Rate Compatible Joint Source-Channel Coding for Point-to-Point and Multiple Access Channels

Rate Compatible Joint Source-Channel Coding for Point-to-Point and Multiple Access Channels

by

Imanol Granada Echeverría

Dissertation Submitted for the Degree of Doctor of Philosophy

Under the supervision of:

Professor Pedro M. Crespo Bofill
Associate Professor Xabier Insausti Sarasola

University of Navarra–TECNUN
Faculty of Telecommunication Engineering
2021

Doctoral Dissertation by the University of Navarra-TECNUN

©Imanol Granada Echeverría, Jun 2021

Printed in the Printing Office, TECNUN
Donostia-San Sebastián

A mi familia; mis padres Rafael y Antxoni y mi hermana Isabel. A los que nos han dejado ya. A mi familia en Madrid a la que espero volver a ver pronto.

A mi futura familia, Virginia.

Acknowledgements

Me gustaría empezar desde el principio y darle las gracias a la gente que me acompañó desde mis inicios como embrión de ingeniero: Mireia, Doncel, Oier, Huete, Andres, (Pa)Pagolazo, Dani, Javoo, Mikel... por nuestras historias pasadas y otras muchas por hacer.

Muchas gracias a Pedro, Xabi, Adam y Jesús, por haber sabido captar mi atención como profesores de matemáticas, procesado de señal y teoría de la información. Especialmente a Pedro por haberme dado la oportunidad de realizar esta Tesis doctoral.

Muchas gracias a mis compañeros de faena Luisvi, Xabi, Camila, Maria, Ane Miren... por muchos cafés y charlas; y alguna que otra tarta y/o tortilla. A mis joven padawans Josu y Patrick, espero que sigáis haciéndolo tan bien en el Quantum, que no os quedéis sin leche en mi ausencia... *$Safemoon to the mooooon!!*. Por último a Fernando, que después de una década por fin te libras de mi.

Muchas gracias también a Javier por toda su ayuda con los artículos y por invitarme a Delaware, ojalá hubiera transcurrido en unas circunstancias mas normales.

Muito obrigado Mariano por me convidar a sua casa para ver um filme no momento mais difcil da pandemia.

Muchas gracias a mi suegra Socorro por alimentarme con buenos pucheros.

Muchas gracias a todos *los Chacales*; y en especial a Germán, Ander, Iker y Pedro por haberme acompaado en mis mejores aventuras de estos años de Tesis.

Summary

In this Thesis we consider high-throughput rate compatible Joint Source-Channel Coding (JSCC) schemes based on Rate Compatible Modulation (RCM) codes. These codes achieve JSCC capabilities by embedding source compression into modulation through the generation of multi-level symbols from weighted linear combinations of the input bits. The smooth rate adaptation is achieved seamlessly by varying the number of generated symbols. These two properties make them advantageous over conventional Adaptive Coded Modulation (ACM) techniques, which usually rely on unrealistic instant and accurate channel estimations, and a limited set of coding and modulation combinations to chose from.

Unfortunately, RCM codes experience performance degradation due to the presence of error floors at high Signal-to-Noise Ratios (SNRs). These error floors can be substantially improved by introducing Low Density Generator Matrix (LDGM) codes in parallel. The main idea is to substitute a few RCM symbols by LDGM coded bits, forming an hybrid coding scheme composed by the parallel concatenation of an analog RCM and a digital LDGM code. The goal of the LDGM symbols is to correct residual errors produced by the RCM. If properly designed, the resulting parallel RCM-LDGM coding scheme achieves good performances while preserving the JSCC and rate adaptation capabilities of RCM codes.

This work investigates new applications and design techniques of these family of codes for point-to-point and multi-user communications. In that regard, this dissertation can be divided in two different parts:

Point-to-point communications.

We begin the first part of the dissertation by proposing an EXIT chart analysis and a bit error rate prediction procedure suitable for implementing RCM-LDGM codes. The developed EXIT charts speed up the design

method of good codes, which otherwise requires the use of time-consuming simulations.

We continue by considering the problem of implementing high-throughput JSCC schemes for the transmission of binary sources with memory over AWGN channels. We propose a coding scheme that makes use of the Burrows-Wheeler Transform (BWT) and the rate compatible RCM-LDGM codes. The BWT is utilized to convert the original source with memory into a set of independent non-uniform discrete memoryless binary sources, that are then separately RCM-LDGM encoded, with optimal rates.

Finally, we consider powerline communications modelled as an additive impulsive noise channel. For the first time in the literature, we propose the use of RCM-LDGM codes for this type of environment. We assume that the channel state information is available at the receiver and adapt the decoder accordingly.

Multi-user communications.

The second part of the Thesis pays attention to the transmission of multiple information sources over non-orthogonal Multiple Access Channels (MACs). We begin by considering that the information sources are uncorrelated and propose a new coding scheme based on the use of an irregular RCM encoder for each user. By properly designing the encoders and taking advantage of the additive nature of the MAC, the proposed scheme allows the simultaneous transmission of a large number of uncorrelated users at high rates, while the decoding complexity is the same as that of standard point-to-point RCM schemes.

In the last part of the dissertation, we study the multi-user communication scenario in which the transmitted information sources are spatially correlated. For this case, we propose another novel coding scheme that is also comprised of a set of RCM codes, but designed differently. In order to optimize performance, these codes are constructed aiming to preserve some source correlation in the codewords. This is achieved by building codes that share the same random structure and whose jointly designed weights differ but achieve positive symbol reinforcement. Finally, we also extend the use of LDGM codes in parallel with the proposed RCM systems in the MAC.

Glossary

A list of the most repeated acronyms is provided below.

RCM	*Rate Compatible Modulation*
LDGM	*Low Density Generator Matrix*
LDPC	*Low-Density Parity-Check*
JSC(C)	*Joint Source-Channel (Coding)*
AMC	*Adaptive Coded Modulation*
AWGN	*Additive White Gaussian Noise*
EXIT	*Extrinsic Information Transfer*
BER	*Bit Error Rate*
PER	*Packet Error Rate*
MAC	*Multiple Access Channel*
CIS	*Correlated Information Sources*
PMF	*Probability Mass Function*
PDF	*Probability Density Function*
MC	*Markov Chain*
HMM	*Hidden Markov Model*
MAP	*Maximum A Posteriori*
SP(A)	*Sum-Product (Algorithm)*
FG	*Factor Graph*
SNR	*Signal-to-Noise Ratio*
LLR	*Log-Likelihood Ratio*
CN(D)	*Check Node (Decoder)*
VN(D)	*Variable Node (Decoder)*
(N)OMA	*(Non) Orthogonal Multiple Access*
MU(D)	*Multi-User (Detection)*
SIC	*Successive Interference Cancellation*
CSI	*Channel State Information*
r.v.	*random variable*
i.i.d.	*independent and identically distributed*

Notation

Although all symbols are defined at their first appearance, some are repeated throughout the dissertation. A list of the most frequent symbols is provided below.

$\mathbb{N}$	The set of natural numbers (positive integers).
$\mathbb{R}$	The set of (finite) real numbers.
$\mathbb{C}$	The set of (finite) complex numbers.
$\mathbb{R}^{m\times p}$	The set of all real matrices of dimension $m \times p$.
$\mathbb{R}^N$	N-dimensional euclidean space.
$\{0, 1\}$	The binary set.
$[\cdot]_{i,j}$	Element on the i^{th} row and on the j^{th} column.
$a \mod b$	The remainder of dividing a with b.
$\mathcal{N}(\cdot\ ,\ \cdot)$	Gaussian distribution.
$\mathbb{E}[\cdot]$	Expectation operator.
$\lvert\cdot\rvert$	Absolute value of a number or carnality of a set.
$\mathcal{C}$	Capacity.
H	Entropy.
$\mathcal{H}$	Entropy rate.
$P_X(x)$	Probability mass function of the discrete random variable X.
$p_X(x)$	Probability density function of the continuous random variable X.
$\hat{a}$	Estimate of a.
$\sum_{\sim a_i}$	The sum over all variables except a_i.
T_j	Check (or function) node j.
$r_{j,k}$	The messages passed from the j^{th} CN to the k^{th} VN.
$q_{k,j}$	The messages passed from the k^{th} VN to the j^{th} CN.
$n(T_j)$	The set of VNs connected to CN T_j.
$n(U_k)$	The set of CNs connected to VN U_k.
$n(U_k) \setminus T_j$	The set of CNs connected to VN U_k without considering CN T_j.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

When Claude Shannon published his groundbreaking work *A Mathematical Theory of Communication* [1], which devised a mathematical framework that characterized the fundamental limits of reliable communications in point-to-point links, it established a new frontier in the field of communications. His manuscript, that founded Information Theory, continues to lay the foundation for the development of communications, data processing and storage, and other information technologies.

He considered the transmission of information generated by a source to a distant receiver connected by a noisy channel and showed that this problem can be separated into two independent sub-problems. First, he characterized the actual amount of information H, denoted as entropy, that the source generates. There are a set of algorithms, known as source compression, that exploit the statistical redundancy of the source symbols to represent the same information in fewer bits. Shannon's theory predicted that the limit of this compression is exactly the entropy, given in bits per source symbol. Second, he established the maximum number of bits per channel symbol that any transmitter can accept if these information bits are to be recovered at the receiver side with low probability of error. This measure is commonly referred as the capacity, $\mathcal{C}$, of the channel, and is given by the maximum of the mutual information between the input and output of the channel, where the maximization is over all input distribution. The algorithms to achieve near the capacity transmission rates are

called channel codes. As the Separation Theorem states, reliable communication of the information generated by a source is possible if and only if $H < \mathcal{C}$ and it is optimal to solve each of the sub-problems individually. (Un)Fortunately for the research community, Shannon established the previous limits, but did not provide practical ways to construct such algorithms. The research community has been tirelessly looking for capacity achieving source/channel codes ever since.

While the separation principles simplifies the original problem by dividing it into two independent and more manageable sub-problems, it does not take into consideration practical issues like complexity and delay. It is unpractical to require all applications to compress their data before transmission, and the physical layer could lack of sufficient statistical information to apply conventional compression techniques. Therefore, Joint Source-Channel Codes (JSCC) are becoming more and more popular lately. These algorithms treat the source compression and channel coding in a joint fashion, and consequently, tend to perform better when complexity is an issue. Moreover, there exist multi-user communication systems for which the separation principle does not hold, and JSCC usually exhibits an improved performance. For example, the transmission of correlated information sources over a multiple access channel falls into this category if the energy constraint is defined at the transmitter. Interestingly, the theoretical limit in that case is still not known.

1.1 MOTIVATION AND OBJECTIVES

The wireless channel is time-varying and subject to disturbances such that fading, additive noise and interference. The physical layer has to adapt the transmission rate to these varying channel conditions. Conventional Adaptive Modulation and Coding (AMC) systems achieve this by selecting the best channel coding and modulation combinations for the estimated channel conditions at the sender. These combinations are selected from a group of coding schemes that have been previously designed for each channel condition. The set of rate/modulation combinations available is limited, which yields to stair-shape rate curves and often results in inefficient use of the spectrum. Moreover, the success of AMC depends on the instant and accurate channel estimation on the sender, which can not be simultaneously obtained.

To solve these two disadvantages, authors in [2, 3, 4] published research works on smooth and blind rate adaptation schemes. The rate adaptation is smooth in the sense that it provides fine-grained rate adjustment and blind since the channel information is not required at the sender. The works in [2, 3] achieve this by proposing the Strider system and the Spinal codes, respectively, whereas authors in [4] presented a novel analog coding scheme, named Rate Compatible Modulation (RCM). The work in [4] was extended in [5] for JSCC applied to the transmission of non-uniform sources. As illustrated in [5], RCM outperforms the throughput obtained by conventional ACM under time-varying wireless channels, overcoming the stair-shaped rate curves, while boosting the communication efficiency by intelligently utilizing the data redundancy.

One major disadvantage of the analog RCM codes is that they present error floors at high signal to noise ratios. In order to solve this drawback, the works in [6, 7] suggested the use of digital Low-Density Generator Matrix (LDGM) codes in parallel with the RCMs, aiming at reducing this error floor. On one hand, the RCM symbols represent the majority of the coded sequence and are able to decode most of the source block correctly. On the other hand, the LDGM coded bit, which make up a small fraction of the total coded sequence, are able to correct those source bits that the RCM is unable to decode correctly. By properly designing these analog and digital hybrid schemes, it is possible to lower the error floor while maintaining the smooth rate adaptation and joint source-channel coding capabilities of the RCM codes. The previously designed RCM-LDGM codes found in the literature [6, 7] were found by a trial-error procedure, something that requires a large amount of simulation time. Their robustness against channel and source variations, together with the fact that smooth rate adaption is possible, makes RCM-LDGM codes excellent candidates in applications where channel and source variations are encountered.

This Thesis aims to develop new RCM-LDGM design tools and propose their use in a variety of novel communication scenarios that include point-to-point and multi-user communications.

1.2 OUTLINE AND CONTRIBUTIONS OF THE THESIS

This Thesis begins by detailing in Chapter 2 the required fundamental concepts of communication theory and factor graphs, as well as providing

an overview of RCM, LDGM and parallel RCM-LDGM codes. The rest of the dissertation is divided in two different parts: Part I (Chapters 3, 4 and 5) covers the transmission of a single source through point-to-point communication channels, whereas Part II (Chapters 6 and 7) considers multi-user communications.

1.2.1 CHAPTER 2

This chapter summarizes the required background to understand the presented technical work. It includes some basic communication and graph theory utilized in iterative decoders, and also introduces the reader to RCM and parallel RCM-LDGM codes. This chapter aims to make the dissertation self-contained. However, for the sake of brevity and because the examples considered here are generally well-known, we do not include the Information Theory required to establish the fundamental limits of the considered communication problems. If needed, we refer the reader to [8] for further detail in this regard.

1.2.2 CHAPTER 3 (PART I)

Chapter 3 looks at the problem of designing high rate parallel RCM-LDGM codes over AWGN and fast fading Rayleigh channels for the transmission of both uniform and non-uniform sources. The previous designs of these codes are based on brute force approach, which requires time-consuming Monte Carlo simulations. In this chapter we propose an EXIT chart analysis and an asymptotic bit error rate prediction method capable of speeding up the design of this hybrid scheme.

The research developed in this chapter has been published in the following journal paper:

Granada, I., Crespo, P.M. & Garcia-Frias, J. "Asymptotic BER EXIT chart analysis for high rate codes based on the parallel concatenation of analog RCM and digital LDGM codes". *J Wireless Com Network 2019, 11 (2019).*

1.2.3 CHAPTER 4 (PART I)

This chapter considers the problem of implementing high-throughput (larger than 2 bits per complex channel symbol) JSCC schemes for the transmission of binary sources with memory over AWGN channels. The sources are modeled either by a Markov chain or a hidden Markov model. We explore the combined use of the Burrows-Wheeler Transform (BWT), a block-sorting algorithm frequently used in source compression, and the rate-compatible RCM-LDGM codes. The BWT is first used to convert the original source with memory into a set of independent non-uniform discrete memoryless binary sources, which are then separately RCM-LDGM encoded at optimal coding rates.

The scheme proposed in this chapter has been published in the following journal paper:

Granada, I.; Crespo, P.M.; Garcia-Frias, J. "Combining the Burrows-Wheeler Transform and RCM-LDGM Codes for the Transmission of Sources with Memory at High Spectral Efficiencies". *Entropy 2019, 21, 378.*

1.2.4 CHAPTER 5 (PART I)

Chapter 5 studies power-line communications, which are considered harsh environment for data transmission due to the appearance of noise bursts. These channels are commonly known as additive impulsive noise. We propose the use of RCM-LDGM codes for impulsive noise channels for the first time in the literature. In order to asses their performance, we also propose modified versions of Middleton class A noise and Markov-Middleton impulsive noise for which the Shannon limit is given by the AWGN capacity.

The work developed in this chapter has been published in the following conference paper:

I. Granada and X. Insausti, "On the Suitability of RCM-LDGM Codes for Sending High Data Rates over Impulsive Noise Channels," *GLOBECOM 2020 - 2020 IEEE Global Communications Conference, Taipei, Taiwan, 2020, pp. 1-6.*

1.2.5 CHAPTER 6 (PART II)

The literature lacks of coding schemes suitable for non-orthogonal multiple access operating at high spectral efficiency and capable of accommodating a large number of users, properties highly demanded by modern wireless networks. In this chapter we propose a novel non-orthogonal multiple access coding scheme that is based on RCM codes, one constructed for each user. The system is designed to take advantage of the linear nature of theses codes and the multiple access channel to construct a MAC coding that allows the simultaneous transmission of a large number users at high sum-rates, while maintaining low decoding complexity.

The scheme proposed in this chapter has been published in the following journal paper:

I. Granada, P. M. Crespo and J. Garcia-Frias, "Rate Compatible Modulation for Non-Orthogonal Multiple Access," in *IEEE Access, vol. 8, pp. 224246-224259, 2020.*

1.2.6 CHAPTER 7 (PART II)

Chapter 7 considers the non-orthogonal multiple access scenario of Chapter 6 but whose users transmit spatially Correlated Information Sources (CIS). In this case, as well known, the system performance is optimized if the codewords are designed to take advantage of the correlation among sources in the multiple access channel. The sparse and linear nature of the RCM codes makes them advantageous to design such codewords. Therefore, in this chapter we propose a MAC coding system that is based on RCM schemes.

Moreover, in order to lower the error floor of the MAC coding schemes presented in Chapters 6 and 7, we also propose a parallel LDGM concatenation suitable for the MAC, similar to what is done in point-to-point links.

The research developed in this chapter has been published in the following journal paper:

I. Granada, P. M. Crespo, M. E. Burich and J. Garcia-Frias, "Rate Compatible Modulation for Correlated Information Sources," in *IEEE Access, vol. 9, pp. 65449-65465, 2021.*

1.2.7 CHAPTER 8

This last chapter concludes this dissertation summarizing the main conclusions and providing guidelines of future research lines of interest.

1.3 READING THIS THESIS

The reading of this dissertation does not need to be sequential, and each of the parts (or even chapters, specially for Part I) can be read independently. However, we highly recommend starting with Chapter 2.

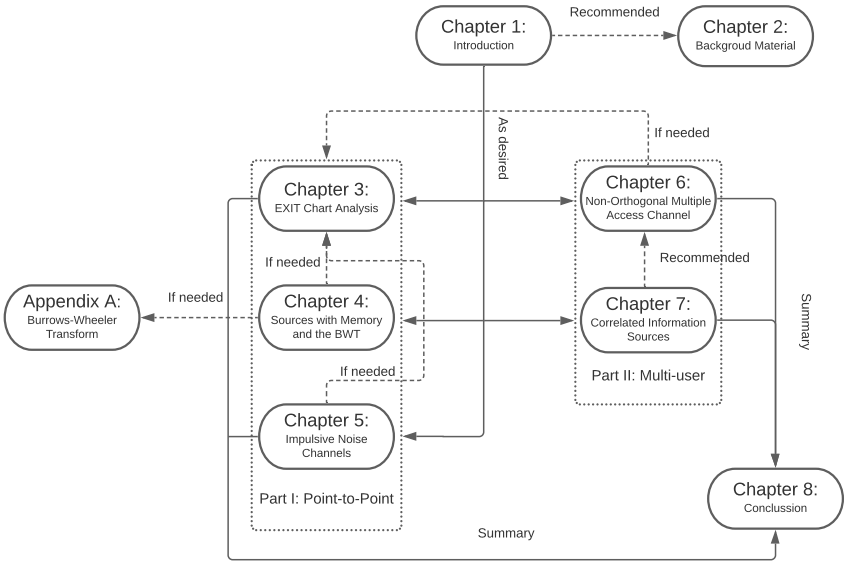


Figure 1.1: *Block diagram detailing the dependencies between chapters.*

We provide in Fig. 1.1 a block diagram with the logical dependencies among the chapters in case the reader decides not to read the dissertation sequentially. A general relationship between chapters is established. Chapters 3, 4 and 5 cover point-to-point communication systems (Part I), whereas chapters 6 and 7 (Part II) consider multi-user communication sys-

tems. Each connecting line is labeled with keywords (Recommended, As desired, If needed, and Summary) indicating the reader how to connect chapters. Each part contains a short common introduction aiming to provide additional context.

CHAPTER 2

Background Material

This chapter provides the reader background material needed for the following chapters. To do so, we review some basic concepts of Communications Theory from a Factor Graph's perspective and present the joint source-channel codes used throughout the Thesis. We begin considering a point-to-point communication system.

2.1 JOINT SOURCE-CHANNEL CODING AND DECODING OF POINT-TO-POINT COMMUNICATIONS

Let $(\Omega, \beta, \mathcal{P})$ be the underlying probability space where all the random variables (r.v.) are defined. We use uppercase when referring to r.v. and lowercase when referring to realizations of r.v.. In addition, we use boldface when referring to vectors. A r.v. X is *discrete* if it takes values on a finite set $\mathcal{X}$ (with $|\mathcal{X}| < \infty$), that we denote as the alphabet, and is *continuous* if it takes value in infinitely many values. For discrete r.v's we define the Probability Mass Function (PMF) of X as $P_X(x) \triangleq \mathcal{P}\{\omega \in \Omega : X(\omega) = x\}$. For continuous r.v's we define the Probability Density Function (PDF) of X as $p_X(x) = \frac{d}{dx} F_X(x)$, where $F_X(x) \triangleq \mathcal{P}\{\omega \in \Omega : X(\omega) \leq x\}$. Whenever the context allows the simplification, we use $P(x)$ and $p(x)$ for PMF and PDF, respectively.

We consider the communications scenario depicted in Figure 2.1. In this work, it is assumed that the source is modeled as a discrete binary K-dimensional random vector. A message $\mathbf{u} \in \{0,1\}^K$ is formed by a block of K source symbols.

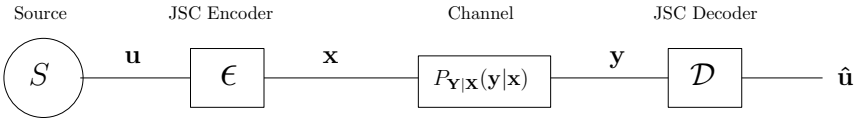


Figure 2.1: *Communications scenario with joint source-channel encoding and decoding functions. Note that modulation is embedded into the encoding.*

The encoder ϵ is assumed to be an injective mapping that assigns each message $\mathbf{u} \in \{0,1\}^K$ to a codeword $\mathbf{x} \in \mathcal{X}^N$, i.e.,

$$\epsilon : \{0,1\}^K \to \mathcal{X}^N, \tag{2.1}$$

where $\mathcal{X} \subset \mathbb{R}$ is the coded symbol's alphabet. We consider the discrete memoryless channel

$$\{\mathcal{X}, P_{Y|X}(y|x), \mathcal{Y}\}$$

modeled by a discrete conditional probability distribution with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y} \subset \mathbb{R}$. The channel is stationary and memoryless in the sense that after it is used t times with inputs $\{X_1, \ldots, X_t\}$, the output at time t ($Y_t \in \mathcal{Y}$) given $\{X_1, \ldots, X_t, Y_1, \ldots, Y_{t-1}\}$ is distributed according to

$$P(y_t|x_1, \ldots, x_t, y_1, \ldots, y_{t-1}) = P(y_t|x_t). \tag{2.2}$$

It can be shown that when this channel is used N times without feedback, the memoryless property implies

$$P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) \triangleq P_{Y_1 \ldots Y_N|X_1 \ldots X_N}(y_1 \ldots y_N|x_1 \ldots x_N) = \prod_{t=1}^{N} P_{Y|X}(y_t|x_t). \tag{2.3}$$

The joint probability of $\mathbf{U}$, $\mathbf{X}$ and $\mathbf{Y}$ can be factorized as

$$P(\mathbf{u}, \mathbf{x}, \mathbf{y}) = P(\mathbf{u}) \cdot P(\mathbf{x}|\mathbf{u}) \cdot P(\mathbf{y}|\mathbf{u}, \mathbf{x}), \tag{2.4}$$

and since by construction $\mathbf{U} \leftrightarrow \mathbf{X} \leftrightarrow \mathbf{Y}$ forms a Markov chain ($P(\mathbf{y}|\mathbf{u}, \mathbf{x}) = P(\mathbf{y}|\mathbf{x})$), the joint probability $P(\mathbf{u}, \mathbf{x}, \mathbf{y})$ is simplified to the factorized chain

given by

$$P(\mathbf{u},\mathbf{x},\mathbf{y}) = \underbrace{P(\mathbf{u})}_{\text{Source}} \cdot \underbrace{P(\mathbf{x}|\mathbf{u})}_{\text{JSCC}} \cdot \underbrace{P(\mathbf{y}|\mathbf{x})}_{\text{Channel}}, \tag{2.5}$$

where the first factor depends directly on the source, the second on the joint source-channel code, and the third on the channel. Finally, the decoder block $\mathcal{D}$ maps the channel output $\mathbf{y}$ to a message sequence $\hat{\mathbf{u}}(\mathbf{y}) = \mathcal{D}(\mathbf{y})$, that is,

$$\mathcal{D} : \mathcal{Y}^N \to \underbrace{\{0,1\}^K}_{\hat{\mathbf{u}}(\mathbf{y})}. \tag{2.6}$$

Given the received sequence $\mathbf{y}$, the decoding function that minimizes the block error probability $\mathcal{P}\{\hat{\mathbf{u}} \neq \mathbf{u}\}$ is the block-wise Maximum a Posteriori (MAP) rule and is given by

$$\hat{\mathbf{u}}(\mathbf{y}) = \underset{\mathbf{u}\in\{0,1\}^K}{\arg\max}\, P(\mathbf{u}|\mathbf{y}) = \underset{\mathbf{u}\in\{0,1\}^K}{\arg\max}\, \frac{P(\mathbf{u},\mathbf{y})}{P(\mathbf{y})}. \tag{2.7}$$

Notice that the denominator $P(\mathbf{y})$ is independent of $\mathbf{u}$, and the numerator $P(\mathbf{u},\mathbf{y})$ is the marginal of the sought joint probability distribution of all variables[1] involved in the communication system. Thus, the previous expression can be rewritten as

$$\hat{\mathbf{u}}(\mathbf{y}) = \underset{\mathbf{u}\in\{0,1\}^K}{\arg\max} \sum_{\mathbf{x}\in\mathcal{X}^N} P(\mathbf{u},\mathbf{x},\mathbf{y}). \tag{2.8}$$

The solution of this problem is often computationally too expensive and it is often relaxed. In those cases, the objective becomes to minimize the symbol-wise error probability, i.e. $\mathcal{P}\{\hat{u}_k \neq u_k\}$ for each $k \in \{1,\ldots,K\}$, given by the symbol-wise MAP decoding rule

$$\begin{aligned}
\hat{u}_i(\mathbf{y}) &= \underset{u_i\in\{0,1\}}{\arg\max}\, P(u_i|\mathbf{y}) \\
&= \underset{u_i\in\{0,1\}}{\arg\max}\, \frac{P(u_i,\mathbf{y})}{P(\mathbf{y})} = \underset{u_i\in\{0,1\}}{\arg\max}\, P(u_i,\mathbf{y}) \\
&= \underset{u_i\in\{0,1\}}{\arg\max} \sum_{u_1}\cdots\sum_{u_{i-1}}\sum_{u_{i+1}}\cdots\sum_{u_K}\sum_{\mathbf{x}\in\mathcal{X}^N} P(\mathbf{u},\mathbf{x},\mathbf{y}) \\
&= \underset{u_i\in\{0,1\}}{\arg\max} \sum_{\sim u_i} P(\mathbf{u},\mathbf{x},\mathbf{y}) \qquad \text{for } i = 1,\ldots,K,
\end{aligned} \tag{2.9}$$

[1]The channel outputs $\{Y_t\}_{t=1}^N$ are actually deterministic values.

where $\sum_{\sim u_i}$ denotes the sum over all variables except u_i. In this work we will exclusively consider the symbol-wise decoding. In order to compute it efficiently, we need to represent the joint probability $P(\mathbf{u}, \mathbf{x}, \mathbf{y})$ as a factor graph and run the Sum-Product Algorithm. The next section presents factor graphs and the Sum-Product Algorithm in detail.

2.2 FACTOR GRAPH AND THE SUM PRODUCT ALGORITHM

2.2.1 FACTOR GRAPHS

Algorithms that must deal with complicated multi-variable global functions often exploit the manner in which the given functions factor as a product of local functions, each of which depends on a subset of the variables. Such a factorization can be visualized with a bipartite graph that is generally denoted as a factor graph [9]. As an example, let $g(x_1, x_2, x_3, x_4, x_5)$ be a multi-variable global function of a set of variables $\{X_1, X_2, X_3, X_4, X_5\}$. This function is represented in Fig. 2.2, where the *function* and *variable nodes* are represented in squares and circles, respectively. Function nodes and variable nodes are connected with an edge if and only if the corresponding variable is an argument of the function. Note that a factor graph is always bipartite, i.e., edges are only allowed between vertices of different types.

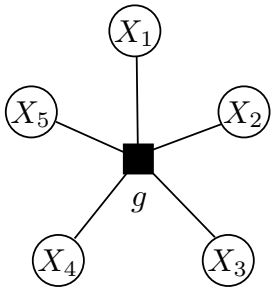


Figure 2.2: *Factor graph representing the global multi-variable function* $g(x_1, x_2, x_3, x_4, x_5)$.

A Factor graph becomes interesting if the represented global function can be factorized as a product of several local functions. Let us assume

that the function $g(x_1, x_2, x_3, x_4, x_5)$ factors as

$$g(x_1, x_2, x_3, x_4, x_5) = f_A(x_1) \cdot f_B(x_2) \cdot f_C(x_1, x_2, x_3) \cdot f_D(x_3, x_4) \cdot f_E(x_3, x_5).$$

In this case, the factor graph can be expanded to represent the local

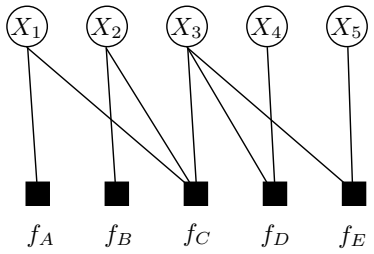


Figure 2.3: *Factorized factor graph of* $g(\cdot)$

functions f_A, f_B, f_C, f_D, f_E. That expansion is depicted in Fig. 2.3 with all variable nodes in the top layer and the function nodes in the bottom layer. Note that this bipartite graph is unique but can be redrawn as desired. For convenience, let us rearrange it as in Fig. 2.4.

Factor graphs are so often used in communication systems due to the existence of the Sum-Product Algorithm (SPA). It consist of a set of rules and procedures for message passing over factor graphs that allows decoding by efficiently calculating the marginals of the joint probability distribution that characterizes the whole communication system.

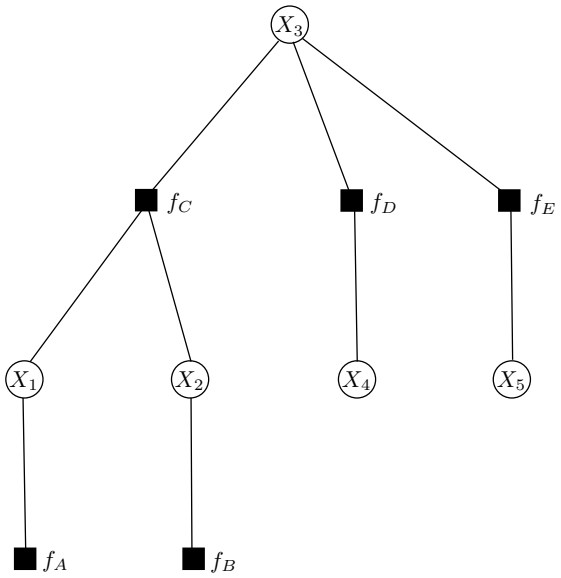


Figure 2.4: *Rearrangement of Fig. 2.3.*

2.2.2 SUM-PRODUCT ALGORITHM

The SPA is essentially a set of rules for transferring messages through the edges of a factor graph with the goal of computing the marginals of the global function described in it in a simply manner. Let us begin by describing the application of the SPA over a cycle-free graph, commonly denoted by factor tree.

The Sum-Product Algorithm over Non-cyclic Factor Graphs

Let us revisit the previous global function $g(\cdot)$ depicted in Fig. 2.4 factorized as

$$g(x_1, x_2, x_3, x_4, x_5) = f_A(x_1) \cdot f_B(x_2) \cdot f_C(x_1, x_2, x_3) \cdot f_D(x_3, x_4) \cdot f_E(x_3, x_5),$$

and assume that we are interested in calculating the marginal function $\hat{g}_1(x_1)$, given by

$$\hat{g}_1(x_1) = \sum_{x_2, x_3, x_4, x_5} g(x_1, x_2, x_3, x_4, x_5) = \sum_{\sim x_1} g(x_1, x_2, x_3, x_4, x_5). \quad (2.10)$$

The SPA exploits the factorized structure of $g(\cdot)$ in such a way that the computation of the marginals is efficiently achieved by exchanging messages among the nodes of its factor graph under the fulfillment of certain rules. Let us begin by defining the following terms:

- We denote by $\psi(X_i)$ to the set of function nodes f_j connected to variable node X_i. Similarly, $\psi(X_i) \setminus f_k$ denotes the set of function nodes f_j connected to variable node X_i except f_k. For example, $\psi(X_1) = \{f_A, f_C\}$ and $\psi(X_1) \setminus f_A = \{f_C\}$ in Fig. 2.4.
- We denote by $\mu_{X_i \to f_j}(x_i)$ and $\mu_{f_j \to X_i}(x_i)$ to the messages transmitted from the variable node X_i to function node f_j and from f_j to X_i, respectively. Note that these messages are only transmitted if $f_j \in \psi(X_i)$ and $X_i \in \psi(f_j)$.

These messages are propagated through all edges of the graph obeying the rules imposed by the algorithm and can be reused to compute different marginal functions. In the special case in which the factor is cycle-free, as the considered case, each message has to be computed only once. We next detail the rules for the computation of variable and function nodes.

Messages transmitted from variable to function nodes: The message from variable node X_i to function node f_j is computed as

$$\mu_{X_i \to f_j}(x_i) = \prod_{f_k \in \psi(X_i) \setminus f_j} \mu_{f_k \to X_i}(x_i). \tag{2.11}$$

Messages transmitted from function to variable nodes: The message from function node f_j to variable node X_i is computed as

$$\mu_{f_j \to X_i}(x_i) = \sum_{X_k \in \psi(f_j) \setminus X_i} f_j(x_k) \cdot \mu_{X_k \to f_j}(x_k). \tag{2.12}$$

Computation of the marginal functions: The desired marginal function $\hat{f}_i(x_i)$ is computed as

$$\hat{f}_i(x_i) = \prod_{f_j \in \psi(X_i)} \mu_{f_j \to X_i}(x_i). \tag{2.13}$$

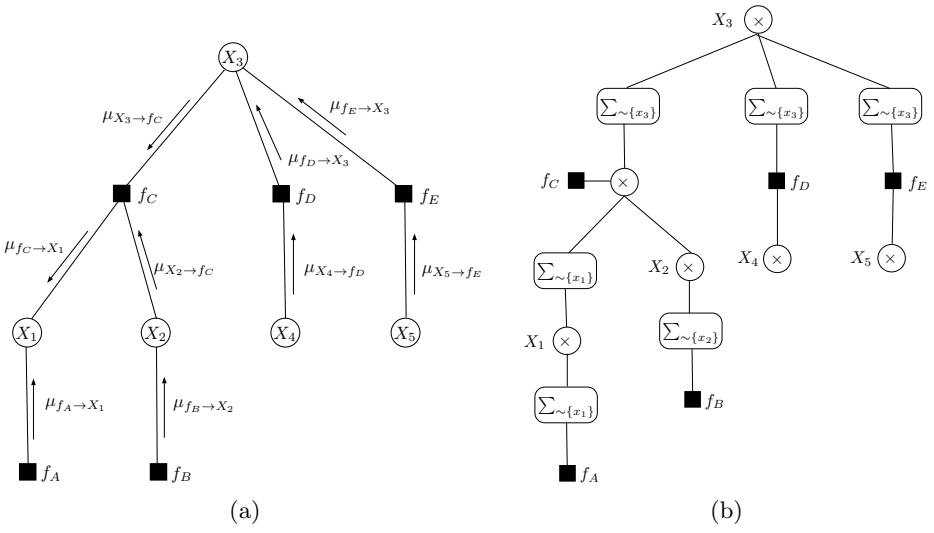


Figure 2.5: ***(a)*** *Messages used for the marginalization of* $\hat{g}_1(x_1)$*.* ***(b)*** *Update rules of the marginalization of* $\hat{g}_1(x_1)$*.*

Figure 2.5 exemplifies the application of the set of rules necessary for calculating the marginal function $\hat{g}_1(x_1)$ by breaking the expression for $\hat{g}_1(x_1)$ down into its compounding factors. Let us observe that

$$\begin{aligned}\hat{g}_1(x_1) &= \sum_{x_2}\sum_{x_3}\sum_{x_4}\sum_{x_5} f_A(x_1)\cdot f_B(x_2)\cdot f_C(x_1,x_2,x_3)\cdot f_D(x_3,x_4)\cdot f_E(x_3,x_5)\\ &= \underbrace{\underbrace{\underbrace{f_A(x_1)}_{\mu_{f_A\to X_1}(x_1)}\cdot\sum_{x_2}\sum_{x_3} f_C(x_1,x_2,x_3)\cdot \underbrace{\underbrace{f_B(x_2)}_{\mu_{f_B\to X_2}(x_2)}}_{\mu_{X_2\to f_C}(x_2)}\cdot \underbrace{\sum_{x_4} f_D(x_3,x_4)\cdot \underbrace{1}_{\mu_{X_4\to f_D}(x_4)}}_{\mu_{f_D\to X_3}(x_3)}\cdot \underbrace{\sum_{x_5} f_E(x_3,x_5)\cdot \underbrace{1}_{\mu_{X_5\to f_E}(x_5)}}_{\mu_{f_E\to X_3}(x_3)}}_{\mu_{X_3\to f_C}(x_3)}}_{\mu_{f_C\to X_1}(x_1)},\end{aligned} \tag{2.14}$$

and the marginalization of X_1 is given by,

$$\hat{g}_1(x_1) = \mu_{f_A\to X_1}(x_1)\cdot\mu_{f_C\to X_1}(x_1). \tag{2.15}$$

The rest of the marginals $\hat{g}_2(x_2)$, $\hat{g}_3(x_3)$, $\hat{g}_4(x_4)$ and $\hat{g}_5(x_5)$ are calculated in a similar fashion, with all needed messages represented in Fig. 2.6. Notice that the execution of the algorithm comprises a finite number of

steps due to the lack of cycles in the underlying tree factor graph. However, many coding factor graphs present cycles that this algorithm does not take into account. In those cases, the algorithm has no natural termination and consequently, the marginalizations are not exact.

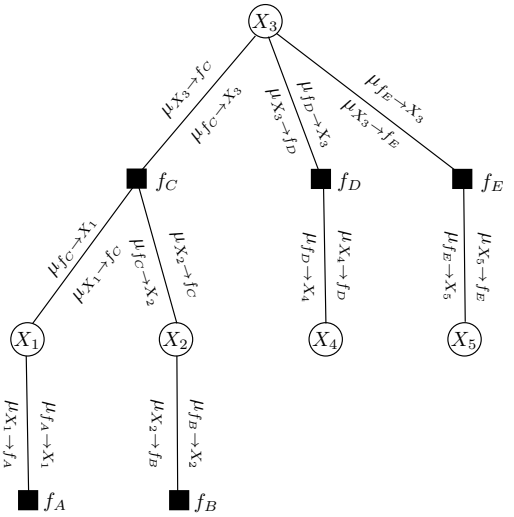


Figure 2.6: *All necessary messages needed for calculating the marginal functions $\hat{g}_1(x_1), \hat{g}_2(x_2)$, $\hat{g}_3(x_3)$, $\hat{g}_4(x_4)$ and $\hat{g}_5(x_5)$. Messages that propagate upwards/downwards are placed on the right/left side of the edges.*

The Sum-Product algorithm over Cyclic Factor Graphs

The majority of the factor graphs used in communication systems have cycles, and therefore, an exact marginalization of the global function is not possible with this algorithm, since there are no finite number of steps in the algorithm [10].

However, the SPA can still be executed on these factor graphs by imposing a termination to the iterative algorithm. This results in a inexact calculation of the marginal functions, yielding to a sub-optimal marginalization algorithm. However, it still represents a good heuristic method to implement iterative decoders of communication systems. This assumption has been shown to hold in many instances, yielding good results in prac-

tice and becoming widely extended in modern coding applications such as Turbo and LDPC coding [11].

All things considered, the decoding of a received noisy sequence proceeds, in general, according to the following scheme:

- All messages are initialized.
- Messages are updated according to an update schedule: (2.11) and (2.12). This schedule may vary from step to step.
- After each step (or a predefined number of steps) the marginal functions of the symbols to be decoded are computed: (2.13).
- A sequence is obtained by taking decisions based on the current marginal function.
- After obtaining the sequence, a decision should be made: continue with the decoding process until a specific number of iterations is achieved or stop if the sequence fulfills some conditions.

2.2.3 FACTOR GRAPHS IN JSC DECODING

This Section presents the applicability of factor graphs and the SPA to the communication problem of Section 2.1. The criteria used for decoding iterative algorithms is the symbol-wise MAP rule, which maximizes the probability $p(\hat{u}_i|\mathbf{y})$ for each symbol $\hat{u}_i$. This is the marginal of the joint probability distribution of the global function (2.9) representing the communication scenario, and is given by

$$\hat{u}_i(\mathbf{y}) = \underbrace{\underbrace{\arg\max_{u_i \in \{0,1\}}}_{\text{Decision}} \underbrace{\sum_{\sim u_i} \underbrace{P(\mathbf{u},\mathbf{x},\mathbf{y})}_{\text{FG (comm. scenario)}}}_{\text{Sum-Product Algorithm}}}_{\text{Symbol-wise decoding}}. \tag{2.16}$$

For large block lengths this sum is computationally infeasible. However, it can be approximated by applying the SPA algorithm to the factor graph representing the global function $P(\mathbf{u},\mathbf{x},\mathbf{y})$ and consequently, by computing efficiently the required marginals for each variable node U_i. Taking into account that the global function $P(\mathbf{u},\mathbf{x},\mathbf{y})$ can be rewritten as in (2.5), we

derive that the factor graph of the communication systems depends directly on the (i) local functions representing the source, the JSCC technique and the characteristics of the channel, and (ii) the variable nodes representing the source symbols $\{U_i\}$, the coded symbols $\{X_i\}$ and the channel observations $\{Y_i\}$. Therefore, expression (2.16) reduces to,

$$\hat{u}_i(\mathbf{y}) = \underset{u_i \in \{0,1\}}{\arg\max} \underbrace{\sum_{\sim u_i} P(\mathbf{u}, \mathbf{x}, \mathbf{y})}_{\text{SPA}} = \underset{u_i \in \{0,1\}}{\arg\max} \underbrace{\sum_{\sim u_i} \underbrace{P(\mathbf{u})}_{\text{Source}} \cdot \underbrace{P(\mathbf{x}|\mathbf{u})}_{\text{JSCC}} \cdot \underbrace{P(\mathbf{y}|\mathbf{x})}_{\text{Channel}}}_{\text{SPA}} . \tag{2.17}$$

Summarizing, the process of performing a symbol-wise MAP decision in the proposed communication scenario follows a three-stage approach:

1. The joint probability distribution $P(\mathbf{u}, \mathbf{x}, \mathbf{y})$ is represented by a factor graph composed of three sub-graphs representing the source, the joint source-channel code and channel.
2. The SPA messages are passed through the factor graph.
3. The MAP decision is based on the calculated marginals.

Observe that the marginals are exact for non-cyclic graphs and approximated in case of cyclic graphs. In the latter case, step 2 and 3 can be performed iteratively, yielding successively refined estimations of $\hat{u}_i$.

In the next sections we present examples of source models $P_U(\mathbf{u})$, channel models $P(\mathbf{y}|\mathbf{x})$ and the JSC codes $P(\mathbf{x}|\mathbf{u})$ used throughout this dissertation.

2.3 FACTOR GRAPH OF SOURCES

In this section we present two examples of the source models $P_{\mathbf{U}}(\mathbf{u})$.

2.3.1 FACTOR GRAPH OF MARKOV SOURCES

We consider the discrete stochastic process with memory in which each random variable depends on the preceding and is conditionally independent of every other preceding r.v.. This temporal dependence is expressed in the following definition.

Definition 1. *A discrete stochastic process* $\{S_k\}_{k=1}^{\infty}$ *is a Markov Chain (MC) if*

$$P_{S_k|S_{k-1},\ldots,S_1}(s_k|s_{k-1},\ldots,s_1) = P_{S_k|S_{k-1}}(s_k|s_{k-1}) \tag{2.18}$$

for $k \in \mathbb{N}$. *Moreover, if the the conditional probability does not depend on* k, *the MC is denoted as time invariant.*

Aiming to keep this dissertation as brief as possible, let us define directly the MC generalization in which the system has unobservable *hidden* states, denoted as Hidden Markov Model (HMM). Concretely, we consider a time-invariant HMM where the r.v's take values in the finite alphabet $\mathcal{S} = \{1,\ldots,M_\lambda\}$ and denote the states of the source at time k as s_k. Complete specification of a binary HMM requires of three probability measures, A, B, and π, defined as:

- $A = [a_{ij}]$ is the state transition probability matrix of dimension $M_\lambda \times M_\lambda$, with a_{ij} the probability of transition from state i to state j, i.e., $a_{i,j} = P(s_{k+1} = j|s_k = i)$ for all k.
- $B = [b_j(u)]$ is the observation symbol probability matrix, with $b_j(u)$ the probability of getting in the binary symbol u in state j, i.e., $b_j(u) = P(u|j)$, $1 \leq j \leq M_\lambda$, $u \in \{0,1\}$.
- π is the initial state distribution vector, with π_j the probability for the initial state to be j, i.e., $\pi_j = P(s_1 = j)$, $1 \leq j \leq M_\lambda$.

For stationary sources, π should be taken as the stationary distribution of the chain, i.e., $\pi = A\pi$. Note that a MC can be defined as a particular example of HMM in which the matrix B has entires 0 and 1, i.e., the observation symbol matrix is a deterministic function.

Let us assume a discrete source with memory modeled by a HMM with parameters A, B, and π, and whose output $\{U_k\}_{k=1}^{K}$ is mapped from a sequence of states $\{S_k\}_{k=1}^{K}$ by means of the mapping $T(\cdot)$ defined as

$$T_k\left(s_{k-1}, s_k, u_k\right) = T(S_{k-1} = s, S_k = s', U_k = u), \tag{2.19}$$

with $s, s' \in \{1,\ldots,M_\lambda\}$ and $u \in \{0,1\}$. With these function and variable nodes, we can plot the source factor graph as in Fig. 2.7

Observe that the considered Markovian model yields a factor graph with no cycles. Thus, the exact factorization of $P_{\mathbf{U}}(\mathbf{u})$ is obtained in a finite number of steps by applying the SPA.

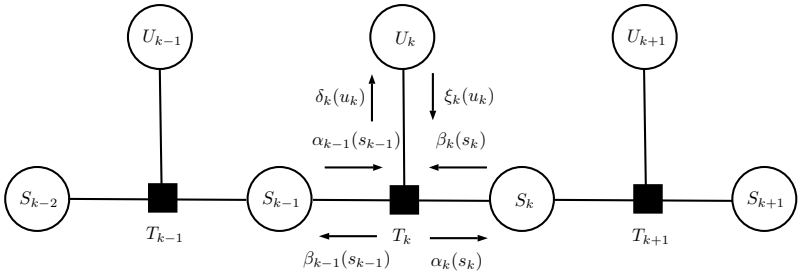


Figure 2.7: *Factor graph modeling a Markov Chain.*

Forward and Backward recursions

The application of the SPA to the factor graph in Fig. 2.7 yields two natural recursions that are often called *forward* and *backward*. The first one $\alpha_k(\cdot)$ as a function of $\alpha_{k-1}(\cdot)$, and the second $\beta_{k-1}(\cdot)$ as a function of $\beta_k(\cdot)$. Before proceeding with the details, let us define the following notation[2]:

- $\xi_k(u_k)$: extrinsic soft information of the variable node U_k.
- $\alpha_k(s_k), \beta_{k-1}(s_{k-1})$: forward and backward variables, that are defined as message from T_k to S_k and from T_k to S_{k-1}, respectively.
- $\delta_k(u_k)$: a posteriori information of the variable node U_k.

And the update rules for the recursions are given by

$$\alpha_k(s_k) = \sum_{\sim s_k} \alpha_{k-1}(s_{k-1}) \cdot T_k(s_{k-1}, s_k, u_k) \cdot \xi_k(u_k) \tag{2.20}$$

$$\beta_{k-1}(s_{k-1}) = \sum_{\sim s_{k-1}} \beta_k(s_k) \cdot T_k(s_{k-1}, s_k, u_k) \cdot \xi_k(u_k), \tag{2.21}$$

where $T_k\left(s_{k-1}, s_k, u_k\right)$ is the k^{th} function node of the HMM. Note that the values $\alpha_0(\cdot)$ and $\beta_K(\cdot)$ are required, and are obtained from the initial

[2]Note that they correspond to the SPA messages $\mu_{U_k \to T_k}(u_k)$, $\mu_{T_k \to S_k}(s_k)$, $\mu_{T_k \to S_{k-1}}(s_{k-1})$ and $\mu_{T_k \to U_k}(u_k)$.

and final conditions of the described MC source. For the HMM sources considered here, $T_k\left(s_{k-1}, s_k, u_k\right)$ is given by

$$T_k\left(s_{k-1}, s_k, u_k\right) = \begin{cases} P(s_k), & \text{for } k = 0 \\ P(u_k|s_{k-1}), & \text{for } k = K \\ P(s_k|s_{k-1})P(u_k|s_{k-1}), & \text{otherwise,} \end{cases} \tag{2.22}$$

and the a posteriori information for the input variable node U_k is computed as

$$\delta_k(u_k) = \sum_{\sim u_k} \alpha_{k-1}(s_{k-1}) \cdot T_k(s_{k-1}, s_k, u_k) \cdot \xi_k(u_k) \cdot \beta_k(s_k). \tag{2.23}$$

Finally, the sought marginal is given by the product of all messages arriving at the node representing the variable to be marginalized, i.e.,

$$P_{U_k}(u_k) = \sum_{\sim u_k} P_{\mathbf{U}}(\mathbf{u}) = \delta_k(u_k) \cdot \xi_k(u_k). \tag{2.24}$$

2.3.2 FACTOR GRAPH OF NON-UNIFORM SOURCES

We consider the simple, yet interesting case, where the source is modeled as a binary independent and identically distributed (i.i.d.) discrete time random process, i.e., whose the output $U_k \in \{0, 1\}$ given $(U_1, \ldots, U_{k-1})$ is distributed according to

$$P_{U_k|U_1,\ldots,U_{k-1}}(u_k|u_1, \ldots, u_{k-1}) = P_{U_k}(u) = \begin{cases} p & u = 0 \\ 1-p & u = 1 \end{cases}, \forall k \in \mathbb{N}. \tag{2.25}$$

By non-uniform we mean that the binary alphabet $\{0, 1\}$ has distribution $(1 - p; p)$. Thus, in this simple case, the a posteriori information of the variable node U_k, $\delta_k(u_k)$, is given by

$$\begin{aligned} \delta_k(u_k = 1) &= p \\ \delta_k(u_k = 0) &= 1 - p. \end{aligned} \tag{2.26}$$

Since it is constant, it is often denoted as a priori probability. Finally, the factor graph of this simple source model is depicted in Fig. 2.9.

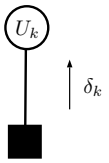


Figure 2.8: *Factor Graph of a non-uniform source.*

2.4 FACTOR GRAPHS OF CHANNELS

In this section we are interested on the conditional probability $P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$.

2.4.1 GAUSSIAN CHANNEL MODEL

Consider the use of a single carrier passband communication system to send i.i.d. real symbols $\{X_t\}_{t\in\mathbb{N}}$ across a white Gaussian passband channel. Assume that the X_t symbols generated by the encoder belong to some one-dimensional real constellation $\mathcal{A}$.

In order to adapt the transmitted symbols to the passband modulator, one needs to consider two consecutive symbols $\overline{X}_t = (X_{2t-1}, X_{2t})$ drawn from the two-dimensional constellation $\mathcal{A}^2 = \mathcal{A} \times \mathcal{A}$. In what follows we use the standard notation

$$\overline{X}_t = (X_t^I, X_t^Q), \tag{2.27}$$

where the upper letters I and Q stand for *in-phase* and *quadrature* components, respectively.

It is well known that the sample vector $\overline{Y}_t \in \mathbb{R}^2$ at output of the passband demodulator, is a sufficient statistic to detect the symbol $\overline{X}_t$ and is given by

$$\overline{Y}_t = \begin{pmatrix} Y_t^I \\ Y_t^Q \end{pmatrix} = \begin{pmatrix} X_t^I \\ X_t^Q \end{pmatrix} + \begin{pmatrix} Z_t^I \\ Z_t^Q \end{pmatrix} = \overline{X}_t + \overline{Z}_t,\ t \in \mathbb{N}, \tag{2.28}$$

where $\{\overline{Z}_t\}_{t=1}^{\infty}$ is an i.i.d. sequence of two-dimensional Gaussian random vectors of zero mean and covariance matrix $N_0 \cdot \mathrm{I}_2$, with I_2 the 2×2 identity

matrix. In here, N_0 denotes the two-sided Power Spectral Density (PSD) level of the additive white Gaussian noise random process.

Therefore, the conditional PDF of the Gaussian vector, $\overline{Y}_t$ given the symbol $\overline{X}_t$ is given by

$$P_{\overline{Y}|\overline{X}}(\overline{y}|\overline{x}) = \frac{1}{N_0\pi}\exp\left(-\frac{\|\overline{y}-\overline{x}\|^2}{N_0}\right). \tag{2.29}$$

Observe that if we use complex notation, expression (2.28) results in what is called the complex AWGN channel

$$\widetilde{Y}_t = \widetilde{X}_t + \widetilde{Z}_t,\ t \in \mathbb{N}, \tag{2.30}$$

where

$$\widetilde{X}_t = X_t^I + i \cdot X_t^Q \tag{2.31}$$

now belongs to the one-dimension complex constellation, $\mathcal{A}_{\mathbb{C}} = \mathcal{A} \times i \cdot \mathcal{A} \subset \mathbb{C}$, and the i.i.d. complex Gaussian random variables $\{\widetilde{Z}_t\}$

$$\widetilde{Z}_t = Z_t^I + i \cdot Z_t^Q \tag{2.32}$$

are circularly symmetric complex random variables.

The energy per two real-dimensions (or one-complex dimension), E_s, is defined as

$$E_s = \mathrm{E}(\|\overline{X}\|^2) = 2 \cdot \mathrm{E}(X^2), \tag{2.33}$$

and the Signal to Noise Ratio is defined as

$$SNR = \frac{E_s}{N_0}, \tag{2.34}$$

that is, the ratio between the energy per two real-dimension 2-D (or one complex-dimension 1-C) and the variance of the noise per 2-D.

Note that due to the independence of the Gaussian noise random components $Z_t^I = Z_{2t-1}$ and $Z_t^Q = Z_{2t}$ at the output of the passband demodulator in (2.28), one can simplify the model (2.28) by considering two consecutive uses of the following one-dimension real AWGN channel

$$Y_t = X_t + Z_t, \tag{2.35}$$

where t takes values in the natural numbers $\mathbb{N}$. In this dissertation, and for the sake of notation, we will use this simplify one-dimensional real AWGN channel, depicted in Fig. 2.9.

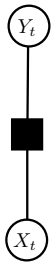


Figure 2.9: *Factor Graph of the considered AWGN channel.*

2.4.2 RAYLEIGH FAST FADING CHANNEL MODEL

In a fast-fading Rayleigh channel the sequence complex symbols $\{\widetilde{X}_t\}$ in the AWGN complex channel (2.30) is pairwise multiplied by an i.i.d. sequence of complex circular symmetric Gaussian random variables $\{\widetilde{H}_t\}$ with variance[3] $E(|\widetilde{H}_t|^2) = 1$, modelling the fading. That is,

$$\widetilde{Y}_t = \widetilde{H}_t \cdot \widetilde{X_t} + \widetilde{Z_t},\ t \in N. \tag{2.36}$$

Therefore, the conditional PDF (in complex notation) given the fading coefficient $\tilde{h}$ and the complex symbol $\tilde{x}$, is given by:

$$P_{\widetilde{Y}|\widetilde{X},\widetilde{H}}(\tilde{y}|\tilde{x},\tilde{h}) = \frac{1}{\pi N_0} \exp\left(-\frac{\left\|\widetilde{y} - \tilde{h}\cdot\widetilde{x}\right\|^2}{N_0}\right). \tag{2.37}$$

Note that when perfect channel state information is known at the receiver (coherent detection), the output of the demodulator can be multiplied by $\widetilde{H}_t^*$, and the above Rayleigh fast fading channel (2.36) becomes a time-varying AWGN complex channel

$$\widetilde{Y}_t = |\widetilde{H}_t|^2 \cdot \widetilde{X_t} + \widetilde{Z_t},\ t \in N, \tag{2.38}$$

where $|\widetilde{H}_t|^2$ is exponential distributed with unit variance, i.e.,

$$p_{|\widetilde{H}_t|^2}(h) = e^{-h} \qquad h > 0, t \in \mathbb{N}. \tag{2.39}$$

[3]The complex circular symmetric Gaussian random variable $\widetilde{H}_t$, $t \in \mathbb{N}$, can be written as $\widetilde{H}_t = \frac{1}{\sqrt{2}} \cdot \left(V_t^I + i \cdot V_t^Q\right)$ where V_t^I and V_t^Q, are independent real Gaussian r.v's of zero mean and variance 1

Note that the average SNR is still the same as before, but the instantaneous SNR at any given time instant is now

$$\mathrm{SNR}_t = |\widetilde{h}_t|^2 \cdot \frac{E_s}{N_0}, \tag{2.40}$$

where $\widetilde{h}_t$ denotes a realization of the random variable $\widetilde{H_t}$.

2.5 EXTENSION TO MULTI-USER CODING AND DECODING

We consider the communications scenario depicted in Figure 2.10, where the information produced by λ discrete binary random vectors is transmitted to the same receiver through a shared channel. The message of user i $\mathbf{u}^i \in \{0,1\}^{K_i}$ is formed by a block of K_i source symbols.

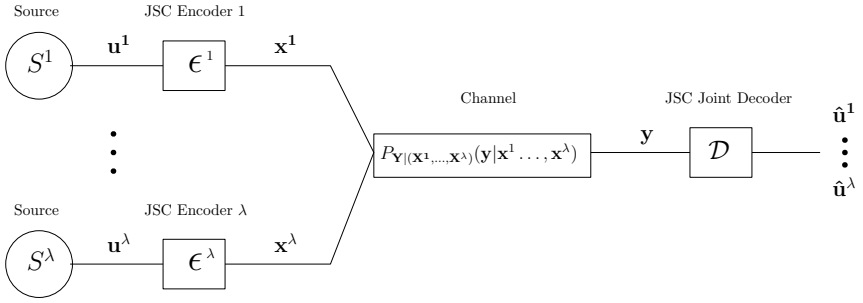


Figure 2.10: *Multi-user communication scenario with independent encoding and joint decoding functions.*

The λ encoders $\{\epsilon^1, \ldots, \epsilon^\lambda\}$ are assumed to be injective mappings that each assigns its message $\mathbf{u}^i \in \{0,1\}^{K_i}$ to a codeword $\mathbf{x}^i \in \mathcal{X}^N$, i.e.,

$$\epsilon^i : \{0,1\}^{K_i} \to \mathcal{X}^N. \tag{2.41}$$

Note that we allow the encoders to encode source blocks of different lengths (and consequently, at different rates). The encoded sequences $\{\mathbf{x}^1, \ldots, \mathbf{x}^\lambda\}$ are then all transmitted through the discrete memoryless channel

$$\{\mathcal{X}^1, \ldots, \mathcal{X}^\lambda, P_{Y|X^1,\ldots,X^\lambda}(y|x^1, \ldots, x^\lambda), \mathcal{Y}\}$$

modeled by a discrete conditional probability distribution with input alphabet $\mathcal{X} \times \ldots \times \mathcal{X} \subset \mathbb{R}^\lambda$ and output alphabet $\mathcal{Y}$. As in Section 2.1, the

channel is stationary, memoryless and used without feedback. Thus, when it is used N times, its distribution is given by

$$P_{\mathbf{Y}|\mathbf{X}^1,\ldots,\mathbf{X}^\lambda}(\mathbf{y}|\mathbf{x}^1,\ldots,\mathbf{x}^\lambda) = \prod_{t=1}^{N} P_{Y_t|X_t^1,\ldots,X_t^\lambda}(y|x_t^1,\ldots,x_t^\lambda). \tag{2.42}$$

The joint probability of $(\mathbf{U}^1,\ldots,\mathbf{U}^\lambda)$, $(\mathbf{X}^1,\ldots,\mathbf{X}^\lambda)$ and $\mathbf{Y}$ can be factorized as

$$\begin{aligned} P(\mathbf{u}^1,\ldots,\mathbf{u}^\lambda,\mathbf{x}^1,\ldots,\mathbf{x}^\lambda,\mathbf{y}) = \\ P(\mathbf{u}^1,\ldots,\mathbf{u}^\lambda) \cdot \prod_{i=1}^{\lambda} P(\mathbf{x}^i|\mathbf{u}^i) \cdot P(\mathbf{y}|\mathbf{u}^1,\ldots,\mathbf{u}^\lambda,\mathbf{x}^1,\ldots,\mathbf{x}^\lambda) = \\ \underbrace{P(\mathbf{u}^1,\ldots,\mathbf{u}^\lambda)}_{\text{Multi-terminal Sources}} \cdot \underbrace{\prod_{i=1}^{\lambda} P(\mathbf{x}^i|\mathbf{u}^i)}_{\text{Ind. Encoders}} \cdot \underbrace{P(\mathbf{y}|\mathbf{x}^1,\ldots,\mathbf{x}^\lambda),}_{\text{Shared channel}} \end{aligned} \tag{2.43}$$

where the last equation holds because $(\mathbf{U}^1,\ldots,\mathbf{U}^\lambda) \leftrightarrow (\mathbf{X}^1,\ldots,\mathbf{X}^\lambda) \leftrightarrow \mathbf{Y}$ forms a Markov Chain, that is, $P(\mathbf{y}|\mathbf{u}^1,\ldots,\mathbf{u}^\lambda,\mathbf{x}^1,\ldots,\mathbf{x}^\lambda) = P(\mathbf{y}|\mathbf{x}^1,\ldots,\mathbf{x}^\lambda)$. The first factor depends on the multi-terminal source, the second on the independent JSC encoders and the last on the multiple access channel.

The joint decoder block $\mathcal{D}$ maps the channel output $\mathbf{y}$ to λ message sequences $\{\hat{\mathbf{u}}_m^1(\mathbf{y}),\ldots\hat{\mathbf{u}}_m^\lambda(\mathbf{y})\}$, that is,

$$\mathcal{D}: \mathcal{Y}^N \rightarrow \underbrace{\{0,1\}^{K_1}}_{\hat{\mathbf{u}^1}(\mathbf{y})} \times \ldots \times \underbrace{\{0,1\}^{K_\lambda}}_{\hat{\mathbf{u}^\lambda}(\mathbf{y})}. \tag{2.44}$$

Given the received sequence $\mathbf{y}$, the decoding objective (the relaxed objective of iterative decoders) is to minimize the symbol-wise error probability, i.e. $\mathcal{P}\{\hat{u}_k^i(\mathbf{y}) \neq u_k^i\}$ for all i and k, given by the symbol-wise MAP decoding rule

$$\begin{aligned} \hat{u}_k^i(\mathbf{y}) &= \arg\max_{u_k^i \in \{0,1\}} P(u_k^i|\mathbf{y}) \\ &= \arg\max_{u_k^i \in \{0,1\}} \frac{P(u_k^i,\mathbf{y})}{P(\mathbf{y})} = \arg\max_{u_k^i \in \{0,1\}} P(u_k^i,\mathbf{y}) \\ &= \arg\max_{u_k^i \in \{0,1\}} \sum_{\sim u_k^i} P(\mathbf{u}^1,\ldots,\mathbf{u}^\lambda,\mathbf{x}^1,\ldots,\mathbf{x}^\lambda,\mathbf{y}). \end{aligned} \tag{2.45}$$

2.5.1 MULTI-TERMINAL SOURCE MODEL

The multi-terminal sources considered in this dissertation are modeled by an stochastic process $\{\mathbf{U_k}\}_{k=1}^{\infty}$ composed by $\lambda-$ dimensional random vectors $\mathbf{U_k} \triangleq (U_k^1, \ldots, U_k^\lambda)$, whose components take values from the alphabet $\{0, 1\}$. Concretely, we focus on the spatial multi-terminal sources that have no memory, and hence, the output vectors have no temporal correlation. Their joint probability distribution is given by

$$P_{\mathbf{U}}\left((u_1^1, \ldots, u_1^\lambda), \ldots, (u_K^1, \ldots, u_K^\lambda)\right) = \prod_{k=1}^{K} P_{\mathbf{U}}(u_k^1, \ldots, u_k^\lambda). \tag{2.46}$$

We will also distinguish two types of multi-terminal sources: the spatially correlated and the uncorrelated ones. A multi-terminal source is said to be uncorrelated (without spatial correlation) if its output components are independent, i.e., their joint probability distribution satisfies

$$P_{\mathbf{U}}\left(u^1, \ldots, u^\lambda\right) \triangleq P_{U^1,\ldots,U^\lambda}\left(u^1, \ldots, u^\lambda\right) = \prod_{i=1}^{\lambda} P_{U^i}\left(u^i\right). \tag{2.47}$$

Conversely, if the above condition does not hold, the multi-terminal source is said to be correlated.

2.5.2 MULTIPLE ACCESS CHANNEL

In this section we are interested on the conditional probability

$$P_{\mathbf{Y}|\mathbf{X}^1,\ldots,\mathbf{X}^\lambda}(\mathbf{y}|\mathbf{x}^1, \ldots, \mathbf{x}^\lambda).$$

2.5.2.1 Gaussian MAC

The Gaussian multiple access channel considered in this dissertation is modeled by the sampled baseband (real) link input-output relation

$$\{X^1, \ldots, X^\lambda\} \to Y \tag{2.48}$$

given by

$$Y_t = \sum_{i=1}^{\lambda} X_t^i + Z_t, \forall t \in \mathbb{N}, \tag{2.49}$$

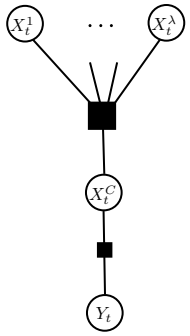


Figure 2.11: *Factor Graph of the multiple access channel at time t.*

where again, the additive Gaussian noise $\{Z_t\}_{t=1}^{\infty}$ are i.i.d Gaussian random variables of zero mean and variance N_0.

For this channel, the conditional PDF is given by

$$P\left(y|x^1,\ldots,x^\lambda\right) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{\left(y-\sum_{i=1}^{\lambda} x^i\right)^2}{N_0}\right). \tag{2.50}$$

For convenience we depict the MAC channel as the Factor Graph in Fig. 2.11, where we define the new variable node $X_t^C = X_t^1 + \ldots + X_t^\lambda$, i.e., the MAC symbol before considering the Gaussian noise.

2.5.2.2 Rayleigh fast fading MAC

We consider the Rayleigh fast fading MAC channel where the complex-domain baseband representation relation between the input and the output

$$\{\widetilde{X^1},\ldots,\widetilde{X^\lambda}\} \to \widetilde{Y} \tag{2.51}$$

is given by

$$\widetilde{Y_t} = \sum_{i=1}^{\lambda} \widetilde{H_t^i} \cdot \widetilde{X_t^i} + \widetilde{Z_t}, \forall t \in \mathbb{N}. \tag{2.52}$$

As for the point-to-point Rayleigh case of Section 2.4.2, the additive Gaussian noise $\{\widetilde{Z_t}\}_{t=1}^{\infty}$ are i.i.d circularly symmetric complex Gaussian

random variables of zero mean and variance N_0. The fading components $\{\widetilde{H_t^i}\}$, for $i = 1, \ldots, \lambda$ and $t \in \mathbb{N}$, are modeled as i.i.d. circularly symmetric complex Gaussian random variables of unit variance (as in Section 2.4.2).

For this channel model, we define the variable node $\widetilde{X_t^C}$ as the complex-domain baseband representation of the MAC channel output symbols before considering the noise, given by

$$\widetilde{X_t^C} = \widetilde{H_t^1} \cdot \widetilde{X_t^1} + \ldots + \widetilde{H_t^\lambda} \cdot \widetilde{X_t^\lambda}.$$

The conditional PDF (in complex-domain representation) is given by

$$P\left(\widetilde{y}|\widetilde{x^1}, \ldots, \widetilde{x^\lambda}, \widetilde{h^1}, \ldots, \widetilde{h^\lambda}\right) = \frac{1}{\pi N_0} \exp\left(- \frac{\left\| \widetilde{y} - \sum_{i=1}^{\lambda} \widetilde{h^i} \cdot \widetilde{x^i} \right\|^2}{N_0} \right). \quad (2.53)$$

Whenever the context allows us to do so, we will drop the *widetilde* in $\widetilde{Y_t}$, $\widetilde{H_t}$ and $\widetilde{Z_t}$ and refer to them by Y_t, H_t and Z_t.

2.6 PARALLEL RCM-LDGM CODES

In this section we present a detailed description of the JSCC $P_{\mathbf{X}|\mathbf{U}}(\mathbf{x}|\mathbf{u})$ that are utilized during the dissertation. For convenience we change the notation from time t to symbol j.

2.6.1 RATE COMPATIBLE MODULATION CODES.

RCM codes were proposed by Cui in [4, 5] to achieve rate adaptation by generating multi-level symbols from weighted linear combination of the input bits as if Fig. 2.12. RCM symbols are used to form the constellation directly, without the need of any labeling, by placing them two-by-two.

In essence, RCM codes embed source compression into modulation through the mentioned generation of multi-level symbols from weighted linear combinations of randomly chosen input bits. This process is realized by using a random mapping matrix with real-valued entries.

The schematic diagram of an RCM is presented in Fig. 2.13 and is represented by a bipartite graph where the edges connecting input bits

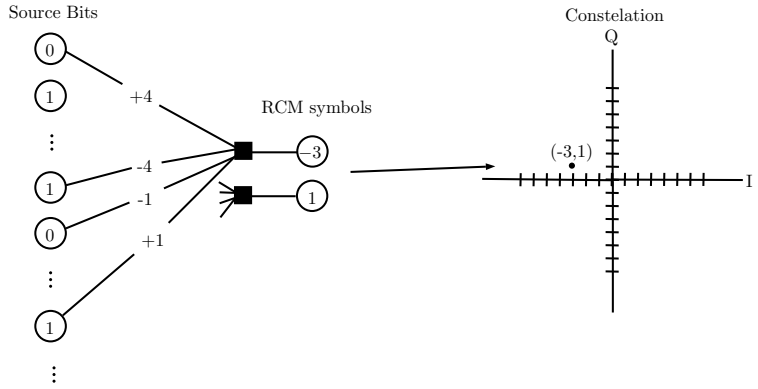


Figure 2.12: *Illustration of RCM: source bits are mapped to multi-level symbols through weighted sum operations. Symbols are paired two-by-two to form a QAM constellation.*

and RCM symbols have associated weights. These symbol-to-bit edges are randomly chosen. Specifically, If $\mathcal{W} = \{w_1, w_2, \ldots, w_{d_r}\}$, with $\{w_i\} \in \mathbb{N}\ \forall i \in \{1, \ldots, d_r\}$, is the design weight multi-set of an RCM code, then, its RCM symbols are generated as the sum of $2 \cdot d_r$ randomly chosen input bits weighted by the values in $\pm\mathcal{W} = \{w_1, w_2, \ldots, w_{d_r}, -w_1, -w_2, \ldots, -w_{d_r}\}$. Notice that a multi-set is a generalization of the concept of a set that, unlike a set, allows multiple instances of the multi-set's elements. The selection of this multi-set determines the characteristics of the mapping and the number of input bits, $2 \cdot d_r$, added in each multi-level symbol.

Let G, of size[4] $M \times K$, be the incidence sparse matrix of the RCM graph, where rows represent coded symbols, columns represent input bits and the nonzero entry (j, k) is the weight associated to the edge connecting symbol j with input bit k. The RCM encoding is performed as follows: First, the source information blocks are partitioned into blocks of length K

[4]Here we use M instead of N when considering only RCM codes, but it also refers to the coded sequence length.

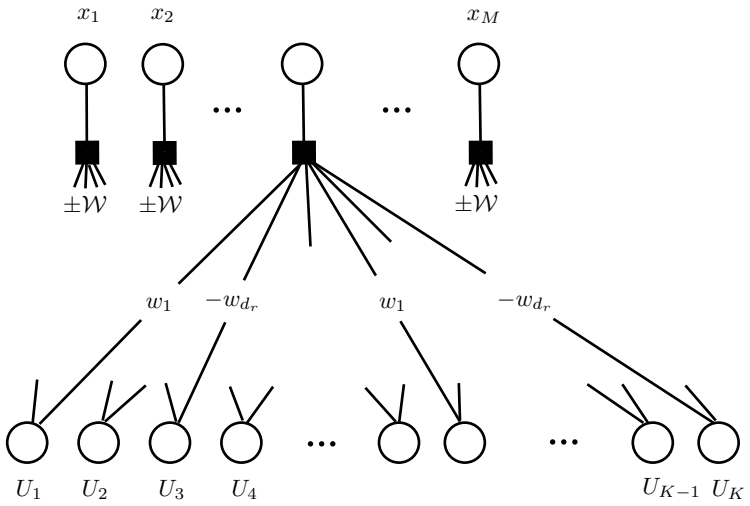


Figure 2.13: *Bipartite graph representation of an RCM code.*

denoted by $\mathbf{u}$, and then, the j^{th} RCM symbol, x_j, is obtained as

$$x_j = [G\mathbf{u}]_j = \sum_{k=1}^{K} g_{j,k} u_k, \qquad j \in \{1, \ldots, M\}, \tag{2.54}$$

where $[\,\cdot\,]_j$ is the element at row j, $g_{j,k}$ denotes entry (j,k) of matrix G, and the operations are in the real field. Each multi-level symbol takes values in the finite symbol alphabet $\Phi = \{\phi_1, \phi_2, \ldots, \phi_{|\mathcal{X}|}\}$.

The design of good RCM codes follows three constraints:

1. **Constraint 1**: G should be regular in rows.
   In order to achieve a fixed constellation, all sum operations of the generated RCM symbols should have the same weights. Therefore, G has to be regular in rows, i.e., each row contains the same number of non-zero entires drawn from a fixed weight multi-set.

   Despite the fact that the weight set is fixed, they should be randomly placed in order to create many symbols enabling the rateless property.

2. **Constraint 2**: G should be as regular as possible in regards to its columns.
   Since RCM codes can be viewed as a coded modulation scheme, it is important to measure its free distance d_{free}. Since the all-zero bit sequence generates an all-zero codeword, d_{free} is given by

$$d_{\text{free}} = \min_{\mathbf{u}\in\{0,1\}^K} \|G\mathbf{u}\|^2. \tag{2.55}$$

   The minimum is usually taken with source sequences $\mathbf{u}$ with only one bit 1. In this case, the minimum is given by $\|G_{[:,i]}\|^2$, where $G_{[:,i]}$ is the i^{th} column of G. Thus, G should be constructed such that $\min_i \|G_{[:,i]}\|^2$ is maximized.

3. **Constraint 3**: The weight set $\mathcal{W}$ should create diverse symbol values. The information carried in the multi-level symbols is upper bounded by the entropy of these symbols. In particular, if we denote by X the random symbol with possible outcomes $\{\phi_1, \phi_2, \ldots, \phi_{|\mathcal{X}|}\}$, which occur with probability $\{P_X(\phi_i)\}_{i=1}^{|\mathcal{X}|}$, at the output of the encoder, then

$$H(X) = -\sum_{i=1}^{|\mathcal{X}|} P_X(\phi_i) \log P_X(\phi_i) \geq R, \tag{2.56}$$

   where $H(X)$ is the entropy of the RCM symbols and R is the RCM code rate.

Construction of RCM encoding matrices

In this section we present the construction of the generation matrix G such that it has good properties at varying values of M through an example. If $d_r = 8$, and assuming K divisible of by d_r, then the construction of matrix G is given by the following steps:

1. Construct the $K/2 \times K$ matrix G_0 as

$$G_0 = \begin{bmatrix} \Pi(A_{w_3}) & \Pi(A_{w_4}) & \Pi(A_{w_1}) & \Pi(A_{w_2}) \\ \Pi(A_{w_1}) & \Pi(A_{w_2}) & \Pi(A_{w_3}) & \Pi(A_{w_4}) \\ \Pi(A_{w_4}) & \Pi(A_{w_3}) & \Pi(A_{w_2}) & \Pi(A_{w_1}) \\ \Pi(A_{w_2}) & \Pi(A_{w_1}) & \Pi(A_{w_4}) & \Pi(A_{w_3}) \end{bmatrix},$$

where $\Pi(\cdot)$ denotes random column permutations of a matrix, and A_{w_l} is a $K/8 \times K/4$ sparse matrix given by:

$$A_{w_l} = \begin{bmatrix} w_l & -w_l & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & w_l & -w_l & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & w_l & -w_l & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_l & -w_l \end{bmatrix}$$

with $w_l \in \mathcal{W}$, for $l \in \{1, \dots, d_r\}$.

2. Vertically stack as many G_0 as needed (and only as many rows as needed from the last one) until the required $M \times K$ G matrix is obtained.

2.6.2 LOW DENSITY GENERATOR MATRIX CODES

Random-like codes with iterative decoding such as Turbo codes [13] and low-density parity check (LDPC) codes [14, 15, 16] have been widely studied and are known to be able to approach capacity. The third and fourth generation (3G and 4G) of communication standards utilizes Turbo codes, while they are replaced by LDPC codes in 5G. Turbo and LPDC codes present high complexity in their decoding and encoding, respectively. This presents some challenges when employing them in some practical applications and leads the way to new random-like codes with low encoding and decoding complexity.

Low-Density Generator Matrix codes, or LDGM codes, are a random-like binary subclass of LDPC codes with a sparse generator matrix G_L, which allows to lower the encoding complexity. Moreover, since they utilize their encoding sparse matrix for decoding, have also low decoding complexity. Unlike standard LDPC codes, regular LDGM codes have excellent performance. However, LDGM codes have been historically disregarded due to their high error floors [16]. Later, authors in [17, 18] showed that LDGM codes can achieve performances similar to the state of the art codes if properly adapted. Moreover, the work in [19] shows that the parallel concatenation of two regular LDGMs codes can reduce the error floor significantly. The idea in [19] is to use a powerful low rate code in parallel with a high rate code. The low rate code is capable of correcting most of the errors, while the high rate code is designed to eliminate as many of the residual errors as possible. In light of their excellent performance, in this thesis we focus on systematic LDGM codes.

We denote by systematic to LDGM codes, whose generator matrices are of the form $G_L = \begin{bmatrix} I_K \\ P \end{bmatrix}$, where I_K is the identity matrix of size K and P is a $I \times K$ random-like sparse matrix whose non-zero entries are 1s. The matrix P is characterized by the pair (d_l, d_c), which denote the number of nonzero elements of a column and of a row, respectively. Given a block of K information bits $\mathbf{u}$, the LDGM coded Binary Phase-Shift Keying (BPSK) modulated sequence $\mathbf{x}$, of length $N = K + I$ and encoded with rate $\frac{K}{K+I}$, is given by

$$\mathbf{x} = 2 \cdot \left(\left(\begin{bmatrix} I_K \\ P \end{bmatrix} \mathbf{u} \mod 2 \right) - \frac{1}{2} \right) = \underbrace{[x_1, x_2, \ldots, x_K}_{systematic}, \underbrace{x_{K+1}, x_{K+2}, \ldots, x_{K+I}}_{non-systematic}]^\top, \tag{2.57}$$

where operations are in the binary field. Notice that the sequence $\mathbf{x}$ consists of the BPSK modulated source bit block $\{x_1, x_2, \ldots, x_K\}$ and of the non-systematic I LDGM coded BPSK modulated symbols given by $\{x_{K+1}, x_{K+2}, \ldots, x_{K+I}\}$.

2.6.3 PARALLEL RCM-LDGM CODES

The parallel concatenation of RCM and LDGM codes was first proposed in [6, 7] in order to lower the error floor exhibited by RCM codes. The main idea is to substitute a few analog RCM coded symbols by LDGM binary symbols so that the LDGM symbols are able to correct residual errors produced by the RCM code. This generates the hybrid scheme of Fig. 2.14 with two different sub-blocks, one them analog (RCM) and the other digital (LDGM). At the receiver, the RCM and LDGM sub-blocks perform decoding jointly using belief propagation/message passing. Interestingly enough, this parallel concatenation preserves the smooth rate adaption of RCM codes, since LDGM codes can also have rateless [20] properties.

A parallel RCM-LDGM code of rate K/N, with $N = M + I$, consists of a parallel concatenation of an RCM code of rate K/M with the I non-systematic coded binary symbols of a systematic LDGM code. Given a block of K information bits $\mathbf{u}$, the RCM-LDGM coded sequence $\mathbf{x}$ is

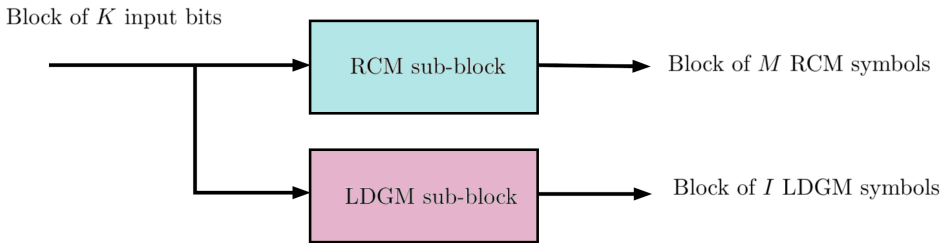


Figure 2.14: *Encoder diagram of the parallel concatenation of an RCM sub-block and an LDGM code.*

obtained as

$$\mathbf{x} = \begin{bmatrix} G\mathbf{u} \\ 2\left((P\mathbf{u} \mod 2) - \frac{1}{2}\right) \end{bmatrix} = \left[\underbrace{x_1, x_2, \ldots, x_M}_{RCM}, \underbrace{x_{M+1}, x_{M+2}, \ldots, x_{M+I}}_{LDGM}\right]^\top, \tag{2.58}$$

where G is the RCM generator matrix of Section 2.6.1 and P is the non-systematic part of the LDGM generator matrix. Finally, the M RCM symbols and the I BPSK modulated LDGM coded bits are grouped two-by-two and transmitted using a (Quadrature Amplitude Modulation) QAM modulator, so that the spectral efficiency, ρ, is

$$\rho = \frac{2 \cdot K}{M + I} \tag{2.59}$$

bits per complex channel symbol.

2.6.4 DECODER OF THE PARALLEL RCM-LDGM CODES

In this section we present the details of the SPA applied to the factor graph sketched in Fig. 2.15 that models the RCM-LDGM parallel concatenation. In order to link the notation utilized in RCM decoding with the previously developed FG theory, let us begin by providing some clarifications and/or definitions, connecting the previously developed graph theory with the notation encountered in the RCM-LDGM literature.

- The Variable Nodes (VN) represent source bits $\{U_1, \ldots, U_K\}$.

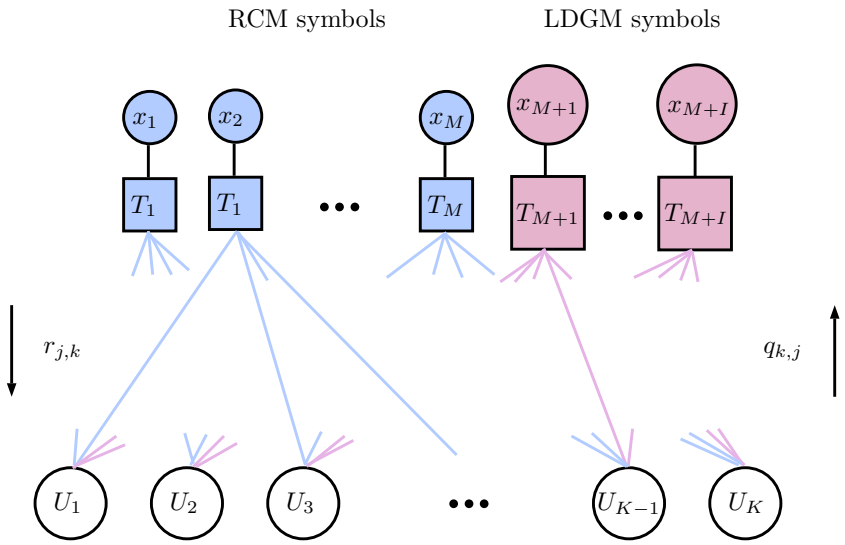


Figure 2.15: *Factor graph of a parallel RCM-LDGM code. Note that the source's and channel's FG are not plotted.*

- The function nodes (also referred as Check Nodes (CN)) represent the coded symbols $\{x_1, \ldots, x_N\}$.
- CNs $\{T_j\}_{j=M+1}^{M+I}$ (in pink) are LDGM and their function $T_j(\cdot)$ is the same as in standard LDGM coding.
- RCM CNs (in blue), $\{T_j\}_{j=1}^{M}$, are more complex and their function $T_j(\cdot)$ requires the use of direct convolutions, due to the arithmetic additions involved in the generation of these symbols.
- Messages $\mu_{T_j \to U_k}$: The messages passed from the j^{th} CN to the k^{th} VN are denoted by $r_{j,k}$.
- Messages $\mu_{U_k \to T_j}$: The messages passed from the k^{th} VN to the j^{th} CN are denoted by $q_{k,j}$.
- Sets $T_j \in \psi(U_k) \setminus T_j$: The set of CNs connected to VN U_k without considering CN T_j are denoted by $n(U_k) \setminus T_j$.
- Sets $U_i \in \psi(T_j) \setminus U_k$: The set of VNs connected to CN T_j without considering VN k are denoted by $n(T_j) \setminus U_k$.

One complication that this decoder addresses is that the messages usually exchanged in RCM are PMF., whereas Log Likelihood Ratio (LLR) messages are normally used in LDGM codes. As the source bits are binary, the PMF. can be represented by LLR values. Thus, the messages can be easily exchanged between the RCM sub-block and the LDGM sub-block by just using LLR values. The LLR values are only converted to a PMF. representation when updating the RCM CNs.

The details of the decoding algorithm are provided below. At each iteration t, and for a predefined number of iterations t_{max}, the sum-product algorithm is implemented by the sequential execution of the following steps:

- **STEP 1. $q_{k,j}^{(t)}$: Message passing from variable nodes, $\{U_k\}_{k=1}^{K}$, to RCM and LDGM check nodes $\{T_j\}_{j=1}^{M+I}$.**

$$q_{k,j}^{(t)} = \sum_{i \in n(U_k) \setminus T_i i} r_{i,k}^{(t-1)} + \log\left(\frac{\delta_k^{(t-1)}(u_k = 1)}{\delta_k^{(t-1)}(u_k = 0)}\right), \tag{2.60}$$

  where $r_{i,k}^{(0)} = 0$ for $k \in \{1, \ldots, K\}$, $i \in n(U_k) \setminus T_j$ and $\delta_k^{(t-1)}(u_k)$ is the posteriori information for the input variable node U_k is computed at time $t-1$ from the source's factor graph.

- **STEP 2. $r_{j,k}^{(t)}$: Message passing from RCM and LDGM check nodes, $\{T_j\}_{j=1}^{M+I}$, to variable nodes $\{U_k\}_{k=1}^{K}$.**

  - **Computation at RCM check nodes $\{T_j\}_{j=1}^{M}$:**
    Observe that the RCM symbol $x_j = \sum_i g_{j,i} u_i$ can be rewritten as the sum of two terms given by $x_j = a_{j,k} + g_{j,k} u_k$, where $a_{j,k} = \sum_{i \sim k} g_{j,i} u_i$ and $\sum_{i \sim k}$ means the sum over all i except k. In AWGN channels, given the received channel symbol $y_j = x_j + z_j = a_{j,k} + g_{j,k} u_k + z_j$, the message $r_{j,k}^{(t)}$ is calculated as

$$r_{j,k}^{(t)} = \log\left(\frac{\sum_{v \in \Phi} P^{(t)}(a_{j,k} = v) \cdot e^{-(y_j - v - g_{j,k})^2 / N_0}}{\sum_{v \in \Phi} P^{(t)}(a_{j,k} = v) \cdot e^{-(y_j - v)^2 / N_0}}\right), \tag{2.61}$$

    where the sum in v is over the RCM symbol alphabet Φ. Notice that $P^{(t)}(a_{j,k} = v)$, the probability of $a_{j,k} = v$ at iteration t, is

calculated in a straightforward manner by convolving the PFM of the terms in the summation, obtained from the LLR messages $q_{k,j}^{(t)}$. An efficient way to implement these convolutions in a zig-zag manner is explained in [5].

– **Computation at LDGM check nodes $\{T_j\}_{j=M+1}^{M+I}$:** As in standard LDGM codes, the LLR message transmitted from the j^{th} check node to the variable node U_k is given by

$$r_{j,k}^{(t)} = -2\text{atanh}\left(\tanh\left(-\frac{\gamma_j}{2}\right)\prod_{i\in n(T_j)\setminus U_k}\tanh\left(-\frac{q_{k,i}^{(t)}}{2}\right)\right), \tag{2.62}$$

where $\gamma_j = \dfrac{(y_j+1)^2-(y_j-1)^2}{N_0}$ is the channel LLR value.

- **Decision:** At the end of the iterations, when $t = t_{max}$, an estimate of u_k can be calculated as

$$\widehat{u_k} = \begin{cases} 1, & \left(\sum_{i\in n(U_k)} r_{i,k}^{(t_{max})} + \log\left(\dfrac{\delta_k^{(t_{max})}(u_k=1)}{\delta_k^{(t_{max})}(u_k=0)}\right)\right) > 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.63}$$

Note that the following algorithm can be used for only RCM codes by just setting $I = 0$.

Part I

Point-to-Point Communications

This first part of the dissertation considers the point-to-point transmission of the information generated by a single source. Concretely, we look for new design methods and applications to the RCM and parallel RCM-LDGM codes presented in Section 2.6.1. Recall that what made these codes interesting in the first place was (i) their joint source-channel coding capabilities, (ii) the way in which they embed the source compression directly into modulation by building simple linear combinations, (iii) the capacity of these codes to adapt their rate seamlessly by varying the number of generated symbols. Before explaining the proposed point-to-point coding schemes or design methods, aiming to provide further context to the utilized codes, let us start by discussing the related work on (i) joint source-channel coding, (ii) coded modulation and (iii) rate adaptation, as well as provide background information on compressive sensing theory, since the core of RCM is a new Random Projection (RP) code inspired by the Compressive Sensing (CS) theory.

Joint Source-Channel Coding

The transmission of information produced by binary single sources is usually separated into two independent sub-problems, named source and channel coding. This procedure simplifies the construction of the system, since the source encoder can be optimized without any knowledge of the channel statistics, while the channel encoder can be optimized irrespectively of the source. If required, each encoder can be adjusted leaving the other unchanged. However, many schemes treat the source compression and channel coding in a joint fashion under a unified resource constraint and/or toward a common optimization problem. This family of algorithms are called Joint Source-Channel Codes (JSCC) and, for point-to-point communications, are classified into two categories.

- Works in the first category apply to multimedia (such as video or image) transmission. For example, authors in [21, 22, 23] exploit the rate distortion behavior of the source and study optimal rate allocation schemes aiming to minimize the expected distortion, whereas the goal in [24] is to optimize the power consumption. Other works, such as [25] pair stronger channel codes to more important content, designing Unequal Error Protection (UEP) schemes for progressively encoded sources. The JSC coding works in this last group studied in the literature include the use of different codes, such as the rate-

compatible LDPC codes in [26], and rate-free Raptor codes, both over GF(2) [27] and GF(4) [28].

- The second category considers the use of the same code for both source compression and protection of a single source. Regular channel codes have been shown to work for source compression if concatenated with the block sorting Burrows-Wheeler Transform (BWT), such as LDPC codes [29], Raptor codes [30] and Turbo codes [31]. Furthermore, authors in [32] use two concatenated independent LDPC codes in the transmitter: one for source coding and the other for channel coding, with a joint belief propagation decoder. The works in [33] and [34] design Turbo codes for the transmission of non-uniform binary sources over noisy channels. However, none of the previous examples can achieve smooth rate adaptation.

Coded Modulation

The RCM code is essentially an enhanced coding modulation scheme for high spectral efficiencies with embedded source compression. Therefore, the following briefly reviews related work on coding and modulation. Ungerboeck [35, 36] was the first to propose an scheme of coded modulation, that he denoted as Trellis-Coded Modulation (TCM). The goal of that scheme is to maximize the minimum free Euclidean distance by introducing trellis coding and set partitioning in the physical layer. Another coded modulation scheme is Multi-Level Coded (MLC) modulation [37, 38], which divides the channel coding into as many independently designed levels as modulation bits. MLC present the advantage that any binary code designed on Hamming distance can be used.

When Rayleigh channels are considered, the code's robustness becomes vital. In order to address this issue [39] proposed Bit-Interleaved Coded Modulation (BICM) and later, authors in [40] presented an information theoretic analysis of BICM. On one hand, bit interleaving increases the channel diversity which is equivalent to increasing the Hamming distance. On the other hand, infinite depth interleaving tackles the mismatch decoding issue [41].

The invention of capacity-achieving turbo codes [42] and the rediscovery of LDPC codes [43, 44] boosted the development of contemporary coded modulation, which are able to approach the Shannon limit at the designed

SNR. However, they lack rate adaptation capabilities, and must rely on separate adaptive modulation to change the rate when the channel changes greatly. In addition, in order to achieve capacity performance, infinite block coding length and high complexity are usually required.

Rate Adaptation

The capacity of communication systems to adapt their coding rate, modulation order and power allocation based on varying channel conditions is denoted as rate adaptation. Among these techniques, the most researched and practically used is Adaptive Modulation and Coding (AMC) [45, 46, 47, 48]. However, ACM suffers from two important disadvantages. On one hand, the rate and modulation selection is based on a predefined set of combinations, whose limitation results in stair-shaped rate curves. On the other hand, the rate selection is highly dependent on an accurate and instantaneous channel estimate, which cannot be obtained simultaneously in reality.

Hybrid automatic repeat request (HARQ) [52] is a supplemental technique to AMC to combat the uncertain wireless conditions. In event of a decoding fail, type-I HARQ [49] discards the received data, whereas the incremental redundancy scheme, named type II HARQ [50, 51], combines it with the re-transmitted data. Type II HARQ is able to provide smoother rate adaptation utilizing either a rate compatible code, such as punctured Turbo code [53] or alternative rate-less codes, such as Raptor code [54]. However, the range of rate adaptation is still limited.

Aiming to achieve smooth and blind rate adaption, Gudipati and Katti [2] proposed a new scheme called Strider, which automatically achieves almost the optimal rate adaptation by adding a Minimum Distance Transformer (MDT) after the existing encoding and modulation. The free distance of the constellation is adapted to suit the estimated channel conditions. In parallel, Spinal codes were proposed in [3], which are able to cope with time-varying channel conditions by sequentially applying a pseudo-random hash function to the message bits to produce a sequence of coded symbols for transmission. However, all these schemes assume that the transmitted data is already compressed. Otherwise, the end-to-end performance would suffer significant degradation.

Compressive Sensing

The RCM coding scheme consists of a bit-to-symbol mapping that converts the input source bits into multi-level symbols through weighted linear combinations. Essentially, this is a Compressive Sensing (CS) [55, 56] scheme applied to a binary input. Thus, in the sequel we review CS theory and emphasize its relation to channel coding.

The CS theory, presented in [57, 56, 58] by Candes, Romberg, Tao and Donoho, states the sparsity of a signal can be exploited to recover it from far fewer samples than required by the NyquistShannon sampling theorem [59], i.e., an n-dimensional signal having a sparse or compressible representation can be reconstructed from m linear measurements even if $m < n$ [60], [56]. The reason is that the actual information of the sparse signal can be represented with only a few significant components. Examples of signals having sparse representation are images or speech signals.

Traditionally, CS considers sparse signals in $\mathbb{R}^N$ and dense projection matrices $G \in \mathbb{R}^M \times \mathbb{R}^N$. The original signal can be reconstructed by applying convex relaxation [61, 62], l_1-minimization [63], matching pursuit [64, 65, 66], iterative threshold methods [67], and subspace pursuit [68]. Although it is possible to extend these algorithms to binary signals and/or sparse projection matrices, their greedy iterative algorithm makes the decoding complexity very high.

CS can be used as a JSCC scheme [69, 70, 71, 72, 73]. The works in [70] and [73] study the energy performance of CS and show its robustness against channel fading and noise. The application of CS to compress correlated sources and protect the measurements under noisy channels was studied in [69, 71, 72], where it was shown that CS decoders can provide trade-offs between rate and decoding complexity. CS-based JSCC schemes have continuous rate-distortion performance in a noisy Gaussian channel. However, these works consider highly-sparse real-valued sensor readings of wireless sensor systems, whose decoding is based on computation-intensive convex optimization, due to the traditional use of CS techniques. Therefore, the applicability of such schemes for the transmission of generic binary symbols seems impractical.

By applying several restrictions to reduce its complexity, a less traditional application of CS is to use them as conventional channel codes. For

example, Sudocodes [74] can be used as erasure codes for real-valued data and reduce the encoding and decoding complexity by limiting the projection matrices to be sparse binary. The first to consider binary sources in CS were Wu et al. [75] and Liu et al. [76], who also derived close-form decoding solutions. However, they only considered the noiseless case. Authors in [77] provided a solution to decode noisy CS based on the belief propagation algorithm, but only considered projection matrices whose entries are drawn from Rademacher distribution.

We begin this part of the Thesis by proposing an EXIT chart analysis for RCM-LDGM codes in Chapter 3, with the purpose of speeding up their design. Chapter 4 considers the transmission of sources with memory, for which we build a joint source-channel coding scheme utilizing the Burrows-Wheeler transform and RCM-LDGM codes. The novelty of the proposed scheme lies in its high throughput and rate compatibility capabilities that it inherits from its constituent codes. Finally, in Chapter 5 we propose the use of RCM-LDGM codes for impulsive noise channels.

CHAPTER 3

EXIT Chart Analysis

In this chapter we propose the design of parallel RCM-LDGM codes using an EXIT chart analysis. The goal of these methods is to speed up the design procedure, avoiding time-consuming Monte Carlo simulations that are currently needed when designing RCM-LDGM codes (see [6, 7]). We consider the transmission of memoryless non-uniform binary sources (refer to Section 2.3.2) over AWGN and fast Rayleigh channels (refer to Section 2.4).

3.1 INTRODUCTION

EXIT charts were first introduced in [78] to study the iterative decoding behavior of two or more constituent decoders. They can be used to predict the asymptotic converge of a code, and to obtain an estimate of the lowest SNR needed for convergence, as well as an indication about the required number of decoding iterations for convergence, resulting in a great tool for designing iterative codes. For example, authors in [79] proposed a curve fitting procedure based on EXIT charts to design an LDPC code valid for modulation and detection. Given the iterative nature of the considered codes, the EXIT charts present a good method to visually explore the iterative exchange of information that occurs in the decoders of these schemes.

It should be mentioned that the authors in [4] were the first to use EXIT charts for RCM codes. However, the EXIT analysis for pure RCM codes is not straightforward to apply to the parallel RCM-LDGM coding scheme, since two different types of check nodes, RCM and LDGM, have to be considered (refer to Fig. 2.15). There are many works in the literature that consider multi-edge type EXIT charts. For example, the work in [80] utilizes this type of EXIT charts to design the bit mapping for LDPC coded BICM schemes, whereas authors in [81] optimize the bit mapping of LDPC coded modulation with APSK constellations. However, RCM-LDGM codes present the added difficulty of mixing analog and digital check nodes, and, therefore, strategies developed in [80, 81] can not be directly applied. In this chapter we present a novel EXIT chart for RCM-LDGM codes and a BER prediction analysis based on EXIT charts that was not previously considered in the literature for this type of codes.

3.2 RCM-LDGM EXIT CHARTS

The model for the EXIT chart is depicted in Fig. 3.1. As it can be seen, it is composed by two constituent decoders: the Variable Node Decoder (VND), composed by all source variable nodes, and the Check Node Decoder (CND), composed by RCM and LDGM check nodes. Note that RCM check nodes work in the real field, whereas LDGM check nodes work in the binary field. As shown in Section 2.6.4, the decoder interchanges LLR messages between VNs and CNs. We model these LLR messages as outcomes of real-valued r.v's E (output from either a VND or a CND) and A (input to either a VND or a CND), which are super-scripted by (vn) or (cn) depending on the decoder. For example, $E^{(vn)}$ models the extrinsic information produced at the VND whereas $A^{(vn)}$ is the a priori information arriving at the VND. Similarly, $E^{(cn)}$ and $A^{(cn)}$ are the (produced) extrinsic and (arrived) a priori information on the CND, respectively. In order to characterize the behavior of each consistent decoder, a function of the mutual information $I(E;U)$ between the extrinsic output E and the binary source U (with distribution $(p_1;p_0)$) is obtained. The input of this function is given by the mutual information $I(A;U)$ between the decoder's a priori input A and U. These mutual information are obtained by numerical

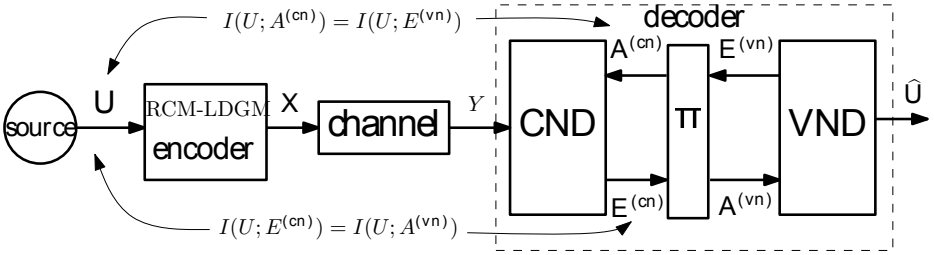


Figure 3.1: *Model for EXIT chart analysis. The Factor Graph of the VND and CND is given in Fig. 2.15.*

integration given by

$$I(L;U) = p_0 \int_{-\infty}^{\infty} f_L(\xi|0)\log_2\left(\frac{f_L(\xi|0)}{p_0 f_L(\xi|0) + p_1 f_L(\xi|1))}\right) d\xi + p_1 \int_{-\infty}^{\infty} f_L(\xi|1)\log_2\left(\frac{f_L(\xi|1)}{p_0 f_L(\xi|0) + p_1 f_L(\xi|1))}\right) d\xi, \quad (3.1)$$

where $L \in \{A, E\}$ and $f_L(\xi|u)$[1], for $u = 0, 1$, is the conditional PDF of L given U, which depends on the decoder under consideration. In the sequel, we will denote $I(L;U)$ for a VND or a CND as $\mathrm{I_{L,VND}} = I(L^{(vn)};U)$ or $\mathrm{I_{L,CND}} = I(L^{(cn)};U)$, respectively. Finally, $f_L(\xi|u)$ depends on the constituent decoder under consideration. In the next sections we show the calculations of $f_{A^{(vn)}}(\xi|u)$, $f_{A^{(cn)}}(\xi|u)$, $f_{E^{(vn)}}(\xi|u)$ and $f_{E^{(cn)}}(\xi|u)$.

Before deriving the EXIT chart for the parallel RCM-LDGM code, let us define $p_{RCM}^{(vn)}$ and $p_{RCM}^{(cn)}$ as the average percentage of links connected to an RCM CN (from the sum of RCM and LDGM edges) arriving at the VND and CND, respectively.

3.2.1 VND EXIT CURVE FOR RCM-LDGM CODES

The EXIT curve of the VND is given by the transfer characteristic between $\mathrm{I_{E,VND}} = I(E^{(vn)};U)$ and $\mathrm{I_{A,VND}} = I(A^{(vn)};U)$. Note that the realizations of r.v's $E^{(vn)}$ and $A^{(vn)}$, are the messages exchanged in the

[1]In this chapter we denote the PDF of the r.v. X by f_X instead of p_X to avoid confusions with p_1 and p_0, which stand for the probabilities of the nonuniform source.

sum-product algorithm, $\{r_{j,k}\}$ and $\{q_{k,j}\}$, respectively. In order to evaluate these mutual informations from (3.1), the conditional PDF of the a priori $A^{(vn)}$ and the extrinsic $E^{(vn)}$ at a variable node decoder, given U, have to be found.

Calculation of $\mathbf{I_{A,VND}}$

For RCM-LDGM codes, we have to consider two types of a priori messages arriving at the VND, $A^{(vn)}_{RCM}$ and $A^{(vn)}_{LDGM}$, which correspond to the messages arriving from an edge connected to an RCM and an LDGM, respectively. For convenience, the work in [82] modeled the considered conditional PDFs as the PDF of the LLR random variable obtained at the output of a virtual AWGN channel when its inputs are uniform binary source symbols U, i.e.,

$$Y = U + N, \quad N \sim \mathcal{N}\left(0, \sigma^2\right). \tag{3.2}$$

Assuming this model, $A^{(vn)}_{RCM}$ can be expressed as

$$A^{(vn)}_{RCM} = \log\left(\frac{P(u=1|Y)}{P(u=0|Y)}\right) = \frac{2U-1}{2\sigma^2} + \frac{N}{\sigma^2}.$$

By assuming the same for $A^{(vn)}_{LDGM}$, $A^{(vn)}_{RCM}$ and $A^{(vn)}_{LDGM}$ are expressed as

$$A^{(vn)}_{RCM}|U \sim \mathcal{N}\left((2u-1)\sigma^2_{R,A}/2, \sigma^2_{R,A}\right) \tag{3.3}$$

$$A^{(vn)}_{LDGM}|U \sim \mathcal{N}\left((2u-1)\sigma^2_{L,A}/2, \sigma^2_{L,A}\right), \tag{3.4}$$

where $\sigma^2_{R,A}$ and $\sigma^2_{L,A}$ represent the inverse of the variance of the two different virtual channels, that act as parameters. The main challenge of having two different types of CNs, is that the mutual information $\mathrm{I_{A,VND}}$ will now depend on two variables, $\sigma^2_{R,A}$ and $\sigma^2_{L,A}$. Notice, however, that although $A^{(vn)}_R|U$ and $A^{(vn)}_L|U$ can be considered independent, their variances are coupled due to the way the SPA generates the messages. Good knowledge (high values of LLR) of messages coming from one type of CN will likely be accompanied of similar messages from the other type of CN. Therefore, aiming to simplify the analysis by obtaining $\mathrm{I_{A,VND}}$ as a function of just one variable, we look for the relation between $\sigma^2_{R,A}$ and $\sigma^2_{L,A}$ and try to express one as a function of the other.

Calculating numerically the exact coupling between $\sigma^2_{R,A}$ and $\sigma^2_{L,A}$, as a function $\sigma^2_{R,A} = f(\sigma^2_{L,A}, \mathrm{SNR})$, for the range of SNR of interest (i.e., the

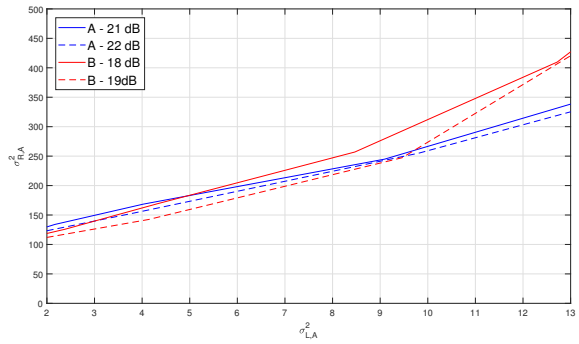


Figure 3.2: *Examples of $\sigma^2_{R,A}$ versus $\sigma^2_{L,A}$ curves obtained by Monte Carlo simulations for two different RCM-LDGM codes considering A (fast fading Rayleigh channel) and B (AWGN channel).*

SNRs belonging to the waterfall region of the code), is computationally expensive, which is counter to the objective of EXIT chart analysis. Fortunately, simulation results like the ones in Fig. 3.2 have shown that this dependency can be approximated linearly.

Therefore, in what follows we will assume that $\sigma^2_{R,A} = f(\sigma^2_{L,A})$ can be approximated in the range of SNR of interest by $\sigma^2_{L,A} = \frac{\sigma^2{}_{R,A}}{\kappa}$, yielding

$$A^{(vn)}_{RCM}|U \sim \mathcal{N}\left((2u-1)\frac{\sigma^2_{R,A}}{2}, \sigma^2_{R,A}\right) \tag{3.5}$$

$$A^{(vn)}_{LDGM}|U \sim \mathcal{N}\left((2u-1)\frac{\sigma^2_{R,A}}{2\cdot\kappa}, \frac{\sigma^2_{R,A}}{\kappa}\right). \tag{3.6}$$

The constant κ scales the variance of the distribution of $A^{(vn)}_{LDGM}$ with respect to the variance of $A^{(vn)}_{RCM}$, and is obtained as explained in Section 3.2.2. Once κ is found, and utilizing (3.5) and (3.6), the corresponding conditional PDF of $A^{(vn)}|U$, is obtained as

$$A^{(vn)}|U \sim f_{A^{(vn)}}(a|u) = \\ f_{A^{(vn)}_{RCM}}(a|u)p^{(vn)}_{RCM} + f_{A^{(vn)}_{LDGM}}(a|u)(1-p^{(vn)}_{RCM}), \tag{3.7}$$

and finally, $\mathrm{I_{A,VND}}$ is calculated from $f_{A_{RCM}^{(vn)}}$ as a parametric expression of $\sigma_{R,A}^2$ by applying (3.1).

3.2.1.1 Calculation of $\mathbf{I_{E,VND}}$

Once $f_{A^{(vn)}}(a|u)$ is calculated, the conditional distribution of the extrinsic r.v. $E^{(vn)}$ at the variable node decoder can be computed. Notice that there are two types of LLR messages: (i) ones passed on an edge connecting a VN to an RCM check node (modeled as a r.v. denoted as $E_{RCM}^{(vn)}$), and (ii) others passed on an edge connecting a r.v. to an LDGM check node (modeled as r.v. $E_{LDGM}^{(vn)}$). From the corresponding connections of the factor graph, we obtain

$$E_{RCM}^{(vn)} = \underbrace{A_{RCM}^{(vn)} + \ldots + A_{RCM}^{(vn)}}_{(\overline{d}_{RCM}^{(v)}-1) \quad \text{i.i.d r.v's}} + \underbrace{A_{LDGM}^{(vn)} + \ldots + A_{LDGM}^{(vn)}}_{d_l \quad \text{i.i.d r.v's}} + \log\left(\frac{p_1}{p_0}\right) \tag{3.8}$$

$$E_{LDGM}^{(vn)} = \underbrace{A_{RCM}^{(vn)} + \ldots + A_{RCM}^{(vn)}}_{\overline{d}_{RCM}^{(v)} \quad \text{i.i.d r.v's}} + \underbrace{A_{LDGM}^{(vn)} + \ldots + A_{LDGM}^{(vn)}}_{(d_l-1) \quad \text{i.i.d r.v's}} + \log\left(\frac{p_1}{p_0}\right), \tag{3.9}$$

where $\overline{d}_{RCM}^{(v)}$ and d_l are the average (and exact) number of connections between each VN and RCM and LDGM check nodes, respectively. The corresponding conditional PDFs of the extrinsic LLR messages, $E_{RCM}^{(vn)}$ and $E_{LDGM}^{(vn)}$, are

$$f_{RCM}(e|u) = \mathcal{N}\left((2u-1)\frac{\sigma_{R,E}^2}{2} + \log\left(\frac{p_1}{p_0}\right), \sigma_{R,E}^2\right)$$

$$f_{LDGM}(e|u) = \mathcal{N}\left((2u-1)\frac{\sigma_{L,E}^2}{2} + \log\left(\frac{p_1}{p_0}\right), \sigma_{L,E}^2\right),$$

respectively, where

$$\sigma_{R,E}^2 = \sigma_{R,A}^2\left(\overline{d}_{RCM}^{(v)} - 1 + \frac{d_l}{\kappa}\right),$$

$$\sigma_{L,E}^2 = \sigma_{R,A}^2\left(\overline{d}_{RCM}^{(v)} + \frac{d_l - 1}{\kappa}\right).$$

Again, since we have two types of extrinsic messages, the overall conditional PDF of the extrinsic LLR random variable $E^{(vn)}|U$ is obtained as

$$E^{(vn)}|U \sim f_{E^{(vn)}}(e|u) = f_{RCM}(e|u)p_{RCM}^{(cn)} + f_{LDGM}(e|u)(1-p_{RCM}^{(cn)}). \quad (3.10)$$

Finally, applying (3.1) $\mathrm{I_{E,VND}}$ is also calculated from $f_{E_{RCM}^{(vn)}}$ as a parametric expression of $\sigma_{R,A}^2$.

3.2.2 COMPUTING κ

κ is computed by Monte Carlo simulation through the following iterative procedure:

1. Start with an initial value of κ in (3.6) (such as $\kappa = 1$), and select a value of $\sigma_{R,A}^2$ so that the corresponding value of the mutual information computed by the PDF in (3.5) is in the range (0.5,0.9). For the value of $\sigma_{R,A}^2$ under consideration, generate the extrinsic messages passed from the VN to the RCM and LDGM check nodes according to (3.8) and (3.9), respectively.

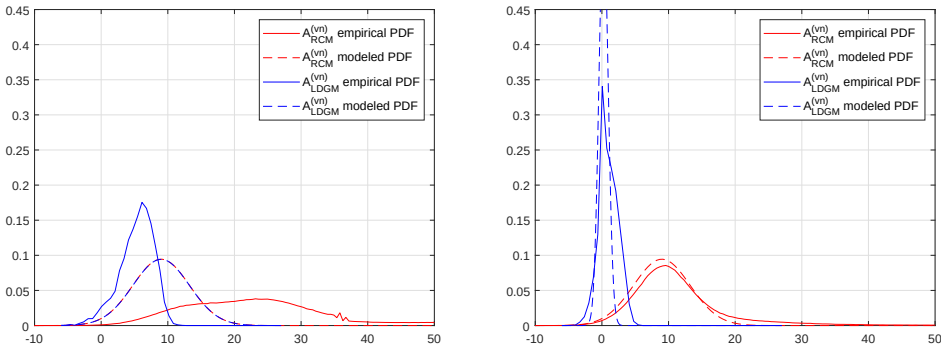
2. Obtain the empirical conditional PDFs of the extrinsic LLR messages passed from each LDGM and RCM check nodes to the VN numerically by running the SPA and using the above generated extrinsic messages.

3. Define κ_1 as the ratio between the variances of the empirical conditional distributions of RCM and LDGM check nodes obtained in step 2.

4. Repeat the previous 3 steps, using κ_1 as the initial value for κ, until the value of κ_1 in step 3 is close enough to the value of κ_1 in the previous iteration.

5. Set $\kappa = \kappa_1$ in the distribution (3.6).

Figure 3.3 shows a graphical example of the steps followed to find κ. The initial empirical conditional PDFs (i.e., when $\kappa = 1$) are shown in Fig. 3.3a. As it can be observed, none of the LLR messages is appropriately modeled at this point, since the initial value for κ ($\kappa = 1$) was chosen arbitrarily.



(a) iteration 1 (b) iteration 2

Figure 3.3: *Example of the iterative procedure to find κ. The histograms (conditioned to $U = 1$) obtained by Monte Carlo simulations are plotted in blue and the corresponding modeled conditional PDFs in red (refer to (3.5) and (3.6)).* **a** *Iteration 1.* **b** *Iteration 2.*

Note that since $\kappa = 1$, the modeled $A^{(vn)}_{RCM}$ is equal to $A^{(vn)}_{LDGM}$. The value of κ obtained in step 3 is 29, and the corresponding empirical conditional distributions are shown in the Fig. 3.3b. Notice that for this value of κ, the messages are better modeled by (3.5), (3.6). This second iteration results in $\kappa = 26$. Although not shown, after continuing to perform iterations we obtain $\kappa = \approx 34$ repeatedly, which will be set as the final value. We have found that in all simulated cases the number of iterations required for $\kappa_0 \approx \kappa_1$ is around three.

3.2.3 CND EXIT CURVE FOR THE RCM-LDGM CODES

From the fact that the a priori information $A^{(cn)}$ at the check node decoder is equal to the extrinsic information $E^{(vn)}$ at the variable node decoder (refer to Fig. 3.1), the PDF $f^{(cn)}_A(a|u)$ of $A^{(cn)}$ is given by the PDF in (3.10), interchanging E with A

$$A^{(cn)}|U \sim f_{A^{(cn)}}(a|u) = \\ f_{RCM}(a|u)p^{(cn)}_{RCM} + f_{LDGM}(a|u)(1 - p^{(cn)}_{RCM}), \tag{3.11}$$

so that $I_{\mathrm{A,CND}} = I_{\mathrm{E,VND}}$.

To compute $I_{\mathrm{E,CND}}$, we need to find the conditional PDF $f_{E^{(cn)}}(e|u)$ of the extrinsic LLR $E^{(cn)}$ at the CND. This is done by running Step 2 of the sum-product algorithm (see Section 2.6.4) and setting $q_{k,i} = a$, where a are realizations of a random variable $A^{(cn)}$ with conditional PDF (3.11). The empirical conditional PDF $f_{E^{(cn)}}(e|u)$ is now found by the histogram of the realizations $\{r_{j,k}\}$.

Remark: Observe that the EXIT Curve of VND only depends on the values of $\overline{d}_{RCM}^{(v)}$, and d_l. On the other hand, the EXIT Curve of the CND depends on all the parameters of the code and source, i.e., $\{\mathcal{W}$, SNR, $\overline{d}_{RCM}^{(v)}$, d_l, p_0, M, $I\}$.

3.2.4 TRAJECTORIES OF ITERATIVE DECODING AND DECODING THRESHOLD

To account for the iterative nature of the decoding process, both the VND and CND transfer characteristics should be plotted into a single diagram. As long as the SNR is large enough so that both transfer curves do not intersect, the iterative process will achieve its maximum mutual information values, given by the source's entropy, $(H(p_0), H(p_0))$, consequently achieving a low BER. The smallest SNR value for which both curves do not intersect is defined as the decoding threshold, and represents the minimum SNR required to decode without errors an infinite length code with the given configuration. Therefore, the code design problem reduces to find a code configuration, i.e., weight sets $\mathcal{W}$, and parameters I and d_l, such that the decoding threshold is as close as possible to the corresponding SNR Shannon limit. Practical codes are finite-length and the SNR required for low BER is larger than the theoretical decoding threshold. However, simulations show that the BER vs SNR behavior of finite-length codes is consistent with the decoding threshold of their infinite length counterpart. For more details on this procedure, we refer to the simulation results section.

Remark: The EXIT chart for a pure RCM code can be calculated as a particular case of the parallel LDGM-RCM by taking $p_{RCM}^{(vn)} = p_{RCM}^{(cn)} = 1$.

3.2.5 PREDICTING THE BER FROM THE EXIT CHART

For those SNR values smaller than the decoding threshold, the EXIT chart can be used to obtain an estimate on the BER after an arbitrary

number of iterations. Following the sum-product algorithm, let us define the decision variable of VN k as s_k, given by the sum of all LLR messages $r_{j,k}$ that were passed over a single edge connecting a CN, j, with the corresponding VN, k, i.e., $s_k = \sum_{j \in n(U_k)} r_{j,k} + \log\left(\frac{p_1}{p_0}\right)$. As assumed in the previous sections, the messages $r_{j,k}$ can be considered to be a realization of the independent Gaussian random variables $A^{(vn)}_{RCM}$ and $A^{(vn)}_{LDGM}$. Thus, the conditional PDF of S_k given U is

$$S|U \sim \mathcal{N}\left(\mu_S(u), \sigma_S^2\right),$$

with $\sigma_S^2 = \sigma_{R,A}^2 \cdot \left(\overline{d}^{(v)}_{RCM} \cdot + \frac{d_l}{\kappa}\right)$, and $\mu_S(u) = (2u-1)\frac{\sigma_S^2}{2} + \log\left(\frac{p_1}{p_0}\right)$. The BER performance is now obtained as

$$\mathcal{P}(\hat{u}_k \neq u_k) = p_0 \cdot \mathcal{P}\left(S > 0 | U = 0\right) + p_1 \cdot \mathcal{P}\left(S \leq 0 | U = 1\right). \tag{3.12}$$

Observe that

$$\mathcal{P}\left(S > 0 | U = 0\right) = 1 - Q\left(\frac{\mu_S(0)}{\sigma_S}\right) \text{ and,}$$
$$\mathcal{P}\left(S \leq 0 | U = 1\right) = Q\left(\frac{\mu_S(1)}{\sigma_S}\right),$$

where $Q(\xi)$ is the Q function

$$Q(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\xi}^{\infty} \mathrm{e}^{-\frac{y^2}{2}} dy.$$

3.3 NUMERICAL RESULTS

In this section, we evaluate the proposed EXIT chart analysis and BER prediction method of Section 3.2 for both AWGN and fast fading Rayleigh channels. We begin by considering the AWGN channel. Subsection 3.3.1 presents some mutual information trajectories of actual codes on the corresponding EXIT charts. In Subsection 3.3.2, we compare the BER predictions obtained using these charts with the BER obtained by Monte Carlo simulations. In Subsection 3.3.3, the EXIT analysis is used to obtain codes that approach the Shannon theoretical limit. Finally, the extension to Rayleigh channels is considered in Subsection 3.3.4.

3.3.1 TRAJECTORIES

We begin in Fig. 3.4 by showing the EXIT chart of a pure RCM code with weight set $\mathcal{W} = \{2, 3, 4, 8\}$ and spectral efficiency $\rho = 7.4$ for three different SNR values, 17, 18 and 20.25 dB, and for a non-uniform source with entropy $H(p_0) = 0.72$ ($p_0 = 0.8$). Notice that the variable node curve (which is valid for all SNRs) ends at the point $(H(p_0), H(p_0))$ as it should be. Also plotted in the figure are the Monte Carlo simulated mutual information trajectories of this code with block length $K = 37000$ (and $M = 10000$). Each trajectory is plotted using the same color as their corresponding SNR's EXIT chart CN curve, and they end where the corresponding CN and VN curves intersect. In addition, the contour lines of the corresponding simulated BERs are also shown. For example, at SNR= 17 dB the BER of the code is $5.5 \cdot 10^{-2}$ and, as observed, the blue curve of the EXIT chart intersects the VN decoder curve very close to the $5.5 \cdot 10^{-2}$ contour line. Similarly, for SNR= 18 dB and 20.25 dB, the simulated BERs are $3.3 \cdot 10^{-2}$ and $2.2 \cdot 10^{-3}$, respectively. Again, the intersections between CN and VN curves occur very close to the corresponding BER's contour lines. Note however, that none of these SNRs allow the channel to be open.

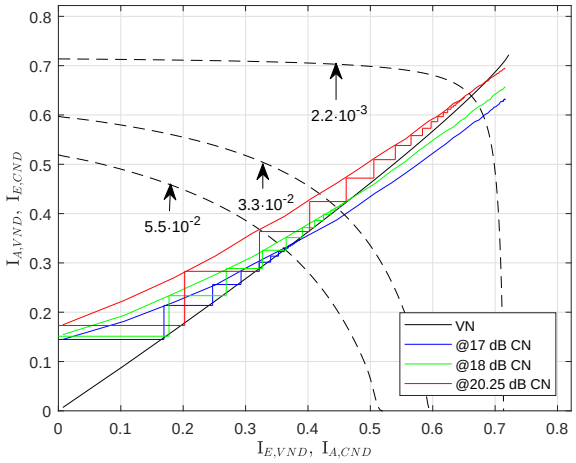


Figure 3.4: *EXIT chart, BER contour lines and mutual information trajectory for a pure RCM code of $\rho = 7.4$ when transmitting a non-uniform source with entropy $H(p_0) = 0.72$ over an AWGN channel.*

Similarly, Fig. 3.5 shows the EXIT chart and mutual information trajectories for the RCM-LDGM code obtained by substituting 200 RCM symbols by 200 LDGM coded binary symbols (with $d_l = 1$) in the previous RCM configuration.

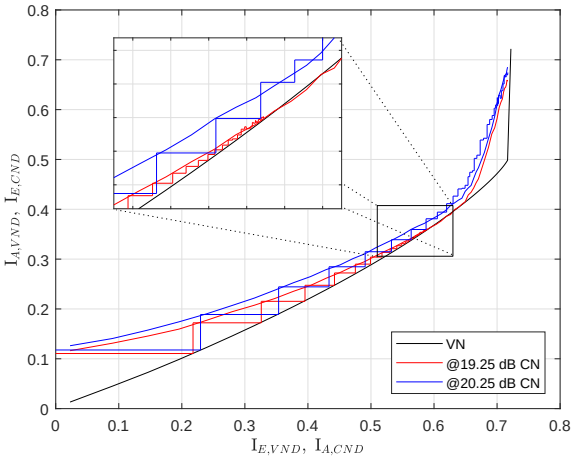


Figure 3.5: *EXIT chart and mutual information trajectory of a RCM-LDGM code when transmitting a non-uniform source with entropy $H(p_0) = 0.72$ over an AWGN channel.*

Observe that by introducing these 200 LDGM coded bits, we avoid the previous intersection of the curves at SNR 20.25 dB, improving in this way the BER at 20.25 dB. The corresponding mutual information trajectory at SNR 20.25 is shown in Fig. 3.5. Since the channel is open, it reaches its maximal value i.e., $(0.72, 0.72)$. It turns out that SNR=20.25 is the smallest SNR that allows the channel to remain open, and as such, it is the corresponding decoding threshold of the given code. In the same figure, the trajectory at SNR=19.25 dB is also shown, but in this case the channel is closed and and does not reach the maximum value.

3.3.2 BIT ERROR RATE FROM THE EXIT CHARTS

As explained in section 3.2.5, an estimated BER can be assigned to each point of the VND curve of the EXIT chart. Therefore, the BER of

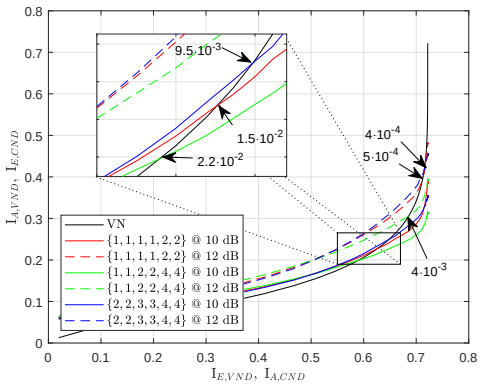
a particular code at a given SNR is obtained from the value of the point where the CND and VND curves intersect.

In this section we will consider two different RCM-LDGM configurations with $\rho = 4$ given by $K = 25000$, $M = 12365$, $I = 135$, with d_l 1 and 2. Moreover, we will consider three different sources with $p_0 = 0.5$, $p_0 = 0.8$ and $p_0 = 0.95$ and three different weight sets $\mathcal{W} = \{1, 1, 1, 1, 2, 2\}$, $\{1, 1, 2, 2, 4, 4\}$ and $\{2, 2, 3, 3, 4, 4\}$. Recall that the VN curve of the EXIT chart depends on M, I, d_l, and $\overline{d}_{RCM}^{(v)}$, whereas the CND curve depends also on the actual values of $\mathcal{W}$ and on the SNR.

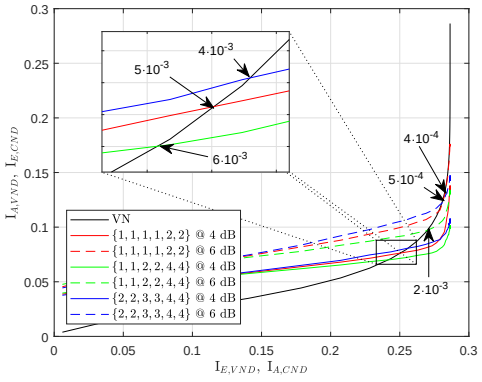
Figure 3.6a shows the EXIT chart of the configuration with $d_l = 1$ for the different weight sets and two SNR values, 10 and 12 dB. The plot shows the BER estimated values at the intersecting points. For example, for the configuration described above with $\{1, 1, 1, 1, 2, 2\}$ the estimated BERs are $1.5 \cdot 10^{-2}$ at SNR= 10 dB and $5 \cdot 10^{-4}$ at SNR= 12 dB. When $\{1, 1, 2, 2, 4, 4\}$, the estimated BERs are $2.2 \cdot 10^{-2}$ at SNR= 10 dB and $3 \cdot 10^{-3}$ at SNR= 12 dB. Finally, for $\{2, 2, 3, 3, 4, 4\}$ we obtain $9.5 \cdot 10^{-3}$ and $4 \cdot 10^{-4}$, respectively. Similarly, Fig. 3.6b shows the EXIT curves at SNR= 4 and 6 dB for the configuration with $d_l = 2$ with their estimated BER values when transmitting a source with $p_0 = 0.95$. Since $H(0.95) = 0.28$ the VN curve ends at point (0.28,0.28).

In order to corroborate our BER predictions, Fig. 3.7 compares the BER curves obtained by Monte Carlo simulation with those obtained by the EXIT charts BER estimation, as it is done in Fig. 3.6. As it can be seen in the figure, the prediction are accurate for both uniform and non-uniform sources.

The parameters of these codes have not been optimized and therefore, they present a large gap to the corresponding Shannon limits given by $10 \cdot \log_{10}(2^{\rho \cdot H(S)} - 1)$, which correspond to 0.91, 8.06 and 11.76 dB for $p_0 = 0.95$, $p_0 = 0.8$ and $p_0 = 0.5$, respectively. In the next section we will obtain near capacity high spectral efficiency codes using the EXIT chart analysis.

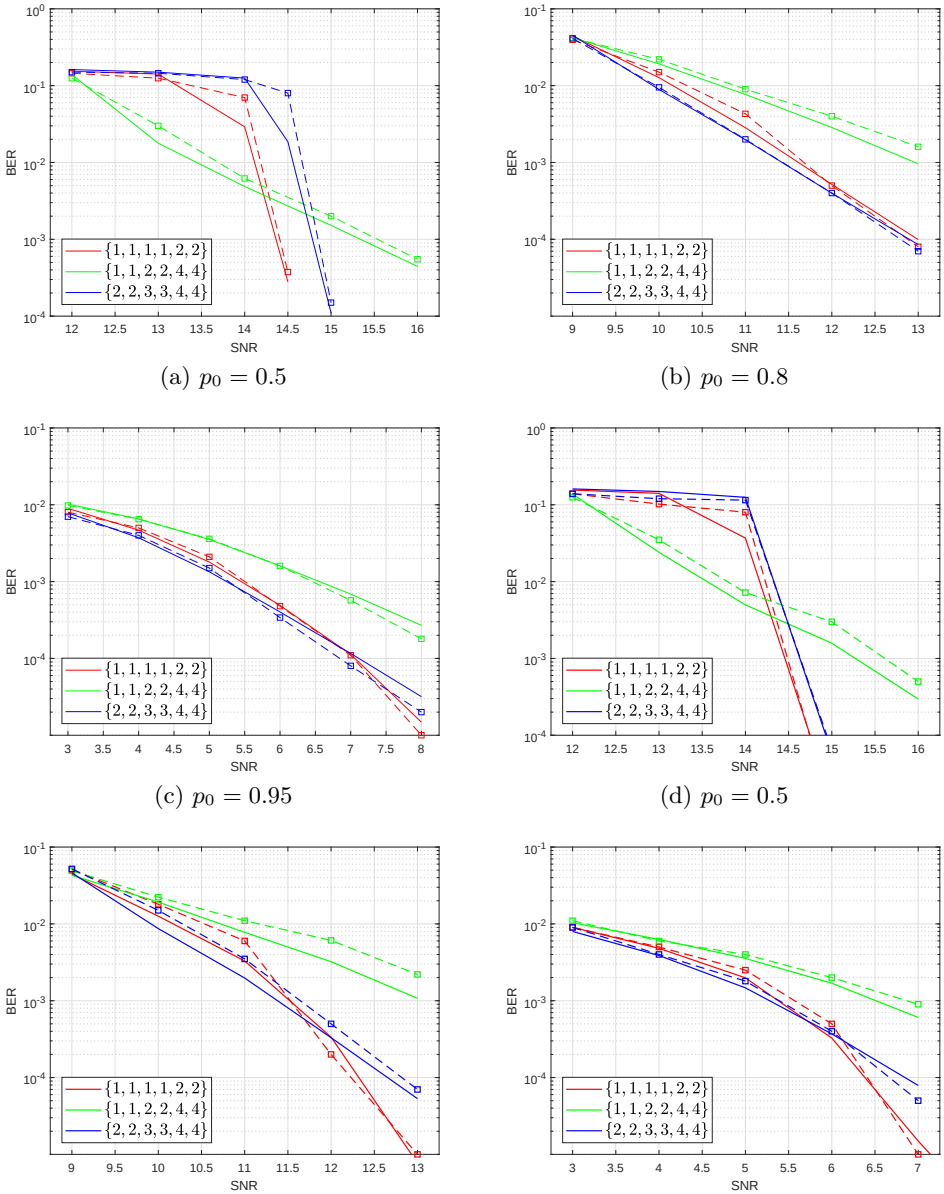


(a)



(b)

Figure 3.6: *EXIT chart and predicted BER at the crossing points for different SNR values.* **a** *for a non-uniform source with* $p_0 = 0.8$ *and the configuration with* $d_l = 1$. **b** *for a non-uniform source with* $p_0 = 0.95$ *and the configuration with* $d_l = 2$.



(a) $p_0 = 0.5$ (b) $p_0 = 0.8$

(c) $p_0 = 0.95$ (d) $p_0 = 0.5$

(e) $p_0 = 0.8$ (f) $p_0 = 0.95$

Figure 3.7: *Predicted (dashed line) and Monte Carlo simulated (continuous line) BER vs SNR curves for the RCM-LDGM codes.* ***a-c*** *The configuration with* $d_l = 1$. ***d-f*** *The configuration with* $d_l = 2$.

3.3.3 DECODING THRESHOLD BASED DESIGN

The idea behind the design method is to start with a pool of possible codes having the required rate, and then obtain the EXIT charts for the source of interest. The codes with the lower decoding threshold or those whose curves intersect closer to the maximum point $(H(p_0), H(p_0)$ are kept. The resulting subset of codes are then tuned-up by slightly changing their designed parameters. We have observed the following trends:

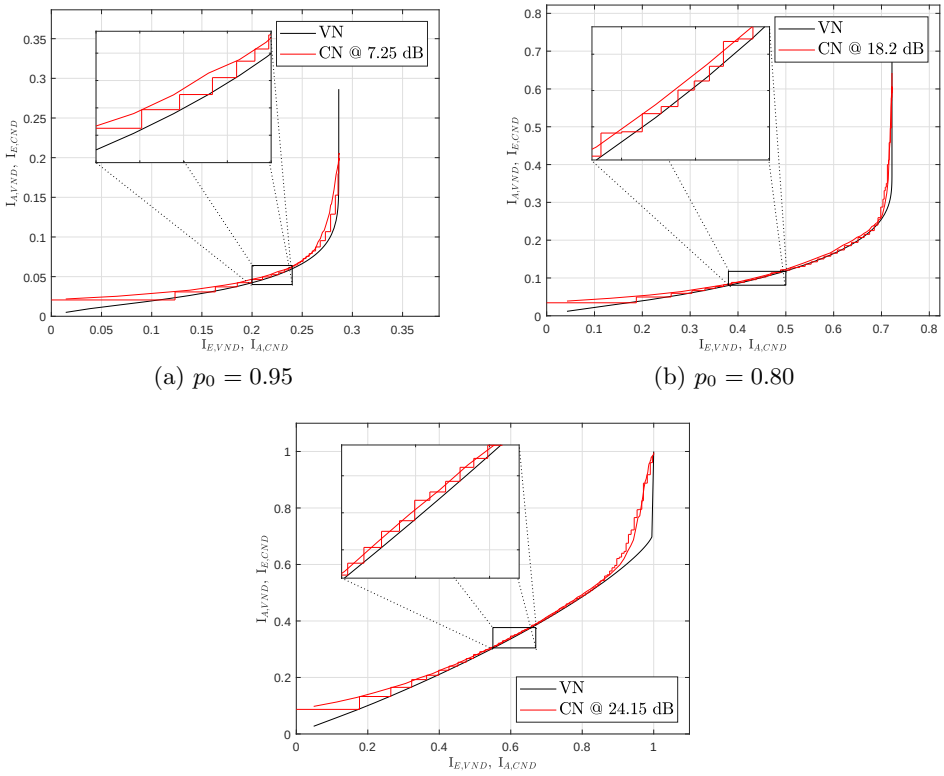
1. For sources with smaller entropy, larger RCM weight sets, $\mathcal{W}$, tend to work better, since the sum-product algorithm is aided by the a priori probability.
2. When designing the LDGM part of the code, there is a trade-off regarding the number I of LDGM binary symbols. By increasing I, more residual errors are corrected in the the waterfall region, making it steeper. However, larger SNRs are required to reach this waterfall region.
3. The range for parameter d_l is between 1 and 5. The larger parameter I is, the larger value for d_l can be selected.

Next, we provide EXIT chart design examples for an RCM-LDGM code with a spectral efficiency $\rho = 7.4$ for the transmission over AWGN channels of three memoryless sources with a priori probabilities $p_0 = 0.5$, $p_0 = 0.8$ and $p_0 = 0.95$. The corresponding Shannon limits are at 22.25 dB, 15.97 dB and 5.24 dB, respectively. Table 3.1 shows the best codes obtained by the EXIT chart analysis for a code length K=37000 bits.

The corresponding EXIT charts and real trajectories of the designed codes are plotted in Fig. 3.8. Note from Fig. 3.8c that when transmitting

Table 3.1: *Best configurations obtained by the EXIT chart analysis for AWGN channels.*

p_0	K	M	I	d_l	$\mathcal{W}$	Decoding Threshold.
0.5	37000	9800	200	1	$\{2, 3, 4, 8\}$	24.15 dB
0.8	37000	9720	280	4	$\{2, 2, 3, 3, 4, 8\}$	18.2 dB
0.95	37000	9200	800	3	$\{1, 1, 1, 1, 1, 1, 1, 1, 1, 1\}$	7.25 dB



(a) $p_0 = 0.95$ (b) $p_0 = 0.80$

(c) $p_0 = 0.5$

Figure 3.8: *EXIT chart and real trajectories of the designed codes for AWGN channels and sources with* ***c*** $p_0 = 0.5$, ***b*** $p_0 = 0.8$ *and* ***a*** $p_0 = 0.95$.

the uniform source ($p_0 = 0.5$), the channel between both EXIT curves remains open at SNR=24.15 dB (1.9 dB from the Shannon limit) and, consequently, the received blocks should be decoded with a low probability of error. However, if this same code were used to transmit the symbols generated by the non-uniform source ($p_0 = 0.8$) by only modifying the a priori probability to 0.8 in the SPA, the decoding threshold would decrease to 20.15 dB (refer Fig. 3.5). This is still 4 dB away from the Shannon limit. The optimized code for this source is given in Table 3.1. Notice that the gap is reduced to 2.23 dB (refer to Fig. 3.8b). This clearly shows that when transmitting binary symbols generated by a non-uniform source, the

channel code behaves like a joint source-channel code and, therefore, it has to be designed according to the source. Finally, Fig. 3.8a plots the EXIT chart of the best found configuration when transmitting the non-uniform source with $p_0 = 0.95$. The decoding threshold is only 2.01 dB away from the Shannon limit.

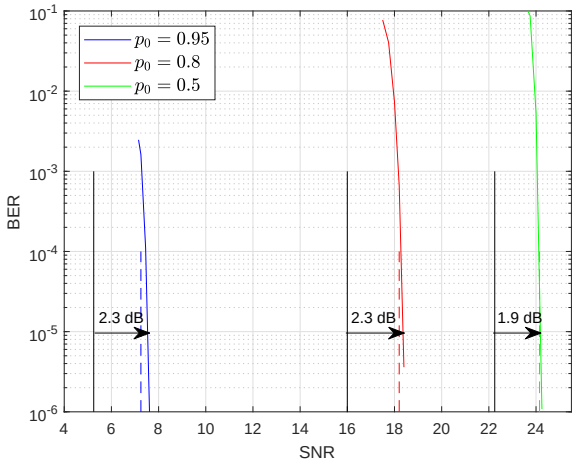


Figure 3.9: *BER vs SNR behavior of the obtained codes for $p_0 = 0.5$, $p_0 = 0.8$ and $p_0 = 0.95$. The Shannon limits (continuous lines) and the predicted decoding threshold (dashed lines) are plotted as vertical lines.*

In order to corroborate that the codes obtained from the EXIT chart analysis perform as expected, Fig. 3.9 plots the BER vs SNR curves obtained by Monte Carlo simulations, as well as the theoretical decoding thresholds for the designed codes (shown as vertical dashed lines), and the corresponding Shannon limits (vertical black lines). Note from the Monte Carlo simulations that the code designed for $p_0 = 0.5$ is 1.9 dB away from its Shannon limit for a BER= 10^{-5}, while the codes optimized for the sources with $p_0 = 0.8$ and $p_0 = 0.95$ present both a gap of 2.3 dB with respect to the Shannon limits. The figure indicates that the decoding threshold obtained from the EXIT chart analysis very accurately predicts the waterfall region. For the source with $p_0 = 0.5$, the gap between the decoding threshold and the waterfall region at BER=10^{-5} is not appreciable, whereas for the sources with $p_0 = 0.8$ and $p_0 = 0.95$ these gaps are 0.1 dB and 0.3 dB,

respectively. An explanation for the gap increase is that for non-uniform sources longer blocks are required to maintain stationarity.

Table 3.2: *Computational time required to predict a BER vs SNR point in Fig. 3.9.*

BER	EXIT chart	Monte-Carlo simulation		reduction factor
		average convergence time per block	blocks for 10 errors	
10^{-3}	10s	229s	1	22.9
10^{-4}	10s	113s	3	33.9
10^{-5}	10s	96s	27	259.2
10^{-6}	10s	85s	270	2295

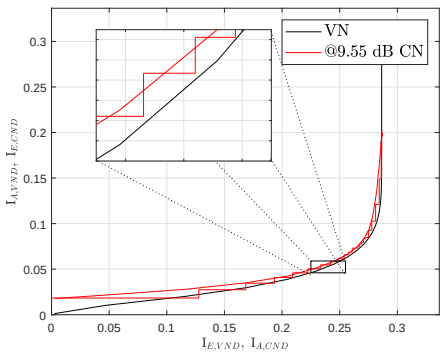
Table 3.2 summarizes the simulation time, as well as the computational time of the EXIT chart analysis, required to obtain the BER vs SNR points of the code $K = 37000$, $M = 9800$, $I = 200$, $d_l = 1$ and $\mathcal{W} = \{2, 3, 4, 8\}$. As shown in the table, the EXIT chart analysis is much faster than the simulations, making the search by trial and error feasible.

3.3.4 EXTENSION TO FAST FADING RAYLEIGH CHANNELS

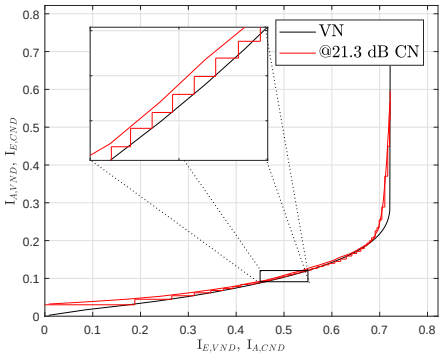
We now look at the behavior of the EXIT chart analysis when considering fast fading Rayleigh channels. Note that the only modification that has to be introduced in this case is in step 2 of the SPA algorithm (see Section 2.6.4). Specifically, since we are assuming perfect CSI at the receiver, the fading factor that multiplies the coded RCM-LDGM symbols (i.e., realizations of i.i.d. exponential random variables of Section 2.4.2) has to be provided to the decoder.

As in the previous AWGN case, we focus on the EXIT chart design for codes of spectral efficiency $\rho = 7.4$ bits per complex dimension and with sources having a priori probabilities $p_0 = 0.5$ $p_0 = 0.8$, and $p_0 = 0.95$. The corresponding SNR Shannon limits are 24.7 dB, 18.3 dB, and 6.8 dB, respectively. Figure 3.10 is similar to Fig. 3.8, except that we now consider fast fading. It plots the EXIT charts and real trajectories of the good codes, specified in Table 3.3, which have been selected by our EXIT chart analysis. The EXIT chart channels are open at SNRs close to the Shannon limits. This is shown in Fig. 3.11, where the BER predictions and the actual Monte Carlo simulations are presented for different values of SNR. Note that the gaps to the Shannon limits are within 3 dBs, and that the BER predictions

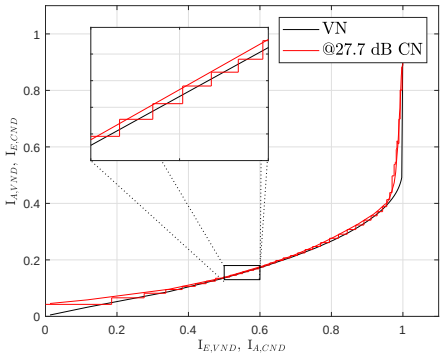
are very close to the results obtained by simulations, corroborating that the proposed EXIT chart analysis is also well suited for fast fading Rayleigh channels.



(a) $p_0 = 0.95$



(b) $p_0 = 0.80$



(c) $p_0 = 0.5$

Figure 3.10: *EXIT chart and real trajectories of the designed codes for fast Rayleigh channel and sources with* **c** $p_0 = 0.5$, **b** $p_0 = 0.8$ *and* **a** $p_0 = 0.95$.

Table 3.3: *Best configurations obtained by the EXIT chart analysis for fast fading Rayleigh channels.*

p_0	K	M	I	d_l	$\mathcal{W}$	Decoding Threshold.
0.5	37000	9600	300	3	$\{2,3,4,4,8\}$	27.7 dB
0.8	37000	9240	760	5	$\{2,2,3,3,4,8\}$	21.3 dB
0.95	37000	9480	520	4	$\{1,1,1,1,1,1,1,1,1,1,1\}$	9.55 dB

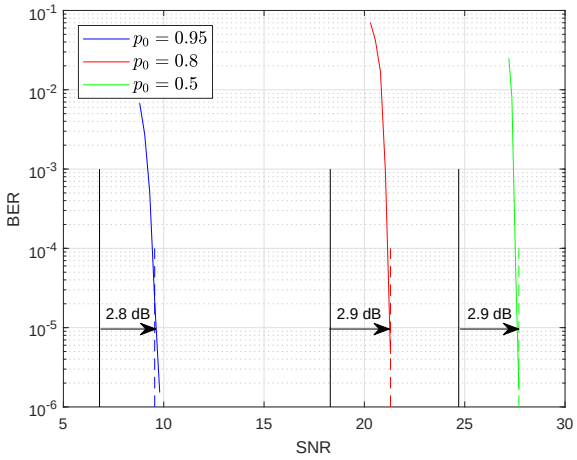


Figure 3.11: *BER vs SNR obtained by Monte Carlo simulations for fast fading Rayleigh channels. The Shannon limits (continuous lines) and the predicted decoding threshold (dashed lines) are plotted as vertical lines.*

3.4 DISCUSSION

We have shown that RCM-LDGM codes are very well suited for implementing high rate joint source-channel coding schemes for transmitting uniform and non-uniform binary memoryless sources in both AWGN and Rayleigh fast fading channels. Good performance is obtained if the codes are designed according to the source's non-uniformity. The search of good design parameters was previously perform by brute force approach, using time-consuming Monte Carlo simulations. In this chapter we have devel-

oped an EXIT chart analysis and a design methodology suitable to speed up this procedure by up to a factor of $\approx 10^3$.

The effectiveness of our proposed EXIT chart and the BER prediction capabilities have been verified through BER simulations, in which the estimated decoding thresholds lie very close to the real waterfall regions and the BER prediction is very accurate for those cases in which both curves intersect.

CHAPTER 4

Sources with Memory and the BWT

This chapter looks at the problem of implementing high-throughput JSCC schemes for the transmission of binary sources with memory, modeled by Markov chain (MC) or a hidden Markov model (HMM) (defined as in Section 2.3.1), and transmitted over AWGN channels (refer to Section 2.4.1).

4.1 INTRODUCTION

When considering sources with memory, Shannon's JSC coding theorem states that reliable transmission is only possible if

$$\mathcal{H}(S) \cdot R \leq \mathcal{C}, \tag{4.1}$$

where $\mathcal{H}$ is the entropy rate of the source in bits per source symbol, defined as

$$\mathcal{H}(S) \triangleq \lim_{k \to \infty} \frac{H(U_1, \ldots, U_k)}{k}$$

and H is the entropy, $\mathcal{C}$ the capacity of the channel in information bits per channel use and R is the JSC code's rate (source symbols per complex channel use).

The proposed coding scheme is based on a combination of the Burrows-Wheeler Transform (BWT) [83] and the rate-compatible parallel RCM-

LDGM codes. The BWT is a block sorting algorithm that has been widely analyzed in [84, 85, 86] and employed for the general problem of data compression [87, 88]. As explained in the introduction of Part I, the applicability of the BWT to coded transmission of sources modeled by Markov chain and transmitted through AWGN channels has been investigated in the literature. Authors in [89] combine the BWT and segmentation of piecewise stationary memoryless sources with recently proposed methods of discrete denoising, whereas [90] utilize the context based redundancy left in channel coded data to improve channel decoding of non-systematic turbo codes.

Later, the work in [31] proposed a source-controlled Turbo coding scheme in which different energy allocation techniques are employed for the coded symbols based on the first order probability distribution of the binary symbols at the output of the BWT. Our scheme is also source-controlled, but different from [31] in the sense that it uses this first order probability distribution to optimize the rates at which the different segments of the BWT output sequences are transmitted. This optimal rate allocation is facilitated by the use of RCM-LDGM codes, which, additionally, provide high-throughput transmission capabilities.

4.2 BURROWS-WHEELER TRANSFORM (BWT)

The BWT is a lexicographical permutation of the characters of a string such that the transformed sequence is easier to be compressed. As explained in detail in the Appendix A, the BWT is obtained from the last column of an array whose rows are all cyclic shifts from the input in dictionary order, which tend to have long runs of identical characters. The BWT, unlike other ordered based transformations, is reversible, i.e., the input string (and all intermediated cyclic shifts) can be recovered from the output string.

Let $\boldsymbol{T} = \{T_k\}_{k=1}^{K}$, $T_k \in \{0, 1\}$, denote the output block of the reversible block-sorting BWT when its input is the block of binary source symbols $\{U_k\}_{k=1}^{K}$, $U_k \in \{0, 1\}$. For sources modeled by MCs with M_λ states, the joint probability mass function, $P_{\mathbf{T}}(\boldsymbol{t})$, of the random block $\boldsymbol{T}$ is approximately memoryless and piecewise stationary [84], in the sense that there exist M_λ index sets, $\mathcal{L}_i = \{w_{i-1} \dots w_i\}$, $i = 1, \dots, M_\lambda$ with $w_0 = 1$ and

$w_{M_\lambda} = K + 1$, and a probability distribution

$$Q_{\mathbf{T}}(\mathbf{t}) = \prod_{i=1}^{M_\lambda} \prod_{k=w_{i-1}}^{w_i - 1} Q_i(t_k) \tag{4.2}$$

such that the normalized divergence between both distributions can be made arbitrarily small for sufficiently large K, i.e.,

$$\frac{1}{K} D(Q_{\mathbf{T}}(\mathbf{t}) \| P_{\mathbf{T}}(\mathbf{t})) \to 0 \tag{4.3}$$

as $K \to \infty$.

As the block length K goes to infinity, the normalized length of the index set in expression (4.2) converges to $c_i \in \mathbb{R}$, i.e., $\lim_{K\to\infty} \frac{|\mathcal{L}_i|}{K} = c_i$.

Definition 2. *Let $\boldsymbol{T}_i$ denote the binary random sequence of length $K_i = c_i K$ at the output of the BWT corresponding to the index set $\mathcal{L}_i$, $i = 1, \ldots, M_\lambda$. That is, $\boldsymbol{T}_i = \{T_k\}_{k \in \mathcal{L}_i}$.*

Observe from (4.2) that for large blocks of length K, the binary random symbols $T_k \in \boldsymbol{T}_i$, with $k \in \mathcal{L}_i$, can be considered i.i.d. with probability distribution

$$Q_i(t_k) \triangleq \left\{ \begin{array}{cc} p_i^0 & \text{if } t_k = 0 \\ p_i^1 = 1 - p_i^0 & \text{if } t_k = 1 \end{array} \right\} \tag{4.4}$$

for some $p_i^0 \in (0, 1)$. Note that these approximations should be understood under the convergence criterium (4.3). Consequently, we model the non-stationary BWT output sequence $\boldsymbol{T}$ as the concatenation of M_λ blocks of length $K_i = c_i K$, $i = 1, \ldots, M_\lambda$ generated by M_λ independent DMS binary sources $\mathcal{S}_1, \mathcal{S}_2 \ldots, \mathcal{S}_{M_\lambda}$, with entropies

$$H_i = -p_i^0 \log p_i^0 - (1 - p_i^0) \log(1 - p_i^0), \;\; i = 1, 2, \ldots, M_\lambda.$$

By the independence of the sources and their symbols, the entropy rate of the original source can be expressed as

$$\mathcal{H}(\mathcal{S}) = \sum_{i=1}^{M_\lambda} \frac{K_i}{K} H_i = \sum_{i=1}^{M_\lambda} c_i H_i. \tag{4.5}$$

4.3 PROPOSED BTW-JSC SCHEME

The proposed BWT-JSC scheme starts by transforming the original source with memory S, into a set of M_λ independent non-uniform memoryless binary sources. This is accomplished by partitioning the source sequence into blocks of length K, $\mathbf{U}^{(l)} = \{U_{l\cdot K+k}\}_{k=1}^{K}$, $l \in \mathbb{N}$, and then applying the BWT to each of these blocks. The corresponding output segment i, inside output block l, is given by

$$\boldsymbol{T}_i^{(l)} = \{T_{l\cdot K+k}\}_{k=w_{i-1}}^{w_i}.$$

Recall from (4.4) that the sequence blocks $\boldsymbol{T}_i^{(l)}$, $i = 0, 1 \ldots, M_\lambda$ can be considered as generated by a non-uniform DMS with entropy H_i, $i = 1, 2, \ldots, M_\lambda$. Thus, we reduce the encoding problem of sources with memory into a simpler one, namely the problem of JSC coding of non-uniform memoryless binary sources, with entropies H_i. Notice that RCM-LDGM codes can now be applied for each of the M_λ independent sources as shown in Fig. 4.1.

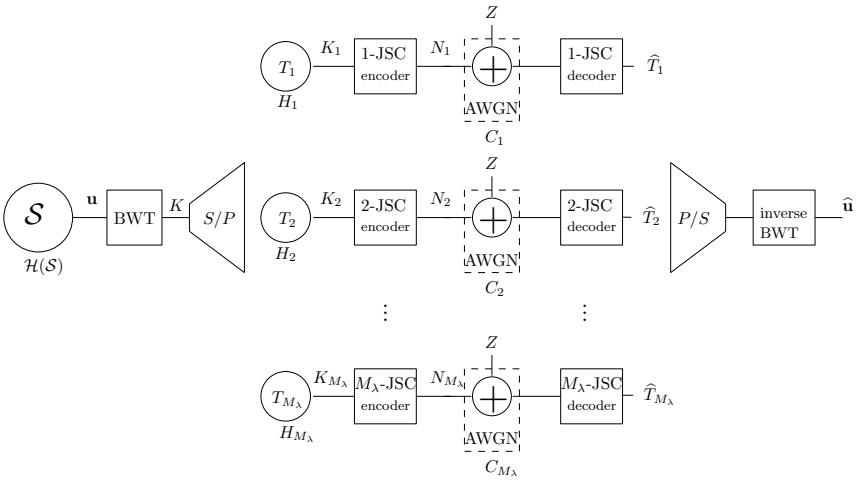


Figure 4.1: *BWT-based proposed communication system.*

More concretely, let us consider a source with memory, $\mathcal{S}$, and with entropy rate $\mathcal{H}(\mathcal{S})$, which generates blocks of K binary symbols to be trans-

mitted at rate $R = K/N$ by the parallel JSC coding system of Fig. 4.1. Let $\mathbf{T}_i$ (refer to Definition 2) be the input sequence to the corresponding i-JSC code of rate $R_i = K_i/N_i$, under the constraint $N = \sum_{i=1}^{M_\lambda} N_i$. Denote by $\{\mathrm{SNR}_i\}_{i=1}^{M_\lambda}$ the set of signal-to-noise ratios allocated to each parallel channel. Define by

$$\overline{\mathrm{SNR}} = \sum_{i=1}^{M_\lambda} \frac{N_i}{N} \mathrm{SNR}_i$$

the average SNR over all parallel channels. The following Theorem proves that the proposed scheme achieves the Shannon limit.

Theorem 1. *Given a target rate R, the minimum overall $\overline{SNR}$ in the coding scheme of Fig. 4.1 is achieved when all the SNR_i's take the same value, given by the SNR Shannon limit from expression (4.1), i.e., $SNR_i^* = 2^{R\mathcal{H}(\mathcal{S})} - 1$. The individual rates R_i are given by $R_i = \frac{R\mathcal{H}(\mathcal{S})}{H_i}$, $i = 1, \ldots, M_\lambda$.*

Proof. Given a set of signal-to-noise ratios $\{\mathrm{SNR}_i\}_{i=1}^{M_\lambda}$, the rates of the JSC encoders in Fig. 4.1 are given by the Shannon's separation theorem as

$$R_i = \frac{K_i}{N_i} = \frac{C(\mathrm{SNR}_i)}{H_i},\ i = 1, \ldots, M_\lambda,$$

where by the BWT hypothesis, $K = \sum_{i=1}^{M_\lambda} K_i$.

We seek to minimize the average signal-to-noise ratio $\overline{\mathrm{SNR}}$ over all the M_λ parallel AWGN channels, i.e.,

$$\overline{\mathrm{SNR}} = \sum_{i=1}^{M_\lambda} \frac{N_i}{\sum_{j=1}^{M_\lambda} N_j} \mathrm{SNR}_i = \sum_{i=1}^{M_\lambda} \frac{K_i}{N} \frac{H_i}{C(\mathrm{SNR}_i)} \mathrm{SNR}_i,$$

under the constraint of achieving a rate

$$R = \frac{K}{N} = \sum_{i=1}^{M_\lambda} \frac{K_i}{\sum_{j=1}^{M_\lambda} N_j}.$$

Please note that since $K = \sum_{i=1}^{M_\lambda} K_i$ is fixed, the constraint in R reduces to the constraint

$$N = \sum_{j=1}^{M_\lambda} N_j = \sum_{j=1}^{M_\lambda} \frac{H_j K_j}{C(\mathrm{SNR}_j)}. \tag{4.6}$$

By applying the Lagrange multipliers method, we define F as

$$F = \sum_{i=1}^{M_\lambda} \frac{K_i H_i}{N} \frac{\mathrm{SNR}_i}{C\left(\mathrm{SNR}_i\right)} + \gamma \left(\sum_{i=1}^{M_\lambda} \frac{K_i H_i}{C\left(\mathrm{SNR}_i\right)} - N \right),$$

and by searching for an extreme of F, we obtain that the optimal SNR_i^* are all equal to some value Γ. Therefore, from constraint (4.6)

$$N = \sum_{i=1}^{M_\lambda} N_i = \sum_i \frac{K_i H_i}{C\left(\Gamma\right)} = \frac{K\mathcal{H}\left(\mathcal{S}\right)}{C\left(\Gamma\right)},$$

where the last equality follows from expression (4.5). Thus, the rate can be written as

$$R = \frac{K}{N} = \frac{C(\Gamma)}{\mathcal{H}\left(\mathcal{S}\right)}.$$

Consequently, the value of Γ is given by the signal-to-noise ratio required to achieve the same rate R in the standard point-to-point communications system. That is,

$$\Gamma = 2^{RH(S)} - 1.$$

We conclude that

$$\overline{\mathrm{SNR}}^* = \mathrm{SNR}_i^* = 2^{R\mathcal{H}(\mathcal{S})} - 1$$

and

$$R_i = \frac{K_i}{N_i} = \frac{C(\overline{\mathrm{SNR}}^*)}{H_i} = \frac{R\mathcal{H}\left(\mathcal{S}\right)}{H_i}. \tag{4.7}$$

□

Remark 1. *Observe that the BWT-JSC is asymptotically optimal in the sense that can achieve the SNR Shannon limit given by the Separation Theorem.*

4.4 CONVENTIONAL NON-BWT APPROACH

In order to compare the performance of the proposed BWT-JSC strategy, we explain in this section the conventional coding scheme that results from using a standard encoder block and modifying the decoder to exploit

the source's statistics. This exploitation is achieved by attaching the factor graph that models the MC source, shown in Fig. 2.7, to the factor graph of the coding scheme (refer to Fig. 2.15). The resulting FG is depicted in Fig. 4.2.

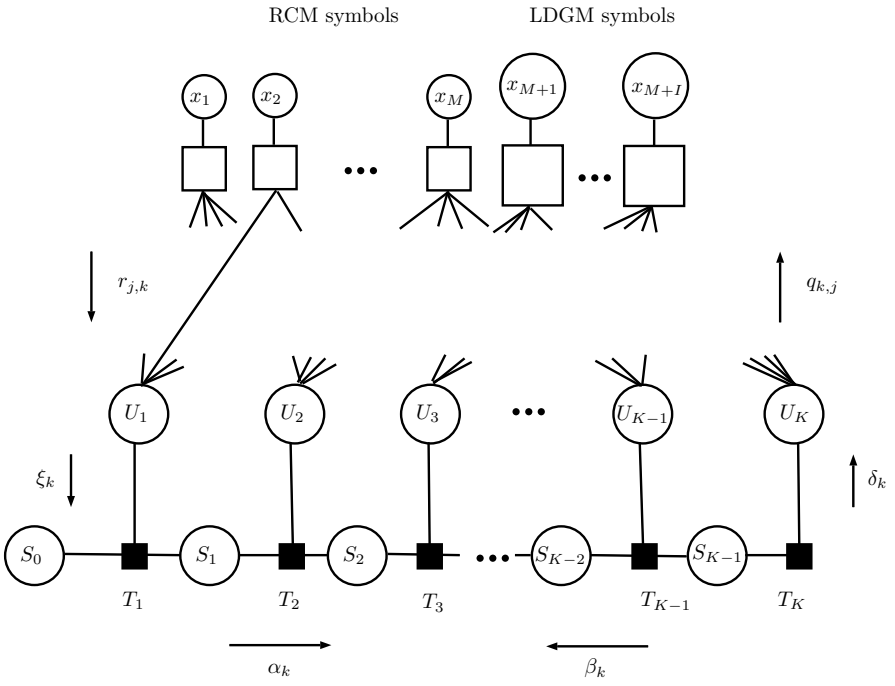


Figure 4.2: *Factor graph of the parallel RCM-LDGM code with the source's factor graph attached.*

Finally, the symbol-wise decoding is performed by applying the SPA over the entire factor graph, where messages ξ_k, α_k, β_k and δ_k are calculated as explained in Section 2.3.1 and $r_{j,k}$ and $q_{k,j}$ as in Section 2.6.4 and interchanged obeying the rules detailed in Section 2.2.2.

4.5 RESULTS

In this section, we evaluate the proposed BWT-JSC scheme, comparing its performance with the conventional approach described in Section 4.4,

that we will denote as NON-BWT-JSC. Without any loss of generality, the spectral efficiency of the communication system has been set to 7.4 binary source symbols per complex channel use, and the source block length to $K = 37000$. Thus, the total number of coded symbols at the output of the JSC encoder is $N = 10000$. We begin by specifying the Markov sources considered in this chapter.

4.5.1 MARKOV SOURCES AND THEIR OUTPUT PROBABILITY PROFILE

We consider three 2-state ($M_\lambda = 2$) Markov sources, the first two modeled by MCs, with entropy rates 0.57 and 0.80 bits per source symbol, and the third modeled by a HMM with entropy rate 0.73. For the sake of notation, they will be referred as S_1, S_2 and S_3. Table 4.1 summarizes their corresponding Markov parameters.

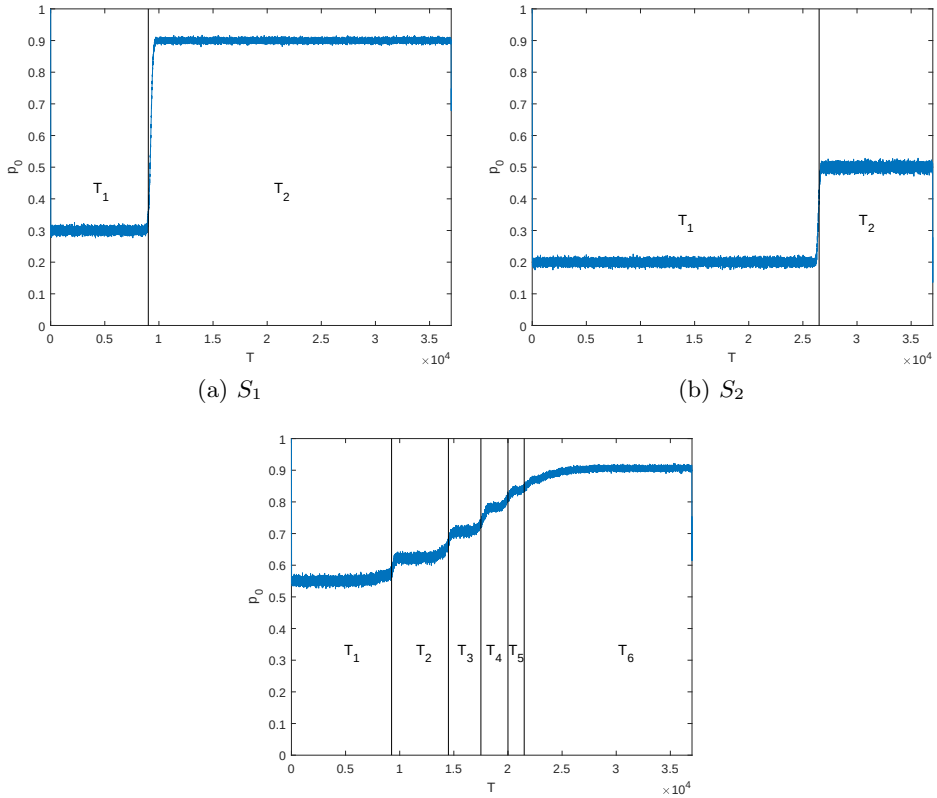
Table 4.1: *Markov Source Parameters.*

Source	Matrix A	Matrix B	Vector π	Entropy $\mathcal{H}$
S_1	$a_{11} = 0.90$ $a_{22} = 0.70$	$b_{11} = 1.0$ $b_{22} = 1.0$	[0.75 0.25]	0.57
S_2	$a_{11} = 0.80$ $a_{22} = 0.50$	$b_{11} = 1.0$ $b_{22} = 1.0$	[0.71 0.29]	0.8
S_3	$a_{11} = 0.90$ $a_{22} = 0.90$	$b_{11} = 0.5$ $b_{22} = 0.995$	[0.5 0.5]	0.73

Figure 4.3 shows the probability mass function $P_{\mathbf{T}}(\mathrm{t})$ (refer to (4.4)) of the binary random block $\boldsymbol{T}$ of length $K = 37000$ at the output of the BWT for sources S_1, S_2, and S_3. Observe that due to the fact that sources S_1 and S_2 follow a 2-state MC behavior, the BWT will produce approximately two i.i.d. segment $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$. This is clearly shown in Figs. 4.3a and 4.3a, with segments of length ($K_1 = 9020$, $K_2 = 27980$) with first order probabilities ($p_0^{(1)} \approx 0.3$, $p_0^{(2)} \approx 0.9$) for S_1 and ($K_1 = 26500$, $K_2 = 1050$) with probabilities ($p_0^{(1)} \approx 0.2$, $p_0^{(2)} \approx 0.5$) for S_2. On the contrary, the source S_3 is characterized by a 2-state HMM, and the hidden property has the effect of increasing the number of apparent states, should the HMM source be approximated by a pure MC. This is observed in Fig. 4.3c, where a 6-state MC source would fairly approximate the statistics of source S_3. In this case, the first order probabilities of segments $\boldsymbol{T}_1 - \boldsymbol{T}_6$ of sizes ($K_1 = 9250$, $K_2 = 5250$, $K_3 = 3000$, $K_4 = 2500$, $K_5 = 1500$, $K_6 = 15500$)

are given by $p_0^{(1)} \approx 0.55$, $p_0^{(2)} \approx 0.63$, $p_0^{(3)} \approx 0.71$, $p_0^{(4)} \approx 0.78$, $p_0^{(5)} \approx 0.84$ and $p_0^{(6)} \approx 0.9$.

Note that all partitions, specially the last one, are decided manually. The property in (4.3) is asymptotic and we consider finite size blocks. Moreover, in some cases the first order probabilities of some consecutive segments (or segment candidates) are very close and it is not worthy to divide further. For example, the segment T_6 of S_3 could possibly be further segmented, but it would lead to negligible benefit or even be counterproductive, since these small segments would have to be encoded.



(a) S_1 (b) S_2

(c) S_3

Figure 4.3: *First order probability profiles of the output blocks of the BWT for example sources* (**a**) S_1, (**b**) S_2 *and* (**c**) S_3.

4.5.2 NUMERICAL RESULTS

In this section, we present the results obtained by Monte Carlo simulation for the proposed BWT-JSC and the conventional NON-BWT-JSC coding schemes. Observe that due to the BWT block, in our proposed scheme a single error at the output of the decoders will be propagated after applying the inverse-BWT. Therefore, to make a fair comparison, the results are presented in the form of Packet Error Rate (PER), i.e., $\mathcal{P}\{\hat{\mathbf{u}} \neq \mathbf{u}\}$. Notice that in the NON-BWT strategy, due to the way in which the the information is forward and backward propagated in the source's factor graph, a bit in error is likely to produce errors in the preceding/following bits, leading to error bursts. Thus, the PER performance comparisons seems reasonable.

It should be mentioned that for the correct recovery of the original transmitted source block, the inverse-BWT at the receiver side needs to know the integer J with the exact position where the original End of File symbol has been moved by the BWT at the transmitted side (refer to Appendix A). Therefore, this additional information should also be transmitted. Notice that for a 37000 block length, this position can be addressed by adding 16 binary symbols. In this work, we have considered this rate loss as negligible, but in real scenarios it must be taken into account.

Figure 4.4 shows the PER vs SNR curves obtained by Monte Carlo simulations for the example sources (**a**) S_1, (**b**) S_2 and (**c**) S_3 when using both the proposed system (BWT-JSC) and the conventional approach (NON-BWT-JSC) as a reference. In the proposed scheme, as stated in Section 4.3, after performing the BWT, each of the resulting M_λ independent non-uniform i.i.d. segments $\boldsymbol{T}_i(p_0^{(i)})$ $i = 1, \dots M_\lambda$ (refer to Fig. 4.3), are encoded by M_λ separated RCM-LDGM JSC codes of rates R_i as given by Theorem 1. The codes used for each DMS in the BWT-JSC approach, as well as the one used in the conventional NON-BWT-JSC scheme are summarized in Table 4.2, which have been designed utilizing the EXIT charts of Chapter 3 and brute force approach, respectively.

Observe from Figs. 4.4a and 4.4b that for sources S_1 and S_2, represented by a MC, our BWT-JSC scheme outperforms the NON-BWT-JSC approach by about 4.2 and 2.3 dB's, respectively. One reason behind this large improvement lies in the fact that in the NON-BWT-JSC system, the FG of the decoder, results from a parallel concatenation of two sub-graphs:

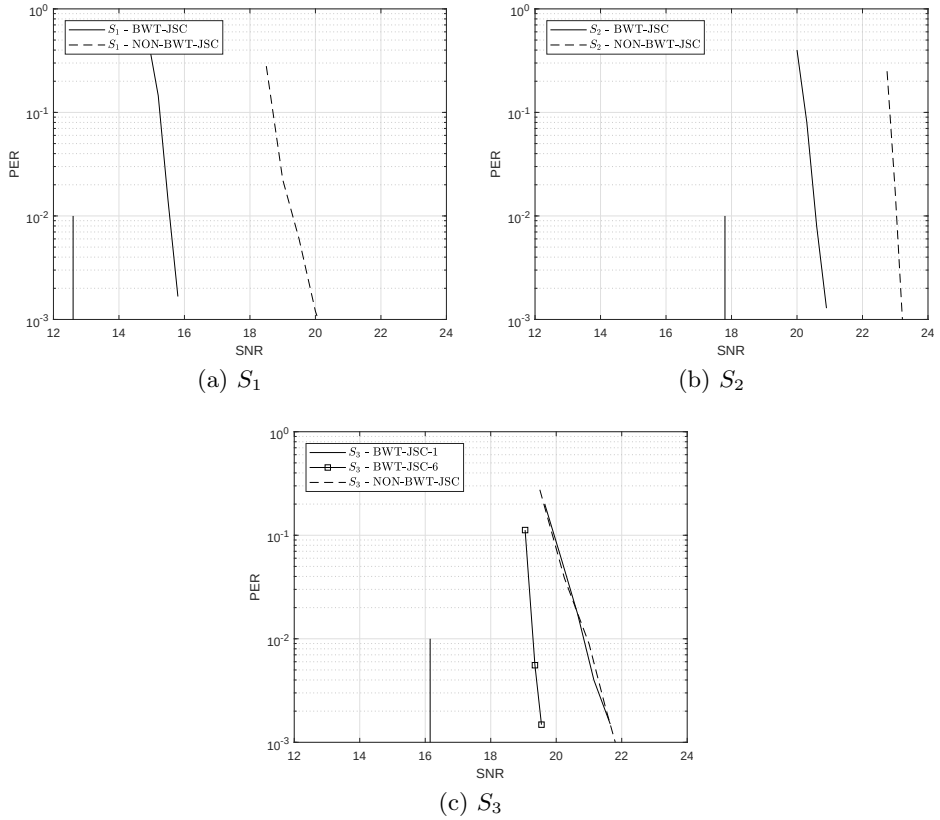


Figure 4.4: *Obtained PER vs SNR curves for the NON-JSC-BWT and JSC-BWT schemes when sources (**a**) S_1; (**b**) S_2 and (**c**) S_3 are considered. The corresponding Shannon limits are plotted in vertical lines.*

The RCM-LDGM code and MC source sub-graphs (refer to Figure 4.2). Consequently, in the overall FG decoder a large amount of cycles between both sub-graphs appear, degrading in this way the performance of sum-product algorithm, even if a large number of iterations is used. However, in the proposed scheme, these cycles do not occur since in this case the sources are memoryless and non-uniform. The contribution of the sources sub-graphs is just to introduce the a priori probabilities of the non-uniform

sources into the variable nodes of the corresponding RCM-LDGM factor sub-graphs.

Let us now consider the HMM source $\mathcal{S}_3$ with entropy rate $\mathcal{H}(\mathcal{S}_3) = 0.73$ and output probability profile as shown in Figure 4.3c. As mentioned before, the BWT output of source S_3 can be approximated by 6 memoryless non-uniform sources $\{\boldsymbol{T}_i\}_{i=1}^{6}$, with blocks of lengths $K_1 \approx 9250, K_2 \approx 5250, K_3 \approx 3000, K_4 \approx 2500, K_5 \approx 1500, K_6 \approx 15000$. Some of these blocks have short lengths, which is detrimental for the performance of the corresponding RCM-LDGM codes. To solve this problem, we build larger segments $\widetilde{\boldsymbol{T}}_i$ that keep the same statistical properties as previous segments. In this approach, named BWT-JSC-κ, we put together κ consecutive output blocks of the BWT to form the new segments as $\widetilde{\boldsymbol{T}}_i^{(l)} = \{\boldsymbol{T}_i^{(l\cdot\kappa)}, \ldots, \boldsymbol{T}_i^{(l\cdot(\kappa+1)-1)}\}$ for $i = 0, 1, \ldots, M_\lambda$ and $l \in \mathbb{N}$. This is, in fact, similar to applying the BWT to source blocks of length $\kappa \cdot K$, but we have opted to joint consecutive output segments due to the quadratic computational complexity of the implemented BWT. The RCM-LDGM codes used to the transmit these segments have the same rate as before, but in this case their input and output block lengths are scaled by κ, i.e., $\widetilde{K}_i = \kappa \cdot K_i$, $\widetilde{M}_i = \kappa \cdot M_i$ and $\widetilde{I}_i = \kappa \cdot I_i$, $i = 0, 1, \ldots, M_\lambda$, respectively.

As before, Figure 4.4c plots the PER versus SNR curves for both strategies BWT-JSC (solid curves) and NON-BWT-JSC (dashed curves). When plotting the performance of the BWT-JSC-κ approach, two different cases have been considered, $\kappa = 1$ and $\kappa = 6$. Please note that when $\kappa = 1$ the scheme is the same as in previous MC examples. On the other hand, by concatenating 6 consecutive BWT output segments ($\kappa = 6$), we force the length of smallest segment to be 9000. Notice that for $\kappa = 6$ the proposed scheme outperforms the conventional approach by 2.3 dB. However, due to the bad performance of the short block-length RCM-LDGM codes, when $\kappa = 1$ the performance is similar to that of the conventional approach. This clearly shows that by concatenating BWT segments the system performance improves thanks to the avoidance of blocks with short lengths.

As summarized in Table 4.3, the proposed scheme clearly outperforms the conventional approach, and the PER vs SNR curves are only about 3 dB away from the Shannon limits.

Table 4.2: *Design parameters (refer to Section 4.4) used for sources (**a**) S_1, (**b**) S_2 and (**c**) S_3.*

		$\mathcal{W}$	M	I	d_l
BWT-JSC	$K_1 = 9020$	{1, 1, 1, 2, 2}	3705	55	5
	$K_2 = 27980$	{1, 1, 1, 1, 2, 2, 2, 2}	6140	100	5
NON-BWT-JSC	$K = 37000$	{2, 2, 3, 3, 4, 7}	9860	140	5
		S_1			
		$\mathcal{W}$	M	I	d_l
BWT-JSC	$K_1 = 26500$	{2, 2, 3, 3, 4, 8}	6310	110	5
	$K_2 = 10500$	{2, 3, 4, 7}	3490	90	5
NON-BWT-JSC	$K = 37000$	{2, 3, 4, 4, 7}	9940	60	3
		S_2			
		$\mathcal{W}$	M	I	d_l
BWT-JSC-κ	$K_1 = 9250$	{2, 3, 4, 4, 7}	3376	29	7
	$K_2 = 5250$	{2, 3, 4, 4, 7}	1839	19	7
	$K_3 = 3000$	{2, 3, 4, 4, 7}	935	37	7
	$K_4 = 2500$	{2, 3, 4, 4, 7}	673	35	6
	$K_5 = 1500$	{2, 2, 3, 3, 4, 7}	351	18	6
	$K_6 = 15500$	{2, 2, 2, 3, 3, 4, 4, 7, 7}	2632	56	6
NON-BWT-JSC	$K = 37000$	{2, 2, 3, 3, 4, 8}	9880	120	3
		S_3			

Table 4.3: *Summary of numerical results. Labels BWT-JSC and NON-BWT-JSC represent the SNR required for a PER of 10^{-3} with each scheme.*

	Entropy Rate	**Shannon Limit**	**BWT-JSC(-κ)**		**NON-BWT-JSC**
S_1	0.57	12.57 dB	15.8 dB		20 dB
S_2	0.80	17.78 dB	20.9 dB		23.25 dB
			1	6	
S_3	0.73	16.15 dB	21.8 dB	19.55 dB	21.8 dB

4.6 DISCUSSION

The BWT is able to transform the source redundancy caused by memory into an equivalent redundancy produced by a non-uniform and non-stationary first order distribution. This fact, paired with sequence segmentation, promotes the idea of employing coding techniques for i.i.d. discrete memoryless sources based on the first order stationary probability distribution of the segmented output of the BWT.

Based on the previous fact, we have proposed a source-controlled coding scheme for the transmission of binary sources with memory over AWGN channels that makes use of the BWT and sequence segmentation technique. The output set of independent non-uniform discrete memoryless binary sources are separately encoded at optimal coding rates based on their source entropy. We design specific RCM-LDGM codes for these non-uniform sources at the optimal rates using the EXIT charts of Chapter III.

CHAPTER 5

Impulsive Noise Channels

In this chapter we address the problem of sending high data rates over additive impulsive noise channels, for which we propose the use of RCM-LDGM codes for the first time in the literature.

5.1 INTRODUCTION

Power-line networks present an interesting alternative for no-new-wires scenarios where there is an electrical power distribution wiring available. These networks were not designed for data communication. Therefore, their electrical characteristics and nature make them a harsh environment for this purpose [91, 92]. The impulsive nature of the noise has a time-varying behavior that produces noise bursts with a duration on a scale from microseconds to milliseconds [93], with significant implications on data transmission, causing problems for the realization of power-line communications with high rate and high reliability. Consequently, unlike other communication channels, the power-line channel cannot be modeled as an AWGN channel [94, 95, 96, 97].

Authors in [98, 99] introduced the Middletons Class A noise model as a statistical model for such impulsive noise. Middleton Class A noise model assumes independent impulse emissions that are Poisson-distributed in time and it is one of the most known impulsive noise models due to its simplicity.

However, the Middleton Class A noise model is not appropriate for modeling practical impulsive noise, which often occurs in bursts. Later, the work in [100] proposed the Markov-Middleton model, which takes into account the noise memory and allows modeling the bursty behavior of impulsive noise.

The statistical feature of these impulsive channel models is much different from that of Gaussian noise. Therefore, the conventional receivers exhibit a degraded performance, since they have been optimized for AWGN channels. Spaulding and Middleton proposed in [101, 102] receivers for coherent and incoherent detectors in additive white class A noise (AWAN) channels. Later, the authors in [103] proposed a turbo decoding suitable for these channels, in which they modified the channel LLR value to suit the class A noise environment. The rest of the original turbo decoder does not need to change. The calculation of these channel values for Class A noise is too difficult, since it involves infinite series of non-linear functions. However, authors in [103] proposed an approximation using Jacobian logarithm [104]. The same strategy is applied in [105] for LDPC codes. In order to improve the spectral efficiency and error performance, authors in [106] study the application of BICM with iterative decoding in the Class-A impulsive noise environment. Since the maximum likelihood decoding of BICM is too complicated to implement in practice, they use a suboptimal iterative method with soft-output demodulator and the single-input single-output decoder. They also show in their following work [107] that BICM-ID is more effective than the Orthogonal Frequency Division Multiplexing (OFDM) technique in combating the impulsive noise. During this chapter, we consider that the channel state information is available at the receiver,and model the impulsive noise wit modified version of the Middleton Class A noise and Markov-Middleton model.

5.2 IMPULSIVE NOISE CHANNEL MODELS

In this chapter we consider the communication scenario of Section 2.1, with noise sample vector $\mathbf{z} = (z_1, \ldots, z_N)^\top \in \mathbb{R}^{N\times 1}$, whose entries are drawn independently but not identically distributed. The entry z_j is modeled as a zero mean Gaussian random variable with variance σ_j^2 for $j \in \{1, \ldots, N\}$. We consider the following two modes of the impulsive noise:

- Modified version of the Middleton Class A noise. The noise sample z_j, for $j \in \{1, \ldots, N\}$, is drawn from a zero-mean Gaussian distribution of variance σ_j^2, where $\sigma_j^2 = \sigma^2$ if there is no impulse at sample j, and $\sigma_j^2 = \nu^2 >> \sigma^2$ if there is an impulse at sample j. The arrival of impulses is modeled as a Poisson-process. Observe that in this case, we refer to the variance of the AWGN noise as σ^2, without impulse. For more details about the Middleton Class A noise model please refer to [98, 99, 108].

- Markov-Middleton model. The noise samples z_j, for $j \in \{1, \ldots, N\}$ are drawn independently but the probability density function of the j-th noise sample is ruled by a four state hidden Markov model. At a given state $s \in \{0, 1, 2, 3\}$, the corresponding noise sample is drawn from a zero-mean Gaussian distribution of variance ν_s^2, with $\nu_0^2 < \nu_1^2 < \nu_2^2 < \nu_3^2$. We consider that the state $s = 0$ corresponds to the absence of impulses and therefore we set $\nu_0^2 = \sigma^2$. For more details on this impulsive noise model please refer to [100].

Notice that both models of impulsive noise considered draw zero-mean Gaussian samples independently, with variances picked from a finite set of possible values depending on whether there is an impulse or not.

From the information theoretic point of view, the channel can be viewed as a set of parallel Gaussian channels, each with noise variance σ_j^2. Considering that there is no channel state information at the transmitter's side, all the channels are assigned the same average energy per symbol E_s and therefore, the capacity C of such a channel is given by (see, e.g., [109, Section 9.4])

$$C = \frac{1}{2N} \sum_{j=1}^{N} \log_2 \left(1 + \frac{E_s}{\sigma_j^2} \right)$$

bits per channel use. Let Υ_N be the random variable that counts the number of impulses per N channel uses, and assume that

$$\lim_{N \to \infty} \frac{E[\Upsilon_N]}{N} = 0. \tag{5.1}$$

It is easy to show that in such a case, for large enough N,

$$C \to \frac{1}{2} \log_2 \left(1 + \frac{E_s}{\sigma^2} \right),$$

that is, the capacity of the channel without impulsive noise. For the remainder of this chapter, let us define the SNR as $\frac{E_s}{\sigma^2}$, i.e., the ratio in the absence of impulsive noise. We will also keep condition (5.1), because this will let us assess the performance degradation of the RCM-LDGM codes due to impulsive noise.

5.3 ADAPTED DECODER

In this preliminary work on impulsive noise channels, we consider that the channel state information is known at the receiver, i.e., the decoder has knowledge about the values σ_j^2 for $j \in \{1, \ldots, N\}$. Taking this into account, the decoding is performed by applying the standard sum-product algorithm of Section 2.6.4 with the particularity that the operations at each CN j are adapted to the Gaussian noise component σ_j^2, for $j \in \{1, \ldots, N\}$.

5.4 NUMERICAL RESULTS

This chapter considers a high data rate scenario with spectral efficiency of 7.4 bits per complex channel symbol, using an RCM-LDGM encoder. In particular, we used an RCM code, of rate K/M and weight set $\mathcal{W} = \{2, 3, 4, 7\}$, in parallel with the non-systematic part (of length I) of a $d_l = 2$ degree LDGM. The parallel concatenation is constructed so that the LDGM coded symbols (I symbols) represent the 1.6% of the total coded sequence of length $N = M+I$. A total of 3 different block lengths are considered with $N = 2000, 10000$ and 30000. The source block length K is appropriately scaled to achieve the target spectral efficiency.

As it was mentioned in Section 5.2, under the assumption given in (5.1) and for a spectral efficiency of 7.4 bits per channel use, the SNR capacity limit for the considered channels is SNR$\geq$ 22.25 dB. Therefore, in order to show the suitability of RCM-LDGM codes over impulsive noise channels, we will compare the performance of these codes in an AWGN channel (without impulsive noise) and in a channel where the noise follows the considered impulsive noises models. For the latter case, we will consider that the decoder of Section 2.6.4 is adapted according to 5.3.

In Figs. 5.1 to 5.5, continuous lines represent the performance of these codes in an AWGN channel. On the other hand, dashed lines represent simulations with Middleton Class A noise (Figs. 5.1 and 5.2) or Markov-

Middleton model noise (Figs. 5.3 to 5.5), marked with circles when the optimal decoder is used and with crosses when the conventional AWGN decoder is used. Finally, the 3 different block lengths, $N = 30000$, $N = 10000$ and $N = 2000$, are differentiated by colors blue, red and black, respectively.

5.4.1 MODIFIED VERSION OF MIDDLETON CLASS A NOISE

We consider a channel with Middleton Class A noise with $\nu^2 = 500\sigma^2$. The arrival of impulses is modeled as a Poisson-process, and the expected number of impulses per N channel uses is set to $E[\Upsilon_N] = N^\alpha$, with $\alpha \in [0, 1)$. In our numerical simulations we have considered 2 cases with $\alpha = 1/2$ and $\alpha = 1/3$.

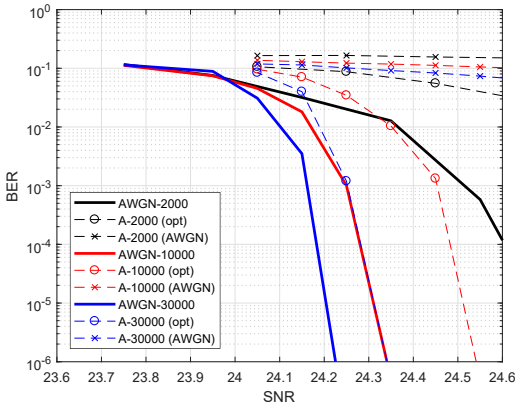


Figure 5.1: *BER vs SNR performance obtained by Monte Carlo simulations for the modified version of Middleton Class A noise with* $\alpha = 1/2$

Figure 5.1 shows the BER vs SNR performance for the case with $\alpha = 1/2$. As it can be seen, the larger N the better performance they achieve even in an AWGN channel. When the Middleton Class A noise is considered, it is observed that for a block length of 2000 symbols there is a big performance degradation when comparing to the AWGN scenario. This gap is closed as we increase the block length, and for $N = 30000$ it is observed that the performance degradation is only 0.1 dB. Similar results are

shown in Fig. 5.2 for $\alpha = 1/3$. In this case the gap is insignificant even for $N = 10000$.

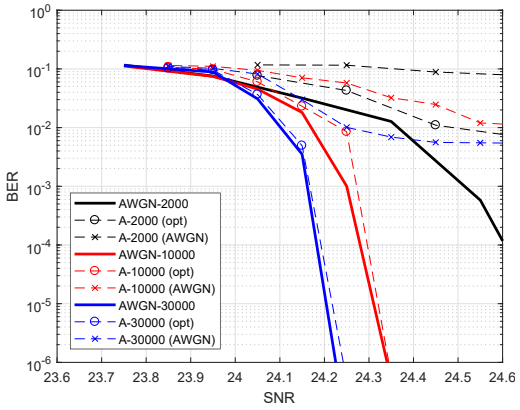


Figure 5.2: *BER vs SNR performance obtained by Monte Carlo simulations for the modified version of Middleton Class A noise with* $\alpha = 1/3$

5.4.2 MODIFIED VERSION OF THE MARKOV-MIDDLETON MODEL

In this subsection we consider that the impulsive noise follows a Markov-Middleton model. As it was introduced in Section 5.2, the noise samples are ruled by a four state HMM with transition matrix

$$\Pi = (1-\epsilon)\mathbf{1}_{4\times 1}\begin{pmatrix} p_0 & p_1 & p_2 & p_3 \end{pmatrix} + \epsilon I_4,$$

where $\mathbf{1}_{4\times 1}$ is the 4×1 matrix of ones, $\epsilon \in (0,1)$ determines the correlation between noise samples,

$$p_s = \frac{\frac{a^s}{s!}}{\sum_{s=0}^{3} \frac{a^s}{s!}}$$

is the probability of entering state s for $s \in \{0,1,2,3\}$ and

$$a = -a_0 \log\left(1 - N^{\frac{1-\beta}{\beta}}\right)$$

with $a_0 > 0$ and $\beta > 1$ being parameters that control the impulse rate. It is easy to show that as long as $\beta > 1$, (5.1) holds. Finally, the noise produced

at state s has variance

$$\nu_s^2 = \sigma^2 \left(1 + \frac{s}{a\gamma}\right) \qquad s \in \{0, 1, 2, 3\},$$

where $\gamma > 0$ controls the power of the impulses.

In our numerical results we have set $a_0 = 0.8$, $\gamma = 0.2$, $\epsilon = 0.9$, and we have considered three different values for β, namely, $\beta = 2.5$, $\beta = 2.1$ and $\beta = 1.7$.

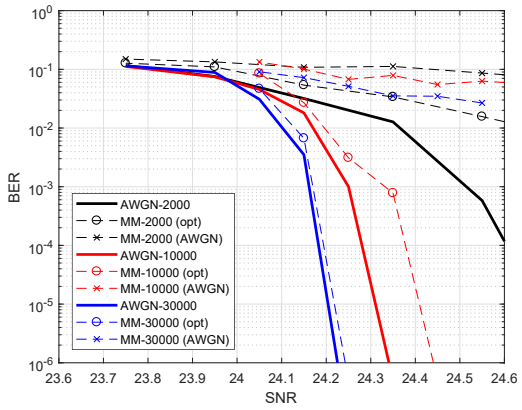


Figure 5.3: *BER vs SNR performance obtained by Monte Carlo simulations for the Markov-Middleton model with* $\beta = 2.5$

Figure 5.3 shows the BER vs SNR performance for the case with $\beta = 2.5$. The AWGN curves are the same as for the previous figures. When Markov-Middleton noise is considered, it is observed that for block length $N = 2000$ there is also a big performance degradation when comparing to the AWGN scenario. For $N = 10000$ the gap is reduced to 0.1 dB. Finally, there is almost no gap for $N = 30000$. The case with $\beta = 2.1$ is shown in Fig. 5.4. In this case, the gap for $N = 10000$ increases to about 0.3 dB at a BER of 10^{-4}.

Finally, we have considered the case with strong impulsive noise by setting $\beta = 1.7$. We have observed that in this case $N = 30000$ was not long enough for the proposed RCM-LDGM code. For that reason, we have slightly modified the RCM-LDGM code so that the LDGM symbols

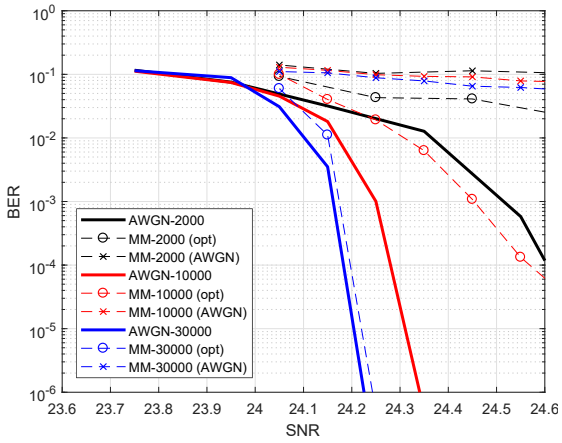


Figure 5.4: *BER vs SNR performance obtained by Monte Carlo simulations for the Markov-Middleton model with* $\beta = 2.1$

represent the 2.8% of N and we have increased the LDGM degree d_l to 3. Results are shown in Fig. 5.5. Notice that this new RCM-LDGM code has a worse BER vs SNR performance in the AWGN scenario, as it is not optimized for AWGN. However, with this stronger LDGM we are able to achieve a gap of 0.2 dB for $N = 30000$ (The gap is 0.4 dB if compared with the optimal RCM-LDGM for an AWGN channel). The shorter block lengths show a bigger gap in this case.

As a bottom line, it can be observed in both cases that the conventional RCM-LDGM decoder (without the adaption at the CNs), designed for an AWGN channel, leads to a huge performance degradation in the presence of impulsive noise.

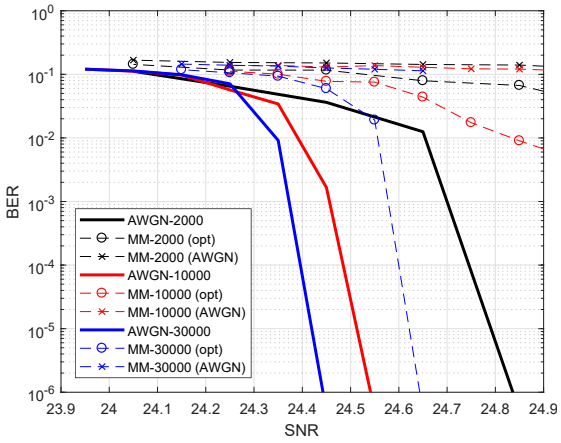


Figure 5.5: *BER vs SNR performance obtained by Monte Carlo simulations for the Markov-Middleton model with* $\beta = 1.7$

5.5 DISCUSSION

This chapter proposes the use of the parallel RCM-LDGM concatenation for harsh communication environments with impulsive noise channels, such as power-line communications. As for other existing codes used in these channels, the receiver has to adapt its decoder according to the channel statistics. Otherwise, the performance exhibits a big degradation.

In order to asses the performance of these codes, we have developed two modified versions of the well-known Middleton class A noise and Markov-Middleton noise models, in which the expected number of impulses per N channel uses tends to 0 as N increases. Under that condition, it is easy to show that the capacity of the impulsive noise channel tends asymptotically to the AWGN capacity. By utilizing this result, we have shown by Monte Carlo simulations that the parallel RCM-LDGM codes are able to perform within 2.5 dB of the Shannon limit, if the block lengths are large enough.

Part II

Multi-User Communications

In this second part of the dissertation we consider multi-user communication systems. Specifically, we extend the RCM and parallel RCM-LDGM coding schemes to the multiple access channel. Before delving into the proposed schemes and with the aim of providing some context, we begin by including a general overview of multiple access and joint source-channel coding techniques.

In the evolution of wireless communications, from the first generation to the actual 5G, the multiple access scheme has been one of the key technology to distinguish different generations. Frequency Division Multiple Access (FDMA) was used in 1G, Time Division Multiple Access (TDMA) in 2G, Code Division Multiple Access (CDMA) in the third generation and Orthogonal Frequency Division Multiple Access (OFDMA) in 4G. In these conventional multiple accesses schemes, different users are assigned to orthogonal resources in time, frequency or code domain to avoid or alleviate inter-user interference.

However, the ever increasing connectivity and throughput demands of current 5G and future 6G is making the spectral efficiency to become one of the key challenges of such explosive data traffic. The rapid development of the Internet of Things (IoT) is forcing wireless systems to handle massive connectivity users (or devices) demanding low latency. Moreover, some of these inexpensive devices require computationally light solutions. It is essential to improve the actual technology to meet these requirements. In this part of the dissertation we focus on the promising Non-Orthogonal Multiple Access (NOMA), that aims to solve some of these challenges for 5G and beyond. What differs NOMA from conventional Orthogonal Multi-Access (OMA) technologies is that it can can accommodate many more users via non-orthogonal resource allocation. Recent research pieces show that NOMA systems outperform OMA schemes in multiple scenarios, such as downlink systems [110], visible light communication systems [111] and short packet communications [112]. Therefore, NOMA is playing an important role in 5G [113, 114, 115] and will be one of the key technologies enabling the upcoming 6G networks [116].

NON-ORTHOGONAL MULTIPLE ACCESS (NOMA)

We focus on NOMA, which achieves massive connectivity and is able to increase system throughput. NOMA allows users are able to trans-

mit their information sharing frequency and time intervals. Most of these techniques fall into one of these two sub-categories [125], namely, power-domain and code-domain NOMA. This last utilizes user-specific spreading sparse sequences or non-orthogonal cross-correlation spreading sequences, and examples that fall into this sub-category are Low Density Diffusion (LDS) [117, 118], Sparse Code Multiple Access (SCMA)[119] and Multi-User Shared Access (MUSA)[120]. On the other hand, power-domain NOMA attains multiplexing by utilizing superposition coding, while Multi-User (MU) detection algorithms are used at the common receiver. The coding schemes considered in this Thesis fall into the power-domain category.

MULTI-USER DETECTION ALGORITHMS

NOMA allows controllable interference by non-orthogonal resource allocation, achieving larger spectral efficiency and low transmission latency. However, it increases the decoding complexity, since the receiver is responsible for detecting the desired signals, utilizing Multi-User (MU) detection algorithms.

We start by discussing the more dominant NOMA scheme, in which different users are allocated different power levels according to their channel conditions to obtain the maximum system throughput. Such power allocation is usually accompanied by the Successive Interference Cancellation (SIC) technique, which exploits differences in signal strength among the signals of interest in order to successively decode information arriving from different users. Starting from the strongest, each signal is decoded by treating the remaining as noise and then subtracting it from the received combined signal. Techniques that make use of SIC are well documented in the literature [121, 122], with recent research in this field focusing specifically on resource allocation. For example, authors in [123] optimize the power allocation in wireless powered communication networks to satisfy the minimum data requirements for each user. In [124] they also studied the optimal power allocation coefficients from a user fairness perspective. However, this widely utilized MU technique also presents some disadvantages. First, it generally requires large differences in the received power of different users, requiring very unequal power allocations. Second, it suffers from a trade-off between performance gain and implementation complexity [125], since the decoding complexity increases with the number of users involved in the communication system.

An alternative MU algorithm that has recently experienced a rise in popularity is non-SIC, in which the multiple users are jointly decoded at the receiver. These MU algorithms present better performance than SIC in several scenarios. For example, authors in [126] show that non-SIC NOMA is more effective than traditional SIC NOMA for sending correlated information sources because, in the latter, the correlated information is not received by all users. Another example is power-balanced scenarios, where SIC does not work well. Authors in [127] presented a novel non-SIC NOMA architecture for near power-balanced scenarios, named Network-Coded Multiple Access, in which they jointly considered physical-layer network coding and the MU detection boosting the throughput. Similarly, [128] combines Bit-Interleaved Coded Modulation (BICM) with Iterative Demapping (ID) and decoding for Multi-User (MU) detection to create MU-BICM-ID systems, which achieve any point on the upper boundary of the MAC capacity region at high rates. However, the complexity of mentioned examples increases substantially with the number of users, yielding to systems with only 3 and 2 simultaneous users.

DISTRIBUTED JOINT SOURCE-CHANNEL CODING OF CORRELATED SOURCES

Unlike in the point-to-point case of Part I, the use of separate source and channel codes can entail significant performance losses if the multiple users transmit the information produced by correlated sources. For example, in the unrealistic approach of defining the energy constraint at the receiver, the transmission can be performed optimally [129] by utilizing separate source and channel coding schemes. However, if the more realistic approach of imposing the energy constraint at the transmitters is considered, separation is not optimal [130] and the theoretical limit has yet to be found. In this dissertation we pay attention to the design of JSC coding schemes with the energy constraint defined at the transmitters.

In this scenarios, it is advantageous to design the user's codewords in such a way that they preserve part of the existing correlation between sources [130]. For example, the work in [131] considers JSC coding for two correlated binary information sequences with turbo codes. The correlation is exploited by exchanging extrinsic information between the constituting decoders as in [132, 133] for turbo codes and [134] for concatenated LDGM codes. The case in which the correlation model is unknown at the decoder

is researched in [135, 136], where the correlation is iteratively estimated during the decoding process. This strategy is applied in [137] with turbo codes and in [138] for LDGM codes. The performance in this research pieces surpasses the theoretical limits obtained when separation between source and channel coding is assumed. [139] introduces a more direct design, namely the parallel connected LDGM code, which allows the correlation between sources to be preserved in the codeword by using the same encoder structure for two users.

There are other works in the literature that have studied this scenario. These works include [140], which utilizes LDPC codes for data collection in Wireless Sensor Networks (WSNs), and [141], which utilizes a single Raptor code for both video compression and packet loss protection for scalable video transmission over wireless networks. However, few existing JSC coding schemes can achieve high transmission rates with good performance, and most results consider only the simplest case of having only 2 correlated sources.

OTHER NOMA SCHEMES

In addition to the power domain and code domain NOMA discussed above, a few other techniques are currently being investigated. Pattern-Division Multiple Access (PDMA) [142] utilizes non-orthogonal patterns, which are designed by maximizing the diversity and minimizing the overlaps among multiple users. This technique can be carried out in the code domain, spatial domain, or a combination. Bit-Division Multiplexing (BDM) [143] is another form of NOMA particularly useful for downlink transmission. Its basic concept is based on hierarchical modulation, and the resources of multiplexed users are partitioned at the bit level. Although strictly speaking the resource allocation of BDM is orthogonal in the bit domain, multi-user signals share the same constellation.

Some other NOMA schemes have also been proposed, such as Interleave Division Multiple Access (IDMA). This technique is a an evolution of Code-Division Multiple Access (CDMA) that relies on interleaving as the only means to distinguish signals and employs parallel interference cancellation to detect the multiple users, inheriting many of the advantages provided by Code-Division Multiple Access (CDMA), but improving its performance in highly loaded systems [144]. Moreover, IDMA allows the accommodation

of a large number of users, up to 100 users [145]. These authors are also able to achieve high power efficiency, with a BER performance 1.4 dB away from the theoretical limit for a 16-user case and a total throughput of 2 bits per channel use. They do so by utilizing complex Turbo-Hadamard codes. Later, authors in [146] proposed the use of low-rate layered LDGM as the core codes of IDMA schemes, simplifying the structure and allowing the sum-product algorithm to be applied for decoding. The scheme presented in [146] outperforms [145] by achieving a gap to the theoretical limit of 1.05 dB for the same total throughput. This is achieved for 20 users with an individual rate of 0.05. In order to accommodate more users, authors in [146] lowered the individual rates, upper bounding the total throughput to 2 bits per channel use. The gap to the the theoretical limit increases with the number of users, regardless of whether the codes have been optimized for the considered number of users. A comparison between IDMA and NOMA when a channel estimator is used was presented in [147]. Authors show that IDMA outperforms NOMA in power-balanced scenarios at the cost of higher complexity, whereas NOMA outperforms IDMA in power imbalanced scenarios. However, they only considered SIC NOMA, which, as mentioned before, is outperformed by non-SIC NOMA in power-balanced scenarios.

In the next chapters we present two novel NOMA coding schemes for the multiple access channel that utilize non-SIC MU algorithms, making them especially valuable for communication systems without large power discrepancies between users. Chapter 6 considers that the information generated by the multiple users is uncorrelated, whereas the correlated case is studied in Chapter 7. Both schemes are based on point-to-point RCM codes, which provide high throughput and rate adaptation capabilities. The main novelty of the proposed MAC schemes lies in the fact that, unlike the previous high rate coding schemes of the literature, these schemes are able to accommodate a large number of users, without affecting decoding performance.

CHAPTER 6

RCM for Non-Orthogonal Multiple Access

This chapter proposes a novel, RCM based, non-SIC NOMA coding scheme for the multiple access channel that allows the transmission of a large number of users at high rates while guaranteeing that the complexity and performance is not negatively affected by the number of simultaneous users. This is an advantage over the coding schemes mentioned previously, which could achieve high data rates or a large number of users, but not both simultaneously.

6.1 REVIEW OF AWGN MAC

We consider the AWGN MAC communication scenario of Section 2.5.2.1, where λ users transmit the information generated by their independent memoryless binary sources to a common receiver. We assume that each user l utilizes an RCM encoder ε_l with generator matrix $G^{(l)} \in \mathbb{R}^{M \times K_l}$, and encodes its information block $\mathbf{u}^l \in \{0,1\}^{K_l}$ at rate R_l, generating the coded sequence $\mathbf{x}^l = G^{(l)}\mathbf{u}^l = x_1^{(l)}, \ldots, x_M^{(l)}$.

We denote the output of the MAC at time j as $y_j = \sum_{l=1}^{\lambda} x_j^{(l)} + z_j = x_j^C + z_j$ and the total transmitted average energy per complex channel use is $E_s = \sum_{l=1}^{\lambda} E_l$, where E_l is the average energy of user l. The capacity

region for this λ-user MAC is given by the set of rate tuples [8] such that

$$\sum_{l\in\mathcal{L}} R_l \leq \log_2\left(1+\frac{\sum_{l\in\mathcal{L}} E_l}{N_0}\right), \quad \text{for every}\mathcal{L}\subseteq[1:\lambda], \tag{6.1}$$

where R_l is the rate of user l. During this dissertation, we will refer as theoretical limit to the equation imposed by the sum-rate

$$R=\sum_{l=1}^{\lambda} R_l \leq \log_2\left(1+\frac{\sum_{l=1}^{\lambda} E_l}{N_0}\right)=\log_2\left(1+\frac{E_s}{N_0}\right). \tag{6.2}$$

Due to the construction method explained later, the average mean energy allocated to each user in the RCM-MAC scheme is proportional to its share of total rate, i.e., $\frac{E_l}{E_s}=\frac{R_l}{R}$ for $l=1,\ldots,\lambda$. It is easy to show that given any set of average energies $\{E_1,\ldots,E_\lambda\}$ with $\sum_{l=1}^{\lambda} E_l = E_s$, the set of rates $\{R_1,\ldots,R_\lambda\}$, $R_l \leq \frac{E_l}{E_s}R^*$ is achievable, where R^* is the theoretical limit for the sum-rate given by (6.2).

Finally, during this chapter we will refer to $\frac{E_s}{N_0}$ as the SNR.

6.2 LINEAR NATURE OF RCM CODES AND THE MAC

The main idea behind the proposed RCM-MAC scheme is to generate weighted linear combinations of each user's own information bits, constructed in such a manner that the overall superimposed coded symbols at the output of the MAC possess the same structure as a point-to-point RCM code. At the receiver, all users are jointly decoded as a point-to-point RCM code, making the decoding complexity independent of the number of users.

Let us begin by exploiting the linear behavior the MAC and RCM codes to interpret the received MAC symbols as if generated from a point-to-point

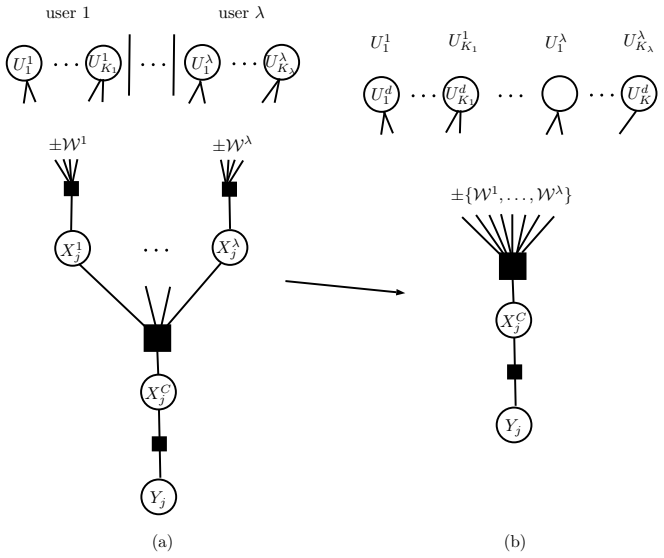


Figure 6.1: *(a) Joint Factor Graph of the MAC and individual RCM codes. (b) Rearrangement of (a) utilizing the linear properties of the MAC and the RCMs.*

RCM code. Each received MAC symbol can be rewritten as

$$\begin{aligned} y_j &= \sum_{l=1}^{\lambda} x_j^{(l)} + Z_j \\ &= \sum_{l=1}^{\lambda} \left(\sum_{k=1}^{K_l} g_{j,k}^{(l)} u_k^{(l)} \right) + Z_j \\ &= \underbrace{\sum_{k=1}^{K} g_{j,k}^{d} u_k^{d}}_{x_j^C} + Z_j, \end{aligned} \tag{6.3}$$

where $\mathbf{u}^d = [\mathbf{u}^1 \ \mathbf{u}^2 \dots \mathbf{u}^\lambda]$ is the sequence generated from concatenating all source blocks, $K = \sum_{l=1}^{\lambda} K_l$ is the total block length of $\mathbf{u}^d$ and the matrix $G^d \in \mathbb{R}^{M \times K}$ is built from the parallel concatenation of all generation

matrices $\{G^{(l)}\}_{l=1}^{\lambda}$ as

$$G^d = \left[G^{(1)} \quad G^{(2)} \quad \dots \quad G^{(\lambda)}\right]. \tag{6.4}$$

Thus, when using RCM codes, the factor graph of the overall MAC scheme can be simplified in such a way that it resembles a point-to-point RCM code generating the symbols $\{x_j^C\}_{j\in\{1,\dots,M\}}$, as shown in Fig. 6.1. This equivalent RCM code has rate R given by the sum-rate of the MAC scheme $R = \sum_l R_l$ and its symbols are given by linear combinations of the input bits of all source sequences $[\mathbf{u}^1\ \mathbf{u}^2 \dots \mathbf{u}^\lambda]$ weighted by the values in $\{\pm\mathcal{W}^1, \dots, \pm\mathcal{W}^\lambda\}$, where $\{\mathcal{W}^l\}$ is the RCM weight set of user $l \in \{1, \dots, \lambda\}$ used in the encoding matrix $G^{(l)}$.

Note that an estimate of $\hat{\mathbf{u}}^d$ can be obtained by running the standard RCM SPA explained in Section 2.6.4 to the simplified factor graph defined by G^d in (6.4). Finally, the estimates of $\{\hat{\mathbf{u}}^l\}_{l=1}^{\lambda}$ are obtained from the corresponding position of $\hat{\mathbf{u}}^d$.

6.3 RCM-MAC SCHEME DESIGN

As mentioned before, given the additive nature of the RCMs and the MAC, the superimposed MAC symbols can be considered as having been generated by a standard point-to-point RCM code encoding the compound source bit block $\mathbf{u}^d = [\mathbf{u}^1\ \mathbf{u}^2 \dots \mathbf{u}^\lambda]$. Following this, the proposed RCM-MAC schemes consist of a set of λ irregular RCM codes, with encoding matrices $G^{(l)}$, one for each individual user. By irregular RCM it is meant that the weight values utilized in the sums of each RCM symbol can vary, i.e., the weight multi-set is dependent on the symbol j. Let us denote by $\mathcal{W}_j^{(l)}$ the design multi-set of user l and RCM symbol j, which generates RCM symbol $x_j^{(l)}$. The j^{th} RCM MAC overall symbol is connected to $\mathbf{u}^d$ by means of the weights in $\{\pm\mathcal{W}_j^{(1)}, \pm\mathcal{W}_j^{(2)}, \dots, \pm\mathcal{W}_j^{(\lambda)}\}$.

The main idea behind the design of RCM-MAC schemes is to specifically select the multi-sets of the irregular RCMs to match, after the multiple access channel, a point-to-point RCM, of rate equal to the sum-rate R, with design multi-set $\mathcal{W}$ by imposing the condition $\{\mathcal{W}_j^{(1)}, \mathcal{W}_j^{(2)}, \dots, \mathcal{W}_j^{(\lambda)}\} = \mathcal{W}$ for all j as sketched in Fig. 6.2.

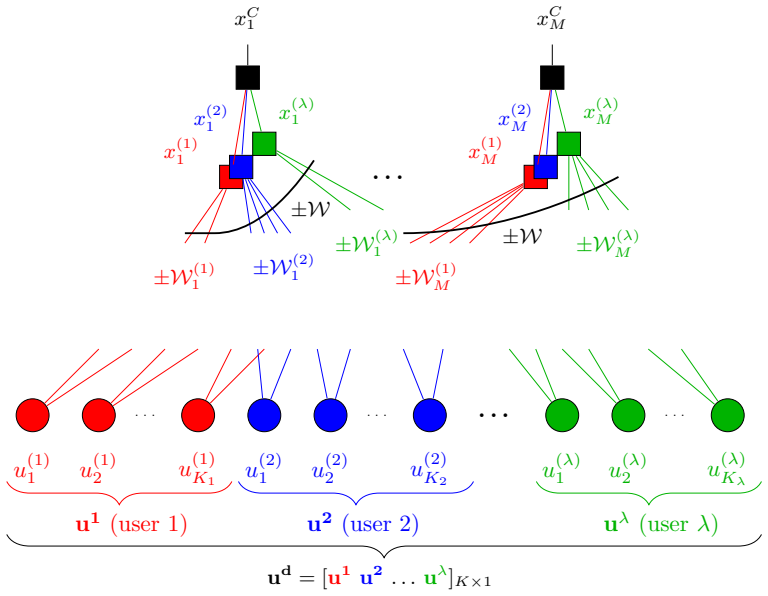


Figure 6.2: *Factor graph of the simplified rearrangement of the Joint Factor Graph of the MAC. The individual RCM codes building the MAC scheme are represented in different colors.*

6.3.1 GENERATION OF INDIVIDUAL RCM CODES

Let R_l, for $l = 1, \ldots, \lambda$, be the rate at which each user transmits, so that the sum-rate is equal to $R = \sum_{l=1}^{\lambda} R_l$. The number of coded symbols per block, M, is fixed and the source block length for user l, K_l, is adjusted to meet the corresponding R_l.

The first step is to find a weight multi-set $\mathcal{W}$ (for the final equivalent RCM code) that i) satisfies the third property mentioned in Section 2.6.1 and ii) leads to good performance in a point-to-point communication system with rate equal to the required sum-rate in the multi-user scenario, i.e., $R = \sum_{l=1}^{\lambda} R_l$. This weight set can be found with the aid of the EXIT chart techniques presented in Chapter III. Once the weight set $\mathcal{W}$ is selected, for each coded symbol at position j we randomly assign each element of $\mathcal{W}$ to one of the λ multi-sets $\mathcal{W}_j^{(l)}$. The probability of being assigned to

the multi-set l is proportional to $\frac{R_l}{R}$, except in some instances in which, as explained in the sequel, the probability is set to 0. Notice that some multi-sets, $\mathcal{W}_j^{(l)}$, may be either left empty or have more than one element.

Next, we detail the construction of encoding matrices $G^{(l)}$, for $l = 1, \ldots, \lambda$, to be assigned to each user.

1. **Step 1: Generation of the intermediate matrices $A^{(l)}$, $l = 1, \ldots, \lambda$.**

   Define matrices $A^{(l)}$, $l = 1, \ldots, \lambda$, as all zeros $M \times K_l$ matrices. We begin by constructing the first row of matrices $A^{(l)}$ in two steps.

   - Randomly assign each element of $\mathcal{W}$ to one of the λ multi-sets $\mathcal{W}_1^{(l)}$ with probability $\frac{R_l}{R}$.
   - For $l = 1, \ldots, \lambda$ substitute the $2 \cdot |\mathcal{W}_1^{(l)}|$ zero matrix entries, $a_{1,1}, \ldots, a_{1,2\cdot|\mathcal{W}_1^{(l)}|}$, in the first row of matrix $A^{(l)}$ by the elements of $\pm\mathcal{W}_1^{(l)}$. If the set $\mathcal{W}_1^{(l)}$ is empty, the corresponding row is left with zeros.

   Let $\kappa(l)$ be the index of the first column of matrix $A^{(l)}$ whose entries are all zero. Note at this point, $\kappa(l) = 1 + 2 \cdot |\mathcal{W}_1^{(l)}|$ for $l = 1, \ldots, \lambda$.

   For $j \geq 2$ until $\kappa(l) > K_l$, $l = 1, \ldots, \lambda$, the row j of matrix $A^{(l)}$ is generated using the following iterative steps:

   (a) Randomly assign each element of $\mathcal{W}$ to one of the λ multi-sets $\mathcal{W}_j^{(l)}$. The probability of being assigned to the multi-set l is proportional to $\frac{R_l}{R}$, unless $\kappa(l) > K_l$. In this case, the probability of assignment to multi-set $\mathcal{W}_j^{(l)}$ is fixed to 0 (i.e., $\mathcal{W}_j^{(l)}$ will be void), and the remaining probabilities are re-normalized. This is done to promote that the condition $\kappa(l) > K_l$, $l = 1, \ldots, \lambda$, be met for all users in a similar number of iterations, with the goal of obtaining final encoding matrices that are regular in columns.

   (b) For $l = 1, \ldots, \lambda$ and when $\kappa(l) + 2 \cdot |\mathcal{W}_j^{(l)}| \leq K_l$ substitute the $2 \cdot |\mathcal{W}_j^{(l)}|$ zero entries, $a_{j,\kappa(l)}, \ldots, a_{j,\kappa(l)+2\cdot|\mathcal{W}_j^{(l)}|}$, in row j of matrix $A^{(l)}$ by the elements of $\pm\mathcal{W}_j^{(l)}$. Otherwise, note that there will be $\left(\kappa(l) + 2 \cdot |\mathcal{W}_j^{(l)}| - K_l - 1\right)$ elements that can not be allocated

in this manner. These remaining elements are then randomly placed in the unoccupied positions (1 to $\kappa(l)-1$) of this row[1].

(c) Set $\kappa(l) = \kappa(l) + 2 \cdot |\mathcal{W}_j^{(l)}|$, $l = 1, \ldots, \lambda$ and go to the next iteration step with $j = j + 1$.

Denote by J_0 the value of j for which the condition $\kappa(l) > K_l$ is met for all l. Then, matrices $A^{(l)}$, $l = 1, \ldots, \lambda$, are obtained by discarding their last $M - J_0$ zero rows. Note that this construction requires J_0 to be smaller than M. However, the expected value of J_0 is $\frac{K}{2 \cdot d_r}$, independent of the number of users, allowing a potentially unlimited number of simultaneous users.

2. **Step 2: Generation of the encoding matrices $G^{(l)}$, $l = 1, \ldots, \lambda$.** Matrix $G^{(l)}$, is obtained by vertically stacking L randomly column-wise permuted $A^{(l)}$ matrices, where $M \leq LJ_0 < M + J_0$. Note that in the last stacked permuted matrix, $A^{(l)}$, only $V = M - (L-1)J_0$ rows should be kept.

The proposed RCM-MAC construction ensures that the superimposed coded symbols at the output of the MAC channel are a weighted linear combination of the input bits proceeding from several users, with the weights belonging to $\pm\mathcal{W}$. Considered as a point-to-point RCM code, an RCM-MAC scheme satisfies the three properties of good RCM codes, as mentioned in Section 2.6.1: i) it is regular in rows, with a "total" design weight set $\mathcal{W}$ by construction; ii) its weight set creates diverse symbol values which can carry the information of all users; and iii) the constructed encoding matrix is also regular in columns.

Notice that due to the use of symmetric weights $\pm\mathcal{W}_j^{(l)}$, we ensure that the output constellation for each user has zero mean, minimizing the average energy per user. In this chapter, we only consider uncorrelated users. Thus, coded symbols from different users remain uncorrelated, and, assuming no noise, the mean energy of the MAC output is the sum of the transmitted mean energies of all users. Since some users may not transmit during certain time slots, superposition coding and time division occur

[1]It can be assumed that the individual block length of the users is always much larger than the length of weight set $K_l >> 2 \cdot |\mathcal{W}| \geq 2 \cdot |\mathcal{W}_j^{(l)}|$. Therefore, there will always be enough unoccupied positions to place these $\left(\kappa(l) + 2 \cdot |\mathcal{W}_j^{(l)}| - K_l - 1\right)$ elements.

implicitly within the proposed RCM-MAC scheme, especially when the number of users increases.

Recall that at the common receiver, an estimate of $\hat{\mathbf{u}}^d$ is obtained by running the standard RCM SPA explained in Section 2.6.4 to the simplified factor graph defined by G^d.

6.4 RAYLEIGH FAST FADING MAC

We consider the fading MAC communication scenario of Section 2.5.2.2. In this section we will consider complex baseband sequences. Let us denote by $\widetilde{\mathbf{x}}^l$ to the output (which is the QAM modulated version of $\mathbf{x}^l$) of transmitter l. The output of the channel at time j is given by $y_j = \sum_{l=1}^{\lambda} \widetilde{x}_j^{(l)} \cdot h_j^{(l)} + z_j$, where the fading factors $\{h_j^{(l)}\}$ are modeled as in Section 2.5.2.2, Z_j are symmetric complex Gaussian r.v.s, and j spans from 1 to $\frac{M}{2}$. The achievable rate bounds are obtained by averaging the inequalities (6.1) with respect to the exponential random variables $|H^{(l)}|^2 \sim e^{-h}$, $l = 1, \ldots, \lambda$, that is,

$$\sum_{l \in \mathcal{L}} R_l \leq E\left[\log_2\left(1 + \frac{\sum_{l \in \mathcal{L}} \left|h^{(l)}\right|^2 E_l}{N_0}\right)\right], \quad \text{for every } \mathcal{L} \subseteq [1:\lambda], \tag{6.5}$$

where again R_l is the rate of user l. Similar to the AWGN channel case, the theoretical limit will be given as

$$R = \sum_{l=1}^{\lambda} R_l \leq E\left[\log_2\left(1 + \frac{\sum_{l=1}^{\lambda} \left|h^{(l)}\right|^2 E_l}{N_0}\right)\right]. \tag{6.6}$$

Applying the Jensen Inequality, it follows that the right-side term in (6.6) is upper-bounded by the corresponding bound in (6.2). By the Law of Large Numbers, it is easy to check that the upper-bounds for the MAC AWGN and MAC Rayleigh fast fading channels are the same when the number of users goes to infinity.

6.4.1 DECODING IN RAYLEIGH FAST FADING MAC

When perfect Channel State Information (CSI) is available at the receiver in point-to-point communications, the fading complex gain h_j of

coded symbol j is canceled by multiplying the received sample by $\frac{\overline{h_j}}{|h_j|^2}$, where $\overline{h_j}$ denotes the conjugate of h_j. Decoding is performed by incorporating the scaling factor $\frac{1}{|h_j|^2}$ to the noise variance of symbol j in the factor graph. However, in the multi-user scenario this is not possible, since the received symbols are combinations of several coded symbols, each with its own fading factor $\{h_j^{(1)}, h_j^{(2)}, \ldots, h_j^{(\lambda)}\}$. In order to decode RCM-MAC schemes using standard belief propagation, the fading factors have to be taken into account in the overall factor graph.

Recall that, unlike the previous AWGN sections, for the Rayleigh fast fading MAC we consider the received symbols y_j as complex (QAM modulated) and denote the complex MAC channel output symbols as $\widetilde{x}_j^C$. We begin by rewriting each of the $\frac{M}{2}$ received complex symbols, for $j = 1, \ldots, \frac{M}{2}$, as

$$\begin{aligned} y_j &= \underbrace{\sum_{l=1}^{\lambda} \widetilde{x}_j^{(l)} \cdot h_j^{(l)}}_{\widetilde{x}_j^C} + z_j = \sum_{l=1}^{\lambda} \left(x_{2j-1}^{(l)} + i \cdot x_{2j}^{(l)} \right) \cdot h_j^{(l)} + z_j \\ &= \sum_{l=1}^{\lambda} \Big(\underbrace{\sum_{k=1}^{K_l} g_{2j-1,k}^{(l)} u_k^{(l)}}_{x_{2j-1}^{(l)}} + i \cdot \underbrace{\sum_{k=1}^{K_l} g_{2j,k}^{(l)} u_k^{(l)}}_{x_{2j}^{(l)}} \Big) \cdot h_j^{(l)} + z_j \\ &= \sum_{l=1}^{\lambda} \left(\sum_{k=1}^{K_l} \left(\left(g_{2j-1,k}^{(l)} + i \cdot g_{2j,k}^{(l)} \right) h_j^{(l)} \cdot u_k^{(l)} \right) \right) + z_j, \end{aligned} \tag{6.7}$$

where $h_j^{(l)}$ is the complex fading factor for the complex symbol j of user l. Note that the QAM modulated transmitted RCM symbols $[\widetilde{x}_j^{(1)}, \widetilde{x}_j^{(2)}, \ldots, \widetilde{x}_j^{(\lambda)}]$ are multiplied in the MAC by the complex fading factors $[h_j^{(1)}, h_j^{(2)}, \ldots, h_j^{(\lambda)}]$ before being added together.

Figure 6.3 illustrates the complex factor graph of the j^{th} received MAC symbol for the λ-user RCM-MAC scheme based on (6.7). Starting from the bottom, the real RCM symbols $x_{2j-1}^{(l)}$ and $x_{2j}^{(l)}$, for $l = 1, \ldots, \lambda$, are obtained as in Fig. 6.2. These RCM real symbols are then paired two-by-two. Following this, each of these transmitted RCM QAM symbols is multiplied by its own complex fading factor $h_j^{(l)}$, making it impossible for the receiver to cancel the fading factors, as is usually done in point-to-

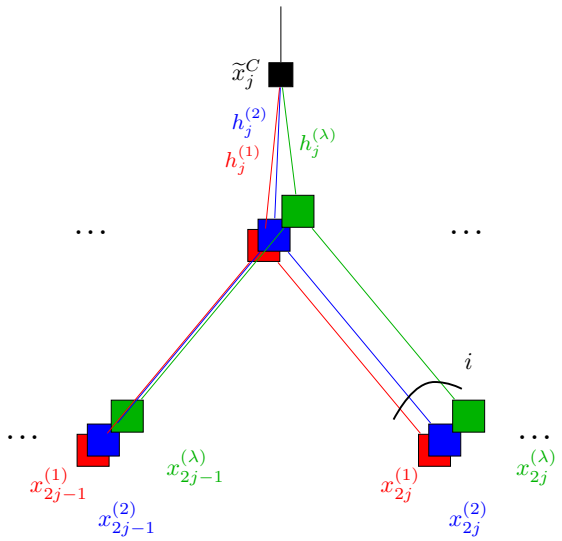


Figure 6.3: *Factor graph of the j^{th} received MAC symbol y_j based on (6.7).*

point fading channels. These complex multiplications mix the real (in-phase) and imaginary (quadrature) part of the signals. In order to avoid the creation of cycles of length 4 that arise when separating the real and imaginary parts, we perform decoding directly over the two-by-two complex modulated fading factor graph, defining the complex decoding matrix $\widetilde{G}^d$ as

$$\widetilde{G}^d = \left[H^{(1)} \cdot \widetilde{G}^{(1)} \quad H^{(2)} \cdot \widetilde{G}^{(2)} \quad \dots \quad H^{(\lambda)} \cdot \widetilde{G}^{(\lambda)} \right]_{\frac{M}{2} \times K}, \tag{6.8}$$

where $H^{(l)}$, for $l = 1, \dots, \lambda$, is a complex diagonal matrix, of size $\frac{M}{2} \times \frac{M}{2}$, whose diagonal elements are the complex fading factors $[h_1^{(l)} h_2^{(l)} \dots h_{\frac{M}{2}}^{(l)}]$. The matrices $\widetilde{G}^{(l)}$, of size $\frac{M}{2} \times K$, are obtained from the matrices $G^{(l)}$ as follows. Each row j, for $j = 1, \dots, \frac{M}{2}$, of $\widetilde{G}^{(l)}$ is obtained by adding the $2j-1$ row in $G^{(l)}$ and the $2j$ row in $G^{(l)}$ multiplied by the imaginary unit i, as in the illustration of Fig. 6.3. From the receiver's perspective, the output of the fading MAC can be seen as a point-to-point (complex) RCM with varying weight multi-sets, which change due to the effect of the fading factors.

Finally, as in the previous AWGN example, the receiver can decode the extended block $\mathbf{u^d}$ of length $K = \sum_{l=1}^{\lambda} K_l$, defined as $[\mathbf{u}^1\ \mathbf{u}^2\ \ldots\ \mathbf{u}^\lambda]$, by executing the sum-product algorithm [9] over the graph describing $\widetilde{G}^d$ in (6.8). The steps of this algorithm are the same of Section 2.6.4 whose equation (2.61) is adapted to the complex field.

6.5 RESULTS

To validate the performance of the proposed RCM-MAC schemes, we next present numerical results obtained by Monte Carlo simulations. We have considered three high rate scenarios with sum-rate, R, equal to 6, 7.4 and 9 bits per complex channel symbol. The total number of information bits transmitted per block K is 37000, and the block length of each user has been adjusted according to the number of users and their rate splits. The decoding is performed by applying the sum-product algorithm of Section 2.6.4, which, for the Rayleigh fading case, is performed over the complex field.

We consider two different rate allocation schemes, and start by studying the case where the λ users transmit with symmetric rate allocation: Each user utilizes blocks of $\frac{37000}{\lambda}$ information bits (or the closest number to 37000 that is divisible by λ). We also consider the case where the λ users transmit in an asymmetric way, while maintaining a total sum-rate R equal to 6, 7.4 and 9 bits per complex channel symbol. In this case, the source block length of user l with rate R_l is given by the closest integer to $37000\frac{R_l}{R}$. Table 6.1 summarizes the 7 different asymmetric rate allocations with **(a)** $\lambda = 2$, **(b)** $\lambda = 3$, **(c)** $\lambda = 4$, **(d)** $\lambda = 5$, **(e)** $\lambda = 9$, **(f)** $\lambda = 13$ and **(g)** $\lambda = 16$ users that are simulated in this chapter.

Notice that independently of the number of users and the rate allocations, for an AWGN MAC with a fixed E_s, the sum-rate can be achieved not only by superposition as in the RCM-MAC scheme, but also with point-to-point RCM codes combined with an adequate time-division multiple access (TDMA) [8] (provided that users are allowed to use an average mean energy E_s during their transmission periods). However, many practical applications impose mean energy limitations (lower than E_s) on each user, making RCM-MAC the best option in such cases.

On the other hand, when Rayleigh fast fading is considered as the number of users increases, the performance of superposition coding overpasses

Table 6.1: *Rate allocations in bits per complex channel symbol for the asymmetric cases considered in this chapter.*

	λ	R_ls for $R = 6$
a	2	$\{2, 4\}$
b	3	$\{2 \times 1.5, 3\}$
c	4	$\{0.5, 2 \times 1.5, 2.5\}$
d	5	$\{2 \times 0.66, 2 \times 1.33, 2\}$
e	9	$\{4 \times 0.4, 4 \times 0.8, 1.2\}$
f	13	$\{4 \times 0.2, 4 \times 0.4, 4 \times 0.6, 1.2\}$
g	16	$\{4 \times 0.15, 4 \times 0.3, 4 \times 0.45, 4 \times 0.6\}$
	λ	R_ls for $R = 7.4$
a	2	$\{2.46, 4.93\}$
b	3	$\{2 \times 1.85, 3.7\}$
c	4	$\{0.62, 2 \times 1.85, 3.08\}$
d	5	$\{2 \times 0.82, 2 \times 1.64, 2.46\}$
e	9	$\{4 \times 0.49, 4 \times 0.98, 1.48\}$
f	13	$\{4 \times 0.24, 4 \times 0.49, 4 \times 0.74, 1.48\}$
g	16	$\{4 \times 0.185, 4 \times 0.37, 4 \times 0.55, 4 \times 0.74\}$
	λ	R_ls for $R = 9$
a	2	$\{3, 6\}$
b	3	$\{2 \times 2.25, 4.5\}$
c	4	$\{0.75, 2 \times 2.25, 3.75\}$
d	5	$\{2 \times 1, 2 \times 2, 3\}$
e	9	$\{4 \times 0.6, 4 \times 1.2, 1.8\}$
f	13	$\{4 \times 0.3, 4 \times 0.6, 4 \times 0.9, 1.8\}$
g	16	$\{4 \times 0.225, 4 \times 0.45, 4 \times 0.675, 4 \times 0.9\}$

the performance of a TDMA scheme, even if users are allowed to use higher powers during their transmission times. In other words, the combination of a point-to-point RCM code and a TDMA scheme with power control does not achieve the same performance attained when using superposition coding with an RCM-MAC scheme (refer to Section 6.5.2).

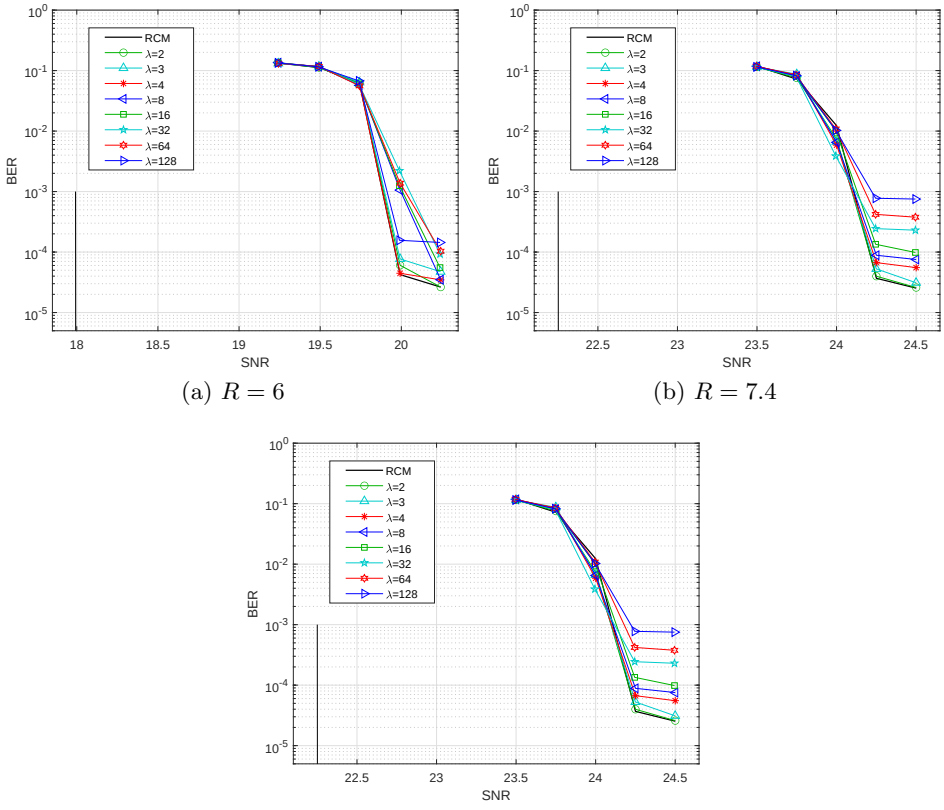
6.5.1 NUMERICAL RESULTS FOR AWGN MAC

We begin by analyzing the performance of the proposed RCM-MAC scheme for the AWGN channel. The corresponding theoretical limits given by (6.2) are SNR = $\frac{Es}{N_0}$ = 18 dB for $R = 6$, SNR = 22.3 dB for $R = 7.4$ and SNR = 27.1 dB for $R = 9$. Fig. 6.4 shows the BER versus SNR for the symmetric multi-user scenario, with λ =2, 3, 4, 8, 16, 32, 64 and 128 users, whereas results for the asymmetric rate allocations of Table 6.1 are shown in Fig. 6.5. To provide further context, we have also depicted the performance of a conventional RCM code with the same rate in a point-to-point AWGN channel. Its performance is equivalent to performing the OMA strategy of TDMA with power control, as discussed previously. The weight sets $\mathcal{W}$ considered in the simulations are $\{3, 3, 4, 5\}$, $\{3, 4, 5, 7\}$ and $\{3, 4, 5, 5, 7\}$ for R =6, 7.4 and 9, respectively. They have been selected using EXIT charts (refer to Chapter 3) of standard point-to-point RCM codes, with a rate equal to the intended sum-rate R.

Notice that in all cases the proposed RCM-MAC scheme presents a waterfall around 2 dBs away from the sum-rate capacity limit (6.2). As the number of users increases, the additional structure that is imposed on the construction of the generation matrices $G^{(l)}$ results in higher error floors compared to the point-to-point code. This degradation of the error floor with the number of users is more pronounced when the sum-rate increases. For example, in the scenario with $R = 9$ and symmetric rate allocation shown in Fig. 6.4c, the error floor goes from near 10^{-5} for $\lambda = 2$, to 10^{-4} and $7 \cdot 10^{-4}$ for $\lambda = 16$ and $\lambda = 128$, respectively.

The increase of the error floor as the number of users increases can be explained by the decrease of the individual block length K_l of each user, which is inversely proportional to the number of users. This effect can be eliminated if the total block length is increased according to λ, so that K_l stays fixed when the number of users increases. This is shown in Fig. 6.6 for the case of $R = 7.4$ with symmetric rate allocation. Figs. 6.6a, 6.6b and 6.6c show the performance when the individual block length of each user K_l, for $l = 1, \ldots, \lambda$, is fixed to 1110, 4440 and 8880, respectively. The coded block length is given by $M = \frac{2K}{R}$, where $K = \lambda \cdot K_l$ is the total number of information bits transmitted per block.

Figs. 6.6a, 6.6b and 6.6c show that the performance depends both on the total block length K and the individual block length K_l. For example,

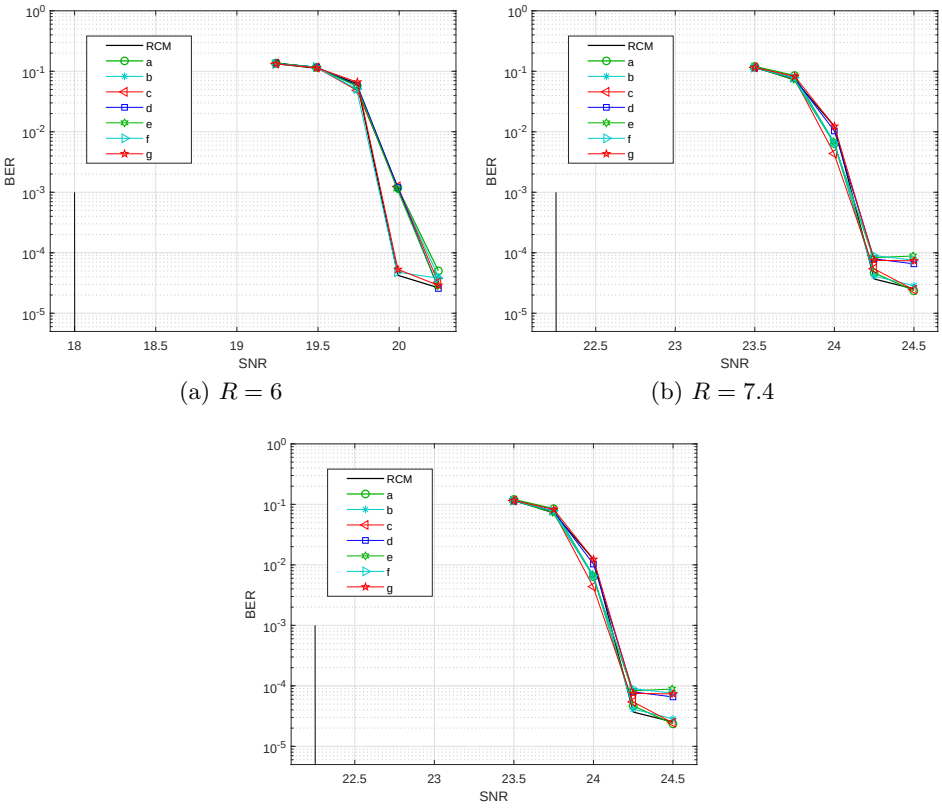


(a) $R = 6$ (b) $R = 7.4$

(c) $R = 9$

Figure 6.4: *Numerical results obtained by Monte Carlo simulations for the symmetric multi-user scenario with* $\lambda \in \{2, 3, 4, 8, 16, 32, 64, 128\}$ *users and for point-to-point (RCM) when the AWGN MAC is considered.*

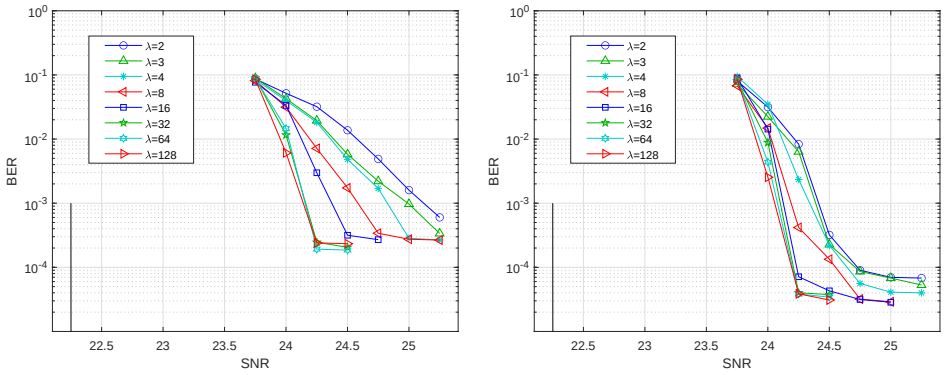
in the case of $K_l = 1110$ there is a minimum error floor of around 10^{-4} that is achieved with $\lambda = 16$, (i.e. a total block length $K = 17760$), and the error floor does not improve any further when λ increases. On the other hand, when $K_l = 8880$ and $\lambda = 2$, with the same total block length $K = 17760$, the error floor improves to $6 \cdot 10^{-5}$. As can be seen in Fig. 6.6c, any number of users can be sent with a low error floor provided that the total and the individual (per user) block lengths are large enough.



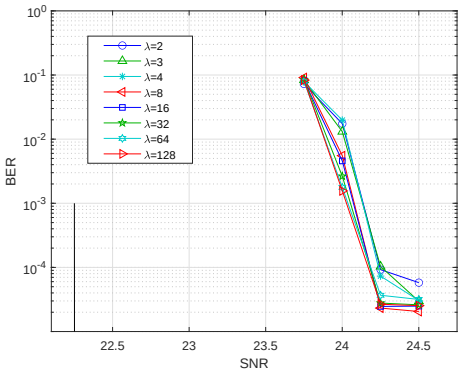
(a) $R = 6$ (b) $R = 7.4$

(c) $R = 9$

Figure 6.5: *Numerical results obtained by Monte Carlo simulations for the asymmetric multi-user scenario with rate distributions as in Table 6.1 and for point-to-point (RCM) when the AWGN MAC is considered.*



(a) $K_l = 1110$ (b) $K_l = 4440$



(c) $K_l = 8880$

Figure 6.6: *Numerical results obtained by Monte Carlo simulations for the symmetric multi-user scenario of $R = 7.4$ with $\lambda \in \{2, 3, 4, 8, 16, 32, 64, 128\}$ when the AWGN MAC is considered.*

6.5.2 NUMERICAL RESULTS FOR THE RAYLEIGH FAST FADING MAC

As mentioned in Section 6.4, when the number of users increases, the sum-rate capacity of the Rayleigh fast fading MAC increases towards its maximum value, namely the sum-rate capacity of the AWGN MAC. Assuming symmetric rate allocation, Table 6.2 shows the SNR Shannon limits for $\lambda = 1, 2, 3, 4$ ($\lambda = 1$ equals to point-to-point RCM) and sum-rates R=6, 7.4 and 9. The table also shows the SNR Shanon limit when the number of users grows asymptotically.

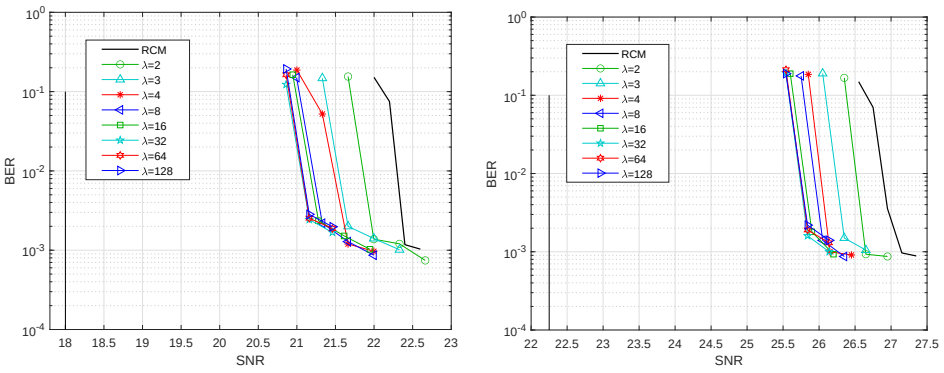
Table 6.2: *Shannon SNR limits (in dB) for the symmetric rate allocation cases with sum-rate R in the Rayleigh fast fading MAC.*

λ	$R = 6$	$R = 7.4$	$R = 9$
1	20.4	24.7	29.6
2	19.1	23.4	28.3
3	18.7	23	27.8
4	18.6	22.8	27.7
∞	18	22.3	27.1

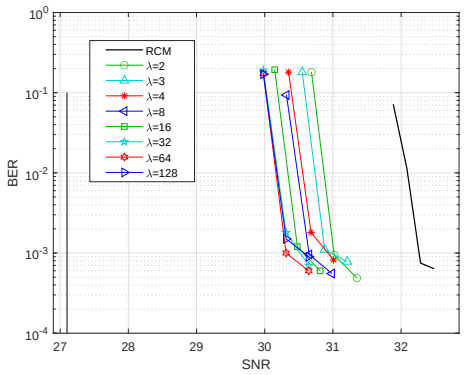
For the simulations, the optimized weight sets $\mathcal{W}$ are $\{2, 3, 4, 4, 7\}$ for $R = 6$ and $\{3, 4, 5, 8, 10\}$ for $R = 7.4$ and 9, independently of the number of users ($\lambda = 1, \ldots, 128$). Fig. 6.7 plots the resulting performance for the symmetric rate scenario, with $\lambda = 2, 3, 4, 8, 16, 32, 64$ and 128, indicated as $\lambda \times 1$ in the figure, and the point-to-point RCM. For the sake of clarity in the figures, only the SNR limits corresponding to a larger number of users (SNR_∞) are plotted as vertical lines, with values given by $\text{SNR}_\infty = 18$ dB, $\text{SNR}_\infty = 22.3$ dB and $\text{SNR}_\infty = 27.1$ dB for $R = 6$, 7.4 and 9, respectively.

Notice that, following the theoretical limits, performance improves as the number of users increases. For example, Fig. 6.7a shows that for $R = 6$ the MAC scenario with $\lambda = 16$ users requires an SNR of around 0.7 dB less than the system with $\lambda = 2$ users, and 1.15 dB less SNR than the point-to-point RCM code, to achieve the same BER. Similarly, we can see in Fig. 6.7b that for $R = 7.4$ and $\lambda = 16$ users the SNR improvement is around 1.15 dB and 0.65 dB compared to the results of the point-to-point RCM code and to the case with $\lambda = 2$ users, respectively. Finally, results

for $R = 9$ are shown in Fig. 6.7c. In this case, the SNR improvement for $\lambda = 16$ users compared to the point-to-point RCM case is increased to around 2 dB. Notice also that for more than 16 users the performance no longer improves significantly. The reason is that the weight sets have length five, and thus, at each RCM-MAC symbol only the averaging effect of up to 10 users applies.

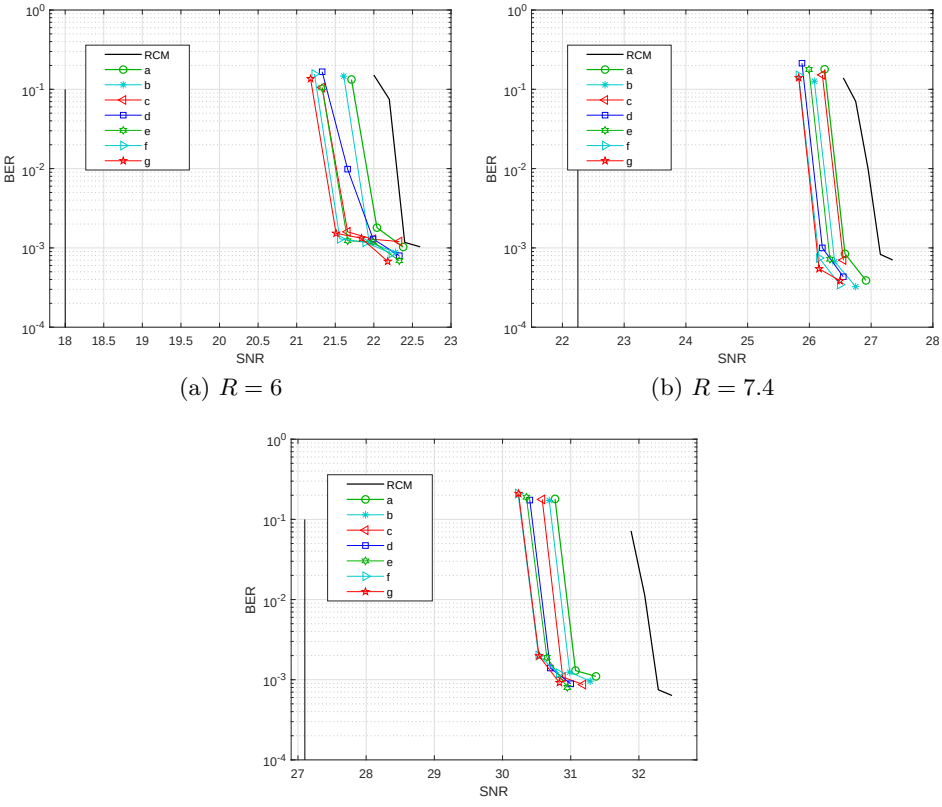


(a) $R = 6$ (b) $R = 7.4$



(c) $R = 9$

Figure 6.7: *Numerical results obtained by Monte Carlo simulations for the symmetric multi-user scenario with* $\lambda \in \{2, 3, 4, 8, 16, 32, 64, 128\}$ *and the point-to-point RCM code (denoted by RCM) when a Rayleigh fast fading MAC is considered.*



(a) $R = 6$ (b) $R = 7.4$

(c) $R = 9$

Figure 6.8: *Numerical results obtained by Monte Carlo simulations for the asymmetric multi-user scenario with rate distributions as in Table 6.1 and the point-to-point RCM code (denoted by RCM) when a Rayleigh fast fading MAC is considered.*

The results for the asymmetric rate allocation case are showed in Fig. 6.8, where the same trends as in the symmetric scenario are observed. To sum up, Table 6.3 presents the SNR performance improvements obtained in the proposed RCM-MAC scheme when the number of users increases and the channel is Rayleigh fast fading. The entries indicate the SNR improvement (in dB, and for BER = 10^{-2}) with respect to the case of 1 user utilizing a point-to-point RCM code.

Table 6.3: *Summary of the SNR performance improvement (in dB) when comparing the RCM-MAC scheme with the point-to-point RCM in a Rayleigh fast fading MAC ($\lambda = 1$).*

	RCM-MAC	$R = 6$	$R = 7.4$	$R = 9$
symmetric	$\lambda = 2$	0.45	0.5	1.3
	$\lambda = 3$	0.76	0.65	1.4
	$\lambda = 4$	0.84	0.87	1.6
	$\lambda = 8$	1.1	1.07	1.67
	$\lambda = 16$	1.15	1.13	1.8
	$\lambda = 32$	1.25	1.2	1.95
	$\lambda = 64$	1.25	1.2	1.95
	$\lambda = 128$	1.25	1.2	1.95
asymmetric	a	0.4	0.53	1.15
	b	0.5	0.57	1.25
	c	0.65	0.73	1.35
	d	0.78	0.77	1.53
	e	0.8	0.87	1.55
	f	0.88	0.95	1.68
	g	0.92	0.97	1.68

6.6 DISCUSSION

We have proposed a new high-rate coding scheme, named RCM-MAC, suitable for non-orthogonal multiple access. It is constructed in such a way that it takes advantage of the linear behavior of both RCM codes and the multiple access channel in order to design a MAC scheme that emulates a point-to-point RCM code from the receiver's perspective. As such, it inherits the large system throughputs of RCM codes and is able to accommodate a large number of users without impacting the performance and decoding complexity.

For the AWGM multiple access channel, and a large range of simultaneous users, simulation results show that the BER vs SNR curves of the proposed RCM-MAC scheme are around 2 dB away from the Shannon limit, similar to the expected performance of a point-to-point RCM. On the other hand, when considering the Rayleigh fast fading channel, the Shannon limit

depends on the number of users. The performance of our proposed scheme is around 2.7-3.5 dB away from these limits. It is worth mentioning, that due to the average effect of having multiple users, the proposed RCM-MAC schemes outperform the combination of point-to-point RCM and TDMA with power control by up to 2 dB.

Finally, it is interesting to notice that for all rate allocation scenarios, as in the case of standard point-to-point RCM codes, the proposed RCM-MAC schemes suffer from high error floors. This paves the way for future research on lowering the error floor by concatenating the proposed RCM-MAC to other codes, as has been done for point-to-point RCM codes.

CHAPTER 7

RCM for Correlated Information Sources

In this chapter we present an RCM based joint source-channel coding scheme for the transmission of correlated information sources. Unlike the previous joint source-channel schemes encountered in the literature, the proposed scheme allows high transmission rates while also considering scenarios with multiple correlated sources. In order to lower the error floor of the scheme presented in this chapter and in the previous one, we also extend the RCM-LDGM parallel concatenation to the multi-user scenario.

7.1 JOINT SOURCE-CHANNEL CODING

Consider μ jointly ergodic sources $\left\{\mathbf{U}^l\right\}_{l=1}^{\mu}$ defined over countably infinite alphabets. The Slepian-Wolf [148, 149] result states that they can be losslessly compressed at rate tuples $(R'_1, R'_2, \ldots, R'_\mu)$ such that

$$\sum_{i\in\mathcal{S}} R'_i \geq H\left(\mathbf{U}(\mathcal{S}) \;\middle|\; \mathbf{U}(\mathcal{S}^c)\right) \quad \text{for all } \mathcal{S} \subseteq [1:\mu], \tag{7.1}$$

where $H(\cdot)$ is the entropy, $\mathcal{S}^c$ is the complementary set of $\mathcal{S}$ and $\mathbf{U}(\mathcal{S}) = \{\mathbf{U}^i : i \in \mathcal{S}\}$. Each source is losslessly compressed in a distributed manner and the original sequences are recovered by applying a joint decoder. Authors in [150, 151, 152] reformulate the Slepian-Wolf result as a channel coding problem with side information, enabling the use of well-establish

channel codes to exploit prior knowledge at the decoder. Some coding structures used in the literature include turbo codes [42], Low-Density Parity Check (LDPC) codes [14, 43, 44], and concatenated Low-Density Generator Matrix (LDGM) codes [153, 154].

The problem becomes more challenging if the compressed sources have to be transmitted through noisy channels. Techniques for exploiting the correlation at the access point when the transmission is performed over orthogonal channels are investigated in [153, 155, 156, 157, 158, 159, 160]. In this chapter we focus on the transmission of these sources over non-orthogonal multiple access. Concretely, we consider a non-SIC NOMA as it was shown in [126] that it is more effective than SIC because, in the latter, the correlated information is not received by all users. Thus, for the rest of this chapter (as in the previous chapter), we will consider non-SIC based NOMA schemes.

While the use of separate source and channel codes is well motivated for the point-to-point case, it can entail significant performance losses in the considered scenario with transmitted energy constraint, as explained in the introduction of Part II. In this chapter we propose a coding scheme based on RCM schemes to transmit CIS over the MAC at high rates. The sparse nature of the linear combinations that generate RCM symbols is well suited to design codewords that preserve part of the source correlation. The proposed scheme, which we will denote as RCM-CIS, consists of a set of RCM codes, one per user, with the same random RCM structure. In this way, RCM symbols colliding in the channel are generated from weighted linear combinations of correlated bits. By appropriately choosing the weight set values for each user, high symbol reinforcement, and thus correlation reinforcement, can be achieved. As in the case of RCM for point-to-point communications and of RCM-MAC schemes, the proposed RCM-CIS schemes are able to adapt their rate seamlessly by varying the number of generated symbols. RCM-CIS schemes are different from the schemes in Chapter 6 in the sense that the latter are created by distributing an RCM structure into irregular RCM codes that are assigned to different users, whereas RCM-CIS replicate the same RCM structure for each user and only vary the weight values.

Unfortunately, if conventional binary decoding is used, the overall factor graph resulting from the combination of the correlated sources' factor graph with the replicated RCM random structures will present short length cycles,

and consequently degrade the decoding performance. Moreover, although messages are passed between the correlated source symbols in the sum-product algorithm, not considering these probabilities jointly will lead to some loss of source correlation information. To overcome these problems, we propose the use of a novel non-binary joint decoder for the RCM-CIS schemes. The concept of non-binary decoding has been studied in the literature [161, 162, 163] and has been recently applied to decode LDPC and LDGM codes designed specifically for the quantum communication paradigm [164, 165, 166, 167], but has never been applied to decode RCM codes.

Finally, in this chapter we also we extend the parallel RCM-LDGM concatenation to the MAC and apply it to improve the error floor of RCM-CIS schemes.

7.2 SYSTEM MODEL

We consider the transmission model of Section 2.5.2.1, i.e., the transmission of the information produced by μ[1] correlated sources over an AWGN multiple access channel. In this case, all users generate blocks of the same length and encode the information at the same rate. In other words, each transmitter l sends a block of K information bits $\mathbf{u}^l = (u_1^l, u_2^l, \ldots, u_K^l)^\top \in \{0,1\}^{K\times 1}$ to the same destination by simultaneously using an additive Gaussian noise channel N times. The total transmitted average energy per MAC channel use is $E_s = \mu \cdot E_l$, where E_l is the average energy per symbol of each user, and is equal for all users.

As mentioned before, the theoretical limit when the energy constraint is defined at the transmitter is not known. As it is done in the literature [139], for comparison purposes we will consider the theoretical limit assuming that separation between source and channel coding holds, which is only valid when the correlation is equal to zero. This would be the theoretical limit if the correlated sources are first compressed up to the Slepian-Wolf limit and then encoded with capacity achieving codes. If the transmission rate of each user is R, then the theoretical limit (assuming source-channel

[1] We denote the number of correlated sources by μ instead of λ, which has been used previously for uncorrelated sources.

separation) is given by

$$R \cdot H(\mathbf{U}^1, \mathbf{U}^2, \ldots, \mathbf{U}^\mu) \leq \log_2 \left(1 + \frac{E_s}{N_0}\right). \tag{7.2}$$

Finally, the total system rate R_s is given by $R_s = R \cdot H(\mathbf{U}^1, \mathbf{U}^2, \ldots, \mathbf{U}^\mu)$. We will also refer to $\frac{E_s}{N_0}$ as the SNR in this chapter.

7.2.1 SOURCE MODEL

We assume that the information bit blocks $\mathbf{u}^l = (u_1^l, u_2^l, \ldots, u_K^l)^\top \in \{0,1\}^{K\times 1}$, for $l = 1, \ldots, \mu$, are generated by binary uniform memoryless sources. As indicated in Fig. 7.1, the correlation between the μ sources is modeled in the following way.

- Generate a symmetric i.i.d. base sequence $\mathbf{s}$ of length K with $P(s_k = 0) = P(s_k = 1) = \frac{1}{2}$.
- The source information sequences $\mathbf{u}^l$, for $l = 1, \ldots, \mu$, are obtained as $\mathbf{u}^l = \mathbf{s} \oplus \mathbf{e}^l$, where $\oplus$ indicate modulus 2 addition. The sequences $\mathbf{e}^l$, of length K, are obtained from a non-uniform i.i.d. binary r.v. that takes value 1 with probability $p \in [0,1]$.

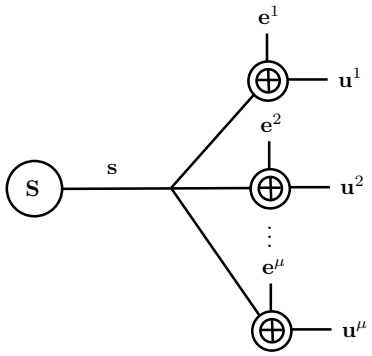


Figure 7.1: *The multi-terminal spatially correlated source model considered.*

Notice that the considered correlation model describes an i.i.d. multi-terminal source (refer to Section 2.5.1). At each time k, the stochastic

process $\{\mathbf{U}_k\}$ is composed by the μ-dimensional random vector $\{\mathbf{U}_k\} = (U_k^1, \ldots, U_k^\mu)$ taking values in $\{0,1\}^{1\times\mu}$ with probability mass function $P_{\mathbf{U}}(u^1u^2\ldots u^\mu)$ given by

$$\begin{aligned} P_{\mathbf{U}}(u^1u^2\ldots u^\mu) =& \frac{1}{2}\cdot\prod_{i=1}^{\mu} p^{u^i}\cdot(1-p)^{(1-u^i)} \\ &+\frac{1}{2}\cdot\prod_{i=1}^{\mu} p^{(1-u^i)}\cdot(1-p)^{u^i}. \end{aligned} \tag{7.3}$$

Although we deal with this simple source model, the ideas proposed in this chapter can be applied to more complex source models.

7.3 THE PROPOSED RCM-CIS SCHEME AND ITS JOINT DECODER

In order to achieve good performance when correlated information sources are transmitted over a multiple access channel, the generated codewords should preserve part of the existing source correlation. The proposed scheme takes advantage of the sparse and linear nature of RCM codes to achieve this goal. Specifically, the introduced RCM-CIS scheme utilizes μ RCM encoders, each encoding one of the users. All the μ RCM encoders have the same graphical structure, but different weight multi-sets $\mathcal{W}^l$, $l = 1, \ldots, \mu$. By same graphical structure we mean that for all l, every symbol x_j^l is connected to the same $2\cdot d_r$ source symbols (each of its own source). This property can be defined from the encoding matrix perspective as follows. If we denote as $\{G^1, \ldots, G^\mu\}$ the encoding RCM matrices for each user, then the non-zero entries of these sparse matrices are all placed in the same row-column positions. This is achieved by utilizing a mother RCM structure, which is replicated in the generation of individual RCMs composing the RCM-CIS scheme.

Notice that in the proposed scheme, the MAC output symbol x_j^C is generated from linear combinations of $2\cdot d_r$ groups of μ correlated source symbols. Specifically, if for user l the input binary symbol in position k, u_k^l, participates in x_j^C, the input bits in position k for all other users also participate. If we design the weights properly, a high degree of symbol reinforcement can be achieved taking advantage of the correlation. An example of such a design, which we will use in this chapter, would be to

jointly select the weights so that correlated source symbols are connected to symbol x_j^C with weights of equal sign.

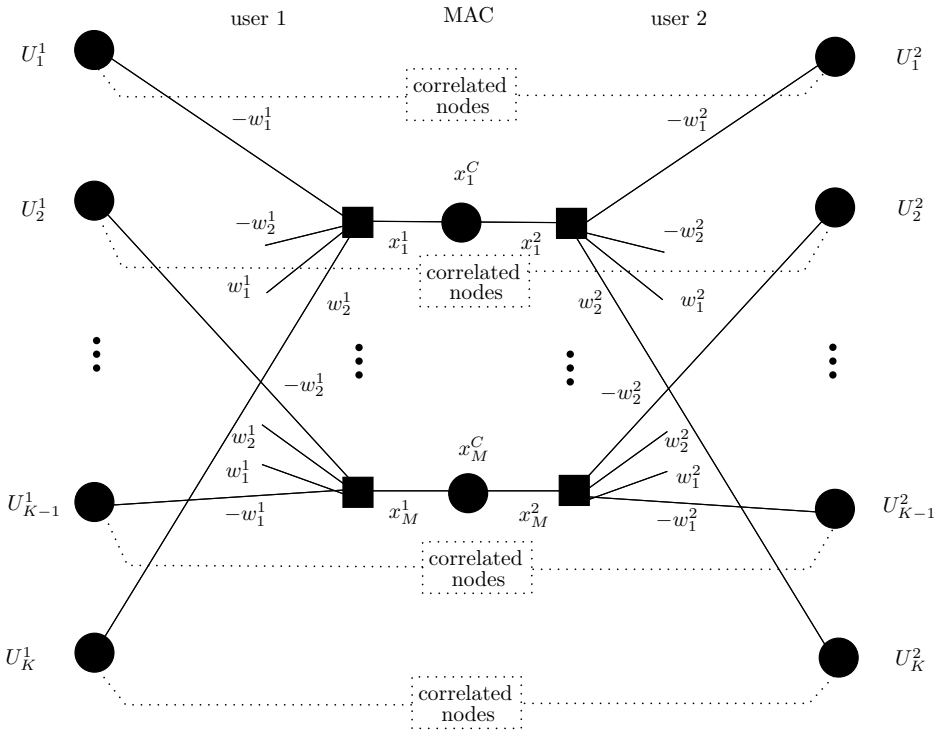


Figure 7.2: *Proposed RCM-CIS scheme with two correlated sources.*

Before detailing the construction of this scheme, let us consider an example with $\mu = 2$ users. As shown in Fig. 7.2, the factor graph of the entire RCM-CIS scheme consists of the individual RCM codes, the MAC and the correlation model. Notice that the graphical structures of the RCM encoding each user are identical. For example, the RCM symbol x_1^1 is connected to U_1^1 and U_K^1 with the weights $-w_1^1$ and w_2^1, whereas the RCM symbol x_1^2, of user 2, is connected to U_1^2 and U_K^2 with the weights $-w_1^2$ and w_2^2, respectively. This is repeated for every symbol. Therefore, the MAC symbol $x_1^C = x_1^1 + x_1^2$ is generated as

$$x_1^C = \underbrace{-w_1^1 \cdot u_1^1 - w_1^2 \cdot u_1^2}_{\text{correlated}} + \underbrace{w_2^1 \cdot u_K^1 + w_2^2 \cdot u_K^2}_{\text{correlated}} + \dots . \tag{7.4}$$

Notice that by RCM definition, the weight values w_j^l are positive. This will produce a symbol reinforcement at the receiver, which increases the average received energy with respect to the total transmitted energy E_c, which is the constraint of the overall communication system. Finally, note that the source symbol variables of both users are connected through the joint source's graph.

7.3.1 GROUPING THE INPUT SYMBOLS FOR ALL USERS

The conventional decoding approach for the example of Fig. 7.2 would be to perform the symbol-wise MAP by applying the sum-product algorithm to the entire factor graph. Note that this FG contains short length cycles at each node. For example, U_1^1 is connected through the source model to U_1^2, which at the same time is linked to x_1^2, x_1^2 to x_1^1, and finally, x_1^1 connects back to U_1^1. This is repeated for all symbols by construction, resulting in the presence of many short length cycles, which ultimately degrades the decoding performance [168, 169]. Moreover, as stated before, the binary decoding losses part of the correlation information by considering the source symbols individually, despite sum-product messages being passed through the source factor graph.

In this Thesis we propose to solve the short length cycle problem and to fully exploit the source correlation by rearranging the factor graph of Fig. 7.2 in such a way that we consider the joint source probabilities, eliminating the mentioned short length cycles. The idea of clustering or grouping nodes to cope with cycles in probabilistic graphs was first proposed by J. Pearl in [170]. In this work, we group the paired correlated source symbols together as non-binary compound variables $U_k^C = \{U_k^1, U_k^2\} \in \{0,1\}^2$ given by the Cartesian product of the individual variables. Note that this new non-binary input variable can take 4 possible values given by $\{00, 01, 10, 11\}$, each with a different probability that depends on the source correlation model. We also directly consider the output MAC symbol x_j^C. This rearranged factor graph is shown in Fig. 7.3. Note that its graphical structure is the same as that of the two previous individual RCMs encoding each user. The difference lies in the fact that variables in Fig. 7.3 are compound: the MAC output symbols, and the Cartesian product of the source symbols. The edges connecting these variables have two weight values. Notice that although we eliminate the short length cycles mentioned before, the resulting non-binary factor graph still contains the same cycles that occur in conventional RCM codes. However, the number of these cycles

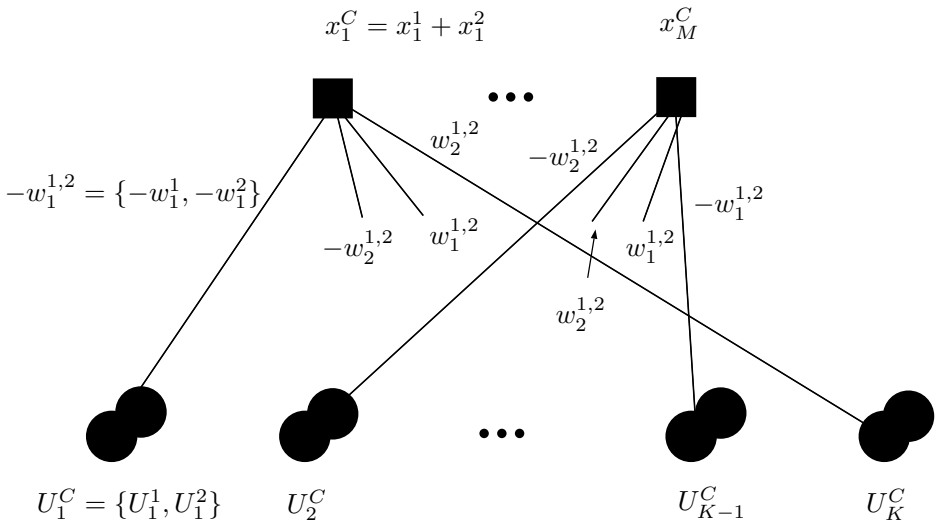


Figure 7.3: *Grouping the input symbols in Fig. 7.2. The notation* $w_j^{1,2}$ *refers to* $\{w_j^1, w_j^2\}$.

is far smaller than in the previous case, and similar to conventional RCM codes, will have very limited impact on the decoding performance.

Finally, we provide another perspective of the rearranged scheme, which will help understand the non-binary decoder used over the rearranged factor graph. Let us label the 4 possible values that the input compound variable U_1^C can take as $0, 1, 2, 3$ which correspond to the individual source symbol combinations $\{U_1^1, U_1^2\} = \{00, 01, 10, 11\}$. Following the example in Fig. 7.3, U_k^C is connected to x_1^C through $-w_1^{1,2} = \{-w_1^1, -w_1^2\}$. Thus, the contribution of U_1^C to x_1^C, given by $-w_1^1 \cdot u_1^1 - w_1^2 \cdot u_1^2$, can be seen as the non-binary mapping of Table 7.1. For example, if the two input symbols take value 1, then, U_1^C adds the factor $-w_1^1 - w_1^2$ to the total sum of x_1^C. The non-binary messages passed through the non-binary RCM decoder will contain the probabilities of the 4 possible values that U_1^C can take considering the symbol x_1^C.

Table 7.1: *Non-binary RCM mapping showing the contribution of u_1^1 and u_1^2 to x_1^C in Fig. 7.3.*

$\{U_1^1, U_1^2\}$	label (U_1^C)	$f(U_1^1, U_1^2)$
$\{0,0\}$	0	0
$\{0,1\}$	1	$-w_1^2$
$\{1,0\}$	2	$-w_1^1$
$\{1,1\}$	3	$-w_1^1 - w_1^2$

7.3.2 DESIGN OF RCM-CIS ENCODERS

The design process of an RCM-CIS scheme begins by building a $M \times K$ RCM random structure with $2 \cdot d_r$ entries $\{w_1, \ldots, w_{d_r}, -w_1, \ldots, -w_{d_r}\}$ as in Fig. 2.13. This mother structure is replicated by the RCM codes of all users, and its entries will be used as indexes for the weight set of each user. Recall that working on a factor graph or its incident matrix (G in this case) is equivalent.

The μ RCM encoders $\{G^1, G^2, \ldots, G^\mu\}$ of the RCM-CIS scheme are all obtained from this RCM mother structure as follows. The RCM encoder of user l, with weight multi-set $\mathcal{W}^l$, is obtained by replacing the index entries $\{w_i\}$ in the mother RCM structure (refer to Fig. 2.13) by its own weight values $\{w_i^l\}$ to form the encoding matrix G^l. The weights are replaced in the specific order of appearance in the multi-sets, which determine the difference between individual RCM encoders.

Therefore, the RCM-CIS scheme is defined by the μ weight multi-sets composed by d_r values in a specific order. The set of μ weight multi-sets $\{\mathcal{W}^1, \ldots, \mathcal{W}^\mu\}$ is jointly designed and can be rewritten as the following compound weight matrix $\pm\mathcal{W}^C$, of size $\mu \times 2 \cdot d_r$,

$$\pm\mathcal{W}^C = \begin{bmatrix} \pm\mathcal{W}^1 \\ \pm\mathcal{W}^2 \\ \vdots \\ \pm\mathcal{W}^\mu \end{bmatrix} = \begin{bmatrix} w_1^1 \; w_2^1 \ldots w_{d_r}^1 - w_1^1 - w_2^1 \ldots - w_{d_r}^1 \\ w_1^2 \; w_2^2 \ldots w_{d_r}^2 - w_1^2 - w_2^2 \ldots - w_{d_r}^2 \\ \vdots \\ w_1^\mu \; w_2^\mu \ldots w_{d_r}^\mu - w_1^\mu - w_2^\mu \ldots - w_{d_r}^\mu \end{bmatrix}, \quad (7.5)$$

where row l represents the weights that user l is going to substitute from the mother RCM structure, $\{w_i\} \rightarrow \{w_i^l\}$.

After the symbols are superimposed with the channel, a output MAC symbol, $x_j^C = x_j^1 + x_j^2 + \ldots + x_j^\mu$ is a linear combination of $2 \cdot d_r \cdot \mu$ binary input symbols, weighted by all values in matrix $\pm\mathcal{W}^C$. The weights in the same row connect x_j^C to source symbols of the same user , whereas the weights in the same column connect x_j^C to μ correlated source symbols, one of each user. Since the weight values $\{w_i^l\}$ are positive, a symbol reinforcement will be achieved by this joint design.

The weight set matrix $\pm\mathcal{W}^C$ should be such that the resulting random symbol X^C at the output of the MAC has entropy larger than the total system rate R_s. That is, if we denote by $\Phi^C = \{\phi_1^C, \phi_2^C, \ldots\}$ the constellation points at the output of the MAC and $P_{X^C}(\phi_i^C)$ their probabilities of occurrence, then[2]

$$H(X^C) = -\sum_i P_{X^C}(\phi_i^C) \log P_{X^C}(\phi_i^C) \geq R_s. \tag{7.6}$$

In order to guarantee that users transmit with equal mean energy per symbol, in this chapter we will consider that every row of the matrix $\pm\mathcal{W}^C$ is constructed with the same weights but in different order, i.e., the weight multi-sets of each users are permutations of each other.

Finally, let us consider the non-binary rearrangement of the constructed RCM-CIS scheme with μ correlated users. The compound variables $U_k^C \in \{0,1\}^\mu$ are the Cartesian product of the μ source variables $\{U_k^1, \ldots, U_k^\mu\}$ that we will label from 0 to $2^\mu - 1$ according to their decimal representation (as in Table 7.1). Recall that U_k^C will contribute (if connected through the FG) to the MAC output symbol x_j^C through the weighted sums of the weights in one column of $\pm\mathcal{W}^C$. These weights are multiplied by the individual source symbols.

7.3.3 PARALLEL CONCATENATION WITH LDGM CODES

In order to correct the residual errors produced by the RCM-CIS scheme, we propose the use of LDGM codes in parallel with the RCM codes used for each user (see Fig. 7.4), as it is done for point-to-point communications. The idea is to correct the possible residual errors not corrected by the RCM codes. The μ LDGM generator matrices $\{P^1, P^2, \ldots, P^\mu\}$, of

[2]It should be noted that the probability mass function $P_{X^C}(\phi^C)$ is function of the weight matrix $\mathcal{W}^C$ and the source correlation.

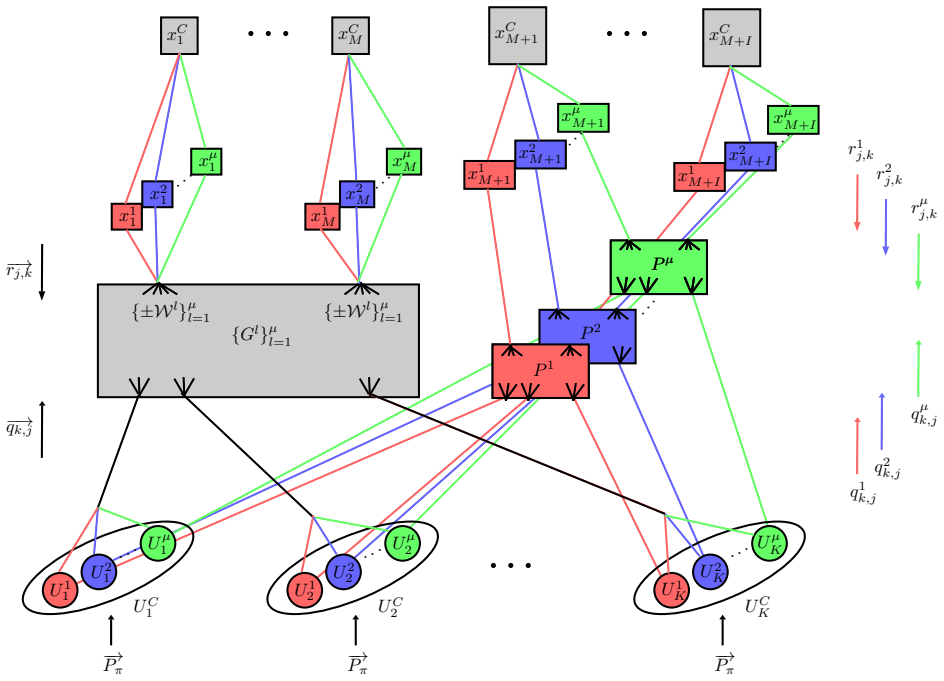


Figure 7.4: *Factor graph of the proposed RCM-CIS scheme with a parallel LDGM concatenation for sending μ correlated information sources.*

size $I \times K$, are all obtained from random row-wise permutations of the non-systematic part P of the same LDGM generator matrix, that produces I non-systematic LDGM symbols with design parameter d_l. These generator matrices will produce LDGM coded bits, which are then BPSK (± 1) modulated and also sent through the MAC. At the receiver, if the superimposed LDGM MAC symbols present mutual reinforcement (are all 1 or -1) it is easy to estimate the LDGM symbols of each user. However, since the source correlation is lost in these LDGM symbols (they are produced by different LDGM encoding matrices), many received symbols are formed by the addition of non-equal LDGM symbols, and therefore, they could be produced by different combinations of LDGM symbols. This causes an ambiguity that may render estimation errors in the individual LDGM symbols, complicating the decoding process.

With the goal of improving the decoding task, we propose to multiply the LDGM symbols by modulation factors before they are transmitted through the MAC. These modulation factors take μ fixed values, denoted as $\Theta = \{\theta_1, \theta_2, \ldots, \theta_\mu\}$, which are chosen to reduce the ambiguity of the LDGM MAC symbols. In order to allocate the mean energy and the symbol protection evenly across all users, these fixed modulation values are each assigned to one user for every LDGM symbols in a rotating manner. If we denote by α_j^l the modulation factor of user l and LDGM coded symbol j, then $\alpha_j^l = \theta_\upsilon$, where $\upsilon = (l + j \mod \mu) + 1$ and mod is the modulus. For example, for $\mu = 2$ and fixed values $\{1, 2\}$, the modulation factors would be $\alpha_j^1 = 1$ and $\alpha_j^2 = 2$ for j even and $\alpha_j^1 = 2$ and $\alpha_j^2 = 1$ for j odd. Consequently, the LDGM MAC output symbol x_j^C, for $j = M + 1, \ldots, M + I$, is given by

$$x_j^C = \sum_{l=1}^{\mu} \alpha_j^l \cdot x_j^l, \tag{7.7}$$

where $x_j^l \in \{-1, 1\}$ is the LDGM coded BPSK modulated symbol j of user l generated with the LDGM encoding matrix P^l.

To sum up, the encoder $\mathcal{E}_l$ of rate $R = \frac{K}{M+I}$, for $l = 1, \ldots, \mu$, is composed by an RCM encoder matrix G^l, of size $M \times K$, and an LDGM encoder matrix P^l, of size $I \times K$. The RCM encoder generates the output symbols $j \in [1 : M]$, while the LDGM encoder generates coded symbols $j \in [M + 1 : M + I]$. Finally, the LDGM BPSK modulated symbol j of user l is multiplied by the modulation factor α_j^l before being transmitted. Note that if $N = M$ $(I = 0)$, the proposed scheme is just an RCM-CIS.

7.3.4 JOINT DECODER

As we showed in the previous section, the RCM sub-graph is rearranged in a non-binary manner in order to avoid the creation of short length cycles, and non-binary decoding is applied by considering the correlated source binary symbols of all users jointly. However, the decoding of the non-linear LDGM sub-graph remains binary. Therefore, the resulting decoder is based on a mixed message passing schedule that uses non-binary and binary Log-Likelihood Ratio (LLR) messages. The non-binary probabilities of U_k^C and the binary probabilities of $\{U_k^1, \ldots, U_k^\mu\}$ are converted to each other by discretization or joining the corresponding PMFs. For the sake of clarity in

the explanation of the joint decoder, let us start by defining the following terms that are used in the iterative decoder:

- Variable Nodes (VN) U_k^l: We refer to the individual variable node corresponding to input binary symbol k of the user l as $\{U_k^l\}$.
- Compound VN U_k^C: We define the k^{th} compound VN as the Cartesian product of the μ VNs of the source symbols k given by $U_k^C = \{U_k^1, U_k^2, \ldots, U_k^\mu\}$. These compound VNs are non-binary variables that can take 2^μ possible values, which we will label from 0 to $2^\mu - 1$ according to the equivalent decimal number in the binary representation defined by $\{U_k^1, U_k^2, \ldots, U_k^{\mu-1}, U_k^\mu\}$ as in Table 7.2.

Table 7.2: *Labeling of the compound VNs U_k^C.*

label	U_k^1	U_k^2		$U_k^{\mu-2}$	$U_k^{\mu-1}$	U_k^μ
0	0	0	...	0	0	0
1	0	0	...	0	0	1
2	0	0		0	1	0
⋮			⋮			⋮
$2^\mu - 2$	1	1	...	1	1	0
$2^\mu - 1$	1	1	...	1	1	1

- Based on the previous labeling, we denote by χ_i, $i = 0, \ldots, 2^\mu - 1$, the subset of users whose entry for label i in Table 7.2 is 1. For example, $\chi_0 = \{\}$, $\chi_1 = \{\mu\}$ and $\chi_{2^\mu - 1} = \{1 : \mu\}$.
- MAC Check Node (CN) T_j^C: We define the j^{th} MAC CN T_j^C as the check node representing the symbol x_j^C at the output of the MAC. $\{T_j^C\}_{j\in[1:M]}$ are RCM nodes, whereas $\{T_j^C\}_{j\in[M+1:M+I]}$ are LDGM nodes. For the later case, we also define the individual LDGM CNs as $\{T_j^l\}_{j\in[M+1:M+I]}$, $l = 1, 2, \ldots, \mu$, i.e., the CN j at the output of the LDGM encoder of user l, before considering the multiple access channel.
- We denote as $\overrightarrow{r_{j,k}}$ and $\overrightarrow{q_{k,j}}$ the vector messages passed from the j^{th} RCM MAC CN T_j^C to the k^{th} compound VN U_k^C and the vector messages passed from the k^{th} compound VN U_k^C to the j^{th} RCM MAC

CN T_j^C, respectively. These messages contain the vector, of length 2^μ, of probabilities associated to each value that the compound VN U_k^C can take, i.e., $\{P_{0\ldots00}, P_{0\ldots01}, P_{0\ldots10}, \ldots, P_{1\ldots11}\}$.

- For the LDGM sub-part, we denote by $r_{j,k}^l$ the LLR messages passed from the j^{th} individual LDGM CN T_j^l of user l to the k^{th} individual VN U_k^l of the same user. Similarly, $q_{k,j}^l$ represents the LLR messages passed from the k^{th} individual VN U_k^l to the j^{th} individual LDGM CN T_j^l.

- $n(U_k^C) \setminus T_j^C$ and $n(T_j^C) \setminus U_k^C$ denote the set of RCM MAC CNs connected to compound VN U_k^C without considering MAC CN T_j^C, and the set of compound VNs connected to RCM MAC CN T_j^C without considering U_k^C, respectively. For the LDGM sub-graphs, one per user, we denote by $n(U_k^l) \setminus T_j^l$ the set of individual LDGM CNs connected to VN U_k^l without considering the LDGM CN T_j^l and by $n(T_j^l) \setminus U_k^l$ the set of VNs connected to LDGM CN T_j^l without considering U_k^l.

- Finally, we denote by $\overrightarrow{P_\pi}$ the a priori probability vector, of length 2^μ, of each compound VN based on the correlation model. For example, for $\mu = 2$ and the considered source model, $\overrightarrow{P_\pi} = \{1/2 \cdot (1-p)^2 + 1/2 \cdot p^2, p \cdot (1-p), p \cdot (1-p), 1/2 \cdot (1-p)^2 + 1/2 \cdot p^2\}$.

At iteration $t \in \{1, \ldots, t_{max}\}$, the iterative decoder operates by transmitting the following messages:

- **STEP 1.1 $\overrightarrow{q_{k,j}}$: Probability vector message passing from compound VNs, $\{U_k^C\}_{k=1}^K$, to RCM MAC check nodes $\{T_j^C\}_{j=1}^M$.**

$$\overrightarrow{q_{k,j}} = \overrightarrow{P_\pi} \cdot \overrightarrow{P_{\mathrm{LDGM}}(k)} \cdot \prod_{i \in n(U_k^C) \setminus T_j^C} \overrightarrow{r_{i,k}}, \tag{7.8}$$

where $\overrightarrow{r_{i,k}} = \{\frac{1}{2^\mu}, \frac{1}{2^\mu}, \ldots, \frac{1}{2^\mu}\}$ for $k \in \{1, \ldots, K\}$, $i \in n(U_k^C) \setminus T_j^C$ in iteration 1. $\overrightarrow{P_{\mathrm{LDGM}}(k)}$ is the probability message vector, representing $\{P_{0\ldots00}, P_{0\ldots01}, P_{0\ldots10}, \ldots, P_{1\ldots11}\}$, of the compound VN U_k^C, generated from the binary LDGM beliefs $\{r_{i,k}^l\}_{i=M+1:M+I, l=1:\mu}$ that arrive on individual VNs $\{U_k^l\}_{l \in [1:\mu]}$.

- **STEP 1.2 $q^l_{k,j}$: LLR message passing from individual VNs $\{U^l_k\}_{l\in[1:\mu],k\in[1:K]}$ to individual LDGM CNs $\{T^l_j\}_{l\in[1:\mu],j\in[M+1:M+I]}$.**

$$q^l_{k,j} = P^{(l)}_{\mathrm{RCM}}(k) + \sum_{i\in n(U^l_k)\setminus T^l_j} r^l_{i,k}, \tag{7.9}$$

where $r^l_{i,k} = 0$ for $k \in \{1,\ldots,K\}$, $i \in n(U^l_k) \setminus T^l_j$ in the iteration 1. $P^{(l)}_{\mathrm{RCM}}(k)$ is the LLR belief of individual VN U^l_k obtained by marginalizing the RCM probability vector $\prod_{i\in n(U^C_k)} \overrightarrow{r_{i,k}}$.

- **STEP 2.1 $\overrightarrow{r_{j,k}}$: Probability vector message passing from RCM MAC check nodes, $\{T^C_j\}^M_{j=1}$, to compound VNs $\{U^C_k\}^K_{k=1}$.**

Note that the j^{th} RCM MAC output symbol x^C_j is given by $x^C_j = \left(\sum_{l=1}^{\mu}\left(\sum_{i=1}^{K} g^l_{j,i}u^l_i\right)\right)$ and let us define $a_{j,k} = \left(\sum_{l=1}^{\mu}\left(\sum_{i\sim k} g^l_{j,i}u^l_i\right)\right)$, where $\sum_{i\sim k}$ refers to the sum over all $i \in [1:K]$ except k and $g^l_{j,i}$ is the entry (j,i) of the encoding matrix G^l.

By rewriting the j^{th} RCM MAC output symbol as $a_{j,k} + \sum_l g^l_{j,k}u^l_k$ for all $k \in n(T^C_j)$, the 2^μ probability entries $\{\overrightarrow{r_{j,k}}(i)\}^{2^\mu-1}_{i=0}$ of the message vector $\overrightarrow{r_{j,k}}$ are calculated as

$$\overrightarrow{r_{j,k}}(i) = \sum_{z\in\Phi^C} P(a_{j,k} = z)\cdot e^{-\frac{(y_j - z - \sum_{l\in\chi_i} g^l_{j,k})^2}{N_0}}, \tag{7.10}$$

where χ_i is the subset of users with 1's for label i in Table 7.2 previously defined, and the sum is over the values that the RCM MAC output symbol can take, i.e., over Φ^C. $P(a_{j,k} = z)$, the probability of $a_{j,k} = z$ at the given iteration, is calculated by convolving the PMFs of the terms in the summation. Notice that $\sum_{l\in\chi_i} g^l_{j,k}$ directly gives the output of the non-binary mapping of Table 7.1, and each entry i of vector $\{\overrightarrow{r_{j,k}}(i)\}$ is the estimated probability of the non-binary variable U^C_k taking the value in labeled as i in Table 7.2.

- **STEP 2.2 $r^l_{j,k}$: LLR message passing from individual LDGM check nodes $\{T^l_j\}_{l\in[1:\mu],j\in[M+1:M+I]}$ to individual variable nodes $\{U^l_k\}_{l\in[1:\mu],k\in[1:K]}$.**

The LLR message transmitted from the j^{th} individual LDGM CN of user l, T_j^l, to the individual VN k of user l, U_k^l, is calculated by means of the standard formula

$$r_{j,k}^l = -2\text{atanh}\left(\tanh\left(-\frac{\gamma_j^l}{2}\right)\prod_{i\in n(T_j^l)\setminus U_k^l}\tanh\left(-\frac{q_{k,i}^l}{2}\right)\right), \tag{7.11}$$

with the difference that the channel LLR value γ_j^l of the individual CN T_j^l, is not obtained directly from the channel but, at every iteration, it has to be calculated from the received MAC symbol y_j and the estimates of the rest of individual LDGM symbols $\{x_j^i\}_{i\in[1:\mu]\sim l}$, as shown in Fig. 7.5.

We start by calculating the probability of each individual LDGM coded symbol x_j^l, which can take the values α_j^l or $-\alpha_j^l$, from the incoming LLR messages $\{q_{k,i}^l\}_{i\in n(T_j^l)}$ as

$$P(x_j^l = \alpha_j^l) = \frac{e^{\beta_j^l}}{1+e^{\beta_j^l}}, \tag{7.12}$$

where

$$\beta_j^l = -2\text{atanh}\left(\prod_{i\in n(T_j^l)}\tanh\left(-\frac{q_{k,i}^l}{2}\right)\right). \tag{7.13}$$

Once the probabilities $\{P(x_j^i)\}_{i\in[1:\mu]\sim l}$ are calculated, and taking into account the LDGM modulation factors $\{\alpha_j^1, \alpha_j^2, \ldots, \alpha_j^\mu\}$ for the users $\{1, 2, \ldots, \mu\}$, the LLR value γ_j^l is obtained by executing an iteration of the sum-product algorithm over the factor graph depicted in Fig. 7.5. Utilizing a similar strategy as for the RCM, let us rewrite x_j^C as $x_j^C = \sum_i x_j^i \cdot \alpha_j^i = \sum_{i\sim l} x_j^i \cdot \alpha_j^i + x_j^l \cdot \alpha_j^l = a_j^l + x_j^l \cdot \alpha_j^l$, where $\sum_{i\sim l} x_j^i \cdot \alpha_j^i = a_j^l$. Since $x_j^l \in \pm 1$ and denoting the LDGM MAC output symbol alphabet as $\Psi^C = \{\psi_1^C, \psi_2^C, \ldots\}$, the channel LLR value γ_j^l is obtained as

$$\gamma_j^l = \log\left(\frac{\sum_{v\in\Psi^C}\text{P}(a_j^l = v)\cdot e^{-\frac{(y_j - v - \alpha_j^l)^2}{N_0}}}{\sum_{v\in\Psi^C}\text{P}(a_j^l = v)\cdot e^{-\frac{(y_j - v + \alpha_j^l)^2}{N_0}}}\right), \tag{7.14}$$

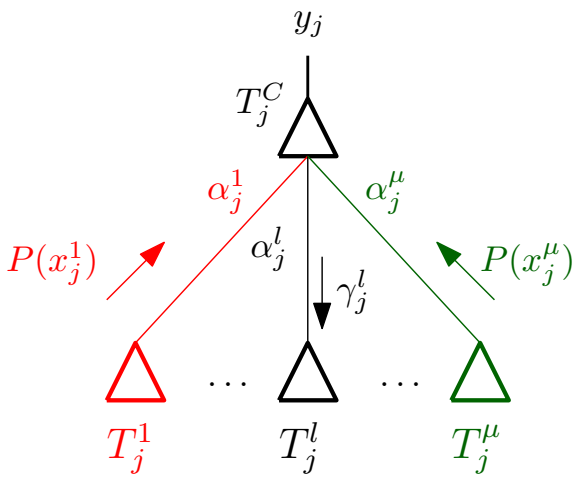


Figure 7.5: *Factor graph of the MAC LDGM check node j with modulation factors $\{\alpha_j^1, \alpha_j^2, \ldots, \alpha_j^\mu\}$*

where the sum is over the values in Ψ^C and $P(a_j^l = v)$ is obtained by convolving the PMFs of the terms in the summation.

- **STEP 3. Decision.**

  At the end of the iterations, the decision probability vector $\overrightarrow{d_k}$, $\{P_{0\ldots00}, P_{0\ldots01}, P_{0\ldots10}, \ldots, P_{1\ldots11}\}$, is calculated as

$$\overrightarrow{d_k} = \overrightarrow{P_\pi} \cdot \overrightarrow{P_{\mathrm{LDGM}}(k)} \cdot \prod_{j \in n(U_k)} \overrightarrow{r_{j,k}}, \tag{7.15}$$

  and the estimates of u_k^l, for $l = 1 \ldots, \mu$, are obtained by marginalizing the decision vector $\overrightarrow{d_k}$.

7.3.5 COMPUTATIONAL COMPLEXITY

In this section we provide a detailed computational analysis of the proposed non-binary decoder. We will focus on the RCM sub-scheme because (i) many schemes will only contain RCM nodes, i.e., when $I = 0$, (ii) even when a parallel LDGM code is used, the LDGM symbols will only represent a small proportion of the total coded sequence, i.e., $M >> I$ and (iii)

the computational complexity of Step 2.2 (LDGM check nodes) is far lower than that of Step 2.1 (RCM check nodes). Thus, the decoder complexity is dominated by the RCM part. Let us start by breaking down one decoding iteration as follows.

- **Step 1.1.** The step 1.1 is performed over K (compound) VNs with an average of $\frac{2d_r \cdot M}{K}$ connections.
  - Each vector $\overrightarrow{r_{i,k}}$ has length 2^μ, yielding to $2d_r \cdot M \cdot 2^\mu$ products.
- **Step 2.1.** The step 2.1 is performed over M (compound) CNs each with $2d_r$ connections. This step calculates the vector probability messages $\overrightarrow{r_{j,k}}$, which are composed by 2^μ elements calculated as (7.10).
  - Each equation (7.10) involves the sum over $z \in \Phi^C$. The cardinality of Φ^C, $|\Phi^C|$, is upper bounded by $2\sum_{l=1}^{\mu}(\sum_{i=1}^{d_r} w_i^l)$ due to its integer nature.
    - Each sum in (7.10) contains 2 products, 1 division, 1 exponential function, and 2 sums. Note that the sums in χ_i can be stored in a table. Therefore, the total number of products, divisions, exponential and sums is given by $\{2,1,1,2\} \cdot (2d_r \cdot M \cdot 2^\mu \cdot \left(2\sum_{l=1}^{\mu}(\sum_{i=1}^{d_r} w_i^l)\right))$, respectively.
  - Calculation of $P(a_{j,k} = z)$ for $z \in \Phi^C$. The PMF of $P(a_{j,k} = z)$ is a vector of at most $2\sum_{l=1}^{\mu}(\sum_{i=1}^{d_r} w_i^l)$ elements and is calculated recursively, with $(2d_r - 1)$ iterations of 2^μ-ary PMF (obtained from the vectors $\overrightarrow{q_{k,j}}$).
    - Each convolution is performed over the accumulated $P(a_{j,k} = z)$ of length $2 \cdot \sum_{l=1}^{\mu}(\sum_{i=1}^{d_r} w_i^l)$ and a 2^μ-ary PMF by shift addition, with $2^\mu \cdot (\sum_{l=1}^{\mu}(\sum_{i=1}^{d_r} w_i^l))$ products and sums. Thus, the total number of operations involved in calculating the $P(a_{j,k} = z)$ is upper bounded by $2d_r \cdot M \cdot 2^\mu \cdot (2d_r - 1) \cdot 2^\mu \cdot 2 \cdot \sum_{l=1}^{\mu}(\sum_{i=1}^{d_r} w_i^l)$ sums and products.

Finally, an approximation of the (worst case) computational complexity per sum-product iteration is presented in Table 7.3. As it can be observed, the complexity is exponential with μ, and it is mainly determined by the convolution operations involved in the calculation of $P(a_{j,k} = z)$.

Table 7.3: *Computational complexity of the decoder (per sum-product iteration).*

	Operations
Products	$\approx 8d_r^2 \cdot M \cdot \left(\sum_{l=1}^{\mu}(\sum_{i=1}^{d_r} w_i^l)\right) \cdot 4^{\mu}$
Sums	$\approx 8d_r^2 \cdot M \cdot \left(\sum_{l=1}^{\mu}(\sum_{i=1}^{d_r} w_i^l)\right) \cdot 4^{\mu}$
Divisions	$\approx 4d_r \cdot M \cdot \left(\sum_{l=1}^{\mu}(\sum_{i=1}^{d_r} w_i^l)\right) \cdot 2^{\mu}$
Exp	$\approx 4d_r \cdot M \cdot \left(\sum_{l=1}^{\mu}(\sum_{i=1}^{d_r} w_i^l)\right) \cdot 2^{\mu}$

Note that calculations in LDGM check nodes also involve performing convolutions. However, in this case $\mathrm{P}(a_j^l = v)$ is calculated recursively with $(\mu - 1)$ iterations of binary PMFs coming from (7.12), yielding $\approx 4 \cdot I \cdot \mu^2 \cdot \sum_{i=1}^{\mu} \theta_i$ sums and products. Finally, the computational complexity of the LDGM check nodes is dominated by (7.11), which has approximately $\frac{\mu \cdot K^2 \cdot d_l^2}{I}$ tanh and product operations. Notice that computations in the LDGM sub-scheme are not exponential with the number of users.

7.4 EXTENSION TO CLUSTERS OF CORRELATED INFORMATION SOURCES BASED ON RCM-MAC SCHEMES

In Chapter 6 we have proposed the RCM-MAC scheme for transmitting λ independent users over a multiple access channel at high sum-rates while maintaining low decoding complexity. In this section we utilize this same idea of converting the MAC scheme into a point-to-point RCM code from the receiver's perspective to the correlated scenario of this chapter.

7.4.1 RCM-CIS SCHEMES FOR CLUSTERS OF CORRELATED INFORMATION SOURCES

RCM-CIS schemes can be generalized to the transmission of λ independent clusters of μ correlated information sources by substituting the point-to-point RCM mother matrix of Fig. 2.13, from which the individual RCM codes of the RCM-CIS scheme are generated, by the set of λ irregular RCM structures of an RCM-MAC scheme. The users of each cluster use each one of these irregular matrices to generate their own RCM matrices.

Note that each row of these irregular matrices will contain only a subset of the index entries $\{w_1, w_2, \ldots, w_{d_r}, -w_1, -w_2, \ldots, -w_{d_r}\}$, which will be replaced by the equivalent subset of weight sets of each user. Due to the way in which these matrices are constructed, the receiver can jointly decode $\mu \cdot \lambda$ users seamlessly, by applying the proposed RCM-CIS decoder.

7.4.2 PARALLEL LDGM SUB-SCHEME FOR CLUSTERS OF CORRELATED SOURCES

Since the parallel LDGM concatenation was not considered in Chapter 6, we begin this section by considering a MAC channel for the transmission of λ independent users. A parallel LDGM-MAC coding sub-scheme for this scenario can be built by distributing the LDGM symbols between the λ users. This is different to the RCM-MAC case, where the distribution was performed in linear weighted connections due to the additive nature of RCM symbols. Since the input block of length K is divided into smaller input blocks of length $\{K_1 K_2 \ldots K_\lambda\}$, such that $K = \sum_l K_l$, the LDGM coded sequence of length I can also be divided into λ LDGM sub-sequences of lengths $I_l = I \cdot \frac{K_l}{K}$, and assigned to its respective user of the MAC as follows

$$\underbrace{[\ \underbrace{x_1^{(1)}, \ldots, x_{I_1}^{(1)}}_{\text{LDGM seq. of user 1}}\ ,\ \ldots\ ,\ \underbrace{x_1^{(\lambda)}, \ldots, x_{I_\lambda}^{(\lambda)}}_{\text{LDGM seq. of user}\lambda}\ ]}_{\text{MAC LDGM sequence (of length I)}}. \tag{7.16}$$

Each user l generates its LDGM sub-sequences with the encoding matrix $P^{(l)}$ of rate $\frac{K_l}{I_l}$, which is equal to the original rate $\frac{K}{I}$. The LDGM-MAC sub-scheme can still be serially concatenated due to the sub-sequence division (sub-sequences of different users are orthogonal, i.e. do not collide in the MAC). As in the RCM-MAC scheme, the parallel LDGM-MAC sub-scheme can be embedded into an LDGM-like block diagonal matrix P^d, of size $I \times K$, as

$$P^d = \begin{bmatrix} P^{(1)} & & \cdots & \\ & P^{(2)} & \cdots & \vdots \\ \vdots & & \ddots & \\ & \cdots & & P^{(\lambda)} \end{bmatrix}_{I \times K},$$

whose diagonal block elements are the matrices $\{P^{(l)}\}_{l=1:\lambda}$ that produce the LDGM sub-sequences. This LDGM-like amalgamation can be considered as

a standard LDGM matrix encoding the concatenated λ source input blocks ($[\mathbf{u}^{(1)}\mathbf{u}^{(2)}\dots\mathbf{u}^{(\lambda)}]$) at the receiver, as has been done for the RCM-MAC scheme.

Finally, the transmission of λ clusters of μ correlated information sources is performed by assigning one sub-sequence to each cluster and generating its LDGM matrix as permutation of the LDGM encoding matrix generating the LDGM sub-sequence associated to its cluster. Note that RCM symbols from the $\mu \cdot \lambda$ users collide in the MAC, whereas only the LDGM symbols of the μ users inside each cluster collide. This yields to a suboptimal use of MAC resources by the users in the clusters. However, it only affects the LDGM subsequences, which represent a small percentage of the total coded sequence.

7.5 NUMERICAL RESULTS

In this section, we illustrate the performance of the proposed schemes in three high throughput scenarios, with $R = 4$, 6 and 9 bits per complex channel symbol, and high amount of correlation between sources (low values of p). The complexity of the non-binary decoder of Section 7.3.4 grows exponentially with the number of correlated users because bits from all of them are considered jointly. Thus, in this dissertation we present numerical results limited to μ =2, 3 and 4 correlated users. Specifically, the rate and correlation combinations considered are summarized in Table 7.4, where three different correlation values, $p = .1$, $p = .01$ and $p = .001$, are considered. It is worth noticing that the total system throughput R_s of the examples is between 6.12 bits per complex channel symbol for the case of 2 users, $R = 6$ and $p = .001$, and 11.93 bits per complex channel symbol for 4 users, $R = 9$ and $p = .01$. For $p = .01$, we have considered rate $R = 4$ because for the cases of $R = 6$ and 9 the total system rate, R_s, is unattainably large.

For the rest of the results, we have fixed the codeword length of each user to $N = 5000$, whereas the source block length K is scaled according to R as $K = \frac{R \cdot N}{2}$. To validate the performance of the proposed schemes, in the sequel we present numerical results obtained by Monte Carlo simulations.

Table 7.4: *Considered rate and correlation combinations in this chapter. The SNR limits (assuming source-channel separation) are in dB.*

$\mu = 2$				
R	p	$H(\mathbf{U}^1, \mathbf{U}^2)$	R_s	SNR Limit (dB)
4	.1	1.68	6.71	20.17
6	.01	1.14	6.83	20.5
6	.001	1.02	6.12	18.36
9	.01	1.14	10.25	30.86
9	.001	1.02	9.19	27.65
$\mu = 3$				
R	p	$H(\mathbf{U}^1, \mathbf{U}^2, \mathbf{U}^3)$	R_s	SNR Limit (dB)
4	.1	2.26	9.06	27.26
6	.01	1.24	7.45	22.4
6	.001	1.033	6.2	18.6
9	.01	1.24	11.16	33.6
9	.001	1.033	9.29	27.95
$\mu = 4$				
R	p	$H(\mathbf{U}^1, \mathbf{U}^2, \mathbf{U}^3, \mathbf{U}^4)$	R_s	SNR Limit (dB)
4	.1	2.8	11.17	33.62
6	.01	1.33	7.95	23.91
6	.001	1.044	6.27	18.8
9	.01	1.33	11.93	35.92
9	.001	1.044	9.37	28.2

7.5.1 NUMERICAL RESULTS FOR RCM-CIS SCHEMES

We begin by analyzing the performance of the proposed RCM-CIS scheme without considering the parallel LDGM concatenation, i.e., $M = N$ and $I = 0$. Fig. 7.6 presents the numerical results for the examples with $\mu = 2$ of Table 7.4. The figure depicts the BER performance versus the Gap (in dBs) to the theoretical limit assuming separation, also shown in Table 7.4. Note that each curve has a different limit despite being plotted together. The design weight matrices we have employed for theses simulations are $\mathcal{W}^C$ =[4, 7, 9; 9, 4, 7] for $R = 4$ and $p = .1$, $\mathcal{W}^C$ =[3, 4, 5, 8; 5, 8, 3, 4] for $R = 6$ and $p = .01, .001$ and $\mathcal{W}^C$ =[3, 4, 5, 5, 7; 5, 7, 4, 3, 5] for $R = 9$ when $p = .01$ and $p = .001$. These weight matrices generate MAC symbols with entropy larger than the system rate R_s (see (7.6)), and have been obtained based on trial and error. For BER $\approx 10^{-3}$, the gap to the theoretical limit assuming separation ranges from 0.6 dB for the case with $R = 9$ and $p = .01$ to -0.4 dB for $R = 6$ and $p = .001$. As mentioned in Section 7.2, source-channel separation is sub-optimal when transmitting correlated information sources and therefore, we are able to outperform the separation limit, especially for the larger correlation values (lower values of p).

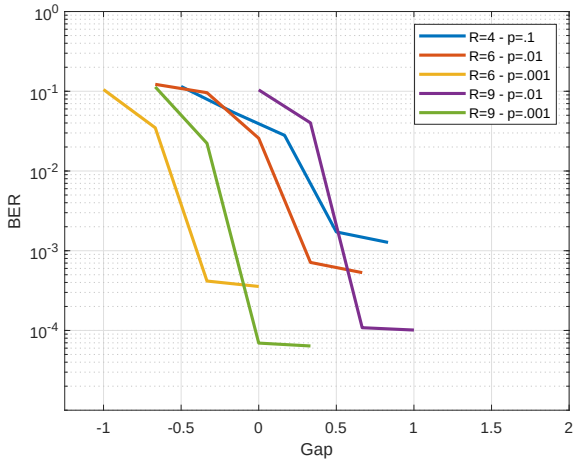


Figure 7.6: *Numerical results obtained by Monte Carlo simulation for the RCM-CIS scheme with $\mu = 2$ of Table 7.4.*

The results for the case of $\mu = 3$ are shown in Fig. 7.7. The utilized design weight matrices are $\mathcal{W}^C$ =[4, 7, 9; 9, 7, 4; 9, 4, 7] for $R = 4$ and $p = .1$, $\mathcal{W}^C$ =[3, 4, 5, 6, 8; 8, 3, 4, 5, 6; 6, 8, 3, 4, 5] for $R = 6$ when $p = .01$ and $p = .001$ and $\mathcal{W}^C$ =[3, 4, 5, 7, 10; 5, 10, 3, 4, 7; 10, 7, 4, 3, 5] for both $p = .01$ and .001 when $R = 9$. In this case, for BER $\approx 10^{-3}$, the gap to the theoretical limit assuming separation ranges from 0.1 to -2 dB.

Finally, the results for the case of $\mu = 4$ are presented in Fig. 7.8. The design weight matrices used in this last case are $\mathcal{W}^C$ =[4, 7, 9; 9, 7, 4; 9, 4, 7; 7, 9, 4] for $R = 4$ and $p = .1$, $\mathcal{W}^C$ =[2, 4, 6, 8, 10; 4, 2, 8, 10, 6; 4, 2, 10, 6, 8; 10, 8, 6, 4, 2] when $R = 6$ and when $R = 9$ with $p = .001$. Finally, the weight matrix for $R = 9$ and $p = .01$ is $\mathcal{W}^C$ =[2, 3, 4, 6, 8; 6, 8, 2, 3, 4; 3, 4, 6, 8, 2; 8, 2, 3, 4, 6]. For BER $\approx 10^{-3}$, the gap to the theoretical limit assuming separation ranges between 0.5 and -3.7 dB.

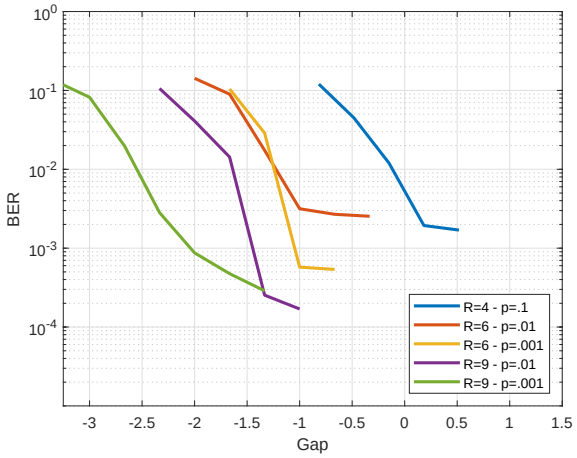


Figure 7.7: *Numerical results obtained by Monte Carlo simulation for the RCM-CIS scheme with $\mu = 3$ of Table 7.4.*

It is worth mentioning that the performance of the considered RCM-CIS examples improves as the number of correlated sources increases. Intuition suggests that the actual joint source-channel limit could be up to $10 \cdot \log_{10}(\mu)$ dB lower than the separation limit in the CIS communication system. This would occur for a fully correlated CIS system ($p = 0$) transmitting over a noiseless MAC, with the μ RCMs designed with the exact same weights. This would result in each RCM symbol being the same for all users and therefore, total symbol reinforcement would be achieved. Thus, the received

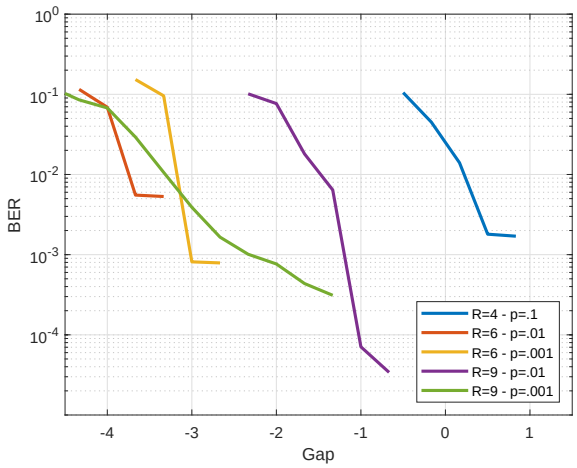


Figure 7.8: *Numerical results obtained by Monte Carlo simulation for the RCM-CIS scheme with $\mu = 4$ of Table 7.4.*

mean energy per symbol would be μ times the sum of the transmitted mean energies in that case. Actually, for correlation parameter values other than $p = 0$ this symbol reinforcement results in a received mean energy less than μ times the transmitted one. Moreover, the weights of the RCM schemes are designed to promote symbol reinforcement in the codewords, but the weights connecting RCM symbols to correlated bits are not identical. The reason is that if all the RCMs were identical, the decoder would be unable to differentiate between users when their source input bits are different. Note that the RCM codes that make up the RCM-CIS schemes considered in this chapter have the same weights, but the key is that they are placed in different order.

To sum up, the performance gap of the proposed RCM-CIS schemes with respect to the theoretical limits assuming source-channel separation is between -0.4 dB and .5 dB for $\mu = 2$, between -2 dB and .1 dB for $\mu = 3$, and between -3.7 dB and .5 dB for $\mu = 4$. These gaps are larger than the ones achieved with lower rate schemes based on powerful binary codes. For example, in [139] the separation limit is outperformed by up to 1.3 dB utilizing turbo-like codes in a system with $\mu = 2$. However, those schemes are limited to low rate codes, which historically perform better than high spectral efficiency codes like RCM or Bit-Interleaved Coded Modulation

(BICM) [5]. Therefore, the results obtained with the proposed RCM-CIS scheme are promising considering the fact that they are obtained for high throughput transmission scenarios, with system rates up to 11.93 bits per complex channel use.

Finally, as observed in all figures, the proposed RCM-CIS scheme suffers from high error floors. In the next section we present the results for the parallel LDGM scheme of Section 7.3.3.

7.5.2 NUMERICAL RESULTS FOR RCM-CIS SCHEMES WITH THE PARALLEL LDGM CONCATENATION

In this case we substitute I of the N (recall $N = 5000$) RCM coded symbols of each encoder by binary LDGM coded symbols generated with design parameter d_l, i.e., each input bit has d_l random connections to the LDGM symbols. As mentioned in Section 7.3.3, the LDGM coded symbols are multiplied by modulation factors before being transmitted through the MAC in order to improve decoding. Thus, each LDGM MAC symbol is a linear combination of the μ binary LDGM symbols weighted by the values $\Theta = \{\theta_1, \theta_2, \ldots, \theta_\mu\}$. This behaves similarly to an RCM code whose entries are $\{\pm 1\}$ instead of $\{0, 1\}$, and therefore, the modulation factors can be selected following the guidelines of RCM design.

Fig. 7.9 shows the numerical results obtained by Monte Carlo simulations for the examples in Table 7.4 with $R = 6$. Continuous lines represent the parallel RCM-CIS LDGM concatenation. For reference, we have also included the curves for the corresponding RCM-CIS schemes in the form of dashed lines. Recall that the parallel RCM-CIS LDGM concatenation maintains the rate of the RCM-CIS schemes because the added LDGM symbols substitute the same amount of RCM symbols. The weight matrices $\mathcal{W}^C$ for the RCM-CIS sub-scheme generating $5000 - I$ symbols are the same as in the previous section. The number of generated LDGM symbols I is equal to $I = 108$ for the cases of $\mu = 2$ with $p = .01$ and $.001$, $I = 204$ for the cases of $\mu = 3$ with $p = .01$ and $.001$, $I = 300$ for the case of $\mu = 4$ with $p = .01$, and finally, $I = 204$ for $\mu = 4$ and $p = .001$. The LDGM node degree parameter d_l has been set to 3 in all cases. The modulation values used in this section are $\Theta = \{1, 2\}$ for $\mu = 2$, $\Theta = \{1, 2, 4\}$ for $\mu = 3$ and $\Theta = \{3, 4, 5, 7\}$ for $\mu = 4$.

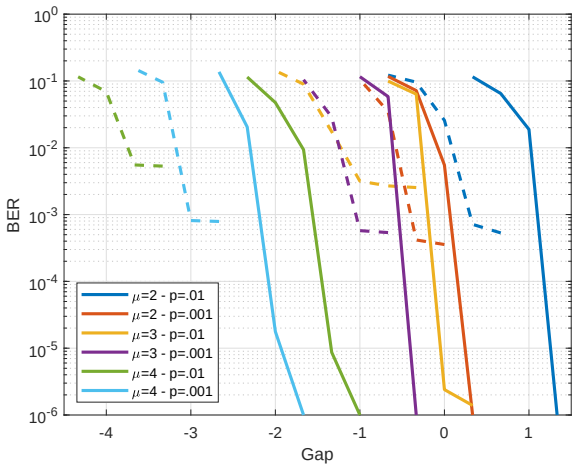


Figure 7.9: *Numerical results obtained by Monte Carlo simulation for the examples in Table 7.4 when $R = 6$ (with $p = .01$ and $.001$), each case in a different color.*

As shown in Fig. 7.9, the MAC parallel concatenation of RCM-CIS schemes with LDGM codes is able to reduce the error floor encountered in RCM-CIS. In this case, the gap to the separation limit varies between 1.25 dB for $\mu = 2$ and $p = .01$ to -2 dB for $\mu = 4$ and $p = .001$. The substitution of RCM symbols by LDGM ones prevents the error floor at the cost of some performance loss. Larger error floors require the use of more LDGM symbols resulting in larger performance loss. For example, in the case of $\mu = 4$ and $p = .01$ the error floor of the RCM-CIS scheme is around $\text{BER} = 5 \cdot 10^{-3}$. The performance degradation when using LDGM concatenation is more than 2 dB, since 300 LDGM symbols have to be used. On the other hand, for the example of $\mu = 2$ and $p = .001$ with an error floor close to 10^{-4}, only 108 LDGM symbols are used. This results in a performance degradation of only 0.5 dB.

7.5.3 NUMERICAL RESULTS FOR TRANSMITTING CLUSTERS OF CORRELATED INFORMATION SOURCES

In Section 7.4, we explained how to convert an RCM-CIS scheme that transmits correlated information blocks of length K, into a system transmitting a set of λ independent clusters of correlated information sources,

each with source blocks of length $\{K_l\}_{l\in[1:\lambda]}$, such that $K = \sum_{l=1}^{\lambda} K_l$. In this results we consider two different cluster examples, the first one where the RCM-CIS scheme is divided into 6 clusters (case $\lambda = 6$) with equal rate $R_l = \dfrac{R}{6}$ for $l = 1, \ldots, 6$, and the second example, in which the RCM-CIS scheme is divided into $\lambda = 3$ clusters (case $\lambda = 3$), the first two with rates $R_1 = R_2 = \dfrac{R}{4}$ and the third one with rate $R_3 = \dfrac{R}{2}$. Figure 7.10 presents the numerical results obtained for the case with total sum-rate (the sum of rates of all clusters) $R = 6$ and $p = .01$. Since the total sum-rate is 6 and we fixed $N = 5000$, the total source block length is $K = 15000$, which is divided between the λ clusters. For example, in the asymmetric case of $\lambda = 3$, the first and second clusters are each comprised of μ correlated users generating blocks of 15000/4 bits, whereas the third cluster contains μ users generating source blocks of 15000/2 bits. Recall that sources of different clusters are independent. Each of the $\mu \cdot \lambda$ ($\lambda = 3$ in this example) transmitters encodes its input bits blocks of length 15000/4 or 15000/2 into 5000 RCM symbols, encoded at rates 1.5 or 3, respectively. These $\mu \cdot \lambda$ coded sequences are all superimposed in the MAC. As explained in Chapter 6 and Section 7.4, the superimposed clusters of correlated information sources can be considered as an RCM-CIS scheme at the receiver and decoded as such.

For context, we have also included in Fig. 7.10 the standard RCM-CIS case ($\lambda = 1$), since the RCM-CIS scheme is just a particular example where all users belong to one unique cluster. The results show that there is almost no performance difference between the transmission of a CIS scenario or a cluster of CIS (for simplicity we have considered that inside every cluster the users are equally correlated, but users from different clusters are independent).

Finally, in order to lower the error floor in Fig. 7.10, we also extend the RCM-CIS LDGM concatenation for the transmission of clusters of correlated information sources, and present the numerical simulation outcomes in Fig. 7.11. Recall that these schemes are designed to resemble a parallel RCM-CIS LDGM scheme after the MAC, and therefore, the design parameters $\mathcal{W}^C$, I and d_l are the same as in the previous examples of Sections 7.5.1 and 7.5.2.

Notice the slight performance degradation when transmitting clusters of correlated information sources, compared to the performance of RCM-

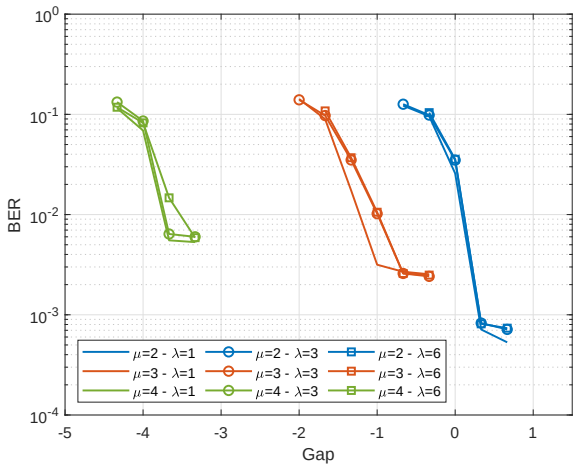


Figure 7.10: *Numerical results obtained by Monte Carlo simulation for the transmission of λ clusters of μ correlated information sources using RCM-CIS schemes when $p = .01$ and the sum-rate is $R = 6$.*

CIS schemes. This effect was also experienced in the RCM-MAC schemes of Chapter 6, where it was shown to occur due to the relatively shorter block lengths resulting from the construction of RCM-MAC schemes. This effect can be avoided by making the block lengths larger.

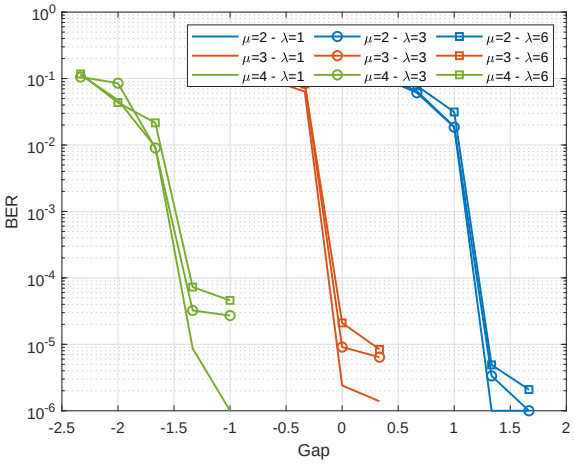


Figure 7.11: *Numerical results obtained by Monte Carlo simulation for the transmission of λ clusters of μ correlated information sources using the parallel concatenation of RCM-CIS schemes with LDGM codes when $p = .01$ and the sum-rate is $R = 6$.*

7.6 DISCUSSION

In this chapter we have considered a joint source-channel coding for the transmission of spatially correlated information sources over the Additive White Gaussian noise multiple access channel. It is well known that the system performance is increased if the designed codewords preserve part of the source correlation. The linear and sparse nature of RCM codes makes them good candidates to fulfill this design condition. Concretely, we have designed a coding scheme based on a set of identically distributed RCM codes whose weight values are different, but jointly designed.

The use of RCM codes with the same structure, in conjunction with correlated sources, produces cycles of degree 4 in case of using a traditional decoder. Therefore, we have been forced to implement a non-binary decoder for the first time in RCMs. The proposed decoder achieves good results, outperforming the separation limit. However, its complexity increases exponentially with the number of users, limiting the number of considered users.

We have also extended the RCM-CIS schemes to the transmission of clusters of correlated information sources by utilizing one irregular RCM structure of an RCM-MAC scheme per cluster, to generate the RCM codes of the users contained in it. From the receiver's point of view, this generalization is seamless and does not increase the complexity.

Finally, as in point-to-point RCM codes, the RCM schemes for the MAC presented here and in Chapter 6 present high error floors. To solve this, we have combined our RCM schemes with an LDGM parallel concatenation methodology that is able to correct those residual errors, as has been done in the literature for conventional RCMs.

CHAPTER 8

Conclusions

In this dissertation, we have explored the use of RCM and parallel RCM-LDGM codes for the transmission of information produced by binary sources over various communication scenarios.The proposed schemes achieve good BER performance while maintaining the high-rate, smooth rate adaptation and joint source-channel coding capabilities that originally made RCM advantageous over traditional AMC techniques. Concretely, we have designed practical encoding and decoding schemes for the transmission of single and multiple sources over point-to-point and multiple access channels, respectively. We next detail the particular communication scenarios or research problems that we focused on:

- **EXIT chart analysis of RCM-LDGM codes.**

Previous works and this dissertation demonstrate that parallel RCM-LDGM codes are very well suited for transmitting uniform and non-uniform binary sources at high rates if they are properly designed. However, the RCM-LDGM codes designed in the literature were based on brute force approach, which required time consuming Monte Carlo simulations, specially for large block lengths. In order to speed up the design process, we successfully developed in Chapter 3 an EXIT chart analysis for these codes, which presented the challenge of the combination of analog and digital check nodes, something not encountered in other works. The presented method obtains very precise EXIT charts for the case of AWGN and fast

fading Rayleigh channels, whose BER vs SNR curve predictions are very close to the results obtained through simulations.

- **High-rate transmission of sources with memory.**

We have proposed in Chapter 4 a new source-controlled coding scheme for high-throughput transmission of binary sources with memory over AWGN channels. The proposed strategy is based on the BWT, which transforms the original source with memory into a set of independent non-uniform discrete memory-less binary sources. These sources are then separately RCM-LDGM encoded at coding rates that are optimized according to the source's entropy. Good RCM-LDGM codes operating at the optimal rates are designed using EXIT chart analysis. Results in Chapter 4 show that the proposed scheme outperforms the traditional strategy of using the source's statistics in the decoding process.

- **RCM-LDGM codes for impulsive noise channels.**

In Chapter 5 we showed the suitability of using of RCM-LDGM codes in impulsive noise channels if the decoder is adapted according to the channel state information of the noise burst. In order to asses the performance of these codes, we considered modified versions of Middleton class A noise and Markov-Middleton noise models whose expected number of impulses per N channel uses tends to 0 as N increases. Under that condition, the capacity of the impulsive noise channels tends asymptotically to the AWGN capacity and the performance of RCM-LDGM codes is within 2.5 dB of that limit.

- **Transmission of uncorrelated sources over MACs.**

A novel high-rate coding system for the non-orthogonal MAC was presented in Chapter 6. The proposed system, named RCM-MAC, utilizes the linear nature of RCM codes and the multiple access channel to construct irregular RCM codes, one for each user, in such a way that the combination of all generation matrices forms a standard point-to-point RCM code from the receiver's perspective. All users are jointly decoded by applying the standard RCM sum-product algorithm over the factor graph that results from the combination of the individual RCMs. These systems are able to accommodate a large number of users while operating at high sum-rates and maintaining the decoding complexity of point-to-point RCM codes. Simulation results show that the performance is around 2 dB away from the Shannon limit for AWGN MAC channels, independently of the number

of users. Interestingly, when the Rayleigh fast fading channel is considered, the performance of the RCM-MAC system improves as the number of users increases due to the average effect.

- **Transmission of correlated information sources.**

Chapter 7 extends the RCM-MAC system of Chapter 6 to a joint source-channel coding system, named RCM-CIS, that considers the transmission of spatially correlated information sources over AWGN multiple access channels. The proposed RCM-CIS systems consist of a set of RCM codes with the same structure but whose weight values differ following a joint design that intends to preserve the source correlation in the codewords of the different users. Simulation results show that the proposed RCM-CIS system is able to outperform the separation limit established by the combination of the Slepian-Wolf, for the source compression, and Shannon, for the channel coding, theorems.

In order to lower the error floor of RCM codes in multi-user communications, we also propose to combine the RCM scheme with an LDGM parallel concatenation methodology previously used in point-to-point communications. Note that this solution is also valid for the RCM-MAC systems of Chapter 6. Finally, we also extend the RCM-CIS systems to the transmission of clusters of correlated information sources by utilizing one irregular RCM structure of an RCM-MAC system per cluster to generate the RCM codes of the users contained in the cluster. As for the RCM-MAC case, this generalization does not increase the decoding complexity.

8.1 FUTURE RESEARCH LINES

Many of the presented coding systems and results are encouraging and suggest further study and analysis. This section presents some extension/improvement ideas of the work developed in this Thesis.

- **Short block-length RCM-LDGM codes for the BWT strategy.**

The sorting algorithm of the BWT is very sensitive to errors in the transformed sequence, requiring very powerful RCM-LDGM codes, that are usually achieved with large block lengths. However, specially when the source is modeled by a HMM, some BWT output blocks may have short

lengths. We solved this problem in Chapter 4 by building larger segments stacking consecutive blocks. However, this solution leads to decoding delays (and possibly system overhead), and could be improved if efficient short block-length RCM-LDGM were investigated. The EXIT chart analysis of Chapter 3 is not valid for this case since it only considers asymptotic behavior.

- **Impulsive noise channels without channel state information.**

The work in Chapter 5 considered that the channel state information is available at the receiver. This is realistic in some cases, but there are other practical situations in which this information is not available. Works in the literature propose approximations to the input LLR values of the received channel signal suitable for Class A noise environments without this information. However, this is not applicable to the RCM case due to the analog nature of its symbols and the consequential decoding based on direct convolutions.

- **Design methods for RCM-CIS systems.**

Chapter 7 proposed a novel coding system suitable for the transmission of correlated information sources over multiple access channels. However, the codes presented in it are designed based on trial and error. As has been done point-to-point RCM-LDGM, design procedures based on EXIT charts could be built in order to speed up their design. These EXIT charts present some complexity given the non-binary nature of the RCM-CIS decoder. The histogram computation in non-binary systems becomes quite demanding or even impossible for larger index sets and alternative approaches like the index-based extrinsic mutual information [171] are ofter required in the literature.

- **Alternative decoding methods for RCM-CIS systems.**

The maximum number of correlated users considered in Chapter 7 was four, due to the fact that the complexity of the utilized non-binary decoder increases exponentially with the number of users. In order to build systems that allow a larger number of users, alternative binary decoding techniques need to be investigated.

- **Rate Adaptation.**

Throughout this dissertation we have developed different coding strategies based on RCM (and RCM-LDGM) codes. We have focused our results and parameter design on the BER performance of these coding systems at specific rates. As mentioned in the introduction, the RCM and the parallel RCM-LDGM codes have excellent rate adaptation capabilities. Therefore, it would be interesting to study this property on the proposed communication systems and study the parameter optimization towards this objective.

- **Multi-user MIMO RCM systems.**

The very high data rates that are required for future wireless systems do not appear to be feasible with the conventional techniques and architectures. This has resulted in an explosion of research activities to characterize the theoretical and practical issues associated with Multiple-Input Multiple-Output (MIMO) channels and the extension of these concepts to multi-user systems. The already high data rates achieved by the single-input single-output multi-user RCM techniques presented in this dissertation encourages further extension of these RCM system to the multi-user MIMO scenario.

- **More Realistic Correlated Information Source Models.**

The work presented in Chapter 7 presents a simple source model. This works can be extended to more sophisticated models, such as when the correlation between sources has memory and is modeled by a HMM as in [172].

Appendices

APPENDIX A

The Burrows Wheeler Transform (BWT)

The Burrows-Wheeler Transform (also known as Block-Sorting) is at the base of compression algorithms which are the state of the art in lossless data compression. In this appendix we provide a detailed description of the Burrows Wheeler Transform, composed by a direct and reverse transformations.

Definition 3. *Assume a stationary ergodic source S whose output is modeled by the stochastic process $\{U_k\}_{k=1}^{\infty}$, with alphabet $\mathcal{U}$, entropy per single letter $H(U)$ and entropy rate $\mathcal{H}(U)$. The BWT is a reversible transform that operates on blocks of length K, defined as a direct mapping*

$$BWT\left(\{U_k\}_{k=1}^{K}\right) : \mathcal{U}^K \rightarrow \mathcal{U}^K \times \{1, 2, \ldots, K\} \tag{A.1}$$

and an inverse mapping

$$BWT^{-1}\left(\{T_k\}_{k=1}^{K}\right) : \mathcal{U}^K \rightarrow \mathcal{U}^K \times \{1, 2, \ldots, K\}. \tag{A.2}$$

The BWT outputs the following quantities 1) a transformed block $\{T_k\}_{k=1}^{K}$ with the same alphabet $\mathcal{U}$ as the input sequence and 2) an integer $J \in \{1, \ldots, K\}$ which is required for the inverse transform. For the shake of completeness, in the remainder of the appendix we detail this pair of mappings.

A.1 DIRECT TRANSFORMATION

In essence, the direct transformation consist of a sorting algorithm, where all possible cyclic permutations of the input sequence, to which a final string pointer $\star$ (also known as End Of File) is added, are ordered following a lexicographic criteria. The output of the BWT transform is composed by a K-length sequence $\{T_k\}_{k=1}^{K}$, which is the last column of the ordered cyclic permutations, after pointer $\star$ is removed. Additionally, the BWT generates at its output an integer number $J \in \{1, \ldots, K\}$ that indicates the position of the removed pointer $\star$ in the last column.

We next present an example of the BWT applied to a binary sequence, but it may serve as good example to fully understand the transformation procedure. Let us assume a source S with binary alphabet $\{0, 1\}$ generating the block $\{U_k\}_{k=1}^{6} = \{110100\}$. The pointer $\star$ is appended to the block and inserted into the BWT direct transform, whose steps are shown in Table A.1.

Table A.1: *Example of the direct Burrows-Wheeler Transformation.*

Input	Cyclic Permutations	Ordered Permutations	Last Column
110100$\star$	110100$\star$	00$\star$1101	11001$\star$0
	$\star$110100	0100$\star$11	
	0$\star$11010	0$\star$11010	
	00$\star$1101	100$\star$110	
	100$\star$110	10100$\star$1	
	0100$\star$11	110100$\star$	
	10100$\star$1	$\star$110100	

And the output of the BWT direct transformation is given by the sequence $\{T_k\}_{k=1}^{6} = \{110010\}$ and the index $J = 6$.

A.2 INVERSE TRANSFORMATION

Sorting criteria based algorithms are generally not reversible. Intuitively we may justify this remark because the result of lexicographically sorting a given data set is always the same independently of its initial data

order. However, the BWT is a reversible transformation and therefore, the initial data set $\{U_k\}_{k=1}^{K}$ must be possible to recover from the transformed block $\{T_k\}_{k=1}^{K}$ and the index J. Recall that $\{T_k\}_{k=1}^{K}$ and J provide sufficient information to recover the last column of the ordered permutations in Table A.1. The inverse transformation is based on the reconstruction of the entire ordered permutation list. To rebuild it, we will procedure column-wise. Notice that, due to its structure, the first column of the ordered permutation table is a lexicographically sorted copy of the last column, i.e. the transformed sequences $\{T_k\}_{k=1}^{K}$ and the appended pointer $\star$. Then, in order to obtain this first column one has to sort the sequence $\{\{T_k\}_{k=1}^{J-1}, \star, \{T_k\}_{k=J+1}^{K}\}$, formed by adding the final string pointer $\star$ at position J. Once this is done, we have every possible combination of 2 symbols within the sequence, since the first and the last column in the table generate the pairs of symbols given by the cyclic permutation. Thereby, by generating all the possible concatenations and applying to them the lexicographical criteria we will obtain the third column of the ordered permutation table. By repeating this process for the successive concatenations the ordered permutation table is reconstructed. In our example, we will recover the original sequence $\{U_k\}_{k=1}^{6}$ based on the knowledge $\{T_k\}_{k=1}^{6} = \{110010\}$ and $J = 6$. Let us illustrate the process through Table A.2.

Table A.2: *Example of inverse Burrows-Wheeler Transformation.*

Start	Sort 1	Concat. 2	Sort 2	Concat. 3	Sort 3		Sort 8
1	0	10	00	100	00$\star$		00$\star$1101
1	0	10	01	101	010		0100$\star$11
0	0	00	0$\star$	00$\star$	0$\star$1		0$\star$11010
0	1	01	10	010	100	...	100$\star$110
1	1	11	10	110	101		10100$\star$1
$\star$	1	$\star$1	11	$\star$11	110		110100$\star$
0	$\star$	0$\star$	$\star$1	0$\star$1	$\star$11		$\star$110100

We can check that the original sequence $\{U_k\}_{k=1}^{6} = \{110100\}$ is the J-th row of the sorted permutations (after removing the appended $\star$). Obviously, the lexicographical sorting criteria can be arbitrarily chosen accordance to the alphabet $\mathcal{U}$ of the input sequences.

References

[1] C. E. Shannon, "A mathematical theory of communication," in *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379-423, July 1948, doi: 10.1002/j.1538-7305.1948.tb01338.x.

[2] A. Gudipati and S. Katti, "Strider: automatic rate adaptation and collision handling", in *Proc. SIGCOMM Comput. Commun. Rev.* 41, 4 (August 2011), 158169. doi: 10.1145/2043164.2018455

[3] J. Perry, P. A. Iannucci, K. E. Fleming, H. Balakrishnan and D. Shah, "Spinal codes", in *Proc. SIGCOMM Comput. Commun. Rev*, New York, NY, USA, 4960. doi: 10.1145/2342356.2342363

[4] H. Cui, C. Luo, K. Tan, F. Wu and C. W. Chen, "Seamless rate adaptation for wireless networking", in *Proc. 2011 ACM MSWiM*, New York, NY, USA, 437446. doi: 10.1145/2068897.2068971

[5] H. Cui, C. Luo, J. Wu, C. W. Chen and F. Wu, "Compressive Coded Modulation for Seamless Rate Adaptation," in *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 4892-4904, October 2013, doi: 10.1109/TWC.2013.090413.121308.

[6] L. Li and J. Garcia-Frias, "Hybrid analog-digital coding for nonuniform memoryless sources," in *2015 49th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, 2015, pp. 1-5, doi: 10.1109/CISS.2015.7086861.

[7] L. Li and J. Garcia-Frias, "Hybrid analog-digital coding scheme based on parallel concatenation of linear random projections and LDGM codes," in *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, 2014, pp. 1-6, doi: 10.1109/CISS.2014.6814118.

[8] A. E. Gamal and Y-H. Kim. "Network Information Theory". *Cambridge University Press*, 2012. USA, doi: 10.1017/CBO9781139030687

[9] F. R. Kschischang, B. J. Frey and H. -. Loeliger, "Factor graphs and the sum-product algorithm," *in IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498-519, Feb 2001, doi: 10.1109/18.910572.

[10] D. J. Mackay, "Information Theory, Inference, and Learning Algorithms", *Cambridge University Press*, 2003. USA.

[11] N.Wiberg, "Codes and decoding on general graphs", *Ph.D. dissertation*, 1996.

[12] Gallager, R. G., "Information Theory and Reliable Communication", *John Wiley and Sons*, New York, 1968.

[13] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proceedings of ICC '93 - IEEE International Conference on Communications*, Geneva, Switzerland, 1993, pp. 1064-1070 vol.2, doi: 10.1109/ICC.1993.397441.

[14] R. Gallager, "Low-density parity-check codes," in *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21-28, January 1962, doi: 10.1109/TIT.1962.1057683.

[15] R. G. Gallager, "Low-Density Parity-Check Codes," *M.I.T Press*, 1963.

[16] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," in *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 399-431, March 1999, doi: 10.1109/18.748992.

[17] J. Garcia-Frias and Wei Zhong, "Approaching Shannon performance by iterative decoding of linear codes with low-density generator matrix," in *IEEE Communications Letters*, vol. 7, no. 6, pp. 266-268, June 2003, doi: 10.1109/LCOMM.2003.813816.

[18] W. Zhong, H. Lou and J. Garcia-Frias, "LDGM codes for joint source-channel coding of correlated sources," in *Proceedings 2003 International Conference on Image Processing* (Cat. No.03CH37429), Barcelona, Spain, 2003, pp. I-593, doi: 10.1109/ICIP.2003.1247031.

[19] W. Zhong, H. Chai and J. Garcia-Frias, "Approaching the shannon limit through parallel concatenation of regular LDGM codes," in *Proceedings. International Symposium on Information Theory*, 2005. ISIT 2005., Adelaide, SA, Australia, 2005, pp. 1753-1757, doi: 10.1109/ISIT.2005.1523646.

[20] H. Lou and J. Garcia-Frias, "Rate-compatible low-density generator matrix codes," in *IEEE Transactions on Communications*, vol. 56, no. 3, pp. 321-324, March 2008, doi: 10.1109/TCOMM.2008.050140.

[21] R. Hamzaoui, V. Stankovic and Zixiang Xiong, "Optimized error protection of scalable image bit streams [advances in joint source-channel coding for images]," in *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 91-107, Nov. 2005, doi: 10.1109/MSP.2005.1550192.

[22] L. P. Kondi, F. Ishtiaq and A. K. Katsaggelos, "Joint source-channel coding for motion-compensated DCT-based SNR scalable video," in *IEEE Transactions on Image Processing*, vol. 11, no. 9, pp. 1043-1052, Sept. 2002, doi: 10.1109/TIP.2002.802507.

[23] Z. He, J. Cai and C. Wen Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 6, pp. 511-523, June 2002, doi: 10.1109/TCSVT.2002.800313.

[24] Q. Zhang, W. Zhu, Z. Ji and Y. Zhang, "A power-optimized joint source channel coding for scalable video streaming over wireless channel," in *ISCAS 2001. The 2001 IEEE International Symposium on Circuits and Systems* (Cat. No.01CH37196), Sydney, NSW, Australia, 2001, pp. 137-140 vol. 5, doi: 10.1109/ISCAS.2001.922004.

[25] S. S. Arslan, P. C. Cosman and L. B. Milstein, "Generalized Unequal Error Protection LT Codes for Progressive Data Transmission," in *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3586-3597, Aug. 2012, doi: 10.1109/TIP.2012.2195668.

[26] X. Pan, A. Cuhadar and A. H. Banihashemi, "Combined source and channel coding with JPEG2000 and rate-compatible low-density Parity-check codes," in *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1160-1164, March 2006, doi: 10.1109/TSP.2005.863032.

[27] S. Aditya and S. Katti, "FlexCast: graceful wireless video streaming", in *Proceedings of the 17th annual international conference on Mobile computing and networking* (MobiCom '11). Association for Computing Machinery, New York, NY, USA, 277288. doi: 10.1145/2030613.2030645

[28] O. Y. Bursalioglu, G. Caire and D. Divsalar, "Joint source-channel coding for deep space image transmission using rateless codes," in *2011 Information Theory and Applications Workshop*, La Jolla, CA, USA, 2011, pp. 1-10, doi: 10.1109/ITA.2011.5743568.

[29] G. Caire, S. Shamai and S. Verdu, "A new data compression algorithm for sources with memory based on error correcting codes," in *Proceedings 2003 IEEE Information Theory Workshop* (Cat. No.03EX674), Paris, France, 2003, pp. 291-295, doi: 10.1109/ITW.2003.1216751.

[30] G. Caire, S. Shamai, A. Shokrollahi and S. Verdu, "Universal variable-length data compression of binary sources using fountain codes," in *Information Theory Workshop*, San Antonio, TX, USA, 2004, pp. 123-128, doi: 10.1109/ITW.2004.1405286.

[31] J. Del Ser, P. M. Crespo, I. Esnaola and J. Garcia-Frias, "Joint Source-Channel Coding of Sources with Memory using Turbo Codes and the Burrows-Wheeler Transform," in *IEEE Transactions on Communications*, vol. 58, no. 7, pp. 1984-1992, July 2010, doi: 10.1109/TCOMM.2010.07.090141.

[32] M. Fresia, F. Perez-Cruz and H. V. Poor, "Optimized concatenated LDPC codes for joint source-channel coding," in *2009 IEEE International Symposium on Information Theory*, Seoul, Korea (South), 2009, pp. 2131-2135, doi: 10.1109/ISIT.2009.5205766.

[33] G.-C. Zhu and F. Alajaji, "Turbo codes for nonuniform memoryless sources over noisy channels," in *IEEE Communications Letters*, vol. 6, no. 2, pp. 64-66, Feb. 2002, doi: 10.1109/4234.984695.

[34] F. Cabarcas, R. D. Souza and J. Garcia-Frias, "Turbo coding of strongly nonuniform memoryless sources with unequal energy allocation and PAM signaling," in *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1942-1946, May 2006, doi: 10.1109/TSP.2006.872526.

[35] G. Ungerboeck and I. Csajka, "On improving data-link performance by increasing the channel alphabet and introducing sequence coding", in *Proccedings 1976 IEEE ISIT*, pp. 53.

[36] G. Ungerboeck, "Channel coding with multilevel/phase signals," in *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 55-67, January 1982, doi: 10.1109/TIT.1982.1056454.

[37] H. Imai and S. Hirakawa, "A new multilevel coding method using error-correcting codes," in *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp. 371-377, May 1977, doi: 10.1109/TIT.1977.1055718.

[38] U. Wachsmann, R. F. H. Fischer and J. B. Huber, "Multilevel codes: theoretical concepts and practical design rules," in *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1361-1391, July 1999, doi: 10.1109/18.771140.

[39] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," in *IEEE Transactions on Communications*, vol. 40, no. 5, pp. 873-884, May 1992, doi: 10.1109/26.141453.

[40] G. Caire, G. Taricco and E. Biglieri, "Bit-interleaved coded modulation," in *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927-946, May 1998, doi: 10.1109/18.669123.

[41] A. Martinez, A. Guillen i Fabregas, G. Caire and F. M. J. Willems, "Bit-Interleaved Coded Modulation Revisited: A Mismatched Decoding Perspective," in *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2756-2765, June 2009, doi: 10.1109/TIT.2009.2018177.

[42] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proceedings of ICC '93 - IEEE International Conference on Communications*, Geneva, Switzerland, 1993, pp. 1064-1070 vol.2, doi: 10.1109/ICC.1993.397441.

[43] R. Gallager, "Low Density Parity Check Codes" in MIT, Cambridge, 1963.

[44] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes", in *Electronics Letters*, vol. 33, no. 6, pp. 457-458, 1997. doi: 10.1049/el:19970362.

[45] S. Nanda, K. Balachandran and S. Kumar, "Adaptation techniques in wireless packet data services," in *IEEE Communications Magazine*, vol. 38, no. 1, pp. 54-64, Jan. 2000, doi: 10.1109/35.815453.

[46] S. H. Y. Wong, H. Yang, S. Lu and V. Bharghavan, "Robust rate adaptation for 802.11 wireless networks", in *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking*, pp. 146-157. doi: 10.1145/1161089.1161107.

[47] Q. Xia and M. Hamdi, "Smart sender: a practical rate adaptation algorithm for multirate IEEE 802.11 WLANs," in *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, pp. 1764-1775, May 2008, doi: 10.1109/TWC.2008.061047.

[48] Y. Song, X. Zhu, Y. Fang and H. Zhang, "Threshold optimization for rate adaptation algorithms in IEEE 802.11 WLANs," in *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 318-327, January 2010, doi: 10.1109/TWC.2010.01.090459.

[49] E. Y. Rocher and R. L. Pickholtz, "An Analysis of the Effectiveness of Hybrid Transmission Schemes," in *IBM Journal of Research and Development*, vol. 14, no. 4, pp. 426-433, July 1970, doi: 10.1147/rd.144.0426.

[50] S. Lin and P. S. Yu, "A hybrid ARQ scheme with parity retransmission for error control of satellite channels", in *IEEE Transactions on Communications*, vol. COM-30, no. 7, pp. 1701-1719, July 1982. doi: 10.1109/TCOM.1982.1095643.

[51] Yu-Ming Wang and Shu Lin, "A Modified Selective-Repeat Type-II Hybrid ARQ System and Its Performance Analysis," in *IEEE Transactions on Communications*, vol. 31, no. 5, pp. 593-608, May 1983, doi: 10.1109/TCOM.1983.1095873.

[52] S. M. Kim, W. Choi, T. W. Ban and D. K. Sung, "Optimal Rate Adaptation for Hybrid ARQ in Time-Correlated Rayleigh Fading Channels," in *IEEE Transactions on Wireless Communications*, vol. 10, no. 3, pp. 968-979, March 2011, doi: 10.1109/TWC.2011.011111.101027.

[53] D. N. Rowitch and L. B. Milstein, "On the performance of hybrid FEC/ARQ systems using rate compatible punctured turbo (RCPT) codes," in *IEEE Transactions on Communications*, vol. 48, no. 6, pp. 948-959, June 2000, doi: 10.1109/26.848555.

[54] A. Shokrollahi, "Raptor codes," in *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2551-2567, June 2006, doi: 10.1109/TIT.2006.874390.

[55] E. J. Candes, J. Romberg and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," in *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006, doi: 10.1109/TIT.2005.862083.

[56] D. L. Donoho, "Compressed sensing," in *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289-1306, April 2006, doi: 10.1109/TIT.2006.871582.

[57] E. J. Candes, J. Romberg and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," in *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006, doi: 10.1109/TIT.2005.862083.

[58] E. J. Candes and T. Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?," in *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406-5425, Dec. 2006, doi: 10.1109/TIT.2006.885507.

[59] C. E. Shannon, "Communication in the Presence of Noise," in *Proceedings of the IRE*, vol. 37, no. 1, pp. 10-21, Jan. 1949, doi: 10.1109/JRPROC.1949.232969.

[60] E. J. Candes, "Compressive sampling", in *Proceedings of the international congress of mathematicians*, Madrid, August 2230, 2006, pp. 1433-1452. doi: 10.4171/022-3/69

[61] R. Gribonval, M. Nielsen, "Highly Sparse Representations from Dictionaries are Unique and Independent of the Sparseness Measure," in *Applied and Computational Harmonic Analysis* Volume 22, Issue 3, May 2007, Pages 335-355. doi: 10.1016/j.acha.2006.09.003.

[62] J. A. Tropp, "Just relax: convex programming methods for identifying sparse signals in noise," in *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1030-1051, March 2006, doi: 10.1109/TIT.2005.864420.

[63] D. Donoho and Y. Tsaig, "Fast solution of l1-norm minimization problems when the solution maybe sparse", in *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4789-4812, 2008. doi: 10.1109/TIT.2008.929958

[64] J. A. Tropp and A. C. Gilbert, "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit," in *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655-4666, Dec. 2007, doi: 10.1109/TIT.2007.909108.

[65] Y. C. Pati, R. Rezaiifar and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, 1993, pp. 40-44 vol.1, doi: 10.1109/ACSSC.1993.342465.

[66] DeVore, R.A., Temlyakov, V.N., "Some remarks on greedy algorithms", in *Adv Comput* Math 5, 173187 (1996). doi: 10.1007/BF02124742.

[67] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing", in *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265-274, 2009. doi: 10.1016/j.acha.2009.04.002.

[68] W. Dai and O. Milenkovic, "Subspace Pursuit for Compressive Sensing Signal Reconstruction," in *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2230-2249, May 2009, doi: 10.1109/TIT.2009.2016006.

[69] M. R. Duarte, M. B. Wakin, D. Baron and R. G. Baraniuk, "Universal distributed sensing via random projections," in *2006 5th International Conference on Information Processing in Sensor Networks*, Nashville, TN, USA, 2006, pp. 177-185, doi: 10.1145/1127777.1127807.

[70] W. U. Bajwa, J. D. Haupt, A. M. Sayeed and R. D. Nowak, "Joint SourceChannel Communication for Distributed Estimation in Sensor Networks," in *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3629-3653, Oct. 2007, doi: 10.1109/TIT.2007.904835.

[71] S. Feizi, M. Mdard and M. Effros, "Compressive sensing over networks," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, USA, 2010, pp. 1129-1136, doi: 10.1109/ALLERTON.2010.5707037.

[72] S. Feizi and M. Medard, "A power efficient sensing/communication scheme: Joint source-channel-network coding by using compressive sensing," in *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, USA, 2011, pp. 1048-1054, doi: 10.1109/Allerton.2011.6120283.

[73] F. Chen, F. Lim, O. Abari, A. Chandrakasan and V. Stojanovic, "Energy-Aware Design of Compressed Sensing Systems for Wireless Sensors Under Performance and Reliability Constraints," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 3, pp. 650-661, March 2013, doi: 10.1109/TCSI.2012.2215738.

[74] S. Sarvotham, D. Baron and R. G. Baraniuk, "Sudocodes - Fast Measurement and Reconstruction of Sparse Signals," in *2006 IEEE International Symposium on Information Theory*, Seattle, WA, USA, 2006, pp. 2804-2808, doi: 10.1109/ISIT.2006.261573.

[75] F. Wu, J. Fu, Z. Lin and B. Zeng, "Analysis on Rate-Distortion Performance of Compressive Sensing for Binary Sparse Source," in *2009 Data Compression Conference*, Snowbird, UT, USA, 2009, pp. 113-122, doi: 10.1109/DCC.2009.24.

[76] X. L. Liu, C. Luo and F. Wu, "Formulating Binary Compressive Sensing Decoding with Asymmetrical Property," in *2011 Data Compression Conference*, Snowbird, UT, USA, 2011, pp. 213-222, doi: 10.1109/DCC.2011.28.

[77] D. Baron, S. Sarvotham and R. Baraniuk, "Bayesian compressive sensing via belief propagation", in *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 269-280, 2010. doi: 10.1109/TSP.2009.2027773.

[78] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," in *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727-1737, Oct. 2001, doi: 10.1109/26.957394.

[79] S. ten Brink, G. Kramer and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," in *IEEE Transactions on Communications*, vol. 52, no. 4, pp. 670-678, April 2004, doi: 10.1109/TCOMM.2004.826370.

[80] J. Du, L. Yang, J. Yuan, L. Zhou and X. He, "Bit Mapping Design for LDPC Coded BICM Schemes With Multi-Edge Type EXIT Chart," in *IEEE Communications Letters*, vol. 21, no. 4, pp. 722-725, April 2017, doi: 10.1109/LCOMM.2016.2646341.

[81] T. Cheng, K. Peng, J. Song and K. Yan, "EXIT-Aided Bit Mapping Design for LDPC Coded Modulation with APSK Constellations," in *IEEE Communications Letters*, vol. 16, no. 6, pp. 777-780, June 2012, doi: 10.1109/LCOMM.2012.041112.112436.

[82] J. Wu, Z. Teng, H. Cui, C. Luo, X. Huang and H. Chen, "Arithmetic-BICM for Seamless Rate Adaptation for Wireless Communication Systems," in *IEEE Systems Journal*, vol. 10, no. 1, pp. 228-239, March 2016, doi: 10.1109/JSYST.2014.2363081.

[83] Burrows, M., Wheeler, D., "A Block Sorting Lossless Data Compression Algorithm", in *Technical Report, 124, Digital Equipment Corporation*: Gdask, Poland, 1994.

[84] K. Visweswariah, S. Kulkarni and S. Verdu, "Output distribution of the Burrows-Wheeler transform," in *2000 IEEE International Symposium on Information Theory* (Cat. No.00CH37060), Sorrento, Italy, 2000, pp. 53-, doi: 10.1109/ISIT.2000.866343.

[85] M. Effros, K. Visweswariah, S. R. Kulkarni and S. Verdu, "Universal lossless source coding with the Burrows Wheeler transform," in *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1061-1081, May 2002, doi: 10.1109/18.995542.

[86] B. Balkenhol and S. Kurtz, "Universal data compression based on the Burrows-Wheeler transformation: theory and practice," in *IEEE Transactions on Computers*, vol. 49, no. 10, pp. 1043-1053, Oct. 2000, doi: 10.1109/12.888040.

[87] G. Caire, S. Shamai, S. Verdu, "Universal data compression using LDPC codes", in *Proceedings of the International Symposium on Turbo Codes and Related Topics*, Brest, France, 1–5 September 2003.

[88] G. Caire, S. Shamai, A. Shokrollahi, S. Verdu, "Fountain Codes for Lossless Data Compression", in *DIMACS Series in Discrete Mathematics and Theoretical Computer Science; American Mathematical Society:* Providence, RI, USA, 2005; Volume 68.

[89] L. Wang and G. I. Shamir, "Universal Context Based Decoding with Low-Density Parity-Check Codes," in *IEEE Communications Letters*, vol. 11, no. 9, pp. 741-743, September 2007, doi: 10.1109/LCOMM.2007.070338.

[90] Kai Xie and G. I. Shamir, "Context and denoising based decoding of non-systematic turbo codes for redundant data," in *Proceedings. International Symposium on Information Theory*, 2005. ISIT 2005., Adelaide, SA, Australia, 2005, pp. 1280-1284, doi: 10.1109/ISIT.2005.1523548.

[91] J. Mitra and L. Lampe, "Coded narrowband transmission over noisy powerline channels," in *2009 IEEE International Symposium on Power Line Communications and Its Applications*, Dresden, Germany, 2009, pp. 143-148, doi: 10.1109/ISPLC.2009.4913419.

[92] M. Nassar, K. Gulati, Y. Mortazavi and B. L. Evans, "Statistical Modeling of Asynchronous Impulsive Noise in Powerline Communication Networks," in *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, Houston, TX, USA, 2011, pp. 1-6, doi: 10.1109/GLOCOM.2011.6134477.

[93] J. Bilbao, A. Calvo and I. Armendariz, "Fast characterization method and error sequence analysis for narrowband indoor powerline channel," in *2013 IEEE 17th International Symposium on Power Line Communications and Its Applications*, Johannesburg, South Africa, 2013, pp. 309-314, doi: 10.1109/ISPLC.2013.6525869.

[94] S. Gzelgz, H. B. elebi, T. Gzel, H. Arslan and M. K. Mhak, "Time frequency analysis of noise generated by electrical loads in PLC," in *2010 17th International Conference on Telecommunications*, Doha, Qatar, 2010, pp. 864-871, doi: 10.1109/ICTEL.2010.5478758.

[95] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," in *IEEE Transactions on Electromagnetic Compatibility*, vol. 44, no. 1, pp. 249-258, Feb. 2002, doi: 10.1109/15.990732.

[96] M. Nassar, A. Dabak, I. H. Kim, T. Pande and B. L. Evans, "Cyclostationary noise modeling in narrowband powerline communication for Smart Grid applications," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, 2012, pp. 3089-3092, doi: 10.1109/ICASSP.2012.6288568.

[97] M. Katayama, T. Yamazato, and H. Okada, "A mathematical Model of Noise in Narrowband PowerLine Communication Systems," in *IEEE Journal on Selected Areas in Communications*, pp. 1267-1276, July 2006. doi: 10.1109/JSAC.2006.874408.

[98] D. Middleton, "Statistical-physical models of electromagnetic interference," in *IEEE Transactions on Electromagnetic Compatibility*, vol. EMC-19, no. 3, pp. 106–127, 1977.

[99] D. Middleton, "Non-gaussian noise models in signal processing for telecommunications: new methods an results for class a and class b noise models," in *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1129–1149, 1999.

[100] G. Ndo, F. Labeau and M. Kassouf, "A Markov-Middleton Model for Bursty Impulsive Noise: Modeling and Receiver Design," in *IEEE Transactions on Power Delivery*, vol. 28, no. 4, pp. 2317-2325, Oct. 2013, doi: 10.1109/TPWRD.2013.2273942.

[101] A. Spaulding and D. Middleton, "Optimum Reception in an Impulsive Interference Environment - Part I: Coherent Detection," in *IEEE Transactions on Communications*, vol. 25, no. 9, pp. 910-923, September 1977, doi: 10.1109/TCOM.1977.1093943.

[102] A. Spaulding and D. Middleton, "Optimum Reception in an Impulsive Interference Environment - Part II: Incoherent Reception," in *IEEE Transactions on Communications*, vol. 25, no. 9, pp. 924-934, September 1977, doi: 10.1109/TCOM.1977.1093942.

[103] D. Umehara, H. Yamaguchi and Y. Morihiro, "Turbo decoding in impulsive noise environment," in *IEEE Global Telecommunications Conference*, 2004. GLOBECOM '04., Dallas, TX, USA, 2004, pp. 194-198 Vol.1, doi: 10.1109/GLOCOM.2004.1377938.

[104] P. Robertson, E. Villebrun and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proceedings IEEE International Conference on Communications ICC '95*, Seattle, WA, USA, 1995, pp. 1009-1013 vol.2, doi: 10.1109/ICC.1995.524253.

[105] H. Nakagawa, D. Umehara, S. Denno and Y. Morihiro, "A decoding for low density parity check codes over impulsive noise channels," in *International Symposium on Power Line Communications and Its Applications*, 2005., Vancouver, BC, Canada, 2005, pp. 85-89, doi: 10.1109/ISPLC.2005.1430471.

[106] T. Q. Bui and H. H. Nguyen, "Bit-Interleaved Coded Modulation with Iterative Decoding in Impulsive Noise," in *2006 IEEE International Symposium on Power Line Communications and Its Applications*, Orlando, FL, USA, 2006, pp. 98-103, doi: 10.1109/ISPLC.2006.247444.

[107] H. H. Nguyen and T. Q. Bui, "Bit-Interleaved Coded Modulation With Iterative Decoding in Impulsive Noise," in *IEEE Transactions on Power Delivery*, vol. 22, no. 1, pp. 151-160, Jan. 2007, doi: 10.1109/TPWRD.2006.881602.

[108] D. Middleton, "Procedures for Determining the Parameters of the First-Order Canonical Models of Class A and Class B Electromagnetic Interference [10]," in *IEEE Transactions on Electromagnetic Compatibility*, vol. EMC-21, no. 3, pp. 190-208, Aug. 1979, doi: 10.1109/TEMC.1979.303731.

[109] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley, 2006.

[110] N. Jaiswal and N. Purohit, "Performance Evaluation of Non-orthogonal Multiple Access in V2V communications Over Double-Rayleigh Fading Channels," *2019 IEEE Conference on Information and Communication Technology*, Allahabad, India, 2019, pp. 1-5, doi: 10.1109/CICT48419.2019.9066141.

[111] M. B. Janjua, D. B. da Costa and H. Arslan, "User Pairing and Power Allocation Strategies for 3D VLC-NOMA Systems," in *IEEE Wireless Communications Letters*, vol. 9, no. 6, pp. 866-870, June 2020, doi: 10.1109/LWC.2020.2973628.

[112] F. Ghanami, G. A. Hodtani, B. Vucetic and M. Shirvanimoghaddam, "Performance Analysis and Optimization of NOMA with HARQ for Short Packet Communications in Massive IoT," in *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2020.3028434.

[113] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li and K. Higuchi, "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, Dresden, Germany, 2013, pp. 1-5, doi: 10.1109/VTCSpring.2013.6692652.

[114] M. Al-Imari, P. Xiao, M. A. Imran and R. Tafazolli, "Uplink non-orthogonal multiple access for 5G wireless networks," in *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*, Barcelona, Spain, 2014, pp. 781-785, doi: 10.1109/ISWCS.2014.6933459.

[115] N. Otao, Y. Kishiyama and K. Higuchi, "Performance of non-orthogonal access with SIC in cellular downlink using proportional fair-based resource allocation," in *2012 International Symposium on Wireless Communication Systems (ISWCS)*, Paris, France, 2012, pp. 476-480, doi: 10.1109/ISWCS.2012.6328413.

[116] M. Elhattab, M. A. Arfaoui, C. Assi and A. Ghrayeb, "Reconfigurable Intelligent Surface Assisted Coordinated Multipoint in Downlink NOMA Networks," in *IEEE Communications Letters*, doi: 10.1109/LCOMM.2020.3029717.

[117] R. Hoshyar, F. P. Wathan and R. Tafazolli, "Novel Low-Density Signature for Synchronous CDMA Systems Over AWGN Channel," in *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1616-1626, April 2008, doi: 10.1109/TSP.2007.909320.

[118] M. Al-Imari, P. Xiao, M. A. Imran and R. Tafazolli, "Uplink non-orthogonal multiple access for 5G wireless networks," in *2014 11th International Symposium on Wireless Communications Systems (ISWCS)*, Barcelona, 2014, pp. 781-785, doi: 10.1109/ISWCS.2014.6933459.

[119] H. Nikopour and H. Baligh, "Sparse code multiple access," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, London, 2013, pp. 332-336, doi: 10.1109/PIMRC.2013.6666156.

[120] Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang and J. Xu, "Multi-User Shared Access for Internet of Things," *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, Nanjing, 2016, pp. 1-5, doi: 10.1109/VTCSpring.2016.7504361.

[121] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," in *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74-81, September 2015, doi: 10.1109/MCOM.2015.7263349.

[122] S. Verdú, "Multiuser Detection". *Cambridge University Press*, 1998.

[123] J. A. de Carvalho, D. B. da Costa, F. R. M. Lima, R. Oliveira and U. S. Dias, "Non-Orthogonal Multiple Access in Two-Hop Wireless Powered Communication Networks," in *IEEE Wireless Com-*

munications Letters, vol. 9, no. 9, pp. 1398-1402, Sept. 2020, doi: 10.1109/LWC.2020.2992049.

[124] J. A. de Carvalho, D. B. da Costa, L. Yang, G. C. Alexandropoulos, R. Oliveira and U. S. Dias, "User Fairness in Wireless Powered Communication Networks with Non-Orthogonal Multiple Access," in *IEEE Wireless Communications Letters*, doi: 10.1109/LWC.2020.3030818.

[125] S. M. R. Islam, N. Avazov, O. A. Dobre and K. Kwak, "Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721-742, Secondquarter 2017, doi: 10.1109/COMST.2016.2621116.

[126] K. Chung, "NOMA for Correlated Information Sources in 5G Systems," in *IEEE Communications Letters*, doi: 10.1109/LCOMM.2020.3027726.

[127] H. Pan, L. Lu and S. C. Liew, "Practical Power-Balanced Non-Orthogonal Multiple Access," in *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2312-2327, Oct. 2017, doi: 10.1109/JSAC.2017.2727198.

[128] Y. Zhang, K. Peng, S. Chen and J. Song, "A capacity-approaching multi-user BICM-ID scheme for multiple access channel," in *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, Dubrovnik, Croatia, 2015, pp. 852-856, doi: 10.1109/IWCMC.2015.7289194.

[129] M. Gastpar, "Gaussian multiple-access channels under received-power constraints," in *Information Theory Workshop, San Antonio, TX*, USA, 2004, pp. 452-457, doi: 10.1109/ITW.2004.1405346.

[130] T. Cover, A. E. Gamal and M. Salehi, "Multiple access channels with arbitrarily correlated sources," in *IEEE Transactions on Information Theory*, vol. 26, no. 6, pp. 648-657, November 1980, doi: 10.1109/TIT.1980.1056273.

[131] J. Garcia-Frias, "Joint Source-Channel Decoding of Correlated Sources over Noisy Channels," in *Proceedings Data Compression Conference*, 2001, Snowbird, Utah, March 2001. doi: 10.1109/DCC.2001.917159

[132] W. Zhong, Y. Zhao and J. Garcia-Frias, "Turbo-like codes for distributed joint source-channel coding of correlated senders in multiple access channels," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003, Pacific Grove, CA, USA, 2003, pp. 840-844 Vol.1, doi: 10.1109/ACSSC.2003.1292031.

[133] J. Garcia-Frias and Ying Zhao, "Near-Shannon/Slepian-Wolf performance for unknown correlated sources over AWGN channels," in *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 555-559, April 2005, doi: 10.1109/TCOMM.2005.844959.

[134] W. Zhong and J. Garcia-Frias, "LDGM Codes for Channel Coding and Joint Source Channel Coding of Correlated Sources," in *EURASIP Journal on Applied Signal Processing*, pp. 942-953, May 2005. doi: 10.1155/ASP.2005.942

[135] W. Zhong and J. Garcia-Frias, "Joint Source-Channel Coding of Correlated Senders over Multiple Access Channels", Allerton Conference on Communication, Control, and Computing, October 2004, Allerton, Illinois (invited paper).

[136] W. Zhong and J. Garcia-Frias, "LDGM Codes for Transmission of Correlated Senders over MAC", Allerton Conference on Communication, Control, and Computing, October 2005, Allerton, Illinois (invited paper).

[137] Y. Zhao, W. Zhong, and J. Garcia-Frias, "Transmission of Correlated Senders over a Rayleigh Fading Multiple Access Channel," in *Signal Processing*, vol. 86, pp. 3150-3159, November 2006. doi: 10.1016/j.sigpro.2006.03.014.

[138] W. Zhong and J. Garcia-Frias, "Parallel LDGM Codes for the Transmission of Highly Correlated Senders over Rayleigh Fading Multiple Access Channels," in *2006 40th Annual Conference on Information Sciences and Systems*, Princeton, NJ, USA, 2006, pp. 1230-1235, doi: 10.1109/CISS.2006.286653.

[139] J. Garcia-Frias, Y. Zhao and W. Zhong, "Turbo-Like Codes for Transmission of Correlated Sources over Noisy Channels," in *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 58-66, Sept. 2007, doi: 10.1109/MSP.2007.904813.

[140] M. Sartipi and F. Fekri, "Source and channel coding in wireless sensor networks using LDPC codes," in *2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, 2004. IEEE SECON 2004., Santa Clara, CA, USA, 2004, pp. 309-316, doi: 10.1109/SAHCN.2004.1381931.

[141] Q. Xu, V. Stankovic and Z. Xiong, "Distributed joint source-channel coding of video using raptor codes," in *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 4, pp. 851-861, May 2007, doi: 10.1109/JSAC.2007.070520.

[142] X. Dai et al., "Successive interference cancelation amenable multiple access (SAMA) for future wireless communications," in *2014 IEEE International Conference on Communication Systems*, Macau, China, 2014, pp. 222-226, doi: 10.1109/ICCS.2014.7024798.

[143] J. Huang, K. Peng, C. Pan, F. Yang and H. Jin, "Scalable Video Broadcasting Using Bit Division Multiplexing," in *IEEE Transactions on Broadcasting*, vol. 60, no. 4, pp. 701-706, Dec. 2014, doi: 10.1109/TBC.2014.2361471.

[144] K. Kusume, G. Bauch and W. Utschick, "IDMA vs. CDMA: Analysis and Comparison of Two Multiple Access Schemes," in *IEEE Transactions on Wireless Communications*, vol. 11, no. 1, pp. 78-87, January 2012, doi: 10.1109/TWC.2011.111211.100954.

[145] Li Ping, Lihai Liu, Keying Wu and W. K. Leung, "Interleave division multiple-access," in *IEEE Transactions on Wireless Communications*, vol. 5, no. 4, pp. 938-947, April 2006, doi: 10.1109/TWC.2006.1618943.

[146] M. Gonzalez-Lopez, F. J. Vazquez-Araujo, L. Castedo and J. Garcia-Frias, "Interleave-division multiple access (IDMA) using low-rate layered LDGM codes," in *2008 5th International Symposium on Turbo Codes and Related Topics*, Lausanne, 2008, pp. 315-320, doi: 10.1109/TURBOCODING.2008.4658718.

[147] Y. Chen, J. Schaepperle and T. Wild, "Comparing IDMA and NOMA with superimposed pilots based channel estimation in uplink," in *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Hong Kong, 2015, pp. 89-94, doi: 10.1109/PIMRC.2015.7343274.

[148] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," in *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471-480, July 1973, doi: 10.1109/TIT.1973.1055037.

[149] T. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources (Corresp.)," in *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 226-228, March 1975, doi: 10.1109/TIT.1975.1055356.

[150] A.D. Wyner, "Recent Results in Shannon Theory." in *IEEE Transactions on Information Theory*, vol. 20, no. 1, pp. 2 - 10, January 1974.

[151] S. Shamai and S. Verdu, "Capacity of channels with uncoded-message side-information," in *Proceedings of 1995 IEEE International Symposium on Information Theory*, Whistler, BC, Canada, 1995, pp. 7-, doi: 10.1109/ISIT.1995.531109.

[152] J. Garcia-Frias and F. Cabarcas, "Approaching the slepian-wolf boundary using practical channel codes," in *International Symposium on Information Theory*, 2004. ISIT 2004. Proceedings., Chicago, IL, USA, 2004, pp. 330-330, doi: 10.1109/ISIT.2004.1365367.

[153] W. Zhong, H. Lou and J. Garcia-Frias, "LDGM codes for joint source-channel coding of correlated sources," in *Proceedings 2003 International Conference on Image Processing* (Cat. No.03CH37429), Barcelona, Spain, 2003, pp. I-593, doi: 10.1109/ICIP.2003.1247031.

[154] Wei Zhong, Huiqiong Chai and J. Garcia-Frias, "Approaching the shannon limit through parallel concatenation of regular LDGM codes," in *Proceedings. International Symposium on Information Theory*, 2005. ISIT 2005., Adelaide, SA, Australia, 2005, pp. 1753-1757, doi: 10.1109/ISIT.2005.1523646.

[155] X. Zhu, L. Zhang and Y. Liu, "A distributed joint source-channel coding scheme for multiple correlated sources," in *2009 Fourth International Conference on Communications and Networking in China*, Xi'an, China, 2009, pp. 1-6, doi: 10.1109/CHINACOM.2009.5339955.

[156] I. Shahid and P. Yahampath, "Distributed joint source-channel coding of correlated binary sources in wireless sensor networks," in *2011 8th International Symposium on Wireless Communication Systems*, Aachen, 2011, pp. 236-240, doi: 10.1109/ISWCS.2011.6125345.

[157] A. Abrardo, G. Ferrari, M. Martalo, M. Franceschini and R. Raheli, "Optimizing channel coding for orthogonal multiple access schemes with correlated sources," in *2009 Information Theory and Applications Workshop*, San Diego, CA, 2009, pp. 5-14, doi: 10.1109/ITA.2009.5044916.

[158] F. Pujaico and J. Portugheis, "Optimal Rate for Joint Source-Channel Coding of Correlated Sources Over Orthogonal Channels," in *IEEE Communications Letters*, vol. 19, no. 1, pp. 22-25, Jan. 2015, doi: 10.1109/LCOMM.2014.2377237.

[159] A. Abrardo, G. Ferrari, M. Martaló, M. Franceschini and R. Raheli, "Orthogonal Multiple Access With Correlated Sources: Achievable Region and Pragmatic Schemes," in *IEEE Transactions on Communications*, vol. 62, no. 7, pp. 2531-2543, July 2014, doi: 10.1109/TCOMM.2014.2325039.

[160] J. Garcia-Frias and Ying Zhao, "Near-Shannon/Slepian-Wolf performance for unknown correlated sources over AWGN channels," in *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 555-559, April 2005, doi: 10.1109/TCOMM.2005.844959.

[161] D. Declercq and M. Fossorier, "Decoding Algorithms for Nonbinary LDPC Codes Over GF(q)," in *IEEE Transactions on Communications*, vol. 55, no. 4, pp. 633-643, April 2007, doi: 10.1109/TCOMM.2007.894088.

[162] V. Savin, "Min-Max decoding for non binary LDPC codes," in *2008 IEEE International Symposium on Information Theory, Toronto, ON*, 2008, pp. 960-964, doi: 10.1109/ISIT.2008.4595129.

[163] V. S. Ganepola, R. A. Carrasco, I. J. Wassell and S. Le Goff, "Performance study of non-binary LDPC Codes over GF(q)," in *2008 6th International Symposium on Communication Systems, Networks and Digital Signal Processing, Graz*, 2008, pp. 585-589, doi: 10.1109/CSNDSP.2008.4610743.

[164] Z. Babar, P. Botsinis, D. Alanis, S. X. Ng and L. Hanzo, "Fifteen Years of Quantum LDPC Coding and Improved Decoding Strategies," in *IEEE Access*, vol. 3, pp. 2492-2519, 2015, doi: 10.1109/ACCESS.2015.2503267.

[165] P. Fuentes, J. Etxezarreta Martinez, P. M. Crespo, and J. Garcia-Frias, "Approach for the construction of non-Calderbank-Steane-Shor low-density-generator-matrix based quantum codes," in *Phys. Rev. A*, vol. 102, pp. 012423, 2020. doi: 10.1103/PhysRevA.102.012423

[166] P. Fuentes, J. E. Martinez, P. M. Crespo and J. Garcia-Frias, "Performance of non-CSS LDGM-based quantum codes over the misidentified depolarizing channel," in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Denver, CO, USA, 2020, pp. 93-101, doi: 10.1109/QCE49297.2020.00022.

[167] P. Fuentes, J. Etxezarreta Martinez, P. M. Crespo, and J. Garcia-Frias, "Design of low-density-generator-matrixbased quantum codes for asymmetric quantum channels", in *Phys. Rev. A*, vol. 103, 2021. doi: 10.1103/PhysRevA.103.022617

[168] Tao Tian, C. R. Jones, J. D. Villasenor and R. D. Wesel, "Selective avoidance of cycles in irregular LDPC code construction," in *IEEE Transactions on Communications*, vol. 52, no. 8, pp. 1242-1247, Aug. 2004, doi: 10.1109/TCOMM.2004.833048.

[169] J. A. McGowan and R. C. Williamson, "Loop removal from LDPC codes," in *Proceedings 2003 IEEE Information Theory Workshop (Cat. No.03EX674)*, Paris, France, 2003, pp. 230-233, doi: 10.1109/ITW.2003.1216737.

[170] J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", *Morgan Kaufmann Publishers Inc.*, San Francisco, CA, USA, 1988.

[171] A. Bennatan and D. Burshtein, "Design and analysis of nonbinary LDPC codes for arbitrary discrete-memoryless channels," in *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 549-583, Feb. 2006, doi: 10.1109/TIT.2005.862080.

[172] Ser, J.D., Crespo, P.M. & Galdos, O., "Asymmetric Joint Source-Channel Coding for Correlated Sources with Blind HMM Estimation at the Receiver", in *J Wireless Com Network 2005*, 357402 (2005). https://doi.org/10.1155/WCN.2005.483