Thesis for the Degree of Doctor of Philosophy

Bit-Wise Decoders for Coded Modulation and Broadcast Coded Slotted ALOHA

Mikhail Ivanov

CHALMERS

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology

Gothenburg, Sweden, 2016

Bit-Wise Decoders for Coded Modulation
and Broadcast Coded Slotted ALOHA
Mikhail Ivanov

ISBN 978-91-7597-385-2
Doktorsavhandlingar vid Chalmers tekniska högskola
Series No. 4066
ISSN 0346-718X

This thesis has been prepared using LaTeX and BibTeX.

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Telephone: + 46 (0)31-772 1000
www.chalmers.se

Printed by Chalmers Reproservice,
Gothenburg, Sweden, August 2016.

Preface

It happened that this thesis consists of two, seemingly unrelated, topics. Let me explain why. When I applied for my PhD position at Chalmers, my research project was described in the advertisement as follows.

"This research project deals with the evaluation and optimization of coded modulation systems, with particular emphasis on BICM and BICM-MIMO. The systems will be optimized theoretically and numerically and their performance will be evaluated by analysis and simulations. The project is aimed at developing information-theoretic and coding bounds to analyze BICM and BICM-MIMO. The theoretical analysis will be used to derive optimization criteria for the design of practical low-complexity schemes achieving high diversity and high coding gain."

The bit-interleaved coded modulation (BICM) part in the advertisement was due to Alex Alvarado, a PhD student at Chalmers at that time, who was very active in the area of coded modulation. The multiple-input multiple-output (MIMO) part was there since Fredrik Brännström, my main supervisor, was involved in a few projects related to vehicular communications. MIMO seemed to be a promising technique to improve the channel between cars by providing spacial diversity.

I started to work on BICM planning to extend the results to BICM-MIMO for the use in vehicular networks later on. This, however, did not happen. After defending my Licentiate entitled "On bit-wise decoders for coded modulation", came the time to think about extending the results to the vehicular scenario. Unfortunately, the lack of channel models that could be used for designing MIMO schemes rendered the MIMO track not very appealing. So we started to look for alternative ways to increase diversity and resorted to channel coding. For example, a rate-1/2 repetition code when the data is repeated twice provides the diversity order of two.

Following this track leads to lowering the code rate, which results in the increase of the channel load. In vehicular communications, this is highly undesirable since the channel is expected to be heavily loaded. Hence, exploiting time diversity in a vehicular network calls for the design of a new medium access control protocol. Reasoning along these lines, we came to the conclusion that coded slotted ALOHA (CSA) is the solution that can provide the desired diversity and handle the increased channel load. Adapting CSA to the vehicular scenario became the second part of my PhD and resulted in broadcast coded slotted ALOHA.

Abstract

This thesis deals with two aspects of wireless communications. The first aspect is about efficient point-to-point data transmission. To achieve high spectral efficiency, coded modulation, which is a concatenation of higher order modulation with error-correcting coding, is used. Bit-interleaved coded modulation (BICM) is a pragmatic approach to coded modulation, where soft information on encoded bits is calculated at the receiver and passed to a bit-wise decoder. Soft information is usually obtained in the form of log-likelihood ratios (also known as L-values), calculated using the max-log approximation. In this thesis, we analyze bit-wise decoders for pulse-amplitude modulation (PAM) constellations over the additive white Gaussian noise (AWGN) channel when the max-log approximation is used for calculating L-values.

First, we analyze BICM systems from an information theoretic perspective. We prove that the max-log approximation causes information loss for all PAM constellations and labelings with the exception of a symmetric 4-PAM constellation labeled with a Gray code. We then analyze how the max-log approximation affects the generalized mutual information, which is an achievable rate for a standard BICM decoder. Second, we compare the performance of the standard BICM decoder with that of the ML decoder. We show that, when the signal-to-noise ratio goes to infinity, the loss in terms of pairwise error probability is bounded by 1.25 dB for any two codewords. The analysis further shows that the loss is zero for a wide range of linear codes.

The second aspect of wireless communications treated in this thesis is multiple channel access. Our main objective here is to provide reliable message exchange between nodes in a wireless ad hoc network with stringent delay constraints. To that end, we propose an uncoordinated medium access control protocol, termed all-to-all broadcast coded slotted ALOHA (B-CSA), that exploits coding over packets at the transmitter side and successive interference cancellation at the receiver side. The protocol resembles low-density parity-check codes and can be analyzed using the theory of codes on graphs. The packet loss rate performance of the protocol exhibits a threshold behavior with distinct error floor and waterfall regions. We derive a tight error floor approximation that is used for the optimization of the protocol. We also show how the error floor approximation can be used to design protocols for networks, where users have different reliability requirements. We use B-CSA in vehicular networks and show that it outperforms carrier sense multiple access currently adopted as the medium access control protocol for vehicular communications. Finally, we investigate the possibility of establishing a handshake in vehicular networks by means of B-CSA.

Keywords: All-to-all broadcast, bit-interleaved coded modulation, error floor, medium access control, mutual information, slotted ALOHA, vehicular communications.

Acknowledgements

This thesis is a short summary of the five amazing years I spent at Chalmers. Doing a PhD was a fascinating experience of learning new things and meeting nice people. I owe a dept of gratitude to many of them for the great time I had, the support I received throughout the years, and the invaluable contributions to this thesis.

First of all, I want to express my sincere gratitude to my main supervisor Professor Fredrik Brännström and my co-supervisor Professor Alexandre Graell i Amat. If not for them, I would not have become a PhD student and this thesis would not have been written. I would like to thank them for the excellent supervision, their patience, and their open-mindedness. They have never said "no" to me, whether I was planning to change research topics or to stay at home with my newborn son while writing the thesis. I would also like to thank Professor Erik Agrell and Dr. Alex Alvarado for introducing me to the world of BICM. Alex and Erik were my co-supervisors for the first part of this thesis related to coded modulation.

Furthermore, I want to express my gratitude to Professor Petar Popovski for invaluable contributions to the second part of this thesis. Petar is a person with infinite energy and millions of ideas, which he freely shares with others. I am very grateful for the chance to work with him and visit him at Aalborg University. Also, I would like to thank Dr. Gianluigi Liva for inviting me to visit his group at DLR, providing many brilliant insights into iteratively decoded codes, and just being so nice.

I would further like to thank all former and present members of the Communication Systems Group at Chalmers for the amazing work environment. It was a pleasure to come to work every day. I would like to give special thanks to my office mate Dr. Christian Häger for contributing to this thesis and being a great friend I can always count on. I would also like to thank the administrative staff at S2 for making my life at Chalmers very smooth and hassle-free. Finally, I would like to thank Petar's group at Aalborg University and Gianluigi's group at DLR for the fun time I had when I visited them.

Mikhail Ivanov
Gothenburg, July 2016

This work has been supported in part by the Swedish Research Council, under Grants No. 2011-5950 and 2011-5961, Chalmers Antenna Systems Excellence Center in the project 'Antenna Systems for V2X Communication', and the European Research Council, under Grant No. 258418 (COOPNET). I would also like to thank the Ericsson's Research Foundation for supporting my research visits to the University of Cambridge, Aalborg University, and the German Aerospace Center (DLR).

List of Publications

The thesis is based on the following publications:

Paper A
M. Ivanov, C. Häger, F. Brännström, A. Graell i Amat, A. Alvarado, and E. Agrell, "On the Information Loss of the Max-Log Approximation in BICM Systems," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3011–3025, June 2016.

Paper B
M. Ivanov, A. Alvarado, F. Brännström, and E. Agrell, "On the Asymptotic Performance of Bit-Wise Decoders for Coded modulation," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2796–2804, May 2014.

Paper C
M. Ivanov, F. Brännström, A. Graell i Amat, and P. Popovski, "Broadcast Coded Slotted ALOHA: A Finite Frame Length Analysis," submitted to *IEEE Trans. Commun.*

Paper D
M. Ivanov, F. Brännström, A. Graell i Amat, and G. Liva, "Unequal Error Protection in Coded Slotted ALOHA," submitted to *IEEE Wir. Commun. Lett.*

Paper E
M. Ivanov, P. Popovski, F. Brännström, A. Graell i Amat, and Č. Stefanović, "Probabilistic Handshake in All-to-all Broadcast Coded Slotted ALOHA," in *Proc. IEEE Int. Work. Signal Proc. Advances Wir. Commun.*, Stockholm, Sweden, June 2015.

Related publications by the author not included in the thesis:

M. Ivanov, F. Brännström, A. Alvarado, and E. Agrell, "General BER Expression for One-Dimensional Constellations," in *Proc. IEEE Global Commun. Conf.*, Anaheim, CA, Dec. 2012.

M. Ivanov, F. Brännström, A. Alvarado, and E. Agrell, "On the Exact BER of Bit-Wise Demodulators for One-Dimensional Constellations", *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1450–1459, Apr. 2013.

M. Ivanov, F. Brännström, A. Graell i Amat, and P. Popovski, "Error Floor Analysis of Coded Slotted ALOHA over Packet Erasure Channels," *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 419–422, Mar. 2015.

M. Ivanov, F. Brännström, A. Graell i Amat, and P. Popovski, "All-to-all Broadcast for Vehicular Networks Based on Coded Slotted ALOHA," in *Proc. IEEE Int. Conf. Commun. Work.*, London, UK, June 2015.

Contents

Part I

Introduction

Chapter 1

Preliminaries

1.1 Coded Modulation

Coded modulation (CM) is referred to as a technique that combines error-correcting coding with higher-order modulation in order to achieve high spectral efficiency. Its story began in 1948, when Claude Shannon introduced a general communication system model for point-to-point communications, formulated the problem of reliable data transmission, and provided a fundamental limit of error free transmission called channel capacity [1]. In his work that marked the beginning of information theory, Shannon already introduced the technique we now call CM. He even suggested a hypothetical solution for CM to attain the channel capacity. His solution included four main ingredients: a constellation, a code (or codebook), an assignment of the messages to the codewords (or encoding), and a decoding algorithm. The four ingredients described by Shannon lead to solutions which are highly complex and impractical. Therefore, coding theorists and communication engineers have spent a great deal of effort to propose alternative ways to approach capacity.

The first step in order to make the coding scheme more realistic was to use discrete constellations. Once a discrete constellation is chosen, labels (binary or nonbinary) can be assigned to the constellation points. By doing so, the codebook construction and encoding can be done in two stages, which are usually regarded as error-correcting coding and modulation.

For a long time engineers were mainly concerned with the power-limited regime and the main focus was on achieving capacity using binary phase shift keying (BPSK). With the increase of the demand for high date rates, the bandwidth-limited regime came into play. The first approaches to combine higher-order modulation, where binary labels are assigned to the constellation points, with classical codes designed for binary transmission showed very disappointing results, which are well described in Ungerboeck's paper [2]. This even resulted in questioning whether coding is relevant for high spectral efficiencies [3]. The problem with those schemes was actually not a poor design but rather bad decoding techniques adopted from algebraic coding theory. In fact, the design of the scheme that Ungerboeck describes as a "bad scheme" follows the bit-interleaved coded modulation (BICM) approach and corresponds to a good trellis-coded modulation (TCM) scheme.

The second step towards a good practical CM scheme was to impose a very regular structure on the codebook. In [4] and [5], the technique called TCM was proposed, where

convolutional codes (CCs), appropriately tailored to the constellation, are used to obtain codes with a trellis structure. This resulted in easy encoding and, most importantly, it enabled implementation of a "complicated" maximum likelihood (ML) decoder with a reasonable complexity by means of the Viterbi algorithm [6]. In those works, the importance of the Euclidean distance (ED) between codewords for the code performance was highlighted and a set of rules to design TCM systems with a large minimum ED was formulated based on the so-called set partitioning (SP) technique [5]. However, very structured codes are in general unable to achieve the performance predicted by information theory.

The third and most successful step to make a good practical CM scheme was to allow the code to have little structure and to develop good suboptimal decoding algorithms. One of the first examples of such CM systems is multilevel coded modulation (MLCM) proposed by Imai and Hirakawa [7]. The main idea was to use different binary codes for different bit positions in the constellation, which made the code less structured compared to TCM, together with a sub-optimal multi-stage decoder, where the binary decoders for the corresponding binary codes are allowed to exchange information between one another.

Another important example of the development in this direction is BICM [8]. Zehavi suggested in [8] to place bit-wise interleavers between the encoder and the modulator and to use a suboptimal bit-wise decoder. The original motivation behind introducing interleavers was to improve the performance over fast fading channels. For such channels, the most important parameter of the code is the so-called code diversity whereas the ED is secondary. The code diversity is the minimum number of different symbols between any two codewords of the code and, in a sense, is equivalent to time diversity. Interleavers were shown to increase the code diversity making BICM very attractive for fading channels. At the same time, BICM appeared to be very powerful over the AWGN channel. It is currently used in various communications standards, e.g., [9–11] to name a few.

Further developments in coding theory led to iterative decoding algorithms, powerful binary and nonbinary turbo [12] and low-density parity-check (LDPC) [13, 14] codes. This, in turn, resulted in powerful and efficient CM schemes, such as turbo TCM [15, 16] and BICM with iterative decoding (BICM-ID) [17, 18]. Even though BICM has been a research topic for more than 20 years, there are still many unsolved problems, e.g., those related to interleaver design, performance analysis for finite interleaver length, maximum achievable rates, etc. In early BICM systems, where convolutional codes were used, as well as in modern coded systems that use, for instance, LDPC codes, the calculation of log-likelihood ratios (LLRs or L-values) is one of the key operations. The knowledge about L-values and their properties can be used to obtain precise performance estimations, new design criteria, and may help to solve some of the open problems. In the following chapters, we take a closer look at different properties of the L-values.

1.2 Coded Random Access

It is interesting to see the change that is currently happening in wireless communications.

Modern wireless standards, such as 4G are capable of satisfying most of the demands for high data rates of classical applications, which include voice, video, and data transmission, as well video streaming. However, the technological advances give rise to new applications with new requirements that put the data rate aside. Future communications systems are expected to include different types of machines, sensors and devices creating the so-called Internet of Things. Hence, new wireless systems need to support a massive number of devices satisfying stringent reliability and latency constraints [19–21].

One such application is vehicular communications (VCs), in which cars (and roadside units) exchange messages in order to prevent accidents, improve transportation systems, and provide a greener environment. Although it has been a research topic for a long time (first standards are dated 1999 [22]), it seems like only now we start to get the tools to address all the challenges of VCs.

When a wireless communication network consists of more than one transmitter, referred to as users in the following, the problem of managing multiple transmissions arises. The wireless medium is by nature broadcast, hence, transmitted signals are received by all receivers within the reach. Moreover, simultaneous transmissions will collide at a receiver potentially preventing successful reception. Such media are usually referred to as a multiple access channel. The practical aspects of multiple access are handled by medium access control (MAC) protocols that specify rules for transmission over a multiple access channel. In this thesis, we deal with MAC protocols for wireless ad hoc networks. The application that we have in mind here is an exchange of safety-critical messages in vehicular networks.

MAC protocols can be generally divided into two groups [23]. The first group consists of coordinated protocols where a central unit moderates the access to the channel. The second group includes uncoordinated protocols based on random access. In vehicular networks, assigning the functions of the central unit to one of the nodes in the network is not possible making random access the only option.

Random access protocols can be further divided into two groups [24]. The first one is based on carrier sense [25], where a user senses the channel before transmission trying to avoid collisions. The second group of random access protocols is based on ALOHA principles [26] where users transmit at random hoping that their packet gets through. Up until recently, protocols based on ALOHA could not provide satisfying performance and carrier sense multiple access (CSMA) with collision avoidance was adopted as a MAC protocol for vehicular networks [27].

Indeed, the first version of the ALOHA protocol [26] provided very low utilization with only up to 18 percent of the time used for successful transmissions. In ALOHA, users transmit immediately when they have data to be sent. Introducing a slotted time structure and forcing users to differ their transmission until the beginning of the next slot led to slotted ALOHA [28]. The slotted structure avoids partial collisions of packets and improves the performance up to 36 percent utilization of the time. Further improvement was achieved in diversity slotted ALOHA [29], where users select several slots at random, thus increasing the probability of a successful transmission. The gains of diversity slotted

ALOHA are however quite modest.

Recent advantages in signal processing breathed new life into ALOHA protocols by allowing to exploit collisions by means of successive interference cancellation (SIC). This gave rise to a family of protocols referred to as coded slotted ALOHA (CSA) in this thesis. The general idea of CSA is as follows. As in diversity slotted ALOHA, a user transmits several packets. Every packet contains pointers to its copies, so that, once a packet is successfully decoded, full information about the location of the copies is available. The receiver decodes the packets without collision, obtains the location of their copies, and subtracts the interference caused by the copies. The process continues until the receiver cannot find decodable packets without collision.

The first protocol following this principle was proposed in [30] and is called contention resolution diversity slotted ALOHA, in which the repetition factor (typically 2 or 3) is fixed for all users. In [31], it was noted that allowing different repetition rates for different users leads to great performance improvements. The proposed protocol, termed irregular repetition slotted ALOHA, is shown to potentially provide full utilization of the resources at the expense of a large number of transmissions. Finally, a more general protocol that includes the aforementioned ones and dubbed CSA was proposed in [32]. In this protocol, the transmitted packets are generated using a rate-k/n code. To successfully receive the data and reconstruct the copies, the receiver needs to receive k out of n packets. CSA allows to achieve high throughputs under repetition rate constrains by assigning different code rates to different users.

In this thesis, we study the potential of CSA when used in VC. We show that CSA fits very well the type of communication and traffic in vehicular networks and provides high reliability with predictable delays.

1.3 Structure of the Thesis

A doctoral student is given two options when writing a PhD thesis: a classical monograph or a collection of papers published by the student. This thesis is written as a collection of papers and is divided into two parts. Part I gives a general overview of CM techniques and MAC protocols and prepares the reader for the papers that come in Part II. Part I is organized as follows. Chapter 2 presents an information theoretic basis of BICM. In Chapter 3, practical aspects of CM are discussed. Chapter 4 introduces VCs and Chapter 5 provides the basics of CSA. Finally, Chapter 6 gives an overview of the contributions made by the author.

1.4 Notational Convention

The following notation is used in Part I of the thesis. Lowercase letters x denote real or complex scalars and boldface letters $\boldsymbol{x} = [x_1, x_2, \ldots, x_N]$ denote row vectors of scalars of length N. Blackboard bold letters $\mathbb{X}$ denote matrices with elements $x_{i,j}$ in the ith row

and the jth column and $(\cdot)^\intercal$ denotes transposition. Calligraphic capital letters $\mathcal{X}$ denote sets, where $|\mathcal{X}|$ denotes its cardinality. The set of real numbers is denoted by $\mathcal{R}$. Binary addition (exclusive-OR) of two bits a and b is denoted by $a \oplus b$. Random variables (RVs) are denoted by capital letters X, probabilities by $\Pr\{\cdot\}$, the probability density function (PDF) of a continuous RV X by $f_X(x)$, and the conditional PDF of Y conditioned on $X = x$ by $f_{Y|X}(y|x)$. The probability mass function (PMF) of a discrete RV X is denoted by $p_X(x)$. The Gaussian distribution with mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$ and the Gaussian Q-function is denoted by $Q(\cdot)$. The expected value of a function $f(\cdot)$ of a RV X is denoted by $\mathrm{E}\{f(X)\}$. Please note that the appended papers do not always follow the introduced convention.

Chapter 2

Information Theoretic Approach to Coded Modulation

2.1 System Model

In this chapter, we analyze a point-to-point communication system model introduced by Shannon in [1], and shown in Fig. 2.1. The communication problem consists in transmitting a message m, represented by an integer, through a noisy channel to the receiver that produces an estimate of the message $\hat{\mathsf{m}}$ based on the observation vector $\boldsymbol{y}$. The message is encoded into a vector $\boldsymbol{x}_{\mathsf{m}}$ (codeword) of N symbols from a certain alphabet $\mathcal{S}$ (constellation), where N is the number of channel uses. The messages are selected by the transmitter uniformly at random from the set $\mathcal{M}$. Each message m is labeled with a unique binary vector $\boldsymbol{c}_{\mathsf{m}}$ of length $K = \lceil \log_2(|\mathcal{M}|) \rceil$ and is said to convey K bits of information. The data rate at which the system transmits information is defined as

$$R = \frac{K}{N} \tag{2.1}$$

and is measured in bits per channel use (bpcu). The collection of all codewords is called a codebook. Together with the mapping $\mathsf{m} \to \boldsymbol{x}_{\mathsf{m}}$ and the decoding algorithm $\boldsymbol{y} \to \hat{\mathsf{m}}$ the codebook forms a coding scheme.

We define two performance measures for the described system. The codeword error rate (CER) is defined as

$$\mathrm{CER} = \Pr\{\hat{\mathsf{M}} \neq \mathsf{M}\} = \frac{1}{|\mathcal{M}|} \sum_{\mathsf{m} \in \mathcal{M}} \sum_{\hat{\mathsf{m}} \in \mathcal{M} \setminus \mathsf{m}} \Pr\{\hat{\mathsf{M}} = \hat{\mathsf{m}} | \mathsf{M} = \mathsf{m}\} \tag{2.2}$$

and shows the probability that the estimated message does not coincide with the transmitted one. In this chapter, we focus on the CER. In practical applications, however, it is often easier to measure the probability that an information bit is in error given by the bit-error rate (BER) and defined as

$$\mathrm{BER} = \frac{1}{K|\mathcal{M}|} \sum_{\mathsf{m} \in \mathcal{M}} \sum_{\hat{\mathsf{m}} \in \mathcal{M} \setminus \mathsf{m}} d_{\mathrm{H}}(\boldsymbol{c}_{\mathsf{m}}, \boldsymbol{c}_{\hat{\mathsf{m}}}) \Pr\{\hat{\mathsf{M}} = \hat{\mathsf{m}} | \mathsf{M} = \mathsf{m}\}, \tag{2.3}$$

where $d_{\mathrm{H}}(\boldsymbol{b}_{\mathsf{m}}, \boldsymbol{b}_{\hat{\mathsf{m}}})$ is the Hamming distance between the two binary vectors.

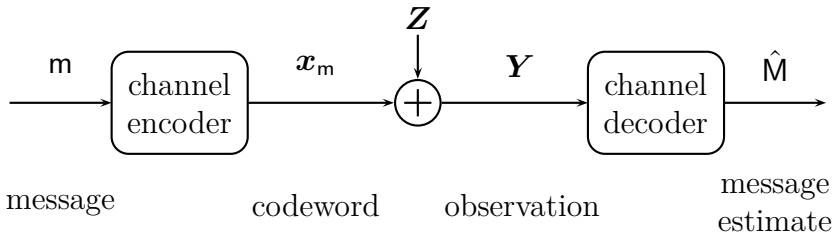


Figure 2.1: A block diagram of a general digital communication system.

The main objective of channel coding is to design a coding scheme that provides high data rate with as small probability or error as possible under various constraints. In this chapter, we consider the average power constraint, i.e., the energy per channel use E_{s}, defined as

$$E_{\mathrm{s}} = \frac{\sum_{\mathsf{m}\in|\mathcal{M}|} \|\boldsymbol{x}_{\mathsf{m}}\|^2}{N\mathcal{M}}, \tag{2.4}$$

has to satisfy

$$E_{\mathrm{s}} \leq 1. \tag{2.5}$$

2.1.1 Channel Model

We consider a discrete time additive white Gaussian noise (AWGN) channel with output $Y = x+Z$, where $x \in \mathcal{S}$ and the noise sample Z is a zero-mean Gaussian RV with variance σ^2. The conditional PDF of the channel output, given the channel input, is

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(y-x)^2}{2\sigma^2}}, \tag{2.6}$$

and we define the signal-to-noise ratio (SNR) as

$$\gamma \triangleq E_{\mathrm{s}}/\sigma^2. \tag{2.7}$$

We remark that this definition of the SNR is not unique and it may change in the appended papers depending on the considered system or problem.

2.1.2 Mutual Information

A rate R is said to be achievable if for $\lfloor 2^{RN} \rfloor$ messages and any $\epsilon > 0$ there exists a coding scheme such that CER $\leq \epsilon$ when N is sufficiently large. It was shown in [1] that for a memoryless channel the rate given by

$$I(X;Y) = \mathrm{E}\left\{\log_2\left(\frac{f_{Y|X}(Y|X)}{f_Y(Y)}\right)\right\} \tag{2.8}$$

is achievable under ML decoding, where X is a RV following a given distribution $f_X(\cdot)$. An ML decoder implements the following rule

$$\hat{\mathsf{m}} = \arg\max_{\mathsf{m}} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}_{\mathsf{m}}). \tag{2.9}$$

As follows from its definition, the ML decoder requires the maximization over all possible messages and is often impractical.

The quantity in (2.8) is called mutual information (MI) and characterizes the amount of information about X contained in Y. The proof of the fact that the rate given by the MI is achievable relies on a random coding argument [1]. The key element of this argument is that the entries of the codebook are independent RVs generated according to the distribution $f_X(\cdot)$.

The capacity of the channel $f_{Y|X}(\cdot|\cdot)$ is the supremum over all possible distributions of all achievable rates, i.e.,

$$\mathcal{C} = \sup_{f_X(\cdot)} I(X;Y), \tag{2.10}$$

and defines the maximum achievable rate, i.e., no coding scheme can achieve a rate larger than the channel capacity. For the AWGN channel in (2.6) with the average power constraint in (2.5) being the only limitation, the capacity is given by [1]

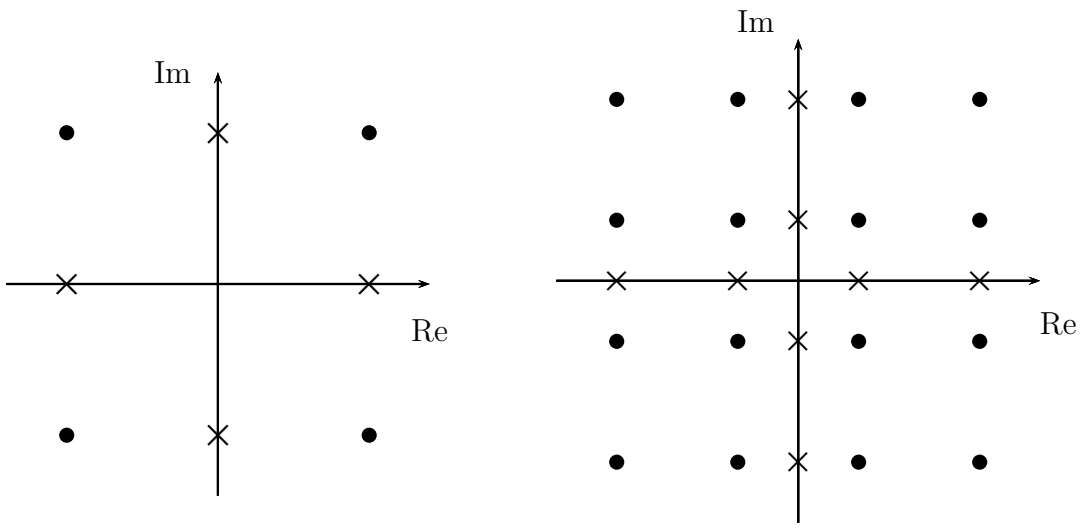
$$\mathcal{C} = \frac{1}{2}\log_2\left(1+\gamma\right). \tag{2.11}$$

To achieve the rate in (2.11), the codebook needs to be generated according to the Gaussian distribution $\mathcal{N}(0, E_{\mathrm{s}})$.

2.1.3 Constellations

To achieve the capacity of the AWGN channel, a continuous constellation needs to be used to generate the codebook. However, such constellations are impractical. Instead, discrete constellations are always implemented in real systems. Furthermore, equally spaced constellations, such as M-ary phase shift keying (PSK), pulse amplitude modulation (PAM), or quadrature amplitude modulation (QAM) formed as the Cartesian product of two PAM constellations, are usually preferred in practice. Examples of PAM constellations are shown in Fig. 2.2 with crosses in each dimension, and the corresponding QAM constellations obtained as the product are shown with dots.

In this thesis, we consider one-dimensional PAM constellations $\mathcal{S} = \{s_1, \ldots, s_M\}$, where $s_1 < \cdots < s_M$. The size of the constellation M is chosen to be a power of two. We further require that the constellation is equally spaced and symmetric around zero. For such constellations, we can write $s_i = -d(M-2i+1), i = 1, \ldots, M$, where d is half the distance between the constellation points and determines the average symbol energy. We remark that the results presented in this thesis apply directly to multi-dimensional constellations obtained as a product of one-dimensional constellations labeled with a product labeling [33, Sec. X].



(a) quadrature phase shift keying (QPSK) or 4-QAM constellation.

(b) 16-QAM constellation.

Figure 2.2: PAM and QAM constellations in 2-dimensions. Crosses in each dimension show PAM constellations and dots show QAM constellations.

Due to the error-correcting code design, the simplest probability distribution to impose on constellation points is a uniform distribution. In this thesis, we assume the constellation points to be equally likely, i.e., $f_X(s_i) = 1/M$ for $i = 1, \ldots, M$. The average symbol energy can then be expressed as

$$E_{\mathrm{s}} = \frac{M^2 - 1}{3} d^2. \tag{2.12}$$

For the rest of this chapter, we assume that the constellation is normalized to unit average energy, i.e., $E_{\mathrm{s}} = 1$.

Discrete constellations with uniform distributions used for the codebook construction generally lead to a loss in the achievable rate, which is clearly seen in Fig. 2.3. This figure shows the MI for M-PAM constellations and demonstrates that such constellations provide rates near the capacity in the so-called power-limited regime. However, in the bandwidth-limited regime (high SNR values) the loss is non-negligible and is equal to 1.53 dB when $M \to \infty$ due to the uniform distribution [34]. The MI shown in Fig. 2.3 is often referred to as CM capacity [35], meaning that it gives the maximum achievable rate for the specified constellation under the uniform distribution constraint. We denote the CM capacity by $\mathcal{C}^{\mathrm{CM}}$.

One could optimize the positions of the constellation points, for instance, in terms of bit error rate [36] or MI [37], which is referred to as geometrical shaping. Probabilistic shaping, i.e., changing the distribution of constellation points, could also be used to reduce this 1.53 dB gap [38–40].

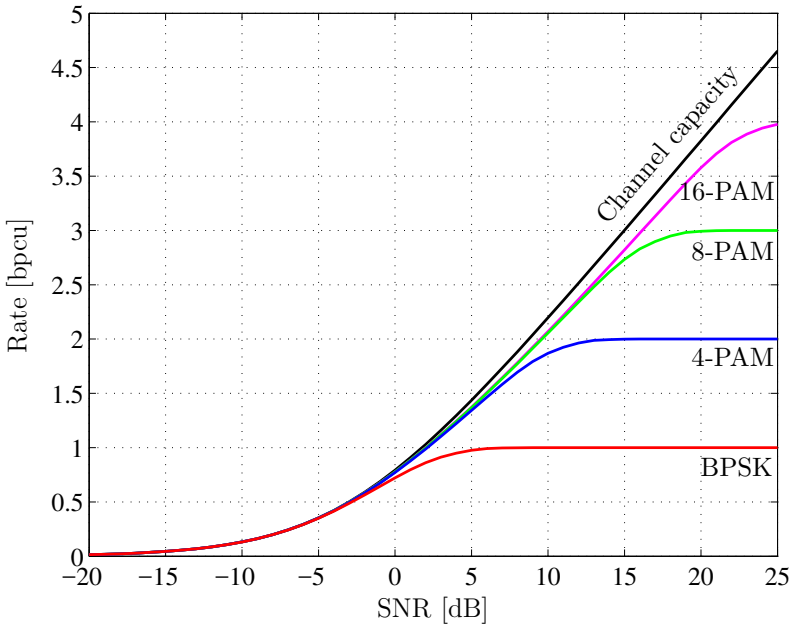


Figure 2.3: MI for equally spaced M-PAM constellations together the channel capacity for the AWGN channel.

2.2 Bit-Interleaved Coded Modulation

Once a discrete constellation is chosen, a unique binary label of length $m = \log_2(M)$ can be assigned to each constellation point. By doing so, the encoding rule $\mathsf{m} \to \boldsymbol{x}_{\mathsf{m}}$ can be performed in two steps. First, the message is mapped into a binary codeword $\boldsymbol{b}_{\mathsf{m}}$ of length mN and then each m bits $b_{m(k-1)+j}$, $k = 1, \ldots, N$, $j = 1, \ldots, m$ are mapped to a constellation point according to the chosen binary labeling. At the receiver side a length-mN vector of L-values $\boldsymbol{l}$ is calculated and then passed to a binary decoder. L-values are log-likelihood ratios defined as

$$L_{m(k-1)+j}(Y) \triangleq \log \frac{f_{Y_k|B_{m(k-1)+j}}(Y_k|1)}{f_{Y_k|B_{m(k-1)+j}}(Y_k|0)}. \tag{2.13}$$

The channel in Fig. 2.1 can then be replaced by the new one shown in Fig. 2.4, in which a modulator Φ is fed with m bits B_j, $j \in 1, \ldots, m$, and maps them to one of $M = 2^m$ possible constellation points. Since the constellation points are assumed to be equally likely, the incoming bits are independent and distributed according to $p_{B_j}(u) = 1/2, \forall j$ and $u \in \{0, 1\}$. The demodulator Φ^{-1} calculates the L-values as

$$L_j(Y) = \log \frac{\sum_{s \in \mathcal{S}_{j,1}} \mathrm{e}^{-\gamma(Y-s)^2/2}}{\sum_{s \in \mathcal{S}_{j,0}} \mathrm{e}^{-\gamma(Y-s)^2/2}}, \tag{2.14}$$

where $j = 1, \ldots, m$ and $\mathcal{S}_{j,u} \subset \mathcal{S}$ is a subconstellation of points whose labels have bit u in the jth bit position. The rest of this chapter is devoted to studying the new chan-

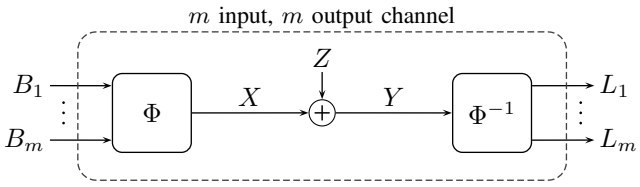


Figure 2.4: Bit-level channel.

nel Fig. 2.4 under the assumption that the input bits are independent and uniformly distributed.

2.2.1 Labelings and Patterns

A binary labeling is specified by a matrix $\mathbb{C}$ of dimensions M by m, where the ith row is the binary label of the constellation point s_i. It follows from (2.14) that the L-values depend on the subconstellations $\mathcal{S}_{j,u}$, which are determined by the bits in the jth column of the matrix $\mathbb{C}$. An example of subconstellations for 8-PAM with the binary reflected Gray code (BRGC) is shown in Fig. 2.5. When studying different properties of higher-order modulation, such as uncoded BER or bit-level MI, it can be convenient to consider the columns of the matrix separately. We refer to these columns as *patterns*, which are formally defined below.

Two commonly used binary labelings for the 8-PAM constellation are

$$\mathbb{C}_{\text{BRGC}} \triangleq \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}^{\mathsf{T}}, \quad \mathbb{C}_{\text{NBC}} \triangleq \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}^{\mathsf{T}}. \tag{2.15}$$

The first labeling is the BRGC, which minimizes the uncoded BER at high SNR and is widely used in noniterative BICM schemes. The second one is the natural binary code (NBC), which follows Ungerboeck's SP principle, and, hence, is popular in TCM schemes.

We define a bit pattern as a length-M binary vector $\boldsymbol{p} = [p_1, \ldots, p_M] \in \{0,1\}^M$ with $M/2$ ones and $M/2$ zeros. The labeling $\mathbb{C}$ can now be defined by m patterns, each corresponding to one column of $\mathbb{C}$. The patterns are indexed as $\boldsymbol{p}_w$ with w being the decimal representation of the vector $\boldsymbol{p}$, i.e., $w = \sum_{i=1}^{M} 2^{M-i} p_i$. The set of m indices of patterns in the labeling $\mathbb{C}$ is denoted by $\mathcal{W}$. The BRGC can be written in terms of patterns as $\mathbb{C}_{\text{BRGC}} = [\boldsymbol{p}_{15}^{\mathsf{T}}, \boldsymbol{p}_{60}^{\mathsf{T}}, \boldsymbol{p}_{102}^{\mathsf{T}}]$, where

$$\begin{aligned} \boldsymbol{p}_{15} &= [0,0,0,0,1,1,1,1], \\ \boldsymbol{p}_{60} &= [0,0,1,1,1,1,0,0], \\ \boldsymbol{p}_{102} &= [0,1,1,0,0,1,1,0], \end{aligned}$$

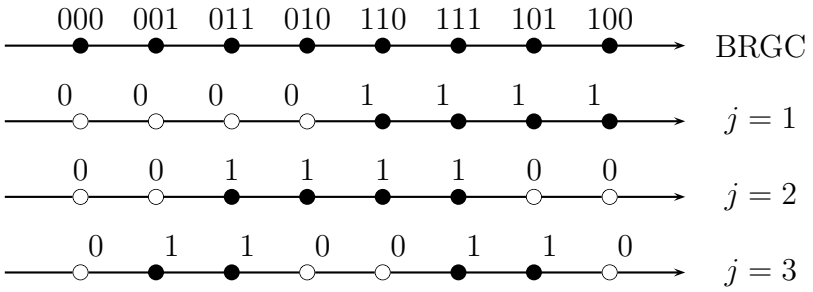


Figure 2.5: Subconstellations for 8-PAM with the BRGC. Subconstellations $\mathcal{S}_{j,0}$ are shown with empty circles and $\mathcal{S}_{j,1}$ are shown with filled circles.

and $\mathcal{W}_{\text{BRGC}} = \{15, 60, 102\}$. The NBC can be represented as $\mathbb{C}_{\text{NBC}} = [\boldsymbol{p}_{15}^{\text{T}}, \boldsymbol{p}_{51}^{\text{T}}, \boldsymbol{p}_{85}^{\text{T}}]$ with $\mathcal{W}_{\text{NBC}} = \{15, 51, 85\}$ and

$$\begin{aligned}\boldsymbol{p}_{15} &= [0,0,0,0,1,1,1,1],\\ \boldsymbol{p}_{51} &= [0,0,1,1,0,0,1,1],\\ \boldsymbol{p}_{85} &= [0,1,0,1,0,1,0,1].\end{aligned}$$

The number of different patterns for equally spaced one-dimensional constellations has been recently analyzed in [41,42]. We note that not all combinations of patterns can form a valid labeling, as all rows of $\mathbb{C}$ must be different.

2.2.2 Mutual Information

Achievable rates of a BICM system are usually analyzed under the ideal interleaver assumption [43], which turns the channel in Fig. 2.4 into m parallel channels. The achievable rate is then given by the sum of the bit-level mutual informations, i.e.,

$$\mathcal{C}^{\text{BICM}} = \sum_{j=1}^{m} I(B_j, L_j), \tag{2.16}$$

which is often referred to as the BICM capacity [35, 43]. Since the L-value L_j forms a sufficient statistics for guessing B_j based on Y [44, Ch. 22.3], the BICM capacity can be expressed as

$$\mathcal{C}^{\text{BICM}} = \sum_{j=1}^{m} I(B_j, Y). \tag{2.17}$$

The BICM capacity in (2.17) has been thoroughly studied, see, e.g., [43, 45, 46]. In Fig. 2.6 we show the BICM capacity for 8-PAM labeled with the BRGC and the NBC. The figure shows that the BICM capacity is upperbounded by the CM capacity, i.e., $\mathcal{C}^{\text{BICM}} \leq \mathcal{C}^{\text{CM}}$ for any labeling and constellation. The figure also demonstrates that the BICM capacity strongly depends on the labeling. It is commonly assumed that the

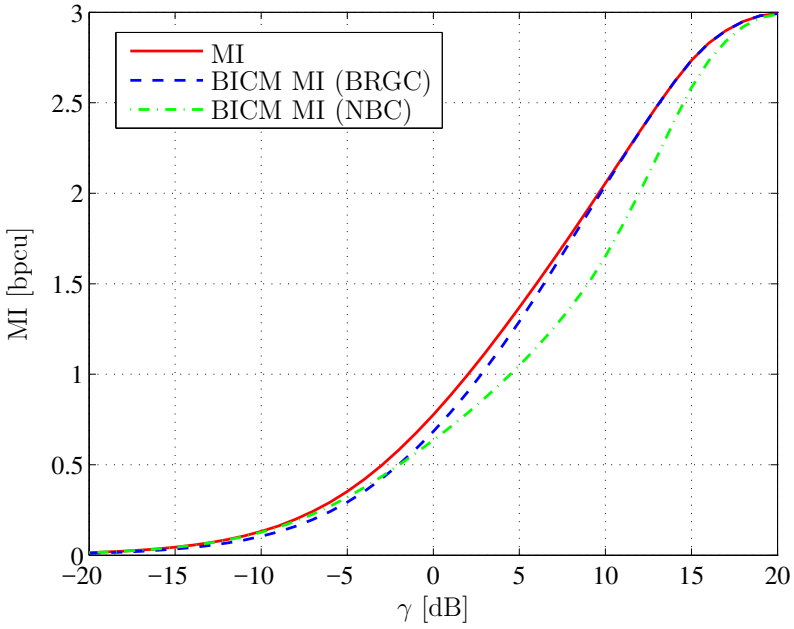


Figure 2.6: BICM MI for uniform input bits for the BRGC and the NBC labelings. The solid line shows the 8-PAM MI for uniform input symbols.

BRGC labeling maximizes the BICM MI for moderate and high SNR. The asymptotic high SNR optimality of the BRGC was shown in [47].

2.2.3 BICM Generalized Mutual Information

More recent studies [48, 49] based on the framework of mismatched decoding and generalized mutual information (GMI) showed that the ideal interleaving assumption is not necessary for achieving the rates predicted by the BICM capacity. The mismatched decoder considered in this thesis is inspired by the ML decoder for binary transmission and can be written as

$$\hat{\mathsf{m}} = \arg\max_{\mathsf{m}\in\mathcal{M}} \boldsymbol{b}_{\mathsf{m}}\boldsymbol{l}^{\intercal}, \tag{2.18}$$

i.e., the decoder finds the codeword that maximizes the correlation with the vector of the observed L-values. We refer to this decoder as the bit-wise decoder. Since the bit-wise decoder does not necessarily produce an ML estimate of the transmitted codeword, the MI analysis is not applicable to predict achievable rates.

It was shown in [48, Sec. III] that the rate given by the GMI is achievable by the bit-wise decoder over the channel in Fig. 2.4. Using the introduced notation, the GMI can we written as [50]

$$\mathrm{GMI} = m - \inf_{s\geq 0} \sum_{j=1}^{m} \mathrm{E}\left\{\log_2\left(1+e^{-s\check{B}_j L_j}\right)\right\}, \tag{2.19}$$

where $\check{B}_j = 2B_j - 1$. Furthermore, it was shown that $\mathrm{GMI} = \mathcal{C}^{\mathrm{BICM}}$ implying that the

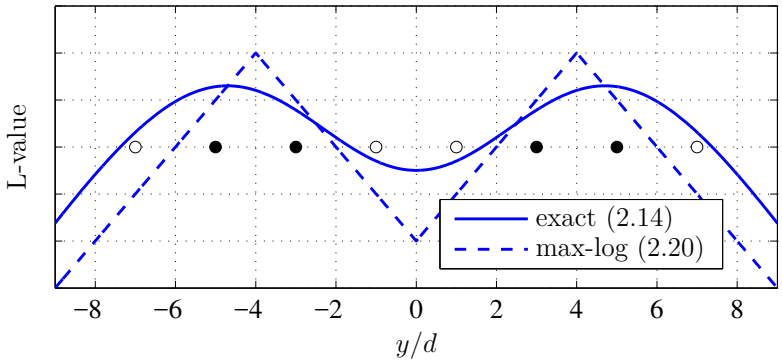


Figure 2.7: Exact and max-log L-values as functions of the (normalized) observation y for the third bit position of 8-PAM with the BRGC for $\gamma = 0$ dB. Empty and filled circles show constellation points labeled with 0 and 1, respectively.

bit-wise decoder is an ML decoder for the L-values in (2.14) under the ideal interleaver assumption. The maximum rate that can be achieved with the bit-wise decoder over the channel in Fig. 2.4 is still an open question. If the ideal interleaver assumption is lifted, one can use a mapper from the L-values $l_{m(k-1)+j}$, $j = 1, \ldots, m$ back to the observation $\boldsymbol{y}_k$ together with the ML decoder to achieve the CM capacity. In the following chapter, we show that the performance of the bit-wise decoder is "nearly" optimal, suggesting that rates larger than (2.19) can be achieved.

2.2.4 Max-Log L-values

The calculation of the L-values in their exact form (2.14) is complicated, especially for large constellations, as it requires calculation of the logarithm of a sum of exponentials. To overcome this problem, approximations are usually used in practice. The most common approximation is the so-called max-log approximation ($\log \sum_i \mathrm{e}^{\lambda_i} \approx \max_i \lambda_i$) [8, eq. (3.2)], [43, eq. (9)], [51, eq. (5)], [52, eq. (8)], which used in (2.14) gives

$$\tilde{L}_j(Y) \triangleq \frac{\gamma}{2}\left[\min_{x\in\mathcal{S}_{j,0}} (Y-x)^2 - \min_{x\in\mathcal{S}_{j,1}} (Y-x)^2\right]. \tag{2.20}$$

The use of the max-log approximation transforms the nonlinear relationship (2.14) into a piece-wise linear relationship (2.20), as previously shown in, e.g., [53, Fig. 3], [54, eqs. (11)–(14)]. The exact and the max-log L-values are shown in Fig. 2.7. The exact L-value approaches the max-log L-value when the SNR increases.

When max-log L-values are considered, most of the previous work concentrates on the correction of the "suboptimal" L-values in order to either maximize the BICM GMI [49,55] or minimize the error probability, [56] and references therein. In [57], the exact and the max-log L-values were compared from an uncoded performance viewpoint. However, a rigorous comparison of achievable rates in terms of the BICM capacity and the GMI has not yet carried out. In [PaperA], we present a comprehensive comparison of the two

L-values in terms of information rate loss for M-PAM (not necessarily equally spaced or symmetric) constellations. The numerical results presented in the paper are based on the distribution of max-log L-values discussed in Section 3.3.2.

Chapter 3

Practical Approaches to Coded Modulation

3.1 Trellis-Coded Modulation

TCM introduced in [4, 5] was one of the first practical CM schemes with reasonably good performance. It was rather quickly implemented in modem standards [58] in the beginning of the 90s. Since then, coding theory has made a huge progress and TCM is not widely used in modern (especially wireless) communication systems nowadays. However, it gives insights into how a binary code and modulation interact. It is also a good example of a coded scheme where the optimal ML decoder can be implemented with affordable complexity. These are the reasons why TCM is considered in this thesis.

The system model of TCM is a special case of the system model in Fig. 2.1. The message m is represented by a length-K vector of information bits $\boldsymbol{c} \in \{0,1\}^K$. This vector is fed to a convolutional encoder[1], which produces a vector of coded bits $\boldsymbol{b} \in \{0,1\}^{N_\mathcal{B}}$ of length $N_\mathcal{B}$. All possible codewords form a binary code $\mathcal{B}$. The rate of the CC is $R_\mathcal{B} = \frac{K}{N_\mathcal{B}}$. The encoder is defined by a generator matrix $\mathbb{G}$ as in [59, Ch. 4], where the elements of the generator matrix are polynomials over the binary field given in octal form. The coded bits are fed to the modulator, defined in the previous chapter, which outputs a vector of symbols $\boldsymbol{x} \in \mathcal{S}^N$ of length N. The vector $\boldsymbol{x}$ is called a TCM codeword. All possible TCM codewords form a TCM code $\mathcal{X}$. The overall rate (or spectral efficiency) is therefore $R = \frac{K}{N}$ and defines how many information bits are transmitted per channel use. The equality $N_\mathcal{B} = mN$ should hold to match the lengths of the vectors. The rate of the CC can be then expressed as $R_\mathcal{B} = \frac{K}{mN}$.

Examining the achievable rates in Fig. 2.3, Ungerboeck suggested to use CCs with rate $R_\mathcal{B} = \frac{m-1}{m}$. For such a rate, a spectral efficiency of $m-1$ bpcu can be achieved with a 2^m-point constellation at SNRs very close to the channel capacity. For instance, a spectral efficiency of 2 bpcu can be achieved with 8-PAM and a rate-2/3 code at around 10 dB. Increasing the constellation cardinality further does not give any significantly noticeable gain for the desired spectral efficiency.

CCs allow a reasonably simple implementation of the ML decoder by means of the Viterbi algorithm [6]. The choice of the rate $R_\mathcal{B} = \frac{m-1}{m}$ facilitates the decoding because m

[1] We consider only feedforward encoders.

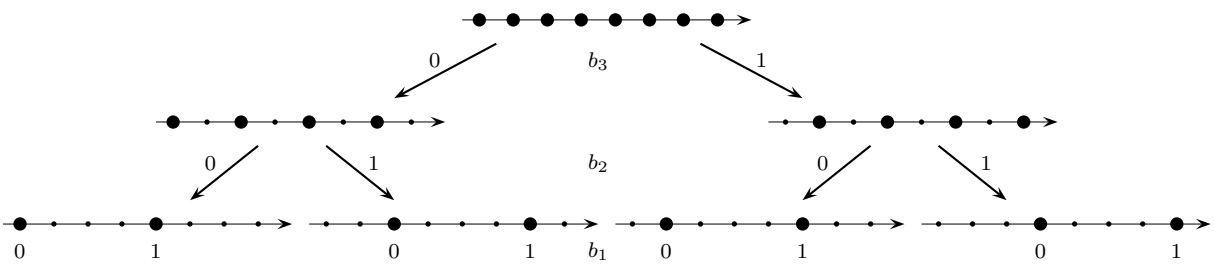


Figure 3.1: Set partitioning for 8-PAM that results in the NBC labeling.

coded bits on each trellis section of the CC are replaced by one symbol. Other rates could be implemented by using multidimensional TCM schemes [60], where one trellis section of the code comprises several symbols. The decoder is formally described in Section 3.3, where symbol-wise and bit-wise decoders are compared.

As pointed out by Ungerboeck, the figure of merit for a TCM system is the ED between the codewords of the TCM code rather than the Hamming distance between the codewords of the underlying CC. For example, the minimum Hamming distance $d_{\min}$, i.e., the minimum Hamming weight of a nonzero codeword of a linear code, is an important property of a CC. Its TCM counterpart is the minimum ED between two codewords of a TCM code. Even though these two quantities are a measure of distance between codewords, there is no explicit relation between them.

Ungerboeck also suggested a set of rules to design good TCM schemes. The key element in his design is the SP technique for constructing labelings for constellations. The main idea is to partition the constellation into subconstellations with nondecreasing minimum ED between the subconstellation points. The SP procedure for the 8-PAM constellation is shown in Fig. 3.1. It is easy to see that for this constellation, each step of SP doubles the minimum ED. After appropriately labeling all the points, we can verify that the NBC defined in (2.15) follows the SP principle.

Consider a TCM code $\mathcal{X}$ obtained by concatenating a labeling $\mathbb{C}$ with a CC $\mathcal{B}$. Recent results in [61] showed that the same TCM encoder, i.e., the same TCM code $\mathcal{X}$ and the same mapping between the information bits and the codewords, can be implemented by using other labelings with properly modified binary encoders. This is not surprising since the labeling is not important once the trellis structure of the code is fixed and the symbols are assigned to the trellis branches. As [61] reveals, different labelings can be grouped into equivalence classes. Any labeling $\mathbb{C}'$ within the same equivalence class as $\mathbb{C}$, together with a properly modified CC of the same memory, can be used to obtain exactly the same TCM encoder. We denote the new binary code by $\mathcal{B}'$. It turns out that for one-dimensional constellations, the SP (or NBC) and the BRGC belong to the same class.

As mentioned earlier, the labeling is not important for TCM and any other labeling outside the equivalence class can be used as well. However, the binary code $\mathcal{B}'$ in this case may no longer be a CC (for instance, it may not contain the all-zero codeword and

may therefore be nonlinear).

As an example of equivalent TCM schemes, we consider 8-PAM labeled with the NBC and a rate-2/3 CC defined by the generator matrix [59, Ch. 4]

$$\mathbb{G} = \begin{bmatrix} g_{1,1} & g_{1,2} & g_{1,3} \\ g_{2,1} & g_{2,2} & g_{2,3} \end{bmatrix}.$$

By looking at the patterns for the two labelings in (2.15), it is easy to establish the relations

$$\boldsymbol{p}_{60} = \boldsymbol{p}_{51} \oplus \boldsymbol{p}_{15}, \tag{3.1}$$

$$\boldsymbol{p}_{102} = \boldsymbol{p}_{85} \oplus \boldsymbol{p}_{51}, \tag{3.2}$$

i.e., all bit patterns in the BRGC can be obtained from the patterns of the NBC. To obtain the same TCM scheme when switching labelings, the new generator matrix $\mathbb{G}'$ is given by

$$\mathbb{G}' = \begin{bmatrix} g_{1,1} & g_{1,2} + g_{1,1} & g_{1,3} + g_{1,2} \\ g_{2,1} & g_{2,2} + g_{2,1} & g_{2,3} + g_{2,2} \end{bmatrix}, \tag{3.3}$$

where addition is performed for the corresponding polynomials over the binary field.

Let us consider an example of a TCM scheme with the CC defined by $\mathbb{G}_{\text{NBC}} = [7, 6, 2; 7, 3, 0]$ and the NBC labeling $\mathbb{C}_{\text{NBC}}$. We note that this TCM scheme is not optimal in terms of minimum ED and also does not posses the structure devised by Ungerboeck, where the most protected bit position is uncoded. However, this example makes the performance analysis for TCM and BICM intuitive and illustrative. After applying the linear operations on the generator polynomials, we conclude that the same TCM scheme can be implemented by using the BRGC and the CC with the generator matrix

$$\mathbb{G}_{\text{BRGC}} = \begin{bmatrix} 7 & 1 & 4 \\ 7 & 4 & 3 \end{bmatrix}. \tag{3.4}$$

We denote this configuration by the tuple $(\mathbb{C}_{\text{BRGC}}, \mathbb{G}_{\text{BRGC}})$ and it is used later on for illustration purposes. The steps of this "transformation" are schematically shown in [61, Fig. 1].

We repeated the same procedure for 4-PAM with the NBC and the rate-1/2 codes specified in [5, Table I] up to memory $\nu = 9$. The results are summarized in Table 3.1, where the generator matrices together with the corresponding minimum Hamming distances $d_{\min}$ are shown. The second column shows the maximum Hamming distance for CCs of a given memory [62]. An interesting observation is that codes that provide good TCM schemes when used with the BRGC have in general a large minimum Hamming

Table 3.1: Convolutional encoders for equivalent TCM schemes. $\mathbb{G} = [g_1, g_2]$ corresponds to the encoder for the NBC, whereas $\mathbb{G}' = [g_1, g_1 + g_2]$ is the encoder for the BRGC. The polynomials are given in octal form. The minimum Hamming distances for the codes $\mathbb{G}$ and $\mathbb{G}'$ are denoted by $d_{\min}$ and $d'_{\min}$, respectively. The second column shows the MFD for a given code memory ν.

ν	MFD	$\mathbb{G}$	$d_{\min}$	$\mathbb{G}'$	$d'_{\min}$
2	5	[5, 2]	3	[5, 7]	5
3	6	[13, 4]	4	[13, 17]	6
4	7	[23, 4]	4	[23, 27]	7
5	8	[45, 10]	4	[45, 55]	7
6	10	[103, 24]	5	[103, 127]	8
7	10	[235, 126]	8	[235, 313]	10
8	12	[515, 362]	8	[515, 677]	12
9	12	[1017, 342]	8	[1017, 1355]	12

distance $d_{\min}$. In fact, most of the codes have the maximum possible minimum Hamming distance $d_{\min}$. When the BRGC labeling is used, this result is intuitively expected as a code with a large minimum Hamming distance $d_{\min}$ is likely to result in a TCM code with a large minimum ED. However, the number of different bits (Hamming distance) between the codewords is not the only parameter to define the ED. The way the bits are grouped also plays an important role and sometimes it could be beneficial to use a code with a slightly smaller minimum Hamming distance $d_{\min}$ but better "bit grouping". This is the case for the codes with $\nu = 5, 6$ in Table 3.1.

In this chapter, we study the performance of various CM schemes in terms of the BER defined in (2.3). To that end, we define the pairwise error probability (PEP) for the codewords $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ as

$$\mathrm{PEP}\{\hat{\boldsymbol{X}} = \boldsymbol{x}_j | \boldsymbol{X} = \boldsymbol{x}_i\} = \mathrm{Q}\left(\sqrt{\frac{d^2(\boldsymbol{x}_i, \boldsymbol{x}_j)}{4\sigma^2}}\right), \tag{3.5}$$

where $d(\boldsymbol{x}_i, \boldsymbol{x}_j) = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|$ denotes the ED between the two codewords. The exact BER in (2.3) is usually too complex to calculate and union bounds are considered in most cases. The union bound for the BER is given by

$$\mathrm{BER}_{\mathrm{UB}} = \frac{1}{K2^K} \sum_{\boldsymbol{x}_i \in \mathcal{X}} \sum_{\substack{\boldsymbol{x}_j \in \mathcal{X}, \\ \boldsymbol{x}_j \neq \boldsymbol{x}_i}} d_{\mathrm{H}}(\boldsymbol{c}_i, \boldsymbol{c}_j) \mathrm{Q}\left(\sqrt{\frac{d^2(\boldsymbol{x}_i, \boldsymbol{x}_j)}{4\sigma^2}}\right), \tag{3.6}$$

where $\boldsymbol{c}_i$ is the information vector assigned to the codewords $\boldsymbol{x}_i$. By combining Q-functions

with the same arguments, (3.6) can be rewritten as in [61, eq. (22)] in terms of the distances $\mathcal{D} = \{d(\boldsymbol{x}_i, \boldsymbol{x}_j) : \boldsymbol{x}_i, \boldsymbol{x_j} \in \mathcal{X}, \boldsymbol{x}_i \neq \boldsymbol{x}_j\}$, i.e.,

$$\mathrm{BER}_{\mathrm{UB}} = \sum_{u \in \mathcal{D}} B_u \mathrm{Q}\left(\sqrt{\frac{u^2}{4\sigma^2}}\right), \tag{3.7}$$

where B_u is the bit multiplicity for the distance $u \in \mathcal{D}$. The distances in $\mathcal{D}$ with the corresponding multiplicities form the so-called distance spectrum.

CM codes in general and TCM codes in particular are nonlinear codes, i.e, the sum of two codewords does not have to be another codeword. This makes the performance analysis of TCM more complicated compared to linear codes, where a performance analysis can be based on the assumption that the all-zero codeword is transmitted. In the case of TCM, the distance spectrum cannot be obtained based on a state machine as for binary CC [63, Ch. 8].

For high SNR, the performance is dominated by pairs of codewords at minimum ED. We illustrate the high-SNR analysis for the setup $(\mathbb{C}_{\mathrm{BRGC}}, \mathbb{G}_{\mathrm{BRGC}})$. The minimum Hamming distance of the code $\mathbb{G}_{\mathrm{BRGC}}$ is $d_{\min} = 5$. We can identify binary codewords that give two TCM codewords at minimum ED. The binary codewords are

$$\begin{aligned} \boldsymbol{b} &= [0, 0, 0, 0, 0, 0, 0, 0, 0 \ldots], \\ \hat{\boldsymbol{b}} &= [0, 1, 1, 0, 0, 1, 0, 1, 1 \ldots], \end{aligned} \tag{3.8}$$

where dots represent some arbitrary common bits of the two codewords. For simplicity, we assume these bits are zero. Using Fig. 2.5, we conclude that the corresponding TCM codewords are

$$\begin{aligned} \boldsymbol{x} &= [s_1, s_1, s_1, \ldots], \\ \hat{\boldsymbol{x}} &= [s_3, s_2, s_3, \ldots], \end{aligned} \tag{3.9}$$

where dots represent some common symbols. The ED between the codewords $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ can be calculated as

$$d^2(\boldsymbol{x}, \hat{\boldsymbol{x}}) = 2d^2(s_1, s_3) + d^2(s_1, s_2) = 36d^2.$$

The Hamming distance between the codewords happens to be 5, i.e., it is equal to the minimum Hamming distance of the code. We note that this is not always the case and binary codewords at minimum Hamming distance may result in TCM codewords with the ED larger than the minimum ED of the code.

For the code $\mathbb{G}_{\mathrm{BRGC}}$, the output distribution of the symbols is uniform[2] and this allows to relate the average symbol energy to the constellation parameter d using (2.12)

$$E_{\mathrm{s}} = 21d^2.$$

[2]This is not generally true for an arbitrary CC.

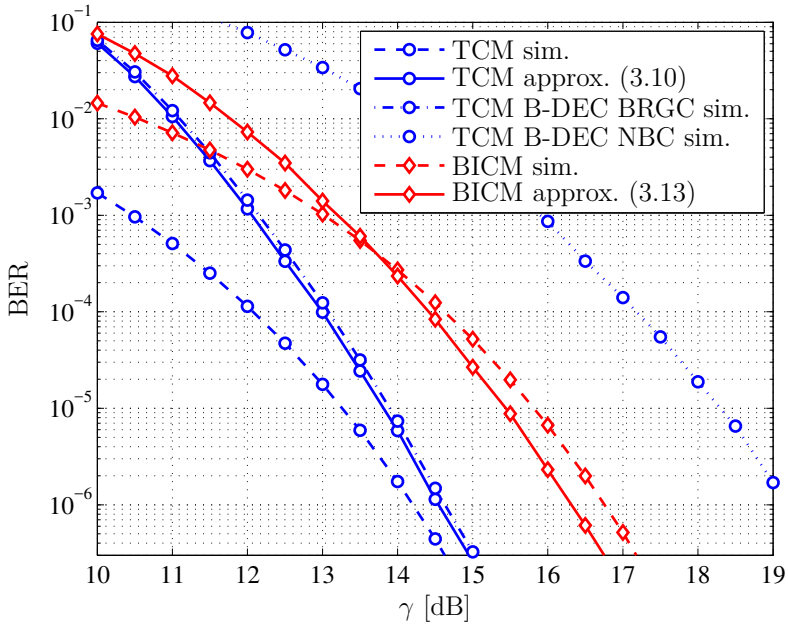


Figure 3.2: The BER performance for a coding scheme $(\mathbb{C}_{\text{BRGC}}, \mathbb{G}_{\text{BRGC}})$ without and with interleaver (see Section 3.2).

The high-SNR approximation of the BER can thus be written as

$$\text{BER} \approx \text{Q}\left(\sqrt{\frac{9d^2}{\sigma^2}}\right) = \text{Q}\left(\sqrt{\frac{9}{21}\gamma}\right). \tag{3.10}$$

The high-SNR approximation in (3.10) is shown in Fig. 3.2 with the dashed line marked with circles. The dashed line marked with circles shows the simulation results for the TCM scheme $(\mathbb{C}_{\text{BRGC}}, \mathbb{G}_{\text{BRGC}})$ and demonstrates a good agreement with the theoretical prediction at high SNR. A better agreement could be achieved if more terms of the distance spectrum with their multiplicities are taken into account according to (3.7). The figure also shows results for a BICM system, described in Section 3.2, and the performance curves for the bit-wise decoder, which is discussed in Section 3.3.

3.2 Bit-Interleaved Coded Modulation

BICM was originally proposed in [8] as a CM scheme for fast fading channels. Two new elements compared to TCM were introduced in a BICM system by Zehavi in [8], i.e., the bit-wise interleaver at the transmitter and the bit-wise decoder at the receiver. It is not always clear which one is an intrinsic part of a BICM system. A system without interleaver but with the bit-wise decoder may still be called BICM, for instance, when a CM scheme uses LDPC codes [64].

The system model of a BICM scheme is shown in Fig. 3.3. After the information vector

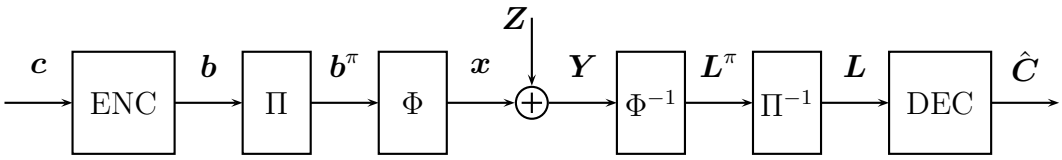


Figure 3.3: BICM system model.

$\boldsymbol{c}$ is encoded into a vector of coded bits $\boldsymbol{b}$, a bit-wise interleaver Π produces a permuted version of the coded bits $\boldsymbol{b}^\pi$. The modulator Φ maps the interleaved bits to symbols and sends the vector of the symbols to the channel. At the receiver, the demapper Φ^{-1} calculates the L-values $\boldsymbol{L}^\pi$ for the coded bits using (2.14) or (2.20). After deinterleaving, the L-values $\boldsymbol{L}$ are passed to a bit-wise decoder (2.18). When the binary code is a CC, such a decoder can be efficiently implemented using a soft-input Viterbi decoder for the binary code. This is the reason we consider CCs in this chapter. The decoder is discussed in detail in the following section. Here, we try to gain some intuition into how the interleaver affects the performance.

We study the effects of the interleaver under a hypothetical ML decoder by considering a simple example. Note that this "analysis" is far from being rigorous and complete but it gives intuition into many phenomena of BICM. Assume the setup $(\mathbb{C}_{\mathrm{BRGC}}, \mathbb{G}_{\mathrm{BRGC}})$ and the pair of binary codewords in (3.8). The binary code is assumed to be sufficiently long and the interleaver matches the length of the code, which corresponds to Fig. 3.3. The number of ones in $\hat{\boldsymbol{b}}$ is 5, which corresponds to the minimum Hamming distance of $\mathbb{G}_{\mathrm{BRGC}}$. Since the number of zeros is much larger than the number of ones, a vector of interleaved bits $\hat{\boldsymbol{b}}^\pi$ is very likely to have a structure, where each triplet of bits contains at most one nonzero bit. For instance, after interleaving, the two codewords may look like

$$\begin{aligned}\boldsymbol{b}^\pi &= [\ldots, 0, 0, 0, \ldots, 0, 0, 0, \ldots, 0, 0, 0, \ldots, 0, 0, 0, \ldots, 0, 0, 0, \ldots],\\ \hat{\boldsymbol{b}}^\pi &= [\ldots, 0, 0, 1, \ldots, 0, 0, 1, \ldots, 0, 0, 1, \ldots, 0, 0, 1, \ldots, 0, 0, 1, \ldots].\end{aligned} \tag{3.11}$$

After the modulator, the corresponding CM codewords have the structure

$$\begin{aligned}\boldsymbol{x} &= [\ldots, s_1, \ldots, s_1, \ldots, s_1, \ldots, s_1, \ldots, s_1, \ldots],\\ \hat{\boldsymbol{x}} &= [\ldots, s_2, \ldots, s_2, \ldots, s_2, \ldots, s_2, \ldots, s_2, \ldots].\end{aligned} \tag{3.12}$$

As we can see, the number of different symbols is equal to the Hamming distance between the binary codewords. For fast fading channels, the minimum number of symbols between all pairs of codewords is the so-called code diversity. Comparing the codewords in (3.9) with (3.12), we conclude that the interleaver increases the code diversity and makes it equal to the minimum Hamming distance of the code. This was the original motivation behind the BICM scheme proposed in [8] for fading channels. We however concentrate on the analysis for the AWGN channel.

Obviously, there are many other possible interleaved versions of the codewords that have the same structure (i.e., each triplet has at most one nonzero bit), but the pair shown in (3.11) is special. The squared ED between the codewords in (3.12) is $d^2(\boldsymbol{x}, \hat{\boldsymbol{x}}) = 5d^2(s_1, s_2) = 5(2d)^2 = 20d^2$. It is easy to verify that, if the BRGC is used, this is the minimum ED among all possible interleaved versions of all codewords, which is in agreement with [43, eq. (64)]. Analogously to TCM, the high-SNR approximation of the BER can be written as

$$\text{BER} \approx \text{Q}\left(\sqrt{\frac{5d^2}{\sigma^2}}\right) = \text{Q}\left(\sqrt{\frac{5}{21}\gamma}\right) \tag{3.13}$$

and it is shown in Fig. 3.2 with the dashed line marked with diamonds. The dashed-dotted line marked with diamonds shows the simulation results for the BICM scheme $(\mathbb{C}_{\text{BRGC}}, \mathbb{G}_{\text{BRGC}})$ with an interleaver. We note that this rough analysis predicts the performance of the BICM system and should agree well at asymptotically high SNR. A better agreement for medium SNR could be achieved by considering more terms of the distance spectrum with their multiplicities in an expurgated bound type of analysis [43].

Comparing the codewords in (3.9) with (3.12), it can be seen that distributing the bits over different symbols decreases the minimum ED when the BRGC is used. This explains the improved performance of BICM systems with CCs when the interleaver is removed, first observed in [65] and then analyzed in [66], as bit differences between codewords of a CC are always "closely situated".

We note that the results for this simple example are based on the assumption of the ML decoder and do not account for the bit-wise decoder which is usually used in practice. In [Paper B], we show that, when the BRGC is used, the performance of the bit-wise decoder in many cases is very similar to that of the optimal ML decoder. This is also illustrated in Section 3.3.2. This makes the intuition for this simple example hold even when the suboptimal bit-wise decoder is used.

3.2.1 BICM-ID

BICM-ID was originally presented in [17,18] and it was shown to be a powerful technique for both AWGN and fading channels. It was also shown that non-Gray labelings may give better performance when BICM-ID is considered. In BICM-ID, the decoder is allowed to exchange information about the coded bits with the demapper. The resulting iterative decoder is outside the scope of this thesis and it is not discussed at all here.

In this short section, we only try to understand what happens when a non-Gray labeling is used in BICM under an ML decoding assumption. To this end, we introduce

the labeling

$$\mathbb{C}_{\text{AGC}} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}^{\intercal}. \tag{3.14}$$

For 8-PAM, this labeling has the largest possible Hamming distance between the closest constellation points. This labeling is called the anti-Gray code (AGC) [18]. We further consider the 8-PAM constellation labeled with the AGC. The most important observation is that the closest points differ by either 2 or 3 bits.

We consider the CC given by $\mathbb{G}_{\text{BRGC}}$ in (3.4) and the two codewords in (3.8). As pointed out in the previous section, the most likely vectors of interleaved bits $\hat{\boldsymbol{b}}^{\pi}$ have a structure where each triplet of bits contains at most one nonzero bit, e.g., as in (3.11). If the AGC is used, the modulated codewords are

$$\begin{aligned} \boldsymbol{x} &= [\ldots, s_2, \ldots, s_2, \ldots, s_2, \ldots, s_2, \ldots, s_2, \ldots], \\ \hat{\boldsymbol{x}} &= [\ldots, s_4, \ldots, s_4, \ldots, s_4, \ldots, s_4, \ldots, s_4, \ldots]. \end{aligned} \tag{3.15}$$

The ED between these codewords $d^2(\boldsymbol{x}, \hat{\boldsymbol{x}}) = 5d^2(s_2, s_4) = 5(4d)^2 = 80d^2$ is much larger compared to the distance between the codewords in (3.12). Since for the AGC the closest points cannot differ in one bit, we can conclude that, with very high probability, interleaved binary codewords result in CM codewords, which on average have larger ED than the same codewords would have if the BRGC was used. Note that such codewords differ in 5 symbols, i.e., the code diversity is not affected when the labeling is changed. We have to note, however, that the interleaver may result in an "unfortunate" pair of codewords[3], e.g.,

$$\begin{aligned} \boldsymbol{b}^{\pi} &= [\ldots, 0, 0, 0, \ldots, 0, 0, 0, \ldots], \\ \hat{\boldsymbol{b}}^{\pi} &= [\ldots, 1, 1, 1, \ldots, 1, 1, 0, \ldots], \end{aligned} \tag{3.16}$$

with the CM codewords of the form

$$\begin{aligned} \boldsymbol{x} &= [\ldots, s_2, \ldots, s_2, \ldots], \\ \hat{\boldsymbol{x}} &= [\ldots, s_1, \ldots, s_3, \ldots]. \end{aligned} \tag{3.17}$$

The ED between these codewords $d^2(\boldsymbol{x}, \hat{\boldsymbol{x}}) = d^2(s_2, s_1) + d^2(s_2, s_3) = 8d^2$ is even smaller than the minimum ED if the BRGC would have been used. This formally reduces both the minimum ED and the code diversity. Luckily, the probability to observe such pairs of codewords is very small. Therefore, most of the codewords are separated by a large ED even though there are also codewords very close to one another. This effect is very

[3] If the BRGC was used, this pair of codewords would be very "fortunate", as the squared ED between the codewords would be $d^2(\boldsymbol{x}, \hat{\boldsymbol{x}}) = d^2(s_6, s_1) + d^2(s_5, s_1) = 164d^2$.

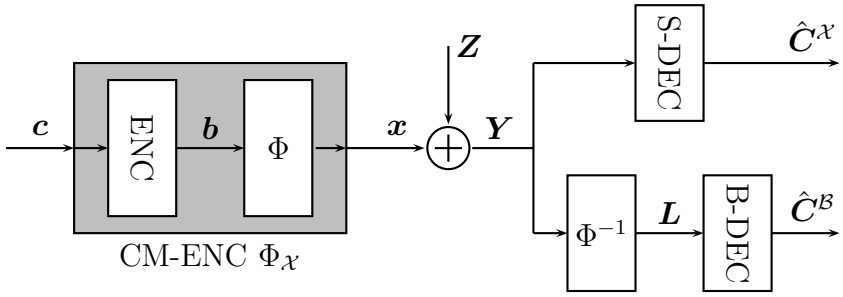


Figure 3.4: Block diagram of a general CM system. The CM encoder $\Phi_{\mathcal{X}}$ is used at the transmitter. At the receiver, two decoders are considered: the ML symbol-wise decoder S-DEC or a suboptimal bit-wise decoder B-DEC.

similar to the so-called "spectral thinning" [67] observed for turbo codes. Codewords far apart explain the good performance of BICM-ID at moderate SNR and codewords close to one another cause flattening of the BER at high SNR, which is usually referred to as the error floor (EF).

The performance predicted based on the ML decoder cannot be achieved if a standard bit-wise decoder is used in combination with a non-Gray labeling, as we illustrate in Section 3.3.2. This motivates the use of iterative decoders for non-Gray labelings.

3.3 Symbol-Wise Decoder Versus Bit-Wise Decoder

In this section, we generalize the system models of TCM and BICM and study the CM scheme shown in Fig. 3.4. A CM encoder (CM-ENC) maps information vectors $\boldsymbol{c} \in \{0,1\}^K$ to codewords $\boldsymbol{x} \in \mathcal{X}$, i.e., $\Phi_{\mathcal{X}} : \{0,1\}^K \to \mathcal{X}$, where no particular structure of the CM code is assumed. The inverse mapping is denoted by $\Phi_{\mathcal{X}}^{-1} : \mathcal{X} \to \{0,1\}^K$. By assigning binary labels to the constellation points, the CM encoder can be seen as a concatenation of a mapper Φ and a binary encoder (ENC in the figure) that performs a mapping of information vectors $\boldsymbol{c}$ to binary codewords $\boldsymbol{b} \in \mathcal{B}$. The binary encoder is defined as the function $\Phi_{\mathcal{B}} : \{0,1\}^K \to \mathcal{B}$ with the corresponding inverse function $\Phi_{\mathcal{B}}^{-1} : \mathcal{B} \to \{0,1\}^K$. The described coding scheme is equivalent to TCM if $\mathcal{B}$ is a CC, and it is equivalent to BICM in Fig. 3.3 if $\mathcal{B}$ is an interleaved version of another binary code.

We consider the AWGN channel model and, as the symbols in the CM code are not required to be equiprobable, we use the ratio d/σ as a measure of SNR. In the following, we describe two decoding algorithms adopted from TCM and BICM and show how they can be compared.

3.3.1 Symbol-Wise Decoder

The symbol-wise decoder (S-DEC) in Fig. 3.4 is an ML decoder. The ML decoding rule for CM over the AWGN channel is given in (2.9). Using the introduced notation, the

estimate of the transmitted data can be obtained as

$$\hat{\boldsymbol{C}}^{\mathcal{X}} \triangleq \Phi_{\mathcal{X}}^{-1}\left(\arg\min_{\boldsymbol{u}\in\mathcal{X}}\left\{\|\boldsymbol{Y}-\boldsymbol{u}\|^2\right\}\right). \tag{3.18}$$

The argument of the minimization function is the squared ED between the observation and a codeword. For a given observation $\boldsymbol{Y} = \boldsymbol{y}$, the squared ED is a function of a codeword $\boldsymbol{u} \in \mathcal{X}$ and it can be written as

$$D^{\mathcal{X}}(\boldsymbol{u}) = \sum_{k=1}^{N} (y_k - u_k)^2.$$

For the transmitted codeword $\boldsymbol{x}$, the squared ED is

$$D^{\mathcal{X}}(\boldsymbol{x}) = \sum_{k=1}^{N} (y_k - x_k)^2 = \sum_{k=1}^{N} (x_k + z_k - x_k)^2 = \sum_{k=1}^{N} z_k^2.$$

For a competing codeword $\hat{\boldsymbol{x}} = [\hat{x}_1, \ldots, \hat{x}_N] \in \mathcal{X}$, the squared ED can be expressed as

$$\begin{aligned} D^{\mathcal{X}}(\hat{\boldsymbol{x}}) &= \sum_{k=1}^{N} (y_k - \hat{x}_k)^2 = \sum_{k=1}^{N} (x_k + z_k - \hat{x}_k)^2 \\ &= \sum_{k=1}^{N} (x_k - \hat{x}_k)^2 + 2(x_k - \hat{x}_k) z_k + z_k^2. \end{aligned}$$

The PEP for these two codewords can then be rewritten as

$$\mathrm{PEP}^{\mathcal{X}}(\boldsymbol{x}, \hat{\boldsymbol{x}}) = \Pr\{D^{\mathcal{X}}(\hat{\boldsymbol{x}}) - D^{\mathcal{X}}(\boldsymbol{x}) < 0\}. \tag{3.19}$$

We therefore take a closer look at the difference $\Delta^{\mathcal{X}}(\boldsymbol{x}, \hat{\boldsymbol{x}}) \triangleq D^{\mathcal{X}}(\hat{\boldsymbol{x}}) - D^{\mathcal{X}}(\boldsymbol{x})$ between the squared ED for the codewords $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$,

$$\Delta^{\mathcal{X}}(\boldsymbol{x}, \hat{\boldsymbol{x}}) = \sum_{k=1}^{N} (x_k - \hat{x}_k)^2 + 2(x_k - \hat{x}_k) z_k = 4d \sum_{k=1}^{N} \Lambda^{\mathcal{X}}(x_k, \hat{x}_k), \tag{3.20}$$

where

$$\Lambda^{\mathcal{X}}(x_k, \hat{x}_k) \triangleq \left(\frac{x_k - \hat{x}_k}{2d}\right)^2 d + \frac{x_k - \hat{x}_k}{2d} z_k \tag{3.21}$$

is referred to as the symbol metric difference (SMD). Since Z_k is a Gaussian RV with zero mean and variance σ^2, the SMD in (3.21) is also a Gaussian RV with mean $\mu_{\mathcal{X}}(x_k, \hat{x}_k)d$ and variance $\sigma^2_{\mathcal{X}}(x_k, \hat{x}_k)\sigma^2$, where

$$\mu_{\mathcal{X}}(x_k, \hat{x}_k) = \sigma^2_{\mathcal{X}}(x_k, \hat{x}_k) = \left(\frac{x_k - \hat{x}_k}{2d}\right)^2. \tag{3.22}$$

Using (3.19)–(3.20) and the fact that the SMDs in (3.21) are independent Gaussian RVs for different $k = 1, \ldots, N$, the PEP for the codewords $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ can be calculated as

$$\begin{aligned}\mathrm{PEP}^{\mathcal{X}}(\boldsymbol{x}, \hat{\boldsymbol{x}}) &= \Pr\left\{\sum_{k=1}^{N} \Lambda^{\mathcal{X}}(x_k, \hat{x}_k) < 0\right\} \\ &= \mathrm{Q}\left(\frac{\sum_{k=1}^{N} \mu_{\mathcal{X}}(x_k, \hat{x}_k)}{\sqrt{\sum_{k=1}^{N} \sigma^2_{\mathcal{X}}(x_k, \hat{x}_k)}} \frac{d}{\sigma}\right) = \mathrm{Q}\left(\sqrt{\sum_{k=1}^{N} \mu_{\mathcal{X}}(x_k, \hat{x}_k)} \frac{d}{\sigma}\right). \end{aligned} \tag{3.23}$$

It can be easily verified that the PEP expression in (3.23) is equivalent to the one in (3.5). The distribution parameters in (3.22) are the key result in this section.

3.3.2 Bit-Wise Decoder

The bit-wise decoder (B-DEC) in Fig. 3.4 operates on the L-values provided by the demapper Φ^{-1}. The demapper acts independently of the B-DEC and calculates a vector of L-values $\boldsymbol{L}$ using the max-log approximation. For later use, we represent the vector of L-values as a concatenation of length-m vectors, i.e., $\boldsymbol{L} = [\boldsymbol{L}_1, \ldots, \boldsymbol{L}_N]$, where $\boldsymbol{L}_k = [L_{1,k}, \ldots, L_{m,k}]$ are the L-values obtained from the kth observation. We consider the max-log L-values in (2.20) with an additional scaling $\sigma^2(2d)^{-1}$ for notational convenience, i.e.,

$$L_{k,j} \triangleq \frac{1}{4d}\left[\min_{s \in \mathcal{S}_{j,0}} (Y_k - s)^2 - \min_{s \in \mathcal{S}_{j,1}} (Y_k - s)^2\right]. \tag{3.24}$$

The scaling is irrelevant for the B-DEC though it may be important if other decoders are used, e.g., a decoder based on the sum product algorithm [59, Ch. 5]. The calculated L-values are passed to the B-DEC that implements the decision rule (2.18), which can be rewritten using the introduced notation as

$$\hat{\boldsymbol{C}}^{\mathcal{B}} \triangleq \Phi_{\mathcal{B}}^{-1}\left(\arg\max_{\boldsymbol{u} \in \mathcal{B}} \{\boldsymbol{u}\boldsymbol{l}^{\intercal}\}\right). \tag{3.25}$$

For the transmitted codeword $\boldsymbol{b}$, the probability of detecting another codeword $\hat{\boldsymbol{b}}$ is

$$\mathrm{PEP}^{\mathcal{B}}(\boldsymbol{b}, \hat{\boldsymbol{b}}) = \Pr\{(2\boldsymbol{b} - 1)\boldsymbol{L}^{\intercal} - (2\hat{\boldsymbol{b}} - 1)\boldsymbol{L}^{\intercal} < 0\} = \Pr\{\Delta^{\mathcal{B}}(\boldsymbol{b}, \hat{\boldsymbol{b}}) < 0\}, \tag{3.26}$$

where

$$\Delta^{\mathcal{B}}(\boldsymbol{b}, \hat{\boldsymbol{b}}) \triangleq (\boldsymbol{b} - \hat{\boldsymbol{b}})\boldsymbol{L}^{\intercal} \tag{3.27}$$

can be interpreted as a distance between the codewords when the B-DEC is used. In [35, Ch. 4], $-\Delta^{\mathcal{B}}(\boldsymbol{b}, \hat{\boldsymbol{b}})$ was called a pairwise score and the distribution of the pairwise score was analyzed under different assumptions. In this chapter, we are also interested in the distribution of $\Delta^{\mathcal{B}}(\boldsymbol{b}, \hat{\boldsymbol{b}})$. However, our main goal is to compare this distribution with that of $\Delta^{\mathcal{X}}(\boldsymbol{x}, \hat{\boldsymbol{x}})$ in (3.20). Since the mapping between $\boldsymbol{b}$ and $\boldsymbol{x}$ is one-to-one, with a slight abuse of notation, we use $\Delta^{\mathcal{B}}(\boldsymbol{x}, \hat{\boldsymbol{x}})$ instead to highlight the similarity with (3.20).

Table 3.2: SMDs as different combinations of L-values conditioned on different transmitted symbols for the BRGC.

		$\hat{x}_i$							
		$s_1\,(000)$	$s_2\,(001)$	$s_3\,(011)$	$s_4\,(010)$	$s_5\,(110)$	$s_6\,(111)$	$s_7\,(101)$	$s_8\,(100)$
x_i	$s_1\,(000)$	0	$-L_3$	$-L_2-L_3$	$-L_2$	$-L_1-L_2$	$-L_1-L_2-L_3$	$-L_1-L_3$	$-L_1$
	$s_2\,(001)$	L_3	0	$-L_2$	$-L_2+L_3$	$-L_1-L_2+L_3$	$-L_1-L_2$	$-L_1$	$-L_1+L_3$
	$s_3\,(011)$	L_2+L_3	L_2	0	L_3	$-L_1+L_3$	$-L_1$	$-L_1+L_2$	$-L_1+L_2+L_3$
	$s_4\,(010)$	L_2	L_2-L_3	$-L_3$	0	$-L_1$	$-L_1-L_3$	$-L_1+L_2-L_3$	$-L_1+L_2$
	$s_5\,(110)$	L_1+L_2	$L_1+L_2-L_3$	L_1-L_3	L_1	0	$-L_3$	L_2-L_3	L_2
	$s_6\,(111)$	$L_1+L_2+L_3$	L_1+L_2	L_1	L_1+L_3	L_3	0	L_2	L_2+L_3
	$s_7\,(101)$	L_1+L_3	L_1	L_1-L_2	$L_1-L_2+L_3$	$-L_2+L_3$	$-L_2$	0	L_3
	$s_8\,(100)$	L_1	L_1-L_3	$L_1-L_2-L_3$	L_1-L_2	L_2	$-L_2-L_3$	$-L_3$	0

We can rewrite (3.27) as

$$\Delta^{\mathcal{B}}(\boldsymbol{x},\hat{\boldsymbol{x}})=\sum_{k=1}^{N}\Lambda^{\mathcal{B}}(x_k,\hat{x}_k), \tag{3.28}$$

where

$$\Lambda^{\mathcal{B}}(x_k,\hat{x}_k)\triangleq(\Phi_{\mathcal{S}}^{-1}(x_k)-\Phi_{\mathcal{S}}^{-1}(\hat{x}_k))\boldsymbol{L}_k^{\intercal} \tag{3.29}$$

is the SMD[4] for the B-DEC. Since the channel is memoryless, the SMDs for different $k=1,\ldots,N$ are independent. Once the distributions of all SMDs are known, the PDF of $\Delta^{\mathcal{B}}(\boldsymbol{x},\hat{\boldsymbol{x}})$ in (3.28) can be found as a convolution of the PDFs of the summands. The PEP in (3.26) is then given by the integral of the negative tail of this PDF. In the following, we discuss distributions of SMDs for a given time instant k and we omit k for clarity.

In order to proceed further, we need to specify the labeling used by the demapper since it greatly affects the SMD in (3.29). We concentrate on the 8-PAM constellation labeled with the BRGC. Different SMDs $\Lambda^{\mathcal{B}}(x,\hat{x})$ can then be expressed in terms of the L-values L_j, $j=1,2,3$ as shown in Table 3.2. The binary labels for the constellation symbols are also shown so that the entries of the table become evident.

The performance of the B-DEC is highly dependent on the distributions of the L-values. The distribution of the L-values in a certain bit position may depend on whether zero or one was transmitted. This complicates the analysis and some simplifications are usually made. In [43], a time-varying labeling was proposed in order to symmetrize the channel. The same effect can be achieved by using a random scrambler as in [35] or [68]. This makes the SMDs for the pairs of symbols whose labels differ in the same bits to have the same distribution. For example, the SMDs on the main anti-diagonal $\Lambda^{\mathcal{B}}(s_1,s_8)$, $\Lambda^{\mathcal{B}}(s_2,s_7),\ldots,\Lambda^{\mathcal{B}}(s_8,s_1)$ in Table 3.2 should all have the same distribution under such an assumption. This, however, is not generally true. The distributions of the aforementioned SMDs are given by the distribution of L_1 conditioned on different transmitted symbols, and hence, are different. The assumption of a random scrambler gives good predictions

[4] $-\Lambda^{\mathcal{B}}(x_k,\hat{x}_k)$ is called a symbol score in [35, Ch. 4].

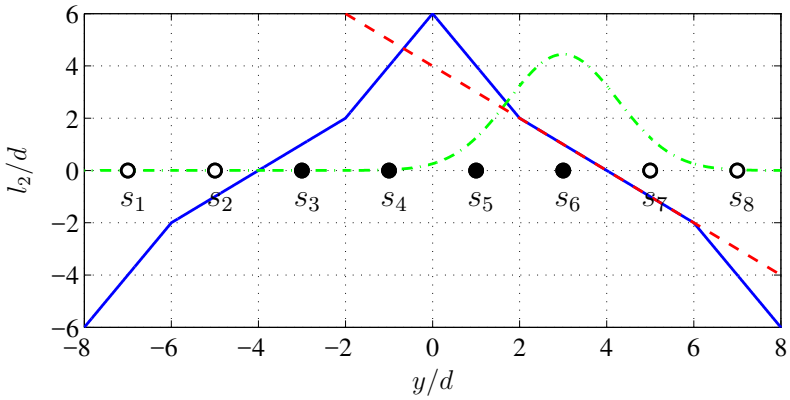


Figure 3.5: Approximation of L-values. The solid line shows (normalized) l_2 as a function of the (normalized) observation y. The dashed line shows the approximation (consistent and ZC) of the L-value. The dash-dotted line shows the distribution $f_{Y|X}(y|s_6)$ for $d/\sigma = -5$ dB; the distribution is scaled for illustration purposes. Empty and filled circles show constellation points labeled with 0 and 1, respectively.

of the PEP but cannot be used when the S-DEC and the B-DEC are compared. Here, we do not use this assumption and treat every SMD in Table 3.2 individually.

Distribution of L-values

As mentioned in Section 2.2.4, the max-log L-values are linear functions of the observation. This implies that the SMDs in Table 3.2 are linear functions of the observation as well. In this section, we show how to obtain the PDFs of two SMDs in Table 3.2 and how these PDFs can be approximated. The distributions of all other SMDs in Table 3.2 can be obtained and approximated in a similar fashion.

It follows from Table 3.2 that the SMD $\Lambda^{\mathcal{B}}(s_6, s_7) = L_2$ and hence, its distribution is the distribution of L_2 given that symbol s_6 was transmitted. Fig. 3.5 shows the piecewise linear function $l_2(y)$ together with the distribution of the observation conditioned on the transmitted symbol s_6 for $d/\sigma = -5$ dB. The PDF $f_{L_2|X}(l|s_6)$ is thus a sum of piecewise Gaussian functions with mean and variance determined by the parameters of the corresponding linear pieces of $l_2(y)$. The PDF $f_{L_2|X}(l|s_6)$, or equivalently, the PDF of $\Lambda^{\mathcal{B}}(s_6, s_7)$, is shown in Fig. 3.6 with the solid line.

The PDFs of the SMDs are easy to obtain but they become analytically intractable when several PDFs need to be convolved, and hence, approximations of the distributions are usually used. We consider two ways to approximate such PDFs. The first one is the so-called consistent approximation [66] for which the L-value is approximated by a linear function corresponding to the linear piece over the Voronoi region of the transmitted symbol. This approximation is shown with the dashed line in Fig. 3.5. The approximated PDF is thus a single Gaussian function that approximates the exact PDF "at the mean" of the L-value and it is shown with the dashed line in Fig. 3.6. The second approach approximates the negative tail of the PDF, which is more important when analyzing

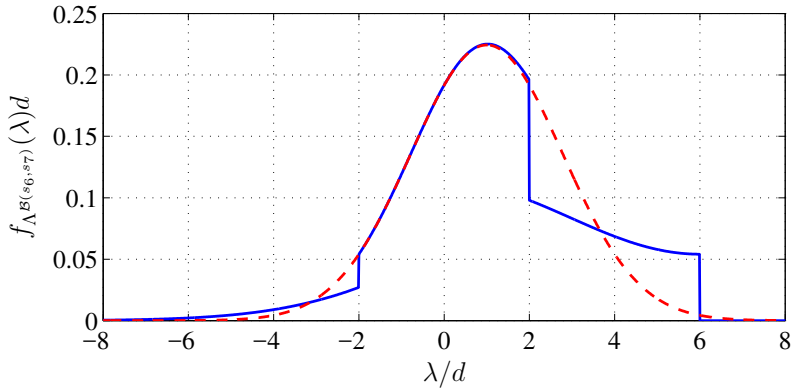


Figure 3.6: Distribution of the SMD $\lambda^{\mathcal{B}}(s_6, s_7)$ for $d/\sigma = -5$ dB. The solid line represents the exact PDF and the dashed line shows the approximated PDF. The consistent and the ZC approximations give the same results.

the PEP. For that, the L-value is approximated using the so-called zero-crossing (ZC) approximation [69], i.e., the L-value is approximated by a linear function corresponding to the linear piece of the L-value function intersecting the zero-level at the closest point to the transmitted symbol. For the example in Fig. 3.5, the ZC approximation coincides with the consistent approximation (dashed line) and gives the same approximated PDF, which is shown with the dashed line in Fig. 3.6. However, the two approximations are not always the same. In the following, we give an example showing that the ZC approximation should be preferred when analyzing the probability of error.

As shown in Table 3.2, the SMD $\Lambda^{\mathcal{B}}(s_6, s_8) = L_2 + L_3$. It is shown in Fig. 3.7 as a piece-wise linear function of the observation y together with the conditional distribution of the observation $f_{Y|X}(y|s_6)$. The exact PDF of $\Lambda^{\mathcal{B}}(s_6, s_8)$ is a sum of piece-wise Gaussian functions shown in Fig. 3.8. Due to the horizontal parts of the SMD in Fig. 3.7, the exact PDF contains a Dirac delta function with amplitude $A = 1-2\mathrm{Q}(d/\sigma)+\mathrm{Q}(5d/\sigma)-\mathrm{Q}(7d/\sigma)$ at $\lambda = 2d$. The consistent approximation approximates the SMD by a linear function $\lambda = 2d$ and hence, the approximated PDF is a Dirac delta function of unit amplitude at $\lambda = 2d$. Clearly, such an approximation gives poor results when the PEP is analyzed.

The ZC approximation of the SMD $\Lambda^{\mathcal{B}}(s_6, s_8)$ is shown with the dashed line in Fig. 3.7 and it results in a Gaussian distribution depicted with the dashed line in Fig. 3.8. It is possible to show that the ZC approximation is asymptotically tight (when $\gamma \to \infty$) in terms of PEP, if $\Delta^{\mathcal{B}}(\boldsymbol{x}, \hat{\boldsymbol{x}})$ in (3.28) consists of only one SMD. We also considered different $\Delta^{\mathcal{B}}(\boldsymbol{x}, \hat{\boldsymbol{x}})$ consisting of several SMDs, for instance $\Delta^{\mathcal{B}}(\boldsymbol{x}, \hat{\boldsymbol{x}})$ for the codewords in [Paper B, Fig. 2], and the asymptotic tightness of the ZC approximation was successfully verified. For a general case, where $\Delta^{\mathcal{B}}(\boldsymbol{x}, \hat{\boldsymbol{x}})$ is a summation of an arbitrary number of different SMDs, the tightness of the ZC approximation remains an open question.

Once the parameters of the approximated distributions (the mean and the variance) for all entries of Table 3.2 are known, they can be compared to those of the S-DEC. We performed this analysis for 4-PAM labeled with a Gray code in [Paper B]. The results showed that the distribution parameters are very similar for the two decoders, which

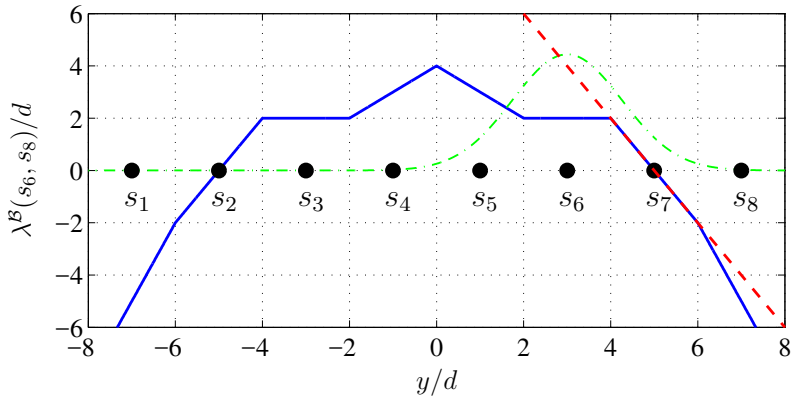


Figure 3.7: Approximation of the SMD. The solid line shows the (normalized) SMD $\lambda^{\mathcal{B}}(s_6, s_8)$ as a function of the (normalized) observation y. The dashed line shows the ZC approximation of the SMD. The dash-dotted line shows the distribution $f_{Y|X}(y|s_6)$ for $d/\sigma = -5$ dB; the distribution is scaled for illustration purposes.

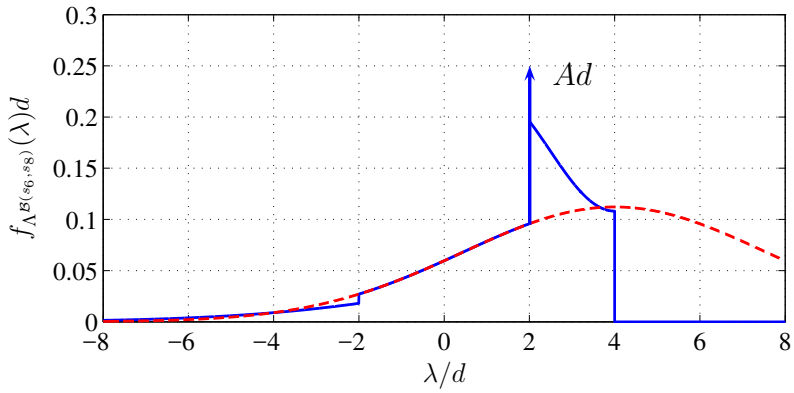


Figure 3.8: Distribution of the SMD $\lambda^{\mathcal{B}}(s_6, s_8)$ for $d/\sigma = -5$ dB. The solid line represents the exact PDF and the dashed line shows the approximated PDF using the ZC approximation. The PDF obtained based on the consistent approximation is a Dirac delta function with amplitude 1 at $\lambda = 2d$ and is not shown in the figure.

allowed us to bound the asymptotic performance loss of the B-DEC compared to the S-DEC. The analysis also showed that for a wide range of codes, the performance loss is equal to zero. This phenomenon is illustrated in Fig. 3.2 for 8-PAM. The dash-dotted line marked circles shows the performance of the B-DEC with the BRGC for the TCM scheme $(\mathbb{C}_{\text{BRGC}}, \mathbb{G}_{\text{BRGC}})$. The B-DEC causes a loss of fractions of a dB compared to the S-DEC shown with the solid line marked with circles.

A similar analysis could be performed for other than Gray labelings. For example, we derived the distribution parameters of the SMDs for 4-PAM labeled with the NBC. Unfortunately, the results do not allow us to bound the performance loss and draw quantitative conclusions. However, they clearly indicate that the loss can be large when the B-DEC is used with the NBC. We illustrate this in Fig. 3.2 for 8-PAM. The dotted line marked with circles shows the performance of the B-DEC with the NBC for the TCM

scheme $(\mathbb{C}_{\mathrm{BRGC}}, \mathbb{G}_{\mathrm{BRGC}})$. For that, the demapper for the NBC together with a decoder for the binary code $\mathbb{C}_{\mathrm{NBC}}$ is used at the receiver. As we can see, the loss is approximately 4.5 dB compared to the S-DEC. This gives an intuition on why the B-DEC performs bad when used with a non-Gray labeling and motivates the use of iterative decoders to obtain a near-optimal performance.

Chapter 4

Vehicular Communications

4.1 Vehicular Ad Hoc Networks

VCs refers to vehicle-to-vehicle (V2V) or vehicle-to-infrastructure (V2I) communications used to exchange messages containing safety and traffic information. VCs is the prerequisite for cooperative intelligent transportation systems (C-ITSs) intended to "increase safety, sustainability, efficiency, and comfort" of transportation systems [70, Ch. 2]. Although the standardization of VCs started in 1999, to date, solutions that can provide reliable message exchange between vehicles are not available. The main problems of VCs are highly mobile networks with rapidly changing topologies and large numbers of users and poor channel qualities. Moreover, safety-critical applications put stringent reliability and latency requirements on the physical (PHY) and the MAC layers.

Two different approaches to reliable VCs are considered in the literature. The first one is based on cellular technologies. The current standards, such as 3G and LTE, already provide vehicles with different services and traffic information. However, they cannot satisfy the stringent delay and reliability requirements of safety-critical applications. There is an ongoing effort to provide high reliability and low latency communications in the upcoming 5G standard [20] to support VCs [71]. The second approach is based on the direct short-range communication between nodes in a vehicular network. Such networks are referred to as vehicular ad hoc networks (VANETs). The name highlights the fact that there is no infrastructure or coordination in such networks. The short-range communication for VANETs is specified in the IEEE standard [72], which is usually referred to as 802.11p and is the de facto standard for VCs.

Different network topologies and types of communications are considered for VANETs (see [27, Ch. 2.2]). For example, in point-to-point communications, a transmitter sends a message to a specific receiver potentially using other nodes in the network for relaying. In point-to-multipoint communications, the transmitter broadcasts a message to its neighbors, who then relay it further. In other scenarios, users transmit their message to a specific geographical area through other nodes. In this thesis, we consider the one-hop broadcast scenario [73, 74], where every user intends to communicate with its neighbors, i.e., all nodes within its transmission range. We refer to this type of communications as *all-to-all broadcast*. An example of a VANET is shown in Fig. 4.1, where each rectangle represents a user in the network, the grey circle shows the transmission range of user 1,

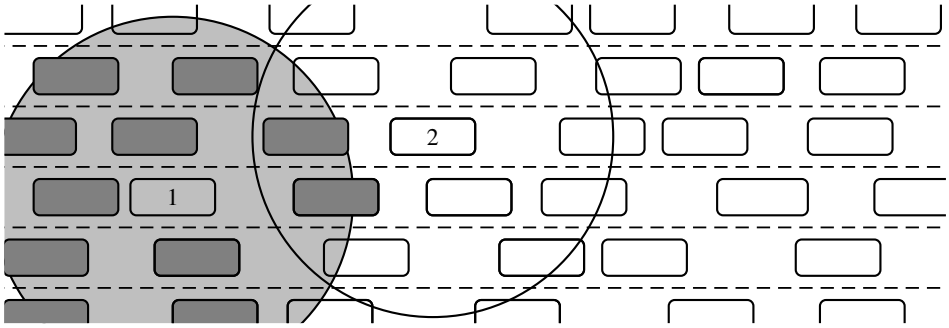


Figure 4.1: VANET model.

and grey rectangles show the neighbors of user 1.

There are two types of packets in VCs, namely decentralized environmental notification messages (DENMs) [75] and cooperative awareness messages (CAMs) [76]. CAM packets are sent periodically and contain information about the vehicle such as its position, speed, direction, etc. On the other hand, DENM packets are event-driven, i.e., they are sent in the case of an emergency or a dangerous situation. This also implies that these packets have higher priority and reliability requirements than those of CAM packets. Ideally, these packets should not be sent often. In this thesis, we consider transmission of CAM packets.

The PHY layer and the MAC layer specified in 802.11p [72] are slight modifications of those used in Wi-Fi for indoor wireless communications. The PHY layer is mainly criticized for not being able to cope with time-varying channels. However, with modified pilot patterns [77] and/or enhanced channel estimation techniques [78] it seems to satisfy the needs of VCs.

The 802.11p standard [72] uses CSMA [25] with collision avoidance as the MAC protocol. CSMA is a simple uncoordinated protocol that performs well for small networks and large amounts of information to be transmitted by each user [79]. However, CSMA cannot provide access with predictable delay and it scales badly with the number of users. To make the CSMA protocol more scalable, power control [74, 80] or transmission rate control [81] algorithms can be used to adjust the number of neighbors seen by each user. CSMA also severely suffers from the hidden terminal problem, discussed in the following section, which results in low reliability. Moreover, the overhead due to sensing becomes nonnegligible when the size of the packet becomes small. Thus, CSMA cannot satisfy the requirements of safety-critical applications in VCs [27].

Other MAC protocols proposed for VCs are based on self-organized time division multiple access (TDMA) [82, 83] and designed to avoid collisions. In these protocols, the time resource is divided into frames, which are further divided into slots, and users dynamically occupy free slots based on the information provided in a separate channel. In [27], collisions are allowed but scheduled for users separated far apart in the space to cope with overloaded networks. Self-organized TDMA protocols require a learning phase, which increases the channel access delay and may not work for fast changing networks.

Moreover, transmission errors during the learning phase render such protocols unusable.

In the following, we list the properties of VCs which need to be taken into consideration when designing a MAC protocol:

- Fast changing topologies (i.e., no coordination can be used).
- Strict latency constraints.
- High reliability requirements.
- Bad channel quality (hence, it is desired that the MAC protocol provides time diversity).
- All-to-all broadcast (i.e., acknowledgements or automatic repeat request (ARQ) schemes cannot be used).
- Large number of users, i.e., the protocol should scale well.
- The protocol should provide fair access and support different priorities.

On the other hand, some aspects of wireless communications can be taken into account less seriously [24]:

- Energy efficiency. It is not very important since there are virtually no battery constraints in VCs.
- It is safe to assume that users in a VANET are time synchronized, which could be achieved by means of global positioning service (GPS).
- Computational constraints since powerful hardware can be installed in a vehicle.
- Data rate. It is not critical since transmitted packets contain little information.

4.2 Carrier Sense Multiple Access

The basic idea of CSMA with collision avoidance is to listen to the medium before transmission. If the channel is sensed idle, a user transmits its packet. If the channel is busy, the user waits until it becomes idle and then defers its transmission by a random delay. If during this time the channel remains idle, the user transmits its packet.

CSMA relies on the fact that users can determine whether the channel is busy or not. This requires every node in the network to be in the range of all other nodes. Obviously, for a VANET in Fig. 4.1, this is not true. This gives rise to the so-called hidden terminal problem [84]. Consider the scenario in Fig. 4.1, where user 1 broadcasts its message to its neighbors, represented by gray rectangles. Assume that user 2 also wants to transmit a message to its neighbors. Since user 2 is outside the range of user 1, it senses the channel idle and proceeds with transmission, thus creating destructive interference for the users in the intersection of the two transmission ranges.

The request to send/clear to send (RTS/CTS) mechanism [84] is designed to mitigate the hidden terminal problem in CSMA. Prior to the data transmission, a transmitter

Table 4.1: The PHY parameters from [72].

Parameter	Variable	Value		Units
Period	t_{period}	100		ms
Data rate	r_{data}	6		Mbps
PHY preamble	t_{pream}	40		μs
CSMA slot duration	t_{slot}	13		μs
AIFS time	t_{aifs}	58		μs
Packet size	d_{data}	200	400	byte
Packet duration	t_{pack}	312	576	μs

sends an RTS packet notifying the intended recipient about the upcoming data transmission and its duration. If successfully received, the recipient replies with a CTS packet that is overheard by every user in its vicinity, thus notifying them about the upcoming transmission. After establishing a virtual point-to-point link, the transmitter sends the packet. However, due to the all-to-all broadcast nature of VCs and fast changing network topologies, this mechanism cannot be used in VANETs. Despite being vulnerable to the hidden terminal problem, CSMA was adopted for VC in [72]. In the following, we describe in detail how it works.

CSMA in 802.11p

We consider transmission of CAM packets [76] that are sent periodically every t_{period} seconds. We assume that all packets contain d_{data} bits and a PHY layer preamble of t_{pream} seconds. The packet duration t_{pack} can then be calculated as $t_{\text{pack}} = t_{\text{pream}} + d_{\text{data}}/r_{\text{data}}$, where r_{data} is the data rate. The parameters used in this thesis are taken from the PHY layer in [72] and are presented in Table 4.1.

The CSMA protocol is succinctly shown in Algorithm 1 and described in the following. Let us denote time by t. Assume that a user wants to transmit a packet at $t = \tau$, which needs to be transmitted within the next t_{period} seconds. If not, the packet is discarded. First, the user listens to the channel for the arbitration inter-frame spacing (AIFS) period t_{aifs}. This listening time is used to determine whether the channel is busy and provide priorities to different classes of messages. The shorter the AIFS period, the faster the access to the channel. In this thesis, all messages are equally important and we choose t_{aifs} corresponding to the highest priority (voice transmission) in 802.11p [22, Table 11]. If the channel is free, the user sends the packet and waits for the next one to arrive.

If the channel is busy or becomes busy during listening, the user invokes the backoff procedure. First, it draws a backoff number b uniformly at random from the interval

```
start Get a packet at t = τ.
listen during t_aifs;
if the channel is idle then
    transmit;
else
    draw b uniformly at random from [0, w];
    while b > 0 do
        wait until the channel becomes idle;
        if t > τ + T − t_aifs − b t_slot then
            drop the packet;
        else
            listen during t_aifs;
            while the channel is idle do
                if b > 0 then
                    b = b − 1 ;
                    listen during t_slot;
                else
                    transmit;
                end
            end
        end
    end
end
```

Algorithm 1: CSMA protocol.

$[0, w]$, where w is the contention window size specified in 802.11p [22, Table 11] for different priority classes. w specifies the number of CSMA slots of duration t_{slot} that the user will have to wait before transmitting.[5] Smaller values of w provide faster access to the channel. For the highest priority class, $w = 3$. However, small values of w increase the probability of collision. In [85], we optimized the contention window size and find that $w = 511$ provides a good balance between the number of collision and the number of dropped packets for the parameters in Table 4.1.

The next step is to wait until the channel becomes idle. The user checks if the channel is idle during t_{aifs}. If it is, the user decrements the backoff number b and listens during the time period t_{slot}. The user continues to decrement the backoff number b and to listen during t_{slot}, until $b = 0$ after which the user transmits the packet. If the channel at some point becomes busy, the user waits until it becomes idle and then continues with the backoff procedure.

[5] We remark that the slot in CSMA is different from that of TDMA schemes and is much smaller as compared to the packet duration.

4.3 Simplified System Model

Estimating the performance of the CSMA protocol for the network in Fig. 4.1 is computationally hard. The main reason for that is that estimating the performance of just one user requires the simulation of the entire network. In this section, we present a simplified model that we used in [85] to obtain "optimistic" bounds on the performance of a real system.

We consider a network of m users, indexed by $j = 1, \ldots, m$. The network is fully connected, i.e., every user is within all other users' transmission range. This assumption eliminates the hidden terminal problem and makes the model advantageous to the actual network. We assume that the network does not change over time. Collisions are considered destructive, i.e., none of the colliding packets can be decoded. On the other hand, packets are successfully decoded by all users if no collision occurs. Such a channel is called the collision channel [86].

Each user wants to send a packet at $T_j + it_{\text{period}}$, where T_j, $j = 1, \ldots, m$, are independent and uniformly distributed RVs over the interval $(0, t_{\text{pack}}]$ and i is an integer. The number of packets that need to be sent every t_{period} seconds is m. Therefore, we define the channel load as

$$g = \frac{mt_{\text{pack}}}{t_{\text{period}}}. \tag{4.1}$$

The throughput shows how efficiently the MAC protocol uses the time resource and is defined as follows. During the period t_{period}, every user wants to transmit one packet, which all other users want to receive. Hence, there are $m(m-1)$ packets that need to be communicated. Every successfully transmitted packet amounts to $(m-1)$ successfully received packets. We define the throughput as the average number of successfully received messages during t_{period} over the total number of packets that need to be received, i.e.,

$$\mathcal{T} = \frac{\bar{r}(m-1)}{m(m-1)} = \frac{\bar{r}}{m}, \tag{4.2}$$

where $\bar{r}$ is the average number of successfully transmitted packets during t_{period}. Under the collision channel model, the achievable throughput is upperbounded by one. Higher throughputs can potentially be achieved if other channel models with the capture effect are considered [73, 87].

In [Paper C], we additionally consider the packet erasure channel (PEC), i.e., we assume that a packet may not be decoded by a receiver with a certain probability ϵ. To combat the effect of packet erasures, users attempt several transmissions of the same packet, which provides time diversity at low channel loads. At moderate and high channel loads (say $g > 0.4$) repeating packets congests the channel and increases the probability of a collision.

In this thesis, we propose an alternative to CSMA, which addresses all the properties of VCs in Section 4.1. The proposed protocol, termed all-to-all broadcast coded slotted ALOHA (B-CSA), is based on CSA. As the TDMA-based protocols, it also uses the

slotted time structure but does not require any learning phase. Instead, users are allowed to transmit several times within a frame and SIC is performed at the receiver. Clearly, this protocol provides low energy efficiency because of the packet repetition and the implementation of SIC entails high complexity. However, as discussed in Section 4.1, this can be afforded in VCs. In [Paper C], we present the main idea of B-CSA and compare it with CSMA. [Paper D] extends the model to the case where packets have different reliability requirements. In [Paper E], we present the idea of virtual acknowledgements enabled by B-CSA. The next chapter provides the background of the CSA protocol.

Chapter 5

Coded Slotted ALOHA

CSA is an uncoordinated multiple access protocol that can provide large throughputs close to those of coordinated schemes. It was originally proposed for satellite communications, where CSMA cannot be used due to large propagation delays [30]. CSA is also appealing for other scenarios where carrier sense fails to provide an efficient channel access, for instance, when the number of contending users is large and the size of transmitted packets is small, or when the mechanisms against the hidden terminal problem cannot be invoked.

As discussed in Section 1.2, CSA is based on the ideas of slotted ALOHA [26, 28], where time is divided into slots and users choose slots at random for their transmission. Additionally, users in CSA employ coding over packets and SIC is used at the receiver.

In this thesis, we consider a framed version of CSA, in which n slots are grouped into a frame and contention is carried out on a frame basis. The contending users encode their messages into multiple packets, which are transmitted to a common receiver in randomly selected slots within a frame. The receiver buffers the entire frame, decodes the packets from the slots without collision and attempts to reconstruct the packets in collision exploiting coding over packets. The interference caused by the reconstructed packets is then subtracted from the buffered signal. Decoding repeats until all users are resolved or there are no slots without collision left.

When energy efficiency is an important issue, high-rate codes need to be used. However, high throughputs cannot be achieved in this case [32]. As discussed in the previous section, energy efficiency is not crucial in VCs, hence, we consider low rate codes that can asymptotically provide throughput one [88]. More specifically, we focus on repetition coding that has trivial encoding and decoding at the transmitter and the receiver side, respectively. In particular, encoding amounts to repeating a packet by a transmitter, and hence, decoding is complete once at least one copy is successfully decoded.

Apart from the framed CSA, other variations of CSA can be found in the literature. In [89], a frameless CSA was proposed, in which the receiver initiates and terminates the contention by sending a pilot signal. The frame length is therefore not fixed a priori but is adjusted depending on the number of contending users. An unslotted version of CSA was proposed in [90], where users can transmit packets at any point in time within a frame. It is assumed that partial collisions can be decoded if the interference does not exceed a certain threshold. Further improvements were obtained in [91] under the assumption that multiple copies of a packet, although in collision, can be combined for decoding. Frame asynchronous CSA was considered in [92, 93], where users still operate on a frame basis

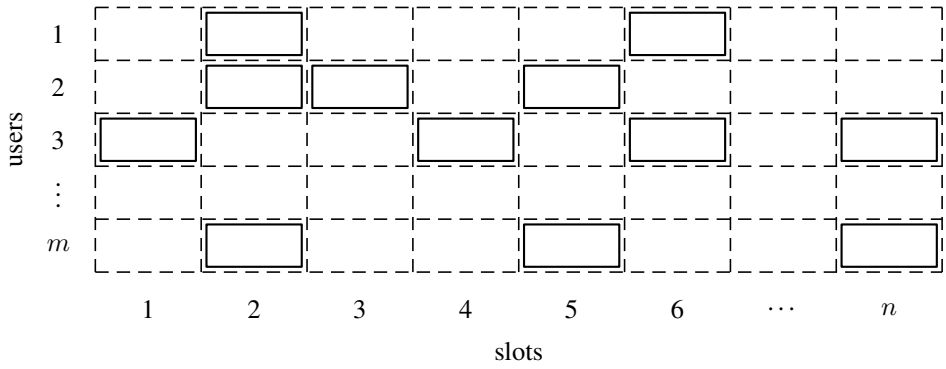


Figure 5.1: User transmissions in CSA. Rectangles represent transmitted packets.

but frames of different users are not aligned. In [94], CSA was extended to a multiple receiver case, where different cooperation strategies between the receivers are analyzed.

Other works that exploit a similar idea to CSA are worth mentioning. In [95] (see also [96]), a mechanism to combat the hidden terminal problem in CSMA, termed ZigZag was proposed. It makes use of several collisions of the same packets with different delays. Since packets do not overlap perfectly, parts of colliding packets can be decoded and SIC can be used to recover packets piece by piece. In [97], the idea of multiple packet transmissions and SIC was combined with network coding.

5.1 System Model

When analyzing MAC protocols, it is usually assumed that the number of contending users is a RV [32], [89]. For the sake of simplicity, in this thesis we consider a fixed number of users, m, which transmit to a common receiver over a shared medium. We assume that time is divided into frames and each frame is divided into n slots of equal duration. The users are assumed to be frame- and slot-synchronized, which can be achieved by means of, e.g., GPS. Each user transmits one packet per frame and no acknowledgement is used by the receiver to confirm a successful reception. This models the periodic message exchange in VCs discussed in Section 4.1. We therefore do not consider buffering at the transmitter side, as it is usually done for ALOHA protocols [98], and hence, stability is not an issue in this thesis.

In the beginning of each frame, all contending users draw a random number l according to a predefined PMF and repeat their packet l times in randomly and uniformly selected slots within a frame. l is called the repetition factor or the degree of a user. Every packet contains pointers to its copies, so that, once a packet is successfully decoded, full information about the location of the copies is available. Since frames are independent, it is sufficient to consider the system over one frame. An example of user transmissions within one frame is shown in Fig. 5.1.

We consider the following channel model. The received signal buffered by the receiver

in slot i, $i = 1, \ldots, n$, is

$$y_i = \sum_{j \in \mathcal{U}_i} h_{i,j} a_j,$$

where a_j is a packet of the jth user, $h_{i,j}$ ($|h_{i,j}| > 0$) is the channel coefficient between user j and the receiver, and $\mathcal{U}_i$ is the set of users that transmit in the ith slot. A slot is called a *singleton* slot if it contains only one packet. If it contains more packets, we say that a collision occurs in the slot.

Decoding proceeds as follows. The receiver decodes the packets in singleton slots and obtains the location of their copies. We assume that packets in singleton slots are always successfully decoded. The channel coefficients corresponding to the copies are then estimated using data-aided methods (since the entire packet can be used as a pilot) and the interference caused by these copies is subtracted. We assume perfect interference cancellation, which is justified by numerical results in [31]. Decoding proceeds until no further singleton slots are found.

Other channel models were analyzed in the literature as well. In [99], the capture effect was analyzed by means of the signal-to-interference-plus-noise ratio (SINR) channel model. The SINR channel model was also used in [90, 91] for the unslotted version of CSA. The PEC was considered in [100] and [101], where the channel coefficient is equal to zero with a certain probability. A more realistic channel model that includes the PHY layer was analyzed in [102]. We remark that the system relies on decoding packets in singleton slots, hence, the users code rates used at the PHY layer do not have to satisfy the rate constraint for joint decoding [103] over more realistic channel models. However, for low enough code rates the capture effect comes into play.

The PMF that users use for transmission is usually defined in the form of a polynomial

$$\Lambda(x) = \sum_{l=0}^{q} \Lambda_l x^l, \tag{5.1}$$

where Λ_l is the probability of a user to have degree l and q is the maximum degree.[6] The choice of the maximum degree is limited by the overhead due to the pointers as well as the frame length. Usually, $q \ll n$. The key performance parameters of CSA protocols are defined as follows. The channel load

$$g = m/n \tag{5.2}$$

shows how "busy" the medium is. The throughput is defined as

$$\mathcal{T} = r/n, \tag{5.3}$$

where r is the average number of resolved users. The throughput shows how efficiently slots are utilized. For example, the throughput of slotted ALOHA is 0.37 users/slot. The maximum achievable throughput under the considered channel model is equal to 1.

[6]Note that $\lambda(x)$ is used in Papers C and E to denote this PMF.

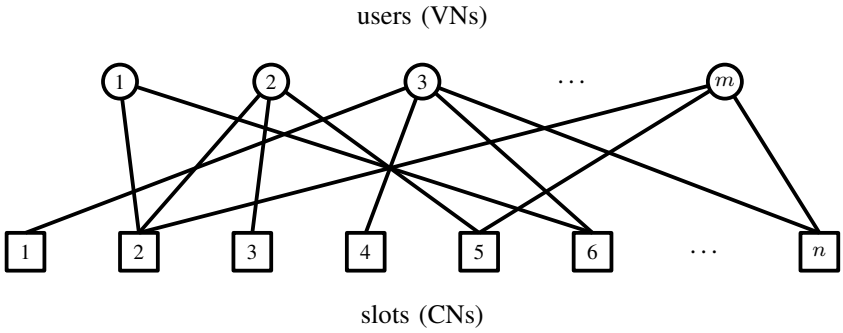


Figure 5.2: Graph representation of CSA.

Higher throughputs can be achieved over channels with capture effect [99]. The packet loss rate (PLR) is the probability that a user is not resolved by the receiver. Since all users are independent, the PLR can be expressed as

$$p = \frac{m - r}{m} = 1 - \frac{\mathcal{T}}{g}. \tag{5.4}$$

In this thesis, we are mostly interested in the PLR, since reliability is one of the key performance parameters in VCs. Finally, the channel access delay shows how long it takes for a user to successfully transmit a packet [27]. For framed CSA, the delay is upperbounded by the frame length. The channel access delay is studied in [Paper D] in more detail.

5.2 Graph Model

It was recognized in [31] that the described decoding algorithm is similar to the peeling decoding of graph-based codes over the binary erasure channel (BEC). The CSA system is similar to an ensemble of low-density generator matrix codes [104] and it can be represented by a bipartite graph, in which each user is represented by a variable node (VN) and each slot is represented by a check node (CN). A VN corresponds to a repetition code and a CN corresponds to a single parity-check code. A bipartite graph is defined by $\mathcal{G} = (\mathcal{V}, \mathcal{C}, \mathcal{E})$, where $\mathcal{V}$, $\mathcal{C}$, and $\mathcal{E}$ represent the sets of VNs, CNs, and edges connecting them, respectively. An edge $e_{j,i}$ connects the jth VN and the ith CN if the jth user transmits in the ith slot. The number of edges connected to a node is called the node degree. Using graph terminology, the PMF in (5.1) is the (node-perspective) VN degree distribution. The graph in Fig. 5.2 represents the transmissions shown in Fig. 5.1.

Check Node Degree Distribution

Similarly to irregular LDPC codes [105], the VN and CN degree distributions have strong influence on the performance of the system. However, unlike LDPC codes, where both distributions can be controlled, the CN degree distribution in CSA cannot be directly

manipulated. From this perspective, CSA resembles Luby Transform (LT) codes [106], where the VN degree distribution is induced by the CN degree distribution. Analogously to (5.1), we define the (node-perspective) CN degree distribution as

$$\Psi(x) = \sum_{l=0}^{m} \Psi_l x^l, \tag{5.5}$$

where Ψ_l is the probability that a CN has degree l. The highest degree, m, corresponds to the case where all users transmit in one slot.

In the following, we derive the CN degree distribution. To that end, we select a CN at random. Let us denote the degree of the jth VN by L_j. We associate with the jth VN a RV X_j that takes on value 1 if the jth VN is connected to the selected CN and 0 otherwise. The corresponding probabilities are L_j/n and $1 - L_j/n$, respectively. The probability generating function of the RV X_j is

$$G_{X_j}(z) = \frac{L_j}{n} + \left(1 - \frac{L_j}{n}\right) z. \tag{5.6}$$

The degree of the CN is then a RV that can be written as

$$Y = \sum_{j=1}^{m} X_j. \tag{5.7}$$

Since X_j are independent for all $j = 1, \ldots, m$, the probability generating function of Y is

$$G_Y(z) = \mathrm{E}\left\{\prod_{j=1}^{m} G_{X_j}(z)\right\} = \mathrm{E}\left\{\prod_{j=1}^{m} \left(\frac{L_j}{n} + \left(1 - \frac{L_j}{n}\right) z\right)\right\}. \tag{5.8}$$

Furthermore, since L_j are independent and identically distributed, we have

$$G_Y(z) = \prod_{j=1}^{m} \mathrm{E}\left\{\frac{L_j}{n} + \left(1 - \frac{L_j}{n}\right) z\right\} = \left(\frac{\bar{l}}{n} + \left(1 - \frac{\bar{l}}{n}\right) z\right)^m, \tag{5.9}$$

where

$$\bar{l} = \mathrm{E}\{L_j\} = \sum_{l=1}^{q} l\Lambda_l = \Lambda'(1) \tag{5.10}$$

is the average number of edges incident to a VN, where the prime denotes the derivative. From (5.9) it follows that the CN degree follows the binomial distribution, i.e., the probability that a node has degree l is

$$\Psi_l = \binom{m}{l} \left(\frac{\bar{l}}{n}\right)^l \left(\frac{\bar{l}}{n}\right)^{m-l} \tag{5.11}$$

and

$$\Psi(x) = \left(\frac{\bar{l}}{n} x + \left(1 - \frac{\bar{l}}{n}\right)\right)^m = \left(1 - \frac{\bar{l}}{n}(1 - x)\right)^m. \tag{5.12}$$

Edge-Perspective Degree Distributions

The distributions in (5.1) and (5.5) specify the probability of a node to have a certain degree. It is also useful to consider the probability that a randomly selected edge is connected to a node of a certain degree. Let λ_l be the probability that an edge is connected to a VN of degree l. The average number of edges in the graph is equal to $\bar{l}m$, whereas the average number of edges connected to degree-l VNs is $l\Lambda_l m$. Hence,

$$\lambda_l = \frac{l\lambda_l}{\bar{l}}. \tag{5.13}$$

For convenience, we define the edge-perspective VN degree distribution as

$$\lambda(x) = \sum_{l=1}^{q} \lambda_l x^{l-1}. \tag{5.14}$$

Analogously, we define the edge-perspective CN degree distribution as

$$\rho(x) = \sum_{l=1}^{m} \rho_l x^{l-1}, \tag{5.15}$$

where

$$\rho_l = \frac{l\Psi_l n}{\bar{l}m} = \frac{l\Psi_l}{\bar{l}g}. \tag{5.16}$$

The degree distributions in (5.1) and (5.5) can be related to those in (5.14) and (5.15) as $\lambda(x) = \Lambda'(x)/\Lambda'(1)$ and $\rho(x) = \Psi'(x)/\Psi'(1)$, respectively. The edge-perspective degree distributions are used in the following section to derive density evolution (DE) equations.

5.3 Asymptotic Analysis

Decoding of LDPC codes is usually done by means of iterative message passing on the graph, i.e., when nodes exchange reliability messages with each other. The performance of a code can be estimated by tracking distributions of these messages over iterations. In the asymptotic case, when $n \to \infty$, this can be done semi-analytically by means of the so-called DE [107]. For the BEC, DE assumes a particularly simple form, where a single parameter, the probability that a message is erased, needs to be tracked. It turns out that this analysis can be readily applied to CSA.

It was shown in [107] that the neighborhood of any node down to a finite depth in a randomly generated graph looks like a tree with high probability. This tree is shown in Fig. 5.3 and is called a computation tree for a randomly chosen VN. The performance of a randomly generated code can then be estimated by means of the and-or tree evaluation [108] run on the computation tree.

Let ϵ_ℓ denote the probability that a CN is not resolved at the ℓth iteration. In other words, ϵ_ℓ is the probability that a CN sends an erasure. A degree-d VN is not resolved at the ℓth iteration if all incoming messages are erasures, which occurs with probability

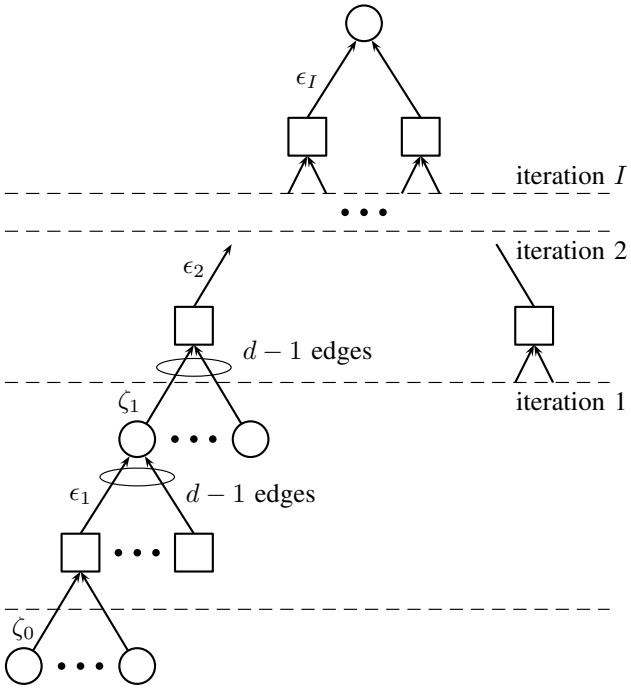


Figure 5.3: Computation tree.

ϵ_ℓ^{d-1}. The outcoming edge is connected to a degree-d VN with probability λ_d, hence, the probability that a VN transmits an erasure is

$$\zeta_\ell = \lambda(\epsilon_\ell). \tag{5.17}$$

A degree-d CN is not resolved at iteration $\ell+1$ if at least one incoming message is an erasure, which occurs with probability $1-(1-\zeta_\ell)^{d-1}$. Since the outcoming edge is connected to a degree-d CN with probability ρ_d, the probability that a CN transmits an erasure is

$$\epsilon_{\ell+1} = \sum_d \rho_d(1-(1-\zeta_\ell)^{d-1}) = 1-\rho(1-\zeta_\ell). \tag{5.18}$$

The probability that a VN is not resolved after ℓ iterations is given by

$$p_\ell = \Lambda(\epsilon_\ell). \tag{5.19}$$

When the frame length goes to infinity (as well as the number of users for a fixed channel load), the CN degree distribution in (5.12) converges to the Poisson distribution

$$\Psi(x) = \exp(-g\bar{l}(1-x)), \tag{5.20}$$

where $\bar{l}$ is defined in (5.10) and the edge-perspective distribution has exactly the same form, i.e.,

$$\rho(x) = \exp(-g\bar{l}(1-x)). \tag{5.21}$$

Hence, using (5.17) in (5.18), we have

$$\epsilon_{\ell+1} = 1 - \exp(-g\bar{l}\lambda(\epsilon_\ell)) = 1 - \exp(-g\Lambda'(\epsilon_\ell)). \tag{5.22}$$

DE allows to predict the performance of a CSA system when the frame length n goes to infinity. The typical performance in this case exhibits a threshold behavior, i.e., most of the users are successfully resolved if the channel load is below a certain value. This value is called the threshold and is denoted by g^*. The threshold can be found as the maximum channel load g such that

$$\epsilon > 1 - \exp(-g\Lambda'(\epsilon)) \tag{5.23}$$

for any $\epsilon \in (0, 1]$. As it can be seen from (5.23), the threshold depends on the VN degree distribution only. It was shown in [88] that the so-called soliton distribution provides the threshold equal to one. However, this distribution is impractical since the maximum VN degree tends to infinity. It is possible to obtain thresholds close to one when the maximum degree is finite. For example, the distribution

$$\Lambda(x) = 0.5x^2 + 0.28x^3 + 0.22x^8, \tag{5.24}$$

optimized under the constraint $q \leq 8$ in [31], gives $g^* = 0.938$. The evolution of the erasure probabilities for $g = 0.9$ is shown in Fig. 5.4 in the form of an EXIT chart [109]. The performance of the system for the distribution in (5.24) and two values of n is shown in Fig. 5.5.

We remark that there is a subtle difference when interpreting the DE results for LDPC codes and CSA. DE gives the performance of the message passing algorithm over a random graph, or the so-called ensemble average performance. The graph in CSA is indeed random and it can be seen as an ensemble. Hence, DE predicts the performance of the CSA system when $n \to \infty$. However, random codes cannot be used in practice and DE for LDPC codes relies on the concentration phenomenon [107], which states that a randomly chosen code with high probability performs very close to the ensemble average if n is large.

5.4 Finite Frame Length Analysis

From the throughput curves in Fig. 5.5(a) it appears that DE predicts the performance for relatively short frame lengths. However, as we can see from Fig. 5.5(b), the PLR performance for finite n is far from the DE results. The main reason for this discrepancy is that the tree assumption is not valid for finite n, and hence, loops are present in the graph. These loops form stopping sets [110] that cause the appearance of the so-called EF region in the PLR performance.

Stopping sets for graph-based codes are well investigated. However, the obtained results are not directly applicable to CSA. As mentioned in the previous section, random

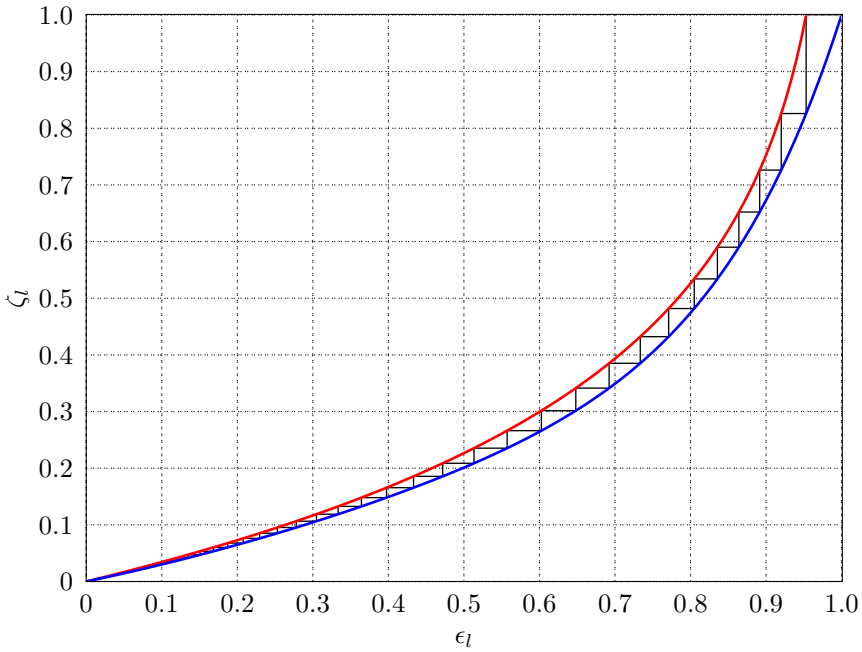
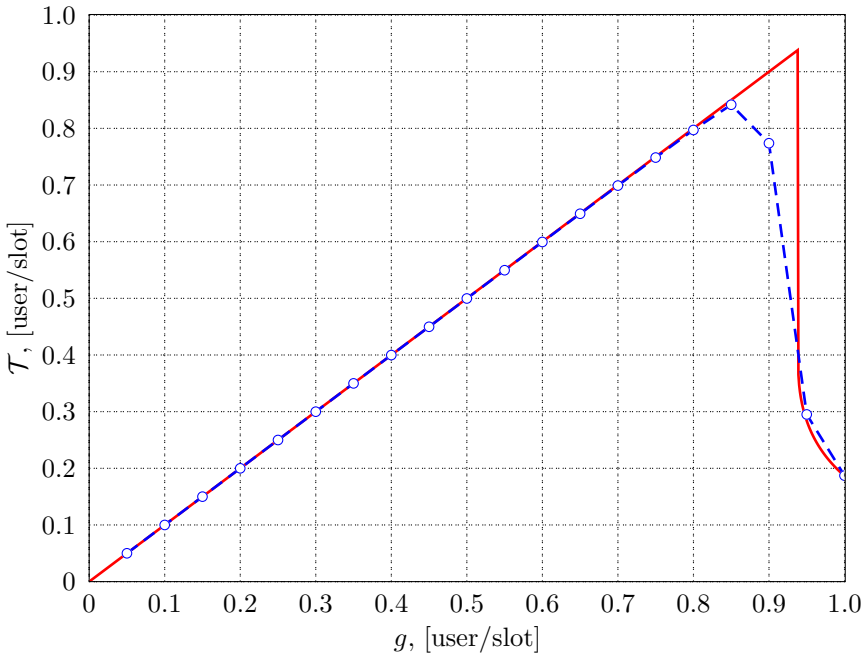


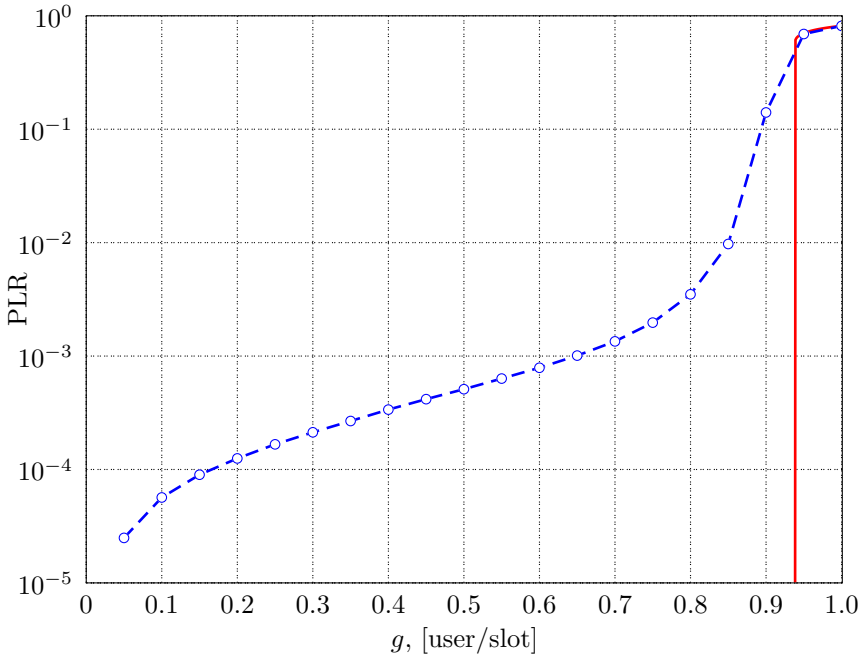
Figure 5.4: Evolution of the erasure probability for the distribution in (5.24) and $g = 0.9$. The blue and the red lines correspond to (5.17) and (5.18), respectively.

codes cannot be used in practice. Hence, the analysis of stopping sets for LDPC codes usually aims at constructing a good code by avoiding stopping sets [111, 112] or analyzing a fixed code [113, 114]. CSA, on the other hand, is intrinsically random and can be seen as a code ensemble.

The exact performance of an ensemble of LDPC codes based on stopping sets was derived in [110], where the authors derive recursive combinatorial equations that take into account all possible stopping sets. This analysis, however, becomes intractable for code lengths of interest and irregular codes. In [115], the distribution of stopping sets for a code ensemble was obtained. This distribution allows to obtain an upper bound on the ensemble average performance. A similar approach was taken in [116] to upperbound the PLR performance of CSA in the EF region. However, the obtained results give loose bounds and have high complexity for irregular graphs. In [101], we proposed a low complexity EF approximation for CSA based on the heuristically determined "dominant" stopping sets. In [Paper C], we extend these results and propose a systematic way of determining dominant stopping sets together with their probabilities. In [Paper D], we apply this analysis to the case where users are assigned different distributions in order to provide unequal error protection.



(a) Throughput.



(b) PLR performance.

Figure 5.5: Performance for the distribution in (5.24). Solid red lines show DE results, i.e., when $n \to \infty$. Dashed blue lines show simulation results for $n = 1000$.

Chapter 6

Contributions and Future Work

6.1 Contributions

This chapter summarizes the contributions of this thesis, which are presented in Part II of the thesis in the form of appended papers. The main objective of [Papers A and B] is to analyze bit-wise decoders that employ the max-log approximation for calculating L-values for coded modulation over the AWGN channel. [Paper A] analyzes the information rate loss caused by the max-log approximation as compared to the exact L-values. [Paper B] studies the performance loss of such decoders in terms of the probability of error. [Papers C-E] are dedicated to CSA in VCs. In [Paper C], we propose the B-CSA protocol for all-to-all broadcast in VANETs and analyze it in the finite frame length regime. [Paper D] extends this analysis to the case where packets have different reliability requirements. Finally, [Paper E] presents the idea of virtual acknowledgements in B-CSA that could be used for handshake in VANETs.

6.1.1 Paper A: "On the Information Loss of the Max-Log Approximation in BICM Systems"

It is widely assumed that the calculation of L-values using the max-log approximation leads to an information loss. In this paper, we study the max-log approximation for one dimensional constellations and show that this assumption is not correct for a symmetric 4-PAM constellation labeled with a Gray code. Further, we prove that for all other combinations of M-PAM constellations and labelings the assumption is indeed correct. We also show that the GMI in (2.19) is too pessimistic for max-log L-values and provide a better bound on the achievable rate of the bit-wise decoder in (2.18), which coincides with the so-called "harmonized" GMI [55].

We then analyze bit-level channel symmetrization and mixing that are often used to simplify the analysis of BICM schemes. Our results show that these operations preserve the GMI but decrease the MI. However, numerical results show that for relatively high SNR the differences between the considered achievable rates become negligible, which justifies the use of the max-log approximation in practical systems and the considered processing techniques for their analysis.

6.1.2 Paper B: "On the Asymptotic Performance of Bit-Wise Decoders for Coded Modulation"

In this paper, we analyze the system model in Fig. 3.4 and compare the bit-wise decoder with the ML decoder for equally-spaced 4-PAM constellations with a Gray labeling. We find the SMD distributions for the bit-wise decoder when max-log L-values are used and compare them with the SMD distributions for the S-DEC. This allows us to bound the asymptotic performance loss in terms of the PEP for any pair of codewords. We show that the asymptotic loss for codes is zero for two out of the four nonequivalent Gray labelings if the underlying binary code $\mathcal{B}$ is linear. For other Gray labelings, including the BRGC, the analysis shows that the asymptotic loss is zero for a wide range of linear codes, including all rate-1/2 convolutional codes. Using the TCM equivalence discussed in Section 3.1 we conclude that, for 4-PAM, the performance of TCM schemes can be achieved by BICM systems without an interleaver.

6.1.3 Paper C: "Broadcast Coded Slotted ALOHA: A Finite Frame Length Analysis"

This paper extends CSA in Section 5 to the all-to all broadcast scenario, where each user acts as both transmitter and receiver in a half-duplex mode. The proposed protocol is termed B-CSA and suits well the transmission of CAM packets in VANETs. We analyze the performance of B-CSA in the finite frame length regime. We provide a general analysis of stopping sets and derive an analytical approximation of the performance in the EF region. In the finite frame length regime, we observe an unequal error protection (UEP) feature from the transmitter perspective, i.e., the more a user repeats its packet, the higher the probability that this packet is decoded by other users. The half-duplex mode gives rise to a similar property from the receiver perspective: the more a user transmits, the lower the probability for this user to decode packets from others. Finally, we compare B-CSA with the CSMA protocol.

6.1.4 Paper D: "Unequal Error Protection in Coded Slotted ALOHA"

This paper aims at exploiting the UEP property from the transmitter perspective in order to provide different protection levels to different classes of users. We extend the analytical EF approximation in [Paper C] to the case of multiple classes of users. We further propose a heuristic approach for the optimization of the degree distributions to provide the required UEP. Finally, we analyze the decoding delay for users in different classes. The presented analysis can be applied to both classical CSA and B-CSA.

6.1.5 Paper E: "Probabilistic Handshake in All-to-all Broadcast Coded Slotted ALOHA"

The half-duplex mode in B-CSA implies that only partial knowledge of the graph is available to users. However, the graph can be reconstructed with a certain level of confidence. This reconstructed graph can be used by a user to conclude whose packets it received. We use this to perform handshakes between users, i.e., each user decides whether its packet was successfully received by the other users. To that end, we consider a fully connected network, where users use B-CSA. The analysis shows that in the rare cases of communication failure between two users, this event can be reliably detected by one of the users. This information can be used for safety-critical applications in VANETs.

6.2 Future Work

In [Paper A], we showed that the exact L-value can be recovered from the max-log L-value in some cases. We can take one step further and recover the entire observation from the vector of the L-values. Obviously, it is possible in some cases. For instance, the L-value (exact or max-log) for the first bit position of any PAM constellation labeled with the BRGC is a monotonic function of the observation, and hence, just one L-value is enough to reconstruct the observation. This means that the calculation of L-values itself does not cause any information loss. It is interesting to see how this observation extends to other constellations and other labelings.

The results of [Paper C] rely on the ZC approximation for max-log L-values. Even though it has been shown numerically in [69], [117], [66] that the ZC approximation is good in terms of predicting the coded BER, the proof for its asymptotic tightness is an open question. Moreover, the ZC approximation has never been thoroughly studied for other than Gray labeling, which could be a future work direction.

Whether B-CSA can provide a reliable channel access in VANETs is still an open question. To answer this question, it is important to understand how accurately the channel estimation can be performed for vehicular channels. We also need to figure out if the PHY layer in 802.11p can be used in conjunction with B-CSA. It is quite possible that in order to make B-CSA work in VANETs, the PHY layer needs to be revisited.

The analysis of CSA in this thesis assumes a fixed number of neighbors in the network. Obviously, the number of neighbors in a VANET is random, which needs to be taken into account. To that end, we need to find good distributions of neighbors for VCs. These distributions will probably depend on the communication scenario (for instance, urban or rural) and on the use of power control algorithms.

The following extension is not related to VCs and can be applied to all CSA schemes. As of now, we have a good understanding of the PLR performance in the EF region, as well as in the waterfall region for large frame lengths. However, the waterfall for short frame lengths needs further investigation. It is therefore interesting to apply finite length scaling [118] of LDPC code ensembles to CSA. In [Papers C and D], we resort to heuristic

criteria based on the EF approximation and the threshold to search for good distributions that provide a required level of reliability. Accurate PLR approximations in the waterfall region will greatly simplify this search.

The results of [Paper E] are obtained under the assumption that the network is fully connected, which is unrealistic in VANETs. Extending these results to partially connected networks and accounting for packet erasures can be of great interest. Finally, the idea of reconstructed graphs in [Paper E] can be exploited in other scenarios. For instance, one can think of a network, where, in addition to contending users, a base station is present in the network. The base station listens to users' transmissions, and thus, it observes the entire graph. By reconstructing graphs observed by other users, the base station can figure out which users missed whose packets. It can therefore take some action, for instance, broadcast the most missed packet to all users.

References

[1] C. E. Shannon, "A mathematical theory of communications," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and Oct. 1948.

[2] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets, Part I: Introduction," *IEEE Communications Magazine*, vol. 25, no. 2, pp. 5–11, Feb. 1987.

[3] J. L. Massey, "Coding and modulation in digital communications," in *Proc. Int. Zurich Seminar on Digital Commun.*, Zurich, Switzerland, Mar. 1974.

[4] G. Ungerboeck and I. Csajka, "On improving data-link performance by increasing channel alphabet and introducing sequence decoding," in *Proc. Int. Symp. Inf. Theory*, Ronneby, Sweden, June 1976, (Book of abstracts).

[5] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 1, pp. 55–67, Jan. 1982.

[6] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. 13, no. 2, pp. 260–269, Apr. 1967.

[7] H. Imai and S. Hirakawa, "A new multilevel coding method using error-correcting codes," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 3, pp. 371–377, May 1977.

[8] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 3, pp. 927–946, May 1992.

[9] ETSI, "Digital video broadcasting (DVB); Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2)," ETSI, Tech. Rep. ETSI EN 302 755 V1.3.1 (2012-04), Apr. 2012.

[10] IEEE 802.11, "Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," IEEE Std 802.11-2012, Tech. Rep., Mar. 2012.

[11] ETSI, "LTE; Evolved universal terrestrial radio access (E-UTRA); Physical channels and modulation," ETSI, Tech. Rep. ETSI TS 136 211 V11.2.0 (2013-04), Apr. 2013.

[12] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *Proc. IEEE Int. Conf. Commun.*, Geneva, Switzerland, May 1993.

[13] R. Gallager, "Low-density parity-check codes," *IRE Trans. Information Theory*, pp. 21–28, Jan. 1962.

[14] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Electronics Lett.*, vol. 33, no. 6, pp. 457–458, Mar. 1997.

[15] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Bandwidth efficient parallel concatenated coding schemes," *Electronics Lett.*, vol. 31, no. 24, pp. 2067–2069, Nov. 1995.

[16] ——, "Parallel concatenated trellis coded modulation," in *Proc. IEEE Int. Conf. Commun.*, Dallas, TX, June 1996.

[17] X. Li and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 1, no. 6, pp. 169–171, Nov. 1997.

[18] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *Proc. IEEE Global Telecommun. Conf.*, Sydney, Australia, Nov. 1998.

[19] P. Popovski, "Ultra-reliable communication in 5G wireless systems," in *Proc. IEEE Int. Conf. 5G for Ubiquitous Connectivity*, Levi, Finland, Nov. 2014.

[20] N. A. Johansson, Y. Wang, E. Eriksson, and M. Hessler, "Radio access for ultra-reliable and low-latency 5G communications," in *Proc. IEEE Int. Conf. Commun.*, London, UK, June 2015.

[21] G. Durisi, T. Koch, and P. Popovski, "Towards massive, ultra-reliable, and low-latency wireless communication with short packets," *Proc. of the IEEE (to appear)*, 2016.

[22] K. Sjöberg, "Medium access control for vehicular ad hoc networks," Ph.D. dissertation, Chalmers University of Technology, Halmstad University, Göteborg, Sweden, 2013.

[23] A. Chandra, V. Gummalla, and J. Limb, "Wireless medium access control protocols," *IEEE Commun. Surveys*, pp. 2–15, 2000.

[24] H. Menouar, F. Filali, and M. Lenardi, "A survey and qualitative analysis of MAC protocols for vehicular ad hoc networks," *IEEE Wireless Commun.*, vol. 13, no. 5, pp. 30–35, Oct. 2006.

[25] L. Kleinrock and F. Tobagi, "Packet switching in radio channels: Part I–carrier sense multiple-access modes and their throughput-delay characteristics," *IEEE Trans. Commun.*, vol. 23, no. 12, pp. 1400–1416, Dec. 1975.

[26] N. Abramson, "Another alternative for computer communications," in *Proc. Fall Joint Computer Conference*, 1970.

[27] K. Bilstrup, E. Uhlemann, E. Ström, and U. Bilstrup, "On the ability of the 802.11p MAC method and STDMA to support real-time vehicle-to-vehicle communication," *EURASIP Journal on Wireless Commun. and Networking*, vol. 2009, pp. 1–13, Apr. 2009.

[28] L. G. Roberts, "ALOHA packet system with and without slot and capture," *SIGCOMM Comput. Commun. Rev.*, vol. 5, no. 2, pp. 28–42, Apr. 1975.

[29] G. Choudhury and S. Rappaport, "Diversity ALOHA-A random access scheme for satellite communications," *IEEE Trans. Commun.*, vol. 31, no. 3, pp. 450–457, Mar. 1983.

[30] E. Casini, R. D. Gaudenzi, and O. Herrero, "Contention resolution diversity slotted ALOHA (CRDSA): An enhanced random access schemefor satellite access packet networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 1408–1419, Apr. 2007.

[31] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted ALOHA," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 477–487, Feb. 2011.

[32] E. Paolini, G. Liva, and M. Chiani, "Coded slotted ALOHA: A graph-based method for uncoordinated multiple access," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6815–6832, Dec. 2015.

[33] E. Agrell, J. Lassing, E. G. Ström, and T. Ottosson, "On the optimality of the binary reflected Gray code," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3170–3182, Dec. 2004.

[34] G. D. Forney, Jr. and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2384–2415, Oct. 1998.

[35] A. Guillén i Fàbregas, A. Martinez, and G. Caire, "Bit-interleaved coded modulation," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 1–2, pp. 1–153, 2008.

[36] G. J. Foschini, R. D. Gitlin, and S. B. Weinstein, "Optimization of two-dimensional signal constellations in the presence of Gaussian noise," *IEEE Trans. Commun.*, vol. 22, no. 1, pp. 28–38, Jan. 1974.

[37] M. Barsoum, C. Jones, and M. Fitz, "Constellation design via capacity maximization," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, June 2007.

[38] A. R. Calderbank and L. H. Ozarow, "Nonequiprobable signaling on the Gaussian channel," *IEEE Trans. Inf. Theory*, vol. 36, no. 4, pp. 726–740, July 1990.

[39] A. Guillén i Fàbregas and A. Martinez, "Bit-interleaved coded modulation with shaping," in *Proc. IEEE Inf. Theory Workshop*, Cairo, Egypt, 2010, pp. 1–5.

[40] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec. 2015.

[41] C. Stierstorfer, "A bit-level-based approach to coded multicarrier transmission," Ph.D. dissertation, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, 2009, available at http://www.opus.ub.uni-erlangen.de/opus/volltexte/2009/1395/.

[42] M. Ivanov, F. Brännström, A. Alvarado, and E. Agrell, "General BER expression for one-dimensional constellations," in *Proc. IEEE Global Commun. Conf.*, Anaheim, CA, Dec. 2012.

[43] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[44] A. Lapidoth, *A Foundation in Digital Communication.* Cambridge University Press, 2009.

[45] A. Guillén i Fàbregas and A. Martinez, "Derivative of BICM mutual information," *IET Electronics Lett.*, vol. 43, no. 22, pp. 1219–1220, Oct. 2007.

[46] A. Martinez, A. Guillén i Fàbregas, and G. Caire, "Bit-interleaved coded modulation in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5447–5455, Dec. 2008.

[47] A. Alvarado, F. Brännström, E. Agrell, and T. Koch, "High-SNR asymptotics of mutual information for discrete constellations with applications to BICM," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1061–1076, Feb. 2014.

[48] A. Martinez, A. Guillén i Fàbregas, G. Caire, and F. Willems, "Bit-interleaved coded modulation revisited: A mismatched decoding perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2756–2765, June 2009.

[49] J. Jaldén, P. Fertl, and G. Matz, "On the generalized mutual information of BICM systems with approximate demodulation," in *Proc. IEEE Inf. Theory Workshop*, Jan. 2010.

[50] A. Ganti, A. Lapidoth, and İ. E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, Nov. 2000.

[51] A. J. Viterbi, "An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 260–264, Feb. 1998.

[52] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proc. IEEE Int. Conf. Commun.*, Seattle, WA, June 1995.

[53] K. Hyun and D. Yoon, "Bit metric generation for Gray coded QAM signals," *IEEE Proc.-Commun.*, vol. 152, no. 6, pp. 1134–1138, Dec. 2005.

[54] M. S. Raju, R. Annavajjala, and A. Chockalingam, "BER analysis of QAM on fading channels with transmit diversity," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 481–486, Mar. 2006.

[55] T. Nguyen and L. Lampe, "Bit-interleaved coded modulation with mismatched decoding metrics," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 437–447, Feb. 2011.

[56] L. Szczecinski, "Correction of mismatched L-values in BICM receivers," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3198–3208, Nov. 2012.

[57] M. Ivanov, F. Brännström, A. Alvarado, and E. Agrell, "On the exact BER of bit-wise demodulators for one-dimensional constellations," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1450–1459, Apr. 2013.

[58] ITU, "A duplex modem operating at data signalling rates of up to 14 400 bit/s for use on the general switched telephone network and on leased point-to-point 2-wire telephone-type circuits," ITU-T Recommendation V.32 bis, Tech. Rep., Feb. 1991.

[59] W. E. Ryan and S. Lin, *Channel codes: Classical and Modern*. Cambridge University Press, 2009.

[60] L.-F. Wei, "Trellis-coded modulation with multidimensional constellations," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 4, pp. 483–501, July 1987.

[61] A. Alvarado, A. Graell i Amat, F. Brännström, and E. Agrell, "On optimal TCM encoders," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2178–2189, Apr. 2013.

[62] K. Larsen, "Short convolutional codes with maximal free distance for rates 1/2, 1/3, and 1/4," *IEEE Trans. Inf. Theory*, vol. 19, no. 3, pp. 371–372, May 1973.

[63] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2000.

[64] Y. Li and W. E. Ryan, "Bit-reliability mapping in LDPC-coded modulation systems," *IEEE Commun. Lett.*, vol. 9, no. 1, pp. 1–3, Jan. 2005.

[65] C. Stierstorfer, R. F. H. Fischer, and J. B. Huber, "Optimizing BICM with convolutional codes for transmission over the AWGN channel," in *Proc. Int. Zurich Seminar on Commun.*, Zurich, Switzerland, Mar. 2010.

[66] A. Alvarado, L. Szczecinski, and E. Agrell, "On BICM receivers for TCM transmission," *IEEE Trans. Commun.*, vol. 59, no. 10, pp. 2692–2707, Oct. 2011.

[67] L. C. Perez, J. Seghers, and D. J. Costello, Jr., "A distance spectrum interpretation of turbo codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 16, pp. 1698–1709, Nov. 1996.

[68] A. Alvarado, L. Szczecinski, and E. Agrell, "On the performance of BICM with trivial interleavers in nonfading channels," in *Proc. IEEE Int. Conf. Commun.*, Kyoto, Japan, June 2011.

[69] M. Benjillali, L. Szczecinski, S. Aissa, and C. Gonzalez, "Evaluation of bit error rate for packet combining with constellation rearrangement," *Wiley Journal Wireless Comm. and Mob. Comput.*, vol. 8, no. 7, pp. 831–844, Sep. 2008.

[70] ISO, "Intelligent transport systems — Cooperative ITS — Part 2: Guidelines for standards documents," ISO, Tech. Rep. TR 17465-2:2015, 2015.

[71] 5G-PPP, "5G automotive vision," 5G-PPP, Tech. Rep. White Paper, Oct. 2015. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White-Paper-on-Automotive-Vertical-Sectors.pdf

[72] IEEE Std. 802.11-2012, "Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," Tech. Rep., Mar. 2012.

[73] M. Torrent-Moreno, S. Corroy, F. Schmidt-Eisenlohr, and H. Hartenstein, "IEEE 802.11-based one-hop broadcast communications: understanding transmission success and failure under different radio propagation environments," in *Proc. Int. Symp. Modeling Analysis and Simulation of Wireless and Mobile Systems*, Terromolinos, Spain, Oct. 2006.

[74] L. Yang, J. Guo, and Y. Wu, "Channel adaptive one hop broadcasting for VANETs," in *Proc. IEEE Int. Conf. Intelligent Transportation Systems*, Beijing, China, Oct. 2008.

[75] ETSI EN 302 637-3: Draft V0.0.7, "Intelligent transport systems (ITS); vehicular communications; basic set of applications; part 3: Specification of decentralized environmental notification basic service," Tech. Rep., June 2012.

[76] ETSI EN 302 637-2: Draft V.0.0.5, "Intelligent transport systems (ITS); vehicular communications; basic set of applications; part 2: Specification of cooperative awareness basic service," Tech. Rep., June 2012.

[77] K. K. Nagalapur, F. Brännström, E. G. Ström, F. Undi, and K. Mahler, "An 802.11p cross-layered pilot scheme for time- and frequency-varying channels and its hardware implementation," *IEEE Trans. Veh. Technol., Connected Vehicles Series. (to appear)*, 2016.

[78] T. Zemen, L. Bernadó, N. Czink, and A. F. Molisch, "Iterative time-variant channel estimation for 802.11p using generalized discrete prolate spheroidal sequences," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1222–1233, Mar. 2012.

[79] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[80] M. Torrent-Moreno, J. Mittag, P. Santi, and H. Hartenstaein, "Vehicle-to-vehicle communication: Fair transmit power control for safety-critical information," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3684–3703, Sep. 2009.

[81] C.-L. Huang, Y. Pourmohammadi, R. Sengupta, and H. Krishnan, "Intervehicle transmission rate control for cooperative active safety system," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 3, pp. 645–658, Sep. 2011.

[82] F. Borgonovo, A. Capone, M. Cesana, and L. Fratta, "RR-Aloha, a reliable R-Aloha broadcast channel for ad-hoc inter-vehicle communication networks," in *Proc. Mediterranean Ad Hoc Networking Workshop*, Baia Chia, Italy, Sept. 2002.

[83] ——, "ADHOC MAC: a new MAC architecture for ad-hoc networks providing efficient and reliable point-to-point and broadcast services," *ACM Wireless Networks*, vol. 10, no. 4, pp. 359–366, July 2004.

[84] S. K. A. P. Jayasumana, "Effect of hidden terminals on the performance of IEEE 802.11 MAC protocol," in *Proc. IEEE Conf. Local Computer Networks*, Lowell, MA, Oct. 1998.

[85] M. Ivanov, F. Brännström, A. Graell i Amat, and P. Popovski, "All-to-all broadcast for vehicular networks based on coded slotted ALOHA," in *Proc. IEEE Int. Conf. Commun. Workshop*, London, UK, June 2015.

[86] J. L. Massey and P. Mathys, "The collision channel without feedback," *IEEE Trans. Inf. Theory*, vol. 31, no. 2, pp. 192–204, Mar. 1985.

[87] J. H. Kim and J. K. Lee, "Capture effects of wireless CSMA/CA protocols in Rayleigh and shadow fading channels," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1277–1286, July 1999.

[88] K. R. Narayanan and H. D. Pfister, "Iterative collision resolution for slotted ALOHA: an optimal uncoordinated transmission policy," in *Proc. Int. Symp. on Turbo Codes & Iterative Information Processing*, Gothenburg, Sweden, Aug. 2012.

[89] C. Stefanovic and P. Popovski, "ALOHA random access that operates as a rateless code," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4653–4662, Nov. 2013.

[90] C. Kissling, "Performance enhancements for asynchronous random access protocols over satellite," in *Proc. IEEE Int. Conf. Commun.*, Kioto, Japan, June 2011.

[91] F. Clazzer, C. Kissling, and M. Marchese, "Exploiting combination techniques in random access MAC protocols: Enhanced contention resolution aloha," Feb. 2016. [Online]. Available: http://arxiv.org/pdf/1602.07636v2.pdf

[92] R. D. Gaudenzi, O. Herrero, G. Acar, and E. Garrido Barrabès, "Asynchronous contention resolution diversity ALOHA: Making CRDSA truly asynchronous," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6193–6206, Nov. 2014.

[93] E. Sandgren, A. Graell i Amat, and F. Brännström, "Asymptotic and finite frame length analysis of frame asynchronous coded slotted ALOHA," 2016. [Online]. Available: http://arxiv.org/pdf/1604.06293.pdf

[94] D. Jakovetić, D. Bajović, D. Vukobratović, and V. Crnojević, "Cooperative slotted aloha for multi-base station systems," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1443–1456, Feb. 2015.

[95] S. Gollakota and D. Katabi, "Zigzag decoding: Combating hidden terminals in wireless networks," *ACM SIGCOMM*, pp. 159–170, 2008.

[96] A. Tehrani, A. Dimakis, and M. Neely, "Sigsag: Iterative detection through soft message-passing," *IEEE J. Sel. Top. Sign. Proces.*, vol. 5, no. 8, pp. 1512–1523, Dec. 2011.

[97] A. ParandehGheibi, J. K. Sundararajan, and M. Medard, "Collision helps - algebraic collision recovery for wireless erasure networks," in *Proc. IEEE Wireless Network Coding Conf.*, Boston, MA, June 2010.

[98] S. Ghez, S. Verdú, and S. Schwartz, "Stability properties of slotted aloha with multipacket reception capability," *IEEE Trans. Automatic Control*, vol. 33, no. 7, pp. 640–649, July 1988.

[99] C. Stefanovic, M. Momoda, and P. Popovski, "Exploiting capture effect in frameless ALOHA for massive wireless random access," in *Proc. IEEE Wireless Commun. and Net. Conf.*, Istanbul, Turkey, Apr. 2014.

[100] A. Munari, M. Heindlmaier, G. Liva, and M. Berioli, "The throughput of slotted ALOHA with diversity," in *Proc. Annual Allerton Conference on Communication, Control, and Computing*, Allerton, IL, Oct. 2013.

[101] M. Ivanov, F. Brännström, A. Graell i Amat, and P. Popovski, "Error floor analysis of coded slotted ALOHA over packet erasure channels," *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 419–422, Mar. 2015.

[102] O. D. R. Herrero and R. D. Gaudenzi, "Generalized analytical framework for the performance assessment of slotted random access protocols," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 809–821, Feb. 2014.

[103] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the gaussian multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 364–375, Mar. 1996.

[104] J. Garcia-Frias and W. Zhong, "Approaching Shannon performance by iterative decoding of linear codes with low-density generator matrix," *IEEE Commun. Lett.*, vol. 7, no. 6, pp. 266–268, June 2003.

[105] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Analysis of low density codes and improved designs using irregular graphs," in *Proc. ACM Symp. on Theory of Computing*, 1998.

[106] M. Luby, "LT codes," in *Proc. IEEE Symp. Foundations Computer Science*, Vancouver, Canada, Nov. 2002.

[107] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[108] M. Luby, M. Mitzenmacher, and A. Shokrollahi, "Analysis of random processes via and-or tree evaluation," in *Proc. ACM-SIAM Symp. Discrete Algorithms*, San Francisco, CA, Jan. 1998.

[109] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[110] C. Di, D. Proietti, I. E. Telatar, T. J. Richardson, and R. L. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1459–1473, June 2002.

[111] T. Tian, C. R. Jones, J. D. Villasenor, and R. D. Wesel, "Selective avoidance of cycles in irregular LDPC code construction," *IEEE Trans. Commun.*, vol. 52, no. 8, pp. 1242–1247, Aug. 2004.

[112] X. Hu, E. Eleftheriou, and D. Arnold, "Regular and irregular progressive edge-growth tanner graphs," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 386–398, Jan. 2005.

[113] C.-C. Wang, S. R. Kulkarni, and H. V. Poor, "Finding all small error-prone substructures in LDPC codes," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 1976–1999, May 2009.

[114] E. Rosnes and O. Ytrehus, "An efficient algorithm to find all small-size stopping sets of low-density parity-check matrices," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4167–4178, Sep. 2009.

[115] A. Orlitsky, K. Viswanathan, and J. Zhang, "Stopping set distribution of LDPC code ensembles," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 929–953, Mar. 2005.

[116] E. Paolini, "Finite length analysis of irregular repetition slotted aloha (IRSA) access protocols," in *Proc. IEEE Int. Conf. Commun. Workshop*, London, UK, June 2015.

[117] A. Alvarado, L. Szczecinski, R. Feick, and L. Ahumada, "Distribution of L-values in Gray-mapped M^2-QAM: Closed-form approximations and applications," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 2071–2079, July 2009.

[118] A. Amraoui, A. Montanari, T. Richardson, and R. Urbanke, "Finite-length scaling for iteratively decoded LDPC ensembles," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 473–498, Feb. 2009.