of IRA ensembles is inferior to that of LDPC ensembles. However, in view of the simplicity of their encoding and decoding, IRA codes, optimized using methods 1 or 3, emerge as a very attractive design alternative.

Fig. 3.8 compares the performance of IRA ensembles obtained via the proposed methods for the BSC. The best codes are those designed with method 3.

	Method 1		Method 2		Method 3		Method 4	
	i	λ_i	i	λ_i	i	λ_i	i	λ_i
	2	0.03545	2	0.04732	2	0.03115	2	0.04657
	3	0.14375	3	0.17984	3	0.14991	3	0.14932
	6	0.03057	9	0.19715	6	0.04630	7	0.07693
	7	0.10963	10	0.06259	7	0.06217	8	0.16249
	9	0.10654	26	0.16429	8	0.08666	20	0.07001
	10	0.02388	27	0.05676	10	0.12644	21	0.20550
	11	0.04856	100	0.29205	17	0.03430	100	0.28919
	12	0.00461			18	0.01506		
	21	0.03035			26	0.00228		
	28	0.22576			27	0.02258		
	29	0.09453			28	0.21774		
	100	0.14635			29	0.08021		
					100	0.12521		
Rate	0.48908		0.49620		0.49226		0.49091	
a	8		8		8		8	
d	8.35724		8.12253		8.25157		8.29627	
p(DE)	0.1091		0.0938		0.1091		0.1009	
p_{gap}(DE)	0.0046		0.0175		0.0035		0.0122	
p_{gap}(approx.)	0.0037		0.0013		0.0026		0.0018	

Table 3.2: IRA codes of rate 1/2, designed with methods 1, 2, 3 and 4, for the BSC, with threshold evaluated with exact DE

3.7 Conclusion

This chapter has tackled the optimization of IRA codes in the limit of large code block length. For the sake of tractable analysis, we have proposed four methods to approximate densities involved in DE with a one-dimensional parameter. These approximations are motivated by recent results in the field of code design (EXIT functions, reciprocal channel approximation, and the non-strict convexity of mutual information) and have led to four optimization methods that can all be solved as a linear program.

We have shown formally that one of the proposed methods (Gaussian approximation, with reciprocal channel approximation) yields a one-dimensional DE approximation with the same stability condition as the one under exact DE, whereas the exact stability condition must be added to the ensemble optimization as an explicit additional constraint for another method (BEC a priori, with reciprocal channel approximation). We have also derived results related to the rates of the codes: in general the Gaussian a priori methods

Method 1				Method 3			
Rate	a	$\bar{d}$	SNR_{gap}	Rate	a	$\bar{d}$	SNR_{gap}
0.10109	2	17.78	0.151	0.10133	2	17.74	0.163
0.20191	3	11.86	0.096	0.20199	3	11.85	0.126
0.30153	4	9.27	0.081	0.30175	4	9.26	0.111
0.40196	6	8.93	0.057	0.40201	6	8.93	0.067
0.50184	8	7.94	0.059	0.50154	8	7.95	0.075
0.60188	11	7.28	0.065	0.60147	11	7.29	0.065
0.70154	16	6.81	0.067	0.70093	16	6.83	0.068
0.79904	29	7.29	0.066	0.79912	29	7.29	0.062
0.89677	61	7.02	0.088	0.89712	61	7.00	0.083

Table 3.3: IRA codes designed with methods 1 and 3 for the BIAWGNC, with threshold evaluated with DE

Method 2				Method 4			
Rate	a	$\bar{d}$	SNR_{gap}	Rate	a	$\bar{d}$	SNR_{gap}
0.09407	2	19.26	0.906	0.09752	2	18.51	0.316
0.19842	3	12.12	0.573	0.19725	3	12.21	0.293
0.29767	4	9.44	0.529	0.29671	4	9.48	0.336
0.39703	6	9.11	0.466	0.39445	6	9.21	0.343
0.49697	8	8.10	0.406	0.49465	8	8.17	0.306
0.59689	11	7.43	0.362	0.59577	11	7.46	0.338
0.69580	16	7.00	0.323	0.69584	16	6.99	0.296
0.79737	26	6.61	0.272	0.79678	26	6.63	0.271
0.89827	56	6.34	0.212	0.89826	56	6.34	0.214

Table 3.4: IRA codes designed with methods 2 and 4 for the BIAWGNC, with threshold evaluated with DE

are optimistic, in the sense that there is no guarantee that the optimized rate is below capacity. On the contrary, the BEC a priori methods always have rates below capacity.

Our numerical results show that, for the BIAWGNC and BSC, the Gaussian a priori approximation is more attractive since the codes designed under this assumption have the smallest gap to Shannon limit. Depending on the desired rate, the EXIT function of the inner decoder has to be computed either with Monte-Carlo simulation (method 3) or with the reciprocal channel approximation (method 1). At least for the BIAWGNC, there is some evidence that the best LDPC codes [9] designed with DE slightly outperform our designed codes. In view of this and the very simple encoding structure

Method 1				Method 3			
Rate	a	$\overline{d}$	p_{gap}	Rate	a	$\overline{d}$	p_{gap}
0.10042	2	17.92	0.0032	0.10137	2	17.73	0.0036
0.19910	3	12.07	0.0037	0.20086	3	11.94	0.0041
0.29573	4	9.53	0.0044	0.29897	4	9.38	0.0031
0.39298	6	9.27	0.0044	0.39621	6	9.14	0.0032
0.48908	8	8.36	0.0046	0.49226	8	8.25	0.0035
0.58590	12	8.48	0.0044	0.58815	11	7.70	0.0040
0.68271	17	7.90	0.0044	0.68409	16	7.39	0.0039
0.78155	28	7.83	0.0038	0.78235	28	7.79	0.0035
0.88437	59	7.71	0.0026	0.88457	63	8.22	0.0025

Table 3.5: IRA codes designed with methods 1 and 3 for the BSC, with threshold evaluated with DE

Method 2				Method 4			
Rate	a	$\overline{d}$	p_{gap}	Rate	a	$\overline{d}$	p_{gap}
0.09406	2	19.26	0.0194	0.09952	2	18.10	0.0121
0.19833	3	12.13	0.0175	0.19842	3	12.12	0.0101
0.29743	4	9.45	0.0190	0.28836	4	9.87	0.0114
0.39650	6	9.13	0.0187	0.38865	6	9.44	0.0149
0.49620	8	8.12	0.0175	0.49091	8	8.30	0.0122
0.59580	11	7.46	0.0155	0.59349	11	7.53	0.0124
0.69559	16	7.00	0.0126	0.69107	16	7.15	0.0116
0.79583	26	6.67	0.0091	0.79283	26	6.79	0.0090
0.89692	56	6.44	0.0049	0.89337	57	6.80	0.0051

Table 3.6: IRA codes designed with methods 2 and 4 for the BSC, with threshold evaluated with DE

of IRA codes, they emerge as attractive design choices.

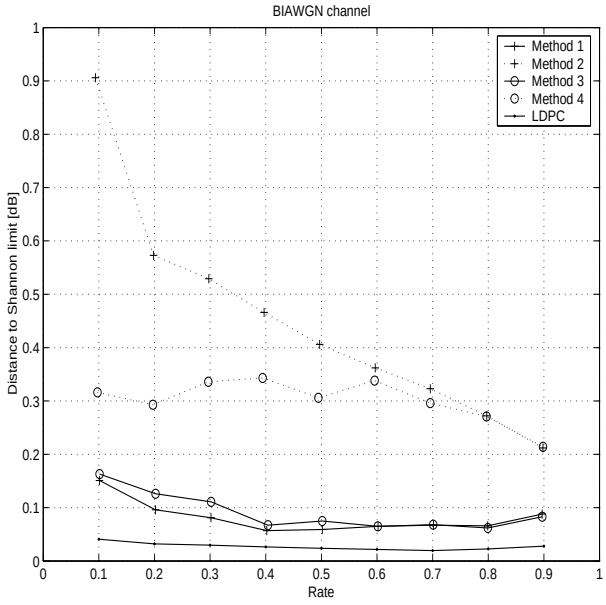


Figure 3.7: Gap to Shannon limit (obtained by DE) vs. rate for BIAWGNC

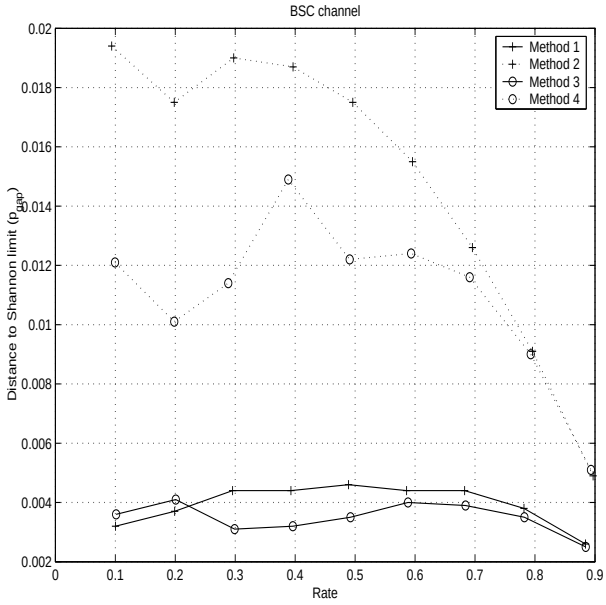


Figure 3.8: Gap to Shannon limit (obtained by DE) vs. rate for BSC

APPENDIX

3.A Proof of Proposition 3.1

First, notice that

$$f_{Y|X=1}(y) = \sum_i^M q_i f_{Y|X=1}^{(i)}(y) = \sum_i^M q_i e^{-y} f_{Y|X=0}^{(i)}(y) = e^{-y} f_{Y|X=0}(y) \quad (3.35)$$

Hence,

$$\begin{aligned}
I(X;Y) &= p\int f_{Y|X=0}(y)\log_2 \frac{f_{Y|X=0}(y)}{pf_{Y|X=0}(y)+(1-p)f_{Y|X=1}(y)}\,dy \\
&+ (1-p)\int f_{Y|X=1}(y)\log_2 \frac{f_{Y|X=1}(y)}{pf_{Y|X=0}(y)+(1-p)f_{Y|X=1}(y)}\,dy \\
&= p\int f_{Y|X=0}(y)\log_2 \frac{1}{p+(1-p)e^{-y}}\,dy \\
&+ (1-p)\int f_{Y|X=1}(y)\log_2 \frac{1}{pe^{y}+(1-p)}\,dy \\
&= p\int \sum_i^M q_i f_{Y|X=0}^{(i)}(y)\log_2 \frac{1}{p+(1-p)e^{-y}}\,dy \\
&+ (1-p)\int \sum_i^M q_i f_{Y|X=1}^{(i)}(y)\log_2 \frac{1}{pe^{y}+(1-p)}\,dy \\
&= \sum_i^M q_i \left(p\int f_{Y|X=0}^{(i)}(y)\log_2 \frac{1}{p+(1-p)e^{-y}}\,dy \right. \\
&\qquad \left. +(1-p)\int f_{Y|X=1}^{(i)}(y)\log_2 \frac{1}{pe^{y}+(1-p)}\,dy \right) \\
&= \sum_i^M q_i \left(p\int f_{Y|X=0}^{(i)}(y)\log_2 \frac{f_{Y|X=0}^{(i)}(y)}{pf_{Y|X=0}^{(i)}(y)+(1-p)f_{Y|X=1}^{(i)}(y)}\,dy \right. \\
&\qquad \left. +(1-p)\int f_{Y|X=1}^{(i)}(y)\log_2 \frac{f_{Y|X=1}^{(i)}(y)}{pf_{Y|X=0}^{(i)}(y)+(1-p)f_{Y|X=1}^{(i)}(y)}\,dy \right) \\
&= I(X;Y|S) \qquad (3.36)
\end{aligned}$$

Therefore, the mutual information is not strictly convex on channels with transition probability $f_{Y|X}^{(i)}$ as in (3.35) . □

Proof of Corollary 3.2 The assertion of Corollary 3.2 follows from Proposition 3.1 since for a collection of binary-input symmetric-output channels with symmetric transition probability

$$\begin{aligned}\forall i, \forall y : p_{Y|X,S}(y|X=1,S=i) &= p_{Y|X,S}(-y|X=0,S=i) \\ &= e^{-y} p_{Y|X,S}(y|X=0,S=i)\end{aligned}$$

□

3.B EXIT Function with Monte Carlo

We want to compute $I(X;L)$, the mutual information between the input binary symbol X and the output log likelihood message L. Note that

$$I(X;L) = 1 - H(X|L) \tag{3.37}$$

where $H(X|L)$ is the conditional entropy of X on L defined as [51]

$$H(X|L) \stackrel{\Delta}{=} -\sum_x \sum_l p(x,l) \log p(x|l) \tag{3.38}$$

The conditional probability of X on $L = l$ is given by

$$p(X=0|L=l) = \frac{e^l}{e^l+1} \tag{3.39}$$

$$p(X=1|L=l) = \frac{1}{e^l+1} \tag{3.40}$$

Then (3.37) becomes

$$\begin{aligned} I(X;L) &= 1 + \sum_l \sum_x p(l)p(x|l) \log p(x|l) \\ &= 1 + \sum_l p(l) \left(\frac{e^l}{1+e^l} \log \frac{e^l}{1+e^l} + \frac{1}{1+e^l} \log \frac{1}{1+e^l} \right) \\ &= 1 + E_L \left[\frac{l e^l}{1+e^l} - \log\left(1+e^l\right) \right] \end{aligned} \tag{3.41}$$

The mean in (3.41) can be obtained by Monte Carlo simulation.

3.C Proof of Theorem 3.7

The local stability condition for the system ((3.17) and (3.19)) is given by the eigenvalues of the Jacobian matrix for the functions $(\phi, \tilde{\phi})$ in the fixed-point $(x, \tilde{x}) = (1, 1)$. The partial derivatives of ϕ and $\tilde{\phi}$ are

$$\frac{\partial \phi}{\partial x}(1,1) = \sum_{i=2}^{d}\sum_{j=1}^{D} \lambda_i p_j (i-1)(a-1) \lim_{\mu \to +\infty} \frac{J'_{v_j}((i-1)\mu)}{J'(\mu)} \quad (3.42)$$

$$\frac{\partial \phi}{\partial \tilde{x}}(1,1) = \sum_{i=2}^{d}\sum_{j=1}^{D} \lambda_i p_j (i-1)2 \lim_{\mu \to +\infty} \frac{J'_{v_j}((i-1)\mu)}{J'(\mu)} \quad (3.43)$$

$$\frac{\partial \tilde{\phi}}{\partial x}(1,1) = \sum_{j=1}^{D} p_j a \lim_{\mu \to +\infty} \frac{J'_{v_j}(\mu)}{J'(\mu)} \quad (3.44)$$

$$\frac{\partial \tilde{\phi}}{\partial \tilde{x}}(1,1) = \sum_{j=1}^{D} p_j \lim_{\mu \to +\infty} \frac{J'_{v_j}(\mu)}{J'(\mu)} \quad (3.45)$$

where

$$J_{v_j}(\mu) \triangleq 1 - \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-z^2} \log_2 \left(1 + e^{-2\sqrt{\mu}z - \mu - v_j}\right) dz. \quad (3.46)$$

Note that $J_0(\mu) = J(\mu)$. Since both limits tend to 0, we derive an asymptotic expansion for $J'_{v_j}(\mu)$ and $J'(\mu)$.

The derivative of J_{v_j} is given by

$$J'_{v_j}(\mu) = \frac{\log_2(e)}{\sqrt{\mu}} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} (z + \sqrt{\mu}) e^{-v_j} \frac{e^{-(z+\sqrt{\mu})^2}}{1 + e^{-2\sqrt{\mu}z - \mu - v_j}} dz$$

Since F_u is symmetric, the sum over j can be rewritten as:

$$\sum_{j=1}^{D} p_j J'_{v_j}(\mu) = p'_0 J'_0(\mu) + \sum_{j=1}^{D'} p'_j \left(J'_{v'_j}(\mu) + e^{-v'_j} J'_{-v'_j}(\mu)\right)$$

Let us define

$$f_0(\mu) = \frac{1}{\log_2(e)} J'_0(\mu) \quad (3.47)$$

and

$$\begin{aligned} f_{v'_j}(\mu) &= \frac{1}{\log_2(e)}\left(J'_{v'_j}(\mu)+e^{-v'_j}J'_{-v'_j}(\mu)\right) \\ &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}(1+\frac{z}{\sqrt{\mu}})e^{-(z+\sqrt{\mu})^2} \\ &\quad \left(\frac{e^{-v'_j}}{1+e^{-2\sqrt{\mu}z-\mu-v'_j}}+\frac{1}{1+e^{-2\sqrt{\mu}z-\mu+v'_j}}\right)dz \end{aligned} \tag{3.48}$$

Following [52], (3.48) can be rewritten as

$$\begin{aligned} f_{v'_j}(\mu) &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{\mu}}\left(z+\frac{\sqrt{\mu}}{2}\right)e^{-\left(z+\frac{\sqrt{\mu}}{2}\right)^2}\left(\frac{e^{-v'_j}}{1+e^{-2\sqrt{\mu}z-v'_j}}+\frac{1}{1+e^{-2\sqrt{\mu}z+v'_j}}\right)dz \\ &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}\frac{z}{\sqrt{\mu}}e^{-z^2-\frac{\mu}{4}-\frac{v'_j}{2}}\left(\frac{1}{e^{\sqrt{\mu}z+\frac{v'_j}{2}}+e^{-\sqrt{\mu}z-\frac{v'_j}{2}}}+\frac{1}{e^{\sqrt{\mu}z-\frac{v'_j}{2}}+e^{-\sqrt{\mu}z+\frac{v'_j}{2}}}\right)dz \\ &\quad +\frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}\frac{1}{2}e^{-z^2-\frac{\mu}{4}-\frac{v'_j}{2}}\left(\frac{1}{e^{\sqrt{\mu}z+\frac{v'_j}{2}}+e^{-\sqrt{\mu}z-\frac{v'_j}{2}}}+\frac{1}{e^{\sqrt{\mu}z-\frac{v'_j}{2}}+e^{-\sqrt{\mu}z+\frac{v'_j}{2}}}\right)dz \\ &= \frac{e^{-\frac{\mu}{4}-\frac{v'_j}{2}}}{4\sqrt{\pi}}\int_{-\infty}^{+\infty}e^{-z^2}\left(\frac{1}{\cosh(\sqrt{\mu}z+\frac{v'_j}{2})}+\frac{1}{\cosh(\sqrt{\mu}z-\frac{v'_j}{2})}\right)dz \\ &= \frac{e^{-\frac{\mu}{4}-\frac{v'_j}{2}}}{4\sqrt{\pi\mu}}\int_{-\infty}^{+\infty}\frac{e^{-\frac{(z-\frac{v'_j}{2})^2}{\mu}}+e^{-\frac{(z+\frac{v'_j}{2})^2}{\mu}}}{\cosh(z)}dz \end{aligned} \tag{3.49}$$

The first equality in (3.49) is obtained by the change of variable $z'=z+\sqrt{\mu}/2$. The third equality is due to the fact that the first and second integrands in the second line of (3.49) are odd and even functions of z, respectively. Then we use the changes of variable $z'=\sqrt{\mu}z+\frac{v'_j}{2}$ and $z'=\sqrt{\mu}z-\frac{v'_j}{2}$.

Lebesgue's dominated convergence theorem completes the proof. Indeed, the sequence of measurable functions verifies:

$$\forall z\in\mathbb{R}, \frac{e^{-\frac{z^2}{\mu}}}{\cosh(z)} \xrightarrow[\mu\to+\infty]{} \frac{1}{\cosh(z)}$$

and these functions are bounded by an integrable function independent of μ:

$$\forall \mu>0, \forall z\in\mathbb{R}, \left|\frac{e^{-\frac{z^2}{\mu}}}{\cosh(z)}\right| \leq \frac{1}{\cosh(z)} \in L^1(\mathbb{R}).$$

Thus Lebesgue's dominated convergence theorem [53] applies and

$$\int_{-\infty}^{+\infty} \frac{e^{-\frac{z^2}{\mu}}}{\cosh(z)} dz \xrightarrow[\mu\to+\infty]{} \int_{-\infty}^{+\infty} \frac{1}{\cosh(z)} dz = [2\arctan(e^z)]_{-\infty}^{+\infty} = \pi$$

Therefore for large μ

$$\begin{aligned} f_{v'_j}(\mu) &\sim \frac{\sqrt{\pi}}{2} \frac{e^{-\frac{\mu}{4}} e^{-\frac{v'_j}{2}}}{\sqrt{\mu}} \\ f_0(\mu) &\sim \frac{\sqrt{\pi}}{4} \frac{e^{-\frac{\mu}{4}}}{\sqrt{\mu}} \end{aligned}$$

And thus, for $n \geq 1$

$$\begin{aligned} \lim_{\mu\to+\infty} \frac{f_{v'_j}(n\mu)}{f_0(\mu)} &= \begin{cases} 2e^{-\frac{v'_j}{2}} & \text{if} \quad n = 1 \\ 0 & \text{if} \quad n > 1 \end{cases} \\ \lim_{\mu\to+\infty} \frac{f_0(n\mu)}{f_0(\mu)} &= \begin{cases} 1 & \text{if} \quad n = 1 \\ 0 & \text{if} \quad n > 1 \end{cases} \end{aligned}$$

The partial derivatives of ϕ and $\tilde{\phi}$ are

$$\begin{aligned} \frac{\partial\phi}{\partial x}(1,1) &= \lambda_2(a-1)(p'_0 + \sum_{j=1}^{D'} 2p'_j e^{-\frac{v'_j}{2}}) \\ &= \lambda_2(a-1)(p'_0 + \sum_{j=1}^{D'} (p'_j e^{-\frac{v'_j}{2}} + p'_j e^{-v'_j} e^{\frac{v'_j}{2}}) \\ &= \lambda_2(a-1) \sum_{j=1}^{D} p_j e^{-\frac{v_j}{2}} \\ &= \lambda_2(a-1)e^{-r} \end{aligned} \tag{3.50}$$

where r is defined in (2.26). Similarly,

$$\frac{\partial\phi}{\partial\tilde{x}}(1,1) = \lambda_2 2e^{-r} \tag{3.51}$$

$$\frac{\partial\tilde{\phi}}{\partial x}(1,1) = ae^{-r} \tag{3.52}$$

$$\frac{\partial\tilde{\phi}}{\partial\tilde{x}}(1,1) = e^{-r} \tag{3.53}$$

The Jacobian matrix is then

$$\mathbf{J} = \begin{bmatrix} (a-1)\lambda'(0) & 2\lambda'(0) \\ a & 1 \end{bmatrix} e^{-r}$$

In order to be stable the eigenvalues of $\mathbf{J}$ should be inside the unit circle. Therefore the stability condition reduces to:

$$\frac{1}{2}\left(1+\lambda'(0)(a-1)+\sqrt{1+2\lambda'(0)+6\lambda'(0)a+\lambda'(0)^2(a-1)^2}\right) < e^r \quad (3.54)$$

Notice from (2.27) and (3.54) that the stability conditions under DE and approximated DE are the same. □

3.D Proof of Proposition 3.8

The Jacobian matrix of the approximated DE (3.27) about the fixed-point $[x, \tilde{x}] = [1, 1]$, for a given input channel distribution F_u, is

$$\mathbf{J} = \begin{bmatrix} (a-1)\lambda'(0) & 2\lambda'(0) \\ a & 1 \end{bmatrix} (1 - \mathcal{I}(F_u)) = \mathbf{A}(1 - \mathcal{I}(F_u))$$

where $\mathbf{A}$ was already defined in (2.24). The stability of the exact DE is given by the eigenvalues of $\mathbf{A}e^{-r}$ (where r is defined in (2.26)) while it is given by those of $\mathbf{A}(1 - \mathcal{I}(F_u))$ for the approximated DE.

Assuming that F_u has a density function denoted f_u, define

$$\begin{aligned} f_1(f_u) &= \textstyle\int_{-\infty}^{+\infty} e^{-z/2} f_u(z) dz \\ f_2(f_u) &= \textstyle\int_{-\infty}^{+\infty} \log_2\left(1+e^{-z}\right) f_u(z) dz \end{aligned}$$

Because F_u is symmetric, we get

$$\begin{aligned} f_1(f_u) &= \int_0^{+\infty} e^{-z/2} f_u(z) dz + \int_{-\infty}^{0} e^{z/2} f_u(-z) dz \\ &= \int_0^{+\infty} 2e^{-z/2} f_u(z) dz \qquad (3.55) \\ f_2(f_u) &= \int_0^{+\infty} \log_2\left(1+e^{-z}\right) f_u(z) dz + \int_{-\infty}^{0} \log_2\left(1+e^{-z}\right) e^{z} f_u(-z) dz \\ &= \int_0^{+\infty} \left(\left(1+e^{-z}\right)\log_2\left(1+e^{-z}\right) + \frac{z}{\log 2} e^{-z} \right) f_u(z) dz \qquad (3.56) \end{aligned}$$

We will show in the following that

$$\forall z \geq 0: \ (1+e^{-z})\log(1+e^{-z}) + ze^{-z} \leq 2(\log 2)e^{-z/2}$$

Letting $x = e^{-z}$, the above statement becomes equivalent to

$$\forall x \in [0,1]: \ f(x) \leq 0$$

where

$$f(x) \stackrel{\Delta}{=} (1+x)\log(1+x) - x\log x - 2\log 2\sqrt{x} \tag{3.57}$$

Noting that

$$\lim_{x\to 0} f(x) = 0 \qquad \text{and} \qquad f(1) = 0$$

then a sufficient condition for $f(x)$ to be negative over the interval $(0,1)$ is that $f(x)$ has a single extremum, which is a negative minimum (see Fig. 3.9). The derivative of $f(x)$ is

$$f'(x) = \log\frac{x+1}{x} - \frac{\log 2}{\sqrt{x}}$$

with

$$\lim_{x\to 0} f'(x) = -\infty \qquad \text{and} \qquad f'(1) = 0$$

The solution of $f'(x) = 0$ is the same as that of $g_1(x) = g_2(x)$ for $0 \leq x \leq 1$, where $g_1(x) = \frac{x+1}{x}$ and $g_2 = 2^{1/\sqrt{x}}$, with

$$\lim_{x\to 0} g_1(x) = \lim_{x\to 0} g_2(x) = +\infty \qquad \text{and} \qquad g_1(1) = g_2(1) = 2$$

Since

- g_1 and g_2 are convex functions[2] on the interval $[0,1]$,
- $g_2(x)$ grows faster to $+\infty$ than $g_1(x)$, as x tends to 0,
- $g_1'(1) = -1$ and $g_2'(1) = -\log 2$, i.e., the slope of g_1 is larger (in magnitude) than the slope of g_2, at $x = 1$ (see Fig. 3.10),

then there exists a unique $x_0 \in (0,1)$ for which $g_1(x_0) = g_2(x_0)$, i.e., $f'(x_0) = 0$. Since $f(x)$ is decreasing in the vicinity of $x = 0$, then the extremum $f(x_0)$ is a negative minimum, and $\forall x \in [0,1], f(x) \leq 0$. Hence

$$\forall F_u \in \mathcal{F}_{sym} \quad f_2(f_u) \leq f_1(f_u)$$

and the conclusion follows. □

[2] $g_1''(x) \geq 0$ and $g_2''(x) \geq 0$ for $0 \leq x \leq 1$.

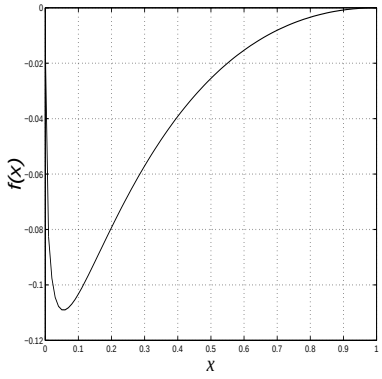


Figure 3.9: Function $f(x)$

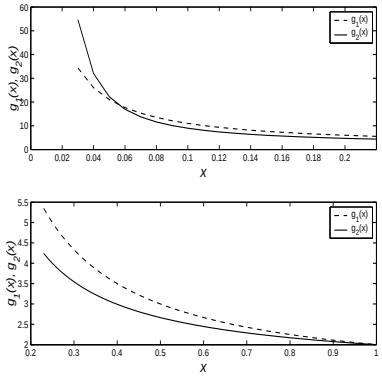


Figure 3.10: Functions $g_1(x)$ and $g_2(x)$

3.E Proof of Lemma 3.3

Because of property 3 of binary-input symmetric-output channels shown in Section 3.3.1, we can assume without loss of generality, that F is symmetric (in the sense of (2.10)). Then, the channel capacity is given by (3.11), namely

$$C(F) = \mathcal{I}(F) = 1 - \int_{-\infty}^{\infty} \log_2(1 + e^{-z}) dF(z)$$

From Proposition 3.8, we find that

$$\mathcal{I}(F) \geq 1 - \int_{-\infty}^{\infty} e^{-z/2} dF(z) = 1 - e^{-r} \tag{3.58}$$

r being the Chernoff bound exponent. Therefore, to show (3.13), it is sufficient to show that

$$1 - e^r \geq \frac{1}{2}\left(r - r^2\right)\log_e 2 \tag{3.59}$$

First note that (3.59) holds for all $r \geq 1$. Hence, we only need to show the inequality for $r \in [0, 1]$. Recall the Taylor series expansion of a real function $f(x)$ around x_0

$$f(x) = \sum_{i=0}^{p} \frac{(x - x_0)^i}{i!} f^{(i)}(x_0) + \frac{(x - x_0)^{p+1}}{(p+1)!} f^{(p+1)}(c)$$

for some $c \in (x_0, x)$ and any integer $p > 0$, where $f^{(i)}(x_0) = \left.\frac{d^i f(x)}{dx^i}\right|_{x=x_0}$ is the i^{th} derivative of $f(x)$ with respect to x. The Taylor series expansion of $1 - e^{-r}$ around 0 yields

$$e^{-r} = 1 - r + \frac{r^2}{2} - \frac{e^{-c} r^3}{3!}$$

for some $c \in [0, r]$, then

$$1 - e^{-r} \geq r - \frac{r^2}{2}$$

Since $\frac{(\log_2 e - 1)}{2} r^2 + (1 - \log_2 e/2) r \geq 0$ for $r \in [0, 1]$, then

$$1 - e^{-r} \geq \frac{1}{2}(r - r^2)\log_2 e$$

thus concluding the proof. □

3.F Proof of Theorem 3.9

Theorem 3.9 follows as a corollary of a result of [42] that we state here for the sake of completeness as Lemma 3.10 below. In order to introduce this result, we consider the model of Fig. 3.11, where $\mathbf{b}$, $\mathbf{x}_1$ and $\mathbf{x}$ are binary sequences and where Channel 1 is the communication channel with output $\mathbf{y}$ and Channel 2 is a BEC channel with output $\mathbf{z}$. Let the decoder be a maximum a posteriori (MAP) symbol-by-symbol decoder, producing for all $i = 1, \ldots, n$, output messages of the form

$$m_{o,i} = \log \frac{P(x_{1,i} = 0 | \mathbf{y}, \mathbf{z}_{[i]})}{P(x_{1,i} = 1 | \mathbf{y}, \mathbf{z}_{[i]})} \tag{3.60}$$

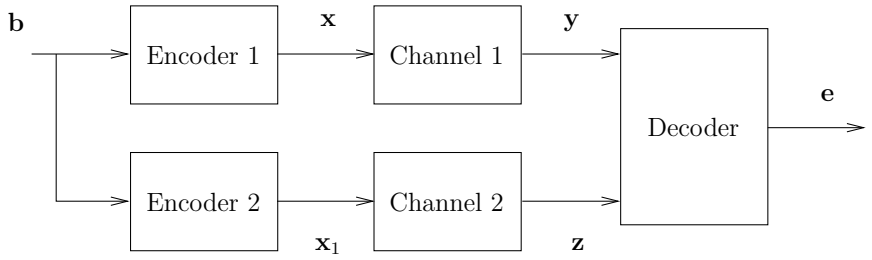


Figure 3.11: General decoding model

where $\mathbf{z}_{[i]} \stackrel{\Delta}{=} (z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n)$. Following [42], we generalize the definition of I_A and I_E given in Section (3.2) to the case of sequences as

$$\begin{aligned} I_A &= \frac{1}{n}\sum_{i=1}^{n} I(x_{1,i}; z_i) \\ I_E &= \frac{1}{n}\sum_{i=1}^{n} I(x_{1,i}; m_{o,i}) \\ &\stackrel{\text{a}}{=} \frac{1}{n}\sum_{i=1}^{n} I(x_{1,i}; \mathbf{y}, \mathbf{z}_{[i]}) \end{aligned} \tag{3.61}$$

where (a) follows from the fact that the decoder is MAP. Again, the decoder EXIT function is the set of points (I_A, I_E) for all $I_A \in [0, 1]$.

For the setup of Fig. 3.11 with the above assumptions, the following result applies:

Lemma 3.10 *[42] Let* $\mathbf{b}$ *be uniformly distributed and i.i.d.. If Encoder 2 is linear with generator matrix having no all-zero columns, then the area under the EXIT characteristic satisfies*

$$\mathcal{A} \stackrel{\Delta}{=} \int_0^1 I_E(z)dz = 1 - \frac{1}{n}H(\mathbf{x}_1|\mathbf{y}) \tag{3.62}$$

We start by proving Theorem 3.9 for the approximated DE of method 4. Consider the IRA encoder of Fig. 2.1 and the Turbo-like decoder of Fig. 3.3.

The inner MAP decoder receives channel observations $\mathbf{u}_p$ for the parity bits and input messages for the symbols of $\mathbf{x}_1$, and produces output messages

for the symbols of $\mathbf{x}_1$. The general decoding model of Fig. 3.11, applied to the inner decoder, yields the model of Fig. 3.12 (a).

The outer MAP decoder receives channel observations $\mathbf{u}_s$ for the information bits and input messages for the symbols of $\mathbf{x}_1$, and produces output messages for the symbols of $\mathbf{x}_1$. The general decoding model of Fig. 3.11, applied to the outer decoder, yields the model of Fig. 3.12 (b).

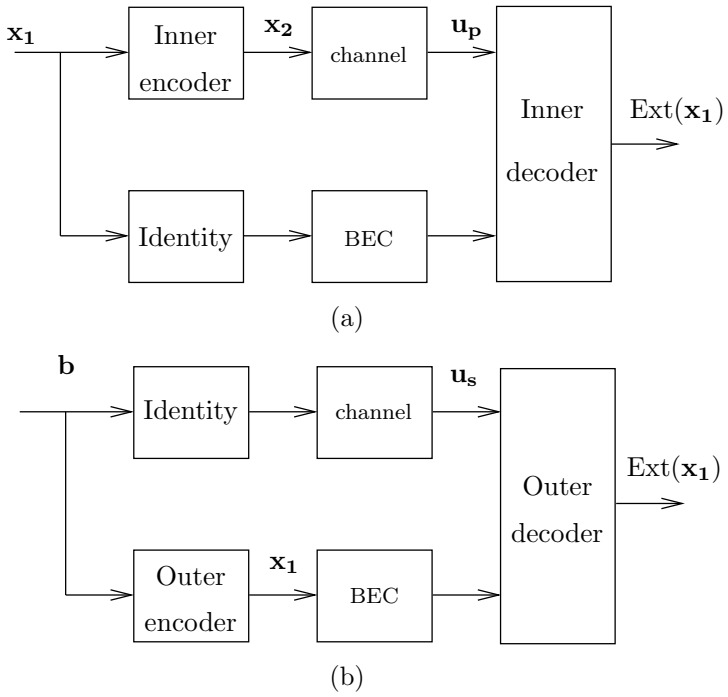


Figure 3.12: Model of inner (a) and outer (b) decoders for method 4

The upper channel is the communication channel with capacity $\mathcal{I}(F_u)$. Since we consider the approximation method 4, we let the lower channel to be a BEC in both Figs. 3.12 (a) and (b). Let k, n and m denote the number of information bits (length of $\mathbf{b}$ and of $\mathbf{u}_s$), the number of repeated information bits (length of $\mathbf{x}_1$) and the number of parity bits (length of $\mathbf{x}_2$ and of $\mathbf{u}_p$), respectively. The inner and outer coding rates are $R_{\text{in}} = n/m$ and $R_{\text{out}} = k/n$, and the overall IRA coding rate (2.3) is given by

$$R = \frac{k}{k+m} = \frac{R_{\text{in}} R_{\text{out}}}{1 + R_{\text{in}} R_{\text{out}}}$$

By applying Lemma 3.10 to the inner code model (Fig. 3.12 (a)), we obtain

$$\begin{aligned}\mathcal{A}_{\text{in}} &= 1-\frac{1}{n}H(\mathbf{x}_1|\mathbf{u}_p)\\ &= 1-\frac{1}{n}(H(\mathbf{x}_1)-I(\mathbf{x}_1;\mathbf{u}_p))\\ &\stackrel{\text{a}}{=} \frac{1}{n}I(\mathbf{x}_1;\mathbf{u}_p)\\ &\stackrel{\text{b}}{=} \frac{m}{n}I(x_{2,i};u_{p,i}) = \mathcal{I}(F_u)/R_{\text{in}} \end{aligned} \tag{3.63}$$

where (a) follows from the fact that, by the model assumption, $\mathbf{x}_1$ is an i.i.d. uniformly distributed binary sequence, and (b) follows from the fact that the accumulator (inner code) generates $\mathbf{x}_2$ with uniform probability (and uniform marginals) if driven by the i.i.d. uniform input sequence $\mathbf{x}_1$ (because the inner code rate a is larger than 1).

By applying Lemma 3.10 to the outer code model (Fig. 3.12 (b)), we obtain

$$\begin{aligned}\mathcal{A}_{\text{out}} &= 1-\frac{1}{n}H(\mathbf{x}_1|\mathbf{u}_s)\\ &= 1-\frac{1}{n}(H(\mathbf{x}_1)-I(\mathbf{x}_1;\mathbf{u}_s))\\ &\stackrel{\text{a}}{=} 1-\frac{k}{n}+\frac{1}{n}I(\mathbf{x}_1;\mathbf{u}_s)\\ &\stackrel{\text{b}}{=} 1-\frac{k}{n}+\frac{k}{n}I(b_i;u_{s,i}) = 1-R_{\text{out}}+R_{\text{out}}\mathcal{I}(F_u) \end{aligned} \tag{3.64}$$

where both (a) and (b) follow from the fact that the repetition code is an invertible mapping, so the entropy $H(\mathbf{x}_1)$ is equal to the entropy of the information sequence $\mathbf{b}$ (equal to k bits) and $I(\mathbf{x}_1;\mathbf{u}_s) = I(\mathbf{b};\mathbf{u}_s) = kI(b_i;u_{s,i}) = k\mathcal{I}(F_u)$.

As seen in Section 3.4.3, the approximated DE has no fixed-points other than $(1,1)$ if and only if $g(x) > h^{-1}(x)$ for all $x \in [0,1)$, where $g(x)$ and $h(x)$ denote the inner and outer decoder EXIT functions. This implies that

$$\mathcal{A}_{\text{in}} > 1-\mathcal{A}_{\text{out}}$$

since

$$\begin{aligned}\mathcal{A}_{\text{in}} &= \textstyle\int_0^1 g(x)dx\\ \mathcal{A}_{\text{out}} &= \textstyle\int_0^1 h(x)dx\end{aligned}$$

By using (3.63) and (3.64), we obtain

$$\begin{aligned} \mathcal{I}(F_u)/R_{\text{in}} &> R_{\text{out}} - R_{\text{out}}\mathcal{I}(F_u) \\ &\Downarrow \\ \mathcal{I}(F_u) &> \frac{R_{\text{in}}R_{\text{out}}}{1 + R_{\text{in}}R_{\text{out}}} = R \end{aligned} \tag{3.65}$$

For method 2, the above derivation still holds, since the communication channel in Fig. 3.11 is replaced by a BEC with erasure probability $\epsilon = 1 - \mathcal{I}(F_u)$. In fact, the inner and outer decoder EXIT functions can be rewritten as

$$h(x) = 1 - (1 - \mathcal{I}(F_u)) \sum_{i=2}^{d} \lambda_i \, (1-x)^{i-1} \tag{3.66}$$

$$g(x) = \frac{x^{a-1}\mathcal{I}(F_u)^2}{(1 - (1 - \mathcal{I}(F_u))x^a)^2} \tag{3.67}$$

and the areas under these functions are again

$$\begin{aligned} \mathcal{A}_{\text{out}} &= \int_0^1 h(x)dx = 1 - (1 - \mathcal{I}(F_u)) \sum_{i=2}^{d} \lambda_i/i = 1 - R_{\text{out}} + R_{\text{out}}\mathcal{I}(F_u) \\ \mathcal{A}_{\text{in}} &= \int_0^1 g(x)dx = \mathcal{I}(F_u)/a = \mathcal{I}(F_u)/R_{\text{in}} \end{aligned}$$

since $R_{\text{in}} = \sum_{i=2}^{d} \lambda_i/i$ and $R_{\text{out}} = a$. Therefore, the final result (3.65) holds also for method 2. □

Chapter 4

Finite Length Repeat Accumulate Codes

In this chapter, we study the performance of finite length regular and irregular repeat accumulate codes, whose Tanner graphs are constructed semi-randomly and satisfy one of the following criteria: (a) the graph is free of cycles up to $2S$, (b) cycles of length up to $2d_{ACE}$ have at least η external edges (S, d_{ACE} and η are design parameters). We compute the input output weight enumerators of IRA codes, and derive an upper bound on the bit error probability of random RA codes under maximum likelihood decoding.

4.1 Finite Length IRA Codes

The performance of the random-like ensemble of systematic IRA codes, in the limit of large block length, optimized in chapter 3 for binary-input symmetric-output channels, is found to be very close to the Shannon limit. The bipartite graphs associated to these codes are free of cycles of any finite length, in the limit of infinite block length. However, finite length bipartite graphs may contain short cycles, and the BP message-passing decoder is no longer optimal, since the local message independence assumption is no longer valid. Randomly constructed IRA codes of finite length may have poor performances

under BP decoding, in the following aspects:

- In the low SNR region, the BER waterfall is far from the code ensemble threshold, hence the gap from channel capacity is not as good as predicted by the infinite length DE analysis;
- In the medium to high SNR region, the BER flattens in an even larger gap from channel capacity for very low BER. This behavior is referred to as an "error floor";
- If $\lambda_2 \neq 0$, as is the case for optimal degree sequence distributions, then the word error rate (WER) is poor.

There exists a trade-off between the threshold SNR and the "error floor" BER of irregular versus regular codes [54]. For short code block lengths, regular codes usually exhibit better error floor BER than their irregular counterparts. On the other hand, for large block lengths, optimized irregular codes largely outperform their regular counterparts in approaching their DE threshold.

The issue of interleaver construction can be tackled in two ways, either adopting a turbo code approach or a LDPC-oriented approach. Different approaches have been adopted for the construction of turbo code interleavers such as increasing the code minimum distance [55, 56, 57] and enhancing the convergence of iterative decoding [58, 59]. These approaches are not suitable for IRA codes, because they do not eliminate short cycles in the bipartite graph, therefore, the performance of the codes under iterative decoding is not improved.

Different approaches have been adopted for the construction of finite-length LDPC codes. [60, 61, 62] propose methods to remove short cycles in order to maximize the length of the shortest cycle, and [63] proposes a method to remove short cycles that contribute to small stopping sets [27].

It is commonly believed (based on heuristic arguments) that the removal of short cycles improves the performance of short length bipartite graph codes under BP. It is also found that maximizing the length of the shortest cycle yields a large minimum distance [60, 64], thus improving the code performance in the high signal to noise ratio (SNR) region. We therefore use the progressive edge-growth algorithm (PEG) [60] to construct regular Repeat Accumulate (regular RA) and IRA codes, whose associated bipartite graph has a shortest cycle length, called *girth*, at least equal to $2S+2$, where S is a design parameter.

Under BP decoding of random LDPC codes, it is found that small stopping sets result in high error floors on the BEC. So, alternatively to girth conditioning, we use the method proposed in [63], to construct IRA codes whose Tanner graphs are free of small stopping sets.

It is instructive to determine the average random IRA performance on the BIAWGNC, under ML decoding. This allows to determine whether a poor performance of the random IRA ensemble is due to the BP decoder or the code ensemble itself. Moreover, it is interesting to compare the performance of the semi-random ensemble performance under BP against that of the random ensemble under ML decoding.

4.2 Construction of Finite Length IRA Codes

Consider the random IRA code ensemble described in chapter 2, whose bipartite graph is depicted in Fig. 2.2. Let there be n_{deg} different repetition degrees r_j. Then, the repetition code can alternatively be defined in terms of a degree distribution $\{0 < f_i \leq 1, i = 1, \ldots, n_{deg}\}$ and a degree set $\{2 \leq d_i \leq d, i = 1, \ldots, n_{deg}\}$, such that $f_i k$ is the number of information bitnodes of degree d_i and $\sum_{i=1}^{n_{deg}} f_i = 1$. The average information bitnode degree is then given by $\bar{d} = \sum_{i=1}^{n_{deg}} i f_i$.

Without loss of generality, assume that $r_1 \leq r_2 \ldots \leq r_k$, i.e., information bitnodes are indexed in the non-decreasing order of repetition degrees. Let $e_1, \ldots, e_N$ denote the edges connected to the information bitnodes $v_1, \ldots, v_k$ so that edges $e_1, \ldots, e_{r_1}$ are connected to information bitnode v_1, edges $e_{r_1+1}, \ldots, e_{r_1+r_2}$ are connected to information bitnode v_2, and so on. Define the mapping V as

$$\begin{aligned} V : \{1, \ldots, N\} &\rightarrow \{1, \ldots, k\} \\ j &\rightarrow i \text{ such that } e_j \text{ is connected to } v_i \end{aligned}$$

We construct the a-to-1 mapping C

$$\begin{aligned} C : \{1, \ldots, N\} &\rightarrow \{1, \ldots, m\} \\ j &\rightarrow i \text{ such that } e_j \text{ is connected to } c_i \end{aligned}$$

on an edge by edge basis. Initially, $C(N)$ can take any value in the list

$$\mathcal{L}_N = \{1, 1, \ldots, 1, 2, 2 \ldots, 2, \ldots, m, m, \ldots, m\}$$

which is formed by repeating every element of the set $\{1, 2, \ldots, m\}$ a times. Once $C(N)$ is selected, the list $\mathcal{L}_{N-1}$ from which $C(N-1)$ can be selected is obtained by removing $C(N)$ from $\mathcal{L}_N$. Likewise, $\mathcal{L}_j$ is obtained by removing $C(j+1)$ from the list $\mathcal{L}_{j+1}$. We construct the mapping C using either one of the following two methods.

Progressive Edge Growth Algorithm [60] The PEG algorithm allows to construct a graph free of cycles of length $4, 6, \ldots, 2S$. Assuming that $C(N), C(N-1), \ldots, C(j+2), C(j+1)$ have all been determined, and that there are no cycles of length smaller than $2S+2$, then $C(j)$ is determined as follows. We first randomly select $C(j)$ from $\mathcal{L}_j$. Then, a local neighborhood originating at $V(j)$ is expanded up to depth S (cf. Fig. 4.1 for $S = 4$). If there is no cycle in the local neighborhood then let $\mathcal{L}_{j-1} = \mathcal{L}_j - \{C(j)\}$ and proceed to edge e_{j-1}, otherwise select another $C(j)$ and redo the previous step.

Edges e_j are assigned in the decreasing order $j = N, N-1, \ldots, 1$, i.e., in the non-increasing order of repetition degrees $r_{V(N)}, r_{V(N-1)}, \ldots, r_{V(1)}$, because it is easier to assign edges connected to high-degree information bitnodes under a girth constraint at the beginning of the algorithm than toward the end.

Stopping Set Maximization (SSMAX) [63] The algorithm proposed in [63] attempts to maximize the size of stopping sets in the graph, by ensuring that cycles of length $4, 6 \ldots, 2d_{ACE}$ have $ACE \geq \eta$, where ACE stands for approximate cycle EMD, and EMD in turn stands for extrinsic message degree. ACE is the number of external edges connected to a cycle, and is given as

$$ACE \triangleq \sum_i (d_i - 2)$$

where d_i is the degree of the i^{th} (information or parity) bitnode in the cycle. Assuming that $C(N), C(N-1), \ldots, C(j+2), C(j+1)$ have all been determined, and that all cycles of length up to $2d_{ACE}$ have $ACE \geq \eta$, then $C(j)$ is determined as follows. We first randomly select $C(j)$ from $\mathcal{L}_j$. Then, a local neighborhood originating at $V(j)$ is expanded up to depth S. If all cycles

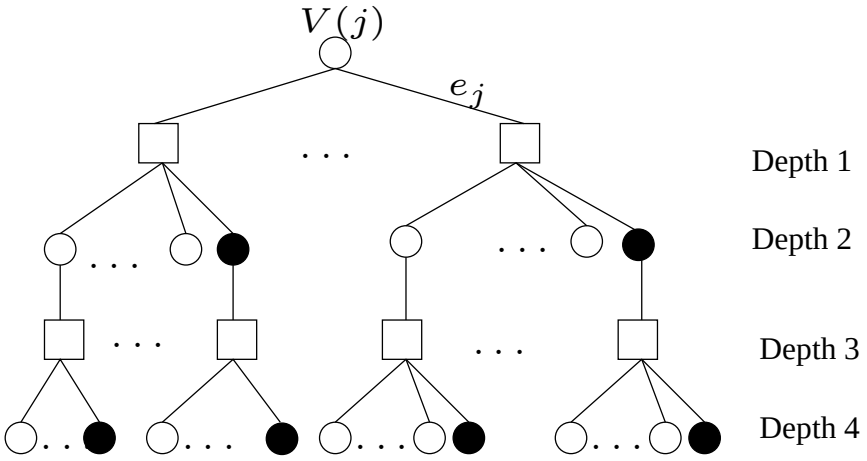


Figure 4.1: Local neighborhood expanded on 4 levels

in the local neighborhood, of length up to $2d_{ACE}$, have $ACE \geq \eta$, then let $\mathcal{L}_{j-1} = \mathcal{L}_j - \{C(j)\}$ and proceed to edge e_{j-1}, otherwise select another $C(j)$ and redo the previous step.

The following proposition shows two simple conditions, that if satisfied, ensure that the resulting graph is free of cycles of length 4.

Proposition 4.1 *A graph satisfying the following two conditions*

1.

$$\forall j, j' \in \{1, \ldots, N\},\ j' \neq j$$

$$if\, V(j') = V(j) \Rightarrow |C(j) - C(j')| \geq 2 \tag{4.1}$$

2.

$$\forall j, j', j_1, j_1' \in \{1, \ldots, N\},\ j' \neq j,\ j_1 \neq j',\ j_1' \neq j_1$$

$$\left.\begin{array}{rl} & V(j') = V(j) \\ if & C(j_1) = C(j) \\ & V(j_1') = V(j_1)\} \end{array}\right\} \Rightarrow |C(j') - C(j_1')| \geq 1 \tag{4.2}$$

is free of cycles of length 4.

Proof: See Appendix 4.A.

4.3 Upper Bound on the Girth of IRA Graphs

An upper bound on the girth of IRA codes can be obtained using results in [65], which were used in [62] to derive an upper bound on the girth of LDPC codes. The average variable node degree α and the average check node degree μ are given as:

$$\begin{aligned} \alpha &= \frac{\bar{d}k(2+a)-a}{k(a+\bar{d})} \\ \mu &= a+2-\frac{1}{n} \end{aligned} \tag{4.3}$$

which can be approximated, for practical information block length k, as

$$\begin{aligned} \alpha &\simeq \frac{\bar{d}(2+a)}{a+\bar{d}} \\ \mu &\simeq a+2 \end{aligned} \tag{4.4}$$

Let g denote the girth of the graph. Then, from [65], the following upper bound can be derived:

$$g \leq 4\frac{\log\left[\frac{k+m}{\mu}((\mu-1)(\alpha-1)-1)+1\right]}{\log\left[(\mu-1)(\alpha-1)\right]} \tag{4.5}$$

then, using (4.4) and $m = k\bar{d}/a$, we get:

$$g \leq 4\frac{\log\left[k(\bar{d}-1)+1\right]}{\log\left[\frac{a+1}{a+d}(\bar{d}a+\bar{d}-a)\right]} \tag{4.6}$$

Using the adjacency matrix of the parity check matrix associated to a Tanner graph [62], we can evaluate the girth of graphs generated by the methods described in the previous section. Let us consider regular RA codes with parameters $a = 4$ and $d = 4$ generated using the progressive edge growth method. The following table shows the minimum k required to obtain a graph with girth g, as well as the corresponding theoretical girth upper bound.

k	g_{bound}	g
21	7.2	6
190	11.0	8
1930	15.0	10

Table 4.1: Theoretical and true girth of short-length regular RA graphs

4.4 Maximum Likelihood Decoding

An upper bound on the BER and WER of the random ensemble of IRA codes, under ML decoding on the BIAWGNC, can be determined using the tangential sphere bound (TSB) [66, 67, 68] detailed in Appendix 4.B. The computation of the TSB on the BER and WER requires the knowledge of the input-output weight enumerators (IOWE) of the random IRA ensemble $A_{w,h}$, $w = 0, 1, \ldots, k$ and $h = 0, 1, \ldots n$. $A_{w,h}$ is the number of output codewords with weight h, generated by information information words of weight w. Using the uniform interleaver technique [69], we can calculate the average $A_{w,h}$, by assuming that the IRA encoder has a uniform interleaver at the output of the repetition encoder, and a second uniform interleaver between the grouping and the accumulator (see Fig. 4.2). Then

$$\begin{aligned} A_{w,h+w} &= \sum_{p=1}^{N} \frac{A^R_{w,p}}{\binom{N}{p}} \sum_{l=0}^{m} \frac{A^G_{p,l} A^A_{l,h}}{\binom{m}{l}}, \\ & w = 0, 1, \ldots, k \quad h = 0, 1, \ldots, m \end{aligned} \tag{4.7}$$

where $A^R_{w,p}, A^G_{p,l}, A^A_{l,h}$ are the IOWE of, respectively, the repetition code, the grouping and the accumulator (2-state rate-1 convolutional code).

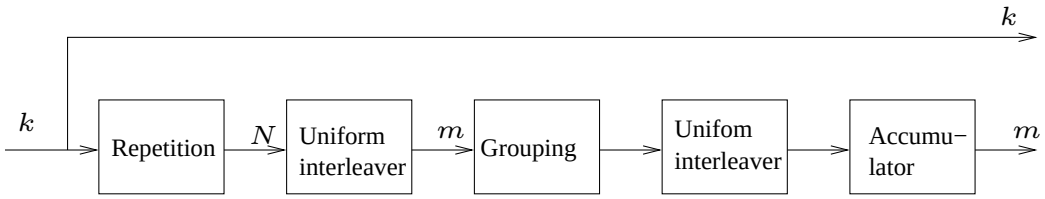


Figure 4.2: Modified IRA encoder with uniform interleavers to compute the IOWE

The following IOWE computation is the extension of the results in [15] to the case $a > 1$ and irregular repetition pattern.

4.4.1 IOWE of Repetition Code

The IOWE of the repetition code is given by

$$A^R_{w,h} = \sum_{\mathbf{w}} A^R_{\mathbf{w},h} \tag{4.8}$$

where

$$A^R_{\mathbf{w},h} = \begin{cases} \prod_{j=1}^{n_{deg}} \binom{f_j k}{w_j} & \text{if} \quad h = \sum_{j=1}^{n_{deg}} d_j w_j \\ 0 & \text{otherwise} \end{cases} \tag{4.9}$$

and the summation is carried out over all ordered integer partitions $\mathbf{w} = [w_1, \ldots, w_{n_{deg}}]$ of w into at most n_{deg} parts, i.e., $\sum_{j=1}^{n_{deg}} w_j = w$ and $w_j = 0, 1, \ldots, w$. Integer partitions of w are generated using algorithms in [70]. The number of these partitions grows at least as $\Theta(w^{n_{deg}-1})$ [70], making the computation of the irregular repetition IOWE intractable for information block lengths $k > 200$.

If $n_{deg} = 1$, the repetition is regular with repetition degree d, and the resulting IOWE is [15]

$$A^R_{w,h} = \begin{cases} \binom{k}{w} & \text{if} \quad h = dw \\ 0 & \text{otherwise} \end{cases} \tag{4.10}$$

4.4.2 IOWE of Grouping

Computing the number of codewords of input weight $w = 0, \ldots, N$ and output weight $h = 0, \ldots, m$ is equivalent to determining the number of ordered partitions of w into m parts, of which exactly h parts are odd, and each part $w_i \leqslant a$. $A^G_{w,h}$ is the coefficient of x^w in $G(x)$ [71], denoted by $\lfloor G(x) \rfloor_w$, where

$$G(x) = \binom{m}{h} 2^{-m} \left((1+x)^a - (1-x)^a\right)^h \left((1+x)^a + (1-x)^a\right)^{m-h} \tag{4.11}$$

Using the binomial theorem to simplify (4.11)

$$(x+y)^m = \sum_{l=0}^{m} \binom{m}{l} x^l y^{m-l} \tag{4.12}$$

the IOWE of grouping with $a = 2$ is

$$\lfloor G(x) \rfloor_w = \begin{cases} 2^h \binom{m}{h} \binom{m-h}{\frac{w-h}{2}} & \text{if } w-h \text{ is even} \\ 0 & \text{if } w-h \text{ is odd} \end{cases} \tag{4.13}$$

and the IOWE of grouping with $a = 4$ is

$$\lfloor G(x) \rfloor_w = \begin{cases} 2^h \binom{m}{h} \sum_{l=0}^{l_{max}} \binom{m-h}{l} 2^{2l} \binom{2m-2l-h}{\frac{w-h}{2}-l} & \text{if } w-h \text{ is even} \\ 0 & \text{if } w-h \text{ is odd} \end{cases} \tag{4.14}$$

where

$$l_{max} = \min\left(2m - \frac{w+h}{2}, m-h, \frac{w-h}{2}\right) \tag{4.15}$$

For $a \neq 2$ and $a \neq 4$, the grouping IOWE cannot be expressed in closed-form. However, noticing that the coefficients of $G(x)$ are positive, the following result [72] provides an upper bound on the grouping IOWE:

$$\lfloor G(x) \rfloor_w \leq \inf_{x>0} \frac{G(x)}{x^w} \tag{4.16}$$

4.4.3 IOWE of Accumulator (without grouping)

As stated in [8], the IOWE of the accumulator is given as:

$$A^A_{w,h} = \binom{m-h}{\lfloor w/2 \rfloor} \binom{h-1}{\lceil w/2 \rceil - 1} \tag{4.17}$$

A section of the trellis of the accumulator is shown in figure 4.3. It consists of a total of $\lceil w/2 \rceil$ detours, where a regular detour is a path starting and ending in the zero state, and a non-regular detour is a path that starts at the zero state but does not return to the zero state (it is the last detour, if the trellis is not terminated) [73]. Computing $A^A_{w,h}$ is equivalent to computing the total number of detours of all possible lengths, for an input weight w and an output weight h. Or equivalently, one can enumerate all possible detour lengths whose sum equals h, then, for each output weight distribution, enumerate the different possible positions. The number of ordered partitions of h into exactly $\lceil w/2 \rceil$ detours can be obtained using the generator function method, and is given by:

$$\binom{h-1}{\lceil w/2 \rceil - 1}$$

which is equivalent to giving the lengths of the detours in all codewords. For each of these configurations, the codeword is completely defined if the

positions of the detours are known. These positions are known as soon as the $m-h$ output zero positions are known. If w is odd, then the last detour is irregular, and all the $m-h$ are distributed among the $\lfloor w/2 \rfloor$ regular detours. Therefore, the number of codewords for each output weight distribution is

$$\binom{m-h}{\lfloor w/2 \rfloor}$$

and hence the result of (4.17).

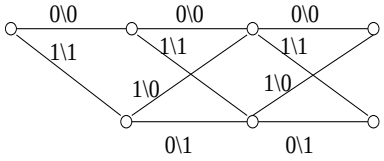


Figure 4.3: A trellis section of the accumulator

4.4.4 Regular RA Code with Grouping Factor $a = 2, 4$

Consider a regular RA code with repetition degree d. Using (4.13) and (4.14), the IOWE of the RA code with grouping factor $a = 2$ is

$$A_{w,w+h} = \frac{\binom{k}{w}}{\binom{dk}{dw}} \sum_t 2^t \binom{m-t}{\frac{dw-t}{2}} \binom{m-h}{\lfloor t/2 \rfloor} \binom{h-1}{\lceil t/2 \rceil - 1} \tag{4.18}$$

and the IOWE with grouping factor $a = 4$ is

$$A_{w,w+h} = \frac{\binom{k}{w}}{\binom{dk}{dw}} \sum_t 2^{2t} \binom{m-h}{\lfloor t/2 \rfloor} \binom{h-1}{\lceil t/2 \rceil - 1} \sum_{l=0}^{l_{max}} 2^{2l} \binom{m-t}{l} \binom{2m-2l-t}{\frac{dw-t}{2} - l} \tag{4.19}$$

where

$$\begin{aligned} w &= 0, \ldots, k \\ h &= 0, \ldots, m \\ l_{max} &= \min(m-t, \tfrac{dw-t}{2}) \end{aligned}$$

and summations are on all integers t such that $dw - t$ is even and $t \leq \min(m, dw)$.

4.5 Simulation Results

4.5.1 Regular RA Codes

Figs. 4.4(a), 4.5(a) and 4.6(a) show the average performances of randomly and semi-randomly PEG constructed regular RA codes, with $S = 2$ and $S = 3$ (graphs are free of cycles of length up to $2S$), on the BIAWGNC under BP decoding with 25 decoder iterations. Also shown is the TSB on the BER and WER of the random regular RA ensemble under ML decoding. The code rate is $R = 1/2$, and the information block lengths are respectively $k = 150$, $k = 256$ and $k = 512$. The code parameters $d = 4$ and $a = 4$ were selected because they correspond to the best DE threshold of the random-like regular RA ensemble at rate $R = 1/2$.

The semi-random regular RA ensemble under BP decoding outperforms its random counterpart under both BP decoding and ML decoding, in the error floor region. We notice that, as expected for such short code lengths, the BER waterfall is far from the DE evolution threshold $\left(\frac{E_b}{N_0}\right)^* = 0.2108$ dB.

Comparing the semi-random regular RA codes with the regular LDPC codes proposed in [62] of the same conditioning level (girth 6 and 8) and the same information block lengths, we note the following. For $k = 150$, the girth 6 semi-random regular RA code outperforms the regular LDPC code by 0.3 dB at a BER of 2×10^{-6}. For $k = 512$, the girth 6 and 8 regular RA codes outperform the LDPC code in the waterfall region (at $E_b/N_0 = 3$ dB, BER of semi-random regular RA code is 5×10^{-7} while the BER of the LDPC code is 5×10^{-6}). In the error floor region, BER performances of RA codes of girth 6 and 8 are comparable to those of LDPC codes(at $E_b/N_0 = 3.5$ dB, the BER of the girth 8 regular RA code is 2×10^{-8} while the BER of the girth 8 LDPC code is 10^{-8}). Note that the complexity of the PEG algorithm (as well as that of the SSMAX) is exponential in the girth and linear in the block length, whereas that of the algorithm proposed in [62] is linear in the girth and polynomial in the block length. Therefore, for short block length and small girth, it is preferable to use the PEG algorithm for girth conditioning.

Comparing the semi-random regular RA codes of girth 6 with the regular LDPC codes proposed by MacKay in [74] of information block length around $k = 256$, we note that their BER performances are similar. At $E_b/N_0 = 3.5$ dB, the BER of the RA code is 2×10^{-6}, and that of the MacKay LDPC code is 10^{-6}. The regular RA codes of girth 8 are found to outperform both

MacKay regular LDPC codes and PEG-constructed LDPC codes of girth 8 [60] for information block length $k = 256$. In fact, at $E_b/N_0 = 3.25$ dB, the BER of the RA code is 5×10^{-7}, while the BER of of the McKay LDPC code is 4×10^{-6} and that of the PEG LDPC code is 2×10^{-6}.

Using the error impulse method proposed in [75, 76], and described in Appendix 4.C, we compute the minimum distances (d_{min}) of 200 realizations of randomly and PEG constructed regular RA codes. Table 4.2 shows the minimum, maximum and average d_{min} thus obtained, and indicates that as the girth of the bipartite graphs increases from 4 to 6, so does the minimum distance of the associated regular RA codes. Figs. 4.4(b), 4.5(b) and 4.6(b) show the performances of rate 1/2 random and PEG codes with respective maximum d_{min}, for information block lengths k equal to 150, 256 and 512, respectively.

k	150			256			512		
Girth	4	6	8	4	6	8	4	6	8
min. d_{min}	2	5	n.a.	2	5	6	2	5	6
max. d_{min}	6	8	n.a.	8	9	11	9	11	13
av. d_{min}	4.11	6.54	n.a.	5.27	7.07	8.97	6.09	8.59	10.49

Table 4.2: Minimum, maximum and average minimum distance d_{min} vs. girth of short-length regular RA codes

4.5.2 Irregular RA Codes

Fig. 4.7 compares the average BER performance of randomly and semi-randomly (PEG and SSMAX) constructed irregular RA codes to that of an LDPC code of the same rate, on the BIAWGNC under BP decoding with 25 decoder iterations. The code rate is $R = 0.5$, and the information and parity block lengths are respectively $k = 5020$ and $m = 4940$. The code degree sequences and grouping factor are obtained by optimizing the code with method 1 in chapter 3, and have a maximum repetition degree of 20.

Although the maximum achievable girth of the considered IRA graph is 12.4, the maximum girth obtained using the PEG algorithm is 6. The curves labeled (d_{cyc}, d_{ACE}, η) in fig. 4.7 represent the average performances of codes constructed with the SSMAX method. (d_{cyc}, d_{ACE}, η) means that the IRA code is free of cycles of length up to $2d_{cyc}$, and all cycles of length up to d_{ACE} have $ACE \geq \eta$.

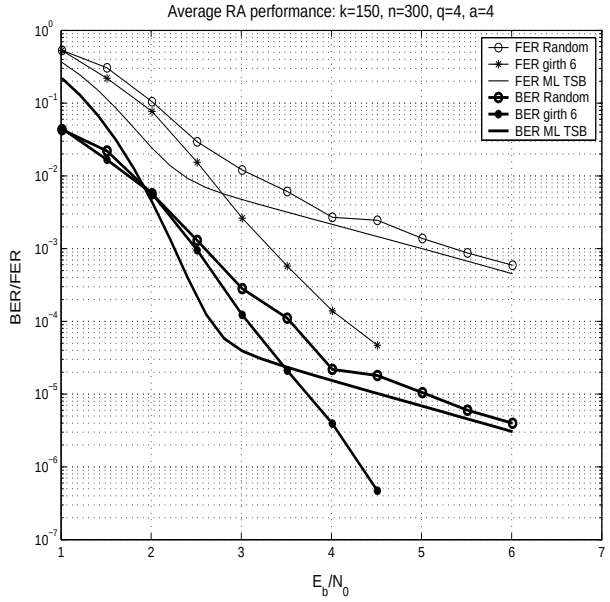
The LDPC code has rate 1/2, a code block length of 10000, and a degree sequence optimized in [9]. It is generated randomly, but the degree-2 nodes are arranged in a single cycle.

Fig. 4.7 shows that the SSMAX method yields a better error floor than the PEG algorithm with $S = 2$. Comparing the $(1, 9, 3)$ code with the randomly constructed LDPC code ensemble, we note that it outperforms the random LDPC in the error floor region. But, its performance is inferior to that of the $(1, 9, 4)$ LDPC code of [63].

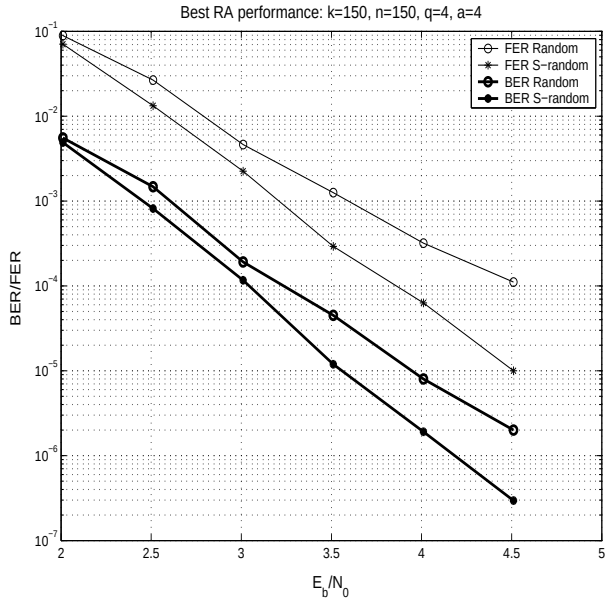
4.6 Conclusion

We have presented a comparative study of the performance of finite length regular and irregular repeat accumulate codes, constructed according to two criteria: girth maximization and stopping set maximization. Our simulations show that girth conditioning yields an improvement in the error floor region of short-length regular RA codes in the BER and WER, as compared to the random regular RA ensemble. The codes thus designed perform as well as the best known LDPC codes [60, 61, 62, 63]. Indeed, increasing the girth of the Tanner graph has a direct effect on the minimum distance of the code, which increases accordingly, therefore improving the code performance in the error floor region. Hence, the performance-complexity trade-off of the constructed regular RA codes is very advantageous.

Large block length irregular RA codes exhibit a better error floor using the stopping set maximization method, as compared to the random and girth-conditioned IRA ensembles. But the average IRA code performance remains inferior to that of LPDC codes with comparable graph conditioning and block length.

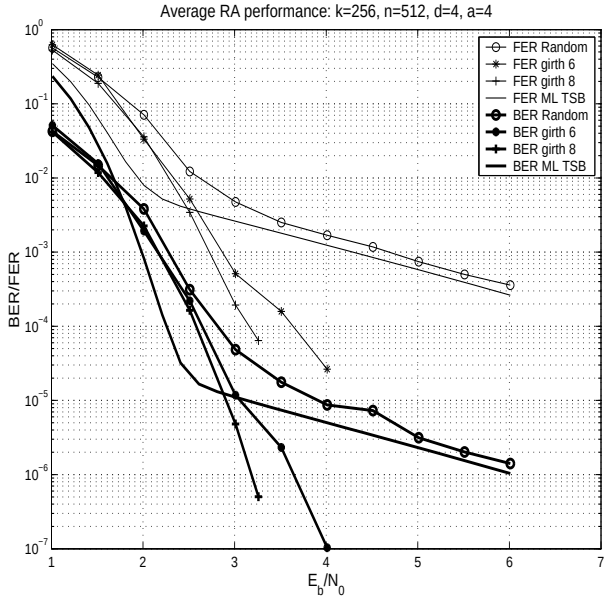


(a)

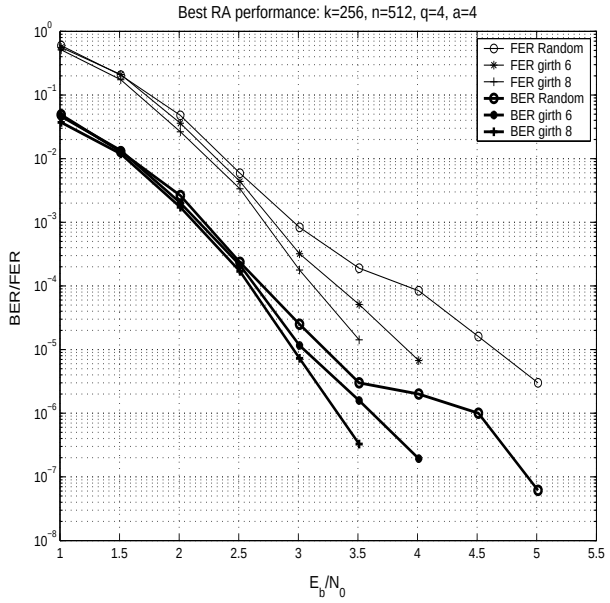


(b)

Figure 4.4: Average (a) and best (b) regular RA performances with $k = 150$, $n = 300$, $d = 4$, $a = 4$

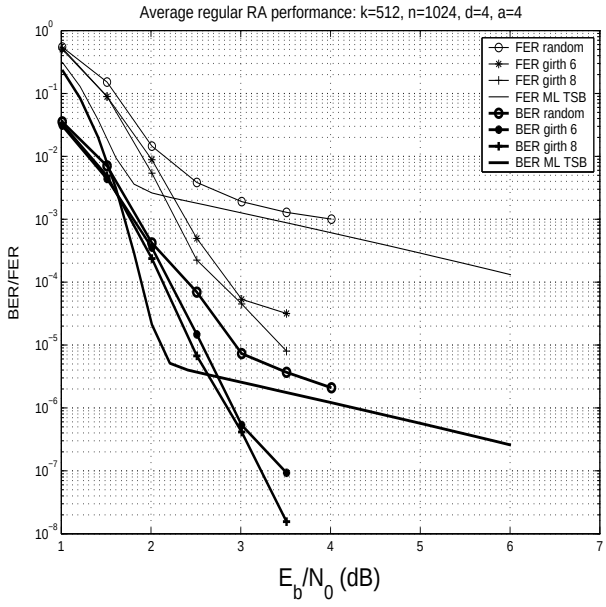


(a)

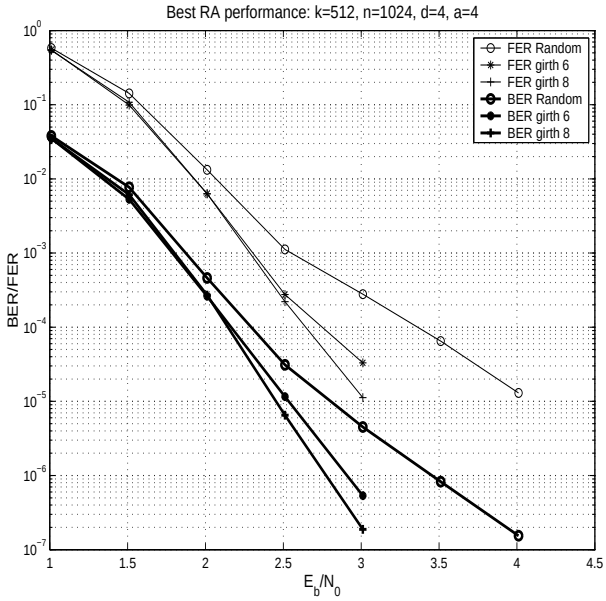


(b)

Figure 4.5: Average (a) and best (b) regular RA performances with $k = 256$, $n = 512$, $d = 4$, $a = 4$



(a)



(b)

Figure 4.6: Average (a) and best (b) regular RA performances with $k = 512$, $n = 1024$, $d = 4$, $a = 4$

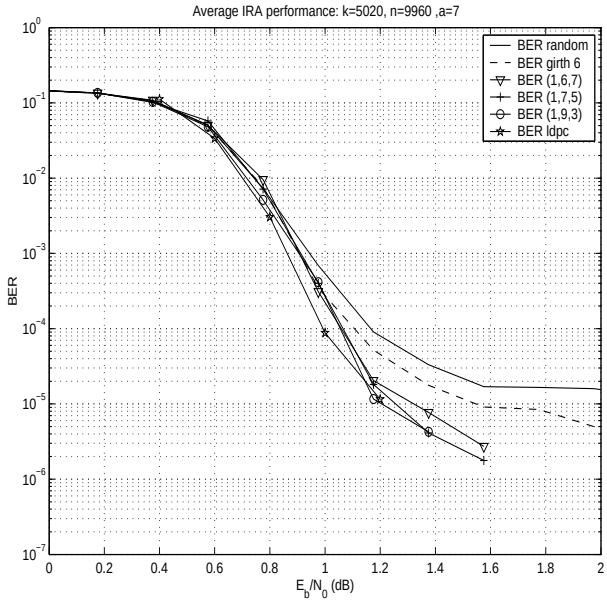


Figure 4.7: Average IRA performance with $k = 5020$, $n = 9960$, $\bar{d} = 6.89$, $a = 7$

APPENDIX

4.A Proof of Proposition 4.1

Suppose conditions 1 and 2 are satisfied, and that the graph contains a cycle of length 4. There are only two ways in which this cycle forms in the graph of an IRA:

1. The cycle is composed of one information bitnode and one parity bitnode (cf. Fig. 4.8(a)). Because of the zigzag pattern of the graph of the accumulator, the two checknodes in the length-4 cycle are adjacent to each other, therefore the distance between them is exactly 1. There remain two edges connecting the information bitnode to the two checknodes. Then, condition 1 is violated.

2. The cycle is composed of two information bitnodes sharing two checknodes (cf. Fig. 4.8(b)). Denoting the four edges of the cycle as j, j_1, j', j'_1, and letting $V(j) = V(j')$ and $V(j_1) = V(j'_1)$, then because condition 1 is met, $|C(j) - C(j')| \geq 2$ and $|C(j_1) - C(j'_1)| \geq 2$. Then because the cycle is of length 4, $C(j) = C(j_1)$ and $C(j') = C(j'_1)$. This is a contradiction with condition 2, which requires one of the two distances $|C(j) - C(j_1)|$ and $|C(j') - C(j'_1)|$ to be at least equal to 1. □

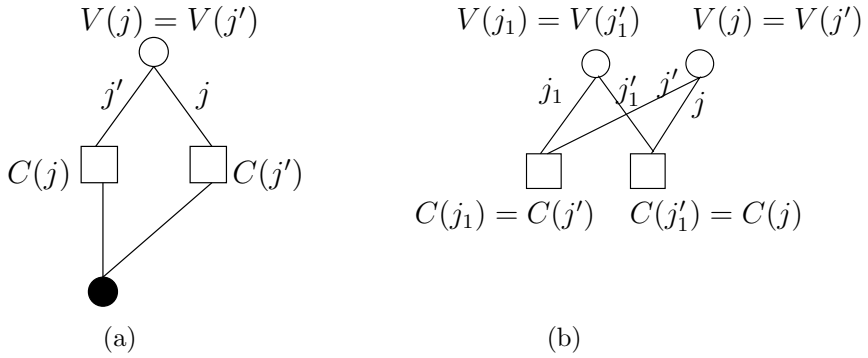


Figure 4.8: Length-4 cycles

4.B Tangential Sphere Bound

Consider an (n, k) linear block code and its input-output weight enumerating function:

$$A(W, H) = \sum_{w=0}^{k} \sum_{h=0}^{n} A_{w,h} W^w H^h \tag{4.20}$$

and define

$$\begin{aligned} A_h &= \sum_{w=0}^{k} A_{w,h} \\ B_h &= \sum_{w=1}^{k} \frac{w}{k} A_{w,h} \end{aligned} \tag{4.21}$$

The tangential sphere upper bound on word error probability is given as [66, 67, 68]

$$\begin{aligned} P_w \leqslant \min_r \int_{-\infty}^{+\infty} \frac{dz_1}{\sigma\sqrt{2\pi}} e^{-z_1^2/2\sigma^2} \Bigg\{ & \\ 1 - \bar{\gamma}(\tfrac{n-1}{2}, \tfrac{r_{z_1}^2}{2\sigma^2}) + & \\ \sum_{h:\delta_h/2<\alpha_h} A_h \left(Q(\frac{\beta_k(z_1)}{\sigma}) - Q((\frac{r_{z_1}}{\sigma})) \right) \bar{\gamma} \left(\frac{n-2}{2}, \frac{r_{z_1}^2 - \beta_h^2(z_1)}{2\sigma^2} \right) \Bigg\} & \end{aligned} \tag{4.22}$$

where

$$\begin{cases} r_{z_1} & = r\left(1 - \frac{z_1}{\sqrt{nE_s}}\right) \\ \beta_h(z_1) & = \frac{\sqrt{nE_s} - z_1}{\sqrt{nE_s - \delta_h^2/4}} \delta_h/2 \\ \alpha_h & = r\sqrt{1 - \frac{\delta_h^2}{4nE_s}} \end{cases} \tag{4.23}$$

and $\bar{\gamma}$ is the normalized incomplete gamma function defined as

$$\bar{\gamma}(a, x) = \frac{1}{\Gamma(a)} \int_0^x t^{a-1} e^{-t} dt \tag{4.24}$$

with $\bar{\gamma}(a, \infty) = 1$. Q is the Q-function (or complementary cumulative distribution function) given by:

$$\begin{aligned} Q : \mathbb{R} &\rightarrow [0, 1] \\ x &\rightarrow \tfrac{1}{\sqrt{2\pi}} \textstyle\int_x^\infty e^{-t^2/2} dt. \end{aligned}$$

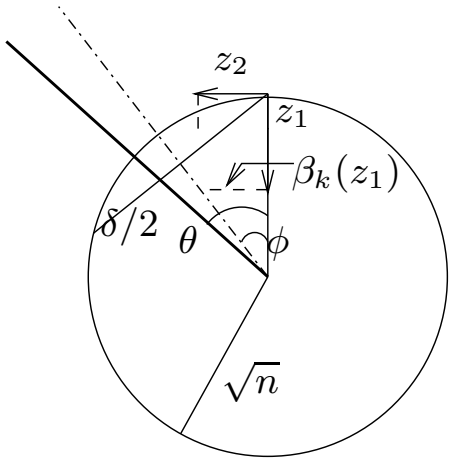


Figure 4.9: Tangential Sphere Bound

The optimal radius r is the solution of the following equation:

$$\begin{cases} \displaystyle\sum_{h:\delta_h/2<\alpha_h} A_h \int_0^{\theta_h} (\sin\varphi)^{n-3} d\varphi = \frac{\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n-1}{2})}\sqrt{\pi} \\ \cos\theta_h = \frac{\delta_h/2r}{\sqrt{1-\frac{\delta_h^2}{4nE_s}}} \end{cases} \tag{4.25}$$

The result in (4.22) is obtained by noticing that $\bar{\gamma}(a,x)$ is an increasing function in x. Since $z_2^2 \geqslant \beta_h^2(z_1)$, then

$$\bar{\gamma}\left(\frac{n-2}{2}, \frac{r_{z_1}^2 - z_2^2}{2\sigma^2}\right) \leqslant \bar{\gamma}\left(\frac{n-2}{2}, \frac{r_{z_1}^2 - \beta_h^2(z_1)}{2\sigma^2}\right)$$

The tangential sphere upper bound on the bit error rate has the same expression as in (4.22), except that A_h is replaced with B_h. This, again, applies to the optimization condition (4.25).

For antipodal modulation, $\delta_h = 2\sqrt{hE_s}$, and the condition $\delta_h/2 < \alpha_h$ becomes:

$$h < \frac{1}{\frac{1}{n} + \frac{1}{r^2}}$$

4.C Minimum Distance Estimation

Consider an (m,k) linear code $\mathcal{C}$, on the BIAWGNC, with BPSK modulation, where the "all-zero" codeword $\mathbf{x}_0 = [-1,\ldots,-1]$ is transmitted. Applying

an impulse error at position i to the all-zero codeword, the input to the ML decoder is $\mathbf{y} = [-1, \ldots, -1, -1 + A_i, -1, \ldots, -1]$. The decoded codeword $\hat{\mathbf{x}}$, under ML decoding is such that

$$\forall x \in \mathcal{C}\ \langle \hat{x}, y \rangle \geqslant \langle x, y \rangle \tag{4.26}$$

where $\langle x, y \rangle$ is the scalar product between x and y. Let $w_H(\mathbf{x})$ be the weight of a codeword $\mathbf{x}$. It can be shown [75, 77] that

$$\hat{\mathbf{x}} \neq \mathbf{x}_0 \quad \Rightarrow \quad A_i \geqslant \min_{\mathbf{x} \in \mathcal{C}, x_i = +1} w_H(\mathbf{x}) \tag{4.27}$$

$$\hat{\mathbf{x}} = \mathbf{x_0} \quad \Rightarrow \quad A_i \leqslant \min_{\mathbf{x} \in \mathcal{C}, x_i = +1} w_H(\mathbf{x}) \tag{4.28}$$

Let

$$\begin{aligned} A_i^* &= \max\{A_i | \hat{\mathbf{x}} = \mathbf{x_0}\} \\ &= \min\{A_i | \hat{\mathbf{x}} \neq \mathbf{x_0}\} \\ &= \min_{\mathbf{x} \in \mathcal{C}, x_i = +1} w_H(\mathbf{x}) \end{aligned} \tag{4.29}$$

then, the minimum distance of the code is the minimum impulse error amplitude, over all positions i, such that the decoded codeword is not the all-zero codeword, i.e.,

$$d_{min} = \min_i A_i^* \tag{4.30}$$

The minimum distance is the minimum weight among all ensembles of codewords $\{bfx\}$ such that $x_i = 0$. In practice, fix a position i, add a small error impulse A_i to the all-zero codeword, and increment A_i until the decoded codeword is no longer the all-zero codeword, for which the error impulse is A_i^*. The code minimal distance is obtained by testing all positions $i = 1, \ldots, m$.

We use the iterative turbo decoder, instead of ML, although the optimality of iterative turbo decoding has not been proved, especially on a non-realistic channel such as the error impulse channel considered here. In fact, the number of decoding iterations becomes an important issue.

Algorithm Assuming that $d_{\min}$ lies in the range $[d_0, d_1]$, the minimum distance is determined following these steps:

- $A_{\min} = d_1 + \epsilon$, $\epsilon << 1$
- for $i = 1 : k$

  - $A = d_0 - \epsilon$
  - flag = 1
  - while (flag==1) and ($A < A_{\min}$)
    - $A = A + 1$
    - $\mathbf{y} = [-1, \ldots, -1, -1 + A, -1, \ldots, -1]$, where $-1 + A$ is at position i
    - Decode $\mathbf{y} \rightarrow \hat{\mathbf{x}}$, using ML or iterative decoding
    - if ($\hat{\mathbf{x}} \neq \mathbf{x_0}$) then flag=0
  - end while
  - $A_{\min} = A$

- $d_{\min} = [A_{\min}]$

Part II

Coded CDMA under Successive Decoding

Chapter 5

Spectral Efficiency of Coded CDMA

We investigate the spectral efficiency achievable by random synchronous CDMA in the large system limit where the number of users, the spreading factor and the code block length go to infinity. We quantify the loss in efficiency incurred by the use of random CDMA (with Gaussian inputs, QPSK inputs and/or sub-optimum linear MMSE decoding) with respect to the capacity of the multiple access channel without spreading.

5.1 Synchronous CDMA Canonical Model

We consider the complex baseband discrete-time channel model

$$\mathbf{y}_i = \mathbf{S}\mathbf{x}_i + \mathbf{n}_i, \qquad i = 1, \ldots, n \tag{5.1}$$

originated by sampling at the chip-rate a synchronous CDMA system [29], where:

1) $\mathbf{y}_i, \mathbf{n}_i \in \mathbb{C}^N$, are the vector of received chip-rate samples and the corresponding additive white Gaussian noise (AWGN) samples $\sim \mathcal{N}_{\mathbb{C}}(0, 1)$ received at time i;

2) $\mathbf{S} \in \mathbb{C}^{N\times K}$ contains the user spreading sequences by columns. Spreading sequences are proper complex[1] [78], known to the receiver, with i.i.d. zero-mean chips, variance $1/N$ and finite fourth order moment;

3) $\mathbf{x}_i \in \mathbb{C}^K$ is the vector of user modulation symbols transmitted at time i, where its k-th component $x_{k,i}$ takes on values in some signal constellation, with given average energy per symbol (i.e., per transmitted N chips) $E[|x_{i,k}|^2] = \alpha_k\,\mathsf{SNR}$. The scaling factors α_k represent *power control* and, without loss of generality, are normalized such that $\frac{1}{K}\sum_{k=1}^{K}\alpha_k = 1$;

4) N, K and n denote the spreading factor, the number of users and the code block length, respectively.

For the purpose of system design it is convenient to consider a system formed by L user classes. The size of class j is K_j, and we denote by $\beta_j = K_j/N$ the "class load" of class j. Thus, the total *channel load* is

$$\beta = \sum_{j=1}^{L}\beta_j \quad \text{users/chip}$$

Users in class j have the same SNR, denoted by γ_j (i.e., $\alpha_k\,\mathsf{SNR} = \gamma_j$ for all users k in class j). Without loss of generality, we assume $\gamma_1 \le \cdots \le \gamma_L$.

The fundamental figure of merit of the CDMA system is the *spectral efficiency*, that we equally denote by ρ or C. It is defined as the number of bits per chip that can be transmitted arbitrarily reliably (number of bits per complex dimension), and is measured in bits per second per Hertz (bits/s/Hz), since each second $\times$ Hertz requires one complex dimension. The total system spectral efficiency is given by

$$\rho = \sum_{j=1}^{L}\beta_j R_j \ \text{ bit/s/Hz} \tag{5.2}$$

where R_j denotes the average rate of users in class j.

[1] This is the generalization of the circular symmetry of a scalar complex random variable to a complex random vector

The user individual E_b/N_0's are different in general. Nevertheless, for the sake of comparison with a reference equal-rate equal-power system, it is convenient to define a "system" E_b/N_0 by

$$\left(\frac{E_b}{N_0}\right)_{\text{sys}} \triangleq \frac{\sum_{j=1}^{L} \beta_j \gamma_j}{\sum_{j=1}^{L} \beta_j R_j} = \frac{\sum_{j=1}^{L} \beta_j \gamma_j}{\rho} \tag{5.3}$$

which coincides with the individual E_b/N_0's in the case where users are dynamically assigned to the classes so that each user belongs to class j for a fraction β_j/β of the time.

5.2 Gaussian Multiple Access Channel

The conventional discrete-time synchronous Gaussian multiple access channel (GMAC) is modeled by

$$y_i = \sum_{k=1}^{K} x_{i,k} + n_i \quad i = 1, \cdots, n \tag{5.4}$$

where $x_{i,k}$ is the symbol transmitted by user k at time i and n_i is an i.i.d. zero-mean Gaussian noise sample $n_i \sim \mathcal{N}_{\mathbb{C}}(0,1)$. Under the assumption that the power of the k^{th} user is constrained to P_k, and letting r_k denote the k^{th} user rate, Cover [79] and Wyner [80] showed that the capacity region of (5.4) is the subset of $\mathbb{R}^K$ containing the rate K-tuples $(r_1, r_2, \cdots, r_K)$ satisfying [51]

$$\forall S \subseteq \{1, \cdots, K\} \qquad \sum_{k \in S} r_k \leq \log_2 \left(1 + \sum_{k \in S} P_k\right) \tag{5.5}$$

An example of the Cover-Wyner capacity region of a 2-user GMAC is shown in Fig. 5.1.

Let us define a vertex (a corner point) of the capacity region as a rate $K-$tuple that satisfies with equality K (out of $2^K - 1$) inequalities (5.5).

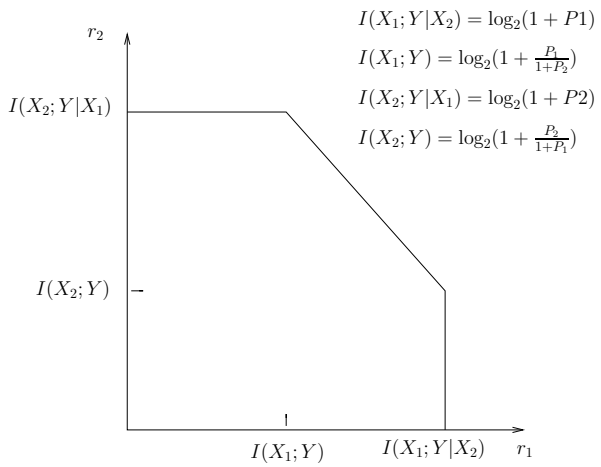


Figure 5.1: Achievable capacity region of a 2-user Gaussian multiple access channel

Then, after some possible re-indexing, vertices satisfy with equality

$$\forall k \in \{1, \cdots, K\} \qquad r_k = \log_2 \left(1 + \frac{P_k}{1 + \sum_{j<k} P_j} \right) \tag{5.6}$$

For these points, the channel can be seen as a set of single-user channels with SNR's $P_k/(1+\sum_{j<k} P_k)$, i.e., interfering users are seen as Gaussian noise, and user k transmits at the maximum single-user rate $\log_2(1+P_K/(1+\sum_{j<K} P_j))$. Then, the decoder of user K decodes considering all users $k = 1, \cdots, K-1$ as Gaussian noise. Since user K has been decoded without error, its contribution can be removed from the received word, and the decoder carries on the successive single-user encoding/decoding and interference cancellation (IC) procedure described above until all users have been decoded. This decoding procedure is known as *successive cancellation*, *stripping* and *onion peeling*. The non-corner points of the Cover-Wyner capacity region are achievable by time sharing or rate/power-splitting [51, 81].

In the absence of spreading, i.e., $N = 1$, the channel (5.1) becomes the conventional Gaussian multiple-access channel

$$y_i = \sum_{k=1}^{K} x_{i,k} + n_i \tag{5.7}$$

where again $n_i \sim \mathcal{N}_{\mathbb{C}}(0,1)$. The Cover-Wyner capacity region of the conventional GMAC [79, 80, 51] applies to this case and is given by

$$\forall S \subseteq \{1, \cdots, K\} \qquad \sum_{k \in S} R_{j(k)} \leq \log_2 \left(1 + \sum_{k \in S} \gamma_{j(k)}\right) \tag{5.8}$$

where $j(k) = j$ if user k belongs to class j. The total maximum spectral efficiency in the absence of spreading is then given by

$$\mathsf{C}^\star = \log_2(1 + \sum_{j=1}^{L} \beta_j \gamma_j) \tag{5.9}$$

where β_j is used *in lieu* of K_j (they are equal in the absence of spreading) for consistency with the results in the case of spreading. The spectral efficiency has to be expressed as a function of $\frac{E_b}{N_0}$ in order to be able to compare different systems with possibly different parameters. If the spectral efficiency reaches the maximum given by (5.9), then

$$\sum_{j=1}^{L} \beta_j \gamma_j = \mathsf{C}^\star \frac{E_b}{N_0}$$

and the spectral efficiency in the absence of spreading is the solution to

$$\mathsf{C}^\star = \log_2(1 + \mathsf{C}^\star \frac{E_b}{N_0}) \tag{5.10}$$

which coincides with the additive white Gaussian noise channel (AWGNC) single-user capacity, implicitly given by

$$\frac{2^{\mathsf{C}^\star} - 1}{\mathsf{C}^\star} = \frac{E_b}{N_0} \tag{5.11}$$

The solution of (5.10) exists if and only if $\frac{E_b}{N_0} \geq \log_e 2 = -1.5917$ dB.

The Cover-Wyner capacity region was generalized in [82] to encompass the case of CDMA with arbitrary signature waveforms. As in the scalar case, all points in the boundary of the capacity region of the CDMA channel can also be achieved by stripping and single-user encoding/decoding, as long as the stripping decoders incorporate MMSE filters against not yet decoded users at each successive cancellation stage [83]. Recall that the MMSE filter of

a particular user maximizes the output Signal to Interference plus Noise Ratio (SINR) at the output of this linear transformation. Key to the optimality of stripping is the use of Gaussian codes of rate arbitrarily close to (but not larger) than the capacity of the channel obtained by removing the already decoded users. In this way, optimal spectral efficiency is achieved by simple single-user coding and decoding, with linear complexity in the number of users.

For given signature waveforms, successive stripping generally requires that every user must transmit at a different rate, or must be received at a different SNR level. This can be avoided by designing the signature waveforms such that the equal-rate point coincides with a vertex of the equal-power capacity region [84]. However, optimizing the signature waveforms (e.g., [85, 86, 87]) is highly impractical in real-life applications, where transmission is usually affected by frequency selective fading channels.

On the other hand, existing nonorthogonal CDMA systems [30, 88] are largely based on pseudo-random waveforms. The maximum spectral efficiency of randomly spread (synchronous) CDMA, in the large system limit, where the number of users and the spreading factor grow without bound while their ratio tends to a constant β, was found in the case of power-constrained inputs in [31, 32], and in the case of binary antipodal inputs in [89].

While, in the power-constrained case, capacity is achieved by Gaussian inputs, practical systems make use of discrete small-size modulation alphabets. Given its widespread application and the fact that QPSK is optimal in the wideband low SNR regime [90] we will restrict our analysis to QPSK-modulated CDMA.

5.3 Spectral Efficiency of Random Synchronous CDMA

In [32], the spectral efficiency (in bit/s/Hz) of random CDMA in the large system limit ($K, N \to \infty$ with $K/N = \beta < \infty$) subject to fading with an input power constraint is found to be

$$\mathsf{C}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathsf{C}^{\text{mmse}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) + \log_2 \frac{1}{\eta} + (\eta - 1) \log_2 e \tag{5.12}$$

where

- $\boldsymbol{\beta} \triangleq (\beta_1, \ldots, \beta_L)$ and $\boldsymbol{\gamma} \triangleq (\gamma_1, \ldots, \gamma_L)$.
- η is the large-system (non-asymptotic) multiuser efficiency [29] of the linear MMSE receiver, given by the positive solution of the Tse-Hanly equation [91]

$$\eta = \left(1 + \sum_{j=1}^{L} \beta_j \frac{\gamma_j}{1 + \eta\gamma_j}\right)^{-1} \tag{5.13}$$

  For later use, (5.13) is written as

$$\eta = f_L(\eta, \beta_L) \tag{5.14}$$

  where we define

$$f_j(\eta, z) \triangleq \left(1 + z\frac{\gamma_j}{1 + \gamma_j\eta} + \sum_{i=1}^{j-1} \beta_i \frac{\gamma_i}{1 + \gamma_i\eta}\right)^{-1} \tag{5.15}$$

- $\mathsf{C}^{\text{mmse}}(\boldsymbol{\beta}, \boldsymbol{\gamma})$ is the achievable spectral efficiency if the decoder is assumed to consist of a bank of MMSE linear filters followed by single-user decoders. $\mathsf{C}^{\text{mmse}}(\boldsymbol{\beta}, \boldsymbol{\gamma})$ is given by

$$\mathsf{C}^{\text{mmse}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{j=1}^{L} \beta_j \log_2(1 + \gamma_j\eta) \tag{5.16}$$

The spectral efficiencies in (5.12) and in (5.16) are achieved with codes whose empirical distributions are Gaussian.

As shown in [32], the supremum of (5.12) over all possible $L, \boldsymbol{\beta}, \boldsymbol{\gamma}$ (for fixed E_b/N_0 and β) is achieved by $L = 1$ (one class only).

This can be readily seen by writing the total spectral efficiency for finite K, N and given $\mathbf{S}$ as [82]

$$\mathsf{C}^{\text{fin}}(K, N, \mathbf{S}) = \frac{1}{N} \log_2 \det(\mathbf{I} + \ \mathsf{SNR}\ \mathbf{S}\mathsf{A}^2\mathbf{S}^H) \tag{5.17}$$

where $\mathsf{A} \triangleq \text{diag}(\sqrt{\alpha_1}, \ldots, \sqrt{\alpha_K})$. Then, we notice that, for $K/N = \beta$ and assuming that the empirical distribution of the scaling factors α_k converges to some fixed (non-random) distribution as $K \to \infty$, the limit

$$\lim_{K\to\infty} \frac{1}{N} \log_2 \det(\mathbf{I} + \ \mathsf{SNR}\ \mathbf{S}\mathsf{A}^2\mathbf{S}^H) = \lim_{K\to\infty} \frac{1}{N} E\left[\log_2 \det(\mathbf{I} + \ \mathsf{SNR}\ \mathbf{S}\mathsf{A}^2\mathbf{S}^H)\right]$$

holds with probability 1 [32]. Finally, by averaging over all $K \times K$ permutation matrices $\mathbf{\Pi}$, by noticing that $\mathbf{S}$ and $\mathbf{S\Pi}$ are identically distributed and by using Jensen's inequality, it follows from the concavity of $\log\det(\cdot)$ on the cone of non-negative definite Hermitian symmetric matrices that

$$\begin{aligned} E\left[\log_2\det\left(\mathbf{I}+\ \mathsf{SNR}\,\mathbf{S}\mathsf{A}^2\mathbf{S}^H\right)\right] &= \frac{1}{K!}\sum_{\mathbf{\Pi}} E\left[\log_2\det\left(\mathbf{I}+\ \mathsf{SNR}\,\mathbf{S\Pi}\mathsf{A}^2\mathbf{\Pi}^H\mathbf{S}^H\right)\right] \\ &\leq E\left[\log_2\det\left(\mathbf{I}+\frac{\mathsf{SNR}}{K!}\sum_{\mathbf{\Pi}}\mathbf{S\Pi}\mathsf{A}^2\mathbf{\Pi}^H\mathbf{S}^H\right)\right] \end{aligned}$$

Then, noting that $\displaystyle\sum_{\mathbf{\Pi}}\mathsf{A}^2\mathbf{\Pi}^H = (K-1)!\,\mathrm{diag}\left(\sum_{k=1}^{K}\alpha_k\right) = K!\,\mathbf{I}$, then we get

$$E\left[\log_2\det\left(\mathbf{I}+\ \mathsf{SNR}\,\mathbf{S}\mathsf{A}^2\mathbf{S}^H\right)\right] = E\left[\log_2\det\left(\mathbf{I}+\ \mathsf{SNR}\,\mathbf{S}\mathbf{S}^H\right)\right]$$

where the last line is achieved when all users are received with the same SNR.

The decoding configuration consisting of MMSE filtering followed by single-user decoders (without IC) is clearly sub-optimal. Indeed, the supremum of the spectral efficiency $\mathsf{C}^{\mathrm{mmse}}$ is achieved for $\beta_{\mathrm{opt}}^{\mathrm{mmse}} < \infty$ with

$$\mathsf{C}^{\mathrm{mmse}}(\beta_{\mathrm{opt}}^{\mathrm{mmse}},\gamma) < \mathsf{C}^{\star}$$

where $\mathsf{C}^{\star}$ is the AWGNC capacity at the corresponding $\frac{E_b}{N_0}$ (see Fig. 5.3, Fig. 5.4 and Appendices 5.C and 5.D). Moreover, the asymptotic behavior of the MMSE receiver

$$\lim_{\beta\to\infty}\mathsf{C}^{\mathrm{mmse}} = \log_2 e - \frac{N_0}{E_b}$$

which is bounded away from $\mathsf{C}^{\star}$.

The supremum of C over β is achieved for $\beta\to\infty$, and coincides with the AWGNC single-user capacity $\mathsf{C}^{\star}$ given by (5.11). The spectral efficiency $\mathsf{C}(\boldsymbol{\beta},\boldsymbol{\gamma})$ can be achieved by single-user decoding with successive stripping and MMSE filtering against not yet decoded users. Suppose that users are decoded one by one, starting from users in class L, then class $L-1$ and so on. Then, $\mathsf{C}(\boldsymbol{\beta},\boldsymbol{\gamma})$ can be written as

$$\mathsf{C}(\boldsymbol{\beta},\boldsymbol{\gamma}) = \sum_{j=1}^{L}\int_0^{\beta_j}\log_2(1+\gamma_j\eta_j(z))dz \tag{5.18}$$

where $\eta_j(z)$ is the solution to $\eta = f_j(\eta, z)$. Notice that stripping of the users one by one implies that users in the same class have different rates. Namely, the user decoded in position $\lfloor K_j z/\beta_j \rfloor$ of class j (where $z \in [0, \beta_j]$), transmits at rate $\log_2(1+\gamma_j \eta_j(z))$. Hence in general a different rate is required for each user. This makes such a system highly impractical from the implementation point of view.

5.4 Approaching the Optimal Spectral Efficiency with QPSK

Information theory teaches us that one way to approach (5.12) is to use single-user capacity approaching codes for the AWGNC, successive interference cancellation and MMSE filtering at each cancellation stage [83]. Furthermore, substantial progress has been made in the last few years in designing binary codes and low-complexity decoders whose rate comes fairly close to single-user capacity at vanishing BER. Among those modern codes are Turbo codes, IRA codes, and LDPC, all of which are decoded by efficient iterative techniques (see the special issue [92] and references therein, as well as [93] and references therein). These code ensembles are characterized by their rate-threshold pair (R, g), such that for $\mathsf{SNR} \geq g$ the BER can be made arbitrarily small in the limit of code block length $n \to \infty$. For carefully optimized code ensembles [39, 19, 22, 43], on the standard single-user additive white Gaussian noise channel, the rate-threshold pairs achieved so far come remarkably close to the curve $R = C_{\text{qpsk}}(\mathsf{SNR})$, where

$$\begin{aligned} C_{\text{qpsk}}(\mathsf{SNR}) &= 2\left(1 - \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \log_2\left(1 + e^{-2\,\mathsf{SNR} - 2\sqrt{2\,\mathsf{SNR}}v}\right) e^{-v^2} dv\right) \\ &= 2J(2\,\mathsf{SNR}) \end{aligned} \tag{5.19}$$

is the QPSK-input AWGNC capacity, as a function of SNR, and $J(\mu)$ is given by (3.15). Fig. 5.2 shows the QPSK capacity (5.19) and rate-threshold pairs corresponding to some LDPC code ensembles from [9].

In the large system limit, under our system assumptions, it is well-known that the residual interference at the output of the MMSE filter at any cancellation stage is a circular symmetric complex Gaussian random variable [94]. Assuming optimal QPSK codes characterized by the rate-threshold pairs $(R, C_{\text{qpsk}}^{-1}(R))$, for $R \in [0, 2]$, (see Fig. 5.2), the spectral efficiency achieved

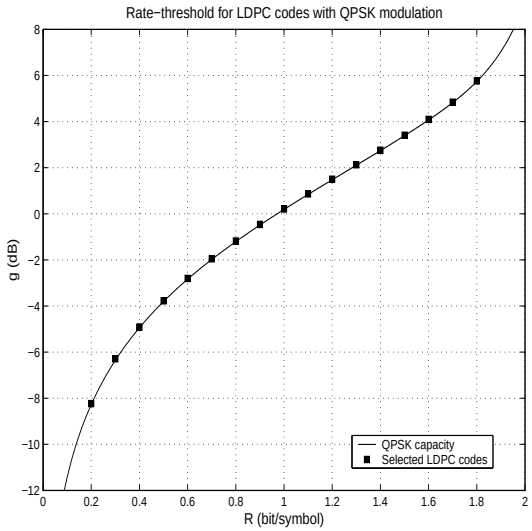


Figure 5.2: Rate-threshold pairs corresponding to QPSK capacity and for some optimized LDPC codes

by a stripping decoder is given by

$$\mathsf{C}_{\mathrm{qpsk}}(\boldsymbol{\beta},\boldsymbol{\gamma}) = \sum_{j=1}^{L} \int_0^{\beta_j} C_{\mathrm{qpsk}}(\gamma_j \eta_j(z))dz \tag{5.20}$$

Fig. 5.3 and Fig. 5.4 show $\mathsf{C}_{\mathrm{qpsk}}(\beta,\gamma)$, $\mathsf{C}^{\mathrm{mmse}}(\beta,\gamma)$ and $\mathsf{C}(\beta,\gamma)$ (for a single-class system, i.e., $L=1$) vs. β, for $E_b/N_0 = 3$ dB and 10 dB, respectively. The corresponding AWGNC capacity $\mathsf{C}^{\star}$ is shown for comparison. We notice that the loss incurred by QPSK codes with respect to Gaussian codes gets more pronounced as E_b/N_0 increases. For any fixed E_b/N_0 and sufficiently large β, the loss vanishes. However, exceedingly large values of β are required to make the loss negligible, for high E_b/N_0.

The following result shows that as the system load grows without bound, QPSK suffers no loss of optimality. The result applies to the general case where the received user SNRs are given by $\mathsf{SNR}|\mathsf{A}_k|^2$, under the mild requirement that, as $K \to \infty$, the empirical distribution of the user amplitudes $|\mathsf{A}_k|$ converges to a given non-random distribution $F_{|\mathsf{A}|}$. For $|\mathsf{A}_k|^2 = \alpha_k$ (deterministic) we obtain the system model defined in Section 5.1, but the result applies also to the case where $|\mathsf{A}_k|^2$ represents the random fading coefficient of user k, as in [32]. As in all capacity results of CDMA in the large system

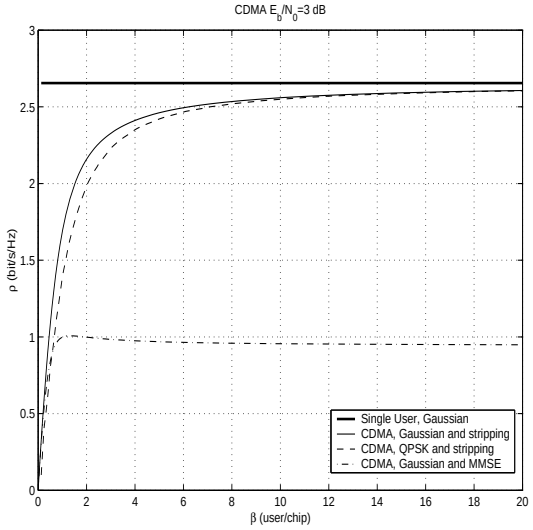


Figure 5.3: Spectral efficiency vs. β for random CDMA, $E_b/N_0 = 3dB$, with Gaussian inputs (stripping decoder vs. MMSE decoder) and QPSK inputs (with stripping decoder)

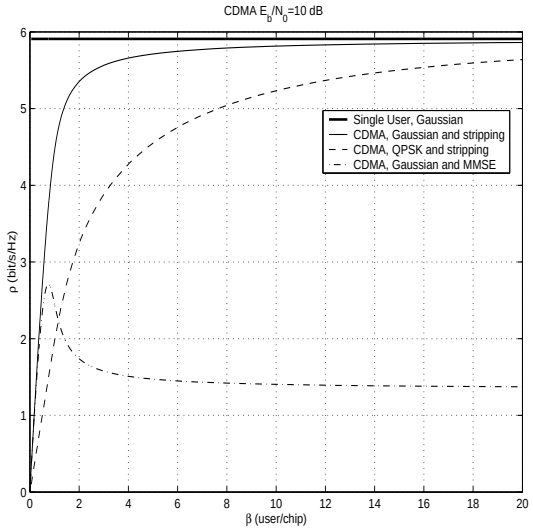


Figure 5.4: Spectral efficiency vs. β for random CDMA, $E_b/N_0 = 10dB$, with Gaussian inputs (stripping decoder vs. MMSE decoder) and QPSK inputs (with stripping decoder)

limit (see for example [31, 32, 91]), achievability implies that the user code block length goes to infinity. The natural and meaningful order of the limits with respect to n and K is: *first* let $n \to \infty$ and *then* let $K \to \infty$. This means that the large-system spectral efficiency is the limit of the spectral efficiencies of systems with increasing K.

Theorem 5.1 *Let*

$$C_Q(\beta,\ \mathsf{SNR}) = \int_0^\beta E[C_{\text{qpsk}}(|\mathsf{A}|^2\ \mathsf{SNR}\,\eta(z,\ \mathsf{SNR}))]dz \tag{5.21}$$

where $\eta(z,\ \mathsf{SNR})$ *is the solution to*

$$\eta + zE\left[\frac{\mathsf{SNR}|\mathsf{A}|^2\eta}{1+\ \mathsf{SNR}|\mathsf{A}|^2\eta}\right] = 1 \tag{5.22}$$

where $|\mathsf{A}| \sim F_{|\mathsf{A}|}$.

Fix β *and* $\frac{E_b}{N_0}$ *and define*

$$\mathsf{C}_Q(\beta, \frac{E_b}{N_0}) = C_Q(\beta,\ \mathsf{SNR}) \tag{5.23}$$

for the SNR *satisfying*

$$\frac{E_b}{N_0} C_Q(\beta,\ \mathsf{SNR}) = \beta\ \mathsf{SNR} \tag{5.24}$$

Then, for all $\frac{E_b}{N_0} \geq \log_e 2$,

$$\lim_{\beta\to\infty} \mathsf{C}_Q(\beta, \frac{E_b}{N_0}) = \mathsf{C}^\star \tag{5.25}$$

Proof: See Appendix 5.A.

Interestingly, the optimality of QPSK in the large β limit proved in theorem 5.1 is different in nature from its wideband optimality proved in [90]. In fact, as a consequence of the rotational invariance of the spreading sequences, theorem 5.1 also holds if the modulation is BPSK. A related result on the optimality of binary inputs in the absence of spreading admits a different proof based on the central-limit theorem [82].

5.5 Conclusion

We have presented the spectral efficiency achievable by a synchronous random CDMA system in the large system limit (infinite number of users, spreading gain and code block length) with finite channel load. In the absence of spreading, the spectral efficiency of the system coincides with the single-user capacity of the AWGNC. The optimum spectral efficiency (with optimum decoder) is achieved by stripping, based on MMSE filtering against not yet demodulated users and successive single-user decoding. We have shown that the use of (non-Gaussian) QPSK inputs incurs no loss on the maximum spectral efficiency achieved for $\beta \to \infty$, when using binary capacity-achieving codes and the stripping decoder. However, for a fixed channel load, the discrepancy between the QPSK-encoded CDMA spectral efficiency and the maximum spectral efficiency gets more important as the bit energy to noise ratio $\frac{E_b}{N_0}$ increases. In view of these observations, it is possible to design a *pragmatic* CDMA system, which can achieve near-optimum spectral efficiency. This system would have low-complexity encoding based on binary error codes and QPSK modulation, as well as low-complexity decoding based on successive stripping, linear MMSE filters and single-user decoding.

APPENDIX

5.A Proof of Theorem 5.1

In view of the result shown in [32, Eq. 163] and since the use of QPSK cannot improve upon the result obtained with Gaussian inputs with the same power, it is enough to show that

$$\lim_{\beta\to\infty} \mathsf{C}_Q(\beta, \frac{E_b}{N_0}) \geq \mathsf{C}^\star \tag{5.26}$$

where $\mathsf{C}^\star$ is given by (5.11) for any $\frac{E_b}{N_0} > \log_e 2$ (notice that $\mathsf{C}^\star(\log_e 2) = 0$).

To show (5.26), we will show that for every β, if $\mathsf{SNR} \to 0$ then

$$C_Q(\beta,\ \mathsf{SNR}) \geq \log_2(1+\beta\ \mathsf{SNR}) - \frac{\kappa(|\mathsf{A}|)\beta\ \mathsf{SNR}^2}{\beta\ \mathsf{SNR}+1}\ \log_2 e, \tag{5.27}$$

where

$$\kappa(|\mathsf{A}|) = \frac{E[|\mathsf{A}|^4]}{(E[|\mathsf{A}|^2])^2}$$

denotes the kurtosis of the distribution of $|\mathsf{A}|$. The bound in (5.27) will be sufficient for our purposes because if we choose the following signal-to-noise ratio

$$\mathsf{SNR}_\beta = \frac{\mathsf{C}^\star}{\beta}\frac{E_b}{N_0} \tag{5.28}$$

then,

$$\begin{aligned} C_Q(\beta,\ \mathsf{SNR}_\beta) &\geq \log_2(1+\beta\ \mathsf{SNR}_\beta) - \kappa(|\mathsf{A}|)\ \mathsf{SNR}_\beta \frac{\frac{E_b}{N_0}\mathsf{C}^\star}{\frac{E_b}{N_0}\mathsf{C}^\star+1}\ \log_2 e \\ &\to \log_2(1+\tfrac{E_b}{N_0}\mathsf{C}^\star) \\ &= \mathsf{C}^\star \end{aligned} \tag{5.29}$$

Furthermore the $\frac{E_b}{N_0}$ required by SNR_β is upper bounded by

$$\frac{\beta\ \mathsf{SNR}_\beta}{C_Q(\beta,\ \mathsf{SNR}_\beta)} \leq \frac{\frac{E_b}{N_0}\mathsf{C}^\star}{\log_2(1+\frac{E_b}{N_0}\mathsf{C}^\star) - \epsilon\mathsf{C}^\star} \tag{5.30}$$

$$= \frac{E_b}{N_0}\frac{1}{1-\epsilon} \tag{5.31}$$

for an arbitrarily small ϵ, provided β is large enough.

To show (5.27) we need the following two inequalities:

$$C_{\text{qpsk}}(x) \geq \left(x - x^2\right) \log_2 e \tag{5.32}$$

$$\eta \geq \frac{1}{1 + \beta\ \mathsf{SNR}} \tag{5.33}$$

where η is the solution to

$$\eta + \beta E\left[\frac{\mathsf{SNR}|\mathsf{A}|^2\eta}{1 +\ \mathsf{SNR}|\mathsf{A}|^2\eta}\right] = 1 \tag{5.34}$$

To show (5.33) rewrite the Tse-Hanly equation as

$$\mathsf{SNR} = \eta\ \mathsf{SNR} + \beta\ \mathsf{SNR}^2\eta E\left[\frac{|\mathsf{A}|^2}{1 +\ \mathsf{SNR}|\mathsf{A}|^2\eta}\right] \tag{5.35}$$

$$\leq \eta\ \mathsf{SNR} + \beta\ \mathsf{SNR}^2\eta \tag{5.36}$$

where the inequality comes from $E[|\mathsf{A}|^2] = 1$. Relation (5.32) is demonstrated in Appendix 5.B.

Now, (5.27) readily follows from

$$C_Q(\beta,\ \mathsf{SNR}) = E[C_{\text{qpsk}}(|\mathsf{A}|^2\ \ \mathsf{SNR}\,\eta(z,\ \mathsf{SNR}))]dz \tag{5.37}$$

$$\geq \int_0^\beta E\left[C_{\text{qpsk}}\left(\frac{|\mathsf{A}|^2\ \ \mathsf{SNR}}{1 + z\ \ \mathsf{SNR}}\right)\right] dz \tag{5.38}$$

$$\geq \int_0^\beta E\left[\frac{|\mathsf{A}|^2\ \ \mathsf{SNR}}{1 + z\ \ \mathsf{SNR}}\right] \log_2 e - E\left[\frac{|\mathsf{A}|^4\ \ \mathsf{SNR}^2}{(1 + z\ \ \mathsf{SNR})^2}\right] \log_2 e\, dz \tag{5.39}$$

$$\geq \int_0^\beta \frac{\mathsf{SNR}}{1 + z\ \mathsf{SNR}} \log_2 e - \frac{\kappa(|\mathsf{A}|)\ \mathsf{SNR}^2}{(1 + z\ \ \mathsf{SNR})^2} \log_2 e\, dz \tag{5.40}$$

$$= \log_2(1 + \beta\ \ \mathsf{SNR}) - \frac{\kappa(|\mathsf{A}|)\beta\ \mathsf{SNR}^2}{1 + \beta\ \ \mathsf{SNR}} \log_2 e \tag{5.41}$$

thus concluding the proof. □

5.B Proof of Relation (5.32)

We want to show

$$C_{\text{qpsk}}(x) \geq g(x) \ \text{ for } \ x \to 0$$

where $g(x) \triangleq (x - x^2)\log_2 e$.

Define

$$f(x,\alpha) \triangleq 2 - \frac{2}{\sqrt{2\pi}}\log_2 e \int_{-\alpha}^{\alpha} \log(1+e^{u(x,z)})e^{-z^2/2}dz \tag{5.42}$$

where $u(x,z) = -2x - 2\sqrt{x}z$. Then, because of the continuity of $f(x,\alpha)$ in x and α, we have

$$\lim_{x\to 0} C_{\text{qpsk}} = \lim_{\alpha\to+\infty}\lim_{x\to 0} f(x,\alpha)$$

Expanding the integrand of $f(x,\alpha)$ about $x_0 = 0$ results in

$$\log(1+e^{u(x,z)}) = \log 2 - x + \frac{x^2}{2} + \sqrt{x}(x-1)z + \frac{x}{2}z^2 - \frac{x^2}{12}z^4 + o(x^{5/2}) \tag{5.43}$$

Hence, the first and second derivatives of $\mathsf{C}_{\text{qpsk}}(x)$ at $x = 0$ are given by

$$\begin{aligned}\left.\frac{dC_{\text{qpsk}}}{dx}\right|_{x=0} &= \lim_{\alpha\to+\infty}\lim_{x\to 0}\frac{\delta f(x,\alpha)}{\delta x}\\ &= \frac{2\log_2 e}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(1-\frac{z^2}{2}\right)e^{-z^2/2}dz\\ &= \log_2 e\end{aligned} \tag{5.44}$$

$$\begin{aligned}\left.\frac{d^2C_{\text{qpsk}}}{dx^2}\right|_{x=0} &= \lim_{\alpha\to+\infty}\lim_{x\to 0}\frac{\delta^2 f(x,\alpha)}{\delta x^2}\\ &= \frac{2\log_2 e}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left(\frac{z^4}{6}-1\right)e^{-z^2/2}dz\\ &= -\log_2 e\end{aligned} \tag{5.45}$$

Since $g'(x) = \log_2 e$ and $g''(x) = -2\log_2 e$, then there exists $\epsilon > 0$ such that

$$\forall x \in (0,\epsilon) \quad C_{\text{qpsk}}(x) \geq \left(x - x^2\right)\log_2 e$$

□

5.C Gaussian Input and MMSE Decoder

For a single class of users with SNR γ and channel load β, the spectral efficiency (5.16) becomes

$$\mathsf{C}^{\text{mmse}} = \beta\log_2(1+S) \tag{5.46}$$

where S is the positive solution of the fixed point equation

$$S = \frac{\gamma}{1 + \frac{\beta\gamma}{1+S}} \tag{5.47}$$

For a given E_b/N_0, the SNR γ is

$$\gamma = \frac{E_b}{N_0} \frac{\mathsf{C}^{\text{mmse}}}{\beta} \tag{5.48}$$

Using (5.47) and (5.48), we obtain the following parametric system, giving C^{mmse} as a function of β through the parameter $A \in [0, \infty]$:

$$\mathsf{C}^{\text{mmse}} = (S+1)\left(\frac{\log_2(1+S)}{S} - \frac{N_0}{E_b}\right) \tag{5.49}$$

$$\beta = \frac{E_b}{N_0} \frac{(S+1)\mathsf{C}^{\text{mmse}}}{S(S+1+\frac{E_b}{N_0}\mathsf{C}^{\text{mmse}})} \tag{5.50}$$

Differentiating the RHS of (5.50) w.r.t. S and equating the derivative to 0, the maximum spectral efficiency corresponds to

$$\frac{E_b}{N_0} = \frac{S^2}{S\log_2(e) - log_2(1+S)} \tag{5.51}$$

for $\frac{E_b}{N_0} \geq 2\log 2$. If $\log 2 \leq \frac{E_b}{N_0} < 2\log 2$, then the derivative is negative, and the maximum spectral efficiency corresponds to $S = 0$. Hence, the maximum spectral efficiency with MMSE followed by single-user decoding is given as

$$\mathsf{C}^{\text{mmse}} = \begin{cases} \log_2 e - \frac{N_0}{E_b} & \text{for } \log 2 \leq \frac{E_b}{N_0} < 2\log 2 \\ (1+S)\left(\frac{\log_2(1+S)}{S} - \frac{N_0}{E_b}\right) & \text{for } \frac{E_b}{N_0} \geq 2\log 2 \end{cases} \tag{5.52}$$

with S given by (5.51).

5.D Gaussian Input and Stripping Decoder

For a single class of users with SNR γ and channel load β, the spectral efficiency (5.12) becomes

$$\mathsf{C} = \beta\log_2(1+S) + \log_2\frac{\gamma}{S} + \left(\frac{S}{\gamma} - 1\right)\log_2 e \tag{5.53}$$

where S is given by the positive solution of (5.47). Following the same steps as in the previous section, for a fixed $\frac{E_b}{N_0}$, the spectral efficiency C is given as a function of β by the following parametric system:

$$f_1(S, \mathsf{C}, \frac{E_b}{N_0}) = f_2(S, \frac{E_b}{N_0}) \tag{5.54}$$

$$\beta = \frac{E_b}{N_0} \frac{(S+1)\mathsf{C}}{S(S+1+\frac{E_b}{N_0}\mathsf{C})} \tag{5.55}$$

where

$$f_1(S, \mathsf{C}, \tfrac{E_b}{N_0}) \triangleq \frac{S+1+\frac{E_b}{N_0}\mathsf{C}}{\mathsf{C}} \log_2 \frac{S+1+\frac{E_b}{N_0}\mathsf{C}}{S+1} - \frac{E_b}{N_0}\mathsf{C}$$

$$f_2(S, \tfrac{E_b}{N_0}) \triangleq S + 1 + \tfrac{E_b}{N_0} \left(\log_2 e - \tfrac{(S+1)\log_2(S+1)}{S} \right)$$

with $S \in [0, S_{\max}]$ and $S_{\max}$ satisfies

$$\frac{S_{max}}{\log_2(1+S_{\max})} = \frac{E_b}{N_0}$$

Chapter 6

Approaching the Optimum with Low Complexity

We consider a pragmatic approach to QPSK-modulated CDMA based on using single-user binary codes and the same stripping decoding approach which would be optimal for Gaussian codes. For this setting, we optimize the spectral efficiency in the large system regime with single-user capacity-achieving binary codes and with the best known LDPC code ensembles [9], in the limit for large code block length, in the cases of equal received SNRs or equal rate users. The proposed equal-power and equal-rate design approaches can be effectively applied to non-asymptotic code block length, and provide a simple tool to dimension CDMA systems for given target BER, user codes, and desired spectral efficiency.

6.1 Optimization of Spectral Efficiency

In the following, we will optimize a CDMA system, in the limit of infinite number of users K and spreading gain N, under successive decoding, assuming error-free decoding at each decoding level when the decoder operates above its threshold SNR. This corresponds to assuming very large code block length, i.e., $n \to \infty$, and using single-user capacity-achieving binary codes.

As in Section 5.4, each code ensemble is characterized by a rate-threshold pair (R, g), such that if SNR $\geq g$, the bit error rate vanishes in the limit of infinite block length. We seek to find the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ so that, at each stripping decoder stage, the threshold requirement of each single-user decoder is satisfied. We refer to this condition as the *successive decodability condition*.

We consider two alternative pragmatic CDMA optimization problems:

Equal-rate system All classes have the same rate R but have different SNR levels γ_i, such that, without loss of generality, $\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_L$.

Equal-power system All classes have the same SNR level γ but each class makes use of a different code ensemble of rate R_i, such that, without loss of generality, $R_1 \geq R_2 \geq \cdots \geq R_L$.

For practical reasons, we constrain our system to have only a finite number of coding rates and received SNR levels. In both configurations, we assume that users in each class $1 \leq i \leq L$ are decoded in parallel by a bank of single-user decoders, while classes are stripped off from L to 1, i.e., in decreasing SNR order for the equal-rate case, or in increasing rate order for the equal-power case (see Fig. 6.1). Notice that our approach is pragmatic in three ways:

- it makes use of QPSK input constellation rather than Gaussian codes;
- it makes use of low-complexity iterative decoding;
- and it performs class-by-class stripping, rather than user-by-user, as implied by expressions (5.18) and (5.20).

6.1.1 Optimization for Equal-Rate Systems

We assume that all users in all classes make use of codes drawn randomly and independently from the same ensemble with rate-threshold pair (R, g). The SINR at the output of the MMSE filter for class i users, assuming that all users in classes $i+1, \ldots, L$ have been perfectly canceled, is given by $\gamma_i \eta_i(\beta_i)$. Hence, the successive decodability condition of all the users in class i, for $1 \leq i \leq L$, is

$$\eta_i(\beta_i) \geq \frac{g}{\gamma_i} \tag{6.1}$$

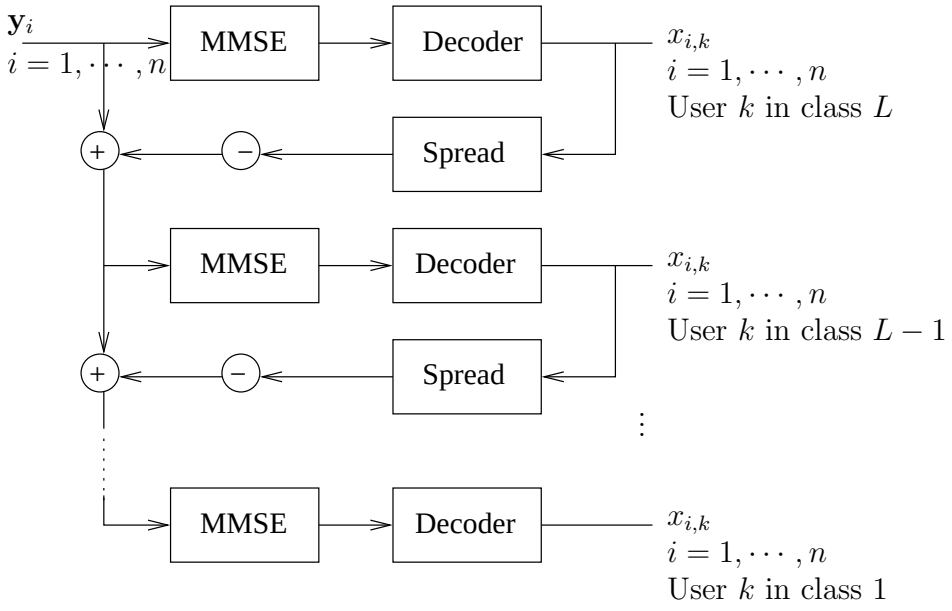


Figure 6.1: Successive decoding class by class in descending order of powers or ascending order of rates

where $\eta_i(z)$ is the solution to $\eta = f_i(\eta, z)$. $f_i(\eta, z)$ was introduced in Section 5.3, and we recall it here for the sake of clarity:

$$f_i(\eta, z) = \left(1 + z\frac{\gamma_i}{1+\gamma_i\eta} + \sum_{j=1}^{i-1}\beta_j\frac{\gamma_j}{1+\gamma_j\eta}\right)^{-1} \tag{6.2}$$

We fix the received power levels $\boldsymbol{\gamma}$, and consider the optimization of the class loads $\boldsymbol{\beta}$. Without loss of generality, we assume $\gamma_1 \geq g$. Indeed, if $\gamma_j < g$ for a given j, then we would get $\beta_j = 0$. The spectral efficiency optimization problem of this setting can be formulated as a linear program as follows. Because of the monotonicity of $f_i(\eta, z)$ in η, and since the solution to $\eta_i(z) = f_i(\eta_i(z), z)$ is unique, we conclude that [91]

$$\forall\ x \in [0, \infty) \qquad x \leq \eta_i(z) \Leftrightarrow x \leq f_i(x, z) \tag{6.3}$$

Accordingly, the successive decodability condition is equivalent to

$$\left(1 + \sum_{j=1}^{i}\beta_j\frac{\gamma_j}{1+\gamma_j\frac{g}{\gamma_i}}\right)^{-1} \geq \frac{g}{\gamma_i}, \quad \forall\ i = 1, \ldots, L \tag{6.4}$$

which can be written in compact form as

$$\mathbf{A}\boldsymbol{\beta} \leq \mathbf{b}$$

where $\mathbf{A}$ is a $L \times L$ lower triangular matrix with non-zero elements

$$a_{i,j} = \frac{(1+g)\gamma_j}{\gamma_i + \gamma_j g} \in (0, 1] \tag{6.5}$$

and $\mathbf{b}$ is a positive vector with elements

$$b_i = \frac{(1+g)(\gamma_i - g)}{\gamma_i g} \tag{6.6}$$

Notice that $a_{i,i} = 1$, $a_{i,j}$ (for $1 \leq j \leq i$) is increasing with j and decreasing with i and b_i is increasing with i.

For a desired spectral efficiency $\rho = \beta R$, the optimal vector $\boldsymbol{\beta}$ which achieves (if possible) arbitrarily small BER with minimal $(E_b/N_0)_{\text{sys}}$ is the solution to the linear program:

$$\begin{cases} \text{minimize} & \displaystyle\sum_{i=1}^{L} \beta_i \gamma_i \\ \text{subject to} & \mathbf{A}\boldsymbol{\beta} \leq \mathbf{b} \\ & \displaystyle\sum_{i=1}^{L} \beta_i \geq \beta \\ & \boldsymbol{\beta} \geq \mathbf{0} \end{cases} \tag{6.7}$$

given by the following result:

Proposition 6.1 *The equation* $\mathbf{Ax} = \mathbf{b}$ *has a unique solution with nonnegative elements* $\boldsymbol{\tau}$*. Furthermore, the feasible set in (6.7) is nonempty if and only if* $\beta \leq \displaystyle\sum_{j=1}^{L} \tau_j$*. The solution of (6.7) is given explicitly by*

$$\beta_i^\star = \begin{cases} \tau_i, & i = 1, \ldots, \hat{L} - 1 \\ \beta - \displaystyle\sum_{j=1}^{\hat{L}-1} \tau_j, & i = \hat{L} \\ 0, & i = \hat{L} + 1, \ldots, L \end{cases} \tag{6.8}$$

where $\hat{L}$ *denotes the minimum* i *for which* $\beta \leq \sum_{j=1}^{i} \tau_j$.

Proof: See Appendix 6.A.

6.1.2 Optimization for Equal-Power Systems

We assume that all users in all classes have the same fixed SNR level γ, but users in each class j make use of a different code ensemble, characterized by the rate-threshold pair (R_j, g_j). Since any good family of codes satisfies the condition that codes with larger rate have larger SNR thresholds, then $g_1 \geq \cdots \geq g_L$. Assuming optimal QPSK codes corresponds to obtaining the pairs (R_j, g_j) by sampling the curve of Fig. 5.2 in given L desired rate values. Without loss of generality, we assume $\gamma \geq g_1$, since for all j such that $\gamma < g_j$, we would have $\beta_j = 0$. Then, the successive decodability condition of all users in class i, for $i = 1, \cdots, L$, is given by

$$\eta_i(\beta_i) \geq \frac{g_i}{\gamma} \tag{6.9}$$

Again, using (6.3), the successive decodability condition(6.9) is equivalent to

$$\left(1 + \sum_{j=1}^{i} \beta_j \frac{\gamma}{1+g_i}\right)^{-1} \geq \frac{g_i}{\gamma} \tag{6.10}$$

and translates into

$$\sum_{j=1}^{i} \beta_j \leq b_i, \qquad i = 1, \ldots, L \tag{6.11}$$

where

$$b_i = \frac{(1+g_i)(\gamma - g_i)}{\gamma g_i} \tag{6.12}$$

Hence, for given rate-threshold pairs (R_j, g_j), the spectral efficiency $\rho = \sum_{i=1}^{L} \beta_i R_i$ maximized over the class loads is obtained as the solution to the

following linear program:

$$\begin{cases} \text{maximize} & \displaystyle\sum_{i=1}^{L} \beta_i R_i \\ \text{subject to} & \mathbf{L}\boldsymbol{\beta} \leq \mathbf{b} \\ & \displaystyle\sum_{i=1}^{L} \beta_i \leq \beta \\ & \boldsymbol{\beta} \geq \mathbf{0} \end{cases} \tag{6.13}$$

where $\mathbf{L}$ is a lower triangular $L \times L$ matrix with $l_{ij} = 1$ for all $i \geq j$ and $\mathbf{b} = (b_1, \ldots, b_L)^T$ with b_i given in (6.12). The solution of the linear program (6.13) is given by the following result:

Proposition 6.2 *The problem (6.13) is always feasible and its solution is given explicitly by*

$$\beta_i^\star = \begin{cases} b_i - b_{i-1}, & i = 1, \ldots, \hat{L} - 1 \\ \beta - b_{\hat{L}-1}, & i = \hat{L} \\ 0, & i = \hat{L} + 1, \ldots, L \end{cases} \tag{6.14}$$

where $b_0 \triangleq 0$*, and* $\hat{L}$ *denotes the minimum* i *for which* $\beta \leq b_i$*.*

Proof: See Appendix 6.B.

6.2 Numerical Examples

In this section we give examples of the equal-rate and the equal-power system designs using the off-the-shelf LDPC code ensembles optimized for the binary-input AWGN channel, found in [9], whose rate-threshold pairs correspond to the marks in Fig. 5.2.

6.2.1 Equal-Rate Design

In Fig. 6.2, the curves labeled "LDPC, R=0.2, 1.0, 1.96" represent the spectral efficiencies achieved by the equal-rate design with the LDPC codes of rate 0.2, 1 and 1.96 bit per QPSK symbol (corresponding to binary rate 0.1, 0.5 and 0.98), whose rate-threshold pairs are shown in Fig. 5.2. Likewise, Figs. 6.3 and 6.4 show the spectral efficiencies achieved by LDPC codes of

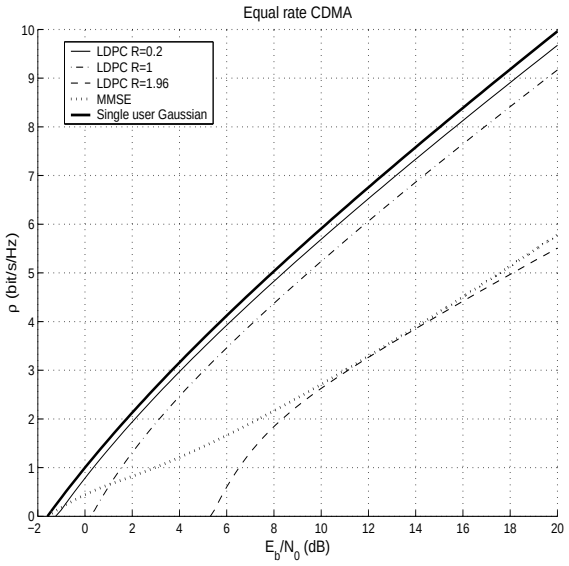


Figure 6.2: Spectral efficiency of some LDPC codes with equal-rate design

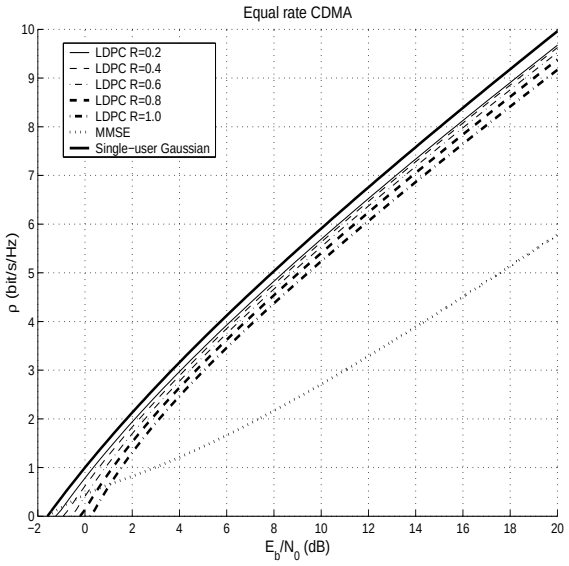


Figure 6.3: Spectral efficiency of LDPC codes with equal-rate design, for rates between 0.2 and 1 bit/channel use

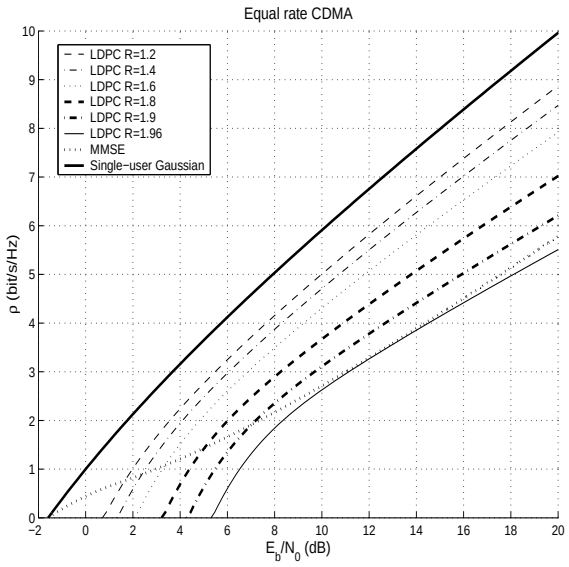


Figure 6.4: Spectral efficiency of high-rate LDPC codes with equal-rate design, for rates between 1.2 and 1.96 bit/channel use

QPSK rates in the intervals $[0.2, 1]$ and $[1.2, 1.96]$ respectively. The equal-rate spectral efficiency curves are obtained considering increasing values of β. For each β, a vector $\boldsymbol{\gamma}$ is obtained by evenly discretizing the interval $[(1+\epsilon)g, \bar{\gamma}(\beta)]$ with a step of 0.1 dB, where ϵ is a small positive real number (chosen here as 10^{-3}), and $\bar{\gamma}(\beta)$ is the minimum γ_L for which the feasible set of (6.7) is non-empty. The single-user capacity $\mathsf{C}^{\star}$, and the maximum spectral efficiency achievable with MMSE decoding and Gaussian input codes C^{mmse}, are shown for comparison. We notice that the equal-rate design is able to closely approach $\mathsf{C}^{\star}$ for low QPSK rate R, at the expense of a very large load β. For example, for $R = 0.2$ at $\rho = 2$ bit/s/Hz, the gap from $\mathsf{C}^{\star}$ is 0.35 dB, with load $\beta = 10$ users/chip. On the other hand, as the code rate increases, the gap to $\mathsf{C}^{\star}$ becomes more important. For example, for $R = 1.96$, the spectral efficiency achievable with the equal-rate design is even below the C^{mmse} curve. Note that the spectral efficiency is 0 for $E_b/N_0 = g/R$, which is the minimum E_b/N_0 required to have a vanishing fraction of users with non-zero rate codes.

Figs. 6.5, 6.6 and 6.7 show the distributions (mass functions) of β_j vs. γ_j corresponding to $\rho = 2$ and LDPC codes of rates $R = 0.2$, $R = 1$ (with 1.4 dB gap from $\mathsf{C}^{\star}$ at $\beta = 2$ users/chip) and $R = 1.96$ (with 6.6 dB gap

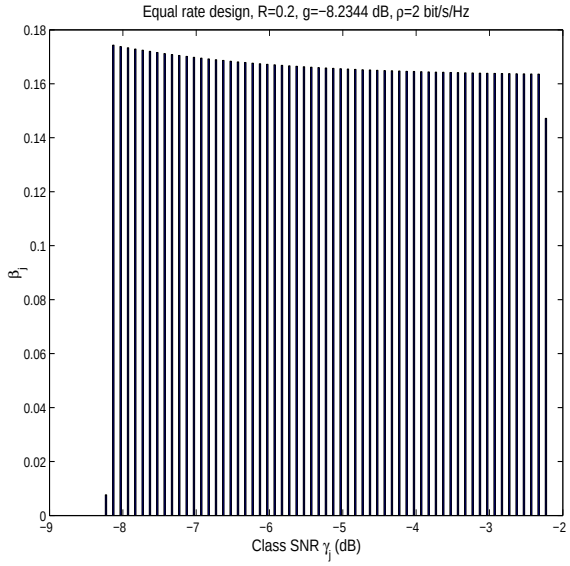


Figure 6.5: Load distribution ($\{\beta_j\}$ vs. $\{\gamma_j\}$) for the equal rate design with LDPC-coded QPSK of rate 0.2 bit/channel use and $\rho = 2$ bit/s/Hz

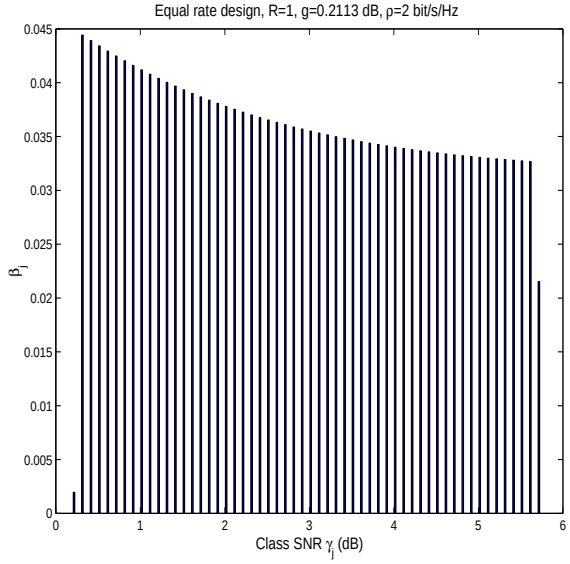


Figure 6.6: Load distribution ($\{\beta_j\}$ vs. $\{\gamma_j\}$) for the equal rate design with LDPC-coded QPSK of rate 1 bit/channel use and $\rho = 2$ bit/s/Hz

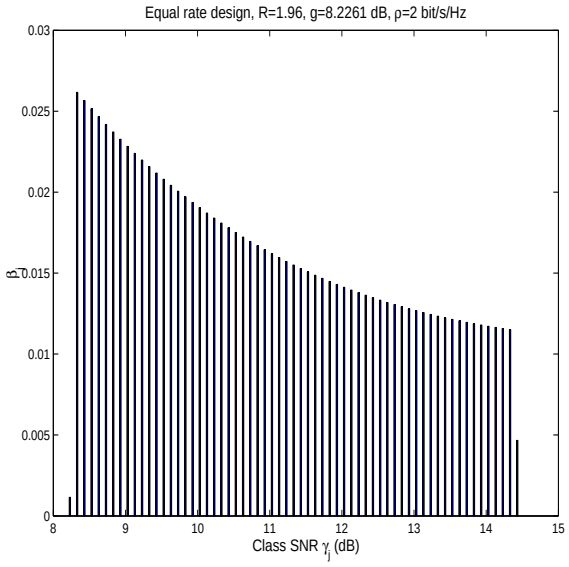


Figure 6.7: Load distribution ($\{\beta_j\}$ vs. $\{\gamma_j\}$) for the equal rate design with LDPC-coded QPSK of rate 1.96 bit/channel use and $\rho = 2$ bit/s/Hz

from $\mathsf{C}^\star$ at $\beta = 1.02$ users/chip). Note that the load distribution β_j is almost a continuous function of the SNR level γ_j, i.e., every class with SNR $\gamma_j \in [(1+\epsilon)g, \bar{\gamma}(\beta)]$ has a non-zero load β_j. This result is different from what is obtained in [95] for iterative IC with convolutional codes, where the load distribution is discrete. Then, one has to make a tradeoff between the gap to the single-user capacity $\mathsf{C}^\star$ and the number of distinct user classes.

6.2.2 Equal-Power Design

Fig. 6.8 shows the spectral efficiency achieved by the equal-power design with LDPC codes found in [9] and optimal QPSK codes. The single-user capacity $\mathsf{C}^\star$ is shown for comparison. The rate-threshold pairs of the curve labeled "LDPC" are those of LDPC code ensembles in [9] with (non-uniformly spaced) QPSK rates between 0.1 and 1.96 bits/s/Hz. The rate-threshold pairs of the curve labeled "discr.QPSK" are obtained by sampling the QPSK capacity curve in Fig. 5.2 from $R = 0.01$ to 1.99 bits/s/Hz with step 0.01. These equal-power spectral efficiency curves are obtained as the upper envelope of the solution of (6.13), over all $\gamma \geq g_L$ and $\beta \in [0, b_L]$, i.e., for all pairs (γ, β) for which the solution (6.14) exists.

The "LDPC" curve does not approach the "discr.QPSK" curve at high $\frac{E_b}{N_0}$ because the largest rate available in the LDPC code family of [9] is limited to 1.96 bits/channel use. It turns out that in order to approach $\mathsf{C}^\star$, it is necessary to have many different classes. Even a relatively finely discretized distribution of optimal (i.e., single-user capacity-achieving) rates, such as curve "discr.QPSK", suffers some loss away from $\mathsf{C}^\star$. The low $(E_b/N_0)_{\text{sys}}$ behavior of the spectral efficiency is dominated by the class with lowest coding rates, hence lowest SNR thresholds as well. In fact, the spectral efficiency is found to be 0 for $E_b/N_0 = g_L/R_L$, which is the minimum E_b/N_0 to have a vanishing fraction of users at non-zero rate. We conclude that in order to approach optimal spectral efficiency with the equal power scheme a wide range of coding rates is needed and, in particular, very high and very low rate codes must be designed. Therefore, similarly to the equal-rate design, one has to make a tradeoff between the gap of the spectral efficiency to $\mathsf{C}^\star$ and the number of distinct user classes.

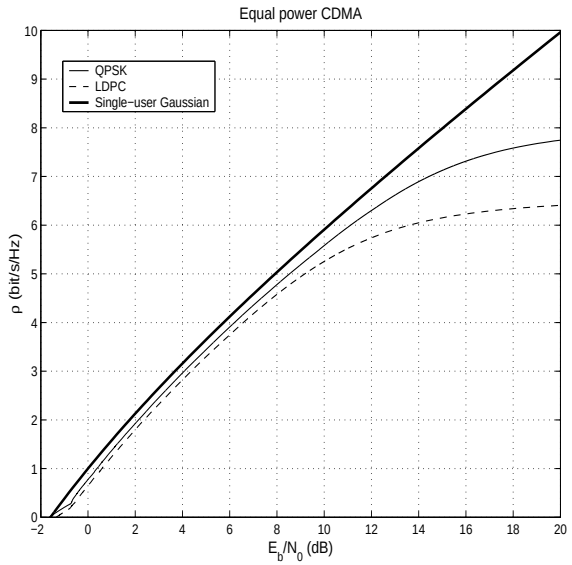


Figure 6.8: Spectral efficiency of LDPC and optimal QPSK codes with equal-power design

6.2.3 Effect of Finite n and K

We discuss the effect of finite block length n and finite K and N on the system achievable performance through an example. For finite code length, the decoding SNR threshold g above which the post-decoding BER is vanishing is not defined (strictly speaking, it is infinite). However, in practical system design, we set a target SNR threshold g_j for each class j, such that if the SINR at the input of the decoders of users in class j is above g_j, then the BER after decoding is so small that it has a negligible impact on the following decoding stages of classes $j-1, j-2, \ldots, 1$. A natural question is whether such a small but non-vanishing BER has a catastrophic effect, preventing the successive stripping decoder from decoding some class of users. We argue that for n sufficiently larger than K, and sufficiently small post-decoding BER, the effect of residual errors is indeed negligible. In fact, assuming random errors (if errors are correlated after decoding, we can use independent interleaving for each user), the expected number of incorrectly decoded symbols interfering with a user of class j is given by $\epsilon \sum_{i>j}^{L} K_i$, where ϵ is the residual BER. At the transition between the waterfall and the error flattening regions, LDPC codes have the property that the BER is $\epsilon = O(1/n) = \omega/n$ for some constant $\omega \ll n$. Turbo-codes have a similar property, where $\epsilon = O(1/I)$ and I is the size of the interleaver of the turbo encoder. This effect is called "interleaving gain" in [96]. Hence, by setting the thresholds g_j as the SNR at the transmission between waterfall and flattening, and by letting $n \gg K$, then the expected number of incorrectly decoded symbols interfering with a user of class j is given by $\omega(\sum_{i>j}^{L} K_i)/n \ll 1$.

Eventually, a *sensible* approach for the design of practical CDMA systems based on successive decoding is to use the large-system optimization methods developed before, while replacing the infinite block length thresholds g_j by some target SNR values chosen according to the BER vs. SNR performance of actual finite-length codes. While this argument provides only a heuristic design approach, extensive simulations show that the resulting systems are very good.

Moreover, in order to improve the robustness of the stripping decoder to residual errors, *soft-stripping* can be used, and successive decoding of the users from class L to class 1 can be iterated more than once. We refer to this approach as the multi-pass soft-stripping decoder, where one decoding pass consists of decoding all the users once. Soft-stripping (see for example [24]) consists of subtracting the minimum mean-square error (MMSE) estimates of

the signals of the already-decoded users from the signal, instead of their hard estimates. When the user decoders are symbol-by-symbol MAP (or approximations via Belief-Propagation, as in LDPC decoding [19, 39]), the MMSE estimate of the i-th symbol of user k, $x_{i,k}$ is obtained as the conditional mean

$$\widetilde{x}_{i,k} = E[x|\mathrm{EXT}_{i,k}]$$

where $\mathrm{EXT}_{i,k}$ denotes the k-th decoder *extrinsic information* for the i-th symbol. For example, with QPSK symbols we have[1]

$$\widetilde{x}_{i,k} = \frac{1}{\sqrt{2}} \tanh(m_{i,k}^{(I)}/2) + \frac{j}{\sqrt{2}} \tanh(m_{i,k}^{(Q)}/2) \tag{6.15}$$

where $m_{i,k}^{(I)}$ and $m_{i,k}^{(Q)}$ denote the extrinsic belief-propagation decoder messages for the variable nodes corresponding to the bits modulated in the in-phase and quadrature components of the i-th QPSK symbol of user k [24]. If decisions are not reliable, i.e., $|m_{i,k}^{(I)}|$ and/or $|m_{i,k}^{(Q)}|$ are small, then soft stripping attenuates the effect of this symbol on the residual interference signal.

Fig. 6.9 shows the average BER performance of irregular LDPC codes of binary rate 1/2 whose degree sequence is shown in [9]. The BER is obtained by averaging over randomly generated parity-check matrices with the given degree distributions with maximum left degree 100, average right degree 11. The belief-propagation decoder performs up to 200 iterations. Fig. 6.10 shows the spectral efficiency achievable by the equal-rate design with $R = 1.0$. The curve denoted by "optimal" corresponds to the threshold $g = 0.186$ dB, of ideal infinite-length QPSK random coding. The curve denoted by "practical" corresponds to the threshold $g = 0.933$ dB, of the finite-length irregular LDPC code ensemble generated from the degree distribution found in [9]. It is obtained by trial-and-error as the minimum value for which the multi-pass successive decoder removes all interference in two decoding passes for block length $n = 5000$. This value might be different for other codes, block length and channel load. From Fig. 6.9 we observe that this threshold corresponds to BER $\simeq 5 \cdot 10^{-3}$ for $n = 5000$ and BER $\simeq 5 \cdot 10^{-4}$ for $n = 10000$.

We have simulated a system with $\rho = 2$ bit/s/Hz, spreading factor $N = 64$ and $K = 128$ users, corresponding to the mark in Fig. 6.10. The evolution of the minimum, maximum and average users' SINR as a function of the

[1] j in the complex symbol context is given by $j^2 = -1$.

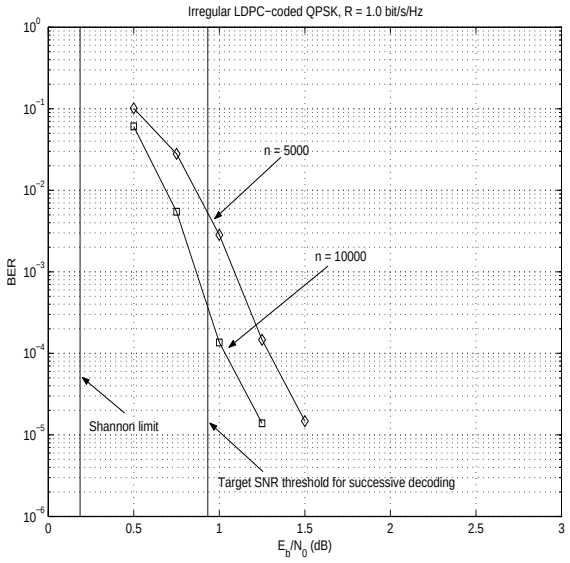


Figure 6.9: Average BER performance of irregular LDPCs of binary rate 1/2 over (single-user) AWGN, block lengths $n = 5000$ and $n = 10000$

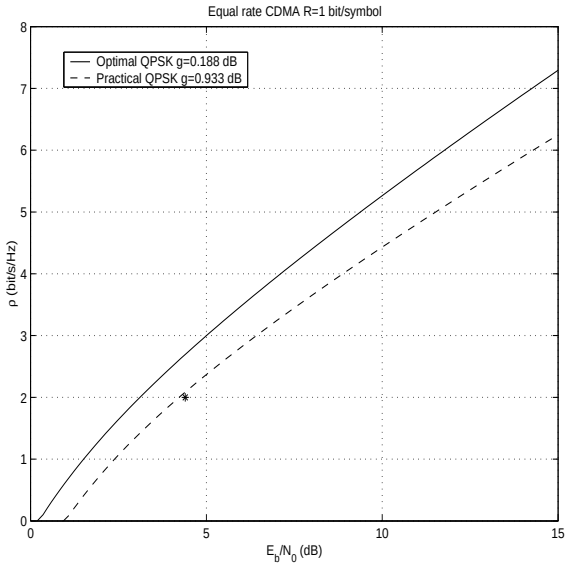


Figure 6.10: Spectral efficiency achievable by optimal and suboptimal QPSK code ensembles of rate $R = 1$

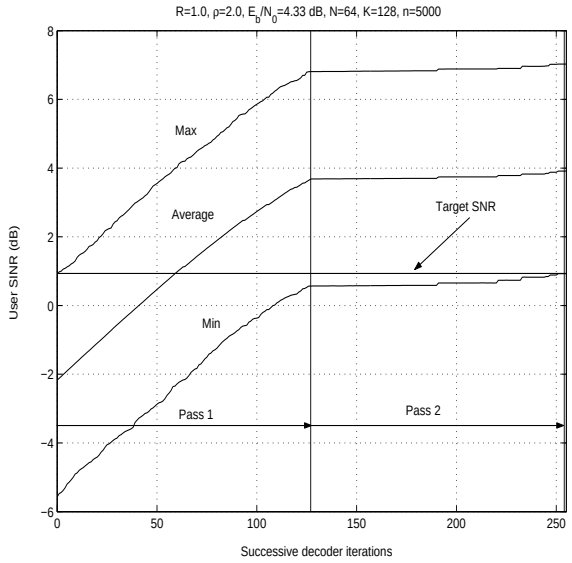


Figure 6.11: Evolution of the user SINR at the LDPC decoder input vs. the successive decoding steps, for the multi-pass soft-stripping decoder with LDPC code length $n = 5000$

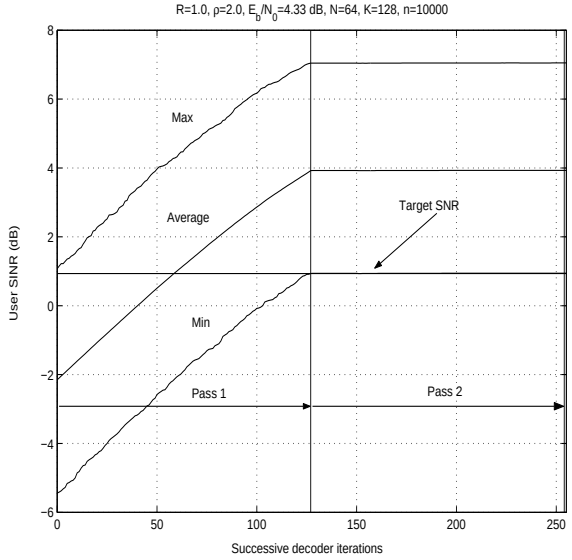


Figure 6.12: Evolution of the user SINR at the LDPC decoder input vs. the successive decoding steps, for the multi-pass soft-stripping decoder with LDPC code length $n = 10000$

successive stripping decoder iterations is shown in Fig. 6.11 for $n = 5000$ and in Fig. 6.12 for $n = 10000$. These curves are snapshots obtained by random generation of the noise, the spreading sequences, the information sequences and the LDPC code graphs. The successive decoder makes use of soft stripping. Each LDPC decoder is run for a maximum of 200 iterations, and a maximum of three interference cancellation passes. A complete interference cancellation pass corresponds to 128 decoding steps, i.e., to the decoding of all $K = 128$ users.

For $n = 5000$ we need to perform two soft-stripping successive decoding passes before all users reach their target SNR threshold. For $n = 10000$, since the threshold is much more conservative, a single pass is sufficient. Equivalently, one could have lowered the threshold and achieved the same spectral efficiency at a smaller E_b/N_0.

As a general remark, the gap between an optimal (infinite block length) system and a practical finite-length system depends almost entirely on the fact that finite-length codes require a significantly larger SNR (0.933 dB for $n = 5000$ and 0.747 dB for $n = 10000$) than the infinite-length decoding threshold. On the other hand, no catastrophic error propagation effect is observed when the system is designed according to the rules outlined above.

6.3 Conclusion

We have considered the optimization of a canonical coded synchronous CDMA system characterized by random spreading and QPSK signaling, in the limit of large number of users, large spreading gain, and large user code block length. Such assumptions may be regarded as "pragmatic", in the sense that they are all motivated by actual CDMA systems. The CDMA system considered here has low complexity, as it assumes successive stripping with MMSE filters. Excellent approximations to the MMSE filters can be precomputed using the large random matrix design approach of [97], with complexity $O(K^2)$. Moreover, powerful long user codes such as LDPC codes can be decoded iteratively, with linear complexity in the block-length. Hence, the overall complexity per decoded information bit of the multiuser decoder is linear in K and constant in the code block length, i.e., comparable with the complexity of standard CDMA systems based on single-user detection and separated single-user decoding. Nevertheless, the proposed system optimization, driven by recent information-theoretic results, yields spectral efficiencies

remarkably close to the optimal (i.e., optimizing also with respect to the user signature waveforms and using Gaussian codebooks).

We have considered two special cases of the general rate and power allocation problem: namely, the optimization of the received SNR distribution for an equal-rate system, and the optimization of the user rate distribution for an equal-power system, subject to the successive decodability condition imposed by the stripping decoder. Both problems yield linear programs that admit closed form explicit solutions.

From a practical point of view, the equal-rate system design seems more attractive than its equal-power counterpart since it can approach optimal spectral efficiency uniformly, for all E_b/N_0's, provided that the individual users coding rate is small. Moreover, controlling the received user SNR is much easier and closer to existing power-control schemes than allocating coding rates (and channel codes) to the users.

Numerical results show that the system optimization carried out in the large-system limit and for infinite code block length can be used effectively to dimension practical systems, provided that the SNR thresholds are chosen according to the actual BER performance of the finite-length user codes. Systems optimized according to the proposed method do not suffer from catastrophic error propagation of the successive stripping decoder even if, in general, finite-length codes have non-vanishing post-decoding BER.

APPENDIX

6.A Proof of Proposition 6.1

A necessary condition for $\boldsymbol{\beta}$ minimizing the objective function in (6.7) is that the constraint $\sum_j \beta_j \geq \beta$ holds with equality. Hence, without loss of generality we rewrite (6.7) in the canonical form

$$\begin{cases} \text{minimize} & \boldsymbol{\gamma}^T \boldsymbol{\beta} \\ \text{subject to} & -\mathbf{A}\boldsymbol{\beta} \geq -\mathbf{b} \\ & \mathbf{1}^T \boldsymbol{\beta} = \beta, \\ & \boldsymbol{\beta} \geq \mathbf{0} \end{cases} \tag{6.16}$$

The dual linear program is given by [98]

$$\begin{cases} \text{maximize} & (-\mathbf{b}^T, \beta) \begin{bmatrix} \mathbf{y} \\ \alpha \end{bmatrix} \\ \text{subject to} & \left[-\mathbf{A}^T, \mathbf{1}\right] \begin{bmatrix} \mathbf{y} \\ \alpha \end{bmatrix} \leq \boldsymbol{\gamma} \\ & \mathbf{y} \geq \mathbf{0} \end{cases} \tag{6.17}$$

where α can be either positive or negative.

From the properties of the coefficients $a_{i,j}$ and b_j we get immediately that $\mathbf{A}$ is invertible and the vector $\boldsymbol{\tau}$ such that $\boldsymbol{\tau} = \mathbf{A}^{-1}\mathbf{b}$ has non-negative components. The vector $\boldsymbol{\beta} \in \mathbb{R}_+^L$ maximizing $\mathbf{1}^T\boldsymbol{\beta}$ and satisfying $\mathbf{A}\boldsymbol{\beta} \leq \mathbf{b}$ is $\boldsymbol{\tau}$ (this is easily shown by contradiction, since $\boldsymbol{\tau}$ is the unique non-negative vector $\boldsymbol{\beta}$ that makes the inequality $\mathbf{A}\boldsymbol{\beta} \leq \mathbf{b}$ component-wise tight). Hence, if $\mathbf{1}^T\boldsymbol{\tau} < \beta$ the primal problem is infeasible. On the other hand, if $\mathbf{1}^T\boldsymbol{\tau} \geq \beta$ the primal problem is feasible, and a feasible point is given by (6.8). In order to show that this is indeed the desired solution, we shall assume that $\mathbf{1}^T\boldsymbol{\tau} \geq \beta$ and find a feasible point for the dual problem. Then, we show that the value of the dual problem at this point is equal to the value of the primal problem at (6.8).

We rewrite the inequality constraint and the objective function in the dual problem (6.17) as

$$\mathbf{A}^T\mathbf{y} \geq \alpha\mathbf{1} - \boldsymbol{\gamma} \tag{6.18}$$

and

$$-\mathbf{b}^T\mathbf{y} + \alpha\beta \tag{6.19}$$

The vector $\alpha\mathbf{1}-\boldsymbol{\gamma}$ has decreasing components. For fixed α, let K_α denote the number of positive elements of $\alpha\mathbf{1}-\boldsymbol{\gamma}$. It is clear from (6.18) and (6.19) that the objective function is maximized by letting the last $L-K_\alpha$ components of $\mathbf{y}$ equal to zero. We introduce the following short-hand notation: for a vector $\mathbf{x}\in\mathbb{R}^L$ and a matrix $\mathbf{M}\in\mathbb{R}^{L\times L}$, we let $\mathbf{x}_\alpha$ and $\mathbf{M}_\alpha$ denote the $K_\alpha\times 1$ subvector of $\mathbf{x}$ formed by its first K_α components, and the $K_\alpha\times K_\alpha$ submatrix of $\mathbf{M}$ formed by its first K_α rows and columns, respectively. Then, a feasible point for the dual problem is the vector $\boldsymbol{\pi}$ such that its first K_α components are given by

$$\boldsymbol{\pi}_\alpha = \alpha(\mathbf{A}_\alpha^T)^{-1}\mathbf{1}_\alpha - (\mathbf{A}_\alpha^T)^{-1}\boldsymbol{\gamma}_\alpha \tag{6.20}$$

and the remaining $L-K_\alpha$ components are equal to zero.

The value of the objective function (6.19) at this point is given by

$$f(\alpha) = \mathbf{b}_\alpha^T(\mathbf{A}_\alpha^T)^{-1}\boldsymbol{\gamma}_\alpha + \alpha\left(\beta - \mathbf{b}_\alpha^T(\mathbf{A}_\alpha^T)^{-1}\mathbf{1}_\alpha\right) \tag{6.21}$$

It is not difficult to see that $f(\alpha)$ is a continuous and piecewise linear function of α, for $\alpha\leq[\gamma_1,\gamma_L]$.

The assumption $\mathbf{1}^T\boldsymbol{\tau}\geq\beta$ can be rewritten as $\beta-\mathbf{1}^T\mathbf{A}^{-1}\mathbf{b}\leq 0$. Hence, for $\alpha>\gamma_s$, for some $1\leq s\leq L$, the slope of $f(\alpha)$ is negative, while for $\alpha<\gamma_s$ the slope is positive. Therefore, the maximum of $f(\alpha)$ with respect to α is achieved for $\alpha=\gamma_s$ and, by definition, s is the minimum index in $1,\ldots,L$ such that $\sum_{j=1}^{s}\tau_j\geq\beta$, i.e., $s=\hat{L}$ defined in (6.8). The primal objective function evaluated at the feasible point (6.8) is given by

$$(\gamma_1,\ldots,\gamma_s,0,\ldots,0)\mathbf{A}^{-1}\mathbf{b} + \gamma_s\left(\beta - (\underbrace{1,\ldots,1}_{s},0,\ldots,0)\mathbf{A}^{-1}\mathbf{b}\right)$$

It is immediate to see that this coincides with the dual objective function $f(\alpha)$ evaluated at $\alpha=\gamma_s$. Hence, we conclude that (6.8) is the sought solution. □

6.B Proof of Proposition 6.2

The proof follows immediately by observing that, for $\beta \leq b_L$, the program (6.13) can be reformulated as the $\hat{L}$-dimensional polymatroid program [99]

$$\begin{cases} \text{maximize} & \displaystyle\sum_{i=1}^{\hat{L}} \beta_i R_i \\ \text{subject to} & \displaystyle\sum_{i \in S} \beta_i \leq r(S), \quad \forall\, S \subseteq \{1, \ldots, \hat{L}\} \\ & \boldsymbol{\beta} \geq \mathbf{0} \end{cases} \tag{6.22}$$

where the rank function $r(S)$ is defined by

$$r(S) = \sum_{i=1}^{\max\{S\}} \Delta_i$$

where $\Delta_i = b_i - b_{i-1}$ for $i = 1, \ldots, \hat{L} - 1$ and $\Delta_{\hat{L}} = \beta - b_{\hat{L}-1}$. Since $(b_0, \ldots, b_{\hat{L}-1}, \beta)$ is increasing, $r(S)$ is submodular. □

Chapter 7

Conclusions and Perspectives

Conclusion

In this thesis, we have proposed various low-complexity coding/decoding schemes to approach the capacity limits of binary-input symmetric-output channels and CDMA channels.

First, we have analyzed the belief propagation decoder of the infinite block length, systematic, random-like IRA code ensemble using density evolution, assuming transmission over a binary-input symmetric-output channel. We have tracked the evolution of message densities over the cycle-free bipartite graph, and studied the local stability condition of the fixed point, corresponding to zero bit error rate, of the resulting two-dimensional non-linear dynamical system.

We have then addressed the optimization of the IRA code ensemble for the class of binary-input symmetric-output channels. The code design makes use of three tools: the mutual information evolution described by the EXIT function, the reciprocal channel approximation and the non-strict convexity of mutual information on the set of binary-input symmetric-output channels. We have proposed four methods to approximate densities involved in density evolution with a one-dimensional parameter, namely mutual information, thus yielding four low-complexity design methods of the IRA code ensemble,

that are formulated as linear programs. Again, local stability conditions of the fixed-points of the approximated density evolution systems have been derived. One of these one-dimensional DE approximation systems, based on Gaussian and reciprocal channel approximations, yields the same stability condition as the one under exact DE, while the exact stability condition has to be added in the optimization problem for other DE approximation methods, such as the one based and BEC a priori with reciprocal channel approximation.

We have optimized IRA codes for a wide range of rates, on the binary-input additive white Gaussian noise channel and the binary symmetric channel, using the four design methods. Comparing the BER thresholds of these codes evaluated by the exact DE, it is found that the best approximations are those based on the Gaussian approximation, and the BER performances of the codes thus designed are comparable to those of the best LDPC codes of the same rate. IRA codes are then an attractive alternative due to the simplicity of their coding/encoding.

Next, we have studied the BER and WER performances of finite length regular and irregular repeat accumulate codes on the BIAWGNC in the error floor region. Two approaches have been a adopted: girth maximization and minimum stopping set size maximization. The BER performances of girth-conditioned short-length regular RA codes are found to be above the performances of the random ML-decoded regular RA ensemble with uniform interleaving. They are also comparable to the performances of the best LDPC codes of the same rate. On the other hand, the BER performances of the large-block length irregular RA codes remain inferior to those of irregular LDPC codes with comparable length and graph conditioning.

After the analysis and design of IRA codes that closely approach the Shannon limit in the single user case, we have tackled the problem of approaching the optimum spectral efficiency of the random CDMA channel in the large system limit. We have shown that the loss in spectral efficiency due to the use of QPSK inputs instead of Gaussian ones vanishes as the channel load becomes large.

We have then considered approaching the optimum CDMA with a low-complexity encoding/decoding setting, using capacity-achieving (approaching) binary error correcting codes, QPSK modulation and stripping decoder. We have optimized the channel load distribution in two cases: the equal-rate system with non-equal SNR levels, and the equal-power system with non-equal rates, subject to the successive decodability condition imposed by the

stripping decoder. With the equal-rate system design, the spectral efficiency approaches the optimal spectral efficiency for low code rates. The equal-rate system design seems more attractive than the equal-power system design which requires a very wide range of rates to get close to the optimal. Numerical results have shown that these design methods can be used to dimension practical systems, and the resulting systems do not suffer from catastrophic error propagation.

Future Work

Irregular repeat accumulate codes can be seen as a special instance of irregular turbo codes, with the constituent convolutional encoder having only two states. Although the threshold of the designed IRA code ensembles in this work are very close to Shannon limit and can hardly be improved any further, we believe that the finite length performance of irregular turbo codes (i.e. convolutional constituent encoders with larger memory) can bring a substantial improvement over that of irregular repeat accumulate codes for medium to large block lengths. This is a direction that should be investigated. It must be noted that the DE analysis of such codes presents much more difficulty than for IRA codes, because of the presence of cycles in the bipartite graph of such an irregular turbo code due to the increased memory. Moreover, DE approximation methods 1 and 2 cannot be used, for the same reasons. However, one can still use methods 3 and 4 to design capacity-approaching irregular turbo codes, because these methods make use of the BCJR algorithm on the cycle-free graph of the convolutional code resulting from the use of hidden variables (states).

Another issue that could be investigated is the design of a low complexity scheme to approach the spectral efficiency of a cellular multiple access channel. In [100], Wyner proposed to model the received signal at a cell site as the sum of the noise, the signals transmitted from within the cell and a scaled version of the sum of signals from adjacent cells with scaling factor $0 \leq \alpha \leq 1$. However, users from adjacent cells do not cause the same inter-cell interference at the cell site, depending on how close they are from the edge of the cell. Therefore one could use different scaling factors that depend on the position of the adjacent interferer and accordingly optimize the power/rate profile of the users to maximize the spectral efficiency.

Publications

Repeat Accumulate Codes

- A. Roumy, S. Guemghar, G. Caire and S. Verdú, *Design Methods for Irregular Repeat-Accumulate Codes*, IEEE International Symposium on Information Theory, ISIT 2003, Yokohama, Japan, June 29$^{\text{th}}$ - July 5$^{\text{th}}$ 2003.

- A. Roumy, S. Guemghar, G. Caire and S. Verdú, *Design Methods for Irregular Repeat-Accumulate Codes*, submitted to IEEE Transactions on Information Theory, October 2002 (revised October 2003).

- S. Guemghar and G. Caire, *On the Performance of Finite Length Irregular Repeat Accumulate Codes*, 5th International ITG Conference on Source and Channel Coding, January 14-16 2004.

Coded CDMA under Successive Decoding

- G. Caire, S. Guemghar, A. Roumy and S. Verdú, *Design Approaches for LDPC-Encoded CDMA*, 39th Annual Allerton Conference on Communication, Control and Computing, Monticello, IL, October 3-5, 2001.

- G. Caire, S. Guemghar, A. Roumy and S. Verdú *Maximizing the Spectral Efficiency of Coded CDMA under Successive Decoding*, IEEE Transactions on Information Theory, Vol. 50, Issue 1, Jan. 2004.

Chapter 8

Résumé Détaillé en Français

8.1 Introduction

8.1.1 Techniques Avancées de Codage

Claude Shannon a montré [1] l'existence de codes permettant une transmission fiable si le taux (rendement) de transmission d'information, mesuré en bits par utilisation du canal, est inférieur à la *capacité* du canal. Un code aléatoirement construit et ayant une grande longueur de bloc a une grande probabilité d'approcher la limite de Shannon. Cependant, la complexité de décodage d'un tel code est exponentielle en la longueur du code, rendant le décodage impossible en pratique. Le but de la théorie de codage est de construire des codes performants, atteignant la limite de Shannon, tout en ayant une complexité de décodage raisonnable. Une importante étape dans cette direction a été franchie par David Forney [2] en introduisant des codes concaténés. Ces derniers se composent d'un code bloc *externe* puissant et d'un code convolutif *interne*. Le code interne est décodé par l'algorithme de Viterbi, suivi du décodeur externe qui fait usage de décisions dures et de décodage algébrique.

En 1993, Berrou, Glavieux et Thitimajshima ont introduit une nouvelle technique de codage: les *turbo codes*. Un turbo code résulte de la concaténation de deux codes convolutifs à travers un entrelaceur. Cette struc-

ture admet un schéma de décodage itératif, basé sur l'estimation récursive de probabilités *a posteriori* (APP) en utilisant l'algorithme BCJR [4], ainsi que l'échange d'informations *extrinsèques* entre les deux décodeurs constituants. La performance des turbo codes approche d'une fraction de dB la limite de Shannon du canal à entrée binaire et sortie Gaussienne.

L'introduction des turbo codes constitue une avancée majeure dans la théorie de codage en engendrant plusieurs travaux de recherche dans le domaine des codes pseudo-aléatoires, et en donnant naissance à la famille de codes *turbo-like*. Nous notons en particulier la redécouverte des codes LDPC (*low density parity check*) initialement proposés dans [5], l'introduction des codes LDPC irréguliers [6, 7] et l'introduction des codes RA (répétition-accumulation ou *repeat-accumulate*) [8]. Les performances de ces codes sont proches des limites de Shannon de plusieurs canaux de grand intérêt. Il a été montré dans [6, 7] que les codes LDPC irréguliers atteignent asymptotiquement la capacité du canal BEC (*binary erasure channel*) en faisant usage d'un décodeur par passage de messages. Bien que le BEC soit le seul canal pour lequel ce résultat existe, des codes LDPC irréguliers avec de très bonnes performances ont été conçus pour d'autres canaux à entrée binaire, tels le BSC (*binary symmetric channel*), le BIAWGNC (*binary-input additive white Gaussian noise channel*) [9] et le canal ISI (*inter-symbol interference*) [10, 11, 12, 13].

La description de ces codes turbo-like et de leur décodage se fait à l'aide du *graphe de Tanner (bipartite)* [14]. Les codes sont décodés itérativement en effectuant des opérations locales aux *sommets*, appelées *contraintes locales*, puis en passant l'information obtenue le long des *arêtes* du graphe. La complexité de décodage dépend de la complexité des contraintes locales , de la complexité du passage de l'information sur les arêtes et du nombre des itérations du décodeur. Nous nous intéressons principalement aux décodeurs par passage de messages pour lesquels les messages représentent des estimations des bits transmis. En outre, nous considérons uniquement l'algorithme somme-produit (*sum-product*) aussi appelé algorithme de propagation des croyances (*belief propagation* ou BP), qui suppose l'indépendance locale des messages *entrants* aux sommets. Dans ce cas, les contraintes locales aux sommets obéissent aux lois de probabilité.

L'introduction des codes LDPC irréguliers a motivé d'autres schémas de codage irréguliers et turbo-like, tels les codes irréguliers RA (IRA) [15] et les turbo codes irréguliers [16, 17]. Les codes IRA, qui sont en fait un type particulier de codes LDPC et de turbo codes irréguliers, atteignent la

capacité du canal BEC. Ces codes sont particulièrement intéressants de par la faible complexité de leur codage/décodage et de leur bonne performance comparable à celle des codes LDPC et turbo.

Un code IRA possède les degrés de liberté suivants: un *profil de répétition* $\{f_i \geq 0, i = 2, 3, \ldots : \sum_{i=2}^{\infty} f_i = 1\}$ et un *facteur de groupement* a. Une fraction f_i des bits d'information est répétée i fois, où $i = 2, 3, \ldots$. La séquence résultant de l'étape de répétition est entrelacée et est passée en entrée d'une machine récursive à états finis, l'*accumulateur*, qui sort un bit tous les a bits entrés.

La machine récursive à états finis est le schéma le plus simple permettant d'obtenir un rendement rationnel arbitraire entre 0 et 1. Dans le présent travail, nous nous intéressons exclusivement aux codes IRA utilisant des codes convolutifs à deux états, bien qu'il soit légitime de penser que l'inclusion de la machine à états finis dans l'ensemble des degrés de liberté pourrait donner lieu à la construction de meilleurs codes. Dans le cas où le nombre d'états est supérieur à deux, le code prend la forme générale d'un turbo code irrégulier.

Les premières tentatives d'optimisation de codes LDPC irréguliers ([18] pour le BEC et [19] pour d'autres canaux) étaient basées sur la technique de l'*évolution de densités* (*density evolution* ou DE), qui consiste à déterminer la performance moyenne d'un ensemble de codes pseudo-aléatoires dans la limite d'une longueur de bloc infinie. Afin de réduire la complexité de l'optimisation de l'ensemble basée sur la DE, plusieurs méthodes d'approximation de la DE ont été proposées. Elles consistent à approximer les densités de probabilité des messages passés sur le graphe par des paramètres unidimensionnels. Ces techniques ne sont exactes que dans le cas du BEC, pour lequel la DE est unidimensionnelle. Les techniques d'approximation les plus utilisées sont basées sur l'*approximation Gaussienne* des messages, qui, alliée à la *condition de symétrie* des densités des messages, permet de représenter la densité des messages par un seul paramètre. Les techniques d'approximation varient selon le paramètre utilisé pour représenter les densités des messages et les fonctions utilisées pour définir le système dynamique qui décrit l'évolution des densités sur le graphe de Tanner de l'ensemble de codes [20, 21, 22, 23, 24, 25, 26]..

Dans cette thèse, nous considérons l'optimisation d'ensembles de codes IRA pour des canaux à entrée binaire et sortie symétrique. Nous proposons quatre méthodes d'approximation de la DE par l'évolution d'un paramètre unidimensionnel, en l'occurrence l'information mutuelle entre les bits transmis et les *rapports de vraisemblances* sur le graphe. Ces méthodes se dis-

tinguent par les différentes façons d'approximer les densités des messages et le calcul des messages aux sommets avec l'algorithme BP. La première méthode fait usage de l'approximation Gaussienne et de l'approximation *réciproque* (*duale*). La deuxième méthode utilise toujours l'approximation réciproque, mais suppose cette fois que les messages sont les sorties d'un canal BEC virtuel dont la capacité est égale à l'information mutuelle calculée au sommet considéré. Ces deux premières méthodes permettent d'exprimer les récursions de la DE de façon analytique, contrairement aux méthodes 3 et 4 qui n'ont pas cette faculté. En effet, ces deux dernières méthodes font usage de simulations Monte Carlo afin de suivre l'évolution de l'information mutuelle sur le graphe de Tanner. La formulation de ces méthodes est telle qu'elle reste valide pour tout canal à entrée binaire et sortie symétrique.

Si la longueur de bloc est finie, le graphe de Tanner contient des cycles et l'hypothèse de l'indépendance locale n'est plus valide. En effet, des codes turbo-like aléatoires de longueur finie peuvent avoir de mauvaises performances avec l'algorithme de décodage BP, et l'écart de la capacité du canal peut se détériorer au passage de la longueur infinie à la longueur finie. La conception de l'entrelaceur devient de ce fait une question cruciale pour la performance de codes sur graphes à longueur finie. Les critères de construction d'entrelaceurs reposent essentiellement sur des arguments heuristiques: l'élimination de cycles courts afin de limiter la propagation de messages à faible fiabilité; la maximisation de la *distance minimale* afin d'éliminer les mots de code de faible *poids de Hamming* responsables de la mauvaise performances pour les hauts rapports signal à bruit; la maximisation de la taille d'ensembles bloquants (*stopping sets*) [27] responsables des erreurs résiduelles de codes LDPC sur le BEC.

Dans cette thèse, nous analysons la performance de codes IRA de longueur finie sur le BIAWGNC, avec des structures d'entrelacement obéissant à l'une des deux contraintes suivantes: (a) la taille du plus petit cycle est supérieure à un certain paramètre, (b) la taille du plus ensemble bloquant est supérieure à un certain paramètre. Nous montrons aussi comment calculer la performance moyenne de l'ensemble aléatoire de codes RA réguliers sur le BIAWGNC, en utilisant un décodeur par maximum de vraisemblance, et la comparons à la performance moyenne de l'ensemble de codes RA réguliers dont le graphe de Tanner obéit au critère (a).

8.1.2 CDMA Codé avec Décodage Successif

Dans un *canal à accès multiple* MAC (*multiple access channel*), plusieurs utilisateurs transmettent sur le même canal physique. Ainsi, la sortie du canal est une version bruitée de la superposition des signaux transmis. Dans ce travail, nous considérons le canal Gaussien à accès multiple (*Gaussian multiple access channel* ou GMAC) où le bruit est additif blanc Gaussien (AWGN pour *additive white Gaussian noise*). La capacité du GMAC étant le nombre total d'utilisateurs pouvant transmettre de façon fiable sur le canal, la théorie de l'information multi-utilisateur [28] montre que la capacité est généralement maximisée par la transmission de signaux qui interfèrent entre eux (non orthogonaux). La principale mesure est l'*efficacité spectrale* qui représente le taux total de transmission de données par unité de largeur de bande (bits par seconde par Hertz ou bits par utilisation de canal).

Dans ce travail, nous nous proposons des schémas de codage/décodage de faible complexité afin d'approcher l'efficacité spectrale maximale du GMAC. Afin d'atteindre la performance optimale, il est nécessaire d'employer des stratégies de codage adéquates, ainsi qu'une structure de décodage permettant de décoder les données transmises et ce de façon arbitrairement fiable. Ceci est accompli à l'aide du décodage successif qui décode un utilisateur donné en considérant tous les autres utilisateurs comme du bruit, puis soustrait la contribution de l'utilisateur décodé du signal reçu et répète ces deux opérations jusqu'à ce que tous les utilisateurs soient décodés avec succès.

Dans cette thèse, nous adoptons la stratégie de codage appelée *accès multiple par répartition de codes* (CDMA pour *code division multiple access*) aussi appelée *accès multiple par étalement de spectre* [29, 30]. En CDMA, chaque utilisateur se voit octroyer une *séquence d'étalement* qui est un vecteur unitaire dans un espace de dimension N et dont les éléments sont appelés *bribes* (*chips*). Le *facteur d'étalement* N est le nombre de bribes par symbole émis, et le rapport entre le nombre d'utilisateurs et le facteur d'étalement est appelé le *gain d'étalement*. Nous supposons que l'étalement est aléatoire, i.e., les séquences d'étalement sont octroyées aléatoirement, et les bribes sont indépendantes et identiquement distribuées (iid). Cette hypothèse est justifiée par:

- L'étalement aléatoire modélise bien les systèmes CDMA employant de longues séquences d'étalement recouvrant plusieurs périodes de symbole.

- L'efficacité spectrale moyennée sur l'ensemble des signatures possibles est une borne inférieure à l'efficacité spectrale optimale obtenue en utilisant des séquences d'étalement déterministes optimales.

L'efficacité spectrale d'un système CDMA synchrone avec bruit Gaussien et étalement aléatoire a été déterminée dans [31, 32], où l'on suppose que le nombre d'utilisateurs ainsi que le facteur d'étalement sont infinis, et que le gain d'étalement est fini. Ces conditions représentent les limites d'un grand système. L'efficacité spectrale maximale est atteinte en utilisant des entrées Gaussiennes, mais les systèmes pratiques emploient des constellations discrètes de petites tailles. Certains travaux récents utilisent l'analyse asymptotique de systèmes CDMA avec étalement aléatoire et différentes structures de réception dans le but de construire des systèmes CDMA pratiques [33].

Nous étudions dans ce travail l'efficacité spectrale maximale d'un système CDMA aléatoire synchrone, dans la limite d'un grand système et avec une taille de bloc infinie, en considérant le cadre suivant: une modulation à déplacement de phase quadrivalente (QPSK pour *quaternary phase shift keying*), des codes binaires correcteurs d'erreurs atteignant (ou approchant) la capacité du canal, un filtre à erreur quadratique moyenne minimale MMSE (*minimum mean square error*) et un décodage successif. Pour des codes binaires donnés, on maximise l'efficacité spectrale du système CDMA avec un décodeur successif, quand tous les utilisateurs sont reçus à la même puissance ou quand tous les utilisateurs transmettent au même rendement. Dans les deux cas de figure, le problème de la maximisation de l'efficacité spectrale peut être formulé comme un programme linéaire dont la solution peut être écrite analytiquement. A l'aide d'exemples utilisant des codes LDPC irréguliers optimisés pour le BIAWGNC, nous montrons que le cas de rendement égal est quasi-optimal en approchant de très près la capacité du GMAC. Dans le cas d'un rendement commun à tous les utilisateurs, nous étudions la performance d'un système CDMA de taille finie basé sur les résultats de l'optimisation, et montrons que notre approche de faible complexité permet de limiter la propagation d'erreurs résiduelles. Notre approche de décodage à faible complexité repose sur un schéma itératif à deux aspects:

- Les codes LDPC binaires sont décodés itérativement.

- Le décodeur successif est itéré afin d'éliminer les erreurs résiduelles à la suite du décodage de chaque utilisateur.

8.1.3 Organisation du Résumé

Cette thèse est composée de deux parties. La première partie étudie les codes IRA, leur codage, leur décodage, leur conception et leur performance. La deuxième partie considère la maximisation de l'efficacité spectrale d'un système CDMA afin d'approcher la capacité du GMAC en utilisant un schéma de codage/décodage de faible complexité.

La section 8.2 présente l'encodeur IRA, le décodeur BP et l'évolution des densités sur le graphe de Tanner, pour les canaux à entrée binaire et sortie symétrique. Dans la section 8.3, nous décrivons l'optimisation des codes IRA sous la forme d'un programme linéaire, en introduisant deux applications de (vers) l'ensemble des distributions symétriques vers (de) l'ensemble des nombres réels. Nous y présentons la fonction EXIT (*extrinsic mutual information transfer function*) et la fonction de capacité de canaux à entrée binaire et sortie symétrique. Nous présentons quatre méthodes pour approximer la DE, ainsi que certaines propriétés analytiques de ces méthodes. Nous finissons cette section en présentant des codes optimisés pour le BIAWGNC et le BSC et leurs seuils de décodage évalués à l'aide de la DE exacte. La section 8.4 étudie la performance de codes RA réguliers et irréguliers sur le BIAWGNC, et dont le graphe de Tanner est dépourvu de cycles ou ensembles bloquants de longueurs inférieures à un certain paramètre.

La section 8.5 présente les limites théoriques qui peuvent être atteintes par le décodage successif sur le canal CDMA aléatoire, dans la limite d'un grand système. Dans la section 8.6, on présente une méthode pour approcher la capacité du GMAC en utilisant une modulation QPSK, des codes binaires correcteurs d'erreur et un décodeur successif.

Codes Irréguliers Répétition-Accumulation

8.2 Codage et Décodage de Codes IRA

Fig. 8.1 montre l'encodeur d'un code IRA systématique. Chaque bit b_i, où $i = 1, \ldots, k$ est répété $2 \leq r_i \leq d$ fois, puis le bloc $\mathbf{b}$ est entrelacé résultant en un bloc de répétition $\mathbf{x}_1$ de longueur N. $\mathbf{x}_1$ est transmis à un accumulateur dont la sortie $\mathbf{x}_2$ est de longueur m. Le rendement de l'accumulateur est le facteur de groupement a et le rendement du code IRA est

$$R = \frac{k}{k+m} = \frac{a\sum_{i=2}^{d} \lambda_i / i}{1 + a\sum_{i=2}^{d} \lambda_i / i} = \frac{a}{a+\bar{d}} \tag{8.1}$$

où $\bar{d}$ est le degré moyen d'un bit d'information et λ_i est la fraction d'arêtes reliées à un bit d'information de degré i.

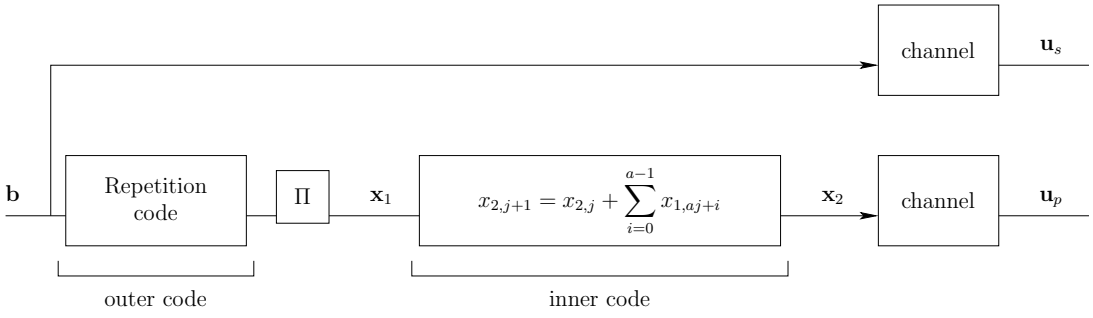


Figure 8.1: Encodeur d'un code IRA systématique

Le graphe de Tanner représentant le code IRA est constitué de trois types de sommet: des sommets de bits d'information (*information bitnodes*), des sommets de bits de parité (*parity bitnodes*) et des sommets de contrle de parité (*checknodes*). Un bitnode est lié à un checknode par une arête.

Chaque bitnode est associé à un message d'observation du canal donné par

$$u_\alpha = \log \frac{p_{Y|X}(y_\alpha | x_\alpha = 0)}{p_{Y|X}(y_\alpha | x_\alpha = 1)} \tag{8.2}$$

L'algorithme de décodage somme-produit est initialisé en mettant les messages des checknodes à zéro. Considérant un bitnode α de degré i, le message sortant de ce sommet est donné par

$$m_{o,\alpha} = u_\alpha + \sum_{j=1}^{i-1} m_j \tag{8.3}$$

où m_j est un message entrant par l'arête j. Considérant un checknode β de degré i, le message sortant de ce sommet est donné par

$$\tanh \frac{m_{o,\beta}}{2} = \prod_{j=1}^{i-1} \tanh \frac{m_j}{2} \tag{8.4}$$

où m_j est un message entrant par l'arête j.

Deux méthodes de séquencement du décodage sont possibles:

- Séquencement LDPC, où tous les bitnodes et tous les checknodes sont activés alternativement et en parallèle.

- Séquencement Turbo, où les bitnodes d'information sont d'abord activés puis l'algorithme BCJR est appliqué au treillis de l'accumulateur.

La technique DE étudie l'évolution des densités des messages propagés sur un graphe aléatoire de taille infinie, où les messages sont localement indépendants. Le *théorème de concentration* [18, 39] garantit que le taux d'erreur de l'algorithme somme-produit appliqué à un code aléatoirement tiré de l'ensemble a une grande probabilité d'être égal au taux d'erreur calculé avec la DE pour une longueur de bloc suffisamment grande.

Les canaux à entrée binaire et sortie symétrique sont caractérisés par une condition de symétrie des distributions de messages. Pour une distribution F possédant une densité de probabilité f, cette condition se traduit par

$$f(x) = e^x f(-x) \tag{8.5}$$

et l'on dénote par $\mathcal{F}_{sym}$ l'ensemble des distributions symétriques.

La DE de l'ensemble de codes IRA aléatoire est donnée par la proposition suivante.

Proposition 8.1 *Considérons que P_ℓ [resp., $\widetilde{P}_\ell$] dénote la distribution moyenne des messages passés d'un information bitnode [resp., parity bitnode] vers un checknode, à l'itération ℓ. Considérons aussi que Q_ℓ [resp., $\widetilde{Q}_\ell$] dénote la distribution moyenne des messages passés d'un checknode vers un information bitnode [resp., parity bitnode], à l'itération ℓ.*

Sous la condition que le graphe n'a pas de cycle, $P_\ell, \widetilde{P}_\ell, Q_\ell, \widetilde{Q}_\ell$ satisfont les récursions suivantes:

$$\begin{aligned} P_\ell &= F_u \otimes \lambda(Q_\ell) && (8.6)\\ \widetilde{P}_\ell &= F_u \otimes \widetilde{Q}_\ell && (8.7)\\ Q_\ell &= \Gamma^{-1}\left(\Gamma(\widetilde{P}_{\ell-1})^{\otimes 2} \otimes \Gamma(P_{\ell-1})^{\otimes(a-1)}\right) && (8.8)\\ \widetilde{Q}_\ell &= \Gamma^{-1}\left(\Gamma(\widetilde{P}_{\ell-1}) \otimes \Gamma(P_{\ell-1})^{\otimes a}\right) && (8.9) \end{aligned}$$

pour $\ell = 1, 2, \ldots$, avec comme condition initiale $P_0 = \widetilde{P}_0 = \Delta_0$, où F_u est la distribution des messages d'observation du canal (8.2), $^\otimes$ représente la convolution de distributions, définie par

$$(F \otimes G)(z) = \int F(z-t)dG(t) \tag{8.10}$$

$^{\otimes m}$ représente la m-ème convolution, $\lambda(F) \triangleq \sum_{i=2}^{d} \lambda_i F^{\otimes(i-1)}$, $\Gamma(F_x)$ est la distribution de $y = \gamma(x)$ (défini sur $\mathbb{F}_2 \times \mathbb{R}_+$), où $x \sim F_x$, et Γ^{-1} dénote la fonction inverse de Γ, i.e., $\Gamma^{-1}(G_y)$ est la distribution de $x = \gamma^{-1}(y)$ si $y \sim G_y$. Dans ce qui précède, nous avons défini $\gamma(x) = \left(sign(x), -\log\tanh\frac{|x|}{2}\right)$.

$(\Delta_\infty, \Delta_\infty)$ est un point fixe des récursions (8.6 – 8.9), et sa stabilité locale est donnée par le résultat suivant.

Théorème 8.1 *Le point fixe $(\Delta_\infty, \Delta_\infty)$ de la DE est localement stable si et seulement si*

$$\lambda_2 < \frac{e^r(e^r-1)}{a+1+e^r(a-1)} \tag{8.11}$$

où $r = -\log(\int e^{-z/2} dF_u(z))$.

8.3 Construction de Codes IRA

Considérons les applications $\Phi : \mathcal{F}_{\text{sym}} \to \mathbb{R}$ et $\Psi : \mathbb{R} \to \mathcal{F}_{\text{sym}}$. Nous dérivons un système dynamique du système (8.6 – 8.9) tel que

$$\begin{aligned} x_\ell &= \Phi\left(F_u \otimes \lambda\left(\mathsf{Q}_\ell\right)\right) && (8.12)\\ \widetilde{x}_\ell &= \Phi\left(F_u \otimes \widetilde{\mathsf{Q}}_\ell\right) && (8.13)\\ \mathsf{Q}_\ell &= \Gamma^{-1}\left(\Gamma\left(\Psi(\widetilde{x}_{\ell-1})\right)^{\otimes 2} \otimes \Gamma\left(\Psi(x_{\ell-1})\right)^{\otimes(a-1)}\right) && (8.14)\\ \widetilde{\mathsf{Q}}_\ell &= \Gamma^{-1}\left(\Gamma\left(\Psi(\widetilde{x}_{\ell-1})\right) \otimes \Gamma\left(\Psi(x_{\ell-1})\right)^{\otimes a}\right) && (8.15) \end{aligned}$$

pour $\ell = 1, 2, \ldots$, avec comme condition initiale $x_0 = \widetilde{x}_0 = \Phi(\Delta_0)$, et où $(x_\ell, \widetilde{x}_\ell)$ sont les variables d'état du système. Ce système peut être écrit de façon équivalente comme

$$\begin{aligned} x_\ell &= \phi(x_{\ell-1}, \widetilde{x}_{\ell-1}) \\ \widetilde{x}_\ell &= \widetilde{\phi}(x_{\ell-1}, \widetilde{x}_{\ell-1}) \end{aligned} \tag{8.16}$$

(8.16) admet un point fixe à (1,1). Afin que ce point fixe soit unique, il est nécessaire (les fonctions Φ et Ψ ayant certaines propriétés) que $x < \phi(x, \widetilde{x}(x)), \quad \forall \ x \in [0, 1)$. Cette condition est la condition de taux d'erreur nul de la DE approximée, et l'on peut par conséquent obtenir une méthode approximée d'optimisation de codes IRA:

$$\begin{cases} \text{maximiser} & a \displaystyle\sum_{i=2}^{d} \lambda_i / i \\ \text{sous les contraintes} & \displaystyle\sum_{i=2}^{d} \lambda_i = 1, \quad \lambda_i \geq 0 \ \forall \ i \\ & x < \phi(x, \widetilde{x}(x)), \quad \forall \ x \in [0, 1) \end{cases} \tag{8.17}$$

Différentes méthodes d'approximation de la DE existent selon le choix des fonctions Φ et Ψ, de l'approximation des distributions Q, $\tilde{\mathsf{Q}}$ et F_u et du séquencement du décodage. Dans ce qui suit, nous citons quatre de ces méthodes pour lesquelles

$$\Phi : F \mapsto \mathcal{I}(F)$$

où

$$\mathcal{I}(F) \triangleq 1 - \int_{-\infty}^{\infty} \log_2\left(1 + e^{-z}\right) dF(z) \tag{8.18}$$

associe une distribution symétrique F à la capacité du canal dont la probabilité de transition est donnée par F. Ceci revient à décrire la DE en termes de l'évolution de l'information mutuelle entre les variables associées aux différents sommets et les messages propagés. L'idée est de décrire le calcul effectué à un sommet pendant le décodage somme-produit par une *fonction de transfert d'information mutuelle* appelée fonction *EXIT* [20, 21, 26, 42, 48, 49, 50]. Les deux premières méthodes résultent en des approximations analytiques tandis que les deux autres font usage de la méthode Monte Carlo.

8.3.1 Méthode 1

Les distributions à chaque itération sont supposées Gaussiennes. Une distribution Gaussienne satisfait la condition de symétrie (8.5) si et seulement si la variance est égale au double de la moyenne, et est dénotée par $\mathcal{N}_{sym}(\mu) \triangleq \mathcal{N}(\mu, 2|\mu|)$.

Nous considérons les approximations suivantes:

$$\begin{aligned} \Psi: \quad x &\mapsto \mathcal{N}_{\text{sym}}(J^{-1}(x)) && (8.19)\\ \mathsf{Q}_\ell &\approx \mathcal{N}_{\text{sym}}(\mu_\ell) && (8.20)\\ \widetilde{\mathsf{Q}}_\ell &\approx \mathcal{N}_{\text{sym}}(\widetilde{\mu}_\ell) && (8.21)\\ F_u &\approx \sum_{j=1}^{D} p_j \Delta_{v_j} && (8.22) \end{aligned}$$

où $D \geq 2$, $v_j \in \mathbb{R}$ et $p_j \in \mathbb{R}_+$ tel que $\sum_{j=1}^{D} p_j = 1$, et

$$J(\mu) \triangleq \mathcal{I}(\mathcal{N}_{\text{sym}}(\mu)) = 1 - \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-z^2} \log_2\left(1 + e^{-2\sqrt{\mu}z - \mu}\right) dz \qquad (8.23)$$

En utilisant l'approximation duale [41, 43], la DE est approximée par le système dynamique:

$$\begin{cases} \mu_\ell &= J^{-1}\left(1 - J\left((a-1)J^{-1}(1-x_{\ell-1}) + 2J^{-1}(1-\widetilde{x}_{\ell-1})\right)\right) \\ \widetilde{\mu}_\ell &= J^{-1}\left(1 - J\left(aJ^{-1}(1-x_{\ell-1}) + J^{-1}(1-\widetilde{x}_{\ell-1})\right)\right) \end{cases} \qquad (8.24)$$

et

$$\begin{cases} x_\ell &= 1-\sum_{i=2}^{d}\sum_{j=1}^{D}\lambda_i p_j \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-z^2}\log_2\left(1+e^{-2\sqrt{(i-1)\mu_\ell}z-(i-1)\mu_\ell-v_j}\right)dz \\ \widetilde{x}_\ell &= 1-\sum_{j=1}^{D} p_j \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-z^2}\log_2\left(1+e^{-2\sqrt{\widetilde{\mu}_\ell}z-\widetilde{\mu}_\ell-v_j}\right)dz \end{cases} \tag{8.25}$$

8.3.2 Méthode 2

Les messages à chaque itération sont supposés provenir d'un canal BEC virtuel tel que

$$\Psi : x \mapsto \mathcal{E}_{\text{sym}}(1-x) \tag{8.26}$$

où $\mathcal{E}_{\text{sym}}(\epsilon) \stackrel{\Delta}{=} \epsilon\Delta_0 + (1-\epsilon)\Delta_\infty$. En utilisant cette approximation et

$$\begin{array}{ll} \mathsf{Q}_\ell &= \mathcal{E}_{\text{sym}}(1-x_{\ell-1}^{a-1}\widetilde{x}_{\ell-1}^2) \\ \widetilde{\mathsf{Q}}_\ell &= \mathcal{E}_{\text{sym}}(1-x_{\ell-1}^{a}\widetilde{x}_{\ell-1}) \end{array} \tag{8.27}$$

l'approximation de la DE est exprimée sous forme d'un système dynamique sur l'ensemble des nombres réels:

$$\begin{cases} x_\ell &= 1-(1-\mathcal{I}(F_u))\sum_{i=2}^{d}\lambda_i\left(1-x_{\ell-1}^{a-1}\widetilde{x}_{\ell-1}^2\right)^{i-1} \\ \widetilde{x}_\ell &= 1-(1-\mathcal{I}(F_u))\left(1-x_{\ell-1}^{a}\widetilde{x}_{\ell-1}\right) \end{cases} \tag{8.28}$$

8.3.3 Méthodes 3 et 4

En supposant que le séquencement de décodage est turbo, les fonctions EXIT des décodeurs interne et externe sont obtenues par des simulations Monte Carlo (cf. Fig. 8.2). Ces fonctions EXIT sont généralement obtenues comme suit:

1. Les messages du canal sont iid selon la distribution F_u.
2. Les messages à l'entrée du décodeur sont iid selon la distribution F_{in}.
3. La distribution F_{out} des messages à la sortie du décodeur est obtenue analytiquement ou par simulation Monte Carlo.

4. La paire $I_A = \mathcal{I}(F_{\text{in}})$, $I_E = \mathcal{I}(F_{\text{out}})$ représente un point sur le graphe de la fonction EXIT.

Les méthodes 3 et 4 consistent à suivre les étapes précitées en supposant que $F_{\text{in}} = \mathcal{N}_{\text{sym}}(J^{-1}(I_A))$ et $F_{\text{in}} = \mathcal{E}_{\text{sym}}(1 - I_A)$ respectivement.

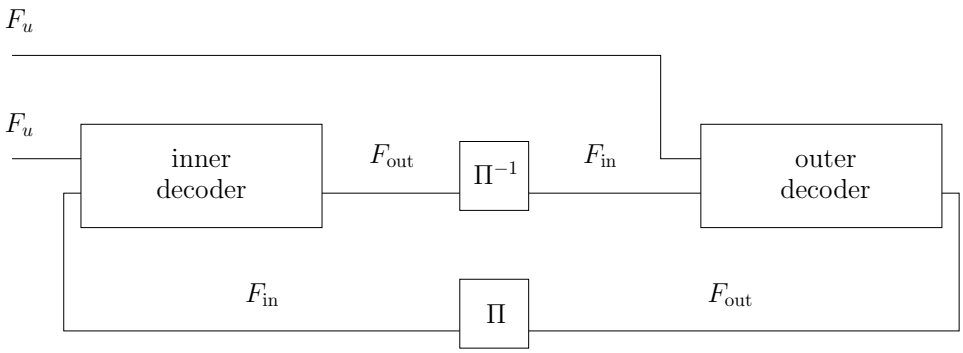


Figure 8.2: Décodeur IRA avec séquencement turbo

8.3.4 Propriétés de l'Evolution de Densités Approximée

Condition de Stabilité (1,1) est un point fixe de la DE approximée, et sa stabilité est donnée par le résultat suivant.

Théorème 8.2 *Le point fixe (1,1) du système (8.24 – 8.25) est stable si et seulement si le point fixe (Δ_∞,Δ_∞) de la DE exacte (8.6 – 8.9) est stable.*

Proposition 8.2 *La condition locale de stabilité de la DE approximée par la méthode 2 est moins contraignante que celle de la DE exacte.*

Points fixes, Rendements et Capacité de Canal Le théorème suivant est un corollaire d'un résultat plus général donné par [42].

Théorème 8.3 *La DE approximée par les méthodes 2 et 4 possède un point fixe unique (1,1) seulement si le rendement R de l'ensemble IRA satisfait l'inégalité $R < C = \mathcal{I}(F_u)$.*

8.3.5 Résultats et Simulations

Des résultats numériques sont donnés dans la section 3.6 pour les canaux BIAWGNC et BSC. Ces résultats montrent que l'approximation Gaussienne est plus intéressante car les codes réalisés avec les méthodes 1 et 3 produisent le plus petit écart à la limite de Shannon. Selon le rendement souhaité, il est préférable de déterminer la fonction EXIT du décodeur interne en utilisant une simulation Monte Carlo (méthode 3) ou l'approximation duale (méthode 1).

Pour le BIAWGNC, les meilleurs codes LDPC [9] ont une meilleure performance que les codes IRA. Néanmoins, les codes IRA représentent une alternative intéressante au vu de la simplicité de leur encodage et décodage.

8.4 Codes Répétition-Accumulation de Longueur Finie

Les graphes de Tanner des ensembles de codes IRA optimisés ne contiennent aucun cycle, dans la limite d'une longueur de code infinie. Par contre, les graphes de codes dont la longueur est finie peuvent contenir des cycles courts. Par conséquent, le décodeur BCJR est sous-optimal car la l'indépendance locale des messages n'est plus valide. En effet, des codes IRA construits de façon aléatoire peuvent avoir de mauvaises performances lors du décodage somme-produit. Ces mauvaises performances se matérialisent comme suit:

1. Dans la région de bas SNR, le taux d'erreur de bits (BER pour *bit error rate*) est loin du seuil de l'ensemble des codes IRA, creusant l'écart à la limite de Shannon pour la longueur infinie.

2. Dans la région de SNR moyen et élevé, le BER s'aplatit et creuse d'autant plus l'écart à la limite de Shannon. Ce comportement est appelé *error floor*.

3. Si $\lambda_2 \neq 0$, le taux d'erreur de mots (WER pour *word error rate*) est mauvais.

8.4.1 Entrelaceurs

La performance de codes IRA de longueur finie peut être améliorée en construisant l'entrelaceur de façon appropriée. Deux stratégies sont adoptées.

La première stratégie consiste à éliminer des cycles courts [60, 61, 62] en utilisant l'algorithme *progressive edge growth* ou PEG [60]. Cette approche a pour effet de maximiser la longueur du plus petit cycle, appelée *girth*. La deuxième stratégie consiste à maximiser la longueur du plus petit ensemble bloquant [27] en utilisant l'algorithme *SSMAX* [63].

8.4.2 Le Girth

Le girth de codes IRA construits avec l'algorithme PEG est comparable à la limite théorique pour des longueurs de codes courtes, mais l'écart se creuse quand la longueur de bloc augmente. La borne théorique sur le girth de codes IRA est donnée par

$$g \leq 4\frac{\log\left[k(\bar{d}-1)+1\right]}{\log\left[\frac{a+1}{a+d}(\bar{d}a+\bar{d}-a)\right]} \tag{8.29}$$

8.4.3 Décodage au Maximum de Vraisemblance

Afin de déterminer l'origine de la mauvaise performance des codes IRA (qui du décodeur somme-produit ou de la famille d'entrelaceurs est en cause), on calcule une borne supérieure au taux d'erreur sous le décodage au maximum de vraisemblance (ML pour *maximum likelihood*), à l'aide du TSB (*tangential sphere bound*). A cet effet, les IOWE (pour *input output weight enumerator* ou *énumérateurs de poids entrée – sortie*) du code de répétition, du groupement et de l'accumulateur sont calculés individuellement. En considérant un code RA régulier avec les paramètres $a = 2$ et $a = 4$, les IOWE du code peuvent être obtenus analytiquement, et sont respectivement donnés par

$$A_{w,w+h} = \frac{\binom{k}{w}}{\binom{dk}{dw}} \sum_t 2^t \binom{m-t}{\frac{dw-t}{2}} \binom{m-h}{\lfloor t/2 \rfloor} \binom{h-1}{\lceil t/2 \rceil - 1} \tag{8.30}$$

$$A_{w,w+h} = \frac{\binom{k}{w}}{\binom{dk}{dw}} \sum_t 2^{2t} \binom{m-h}{\lfloor t/2 \rfloor} \binom{h-1}{\lceil t/2 \rceil - 1} \sum_{l=0}^{l_{max}} 2^{2l} \binom{m-t}{l} \binom{2m-2l-t}{\frac{dw-t}{2}-l} \tag{8.31}$$

où

$$\begin{aligned} w &= 0, \ldots, k \\ h &= 0, \ldots, m \\ l_{max} &= \min(m-t, \tfrac{dw-t}{2}) \end{aligned}$$

où les sommes sont sur tous les entiers t tels que $dw - t$ est pair et $t \leq \min(m, dw)$.

8.4.4 Résultats et Simulations

La méthode d'impulsion [75, 76] est utilisée pour estimer la distance minimale de codes RA réguliers de petite longueur. Des exemples pour des codes RA réguliers de rendement 1/2 montrent que la distance minimale augmente proportionnellement avec le girth.

Nos simulations montrent que la maximisation du girth améliore les performances en BER et WER des codes RA réguliers de petite et moyenne longueur par rapport à l'ensemble RA régulier aléatoire. En effet, l'augmentation du girth a un impact direct sur la distance minimale, résultant en de meilleures performances par rapport à l'ensemble aléatoire et à l'ensemble de codes LDPC de même girth. En conclusion, le compromis performance/complexité des codes RA réguliers est avantageux quand la longueur de bloc est petite à moyenne.

Les codes RA irréguliers de grande longueur ont une meilleure performance avec l'algorithme SSMAX, en comparaison avec les performances de l'ensemble IRA aléatoire et de l'ensemble IRA construit avec l'algorithme PEG. Néanmoins, la performance des codes IRA par SSMAX reste inférieure à celle de codes LDPC construits de la même façon.

CDMA Codé avec Décodage Successif

8.5 Efficacité Spectrale de CDMA codé

8.5.1 Modèle Canonique CDMA

Le canal CDMA complexe discret est modélisé par

$$\mathbf{y}_i = \mathbf{S}\mathbf{x}_i + \mathbf{n}_i, \qquad i = 1, \ldots, n \tag{8.32}$$

obtenu en échantillonnant un système CDMA au rythme de bribe [29], où:

1. $\mathbf{y}_i, \mathbf{n}_i \in \mathbb{C}^N$ sont respectivement le vecteur reçu échantillonné au rythme de la bribe et le vecteur d'échantillons de bruit $\sim \mathcal{N}_{\mathbb{C}}(0,1)$ correspondants;

2. $\mathbf{S} \in \mathbb{C}^{N\times K}$ contient les séquences d'étalement ordonnées par colonne. Les séquences d'étalement sont complexes, connues au récepteur, et les bribes sont iid caractérisées par une moyenne nulle, une variance $1/N$ et un moment fini du quatrième ordre;

3. $\mathbf{x}_i \in \mathbb{C}^K$ est le vecteur des symboles modulés transmis à l'instant i, où le k-ème élément $x_{k,i}$ prend sa valeur dans une certaine constellation et dont l'énergie moyenne (par N bribes) est $E[|x_{i,k}|^2] = \alpha_k$ SNR. Le facteur d'échelle α_k représente un contrle de puissance tel que $\frac{1}{K}\sum_{k=1}^{K}\alpha_k = 1$;

4. N, K et n sont respectivement le facteur d'étalement, le nombre d'utilisateurs et la longueur de bloc.

Les K utilisateurs appartiennent à L classes. Chaque classe j possède k_j utilisateurs qui ont SNR γ_j. Sans perte de généralité, on suppose que $\gamma_1 \leq \cdots \leq \gamma_L$. La *charge d'une classe* j est donnée par $\beta_j = K_j/N$, et la charge totale du canal est

$$\beta = \sum_{j=1}^{L}\beta_j \quad \text{utilisateurs/bribe}$$

.

L'efficacité spectrale totale du système est

$$\rho = \sum_{j=1}^{L} \beta_j R_j \ \text{ bit/s/Hz} \tag{8.33}$$

où R_j est le rendement des utilisateurs de la classe j.

Les rapports signal à bruit par bit des utilisateurs individuels étant généralement différents, on définit un E_b/N_0 pour le système par

$$\left(\frac{E_b}{N_0}\right)_{\text{sys}} \triangleq \frac{\sum_{j=1}^{L} \beta_j \gamma_j}{\sum_{j=1}^{L} \beta_j R_j} = \frac{\sum_{j=1}^{L} \beta_j \gamma_j}{\rho} \tag{8.34}$$

8.5.2 Efficacité Spectrale du Canal CDMA Aléatoire

En l'absence d'étalement, le canal CDMA se réduit à un canal à bruit blanc Gaussien additif dont la capacité est donnée par la solution de

$$\mathsf{C}^\star = \log_2(1 + \mathsf{C}^\star \frac{E_b}{N_0}) \tag{8.35}$$

L'efficacité spectrale d'un système CDMA, avec étalement aléatoire, dans la limite d'une taille infinie ($K, N \to \infty$ avec $K/N = \beta < \infty$, sous la contrainte d'un évanouissement, est donnée par [32]

$$\mathsf{C}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathsf{C}^{\text{mmse}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) + \log_2 \frac{1}{\eta} + (\eta - 1) \log_2 e \tag{8.36}$$

où

- $\boldsymbol{\beta} \triangleq (\beta_1, \ldots, \beta_L)$ et $\boldsymbol{\gamma} \triangleq (\gamma_1, \ldots, \gamma_L)$.
- η est l'efficacité multi-utilisateur grand système (non asymptotique) [29] du récepteur MMSE linéaire. Elle est donnée par la solution positive de l'équation de Tse-Hanly [91]:

$$\eta = \left(1 + \sum_{j=1}^{L} \beta_j \frac{\gamma_j}{1 + \eta \gamma_j}\right)^{-1} \tag{8.37}$$

- $\mathsf{C}^{\text{mmse}}(\boldsymbol{\beta}, \boldsymbol{\gamma})$ est l'efficacité spectrale maximale pouvant être atteinte en utilisant un banc de filtres linéaires MMSE suivis d'un décodeur pour chaque utilisateur. $\mathsf{C}^{\text{mmse}}(\boldsymbol{\beta}, \boldsymbol{\gamma})$ est donnée par:

$$\mathsf{C}^{\text{mmse}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{j=1}^{L} \beta_j \log_2(1 + \gamma_j \eta) \tag{8.38}$$

Le décodeur composé de filtres MMSE suivis de décodeurs mono-utilisateurs est clairement sous-optimal. En effet, le maximum de l'efficacité spectrale C^{mmse} est atteint pour $\beta_{\text{opt}}^{\text{mmse}} < \infty$ avec

$$\mathsf{C}^{\text{mmse}}(\beta_{\text{opt}}^{\text{mmse}}, \gamma) < \mathsf{C}^{\star}$$

où $\mathsf{C}^{\star}$ est la capacité du canal AWGN au $\frac{E_b}{N_0}$ correspondant.

D'un autre cté, le supremum de C est atteint pour $\beta \to \infty$, et concide avec la capacité du canal AWGN. L'efficacité spectrale $\mathsf{C}(\beta, \gamma)$ peut être atteinte par un décodeur qui combine la soustraction successive d'interférences et des filtres MMSE.

8.5.3 Approche de Faible Complexité avec QPSK

Nous avons montré dans la première partie de cette thèse qu'il existe des codes binaires dont le seuil de décodage approche la capacité du canal, tout en ayant un décodeur à faible complexité. Nous proposons donc de faire usage de familles de codes caractérisées par les paires rendement – seuil (R, g).

La capacité d'un canal Gaussien dont l'entrée est QPSK est donnée par

$$\begin{aligned} C_{\text{qpsk}}(\,\mathsf{SNR}) &= 2\left(1 - \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \log_2\left(1 + e^{-2\,\mathsf{SNR} - 2\sqrt{2\,\mathsf{SNR}}v}\right) e^{-v^2} dv\right) \\ &= 2J(2\,\mathsf{SNR}) \end{aligned} \tag{8.39}$$

Dans la limite d'un grand système à accès multiple, l'efficacité spectrale atteinte par un décodeur successif est donnée par

$$\mathsf{C}_{\text{qpsk}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{j=1}^{L} \int_0^{\beta_j} C_{\text{qpsk}}(\gamma_j \eta_j(z)) dz \tag{8.40}$$

Pour un $\frac{E_b}{N_0}$ donné, l'écart entre $\mathsf{C}_{\text{qpsk}}(\beta, \gamma)$ et $\mathsf{C}^{\star}$ disparaît quand $\beta \to \infty$. En effet, le résultat suivant montre que l'usage d'une modulation QPSK est optimal si la charge du canal est infinie.

Théorème 8.4 *Considérons*

$$C_Q(\beta,\ \mathsf{SNR}) = \int_0^\beta E[C_{\mathrm{qpsk}}(|\mathsf{A}|^2\ \mathsf{SNR}\,\eta(z,\ \mathsf{SNR}))]dz \tag{8.41}$$

où $\eta(z,\ \mathsf{SNR})$ *est la solution de*

$$\eta + zE\left[\frac{\mathsf{SNR}|\mathsf{A}|^2\eta}{1+\ \mathsf{SNR}|\mathsf{A}|^2\eta}\right] = 1 \tag{8.42}$$

où $|\mathsf{A}| \sim F_{|\mathsf{A}|}$.

Pour β *et* $\frac{E_b}{N_0}$ *donnés, on définit*

$$\mathsf{C}_Q(\beta, \frac{E_b}{N_0}) = C_Q(\beta,\ \mathsf{SNR}) \tag{8.43}$$

pour le SNR *satisfaisant*

$$\frac{E_b}{N_0} C_Q(\beta,\ \mathsf{SNR}) = \beta\ \mathsf{SNR} \tag{8.44}$$

Donc, pour tout $\frac{E_b}{N_0} \geq \log_e 2$*:*

$$\lim_{\beta\to\infty} \mathsf{C}_Q(\beta, \frac{E_b}{N_0}) = \mathsf{C}^\star \tag{8.45}$$

8.6 Approche de l'Optimum avec une Complexité Réduite

Afin d'approcher l'efficacité spectrale optimale du GMAC avec une complexité raisonnable, nous considérons un système CDMA avec les caractéristiques suivantes:

- une constellation d'entrée QPSK au lieu d'une entrée Gaussienne;
- des codes binaires approchant (ou atteignant) la limite de Shannon;
- un décodeur de canal itératif de faible complexité (linéaire en la longueur de bloc);
- un décodeur successif qui soustrait les interférences classe par classe.

Nous optimisons l'efficacité spectrale dans la limite d'un grand système $(K, N, n \to \infty)$, dans deux cas de figure:

Système à rendement égal Tous les utilisateurs ont le même rendement R, mais chaque classe j a un niveau de puissance différent γ_j tel que $\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_L$.

Système à puissance égale Tous les utilisateurs ont la même puissance γ, mais chaque classe j a un rendement différent R_j tel que $R_1 \geq R_2 \geq \cdots \geq R_L$.

8.6.1 Optimisation d'un Système à Rendement Égal

Une famille de codes est caractérisée par une paire rendement/seuil (R, g). Pour une efficacité spectrale $\rho = \beta R$, le vecteur optimal $\boldsymbol{\beta}$ qui atteint un BER arbitrairement petit à un $E_/N_0$ minimal, est la solution du programme linéaire suivant:

$$\begin{cases} \text{minimiser} & \displaystyle\sum_{i=1}^{L} \beta_i \gamma_i \\ \text{sous les contraintes} & \mathbf{A}\boldsymbol{\beta} \leq \mathbf{b} \\ & \displaystyle\sum_{i=1}^{L} \beta_i \geq \beta \\ & \boldsymbol{\beta} \geq \mathbf{0} \end{cases} \tag{8.46}$$

où $\mathbf{A}$ est une matrice triangulaire dont les éléments sont

$$a_{i,j} = \begin{cases} 0 & \text{si } \ i < j \\ \frac{(1+g)\gamma_j}{\gamma_i + \gamma_j g} \in (0,1] & \text{si } \ i \geq j \end{cases} \tag{8.47}$$

et $\mathbf{b}$ est un vecteur de longueur L dont les éléments sont

$$b_i = \frac{(1+g)(\gamma_i - g)}{\gamma_i g} \tag{8.48}$$

La solution de ce problème est donnée par le résultat suivant.

Proposition 8.3 *L'équation* $\mathbf{Ax} = \mathbf{b}$ *a une solution unique avec des éléments non négatifs* $\boldsymbol{\tau}$*. L'ensemble de solutions faisables de (8.46) est non vide si et*

seulement si $\beta \leq \sum_{j=1}^{L} \tau_j$. *La solution de (8.46) est donc explicitement donnée par*

$$\beta_i^\star = \begin{cases} \tau_i, & i = 1, \ldots, \hat{L} - 1 \\ \beta - \sum_{j=1}^{\hat{L}-1} \tau_j, & i = \hat{L} \\ 0, & i = \hat{L} + 1, \ldots, L \end{cases} \tag{8.49}$$

8.6.2 Optimisation d'un Système à Puissance Égale

Les familles de codes sont caractérisées par des paires rendements/seuils (R_j, g_j), et tous les utilisateurs ont le même SNR γ. L'efficacité spectrale $\rho = \sum_{i=1}^{L} \beta_i R_i$ maximisée sur l'ensemble des charges par classe est obtenue comme la solution du programme linéaire suivant:

$$\begin{cases} \text{maximiser} & \sum_{i=1}^{L} \beta_i R_i \\ \text{sous les contraintes} & \mathbf{L}\boldsymbol{\beta} \leq \mathbf{b} \\ & \sum_{i=1}^{L} \beta_i \leq \beta \\ & \boldsymbol{\beta} \geq \mathbf{0} \end{cases} \tag{8.50}$$

où $\mathbf{L}$ est une matrice triangulaire dont les éléments sont

$$l_{i,j} = \begin{cases} 0 & \text{si } i < j \\ 1 & \text{si } i \geq j \end{cases} \tag{8.51}$$

et $\mathbf{b}$ est un vecteur de longueur L dont les éléments sont

$$b_i = \frac{(1 + g_i)(\gamma - g_i)}{\gamma g_i} \tag{8.52}$$

La solution de ce programme linéaire est donnée explicitement par le résultat suivant.

Proposition 8.4 *Le problème (8.50) a toujours une solution*

$$\beta_i^\star = \begin{cases} b_i - b_{i-1}, & i = 1, \ldots, \hat{L} - 1 \\ \beta - b_{\hat{L}-1}, & i = \hat{L} \\ 0, & i = \hat{L} + 1, \ldots, L \end{cases} \tag{8.53}$$

où $b_0 \triangleq 0$*, et* $\hat{L}$ *dénote le* i *minimal pour lequel* $\beta \leq b_i$*.*

8.6.3 Résultats et Simulations

Les codes binaires ont été choisis parmi les familles de codes LDPC ainsi que les familles de codes QPSK optimaux. L'efficacité spectrale du système optimisé par égalité de rendement est très proche de l'optimal, pour de petits rendements, au prix d'un très grand nombre de classes à différents niveaux de puissance. Par contre, l'efficacité spectrale obtenue par un système à puissance égale reste éloignée de l'optimum, car il faut considérer un très grand nombre de rendements finement espacés, que les codes soient LDPC ou QPSK.

Les résultats d'optimisation peuvent être utilisés pour dimensionner un système fini. Le décodeur successif peut être itéré afin d'éliminer toute erreur résiduelle due à un BER non nul.

8.7 Conclusion

Cette thèse a proposé différents schémas de codage/décodage, de complexité réduite, permettant d'approcher la capacité de canaux à entrée binaire et sortie symétrique ainsi que le canal CDMA.

Nous avons d'abord considéré l'ensemble de codes répétition-accumulation irréguliers IRA aléatoires, systématiques et de longueur de bloc infinie. A l'aide de la méthode d'évolution des densités, nous avons analysé le décodeur par propagation des croyances. Nous avons suivi l'évolution des messages sur le graphe de Tanner (qui ne contient pas de cycles) résultant en un système dynamique bidimensionnel. Nous avons étudié la stabilité locale des points fixes de ce système, correspondant au point BER nul.

Nous nous sommes attelés à l'optimisation de l'ensemble de codes IRA pour la classe de canaux à entrée binaire et sortie symétrique. A cet effet, nous avons utilisé les outils suivants: l'évolution de l'information mutuelle décrite par les fonctions EXIT, l'approximation duale, l'approximation Gaussienne, l'approximation BEC et la convexité non stricte de l'information mutuelle sur l'ensemble des canaux à entrée binaire et sortie symétrique. Nous avons proposé quatre méthodes afin d'approximer les densités impliquées dans l'évolution de densités par un paramètre réel. Les quatre méthodes permettent d'écrire le problème d'optimisation de codes IRA sous la forme de programmes linéaires. Une fois de plus, la condition de stabilité locale des points fixes des approximations a été dérivée. Nous avons optimisé les codes IRA pour un large éventail de rendements et avons montré des exemples pour le canal Gaussien et le canal BSC. Les meilleures approximations en termes de seuil de décodage sont celles basées sur l'approximation Gaussienne. Les performances des codes IRA optimisés sont comparables à celles des meilleurs codes LDPC.

Dans un deuxième temps, nous avons étudié les performances de codes IRA de longueur finie avec différentes structures d'entrelacement. Deux approches ont été adoptées: la maximisation du girth et la maximisation de la taille des ensembles bloquants. La performance des codes RA réguliers de petite et moyenne longueur est comparable à celle des meilleurs codes LDPC. La performance des codes IRA de longueur importante est inférieure à celle des meilleurs codes LDPC.

Nous nous sommes ensuite intéressés à approcher l'efficacité spectrale optimale du canal GMAC avec une complexité réduite. Nous avons montré que la perte en efficacité spectrale occasionnée par l'utilisation d'une constella-

tion QPSK et d'un système CDMA aléatoire au lieu d'une entrée Gaussienne disparaît en augmentant la charge du canal.

Nous avons montré un schéma de codage/décodage de complexité réduite, approchant l'efficacité spectrale optimale. Ce schéma est basé sur l'utilisation de codes binaires approchant (à défaut d'atteindre) la capacité mono-utilisateur, une modulation QPSK et un décodeur successif qui effectue une soustraction successive des interférences. Nous avons optimisé la distribution de la charge du canal dans deux cas: les utilisateurs transmettent avec le même rendement à différents niveaux de puissance; les utilisateurs transmettent avec la même puissance à différents rendements. La contrainte d'optimisation est la satisfaction des conditions de décodage successif. Dans le schéma de rendement égal, l'efficacité spectrale est proche de l'optimum pour des rendements bas. Ce schéma semble plus intéressant que celui qui emploie des rendements différents car ce dernier requiert un très grand nombre de rendements finement espacés afin d'approcher l'optimum. Des résultats numériques montrent que ces méthodes peuvent être utilisées pour dimensionner des systèmes CDMA de taille finie employant des codes binaires tirés de la littérature (LDPC, IRA). Le système ainsi construit est exempt de propagation catastrophique d'erreurs, en permettant que le décodeur successif soit lui-même itératif.

Bibliography

[1] C.E. Shannon. A mathematical theory of communications. *Bell System Technical Journal*, 27:379–423 and 623–656, Jul. and Oct. 1948.

[2] G.D. Forney Jr. *Concatenated Codes*. PhD thesis, Cambridge, MA: MIT Press, 1966.

[3] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon limit error-correcting and decoding: Turbo codes. In *Proceedings of IEEE International Conference on Communication ICC*, pages 1064–1070, Geneva, May 1993.

[4] L.R. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Transactions on Information Theory*, pages 284–287, Mar. 1974.

[5] R.G. Gallager. *Low-Density Parity Check Codes*. PhD thesis, MIT Press, Cambridge, MA, 1963.

[6] M. Luby, M. Mitzenmacher, A. Shokrollahi, and D. Spielman. Practical loss-resilient codes. In *Proceedings of the* 29^{th} *ACM Symposium on Theory of Computing*, pages 150–159, 1997.

[7] M. Luby, M. Mitzenmacher, A. Shokrollahi, and D. Spielman. Efficient erasure correcting codes. *IEEE Transactions on Information Theory*, 47(2):569–584, Feb. 2001.

[8] D. Divsalar, H. Jin, and R. J. McEliece. Coding theorems forv "turbo-like" codes. In *Proceedings of Allerton Conference Comm. Control and Comput.*, pages 201–210, Urbana-Champaign, 1998.

[9] R. Urbanke et al. Web page. http://lthcwww.epfl.ch/research/ldpcopt/, 2002.

[10] N. Varnica and A. Kavcic. Optimized LDPC codes for partial response channels. In *Proceedings of IEEE International Symposium on Information Theory ISIT*, page 197, Lausanne, Switzerland, Jul. 2002.

[11] X. Ma, N. Varnica, and A. Kavcic. Matched information rate codes for binary ISI channels. In *Proceedings of IEEE International Symposium on Information Theory ISIT*, page 269, Lausanne, Switzerland, Jul. 2002.

[12] N. Varnica and A. Kavcic. Optimized low density parity check codes for partial response channels. *IEEE Communications Letters*, 2003.

[13] B.M. Kurkoski, P.H. Siegel, and J.K. Wolf. Joint message-passing decoding of LDPC codes and partial-response channels. *IEEE Transactions on Information Theory*, 48(6):1410–1422, Jun. 2002.

[14] R. M. Tanner. A recursive approach to low complexity codes. *IEEE Transactions on Information Theory*, IT-27:533–547, 1981.

[15] H. Jin, A. Khandekar, and R. McEliece. Irregular repeat-accumulate codes. In *Proceedings of International Symposium on Turbo Codes*, pages 1–8, Brest-France, Sep. 2000.

[16] J. Boutros, G. Caire, E. Viterbo, H. Sawaya, and S. Vialle. Turbo code at 0.03 dB from capacity limit. In *Proceedings of IEEE International Symposium on Information Theory ISIT*, page 56, Lausanne, Switzerland, Jul. 2002.

[17] B.J. Frey and D. MacKay. Irregular turbo codes. In *Proceedings of IEEE International Symposium on Information Theory ISIT*, page 121, Sorrento, Italy, Jun. 2000.

[18] M. Luby, M. Mitzenmacher, A. Shokrollahi, and D. Spielman. Analysis of low-density codes and improved designs using irregular graphs. In *Proceedings of the* 30^{th} *ACM Symposium on Theory of Computing*, pages 249–258, 1998.

[19] T.J. Richardson, M.A. Shokrollahi, and R.L. Urbanke. Design of capacity-approaching irregular low-density parity-check codes. *IEEE Transactions on Information Theory*, 47:619 – 637, Feb 2001.

[20] S. ten Brink. Designing iterative decoding schemes with the extrinsic information transfer chart. *AEÜ Int. J. Electronic. Commun.*, 54(6):389–398, Dec. 2000.

[21] S. ten Brink. Convergence behavior of iteratively decoded parallel concatenated codes. *IEEE Transactions on Communications*, 49(10):1727–1737, Oct. 2001.

[22] Sae-Young Chung, T.J. Richardson, and R.L. Urbanke. Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation. *IEEE Transactions on Information Theory*, 47:657 – 670, Feb 2001.

[23] H. El Gamal and A.R. Hammons. Analyzing the turbo decoder using the Gaussian approximation. *IEEE Transactions on Information Theory*, 47(2):671–686, Feb. 2001.

[24] J. Boutros and G. Caire. Iterative multiuser joint decoding: Unified framework and asymptotic analysis. *IEEE Transactions on Information Theory*, 48(7):1772–1793, Jul. 2002.

[25] F. Lehmann and G.M. Maggio. An approximate analytical model of the message passing decoder of LDPC codes. In *Proceedings of IEEE International Symposium on Information Theory ISIT*, page 31, Lausanne, Switzerland, Jul. 2002.

[26] M. Ardakani and F.R. Kschischang. Designing irregular LPDC codes using EXIT charts based on message error rate. In *Proceedings of IEEE International Symposium on Information Theory ISIT*, page 454, Lausanne, Switzerland, Jul. 2002.

[27] C. Di, D. Proietti, E. Telatar, T. Richardson, and R. Urbanke. Finite length analysis of low-density parity-check codes on the binary erasure channel. *IEEE Transactions on Information Theory*, 48(6):1570–1579, Jun. 2002.

[28] A. El Gamal and T. Cover. Multiple user information theory. *Proceedings of the IEEE*, 68:1463–1483, Dec. 1980.

[29] S. Verdú. *Multiuser Detection*. Cambridge University Press, Cambridge, UK, 1998.

[30] A. J. Viterbi. *CDMA – Principles of Spread Spectrum Communications*. Addison-Wesley, Reading, MA, 1995.

[31] S. Verdú and S. Shamai. Spectral efficiency of CDMA with random spreading. *IEEE Transactions on Information Theory*, 45(2):622–640, Mar. 1999.

[32] S. Verdú and S. Shamai. The impact of frequency-flat fading on the spectral efficiency of CDMA. *IEEE Transactions on Information Theory*, 47(4):1302–1327, May 2001.

[33] E. Biglieri, G. Caire, and G. Taricco. CDMA system design through asymptotic analysis. *IEEE Transactions on Communications*, 48(11):1882 – 1896, Nov. 2000.

[34] A.J. Viterbi and J.K. Omura. *Principles of Digital Communication and Coding*. McGraw-Hill Publishing Company, 1979.

[35] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman, San Mateo, CA, 1988.

[36] R. J. McEliece, D.J.C. MacKay, and J-F Cheng. Turbo decoding as an instance of Pearl's "belief propagation" algorithm. *IEEE Journal on Selected Areas in Communications*, 16:140–152, Feb. 1998.

[37] F.R. Kschischang and B.J. Frey. Iterative decoding of compound codes by probability propagation in graphical models. *IEEE Journal on Selected Areas in Communications*, 16:219–230, Feb. 1998.

[38] J. Hagenauer, E. Offer, and L. Papke. Iterative decoding of binary block and convolutional codes. *IEEE Transactions on Information Theory*, 42(2):429–445, Mar. 1996.

[39] T.J. Richardson and R.L. Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Transactions on Information Theory*, 47(2):599–618, Feb. 2001.

[40] D. Forney. Codes on graphs: Normal realizations. *IEEE Transactions on Information Theory*, 47(2):520–548, Feb. 2001.

[41] A. Ashikhmin, G. Kramer, and S. ten Brink. Extrinsic information transfer functions: A model and two properties. In *Proceedings of* 36^{th} *Annual Conference on Information Sciences and Systems CISS*, Princeton, New Jersey, Mar. 2002.

[42] A. Ashikhmin, G. Kramer, and S. ten Brink. Code rate and the area under extrinsic information transfer curves. In *Proceedings of IEEE International Symposium on Information Theory ISIT*, page 115, Lausanne, Switzerland, Jun. 30^{th} – Jul. 5^{th} 2002.

[43] Sae-Young Chung. *On the Construction of Some Capacity-Approaching Coding Schemes*. PhD thesis, MIT, Sep. 2000.

[44] S.Y. Chung, G.D. Jr Forney, T.J. Richardson, and R. Urbanke. On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit. *IEEE Communications Letters*, 5:58–60, Feb. 2001.

[45] S. ten Brink and G. Kramer. Turbo processing for scalar and vector channels. In *Proceedings of International Symposium on Turbo Codes*, pages 23–30, Brest, Sep. 2003.

[46] S. ten Brink and G. Kramer. Design of repeat-accumulate codes for iterative detection and decoding. *IEEE Transactions on Signal Processing*, 51(11):2764–2772, Nov. 2003.

[47] H. Jin. *Analysis and Design of Turbo-Like Codes*. PhD thesis, California Institute of Technology, May 2001.

[48] S. ten Brink. Exploiting the chain rule of mutual information for the design of iterative decoding schemes. In *Proceedings of Allerton Conference Comm. Control and Comput.*, Oct. 2001.

[49] M. Tuchler, S. ten Brink, and J. Hagenauer. Measures for tracing convergence of iterative decoding algorithms. In *Proceedings of the* 4^{th} *International ITG Conference on Source and Channel Coding*, Berlin, Germany, Jan. 2002.

[50] S. Huettinger and J. Huber. Extrinsic and intrinsic information in systematic coding. In *Proceedings of IEEE International Symposium on Information Theory ISIT*, page 116, Lausanne, Switzerland, Jul. 2002.

[51] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. John Wiley and Sons, Inc., 1991.

[52] T.F. Wong. Numerical calculation of symmetric capacity of Rayleigh fading channel with BPSK/QPSK. *IEEE Communications Letters*, 5(8):328–330, Aug. 2001.

[53] A. Browder. *Mathematical Analysis: An Introduction*. Springer-Verlag, New York, 1996.

[54] D.J.C. MacKay, S.T. Wilson, and M.C. Davey. Comparison of constructions of irregular Gallager codes. In *Proceedings of Allerton Conference Comm. Control and Comput.*, September 1998.

[55] S. Dolinar and D. Divsalar. Weight distributions for turbo codes using random and nonrandom permutations. JPL TDA Progress Report 42-122, August 15, 1995.

[56] W. Feng, J. Yuan, and B.S. Vucetic. A code-matched interleaver design for turbo codes. *IEEE Transactions on Communications*, 50(6):926–937, Jun. 2002.

[57] P.O. Vontobel. On the construction of turbo code interleavers based on graphs with large girth. In *Proceedings of IEEE International Conference on Communication ICC*, volume 3, pages 1408–1412, Apr. 28 - May 2. 2002.

[58] J. Hokfelt, O. Edfors, and T. Maseng. A turbo code interleaver design criterion based on the performance of iterative decoding. *IEEE Communications Letters*, 5(2):52–54, Feb. 2001.

[59] D. Le Ruyet and H. Vu Thien. Design of cycle optimized interleavers for turbo codes. In *Proceedings of International Symposium on Turbo Codes*, pages 335–338, Brest, 2000.

[60] X. Hu, E. Eleftheriou, and D. Arnold. Progressive edge-growth Tanner graphs. In *Proceedings of IEEE Global Telecommunication Conference, Globecom*, volume 2, pages 995–1001, 25-29 Nov. 2001.

[61] Y. Mao and A. H. Banihashemi. A heuristic for good low-density parity-check codes at short block lengths. In *Proceedings of IEEE International Conference on Communication ICC*, volume 1, pages 41–44, 11-14 Jun. 2001.

[62] J. A. McGowan and R. C. Williamson. Loop removal from LDPC codes. In *Proceedings of Information Theory Workshop ITW*, pages 230–233, Paris, France, Mar. 31 - Apr. 4 2003.

[63] T. Tian, C. Jones, J. D. Villasenor, and R. D. Wesel. Construction of irregular LDPC codes with low error floors. In *Proceedings of IEEE International Conference on Communication ICC*, volume 5, pages 3125–3129, 2003.

[64] R. M. Tanner. Minimum distance bounds by graph analysis. *IEEE Transactions on Information Theory*, 47:808–821, Feb. 2001.

[65] S. Hoory. The size of bipartite graphs with a given girth. *Journal of Combinatorial Theory - Series B*, 86(2):215–220, 2002.

[66] H. Herzberg and G. Poltyrev. The error probability of m-ary PSK block coded modulation schemes. *IEEE Transactions on Communications*, 44(4):427–433, Apr. 1996.

[67] G. Poltyrev. Bounds on the decoding error probability of binary linear codes via their spectra. *IEEE Transactions on Information Theory*, 40(4):1284–1292, Jul. 1994.

[68] I. Sason and S. Shamai. Improved upper bounds on the ML decoding error probability of parallel and serial concatenated turbo codes via their ensemble distance spectrum. *IEEE Transactions on Information Theory*, 46(1):24–47, Jan. 2000.

[69] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara. Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding. *IEEE Transactions on Information Theory*, 44(3):909–926, May 1998.

[70] D. L. Kreher and D. R. Stinson. *Combinatorial Algorithms- Generation, Enumeration and Search*. CRC Press, 1999.

[71] R. A. Brualdi. *Introductory Combinatorics*. Prentice Hall, 3^{rd} edition, 1999.

[72] D. Burshtein and G. Miller. Asymptotic enumeration methods for analyzing LDPC codes. Submitted to IEEE Trans. on Information Theory, 2002.

[73] N. Kahale and R.Urbanke. On the minimum distance of parallel and serially concatenated codes. In *Proceedings of IEEE International Symposium on Information Theory ISIT*, page 31, 16-21 Aug. 1998.

[74] D.J.C. MacKay. Online database of low-density parity check codes. http://wol.ra.phy.cam.uk/mackay/codes/data.html.

[75] C. Berrou, S. Vaton, M. Jezequel, and C. Douillard. Computing the minimum distance of linear codes by the error impulse method. In *Proceedings of IEEE International Symposium on Information Theory ISIT*, page 5, Lausanne, Switzerland, Jul. 2002.

[76] G. Zemor, G. Cohen, J. Boutros, C. Berrou, R. Pyndiah, M. Jezequel, and H. Gonzalez. IT: Décodage itératif pour les codes de longueur moyenne. Technical report, GET, Direction Scientifique, 2003.

[77] C. Berrou. Some clinical aspects of turbo codes. In *Proceedings of International Symposium on Turbo Codes*, pages 26–31, Brest - France, 1997.

[78] F.D. Neeser and J.L. Massey. Proper complex random processes with applications to information theory. *IEEE Transactions on Information Theory*, 39(4):1293–1302, Jul. 1993.

[79] T. Cover. Multiple user information theory for the Gaussian channel. In *New Concepts in Multi-User Communication*, number 43 in E:Applied Sciences, pages 53–61. NATO Advance Study Institute Series, J.K. Skwirzynski edition, 1981.

[80] A.D. Wyner. Recent results in the Shannon theory. *IEEE Transactions on Information Theory*, IT-20(2-10), Jan. 1974.

[81] B. Rimoldi and R. Urbanke. A rate splitting approach to the Gaussian multiple-access channel. *IEEE Transactions on Information Theory*, 42(2):364–375, Mar. 1996.

[82] S. Verdú. Capacity region of Gaussian CDMA channels: The symbol-synchronous case. In *Proceedings of Allerton Conference Comm. Control and Comput.*, pages 1025–1034, Oct. 1986.

[83] M. Varanasi and T. Guess. Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple access channel. In *Proceedings of Asilomar Conference on Signals, Systems and Computers*, Nov. 1997.

[84] M. Rupf and J.L. Massey. Optimum sequence multisets for synchronous code-division multiple-access channels. *IEEE Transactions on Information Theory*, 40(4):1261–1266, Jul. 1994.

[85] P. Viswanath and V. Anantharam. Optimal sequences and sum capacity of synchronous CDMA systems. *IEEE Transactions on Information Theory*, 45(6):1984–1991, Sep. 1999.

[86] P. Viswanath and V. Anantharam. Optimal sequences, power control and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers. *IEEE Transactions on Information Theory*, 45(6):1968, Sep. 1999.

[87] W. Yu, W. Rhee, S. Boyd, and J. Cioffi. Iterative water-filling for Gaussian vector multiple access channels. Submitted to IEEE Transactions on Information Theory, Apr. 2001.

[88] 3GPP. TS 25.224 v3.1.0,"3GPP-TSG-RAN-WG1; physical layer procedures (FDD)". ETSI, Dec. 1999.

[89] T. Tanaka. A statistical mechanics approach to large-system analysis of CDMA multiuser detectors. *IEEE Transactions on Information Theory*, 48(11):2888–2910, Nov. 2002.

[90] S. Verdú. Spectral efficiency in the wideband regime. *IEEE Transactions on Information Theory*, 48(6):1319–1343, Jun. 2002.

[91] D.N.C. Tse and S.V. Hanly. Linear multiuser receivers: Effective interference, effective bandwidth and user capacity. *IEEE Transactions on Information Theory*, 45(2):641–657, Mar. 1999.

[92] *Special issue on codes on graphs and iterative algorithms*, volume 47. IEEE Transactions on Information Theory, Feb. 2001.

[93] *Capacity approaching codes, iterative decoding algorithms, and their applications*, volume 41, pages 100–140. IEEE Communications Magazine, Aug. 2003.

[94] E. Chong, J. Zhang, and D. Tse. Output MAI distribution of linear MMSE multiuser receivers in DS-CDMA systems. *IEEE Transactions on Information Theory*, 47(3):1128–1144, Mar. 2001.

[95] G. Caire and R. Müller. The optimal received power distribution for IC-based iterative multiuser joint decoders. In *Proceedings of Allerton Conference Comm. Control and Comput.*, Oct. 2001.

[96] S. Benedetto and G. Montorsi. Unveiling turbo codes: Some results on parallel concatenated coding schemes. *IEEE Transactions on Information Theory*, 42(2):409–428, Mar. 1996.

[97] L. Li, A. Tulino, and S. Verdú. Asymptotic eigenvalue moments for linear multiuser detection. *Communications in Information and Systems*, 1:273–304, Fall 2001.

[98] P. Varaiya. Lecture notes on optimization - web page. http://paleale.eecs.berkeley.edu/∼varaiya/papers_ps.dir/NOO.pdf.

[99] J. Edmonds. Submodular functions, matroids, and certain polyhedra. In *Proc. of Calgary International Conference on Combinatorial Structures and Their Applications*, pages 69–87, New York, 1969. Gordon and Breach.

[100] A. D. Wyner. Shannon-theoretic approach to a Gaussian mutiple-access channel. *IEEE Transactions on Information Theory*, 40(6):1713–1727, Nov. 1994.