Low Density Parity Check Code Designs For Distributed Joint Source-Channel Coding Over Multiple Access Channels

by

Iqbal Shahid

A Thesis submitted to The Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering

The University of Manitoba

Winnipeg, Canada

August 2013

Dedicated to my family - my parents, my wife Amna and my sons, Jafar, Musab and Saad.

Abstract

The efficient and reliable communication of data from multiple sources to a single receiver plays an important role in emerging applications such as wireless sensor networks. The correlation among observations picked-up by spatially distributed sensors in such a network can be exploited to enhance the efficiency and reliability of communication. In particular, information theory shows that optimal communication of information from correlated sources requires distributed joint source-channel (DJSC) coding.

This dissertation develops new approaches to designing DJSC codes based on low density parity check (LDPC) codes. The existence of low complexity code optimization algorithms and decoding algorithms make these codes ideal for joint optimization and decoding of multiple codes operating on correlated sources. The well known EXIT analysis-based LDPC code optimization method for channel coding in single-user point-to-point systems is extended to the optimization of two-user LDPC codes for DJSC coding in multi-access channels (MACs) with correlated users.

Considering an orthogonal MAC with two correlated binary sources, an asymptotically optimal DJSC code construction capable of achieving any rate-pair in the theoretically-achievable two-user rate-region is presented. A practical approach to realizing this scheme using irregular LDPC codes is then developed. Experimental results are presented which demonstrate that the proposed codes can approach theoretical bounds when the codeword length is increased. For short codeword length and high inter-source correlation, these DJSC codes are shown to significantly outperform separate source and channel codes.

Next, the DJSC code design for the transmission of a pair of correlated binary sources over a Gaussian MAC (GMAC) is investigated. The separate source and channel coding is known to be sub-optimal in this case. For the optimization of a pair of irregular LDPC codes,

the EXIT analysis for message passing in a joint factor-graph decoder is analyzed, and an approach to modeling the probability density functions of messages associated with graph nodes which represent the inter-source dependence is proposed. Simulation results show that, for sufficiently large codeword lengths and high inter-source correlation, the proposed DJSC codes for GMAC can achieve rates higher than the theoretical upper bound for separate source and channel coding.

Acknowledgments

I would like to express deepest gratitude to my advisor Professor Pradeepa Yahampath for his guidance and insights throughout this thesis and my education tenure for the last 5 years. He has always been tough but fair in the criticism of my work. I owe sincere gratitude to him for his ultra-prompt and constructive help throughout my program and for his numerous proofreads. This thesis would not have been possible without his direction.

I would like to thank my family for their love and support which made it possible for me to reach this stage. No words can do justice for the support my parents have always lent me. My deepest thanks to my wife Amna for standing by me during these tough academic years. My sons, Jafar, Musab and Saad for bringing so much happiness into my life.

Finally, I express my gratitude to my Creator for the forgiveness and kindness He showered on me all through my life.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
Bi-AWGN	Binary input AWGN
BSC	Binary Symmetric Channel
BP	Belief Propagation
CMA	Collaborative Multiple Access
DJSCC	Distributed Joint Source Channel Coding
DSC	Distributed Source Coding
EEG	Electroencephalography
EXIT	Extrinsic Information Transfer
FG	Factor Graph
JSCC	Joint Source-Channel Coding
LDGM	Low Density Generator Matrix
LDPC	Low Density Parity Check
LLR	Log Likelihood Ratio
MAC	Multiple Access Channel
MAP	Maximum a posteriori Probability
ML	Maximum Likelihood

NJMA	Non collaborative Joint Multiple Access
NSMA	Non collaborative Separate Multiple Access
pdf	probability density function
pmf	probability mass function
SSCC	Separate Source-Channel Coding
SPA	Sum Product Algorithm
SW	Slepian Wolf
UEP	Unequal Error Protection
WSN	Wireless Sensor Networks

Chapter 1

Introduction

Digital communication has revolutionized voice, data and video communications in the past few decades. A key advantage of digital communication over analog communication is the possibility of using source coding for data compression and channel coding for channel error correction. The use of near optimal source and channel codes allows the realization of digital communication systems whose performance can approach the fundamental limits predicted in information theory [5]. This dissertation is concerned with new source and channel code designs for wireless sensor networks which can be used to acquire and communicate information from correlated sources.

1.1 Point to Point Communication

A fundamental model of a digital communication system is shown in Fig. 1.1 for communication between a single information source and a single destination. As shown, the transmitter consists of a source coding module, a channel coding module, and a modulator. The modulated signal is transmitted through a physical channel. The most common single-user channel models include binary symmetric channel (BSC), the additive white Gaussian noise (AWGN) channel, Rayleigh fading channel [6]. The channel output signals are then

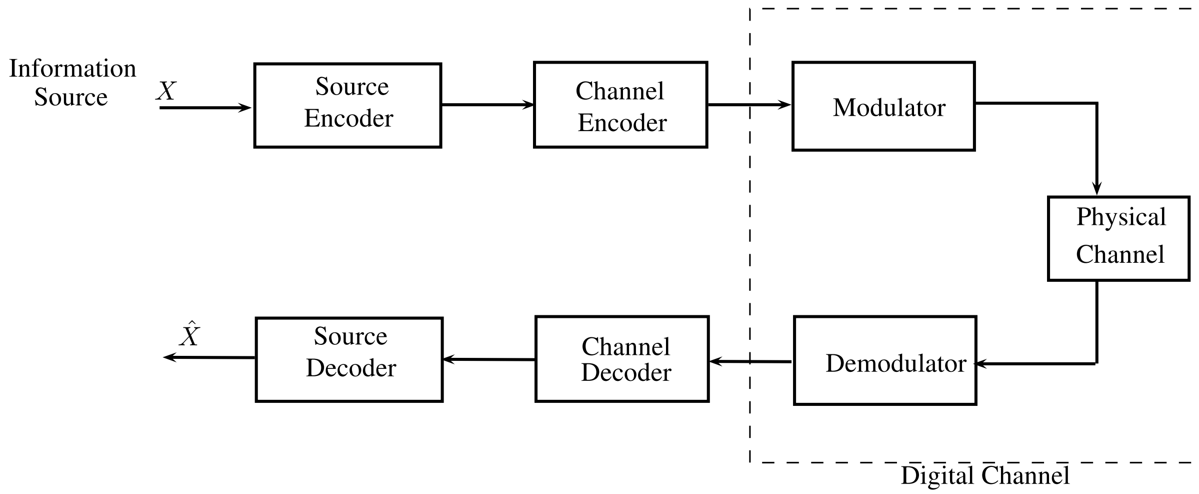

Figure 1.1: A basic model of a digital communication system.

demodulated and decoded to recover the information. We now briefly describe the function of each module.

Source coding (or signal compression) refers to the process of representing an information source in digital form, most commonly as a sequence of binary symbols. The main objective of source coding is to find the most efficient digital representation in terms of utilization of communication resources such as storage, channel bandwidth, *etc*, by elimination any redundancies in the signals generated by the information source. Such a representation is usually referred to as a *code* (and hence the term source coding). When the information source is discrete, the source coding process can be one-to-one in that original information can be exactly recovered from the codewords. The idea of lossless source coding was first presented by Shannon in his landmark paper in 1948 [1]. He associated with each discrete source, a measure referred to as the *entropy*, which is the minimum rate at which a discrete source can be described by a coded representation. When the information source is analog, source coding is preceded by sampling (mapping of an analog signal to discrete-time real values) followed by quantization (representing those sampled real numbers by binary codewords). For analog signals, source coding introduces some distortion and hence termed as *lossy* source coding.

The objective of lossy source coding is to find the most efficient representation for the source which minimizes the average distortion between the original signal and the coded version (in terms of a distortion measure such as the mean squared error). For a given source and a distortion measure, the *rate-distortion function* [7], describes the minimum rate at which the source can be digitally represented at a given level of distortion. For some early work on source coding, see for example [8], [9] and references therein. Recent developments in the field of lossy source coding are summarized in [10]. This dissertation is concerned with the communication of digital sources and hence lossless source coding. The method developed here can however be applied to analog sources after analog-to-digital conversion of the analog signals.

The goal of *channel coding* (or error correction coding) is to introduce a controlled amount of structured redundancy into a sequence of binary digits (bits) transmitted over a noisy channel, so that any errors that occur in the channel can be detected and corrected at the receiver. The pioneering work on channel coding was also done by Shannon in [1], where the theoretical limit of the amount of the minimum amount of redundancy required to achieve near error-free communication over a channel was defined in terms of a quantity called the *channel capacity*. The channel capacity is the highest rate at which the reliable communication of information over the channel is possible. The higher the capacity of a channel, the lesser the redundancy that has to be added to the transmitted bit-stream and vice versa. One of the pioneering practical approaches to channel coding was developed by Hamming [11] in 1950. A recent review of progress in channel coding can be found in [12] and [13].

In general, it is not necessary to have two separate codes for source coding and channel coding. The separation of source coding and channel coding is merely a decomposition used for the purposes of easy analysis and code design. In theory, one could use a single code to perform both source and channel coding. A single code which maps a sequence of symbols or samples from an information source to a sequence of channel input codewords, is referred to as a joint source-channel (JSC) code. However, an important question is whether the concatenation of an optimal source code and an optimal channel code can also be an optimal JSC code. In [1], Shannon proved that for point-to-point communication (as shown in Fig 1.1) of a discrete memoryless source by lossless coding, the individually optimal source and channel codes are also jointly optimal if the channel is discrete and memoryless and if the encoding and decoding can be performed on infinitely long sequences of symbols. That is to say that, separate source and channel codes can be *asymptotically optimal* for point-to-point communication. Under these conditions, reliable communication is possible at the lowest possible bit rate only if the entropy of the source is less than the capacity of the channel. This concept is elegant and attractive since it allows us to design the optimal source coding module independent of the optimal channel coding module. It also gives us the freedom to use the same channel code for many different applications that uses the same communication channel (only the source coding module needs to be changed). However, it is easy to see that in practice separately designed source and channel codes cannot be optimal as it is impractical to design codes which operate on infinitely long bit sequences. Therefore, in practical communication systems where one is essentially restricted to finite length codes, the use of JSC codes may lead to significant performance improvements. This problem is referred to as JSC coding (JSCC). JSCC has been studied extensively in the literature and various practical techniques have also been proposed, see [14–17]. However, much of this work has been focused on point-to-point communication.

1.2 Multiuser Communication

We now focus on an important class of systems in which multiple sources, commonly referred to as users, communicate with multiple destinations (or sinks), possibly over a communication network. The motivation for the research carried out in this dissertation comes from a particular multi-user communication system referred to as a wireless sensor network (WSN) [18, 19]. In a WSN many spatially distributed sensor nodes transmit their respective observations to a central processor (*i.e.* a single destination). Due to the fact that sensors observe a physical process at different points in space, the observations picked up by different sensors are highly correlated. In such a situation the communication efficiency can be increased if the sensors collaborate during the coding process. However, the sensors in a WSN are typically required to be simple devices with very limited battery power and hence cannot have the complex communication functionalities required to communicate with each other. In such a situation an alternative approach to efficient data compression is *distributed source coding* (DSC).

Distributed source coding is a source coding method by which two or more physically separated but statistically correlated information sources are separately encoded (or compressed) and jointly decoded (or decompressed) at a central location. The pioneering theoretical work on DSC was done by Slepian and Wolf [20] in which they considered the communication of multiple correlated discrete sources to a single decoder. For correlated analog sources, a similar work was done by Wyner in [21, 22]. A connection between channel coding and DSC was identified by Wyner in [23] where he showed the possibility of using linear channel codes to practically implement the proposition made by Slepian and Wolf. A detailed review of recent work, including numerous applications of DSC, can be found in [24]. With the advancements in channel coding, many practical techniques for DSC were developed later, and it was found that good channel codes can also be used to design good distributed source codes. DSC is thus considered to be one of the enabling technologies in WSN's. A review of recent work on DSC for WSN's can be found in [24, 25].

Just like the case of single user communication, the channel coding problem in the case of multiple sources and/or multiple sinks has been a topic of considerable interest in information theory literature [23, 26, 27]. The problem of communicating of many sources to a single

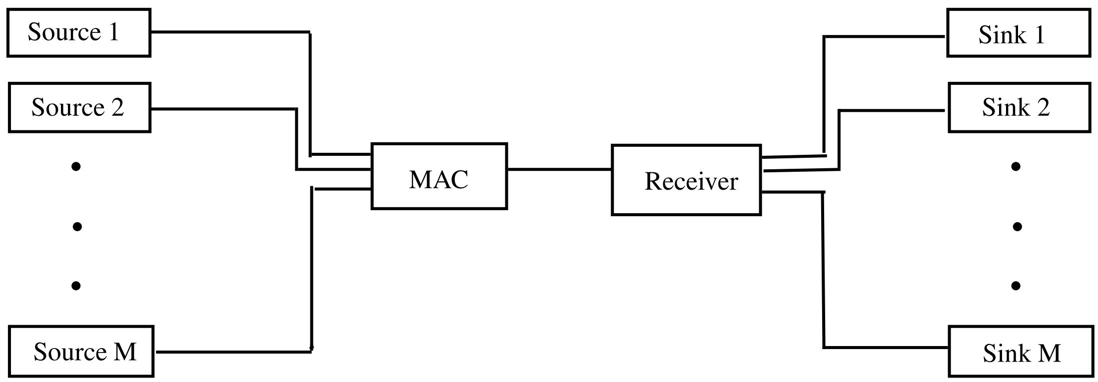

Figure 1.2: A multi-user communication system which uses a MAC.

destination leads us to the study of the *multiple access channel* (MAC) [5], in which many sources share a common channel to transmit their information to a single destination. A simple multi-user communication system model which uses a MAC is shown in Fig.1.2. Examples for MAC include satellite communication with many independent ground stations, a cellular network where many cell phones communicate with a base station, and WSNs. It turns out that the separation of source and channel coding in general does not lead to optimal codes in a MAC scenario even if the codeword lengths are allowed to be infinite [5]. In other words, using an optimal distributed source code followed by an optimal channel code with a MAC is in general, suboptimal (there are some special cases where separation does hold, see for example [28], [29]). A very simple toy example is given in [30] to emphasize this point.

Based on the above observations, it is natural to seek distributed joint source channel (DJSC) coding schemes to achieve optimal performance in a MAC scenario. The focus of this dissertation mainly is to investigate the practically important cases in which DJSC coding for a MAC can strictly yield better performance compared to separately designed source and channel codes, and then to develop practical approaches to design DJSC codes which can approach these performance improvements. The improvement may be in terms of higher bit rates at a given channel noise level, higher level of reliability at a given rate, or a reduction in the computational complexity of the encoding and decoding algorithms at a given rate and channel noise level as compared to a commonly used separation-based designs. One of the challenges in this problem is to design a single JSC code which can simultaneously exploit the correlation among multiple sources for both data compression and channel error correction. While various theoretical aspects of this problem have been studied in previous work (see e.g. [28–31]), practical design of DJSC coding schemes is an area where much work remains to be done. Some previous work on channel code-based JSC code design for single user communications can be found in [17] and channel code design for multiple user communications has been considered in [4]. JSC code designs can be very promising especially in WSN's where the data sensed by different sensor nodes is highly correlated and much can be achieved if this correlation is exploited by a DJSC code for both data compression and channel error correction. Combining ideas from DSC and JSC coding is pivotal in developing efficient DJSC coding schemes.

When developing DJSC coding schemes for a multi-user system with a large number of sources (such as a WSN), an important issue to be considered is the computational complexity of the encoding and decoding algorithms. Linear graph based codes [32] with iterative decoding algorithms describe a class of codes which promises low encoding and decoding complexity while maintaining an excellent level of performance. They have been efficiently used for both source coding and channel coding in single user communication systems [17,33–35]. Linear graph based codes accompanied with iterative decoding algorithms therefore are natural and promising candidates for the implementation of DJSC coding system over a MAC.

1.3 Contributions and Outline of this Dissertation

Chapter 2 presents the background on the problem of source and channel separation in coding and lays the foundation for subsequent new developments carried out in this dissertation. The source-channel separation considered in information-theoretic terms for single user and multiuser communication. Special emphasis is given to the multiuser case where the separation does not hold in general.

In Chapter 3, DJSC coding problem is discussed in detail. Specifically, the focus is place on practical channel coding approaches which have been used to implement DJSC coding schemes developed in this dissertation. The relationship between DJSC coding and conventional channel coding is discussed and it is shown how good channel codes can be used to implement efficient DJSC coding schemes as well.

In Chapter 4, a new approach to DJSC coding of two binary sources over independent (orthogonal) channels is proposed, and algorithms for code design and decoding are developed [36–38]. It is shown that the proposed codes construction is asymptotically optimal in that any rate-pair in the achievable rate region for this problem can be realized, if sufficiently long codes are used. The proposed code construction implicitly rely on time-sharing of the two codes to achieve any given set of rates in the achievable rate region. The practical approach to implementing the proposed code construction based on systematic irregular LDPC codes is then developed. This approach involves the design of a pair of LDPC codes with unequal error protection (UEP) properties. The two codes are jointly decoded by the belief propagation algorithm applied to their joint factor-graph. We rely on low-complexity EXIT analysis commonly used for single user LDPC code to jointly optimize two LDPC codes. However, it is shown that the Gaussian approximation of the density functions of the messages commonly used in EXIT analysis of single-user LDPC code design does not necessarily apply to the joint design of two LDPC codes for correlated sources. It is however established the messages associated with graph nodes representing the dependence between the two sources can be well modeled by Gaussian mixture models. An extensive set of design examples and simulation results are presented to demonstrate the performance advantages of the proposed DJSC codes over separate source and channel coding, as well as other DJSC code designs reported in the literature. Although, theoretically the separation of source and channel coding is optimal for communicating correlated sources over independent channels, the experimental results presented here show that for short to moderate codeword lengths (on the order of few hundreds to few thousand bits), the proposed DJSC codes outperform a setup comprising of a DSC code for both sources followed by independent channel coding of the channel inputs.

In Chapter 5, a new approach to code design for the transmission of two correlated binary sources over a Gaussian MAC (GMAC) is developed. This problem is inherently more difficult to solve as compared to that considered in chapter 4 as the transmissions from the two sources interfere with each other in the channel. As already discussed, in this particular case, the source-channel separation is known to be even theoretically sub-optimal. The approach developed here involves the optimization of two irregular LDPC codes with identical rates, which are decoded using a joint factor-graph that accounts for inter-source dependence. In contrast to codes considered in Chapter 4, the DJSC codes developed in this chapter uses the a priori known inter-source correlation entirely to correct the errors due to noise and mutual interference in the channel (no source coding is therefore achieved). However, in this case, the same performance that can be achieved with independent sources (or when inter-source correlation is ignored as done in conventional source-channel separation-based coding) can be achieved at a lower transmission rate (or improved performance at the same rate). While we use low-complexity EXIT analysis as in Chapter 4 for joint code optimization, it is shown that the messages associated with the nodes in the decoder factor graph that represent the inter-source dependence have a highly skewed probability density function, making the expressions used with EXIT analysis of single-user LDPC codes inapplicable to this case. In order to address this problem, an alternative approach is introduced in which messages associated with nodes representing the inter-source dependence are fitted with Gaussian densities by Monte-Carlo based estimation of the modes of the empirical distributions. The experimental results are presented which demonstrate that the proposed DJSC code designs can outperform conventional separate source and channel codes when communication correlated sources over a GMAC.

Finally, in Chapter 6, conclusions are summarized and possible future directions of research are identified.

Chapter 2

Source-Channel Separation: A Theoretical Perspective

In this chapter, the theoretical optimality of source and channel coding in both single user and multi-user communication is investigated. The theoretical bounds for the achievable rates in source coding and in channel coding are considered and their significance is discussed. These theoretical bounds provide a benchmark for the performance of practical DJSC coding methods which are to be introduced in subsequent chapters.

2.1 Single User Communication Systems

Consider communicating a single source represented by a random variable X to a single destination over a noisy channel.

2.1.1 Lossless Source Coding Theorem [1]:

Theorem 1. *The rate R (in bits per symbol) required to represent the discrete random variable $X \in \mathcal{X}$ is given by*

$$R \geq H(X), \tag{2.1}$$

where, the entropy of the discrete random variable X is defined as

$$H(X) = \sum_x p(x) \log_2 \frac{1}{p(x)}, \tag{2.2}$$

$p(x) = Pr[X = x]$ *for* $x \in \{1, \ldots, |\mathcal{X}|\}$.

The entropy is measured in bits (if natural logarithm is used in the definition of H, the units are nats). The above relation gives a universal lower bound for lossless compression of a discrete source.

Unlike a discrete source, the digital representation of a continuous source is always lossy, since the exact representation of an arbitrary real number requires infinite bits. Therefore, the minimum achievable rate R in compression of a continuous source is given by its *rate-distortion function* [1]. In this case, the minimum rate R required to represent the source with an average distortion D is given by the rate-distortion function $R(D)$. Note that lossless compression is a special case of lossy compression in which $D = 0$. In this dissertation, sources considered are discrete unless explicitly mentioned.

2.1.2 Channel Coding Theorem [1]:

Coming back to the channel coding problem, consider a communication channel whose input and output are given by random variables X and Y respectively. Suppose a sequence of k source symbols are mapped to a sequence of n channel symbols. We say that *asymptotically error-free* or *reliable* communication is possible at rate $R = k/n$ if the probability of decoder error can be made to approach zero in the limit $k, n \to \infty$.

Theorem 2. *The transmission rate R (in information bits per transmission) at which asymptotically error-free communication is possible is upper bounded by [1]*

$$R \leq C, \tag{2.3}$$

where C is the channel capacity given by $\max_{p(x)} I(X;Y)$ *and*

$$I(X;Y) = \sum_{x}\sum_{y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)} \tag{2.4}$$

is the mutual information (in bits) between X and Y.

2.1.3 Source-Channel Separation principle [1]:

Speaking in information theoretic terms, the meeting point for the branches of source coding and channel coding is the *joint source channel coding theorem* [5].

Theorem 3. *Asymptotically error-free communication of a source with entropy H over a channel with capacity C is possible if $H < C$. Conversely if $H > C$, asymptotically error-free communication is not possible.*

An immediate consequence of this theorem is that if $H < C$ then reliable communication can be achieved by separate source coding and channel coding. Therefore, the above theorem is also termed as the *separation theorem* or *separation principle*. This optimality holds if both source and channel are ergodic and memoryless and infinite length codewords are allowed. Another way of putting this is to say that $H < C$ is a necessary *and* sufficient condition for the separation principle to hold. If however, $H < C$ is a sufficient condition but not a necessary condition then in which case the separation principle does not hold.

An informal proof of the joint source channel coding theorem is as follows. Assume that we want to transmit n symbols per transmission from a source with entropy H through a channel with capacity C. The amount of information we have is nH symbols, and hence by using ideal source coding, we can compress the source to nH symbols. Since the capacity of the channel is C, by using an ideal channel coding we can transmit nH bits using nC transmissions. It now follows that a rate of n source symbols per transmission is possible only if $H \leq C$.

In the above example, the source coder is assumed to achieve the bound (2.1) and the channel coder is assumed to achieve the bound (2.3). However this ideal situation cannot be realized in practical scenarios since theoretically source coding and channel coding bounds can be achieved only in the limit of infinite codeword lengths (or at least, for very large codeword lengths) [5]. What if we need to communicate a symbol sequence of finite length (which cannot even be considered as *very large*)? This issue along with issues like complexity (how costly is the encoding/decoding process if the sequence length is very long) and the encoding/decoding delay (how much waiting time is allowed before the actual encoding/decoding process starts) are important due to the ever-increasing demand for high data rates and low latency constraints associated with real-time applications such as video streaming. So under these conditions, a *joint* source-channel code (*i.e.* a code which directly maps the source symbols to channel input symbols) is likely to outperform a combination of separate source and channel codes.

Before we proceed, it is worth noting that the separation theorem has also been considered without the conditions of stationarity and ergodicity, see [39].

2.1.4 Joint Source and Channel Coding

Shannon hinted in his work [1] about exploiting source redundancy for channel error correction. A simple example given in [1] is the English language which has a redundancy of approximately 50%. As a results, a typical telegraph message can be compressed to about half its size by using lossless source coding. This is in practice not done since redundancy helps to combat any noise (in the form of erasure or scrambling of letters) at the receiver. This redundancy allows the correction of any errors as words can be understood from the context even when some letters are missing. Joint source-channel (JSC) coding follows this idea. In ideal JSC coding, a single code which simultaneously achieves both data compression and channel coding is constructed by taking into account both source statistics *and* channel statistics. That is, source messages are directly mapped to codewords which can be transmitted over a noisy channel. On the other hand, a more practical approach is to use source redundancy to facilitate the channel error correction at the decoder. This approach is a subset of JSC coding and is usually referred to as joint source-channel decoding [14]. Several other work exist where the source redundancy (partial or full) is exploited at the channel decoder to improve the error resilience, see for example, [40]. From the source coding perspective, JSC coding is treated as a combination of arithmetic decoding and MAP decoding in [41]. Several universal algorithms (i.e. algorithms which are good for a class of sources) for joint source channel decoding are presented in [42]. Analysis and optimization of JSC codes based on sparse graphs appear in [43] (irregular LDPC codes) and [16] (Raptor codes). Similar practical JSCC schemes using sparse graph-based codes can be found in [44] and references therein. Some of the applications in which ideas of JSC coding has been considered are as follows:

1. In GSM second generation cellular wireless system, the data compressor leaves some residual redundancy prior to channel encoding, which can be exploited by channel encoder and/or decoder [42].

2. The decoding time of the Fountain codes (a class of codes known to perform very well for erasure channels) depends upon the erasure rate. By exploiting the redundancy in the source, decoding time can be improved. This has applications in data communication over the Internet where erasures due to network packet losses are usual [42].

3. The use of maximum a posteriori probability (MAP) decoder in place of maximum likelihood (ML) is more appropriate if the source is not uniformly distributed, i.e., the source has some redundancy. This is typical in decoding based on message passing algorithms [35].

To complete the picture, we also note that there is an overhead involved in making the source statistics available to the decoder. Also, the decoder complexity usually increases when correlation or memory in the source are exploited in JSCC [42].

2.2 Multiple User Communication Systems

In this section, we consider the problem of reliably communicating two or more sources (possibly dependent) to a single destination.

Distributed source coding

Given the recent interest in WSNs, an important extension of the source coding problem is compressing multiple correlated sources in a distributed fashion [45, 46]. In this case, each source has to be compressed by a separate encoder. Ideally, the redundancy within each source as well as the correlation between the sources should be exploited for efficient compression. It appears that for optimal performance, all sources must be encoded jointly. The joint encoding of all the sources not only requires encoders to collaborate employing some form of inter-source communication, but is in general computationally very complex and impractical when the number of sources is large. The joint coding is also impractical when the sources are not co-located. Parallel to the aforementioned single user lossless source coding theorem by Shannon [1], perhaps the most basic result for discrete correlated sources was proven by Slepian and Wolf in their landmark paper [47] in 1973. This theorem shows that joint encoding is not required for optimal communication, provided all the sources can be decoded by a joint decoder.

Theorem 4 (Slepian and Wolf [47])**.** *The achievable rate region for the distributed source*

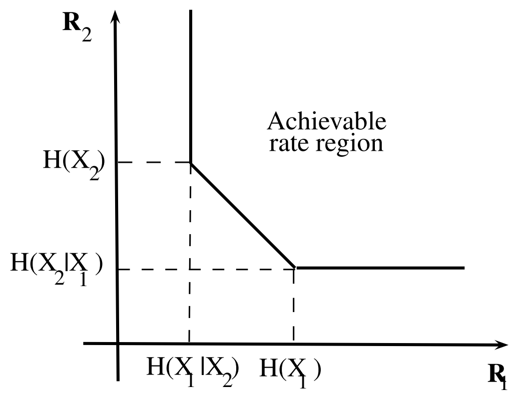

Figure 2.1: Achievable rate region for distributed source coding of two sources X_1 and X_2.

coding of two sources (X_1, X_2) *drawn i.i.d from* $p(x_1, x_2)$ *is given by*

$$\begin{aligned} R_1 &\geq H(X_1 \mid X_2), \\ R_2 &\geq H(X_2 \mid X_1), \\ R_1 + R_2 &\geq H(X_1, X_2), \end{aligned} \tag{2.5}$$

where $H(X_i|X_j) = \sum_{x_j} \sum_{x_i} p(x_i, x_j) \log_2 \frac{1}{p(x_j|x_i)}$ *for* $i, j = 1, 2$ *is the conditional entropy of the random variable* X_i *given the information of random variable* X_j*.*

This is the 2-user equivalent of the Shannon's source coding theorem given by (2.1). The achievable rate-region given by Slepian-Wolf theorem is graphically shown in Fig. 2.1. The above theorem can be easily extended to more than two sources. The dimensionality of the achievable region is then the number of sources involved. The Slepian-Wolf theorem defines a lower bound to the source coding rates for multiple sources. Here it is important to note that if the sources are compressed up to the given limits, the outputs of the source encoders are *independent*. Also interesting is the fact that if the source outputs (or the encoder inputs) are independent, then $H(X_1|X_2) = H(X_1)$ and $H(X_2|X_1) = H(X_2)$ and we get the result in (2.1).

Channel coding

The most basic result on the achievable rates in channel coding for multi-users channels (e.g., MAC) has been given by Cover [5]. Before the theorem is stated, it is appropriate to give a formal definition for a MAC. The following definition is due to Cover [5]:

Definition 1 (Multiple access channel)**.** An m-user multiple access channel is completely characterized by finite or countably infinite input alphabets $\mathscr{X}_1$, $\mathscr{X}_2$, $\cdots$, $\mathscr{X}_m$, output alphabet $\mathscr{Y}$ and a probability transition matrix $p(y|x_1, x_2, ..., x_m)$.

The following theorem gives the achievable channel coding rates for two users [5]:

Theorem 5 (Capacity for a 2-user MAC)**.** *The capacity of a MAC* $(\mathscr{X}_1 \times \mathscr{X}_2, p(y|x_1, x_2), \mathscr{Y})$ *is the closure of the convex hull of all* (R_1, R_2) *satisfying*

$$\begin{aligned} R_1 &\leq I(X_1; Y \mid X_2), \\ R_2 &\leq I(X_2; Y \mid X_1), \\ R_1 + R_2 &\leq I(X_1, X_2; Y) \end{aligned} \tag{2.6}$$

for some product distribution $p(x_1)p(x_2)$ *on* $\mathscr{X}_1 \times \mathscr{X}_2$*.*

The achievable rate region is shown in Fig. 2.2 for the special case of $Y = X_1 + X_2$.

It is important to note that the channel capacity given by the above theorem is defined with respect to the *product* distribution $p(x_1)p(x_2)$ as opposed to the joint distribution $p(x_1, x_2)$. In other words, this gives the MAC capacity that can achieved if the sources are independent.

Source-channel separation

We next attempt to answer the following question: *does the separation principle hold for MAC as it does in the single user case ?* In other words, when communicating correlated

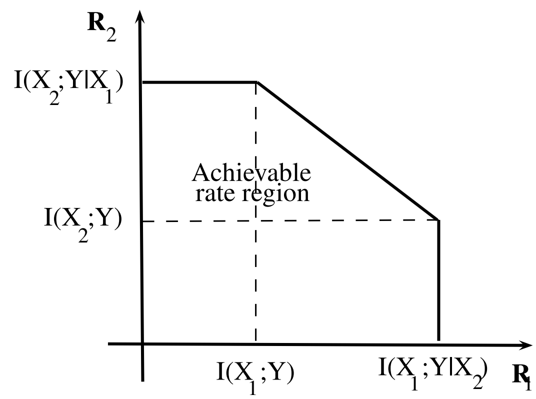

Figure 2.2: MAC Capacity region for two sources X_1 and X_2.

sources over a MAC, can we first apply DSC to the sources and then apply independent channel coding to the output of each source coder ? A step towards the answer is to pose this question differently as follows: *referring back to Theorem 5, is it possible to increase the upper bounds given on the right hand side of the inequalities by using a joint distribution* ? (a negative answer would mean that DSC modules can be optimized independently from the MAC channel coding modules). First, note that the set of all possible distributions $p(x_1, x_2)$ is a superset of the set of all possible product distributions $p(x_1)p(x_2)$. Secondly, the rate bound of X_1 depends upon the X_2, since we have $I(X_1; Y|X_2)$ on the right hand side of the first inequality and vice versa. Now it is easy to see that the maximum rate achieved by a joint distribution may be greater than that achieved by a product distribution. In the language of information theory, we say that the separation principle does not hold in the case of MACs *even* if coding of infinite length sequences are allowed.

The above described situation is highlighted in Fig. 2.3. Fig. 2.3 (a) shows the case of separate source and channel coding which we refer to as *non-collaborative separate multiple access (NSMA)* [29], while Fig. 2.3 (b) shows the case of joint source-channel coding which we refer to as *non-collaborative joint multiple access* (NJMA). The term 'non-collaborative' is used because the encoders of the two sources do not communicate with each other. So on the basis of the above analysis, it can be argued that the system shown in Fig. 2.3 (b) *can*

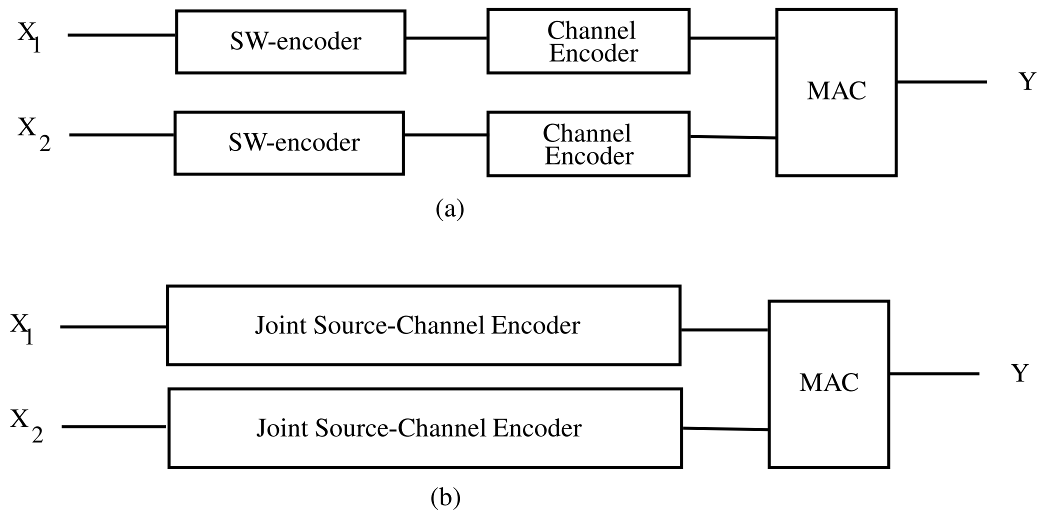

Figure 2.3: (a) NSMA system (b) NJMA system

achieve lower rates than that shown in Fig. 2.3 (a).

2.2.1 Some Specific Cases Where Separation Holds for MAC

Despite the fact that a separation based system may be outperformed by a system with joint source-channel coding, it is noteworthy that the separation principle is very attractive both theoretically and for practical purposes in that it allows easier design of both source coding and channel coding modules and also allows the use of the same channel code for different types of sources. Therefore it is important to identify the conditions under which the separation principle holds for a MAC. Stated in another way, the source-channel separation holds for a MAC when the rates achieved by an NSMA system are equal to those achieved by an NJMA system. This also serves as a sufficient condition for the separation principle to hold [29]. Ray *et al.* [29] considered the communication of two discrete sources over a MAC with additive noise and a MAC with erasures, and showed that if the inputs and the output belong to the same finite set and the erasures/additive noise are independent of the inputs, then the separation principle holds. Gündüz *et al.* [28] showed that the separation principle holds if the receiver has access to side-information about the sources and the two sources are independent conditioned on the receiver side-information. For continuous Gaussian sources, Xiao *et al.* [48] showed that if the MAC is orthogonal, *i.e.,* if the communication links are non-interfering, then separate source and channel coding can outperform uncoded transmission. Here it should be emphasized that an uncoded system is a special form of joint source channel coding since the source outputs are directly transmitted over the channel.

2.2.2 Some Specific Cases Where Separation Does Not Hold for MAC

It is shown in [29] that for discrete sources, if the output and the inputs belong to different alphabets or if they belong to the same alphabet but the additive noise or erasures depend upon the inputs, then separation principle does not hold. In a similar manner, for continuous Gaussian sources, Gastpar [45] showed that if the communication links in the MAC are non-orthogonal (i.e., there is interference), then an uncoded transmission can outperform separate source and channel coding.

2.3 Practical Coding Schemes for Multiple Access

Coding for a MAC is challenging because of the presence of both noise *and* interference in the channel, and due to the requirement of decoding all the sources using only a single channel output. Different authors have considered various forms of MACs to devise schemes for reliable communication. Palanki [49] in his PhD thesis considered the case of a binary adder MAC (defined by $Y = X_1 + X_2$ and $\mathscr{X}_1 = \mathscr{X}_2 = \{0, 1\}, \mathscr{Y} = \{0, 1, 2\}$) and devised a method known as *graph splitting* to split a single low density parity check (LDPC) code into two constituent codes for encoding each user's bit sequence. However, the method works well only when the channel is noise-free. For a two-user noiseless binary adder channel, the authors in [50] investigated convolutional encoding of each user and the ML decoding of both users on a joint trellis.

A scheme referred to as collaborative coding multiple access (CCMA) is discussed in [51], where a real adder MAC with Gaussian noise and an arbitrary number of users is considered. The design of uniquely decodable codes is discussed, and hard-decision ML decoding is used at the receiver. The additive Gaussian MAC in which the bit stream coming from each source is modulated using a suitable modulation scheme such as phase shift keying (PSK) is discussed in [52] where the authors use iterative decoding methods for interference cancellation.

It is noteworthy to mention that aforementioned work on MACs considers only the channel coding problem and assumes independent sources. As was discussed before, exploiting the correlation between the sources in coding can lead to improved performance with MACs. In Chapters 4 and 5 of this dissertation, two new approaches to practical design of DJSC codes are developed, which effectively exploit the correlation among multiple sources for reliable and efficient communication over two-user MACs.

Chapter 3

Distributed Joint Source-Channel Coding

In DJSC coding sequences of symbols from two or more correlated sources are mapped to the inputs of a MAC such that all sources can be reliably decoded at the channel output. The encoders of a DJSC code do not communicate with each other. However, the source-channel mappings of the encoders takes into account the inter-source dependence and the redundancy within each source to generate optimal channel input codewords. Before embarking upon the topic further and discussing the existing work, it is beneficial to present in some detail the practical schemes employed for distributed source coding (DSC). since DSC coding can be considered as a special case of DJSC coding. It will be shown that certain types of DSC code constructions naturally extends to DJSC code constructions. Based on these ideas, two new approaches to the design of DJSC codes are then developed in Chapters 4 and 5.

3.1 Distributed Source Coding

3.1.1 Asymmetric DSC

The proof of the Slepian-Wolf (SW) theorem is non-constructive, in that it does not provide a practical code construction method. The main idea behind the proof is graph coloring or binning [5]. Consider encoding a sequence of symbols of the source X_2 given that X_1 is transmitted error-free, so that it is available at the decoder as side-information for decoding X_2. This referred to as completely *asymmetric* DSC. A distributed source code for X_2 is now constructed as follows. All X_2-sequences which are jointly typical [5] with a given X_1-sequence are put into one bin, where there is a different bin for a each typical X_1-sequence. Since X_1 is assumed to be transmitted in a lossless manner to the decoder (at the rate of $H(X_1)$), the randomness remaining in X_2 will be $H(X_1, X_2) - H(X_1) = H(X_2|X_1)$, which is the minimum rate at which X_2 can be transmitted. However, code construction by binning of typical sequences is not practical as the concept of typicality applies only to infinitely long sequences. In order to practically realize SW codes, two main approaches have been proposed in the literature. Both these approaches rely on techniques from channel coding (a relationship first noticed in 1974 by Wyner [23]). The essence of this dependence on channel coding is that, if two sources are dependent, then one source can be considered as the output of a virtual communication channel whose input is the other source. Both approaches will now be briefly explained. The two approaches are illustrated in Fig. 3.1.

Syndrome Approach

An (n, k) binary channel code maps k information bits to n channel code bits. An (n, k) binary linear channel code $\mathscr{C}$ is completely specified by its $k \times n$ generator matrix $\mathbf{G}$ or by its $(n-k) \times n$ parity check matrix $\mathbf{H}$, *i.e.*, all codewords in $\mathscr{C}$ are known if either $\mathbf{G}$ or $\mathbf{H}$ is known [53]. If $\mathbf{v}$ is an arbitrary n-bit sequence, then the $n-k$ bit sequence $\mathbf{s} = \mathbf{H}\mathbf{v}$ is

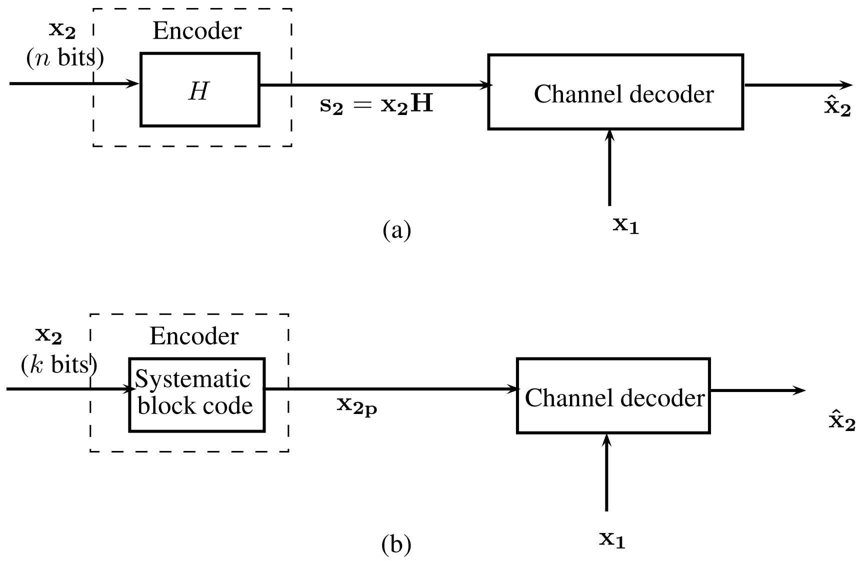

Figure 3.1: Practical schemes for DSC: (a) Syndrome Approach and (b) Parity Approach. $\mathbf{H}$ is the parity check matrix, while $\mathbf{x_{2p}}$ represents the $n-k$ parity bits generated by the systematic channel code with a k bit input $\mathbf{x_2}$. $\hat{\mathbf{x}}_\mathbf{2}$ is the decoded sequence for source X_2.

known as the *syndrome* of $\mathbf{v}$ (so named since it enables us to find the error pattern present in the received sequence). Corresponding to each syndrome, is a set of unique 2^k, n-bit sequences referred to as a *coset* of the code $\mathscr{C}$. Note that the cosets of an (n, k) channel code $\mathscr{C}$ constitute a binning scheme for n-bit sequences and can be used as a distributed source code. Consider encoding an n-bit sequence $\mathbf{x_2}$ (representing source X_2) using a channel code $\mathscr{C}$, given that the n-bit sequence $\mathbf{x_1}$ (representing source X_1) is available as side-information at the decoder. The syndrome approach for DSC [54–56] then compresses $\mathbf{x_2}$ by transmitting the $(n-k)$-bit syndrome $\mathbf{s_2}$ of the coset to which the sequence belongs to. The channel output is decoded by simply choosing that bit sequence in the coset of $\mathbf{s_2}$ which is closest (or has the lowest Hamming distance d_H) to side-information $\mathbf{x_1}$. Clearly, the compression ratio for source X_2 achieved through this method is $\frac{n-k}{n}$. It can be shown that [23], if the channel code $\mathcal{C}$ achieves the capacity of the virtual channel between X_1 and X_2, then this method achieves the minimum transmission rates given by the SW theorem.

Parity Approach

Consider encoding X_2 , given that X_1 is available as side-information at the decoder. Given a systematic channel code $\mathscr{C}$ for the virtual channel between X_1 and X_2, the parity-bit approach transmits for X_2 only the parity bits (represented by $\mathbf{x_{2p}}$) obtained by encoding a k bit sequence $\mathbf{x_2}$ using $\mathscr{C}$. The decoder then determines $\mathbf{x_2}$ by decoding the k bit sequence $\mathbf{x_1}$ using the received parity bits. Since encoding with a systematic (n, k) channel code generates $n-k$ parity bits, the compression ratio in this case is $\frac{n-k}{k}$.

The parity approach is more robust against channel noise compared to the syndrome approach. With the syndrome approach a wrong syndrome can be received due to channel noise in which case the decoding error can be arbitrarily large, since the wrong bin is then used for decoding of the side-information. In contrast, with the parity approach, each bit has some functional relationship with other bits (through the parity check equations) which can be still utilized in the decoder even if some parity bits are received incorrectly. For further discussion on these two approaches, the reader is referred to [2, 57–59].

3.1.2 Non-Asymmetric DSC

The syndrome and parity approaches explained thus far consider the so called *asymmetric* DSC they correspond to the corner points with co-ordinates $(H(X_2), H(X_1|X_2))$ and $(H(X_1), H(X_2|X_1))$ in Fig. 2.1. Note however that the methods described in Fig. 3.1 can be used to achieve any point on the boundary of the SW region by time-sharing of two codes corresponding to the corner points. For example, the middle point on the boundary of the SW rate-region (equal rates for both sources), can be achieved by sending X_1 at the rate $H(X_1)$ for half of the time and X_2 at rate $H(X_2)$ for the remaining half. Achieving any point on the boundary of the SW rate-region other than the corner points is termed as *non-asymmetric* DSC while achieving the middle point on the boundary is called symmetric DSC. Non-asymmetric DSC using the parity approach is discussed in [60] where it is shown that any point between the corner points can be achieved by transmitting some information bits along with a fraction of parity bits. For DSC using the syndrome approach, the reader is referred to [31, 61].

3.2 Practical Schemes for DJSC Coding

In general, much less work on DJSC coding has been reported in the literature compared to DSC, and particularly SW coding. The previous work related to DJSC coding problem can be broadly divided into two categories, (i) those which use inter-source dependence to achieve both compression and channel error control [2, 3, 33, 62–64], and (ii) those which use the inter-source-dependence only for channel error control [65–67]. Almost all work in the first category are based on the *parity-bit* approach to SW coding. In the case of *asymmetric* SW coding (*i.e.,* the compression of one source treating another correlated source as decoder side-information), an encoder based on the parity-bit method transmits only the parity bits obtained by encoding the source using a systematic channel code for the *virtual* correlation channel between the source and the decoder side-information. Compared to the syndrome-approach, the parity-bit approach can be readily extended to account for channel noise, by increasing the number of transmitted parity bits beyond the minimum number required to achieve the SW limit. In [2] parity bits of a turbo code are used to realize a DJSC code for transmitting a binary source over a binary symmetric channel (BSC) with decoder side information. In [33], the design of DJSC codes for the asymmetric problem using a systematic IRA code is investigated. In this case, noise in the channel between the encoder and the decoder is accounted for by designing an IRA code with different channel conditions for systematic bits and parity bits. Although, the optimized IRA codes in [33] appear to outperform turbo-codes in [2], they still require long codewords (*e.g.,* 10^5 bits) to achieve good performance. In a closely related work [64], a DJSC code design for robust video transmission over a packet-erasure channel is presented. A previous work on DJSC coding for

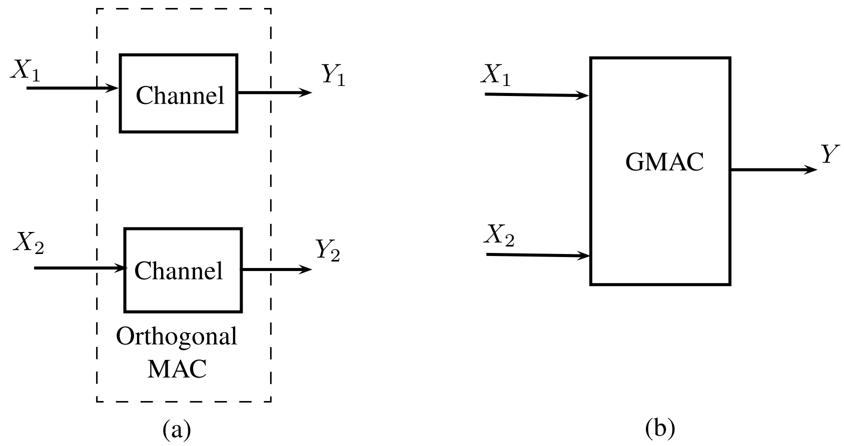

Figure 3.2: Basic input/output relationship for (a) Orthogonal MAC, and (b) GMAC.

the *non-asymmetric* problem (*i.e.,* both sources are compressed) is presented in [62,63], where a turbo code obtained by serial concatenation of two convolutional codes is used for each source to achieve both compression and channel error correction over additive white Gaussian noise (AWGN) channels. An iterative message passing decoder is then used to decode both sources simultaneously. Other related work, including an approach to DJSC coding over a Gaussian multiple-access channel (MAC) using concatenated low density generator matrix (LDGM) codes, can be found in [3].

3.3 Source-Channel Separation For MACs

In Chapters 4 and 5, we will discuss DJSC coding for two-user orthogonal MAC and the two-user GMAC respectively. In the following, we discuss the optimality of the separation of source and channel coding for these channels from an information theoretic perspective.

3.3.1 Two-User Orthogonal MAC

A two-user orthogonal MAC is defined by two separate independent channels (see Fig. 3.2(a)). Since the the two channels are independent, the right hand side in (2.6) becomes,

$$\begin{aligned} I(X_1;Y \mid X_2) &= I(X_1;Y), \\ I(X_2;Y \mid X_1) &= I(X_2;Y), \\ I(X_1,X_2;Y) &= I(X_1;Y) + I(X_2;Y). \end{aligned} \tag{3.1}$$

The maximum rates that can be achieved over this channel is given by (2.6). From (3.1), it is easy to see that the maximum rates are achieved by the product distribution $p(x_1)p(x_2)$ at the inputs of the channel, thus confirming that separation principle holds for a two-user orthogonal MAC. A formal proof can be found in [68].

3.3.2 Two-User GMAC

For a two-user GMAC, the output Y depends on *both* X_1 and X_2 (see Fig. 3.2(b)) . The dependence between X_1 and X_2 is modeled by the parameter $\alpha = Pr(X_1 \neq X_2)$. Channel output Y also depends on additive channel noise which has mean 0 and variance σ^2. We first note that since the channel output Y depends on both X_1 and X_2, $I(X_1, X_2; Y) \neq I(X_1;Y) + I(X_2;Y)$, which means that in order to find the maximum rates as in (2.6), the maximization has to be done over the *joint* distribution $p(x_1, x_2)$. Also note that if $\alpha = 0.5$, X_1 and X_2 are independent. On the other hand when $\alpha \in [0, 0.5)$, X_1 and X_2 are dependent (dependence increasing with the decreasing α). Also note that the maximum rate for any distribution $p(x_1, x_2)$ cannot be greater than $I(X_1, X_2; Y)$, although it is not known whether this maximum can be achieved for dependent sources or not. In general, the capacity of a GMAC with correlated source is not known.Therefore, we use the maximum of $I(X_1, X_2; Y)$ as an upper bound to the capacity in this case.

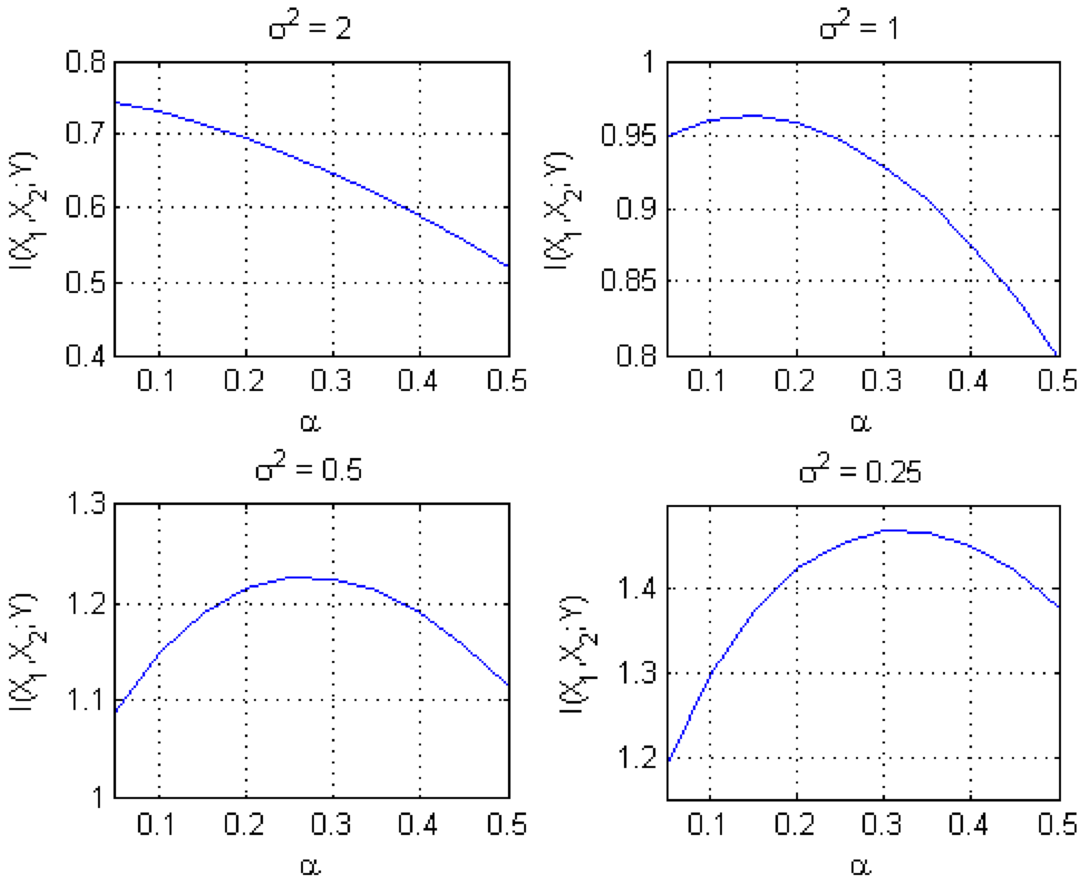

Figure 3.3: $I(X_1, X_2; Y)$ as a function of inter-source correlation parameter α for different values of GMAC noise variance σ^2.

Fig. 3.3 shows $I(X_1, X_2; Y)$ as function of the correlation parameter α for different values of the channel noise variance σ^2. We see that in each case, the maximum of $I(X_1, X_2; Y)$ does not occur at $\alpha = 0.5$ which suggests that the maximum sum-rate that can be achieved when X_1 and X_2 are dependent *can be* higher than the sum-rate that can be achieved when they are independent. We will see in Chapter 5 that this is indeed the case and by making use of source correlation in the code design and in decoding, we can in fact beat the theoretical rate bound for independent sources.

Chapter 4

Code Design for Orthogonal Multiple Access Channels

Distributed source coding (DSC) [24] is considered as one of the enabling technologies for wireless sensor networks (WSNs) in which a set of distributed sensors pick-up correlated information and communicate their readings to a central decoder. The basis for lossless DSC is the Slepian-Wolf (SW) theorem [47] which shows that the rates achievable with the joint lossless encoding of two discrete sources can also be achieved with separate encoding, if joint decoding can be used. After the initial work in [31], a considerable progress has been made in the practical SW code design using cosets and syndromes of linear channel codes, see [69] and references therein. However, such SW codes in practice suffer from two shortcomings: (i) very long codes ($> 10^5$ bits [70]) are required to achieve acceptable performance, and (ii) they can be extremely sensitive to channel noise [71]. Therefore, syndrome-based SW codes are unlikely to be effective when low delay/encoding-complexity is required and communication takes place over noisy channels, such as in WSNs. Even though the source-channel separation is asymptotically (in codeword length) optimal for SW coding over independent channels [68], joint source-channel (JSC) coding can outperform separate source and channel (SSC) coding, when constraints are imposed on the codeword length. This provides the motivation to find constructive methods for designing distributed JSC (DJSC) codes which are good for finite codeword lengths.

This chapter presents an approach to designing distributed joint source channel (DJSC) codes with arbitrary rates for communication of a pair of correlated binary sources over noisy channels. The proposed framework is a generalization of the aforementioned parity-bit SW coding approach. While our goal is to design DJSC codes which perform well for short codeword lengths, the design procedure is based on a code construction which is shown to be asymptotically optimal, *i.e.,* any rate pair in the achievable rate region for the two sources and respective communication channels can be reached. The practical implementation of the proposed codes involves the design of channel codes with a specific unequal error protection (UEP) property which is required to exploit the statistical dependence between the two sources for simultaneous (distributed) compression and channel error correction. To this end we develop a low-complexity linear programming algorithm based on extrinsic information transfer (EXIT) chart analysis [72] for optimizing the degree profiles of a pair of irregular LDPC codes which are jointly decoded by the belief propagation (BP) algorithm. As specific examples of communication channels, we consider both binary-symmetric channel (BSC) and additive white Gaussian noise (AWGN) channel with binary phase shift keying (PSK) inputs (Bi-AWGN channel). Previously, LDPC code design for UEP has been considered in, for example, [73, 74], and irregular LDPC codes which apply UEP for information and parity bits in the context of DSC has been considered in [60]. However, these code designs only assign a higher priority to information bits compared to the parity bits, and the channel capacity is not explicitly taken into account. In contrast, the UEP code design procedure we propose takes the capacity of the channels through which the various bits are transmitted into account, and therefore is better suited for constructing DJSC codes with a short codeword length.

We present experimental results which confirm that the LDPC codes designed by the proposed procedure exhibit the desired UEP property which is key to realizing good DJSC codes. It has been observed that the code optimization procedure consistently produced codes which are better than the best codes found by searching through the possible LDPC degree profiles (as, for example, done in [60, 74]). In order to characterize the rate-loss due to finite and short codeword lengths, we have investigated the rates achievable with practically designed DJSC codes with codeword lengths, particularly in the range of 1000-5000 bits, over both BSCs and Bi-AWGN channels. Importantly, simulation results confirm that, when the constraints are placed on coding block length (or equivalently, the delay and complexity) DJSC codes designed by the proposed method considerably outperform the alternative SSC coding in which SW coding is used for distributed compression and separate channel coding is used for channel error correction. Our simulation results also show that the DJSC codes constructed from LDPC codes with optimized degree profiles outperform the turbo coding schemes of [2] (for BSCs) and [3] (for AWGN channels). While our code design procedure does not take into account the non-uniform distribution of source bits, we demonstrate through simulations that joint decoding of the two sources allows us to exploit the redundancy within each source in the form of non-uniform bit probabilities to further offset the performance loss due to short codeword lengths.

4.1 An Approach to Optimal DJSC Code Construction

Let X_1 and X_2 be two dependent and uniformly distributed binary sources, *i.e.,* $H(X_1) = H(X_2) = 1$ bit. In this case, the dependence between X_1 and X_2 can be modeled by a "virtual" BSC with cross-over probability $p = Pr(X_1 \neq X_2)$, where $H(X_1|X_2) = H(X_2|X_1) = H(p)$. In DJSC coding, each source is encoded by a separate JSC encoder and transmitted to a decoder which jointly decodes both sources. We assume that the encoders for X_1 and X_2 transmit their outputs over independent channels with capacities C_1 and C_2 bits respectively. Let the rates of the two encoders be R_{X_1} and R_{X_2} channel-uses/source-symbol.

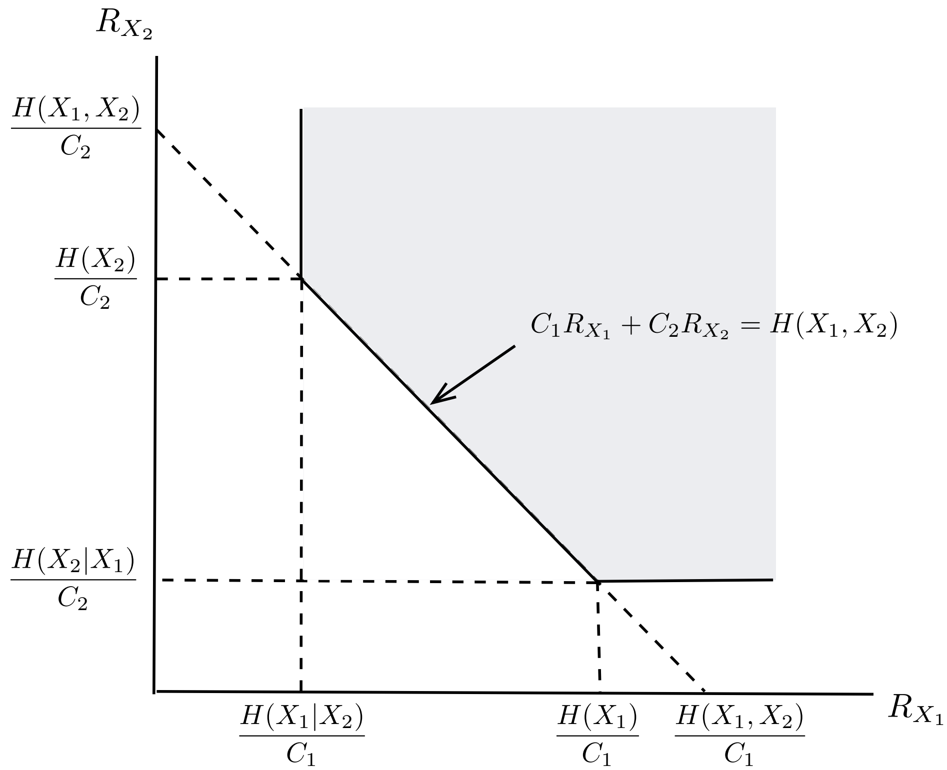

Figure 4.1: Achievable rate-region for DJSC coding of binary sources X_1 and X_2. The rates R_{X_1} and R_{X_2} are in channel-uses/source-symbol.

It is known that source-channel separation is asymptotically optimal for our problem [68]. Hence, by combining Slepian-Wolf and Shannon's theorems we can obtain the achievable rate region for DJSC coding [63]. Suppose $\mathbf{X}_i$ is first encoded by a source code of rate $R_{s,i}$ bits/source-symbol and then by a channel code whose rate is such that the overall joint source-channel coding rate is R_{X_i} channel-uses/source-symbol, $i = 1, 2$. From SW theorem, we have $R_{s,1} \geq H(X_1|X_2)$, $R_{s,2} \geq H(X_2|X_1)$ and $R_{s,1} + R_{s,2} \geq H(X_1, X_2)$, and from channel coding theorem we have $\frac{R_{s,i}}{R_{X_i}} \leq C_i$, $i = 1, 2$. Then, it follows that

$$
\begin{aligned}
R_{X_1} &\geq \frac{H(X_1|X_2)}{C_1}, && (4.1)\\
R_{X_2} &\geq \frac{H(X_2|X_1)}{C_2}, && (4.2)\\
C_1 R_{X_1} + C_2 R_{X_2} &\geq H(X_1, X_2). && (4.3)
\end{aligned}
$$

This region is illustrated in Fig. 4.1.

We will now present a DJSC code construction, which can approach any rate-pair in the achievable rate-region shown in Fig. 4.1. Consider encoding of k-bit sequences $\mathbf{X}_1 = (X_1^{(1)}, \ldots, X_1^{(k)})$ and $\mathbf{X}_2 = (X_2^{(1)}, \ldots, X_2^{(k)})$ from X_1 and X_2. Suppose that the encoder for X_1 transmits the first t bits $(X_1^{(1)}, \ldots, X_1^{(t)})$, while the encoder for X_2 transmits the last $k-t$ bits $(X_2^{(t+1)}, \ldots, X_2^{(k)})$, where $0 \leq t \leq k$. Let the *rate-parameter* $a = t/k$, where $0 \leq a \leq 1$. We will next establish that, if an appropriate number of parity bits generated by encoding $\mathbf{X}_1$ and $\mathbf{X}_2$ using a pair of systematic channel codes are transmitted for the respective sources, then all bits of $\mathbf{X}_1$ and $\mathbf{X}_2$ can be decoded asymptotically error-free. Furthermore, we will prove that by varying a between 0 and 1, this encoding scheme can approach any rate-pair in the achievable rate-region (4.1)-(4.3).

Let the encoder for X_1 transmit m_1 parity bits (in addition to source bits) obtained by encoding $\mathbf{X}_1$ using a rate r_1 systematic channel code and let the encoder for X_2 transmit m_2 parity bits obtained by encoding $\mathbf{X}_2$ using a rate r_2 systematic channel code, where $r_i = k/n_i$ and $n_i = k + m_i$ is the codeword length, $i = 1, 2$. For a given value of a, the JSC coding rates are given by $R_{X_1} = (ak + m_1)/k$ and $R_{X_2} = ((1-a)k + m_2)/k$ (channel-uses/bit). This encoding scheme is illustrated in Fig. 4.2. We refer to those bits which are explicitly transmitted by a given source as *type 1* bits and those which are assumed to be transmitted to the other source over the virtual correlation channel as *type 2* bits. Note that each source can transmit its *type 1* bits reliably to the decoder provided that its channel code contains

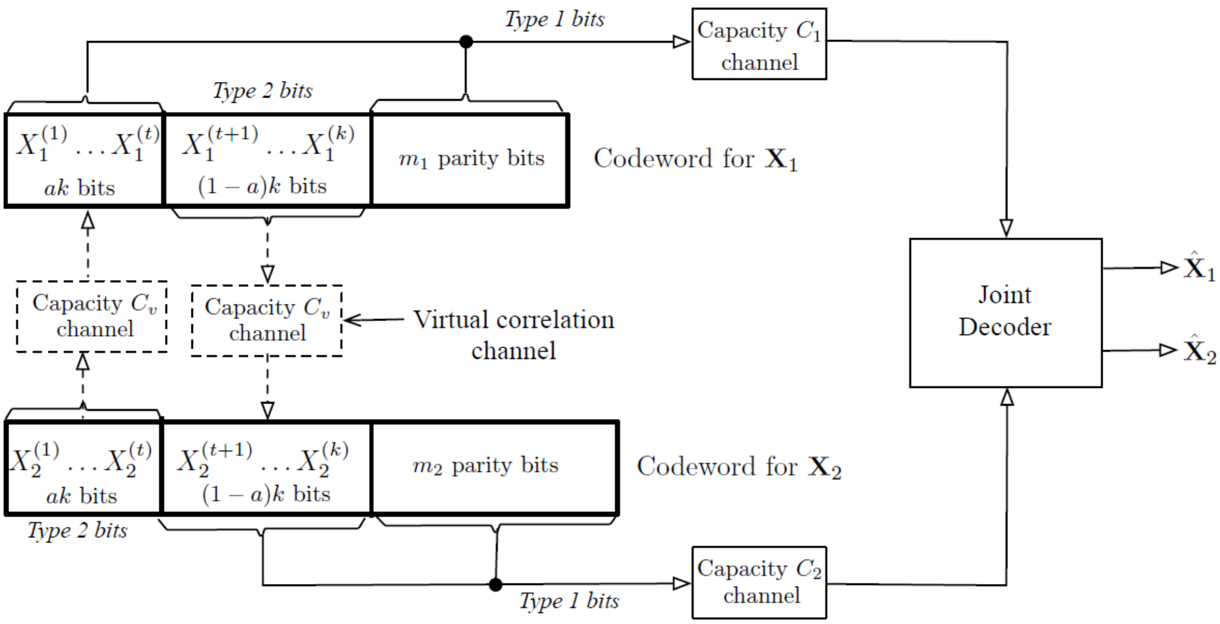

Figure 4.2: The proposed DJSC code structure: codes for source X_1 (top) and X_2 (bottom). By varying a between 0 and 1, any rate-pair in the achievable rate-region can be realized.

a sufficient number of parity bits to correct errors in the communication channel. Then, both sources can be decoded asymptotically error-free at the receiver, if each code also has a sufficient number of parity bits to "implicitly" send *type 2* bits reliably to the other source over the virtual correlation channel whose capacity is $C_v = 1 - H(p)$. To see this, consider the systematic code used for X_i as being composed of two sub-codes, one for *type 1* bits and the other for *type 2* bits, such that $m_i = m_i' + m_i''$, where m_i' and m_i'' respectively are the number parity bits in the sub-codes of X_i, $i = 1, 2$. Then, the codes in Fig. 4.2 can be viewed as follows. A fraction a of time, the encoder for X_1 transmits *type 1* source bits and the parity bits for *type 1* source bits, while the encoder for X_2 only transmits the parity bits for *type 2* source bits. The rest [a fraction $(1 - a)$] of the time, the roles of the encoders for X_1 and X_2 are reversed. Clearly, after decoding *type 1* bits of one source, *type 2* bits of the other source can be decoded. Note that this scheme is equivalent to time-sharing of two corner points as in [47, Table I, Theorem g] where the subcodes for each source consist of an entropy code and an SW code.

Next we determine the minimum number of parity bits m_1' and m_1'' required for coding X_1 at a given value of a. Since the codes are systematic, without loss of generality, consider a single (n_1, k) code with $m_1 = m_1' + m_1''$ parity bits for X_1, used on the equivalent channel formed by the parallel combination of the virtual channel and the communication channel. Out of $n_1 = k + m_1$ code bits, $(ak + m_1)$ bits pass through the communication channel, while $(1 - a)k$ bits pass through the virtual channel. Therefore, we can reliably send a maximum of $(ak + m_1)C_1 + (1 - a)kC_v$ bits of information in n_1 channel uses, or $I_1 = ar_1C_1 + (1 - a)r_1C_v + (1 - r_1)C_1$ bits/channel-use. According to the Shannon's channel coding theorem we require the code-rate $r_1 \leq I_1$ and therefore it follows that

$$r_1 \leq \frac{C_1}{1 + (1-a)(C_1 - C_v)}, \tag{4.4}$$

$$m_1 \geq \frac{[1 - C_1 + (1-a)(C_1 - C_v)]k}{C_1}. \tag{4.5}$$

Similarly, by replacing $(1 - a)$ and C_1 by a and C_2 respectively, it can be shown that for X_2,

$$r_2 \leq \frac{C_2}{1 + a(C_2 - C_v)}, \tag{4.6}$$

$$m_2 \geq \frac{[1 - C_2 + a(C_2 - C_v)]k}{C_2}. \tag{4.7}$$

In this work, we consider channels with binary inputs, such as BSC and Bi-AWGN channel, for which $C_1, C_2 \leq 1$ (while the capacity of the Bi-AWGN channel cannot be expressed in closed-form, it can be evaluated numerically, see [32, pp. 194], [75] for details). From (4.5) and (4.6), we obtain the corresponding lower bounds on achievable JSC coding rates for

$0 \leq a \leq 1$ as

$$R_{X_1} \geq \frac{1-(1-a)C_v}{C_1}, \quad (4.8)$$

$$R_{X_2} \geq \frac{1-aC_v}{C_2}. \quad (4.9)$$

Since $H(X_1|X_2) = H(X_2|X_1) = 1 - C_v$ and $H(X_1) = H(X_2) = 1$, for $a = 0$ we have $R_{X_1} \geq \frac{H(X_1|X_2)}{C_1}$ and $R_{X_2} \geq \frac{H(X_2)}{C_2}$ while for $a = 1$ we have $R_{X_1} \geq \frac{H(X_1)}{C_1}$ and $R_{X_2} \geq \frac{H(X_2|X_1)}{C_2}$. Therefore, with $a = 0$ and $a = 1$, the two corner points A and B of the rate-region in Fig. 4.1 can be achieved with the coding scheme in Fig. 4.2. Furthermore, for $0 \leq a \leq 1$, (4.8) and (4.9) imply

$$C_1 R_{X_1} + C_2 R_{X_2} \geq 2 - C_v. \quad (4.10)$$

Since $H(X_1, X_2) = 2 - C_v$, (4.10) coincides with the sum-rate bound (4.3), *i.e.*, any point on the line AB in Fig. 4.1 can be achieved by our coding scheme for $0 \leq a \leq 1$. Therefore, the proposed DJSC code construction is asymptotically optimal. It should be noted that, when $C_1 = C_2 = 1$ (*i.e.,* channels are noiseless), the above DJSC code reduces to SW codes considered in [60, 76].

Separately decodable codes- An alternative coding scheme is to use a channel code for each source which contains enough parity bits to reliably send *type-2* source bits over the concatenated channel formed by the virtual channel and the communication channel. In this case, each source can be designed separately and decoded by an independent decoder which receives channel outputs for both sources. The rate bounds for this scheme can be obtained by replacing C_v in (4.5) to (4.10) by the appropriate concatenated channel capacity, either C_v' (C_v and C_2) or C_v'' (C_v and C_1). Clearly, this coding scheme is sub-optimal since $C_v', C_v'' \leq C_v$, and the performance gap compared to jointly designed and decoded codes will be significant for $q >> p$.

4.2 Practical Code Design Procedure

The implementation of the optimal DJSC coding scheme in Fig. 4.2 requires a pair of channel codes with a particular UEP property. More specifically, we note that, out of $n_j = k + m_j$ codeword bits of the source X_j, $j = 1, 2$, a fraction α_j (*type-1* bits) is transmitted through a channel with capacity C_j, where

$$\alpha_j = \frac{k}{n_j C_j} \left[1 - \omega_j C_v\right], \tag{4.11}$$

with $\omega_1 = (1 - a)$ and $\omega_2 = a$. The remaining $\bar{\alpha}_j = 1 - \alpha_j$ fraction of bits (*type 2* bits) is assumed to be transmitted through the virtual correlation channel with capacity C_v. Our objective is to design for each source, a UEP code such that the level of protection provided to each bit is commensurate with the capacity of the channel that the bit passes through. In this section, we propose an approach based on irregular LDPC codes to achieve this objective. In an irregular LDPC code, the number of edges incident on variable nodes which represent the bits of a codeword in the bipartite graph of the code are chosen according to some distribution [32]. In general, the code bits which correspond to nodes with a higher *degree* (the number of connected edges) receive more protection (has a lower decoding error probability) compared to those with a lower degree [74]. In this section, we follow this idea to develop a procedure for jointly designing a pair of LDPC codes. In principle, approaching the rate-bounds (4.1)-(4.3) using the DJSC code shown in Fig. 4.2 requires optimal joint decoding of the codewords transmitted for both sources. When a separate LDPC code is used for encoding each source, a close approximation to the maximum likelihood joint decoder can be implemented by using the BP algorithm for iterative message passing between the bipartite graphs of the two LDPC codes. Previously, joint decoding of two LDPC codes in the context of syndrome-based WZ coding has been considered in [77, 78]. The factor-graph of the joint decoder used in our approach is shown in Fig. 4.3, where the two-sub-graphs on

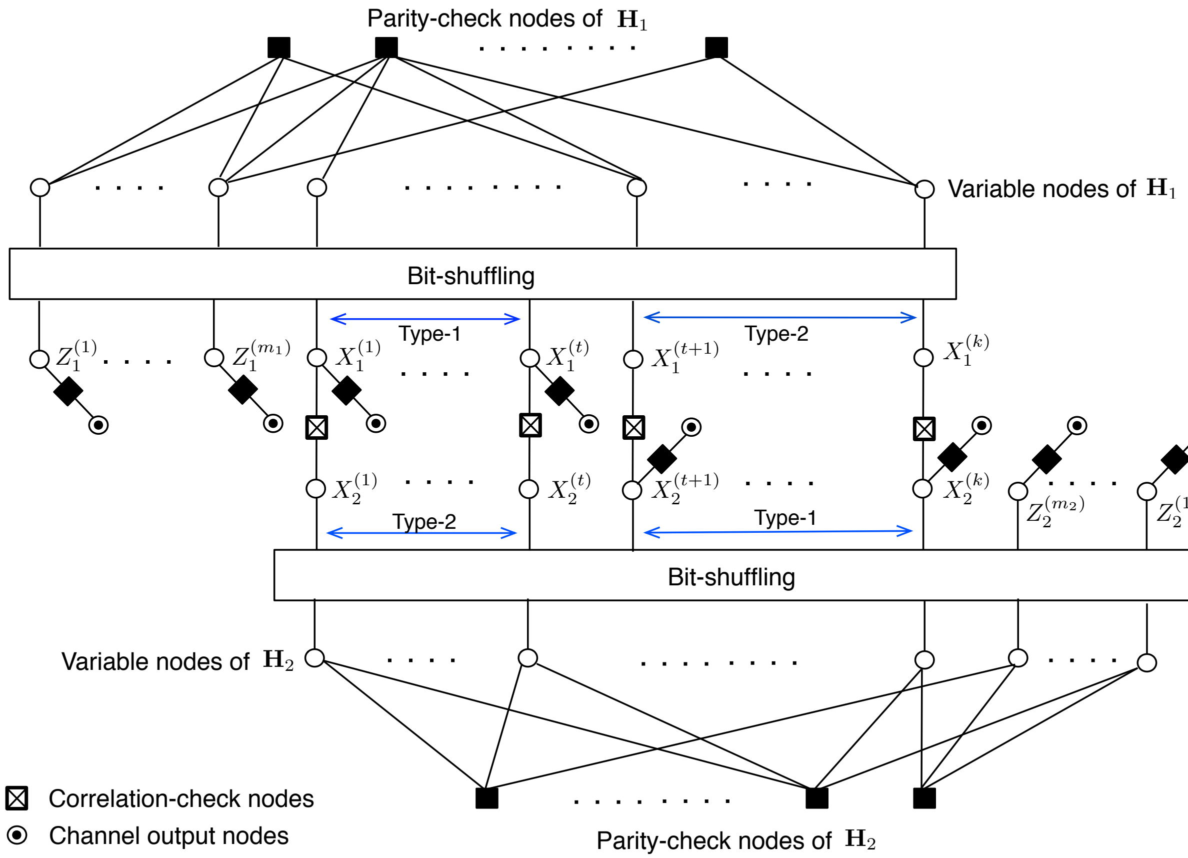

Figure 4.3: Factor graph of the joint decoder. $X_1^{(1)}, \ldots, X_1^{(k)}$ and $X_2^{(1)}, \ldots, X_2^{(k)}$ (see Fig. 4.2) are the systematic information bits of the sources X_1 and X_2 respectively. $Z_j^{(1)}, \ldots, Z_j^{(m_j)}$, $j = 1, 2$ are the parity bits.

either side represent parity-check matrix $\mathbf{H}_1$ and $\mathbf{H}_2$ used for encoding two sources X_1 and X_2 respectively. The factor nodes in this graph represent either parity check constraints of each LDPC code (parity-check nodes) or the joint probability $Pr(X_1, X_2)$ of the corresponding information bits of X_1 and X_2, which we refer to as *correlation-check nodes*. Bit-shuffling operation in Fig. 4.3 is described at the end of this section.

In general, a length n irregular LDPC code with parity-check matrix $\mathbf{H}$ can be completely specified by the parameters $(n, \lambda(x), \rho(x))$, where $\lambda(x) = \sum_{i=1}^{d_{vmax}} \lambda_i x^{i-1}$ and $\rho(x) = \sum_{i=1}^{d_{cmax}} \rho_i x^{i-1}$ are the *edge-perspective* degree polynomials of variable nodes and check nodes respectively, and λ_i (resp. ρ_i) is equal to the fraction of edges connected to variable (resp. check) nodes of degree i, satisfying the constraints $\sum_i \lambda_i = 1$ and $\sum_i \rho_i = 1$ [32]. It is known that a concentrated degree polynomial of the form $\rho(x) = \rho x^{s-1} + (1-\rho)x^s$ for some $s \geq 2$ and $0 < \rho \leq 1$ is sufficient for achieving near optimal performance [79, Theorem 2]. A well known simple method for constructing near-capacity irregular LDPC codes for a given channel parameter (cross-over probability q for a BSC or noise variance σ^2 for a BiAWGN channel) and some fixed $\rho(x)$, is to determine the coefficients λ_i which maximize the rate of the code under BP decoding, subject to Gaussian approximation (GA) for the messages passed in the decoder [79]. The code design in this case is a linear-programming problem of the form [32, Ch. 4]: *maximize*$_{\lambda_i}$ $\sum_{i\geq 2} \lambda_i/i$, *subject to constraints* 1) $\sum_i \lambda_i = 1, (0 < \lambda_i \leq 1)$ (*normalization constraint*), 2) $\lambda_2 < K/\sum_j (j-1)\rho_j$ (*stability condition*), where the constant K depends on the channel parameter, and 3) a linear inequality to ensure the convergence of the BP algorithm [79, Sec. III] (*decoder convergence constraint*).

UEP code structure- In order to achieve the aforementioned UEP property, we define each parity check matrix $\mathbf{H}_j$ using two separate degree polynomials $\lambda_j^{(1)}(x)$ and $\lambda_j^{(2)}(x)$ respectively for *type 1* and *type 2* variable nodes, where $j = 1, 2$. Thus, the code for source X_j is completely specified by $(n_j, \lambda_j^{(1)}(x), \lambda_j^{(2)}(x), \rho_j(x))$, where

$$\lambda_j^{(k)}(x) = \sum_{i=2}^{d_{vmax,j}} \lambda_{j,i}^{(k)} x^{i-1}, \tag{4.12}$$

and $d_{vmax,j}$ is the maximum allowable degree of variable nodes and $0 < \lambda_{j,i}^{(k)} < 1$, $k = 1, 2$ is the fraction (or density) of edges connected to a *type* k variable node of degree i, *i.e.,*

$$\sum_{i=2}^{d_{vmax,j}} \left(\lambda_{j,i}^{(1)} + \lambda_{j,i}^{(2)}\right) = 1, \tag{4.13}$$

for $j = 1, 2$. Let $N_{e,j}$ be the total number of edges in the bipartite graph of $\mathbf{H}_j$. Then, the number of *type* 1 variable nodes is $N_{e,j} \sum_{i=1}^{d_{vmax,j}} \lambda_{j,i}^{(1)}/i$, the number of *type* 2 variable nodes is $N_{e,j}\sum_{i=1}^{d_{vmax,j}}\lambda_{j,i}^{(2)}/i$, and the number of check nodes is $N_{e,j}\sum_{i=1}^{d_{cmax}}\rho_i/i$. It follows that the average variable node degree is

$$\bar{d}_{v,j}=\frac{1}{\sum_i\frac{\lambda_{j,i}^{(1)}}{i}+\sum_i\frac{\lambda_{j,i}^{(2)}}{i}}$$

and the average check node degree is $\bar{d}_{c,j}=1/\sum_{i=1}^{d_{cmax}}(\rho_i/i)$. Therefore, the *design rate* of the code $\mathbf{H}_j$ is given by [32, (3.19)]

$$R_{\text{des,j}}=1-\bar{d}_{v,j}/\bar{d}_{c,j}. \tag{4.14}$$

From (4.5), (4.6), (4.8), and (4.9) it follows that the JSC rates R_{X_1} and R_{X_2} are proportional to $1/r_1$ and $1/r_2$ respectively. Therefore in order to design a code-pair which can achieve the sum rate bound given in (4.10), we equivalently maximize $R_{\text{des},1}+R_{\text{des},2}$. That is, for a given s and ρ, we find the degree distributions $\lambda_j^{(1)}(x)$ and $\lambda_j^{(2)}(x)$, $j=1,2$, which maximize the objective function

$$\mathcal{J}=\sum_{j=1}^{2}\sum_{i=2}^{d_{vmax,j}}\left(\frac{\lambda_{j,i}^{(1)}}{i}+\frac{\lambda_{j,i}^{(2)}}{i}\right), \tag{4.15}$$

subject to appropriate constraints. We now derive the constraints required, in addition to the normalization constraint (4.13).

Stability condition- In LDPC code design, a constraint is required to enforce the decoder stability condition determined by the channel model [32, 79]. While the condition used in [32, 79] is exact for density evolution, it is also a good approximation for EXIT chart-based design when concentrated check-node polynomials are used [79, pp. 664]. In our UEP code design this constraint is different for *type 1* and *type 2* bits. From [32], it follows that for *type 2* bits, which pass through a virtual BSC of cross-over probability p, the stability constraint is

$$\lambda_2^{(2)}<\frac{\bar{\alpha}}{2\sqrt{p(1-p)}\sum_{i=2}^{d_{cmax}}(i-1)\rho_i}. \tag{4.16}$$

For *type 1* bits, this condition depends on the physical communication channels. If the channels are BSCs with error probability q, then we require

$$\lambda_2^{(1)} < \frac{\alpha}{2\sqrt{q(1-q)}\sum_{i=2}^{d_{cmax}}(i-1)\rho_i}, \tag{4.17}$$

whereas if the channels are Bi-AWGN channels with input ± 1 and noise variance σ^2, then we require

$$\lambda_2^{(1)} < \frac{\alpha e^{\frac{1}{\sigma^2}}}{\sum_{i=2}^{d_{cmax}}(i-1)\rho_i}. \tag{4.18}$$

UEP constraint- Different to an LDPC code in which all bits pass through the same channel, an additional constraint is needed in our UEP codes due to the fact that a fraction of variable nodes must be *type 1* and the rest *type 2*. To this end, consider the *node-perspective* degree polynomial of the *type 1* variable nodes in the graph of $\mathbf{H}_j$ $(j = 1, 2)$, $\Lambda_j^{(1)}(x) = \sum_{i=2}^{d_{vmax,j}} \Lambda_{j,i}^{(1)} x^i$, where $\Lambda_{j,i}^{(1)}$ is the number of degree i variable nodes of *type 1*. It is easy to verify that $\Lambda_j^{(1)}(1) = \sum_{i=2}^{d_{vmax,j}} \Lambda_{j,i}^{(1)} = \alpha_j n_j$. Similarly, let $\Lambda_j^{(2)}(x) = \sum_{i=2}^{d_{vmax,j}} \Lambda_{j,i}^{(2)} x^i$ be the node-perspective degree polynomial for *type 2* variable nodes, where $\Lambda_j^{(2)}(1) = \bar{\alpha}_j n_j$. It follows that $\lambda_j^{(1)}(x)$ and $\Lambda_j^{(1)}(x)$ are related by

$$\lambda_j^{(1)}(x) = \frac{\frac{d}{dx}\Lambda_j^{(1)}(x)}{\left[\frac{d}{dx}\Lambda_j^{(1)}(x) + \Lambda_j^{(2)}(x)\right]_{x=1}}. \tag{4.19}$$

A similar relationship exists between $\lambda_j^{(2)}(x)$ and $\Lambda_j^{(2)}(x)$. Since $\lambda_{j,i}^{(1)}$ is the ratio of the number of edges connected to degree i variable nodes of *type 1* to the total number of edges in graph of $\mathbf{H}_j$,

$$\lambda_{j,i}^{(1)} = \frac{i\Lambda_{j,i}^{(1)}}{\sum_{i=2}^{d_{vmax,j}} i\Lambda_{j,i}^{(1)} + i\Lambda_{j,i}^{(2)}}$$

and hence

$$\int_0^1 \lambda_j^{(k)}(z)dz = \frac{\Lambda_j^{(k)}(1)}{\frac{d}{dx}\left(\Lambda_j^{(1)}(x) + \Lambda_j^{(2)}(x)\right)|_{x=1}}, \; k = 1, 2. \tag{4.20}$$

It therefore follows that

$$\frac{\int_0^1 \lambda_j^{(1)}(z)dz}{\Lambda_j^{(1)}(1)} = \frac{\int_0^1 \lambda_j^{(2)}(z)dz}{\Lambda_j^{(2)}(1)}, \tag{4.21}$$

which leads to the following constraint relating $\lambda_j^{(1)}(x)$ and $\lambda_j^{(2)}(x)$:

$$\frac{1}{\alpha_j} \sum_{i=2}^{d_{vmax,j}} \frac{\lambda_{j,i}^{(1)}}{i} = \frac{1}{\bar{\alpha}_j} \sum_{i=2}^{d_{vmax,j}} \frac{\lambda_{j,i}^{(2)}}{i}, \tag{4.22}$$

where $j = 1, 2$.

Joint decoder convergence constraint- Let $I_{AC,j} \in (0, 1]$ be the average extrinsic (*a prior*) information in incoming messages to parity-check nodes of $\mathbf{H}_j$. Also, let $I_{EV,j}(I_{AC,j})$ be the average extrinsic information in outgoing messages from variable nodes. Then, for convergence of the sum-product algorithm on the joint decoder graph, we require the two constraints [72]

$$I_{EV,j}(I_{AC,j}) > I_{AC,j}, \; j = 1, 2. \tag{4.23}$$

Note however that these two constraints are not independent. In deriving these constraints specifically for our joint decoder graph, we need to first differentiate between *type 1* and *type 2* nodes of each code. A given variable node in the decoder graph receives information from the channel output, a set of parity check nodes, and exactly one correlation-check node (see Fig. 4.3). Further note that, the channel information received by *type 1* variables from the channel is $I_{ch,j}^{(1)} = C_j$. However, since the *type 2* bits are not explicitly transmitted, $I_{ch,j}^{(2)} = 0$. Let $I_{EC,j}(I_{AC,j})$ be the average extrinsic information passed from parity-check nodes to variable nodes in $\mathbf{H}_j$, and $I_{CC,j}$ be the extrinsic information in messages passed from correlation-check nodes to variable nodes of $\mathbf{H}_j$. Then, the output extrinsic information from a degree $i+1$ variable node of *type* k in the graph of $\mathbf{H}_j$ is given by

$$I_{EV,j}^{i,k} = J\left(J^{-1}(I_{ch,j}^{(k)}) + (i-1)J^{-1}(I_{EC,j}) + J^{-1}(I_{CC,j})\right) \tag{4.24}$$

for $k = 1, 2$, where $J^{-1}(I_{ch,j}^{(k)})$, $(i-1)J^{-1}(I_{EC,j})$, and $J^{-1}(I_{CC,j})$ are respectively the average means of the pdfs of incoming messages from the channel, other $(i-1)$ parity-check nodes, and the correlation-check node, and $J(\cdot)$ is the mutual information function given by [72, (15)] (both $J(\cdot)$ and $J^{-1}(\cdot)$ can be computed using the approximation given in [75]). The average output extrinsic information from a variable node of $\mathbf{H}_j$ is thus

$$I_{EV,j} = \sum_{k=1}^{2} \sum_{i=2}^{d_{vmax,j}} \lambda_{j,i}^{(k)} I_{EV,j}^{i,k}. \tag{4.25}$$

We note that for messages passed between parity-check nodes and variable nodes of either LDPC code, the usual Gaussian approximation [79] can be applied. Therefore, $I_{EC,j}$ in (4.24) can be computed using the duality approximation [80], [32, pp. 236]

$$I_{EC,j} \approx 1 - \sum_{j=s-1}^{s} \rho_j J\left((j-1)J^{-1}(1 - I_{AC,j})\right). \tag{4.26}$$

However, we will show that the pdf of an outgoing message of a correlation-check node is not well approximated by a Gaussian. Therefore, computing $I_{CC,j}$ requires a different approach as presented below.

As usual let the messages passed between nodes in the factor-graph be in the form of log-likelihood ratios (LLRs). Consider a correlation check-node which receives the message $l_{in}^{(2)}$ from a variable node in $\mathbf{X}_2$ and computes the outgoing message $l_{out}^{(1)}$ to be passed to a variable node in $\mathbf{X}_1$. Let $\beta_{im}^{(1)} = Pr(X_1 = i | X_2 = m)$ and $\beta_{im}^{(2)} = Pr(X_2 = i | X_1 = m)$, where

$i, m \in \{0,1\}$, so that the likelihood-ratio computed in the correlation-check node is

$$\frac{Pr(X_1=1)}{Pr(X_1=0)} = \frac{\beta_{10}^{(1)} Pr(X_2=0) + (1-\beta_{01}^{(1)}) Pr(X_2=1)}{(1-\beta_{10}^{(1)}) Pr(X_2=0) + \beta_{01}^{(1)} Pr(X_2=1)}.$$

Now, by replacing the ratio of prior probabilities by incoming LLR $l_{in}^{(2)}$ from variable node X_2, we obtain the outgoing message of the correlation-check node as

$$l_{out}^{(1)} = \ln\left[\frac{1 + B_1 \exp(l_{in}^{(2)})}{B_2 + B_3 \exp(l_{in}^{(2)})}\right], \tag{4.27}$$

where $B_1 = \frac{(1-\beta_{01}^{(1)})}{\beta_{10}^{(1)}}$, $B_2 = \frac{(1-\beta_{10}^{(1)})}{\beta_{10}^{(1)}}$, and $B_3 = \frac{\beta_{01}^{(1)}}{\beta_{10}^{(1)}}$. As usual, assume that the Gaussian approximation holds for $l_{in}^{(2)}$ which is passed from a variable node to a correlation-check node, *i.e.*, $l_{in}^{(2)}$ is Gaussian with mean μ and variance 2μ [79]. The corresponding density of $l_{out}^{(1)}$ estimated via Monte-Carlo simulations shows that it is uni-modal (approximately Gaussian) for small $|\mu|$ and bi-modal for large $|\mu|$. A closer look at (4.27) shows that for small $|\mu|$, $l_{out}^{(1)}$ is distributed around $\ln\left(\frac{1+B_1}{B_2+B_3}\right)$. On the other hand, for sufficiently large $|\mu|$, $l_{out}^{(1)}$ will be distributed either around $\ln(B_1/B_3)$ (exponentials in numerator and denominator dominate) or $\ln(1/B_2)$ (exponentials are insignificant), giving rise to a bi-modal distribution. Typical histograms are shown in Fig. 4.4 for various values of μ at $p = 0.1$. Therefore, we use a two-component Gaussian mixture approximation for outgoing messages from correlation-check nodes. The extrinsic information in outgoing messages from correlation-check node $l_{out}^{(1)}$ can therefore be given by $I_{CC,1}(I_{ACC,2}) = c_1 J(\mu_1) + c_2 J(\mu_2)$, where μ_1 and μ_2 are the modes of the Gaussian mixture, and c_1 and c_2 are the weights and $I_{ACC,2}$ is the extrinsic information received in $l_{in}^{(2)}$. These values can be estimated via Monte-Carlo-simulation, given the mean $\mu = J^{-1}(I_{ACC,2})$ of $l_{in}^{(2)}$. For the messages passed from correlation-check nodes to variable nodes of $\mathbf{X}_2$, $I_{CC,2}(I_{ACC,1})$ can also be computed in a similar manner.

The objective function given in (4.15) as well as the constraints (4.13), (4.16)-(4.18), (4.22), and (4.23) are linear in the unknown coefficients $\{\lambda_{j,i}^{(1)}, \lambda_{j,i}^{(2)}, 2 \le i \le d_{vmax}, j = 1, 2\}$.

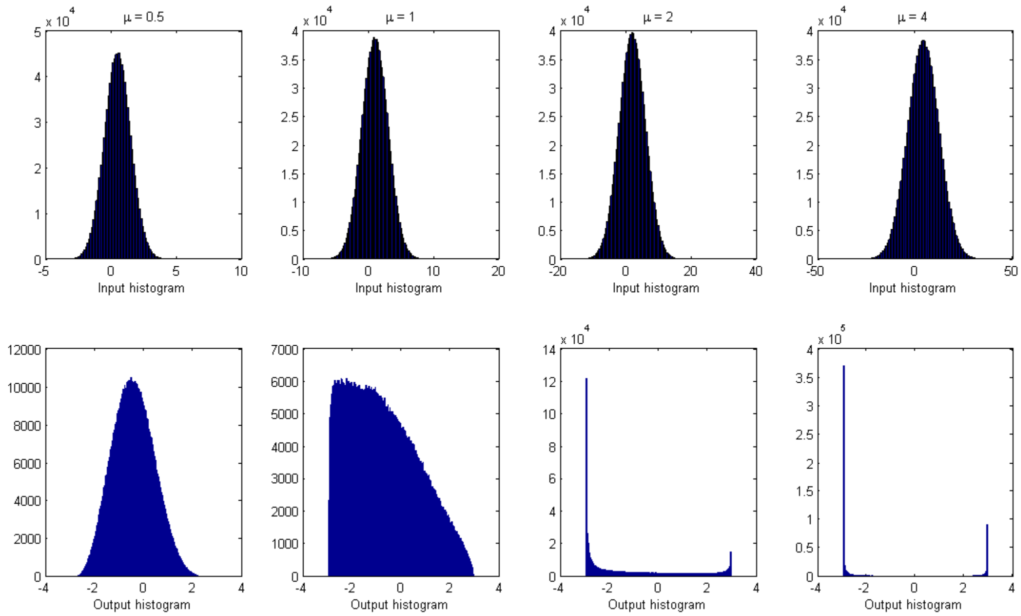

Figure 4.4: Pdf estimates of input messages (top) and output messages (bottom) of the correlation-check nodes for $p = 0.1$.

Therefore, it is straightforward to find the optimum solution using a linear program to design both LDPC codes simultaneously. Note that sparse parity-check matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ obtained through this design procedure does not necessarily correspond to systematic generator matrices. As usual, systematic codes can be obtained by the Gaussian elimination method. However, as such codes have dense parity-check matrices (making BP decoding impractical), a bit-shuffling operation is used in the joint decoder (Fig. 4.3) to rearrange the systematic code-bits so that the codewords correspond to sparse matrices $\mathbf{H}_1$ and $\mathbf{H}_2$.

4.3 Simulation Results

We present simulation results to demonstrate the performance achievable with jointly designed UEP-DJSC codes as proposed in this chapter. As DJSC coding is expected to significantly outperform separate source and channel coding for relatively short codeword lengths, we focus on encoding source sequences of length in the range of $1000 - 5000$ bits. The prac-

Table 4.1: Degree profiles for UEP-LDPC codes generated by the design algorithm for different correlation parameter values p, BSCs with $q = 10p$, and the rate parameter $a = 0.2$.

	p				
	0.002	0.004	0.006	0.008	0.01
$\rho(x)$	$0.3x^9 + 0.7x^{10}$	$0.3x^9 + 0.7x^{10}$	$0.3x^9 + 0.7x^{10}$	$0.3x^9 + 0.7x^{10}$	$0.3x^9 + 0.7x^{10}$
$\lambda_1^{(1)}(x)$	$0.099x + 0.135x^2$	$0.074x + 0.23x^2$ $+0.016x^5$	$0.076x + 0.217x^2$ $+0.148x^7$	$0.113x + 0.168x^2+$ $0.065x^{19} + 0.153x^{20}$	$0.165x + 0.089x^2+$ $0.0514x^{39} + 0.241x^{40}$
$\lambda_1^{(2)}(x)$	$0.325x + 0.441x^2$	$0.217x + 0.110x^2$ $+0.263x^5$	$0.185x + 0.10x^2$ $+0.22x^6 + 0.058x^7$	$0.091x + 0.131x^2+$ $0.262x^7 + 0.019x^8$	$0.058x + 0.137x^2$ $+0.084x^7 + 0.102x^8$ $+0.073x^{27}$
$\lambda_2^{(1)}(x)$	$0.201x + 0.424x^2$ $+0.130x^4 + 0.094x^5$	$0.162x + 0.286x^2+$ $0.036x^7 + 0.247x^8$ $+0.132x^{19} + 0.011x^{20}$	$0.153x + 0.207x^2+$ $0.213x^7 + 0.04x^{23}+$ $0.069x^{24} + 0.212x^{49}$	$0.149x + 0.141x^2+$ $0.130x^4 + 0.484x^{49}$	$0.141x + 0.165x^2+$ $0.6031x^{49}$
$\lambda_2^{(2)}(x)$	$0.076x + 0.075x^6$	$0.041x + 0.085x^7$	$0.02x + 0.004x^2+$ $0.083x^7$	$0.012x + 0.033x^4+$ $0.019x^5 + 0.031x^{49}$	$0.013x + 0.019x^2+$ $0.06x^{49}$
R_1	0.788	0.733	0.707	0.653	0.608
R_2	0.744	0.653	0.577	0.508	0.433
r_1	0.950	0.906	0.862	0.817	0.776
r_2	0.880	0.790	0.712	0.641	0.576
P_e (*Type 1* bits)	9.92×10^{-5}	9.95×10^{-5}	1.03×10^{-4}	9.94×10^{-5}	9.95×10^{-5}
P_e (*Type 2* bits)	1.04×10^{-4}	9.92×10^{-5}	9.91×10^{-5}	1.05×10^{-4}	9.97×10^{-5}

tical code design procedure can be summarized as follows. Suppose we need to design a pair of length n (bits) DJSC code with rates (R_{X_1}, R_{X_2}) channel-uses/source-symbol, specified by a particular value of the rate-parameter a. Given the value of a, the source-correlation parameter p, and the channel noise level (probability q for BSC or noise variance σ^2 for Bi-AWGN channel), the optimization procedure presented in Sec. 4.2 is used to determine the degree polynomials $\lambda_j^{(k)}(x)$, $j, k = 1, 2$ together with the rates of the codes R_1 and R_2 respectively. In choosing $\rho(x)$, we have used $d_{cmax} = 20$ (usually, a high d_{cmax} tend to degrade the performance of an LDPC code [81]) and $\rho \in [0.3, 0.7]$ (we have observed that the choice of the values of ρ is not critical to the performance of the code). Given $\lambda_j^{(1)}(x)$ and $\lambda_j^{(2)}(x)$, a $(n - k_j) \times n$ parity-check matrix $\mathbf{H}_j$ for the code for the source X_i is generated by using a random inter-leaver of length n, where $k_j = nR_j$, $j = 1, 2$. Gaussian elimination is used to convert a parity-check matrix into systematic form, from which the generator matrix of the code is obtained. With a decoding error probability threshold P_d of the order of $10^{-4} - 10^{-5}$, the joint decoder typically converged in less than 100 iterations.

In order to demonstrate the UEP properties of the DJSC codes, we present in Table 4.1 several examples of optimized degree profiles for UEP-LDPC codes for rate-parameter

Table 4.2: Comparison of DJSC and DSC+C codes in terms of codeword lengths (in bits) required to achieve a decoding error probability of 10^{-5} for different values of the correlation parameter p, over BSCs with error rate $q = 0.01$.

p	$R_{X_1} = R_{X_2}$ (channel-uses/bit)	DJSC-JD	DJSC-SD	DSC+C
0.0001	0.78	5000	6910	12770
	1.05	3020	4380	8210
0.001	0.79	5000	6830	12740
	1.05	3130	4530	8420
0.1	1.05	5000	6370	12750

$a = 0.2$ and BSCs. In obtaining these results, we have let the channel error probabilities of BSCs $q = 10p$ so that the capacities of channels through which two types of bits pass through are significantly different. In this table, the code-rates R_1 and R_2 of the final design [calculated by plugging the final values of $\lambda_j^{(1)}(x)$ and $\lambda_j^{(2)}(x)$, $j = 1, 2$ into (4.14)], and the upper-bounds r_1 and r_2 given by (4.5) and (4.6) respectively are also shown. The last two rows of Table 4.1 shows the error probabilities of *type 1* and *type 2* bits in each code design, which confirms the UEP property, *i.e.* both types of bits have nearly the same error probability, despite the fact the *type 1* bits pass through a channel whose error rate is 10 times higher.

Next, we investigate the performance of DJSC code designs, by comparing them against the achievable JSC rate-bounds given by (4.8)-(4.10). We have shown that the rate-bound can be approached by increasing the rates r_1 and r_2 [given by (4.5) and (4.6)] of the constituent channel codes. However, with finite-length LDPC codes, maintaining a given level of decoding error probability at higher code-rates also requires an accompanying increase in block length. Therefore, it is of interest to compare the rates achievable with practical DJSC codes having a given block length n and a decoding error probability. Figs. 4.5 and 4.6 show the JSC rate-pairs achieved by jointly-decodable DJSC (DJSC-JD) code designs for inter-source correlation parameter values $p = 0.002$ (high correlation) and $p = 0.1$ (low correlation) over BSCs and Bi-AWGN channels respectively, at a decoding error probability

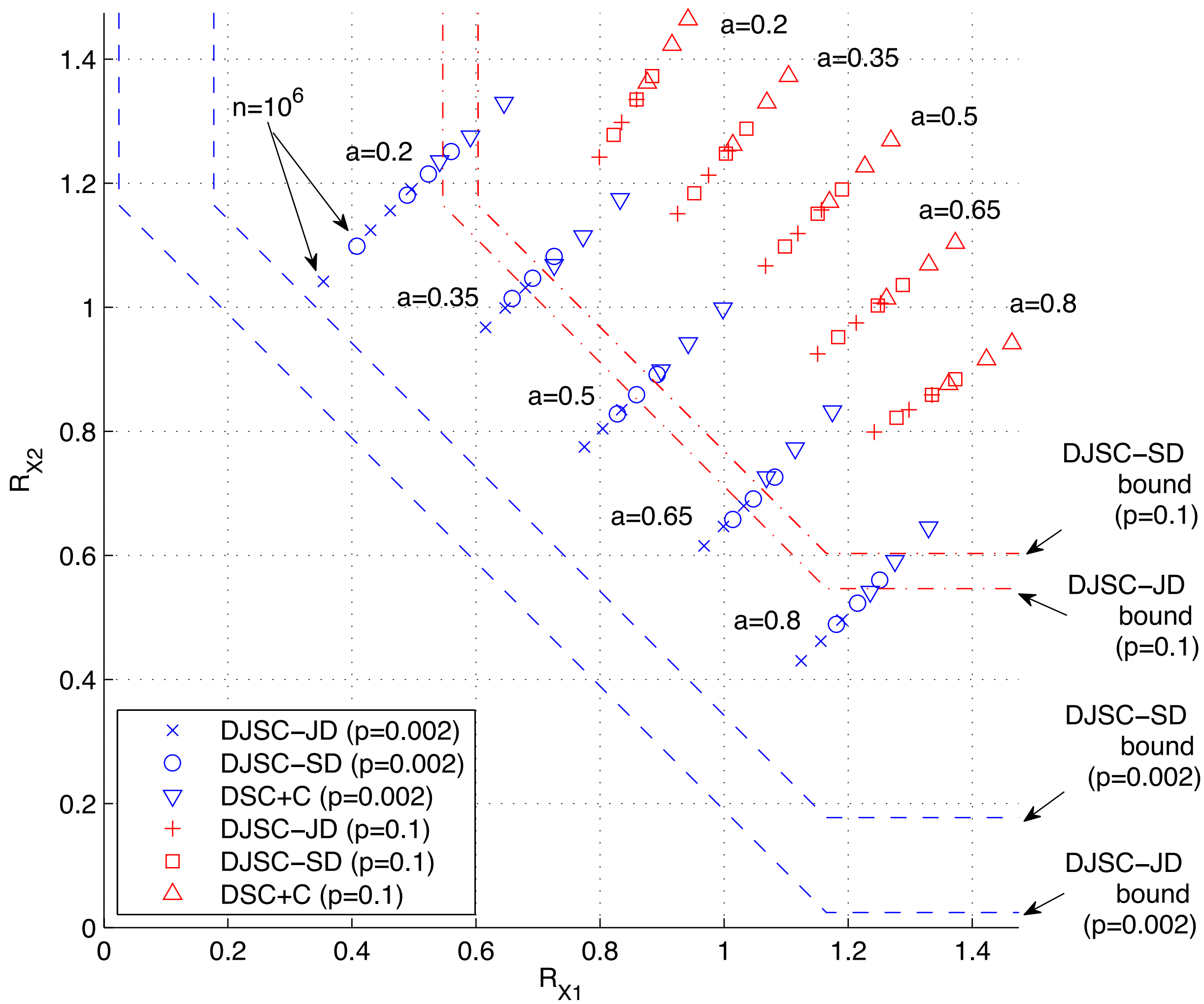

Figure 4.5: JSC rate-pairs (channel-uses/source-bit) achieved by proposed joint code design/optimization (DJSC-JD) compared with separate code design/optimization (DJSC-SD) and tandem DSC/channel coding (DSC+C coding) at different codeword lengths, over BSCs with $q = 0.01$ ($C_1 = C_2 = 0.919$ bits). p is the source-correlation parameter. Except when $(a = 0.2, p = 0.002)$ the codeword lengths are $n = 1000$ bits (furthest from the bound), $n = 2000$ bits, and $n = 5000$ bits (closest to the bound) for all three coding schemes. For $(a = 0.2, p = 0.002)$, the rate-pairs achieved for DJSC-JD and DJSC-SD are also shown for $n = 10^6$ bits.

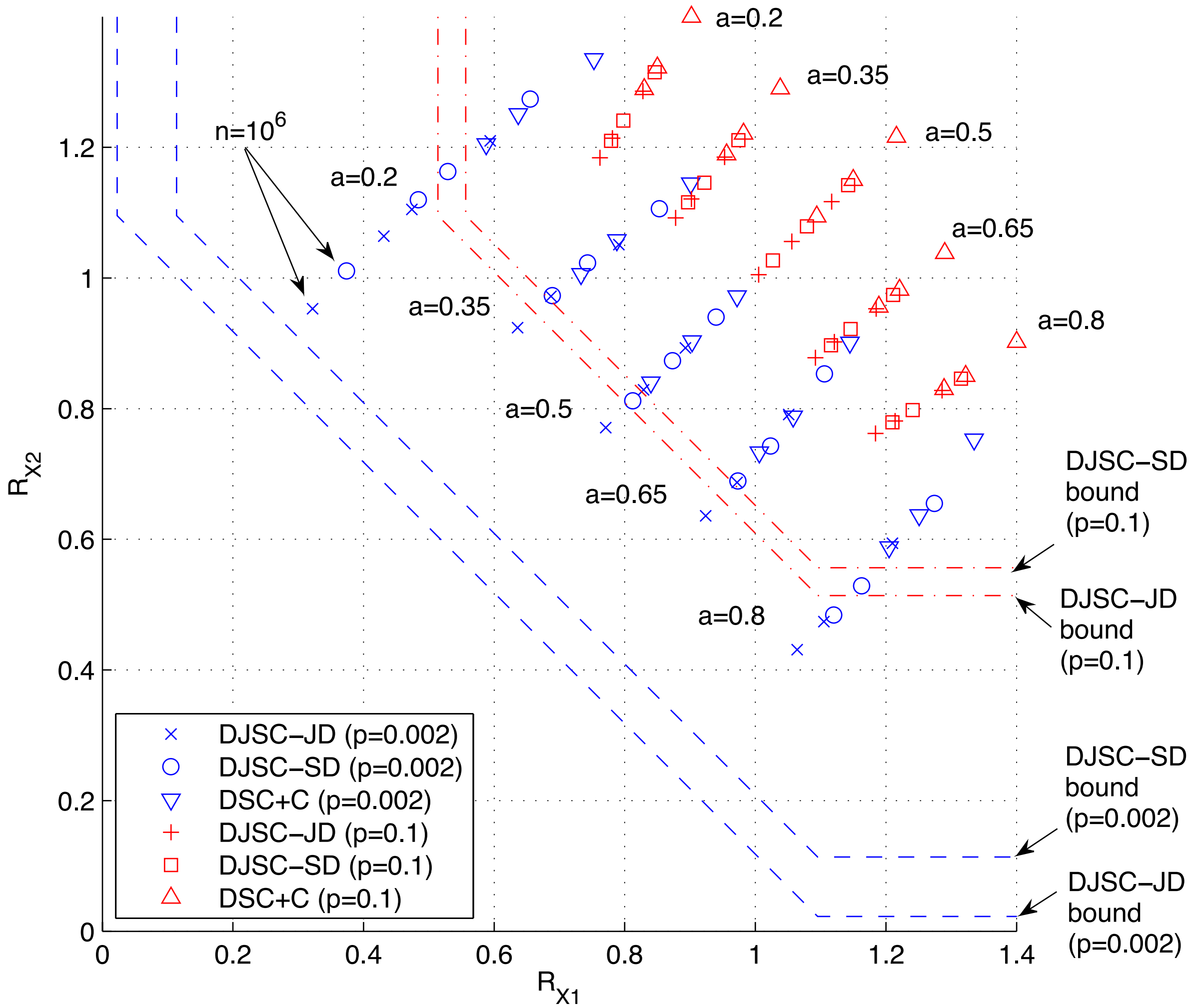

Figure 4.6: JSC rate-pairs (channel-uses/source-bit) achieved by proposed joint code design/optimization (DJSC-JD) compared with separate code design/optimization (DJSC-SD) and tandem DSC/channel coding (DSC+C coding) at different codeword lengths (in bits), over Bi-AWGN channels with a CSNR of 6 dB ($C_1 = C_2 = 0.9128$ bits/channel-use). p is the source-correlation parameter. Except when $(a = 0.2, p = 0.002)$ the codeword lengths are $n = 1000$ bits (furthest from the bound), $n = 2000$ bits, and $n = 5000$ bits (closest to the bound) for all three coding schemes. For $(a = 0.2, p = 0.002)$, the rate-pairs achieved for DJSC-JD and DJSC-SD are also shown for $n = 10^6$ bits.

of 10^{-4}. These plots also show the performance of separately-decodable DJSC (DJSC-SD) codes (see Sec. 4.1) which have been designed by assuming a separate BP decoder for each UEP-LDPC code, *i.e.,* correlation-check nodes of the factor-graph in Fig. 4.3 are ignored and therefore $J^{-1}(I_{CC,j}) = 0$ and $I_{ch,2} = C''_v$, where C''_v is the capacity of the concatenation of C_2 and C_v, used in (4.24). A joint-decoder has however been used in subsequent performance evaluation. These results confirm that DJSC-JD codes indeed outperform DJSC-SD codes when the virtual channel capacity becomes significantly higher than the communication channel capacities ($p << q$), as predicted in Sec. 4.1. In order to demonstrate the benefit of JSC coding, these figures also present the rate-pairs achieved by a separate source-channel coding scheme (DSC+C), in which SW codes are used for DSC and separately designed LDPC channel codes are used for channel error protection such that the overall block length (encoding delay) is the same as in DJSC codes. Note that our DJSC code design procedure generates SW codes for $q = 0$ ($C_1 = C_2 = 1$) and pure channel codes for $p = 0.5$. Hence, the SW code used in the DSC+C code have been obtained by using the same design algorithm for the given value of p and $q = 0$, while the channel codes are designed for the given value of q and $p = 0.5$. The number of parity bits in each channel code is determined such that transmitted codeword length is the same as in DJSC coding. The results in Figs. 4.5 and 4.6 show that the proposed DJSC codes approach the theoretical lower-bound faster (as the block length is increased) than the separate DSC and channel coding. In other words, at low encoding delays, DJSC codes outperform separate source-channel codes. Furthermore, Table 4.2 compares the block-length n required to achieve an error probability of 10^{-5} using DJSC-JD, DJSC-SD, and DSC+C codes at different levels of source correlation, over BSCs with a given noise level. These results not only show the advantage of DJSC codes over DSC+C codes for short block-lengths, but also clearly show that DJSC-JD codes outperform DJSC-SD codes for $p << q$.

In order to further strengthen the claim that the proposed approach to joint optimization

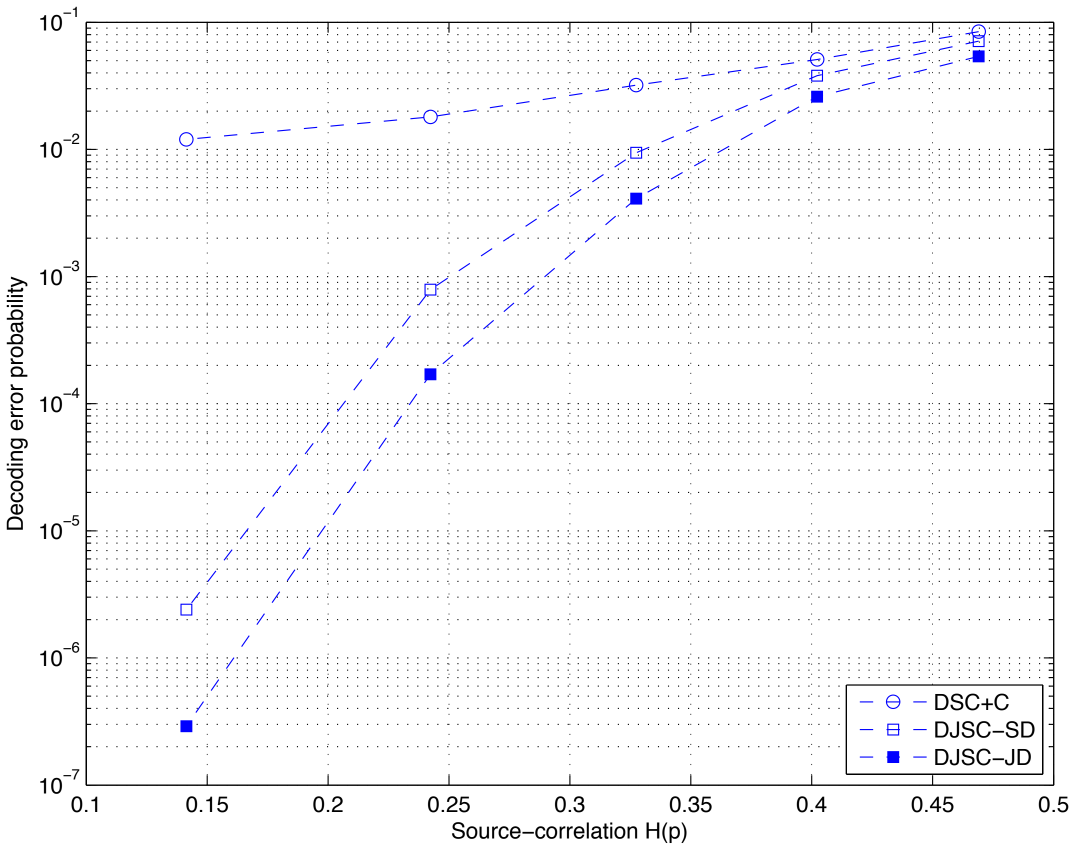

Figure 4.7: Performance comparison of DJSC coded systems with separate source-channel coded (DSC+C) systems with identical JSC coding rates $R_{X_1} = R_{X_2}$ ($a = 0.5$) at a fixed coding block-length of $n = 2000$ bits. The communication channels are BSCs with cross-over probability $q = 0.02$ and the codes have been designed for $p = 0.06$ ($H(p) = 0.327$ bits).

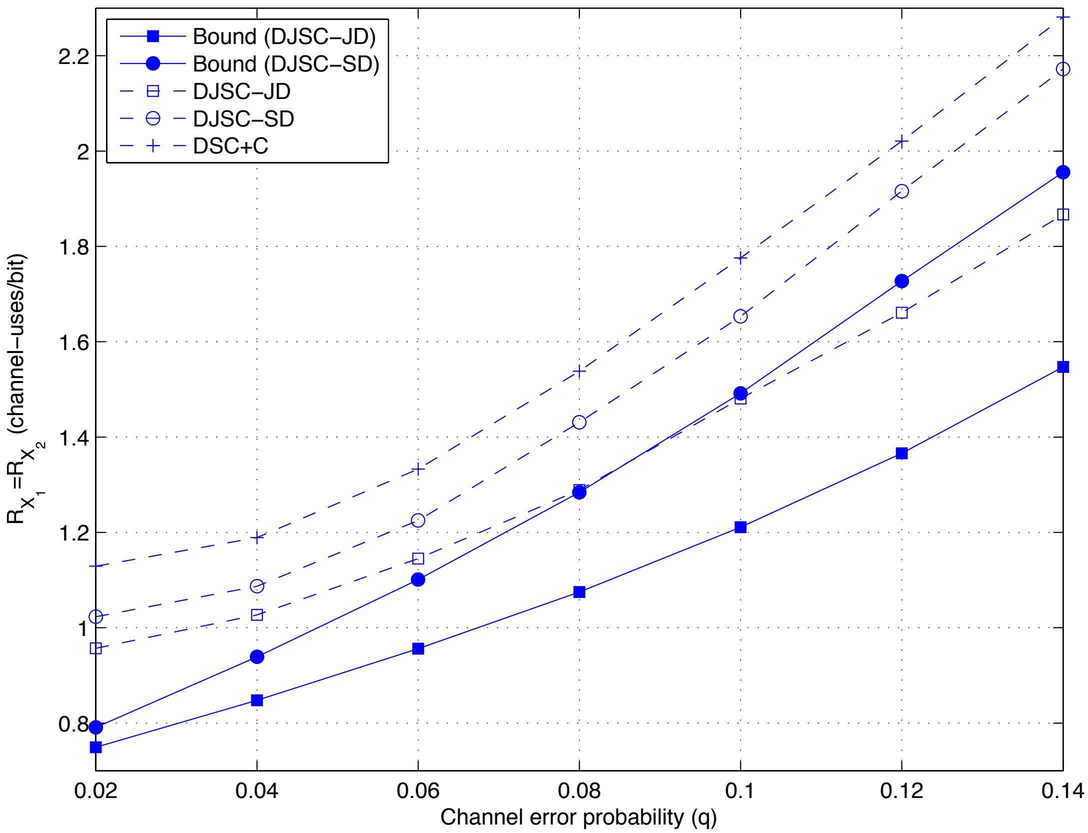

Figure 4.8: Comparison of JSC coding rates R_{X_1} ($=R_{X_2}$) of fixed-length ($n = 2000$ bits) DJSC-JD, DJSC-SD, and DSC+C codes over BSCs at different channel noise levels (source-correlation $p = 0.05$ and decoding error probability of 10^{-5}).

of the degree-distributions produce very good UEP- LDPC codes, we compare in Fig. 4.7 the decoding error probability of DJSC-JD, DJSC-SD, and DSC+C codes at a fixed codeword length of $n = 2000$ bits as a function of inter-source correlation expressed in terms of $H(p)$. On the other hand, Fig. 4.8 compares JSC coding rates achieved at different channel noise-levels q for a given level of inter-source correlation, a decoding error probability 10^{-5} and a codeword length $n = 2000$. The theoretical lower bounds shown in Fig. 4.8 have been obtained from (4.8) [or (4.9)] by using the appropriate values for C_v. Notice that at sufficiently high q, DJSC-JD codes achieve JSC rates lower than the theoretical lower-bound for DJSC-SD codes.

To our knowledge, other results on DJSC coding directly comparable to ours appear sparsely in the literature, apart from the turbo coding approaches in [2] and [3], and the IRA codes in [33]. We compare in Fig. 4.9, the performance of our UEP-LDPC codes with the results reported in [2, Fig. 8] for the case of transmission over BSCs with error probability 0.03 (capacity $C = 0.8$) at $R_{X_1} = 0.5, R_{X_2} = 1$ (corresponds to $a = 0$). In this case, the theoretical limit for error free transmission is $H(X_1|X_2) \leq R_{X_1}C = 0.4$. We decrease the correlation parameter p and design codes until the desired decoding performance is reached. It can be seen that, despite much shorter codeword length, the proposed DJSC codes designed with the UEP property are noticeably better than the turbo codes in [2, Fig. 8]. The codes in [2] are about 0.12 bits away from the theoretical bound with codeword lengths of the order of 10^5 bits, while the proposed DJSC codes based on codeword-lengths of only 5000 bits come within 0.064 bits at 10^{-7} error probability. Fig. 4.10 compares the performance of UEP-LDPC codes with turbo codes presented in [3] for Bi-AWGN channels. Note that the error floor of turbo codes due to interleaving does not occur with UEP-LDPC coding. In [33], a DJSC code design method based on IRA codes is presented for completely asymmetric rates, which corresponds to two special cases in our method *i.e.,* $a = 0$ or $a = 1$. According to the results reported in [33] length 10^5 IRA codes achieved a decoding error

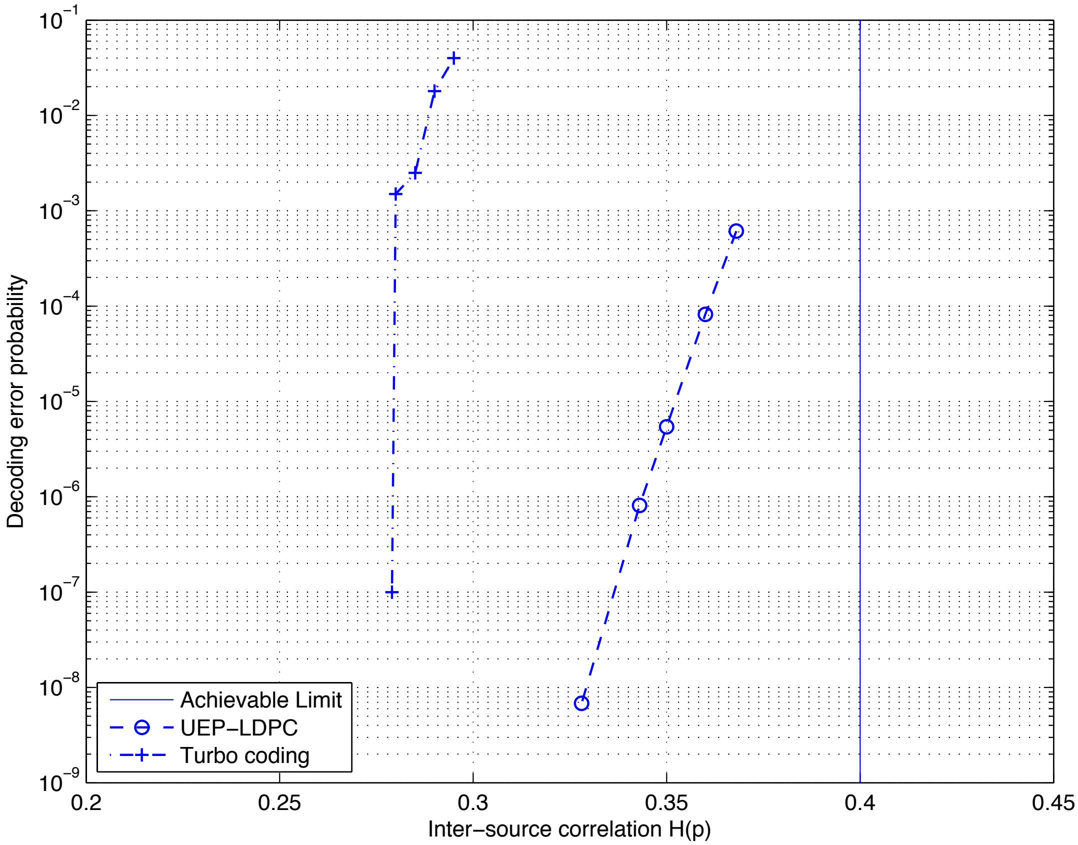

Figure 4.9: Performance comparison of DJSC codes based on UEP-LDPC codes with $n = 5000$ and those based on turbo codes from [2, Fig. 8] with $n = 10^5$, over BSCs.

probability of 10^{-6} for binary sources with correlation parameter $p = 0.066$ [$H(p) = 0.351$ bits]. We found that jointly designed UEP-LDPC codes for $a = 0$ (or $a = 1$) achieved the same performance for $p = 0.069$ [$H(p) = 0.362$ bits] for $n = 10^5$. Furthermore, for the rates $R_{X_1} = 2$ and $R_{X_2} = 1$, IRA codes for $p = 0.3$ (low source correlation) over BiAWGN channels shown in [33, Table I] come within 0.77 dB of the theoretical bound, whereas UEP-LDPC codes come within 0.75 dB. However, as mentioned above, the IRA code design in [33] applies only to corner points.

So far we have considered DJSC coding of uniform binary sources. The proposed DJSC codes are not necessarily optimal for non-uniform sources. As a result, when used with non-

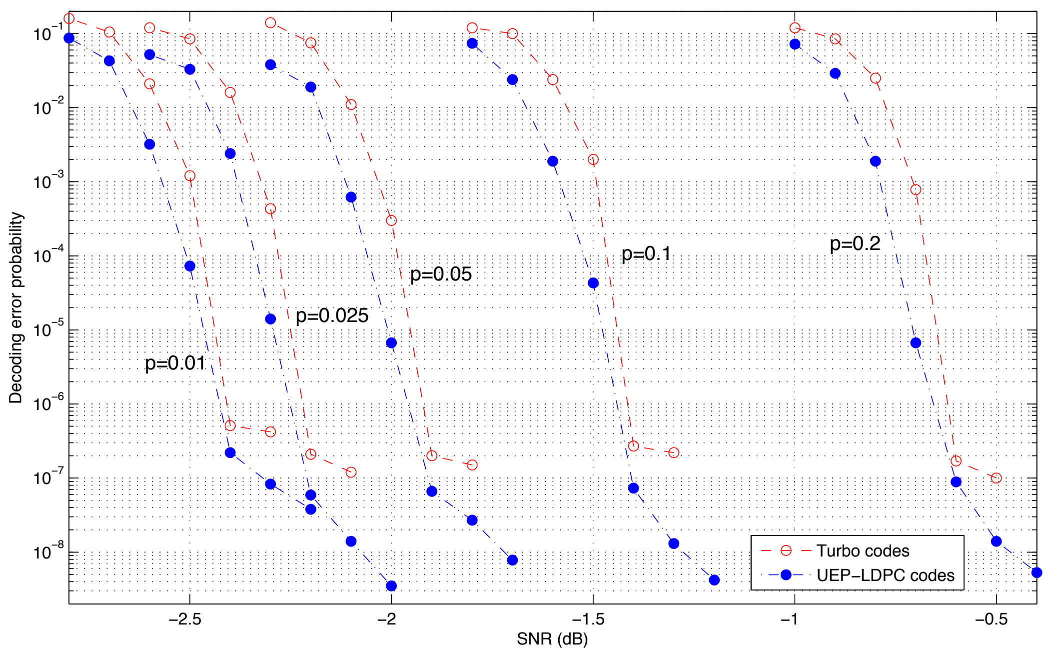

Figure 4.10: Comparison of UEP-LDPC codes and Turbo codes in [3, Figs. 5 and 6] over Bi-AWGN channels. The UEP-LDPC codes for p =0.01, 0.025, 0.05, 0.1, and $p = 0.2$ have rates $R_{X1} = R_{X2}$ ($a = 0.5$) 0.54, 0.58, 0.64, 0.73, and 0.86 respectively, and have been designed for channel SNR (in dB) of -2.7, -2.48, -2.2, -17, and -0.9 respectively. $n = 10^5$ in all cases.

uniform sources, these codes will lead to a residual redundancy in the transmitted bit streams, particularly when the codeword length is short (previous work suggests that with very long codewords, SW codes are less sensitive to marginal source distribution [82]). However, such redundancies can be effectively exploited in the joint decoder to improve the performance of short codes, by using the correct prior probabilities in (4.27). In Fig. 4.11, we consider sources with the correlation parameter $p = 0.01$ and non-uniform marginal probabilities, over BSCs. Note that, for a given value of $Pr(X_1 = 1)$ [or equivalently $H(X_1)$], the value of $Pr(X_2 = 1)$ is determined by the virtual channel relationship between X_1 and X_2. Fig. 4.11 compares the decoding error probability of the decoder which has the *exact prior information* (EPI) with that of the decoder which simply assumes *uniform prior information* (UPI). The codes have been designed for $p = 0.01$ and BSCs with $q = 0.03$. It can be seen that, as the source redundancies increase [*i.e.*, $H(X_1) < 1$], the joint decoding based on exact prior information can yield a substantial improvement in decoding error probability. Furthermore, Fig. 4.11 also shows the decoding error probability of the EPI decoder used with the same codes, but at a twice the channel error rate ($q = 0.06$), i.e., the codes are not matched to the channels. Interestingly, the EPI decoding over the mismatched channel still outperforms the UPI decoding with no channels match, except when the sources are nearly uniform.

4.4 Conclusion

An approach to designing a distributed joint source-channel code for a pair of correlated binary sources transmitted over independent noisy channels at arbitrary rates has been presented in this chapter. In this approach linear programming is used to find the optimal degree polynomials of a pair of irregular LDPC codes which can achieve both distributed compression and channel coding simultaneously via UEP of bits. As evidenced by the experimental results, these codes outperform separate source-channel codes as well as previously reported turbo code designs, when coding block length is relatively short ($1000 \sim 5000$ bits).

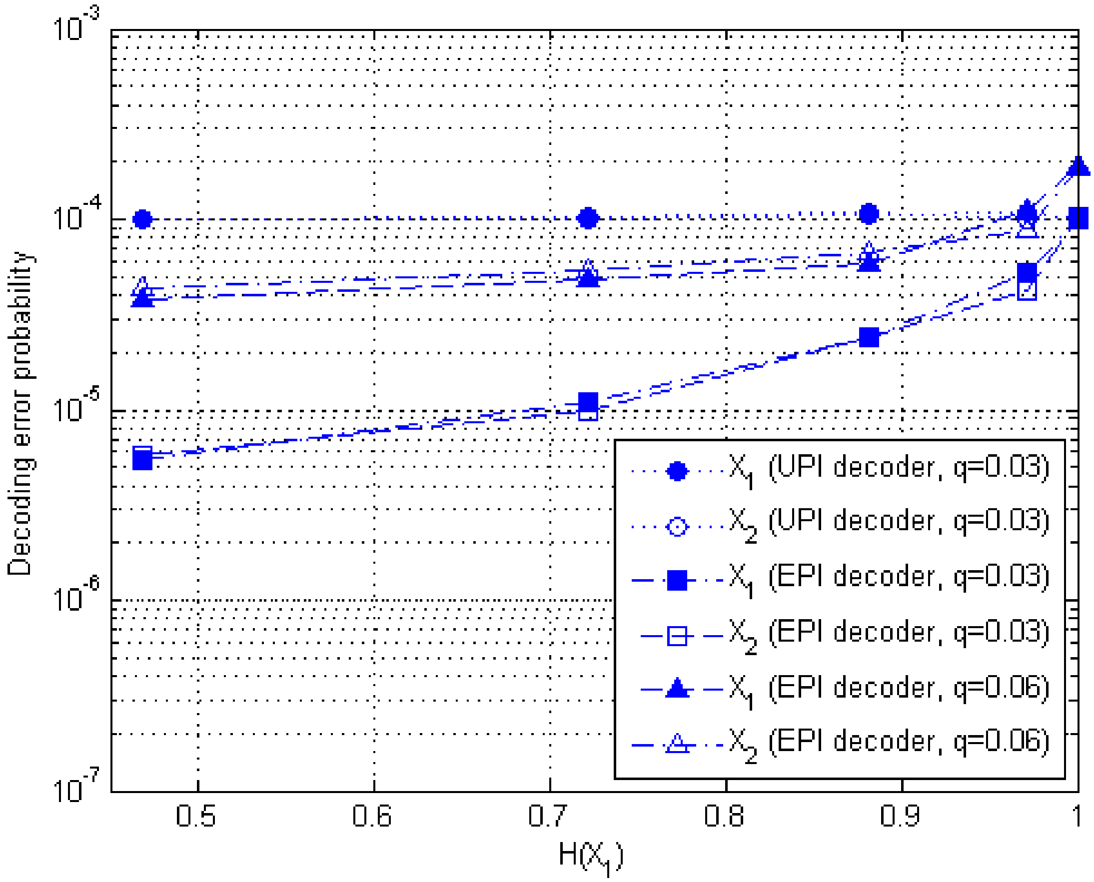

Figure 4.11: Performance of DJSC codes for sources-pair $p = 0.03$ and non-uniform marginal probabilities, as a function of $H(X_1)$. The codes have been designed to achieve a nominal decoding error probability of 10^{-4} over BSCs with $q = 0.03$ (codeword length $n = 2000$). The two curves for $q = 0.06$ show the error probabilities of the same codes decoded over BSCs with twice the error rate $q = 0.06$ using the EPI decoder.

On the other hand, these codes have been shown to approach the achievable rate bound as the block length is increased. DJSC code designs over a non-orthogonal MAC is the focus of the next chapter.

Chapter 5

Code Design for Gaussian MAC

It was established in Section 3.3 that when two dependent information sources are to be communicated over a Gaussian MAC (GMAC), DJSC coding may be used to achieve higher transmission rates than when the sources are independent. In this chapter, a new approach to designing a DJSC code for the transmission of two correlated sources over a binary-input GMAC is developed.

In general, work on the design of codes for GMAC appear sparsely in the literature. Previously, the code design for independent source and a noiseless MAC is considered in [49] where the authors use a graph splitting method for code design. The method exploits the similarity between the MAC and the binary erasure channel, and applies code design methods the already developed for erasure channels to the MAC. However, [49] also shows that the method does not work in the presence of additive noise, since the similarity with the erasure channel can no more be established. A code design procedure for independent sources is presented in [4] where the authors aim to optimize a pair of LDPC codes for BP decoding on a joint graph using EXIT analysis based on the standard Gaussian approximation [34]. However, this work does not naturally extend to correlated sources since the method used to calculate the mutual information in the EXIT analysis applies only to independent sources. This analysis is based on the assumption that one source sends a codeword in which all bits

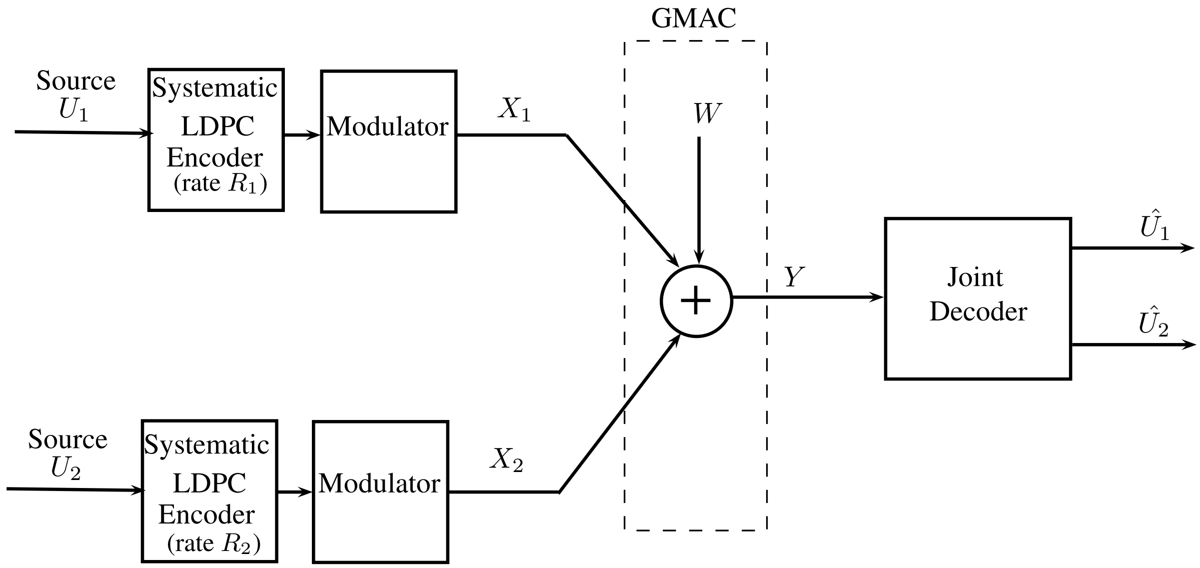

Figure 5.1: Model of the system under consideration. The modulator maps the set $\{0, 1\}$ to $\{+1, -1\}$.

are 1's while the other sends a codeword in which half the bits are 1's and other half -1's. For arbitrary source statistics, this assumption does not hold. In [3], the authors propose the use of LDGM codes over a GMAC in order to preserve the inter source correlation at the channel input. The 'low density' in the generator matrix of the codes helps to preserve the correlation between the parity bits. However, LDGM codes typically suffer from the problem of high error floor. Moreover, no method is presented for the design of the generator matrix of the code.

5.1 System Model and Problem Formulation

A block diagram of the system under consideration is shown in Fig. 5.1. The two binary sources U_1 and U_2 are encoded by channel encoders of rates R_1 and R_2 respectively to generate the channel inputs X_1 and X_2 respectively, where $X_i \in \{+1, -1\}$ for $i = 1, 2$ and the channel output be $Y \in \mathbb{R}$ is given by

$$Y = X_1 + X_2 + W, \tag{5.1}$$

and $W \sim \mathcal{N}(0, \sigma^2)$ is channel noise. Let $Pr(X_1 = (-1)^i, X_2 = (-1)^j) = p_{ij}$ for $i, j \in 0, 1$. Then, the pdf of Y is given by

$$p(y) = p_{00}q_{00}(y) + (p_{01} + p_{10})(q_{01}(y) + q_{10}(y)) + p_{11}q_{11}(y), \tag{5.2}$$

where $q_{ij}(y)$ is the conditional pdf of the output

$$\begin{aligned} q_{ij}(y) &= p(y|x_1 = (-1)^i, x_2 = (-1)^j) \\ &= \mathcal{N}((-1)^i + (-1)^j, \sigma^2) \end{aligned}$$

As in the previous chapter, the dependence between the two sources is modeled by a virtual BSC with correlation parameter $\alpha = Pr(X_1 \neq X_2)$. Note that $p_{00} = p_{11} = (1-\alpha)/2$ and $p_{01} = p_{10} = \alpha/2$.

Recall that, with independent sources the maximum achievable sum rate is the joint mutual information $I(X_1, X_2, Y)$ [5]. In the case of dependent sources, it is not known if this is an achievable upper bound. However, this still provides an upper bound for comparison. For a given value of α, the joint mutual information $I(X_1, X_2; Y)$ for the our MAC coding problem is given by

$$\begin{aligned} I(X_1, X_2; Y|\alpha) &= \sum_{X_1, X_2} \int_y p(x_1, x_2, y) \log_2 \frac{p(x_1, x_2, y)}{p(x_1, x_2)p(y)} dy \qquad (5.3) \\ &= \sum_{X_1, X_2} p(x_1, x_2) \int_y p(y|x_1, x_2) \log_2 \frac{p(y|x_1, x_2)}{p(y)} dy \\ &= \sum_{i,j} p_{ij} \int_y q_{ij}(y) \log_2 \frac{q_{ij}(y)}{p(y)}. \end{aligned}$$

The mutual information can be calculated numerically as done in Section 3.3.

Let each source be encoded by a systematic channel codes of length n bits. The rates of the codes for X_1 and X_2 can be given by $R_1 = k_1/n$ and $R_2 = k_2/n$, where k_1 and k_2 are the number of encoded information bits. The discussion in Section 3.3 suggests that for a given σ^2, the maximum sum rate $(R_1 + R_2)$ can be achieved if the bit streams X_1 and X_2 at the inputs of the GMAC exhibit a specific joint distribution as determined by the value of the correlation parameter α. Consider using a systematic channel code for each source. Then, designing an optimal DJSC for a given GMAC requires finding a pair of channel codes such that those two codes map the source bit stream with joint pmf $Pr(U_1, U_2)$ to a coded bit streams exhibit the optimal channel input distribution $Pr(X_1, X_2)$. While there appears to be no tractable procedure for achieving such a design, the main idea pursued here is to design a systematic channel code for each source so that the correlation between information bits are preserved at the channel inputs, *i.e.,* the two sources are essentially transmitted directly over the GMAC, so that $P(X_1, X_2) = P(U_1, U_2)$ for information bits (the dependence between the parity bits of the two sources cannot be controlled). The two channel codes are then decoded by a joint decoder which exploits the joint probability $Pr(U_1, U_2)$. This approach essentially exploits inter source correlation to correct channel errors caused by noise and mutual interference in the GMAC. In particular, if the two sources are independent, then each channel code require a sufficient number of parity bits to correct the errors due to noise and mutual interference between independent bit streams. However, when the two sources are dependent, the correlation between the information bits of the two codewords provide an additional joint decoding gain and hence the number of parity bits required is reduced, or equivalently, in practical channel codes a lower decoding error probability can be achieved for the same number of parity bits.

Note however that the aforementioned joint code construction can only be applied when $k_1 = k_2$, that is the rates of the two codes must be the same, i.e., $R_1 = R_2$. Furthermore, block codes based on sparse graphs such as IRA codes, LDPC codes, are usually completely specified by their edge degree distribution, and the actual number of information and parity bits play no part in the design process. For this reason, previous work [3, 4, 49] consider only symmetric rates, $R_1 = R_2$. The code design proposed in this paper also applies only to symmetric rates.

The DJSC code approach presented in this paper uses system irregular LDPC codes to encode both sources. This code construction is different to the one proposed for orthogonal channels in the Chapter 4 in several aspects. First, no UEP codes are used. The design of UEP codes in Chapter 4 requires tractable expressions for the capacity of the channels through which different bits of a UEP codes are transmitted. However, no known expression exists for the capacity MACs with correlated sources. Another key difference is the modeling of the pdf of the outgoing messages from the nodes representing joint probabilities $P(X_1, X_2)$ in the joint factor graph used for EXIT analysis. In this Chapter, it is demonstrated that, in the case of the GMAC, the outgoing messages are not well approximated by a Gaussian-mixture model

The LDPC code for X_k, $k = 1, 2$ is defined by parameters $(n, \lambda^{(k)}(x), \rho^{(k)}(x))$ where n is the code length and $\lambda^{(k)}(x)$ ($\rho^{(k)}(x)$) represents the edge-wise degree polynomial associated with the variable (party-check) nodes [35]. Specifically, we have

$$\lambda^{(k)}(x) = \sum_{i=2}^{d_{vmax}} \lambda_i^{(k)} x^{i-1},$$

where $\lambda_i^{(k)}$ is the probability that an edge chosen in a uniformly random manner is connected to a variable node of degree i (a degree of a node is defined as the number of edges connected to it.) d_{vmax} is maximum degree of variable nodes. The edge-wise degree polynomial for the parity-check nodes is similarly defined as

$$\rho^{(k)}(x) = \sum_{i=2}^{d_{cmax}} \rho_i^{(k)} x^{i-1},$$

where d_{cmax} is the maximum parity-check node degree of the code. The parameters d_{cmax} and d_{vmax} are typically chosen in a manner that the sparsity of the corresponding factor graph is maintained (*i.e.*, the edges in the factor graph grow linearly with the codeword length [35]).

5.2 Decoder Factor Graph

Consider MAP decoding of the bits of each channel code, based on the observed channel output sequence. Let $\underline{x}_1 = (x_1^{1)}, \ldots, x_1^{(n)})$ and $\underline{x}_2 = (x_2^{1)}, \ldots, x_2^{(n)})$ be the transmitted channel codes for sources X_1 and X_2 respectively, and $\underline{y}$ be the corresponding channel output. The bit-wise MAP decision rule for the i^{th} bit of X_1 is given by

$$\hat{x_1}^{(i)} = \arg\max_{x_1^{(i)} \in \{\pm 1\}} p(x_1^{(i)} | \underline{y}) \mathbb{I}\{\underline{x_1} \in \mathbb{C}_1\}, \tag{5.4}$$

where $(\hat{x}_1^{(1)}, \ldots, \hat{x}_1^{(n)})$ is the decoded output for X_1, $\mathbb{C}_1(\mathbb{C}_2)$ denotes the code (or codebook) used for encoding $X_1(X_2)$ and $\mathbb{I}\{\ .\ \}$ denotes the indicator function defined by

$$\mathbb{I}\{\text{condition}\} = \begin{cases} 1 & \text{condition} = \text{true}, \\ 0 & \text{condition} = \text{false}. \end{cases} \tag{5.5}$$

The indicator function is used to include the constraints imposed by the parity check equations. By using the law of total probability, we thus get

$$\hat{x_1}^{(i)} = \arg\max_{x_1^{(i)} \in \{\pm 1\}} \sum_{\underline{x_1} \backslash \{x_i^{(i)}\}} \sum_{\underline{x_2}} p(\underline{x_1}, \underline{x_2} | \underline{y}) \mathbb{I}\{\underline{x_1} \in \mathbb{C}_1\} \mathbb{I}\{\underline{x_2} \in \mathbb{C}_2\}, \tag{5.6}$$

where the indicator function is again used to show the relationship of $\underline{x_2}$ with its code $\mathbb{C}_2$. Using the Bayes' rule, we have

$$\begin{aligned} p(x_1^{(j)}, x_2^{(j)} | \underline{y}) &= \frac{p(\underline{y} | x_1^{(j)}, x_2^{(j)}) p(x_1^{(j)}, x_2^{(j)})}{p(\underline{y})} \\ &= \frac{p(\underline{y} | x_1^{(j)}, x_2^{(j)}) p(x_1^{(j)} | x_2^{(j)}) p(x_2^{(j)})}{p(\underline{y})}. \end{aligned}$$

Since $p(\underline{y})$ is a common factor, and sources X_1 and X_2 are uniformly distributed, (5.6) is equivalent to

$$\hat{x_1}^{(i)} = \arg\max_{x_1^{(i)} \in \{\pm 1\}} \sum_{\underline{x_1} \backslash \{x_1^{(i)}\}} \sum_{\underline{x_2}} \Big(\prod_{j=1}^{n} p(y^{(j)} | x_1^{(j)}, x_2^{(j)}) p(x_1^{(j)} | x_2^{(j)}) \Big) \mathbb{I}\{\underline{x_1} \in \mathbb{C}_1\} \mathbb{I}\{\underline{x_2} \in \mathbb{C}_2\}, \tag{5.7}$$

where the product follows from the fact that the channel is memoryless (the j^{th}symbol of $\underline{y}$ only depends on the j^{th} bits of x_1 and x_2). The MAP decoding rule for bits of X_2 can be derived in a similar way and is given by

$$\hat{x_2}^{(i)} = \arg\max_{x_2^{(i)} \in \{\pm 1\}} \sum_{\underline{x_2} \backslash \{x_2^{(i)}\}} \sum_{\underline{x_1}} \Big(\prod_{j=1}^{n} p(y^{(j)} | x_1^{(j)}, x_2^{(j)}) p(x_2^{(j)} | x_1^{(j)}) \Big) \mathbb{I}\{\underline{x_1} \in \mathbb{C}_1\} \mathbb{I}\{\underline{x_2} \in \mathbb{C}_2\}. \tag{5.8}$$

From (5.7), we see that there are three main *factors* taking part in the MAP decision rule.

1. $\phi(x_1^{(j)}, x_2^{(j)}, y) = p(y | x_1^{(j)}, x_2^{(j)}) p(x_1^{(j)} | x_2^{(j)})$,

2. $\mathbb{I}\{\underline{x_1} \in \mathbb{C}_1\}$, and

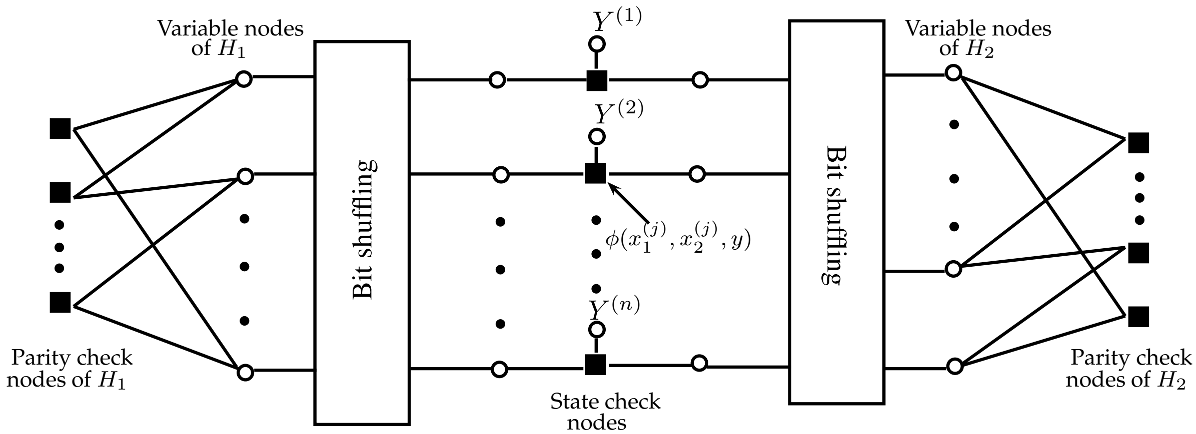

Figure 5.2: Joint decoding graph for the two LDPC codes. Nodes labeled Y represent the channel output.

3. $\mathbb{I}\{\underline{x_2} \in \mathbb{C}_2\}$.

The corresponding factor graph is shown in Fig. 5.2. As usual, the factors $\mathbb{I}\{\underline{x_1} \in \mathbb{C}_1\}$ and $\mathbb{I}\{\underline{x_2} \in \mathbb{C}_2\}$ are represented by factor nodes in the graphs of the respective parity-check matrices $\mathbf{H}_1$ and $\mathbf{H}_2$. The nodes representing the factors $p(y|x_1^{(j)}, x_2^{(j)}) \times p(x_1^{(j)}|x_2^{(j)})$ are referred to as *state-check* nodes.

The LDPC code design algorithm presented below does not ensure a systematic parity check matrix. We thus have to perform Gaussian elimination to convert the parity-check matrix into a systematic form. This results in a dense matrix which makes BP decoding very complex. Hence we perform bit shuffling on the coded bits in the decoder so that the sparse form of the matrix (originally designed) can be used for decoding.

5.3 Joint Decoder

The joint decoding of the two codes is carried out by using the BP algorithm on the factor graph shown in Fig. 5.2. The BP algorithm is an iterative message passing algorithm. In a given iteration, each node computes an outgoing messages to be passed along all edges connected to that node as function of incoming messages on all other edges. We next calculate the message maps associated with different types of nodes in the graph, *i.e.*, how an outgoing message is computed based on incoming messages. A message passed along an edge is in the LLR form. If an edge is connected to the variable node representing $x_1^{(j)}$, then the message passed along that edge is given by

$$\log \frac{Pr(x_1^{(j)} = +1|\underline{v})}{Pr(x_1^{(j)} = -1|\underline{v})}, \tag{5.9}$$

where $\underline{v}$ represents all the information about $x_1^{(j)}$ available from all edges except the edge along which the message is passed. A message along the j^{th} edge of a node is denoted by $m^{(k)}_{node \to node,j}$ where *node* is v, p or s representing variable, parity check or state check nodes respectively, and $k \in \{1, 2\}$ represents the source variable node attached to the edge. Note that each edge is attached to a variable node of exactly one source (see Fig. 5.2). For all outgoing message updates, the incoming messages on different edges are assumed to be independent [34]. The computation of outgoing messages in variable, parity-check, and state-check nodes are presented below.

5.3.1 *Variable Nodes*

An outgoing message from a variable node is an estimate of the pmf of the corresponding code bit based on all the incoming messages (which are independent estimates of the code bit's pmf). Due to this independence, the output pmf will be the product of all incoming pmf's. In LLR terms, the output message will be the sum of all incoming messages. The message from a variable node of source k with degree d_v (the number of edges emanating from it) to a parity check node of the same source on its j^{th} edge is thus given by

$$m^{(k)}_{v \to p,j} = \sum_{\substack{i=1 \\ i \neq j}}^{d_v - 1} m^{(k)}_{p \to v,i} + m^{(k)}_{s \to v}, \tag{5.10}$$

for $k \in \{1, 2\}$. Note that there is no edge index in $m^{(k)}_{s\to v}$ or $m^{(k)}_{v\to s}$ since there is exactly one edge connecting a variable node to a state check node (see Fig. 5.2). The message from a variable node to a state check node is similarly given by

$$m^{(k)}_{v\to s} = \sum_{i=1}^{d_v-1} m^{(k)}_{p\to v,i}. \tag{5.11}$$

5.3.2 Parity Check Nodes

The parity check nodes represent the parity check equations defined by the binary block code (*i.e.* the modulo 2 sum of all the bits connected to a check node must be 0.) The outgoing message update can be calculated by observing the duality between a repetition code (a code which forms a codeword by repeating the information bit a specific number of times) and a single parity check code (a code with only one parity check node connected to all variable nodes). Using the Fourier transform relationship that exists between (p, q) and $(p+q, p-q)$ where $p = Pr(x = +1|\underline{v})$ and $q = Pr(x = -1|\underline{v})$, it can be shown [83] that the outgoing message from a parity check node of degree d_c on its j^{th} edge can be given by

$$m^{(k)}_{p\to v} = 2\tanh^{-1}\left(\prod_{\substack{i=1\\ i\neq j}}^{d_c} \tanh\frac{m^{(k)}_{v\to p,i}}{2}\right) \tag{5.12}$$

for $k \in \{1, 2\}$. An alternate detailed derivation which is not based on Fourier transforms is also shown in [32].

5.3.3 State-Check Nodes

The messages passed between variable nodes and a state check node is shown in Fig. 5.3. We see that an outgoing message from a state check node $m^{(1)}_{s\to v}$ is a function of $m^{(2)}_{v\to s}$ and y. Making use of the distribution $p(y)$ given in (5.2) and the joint pmf p_{ij}, it can be shown

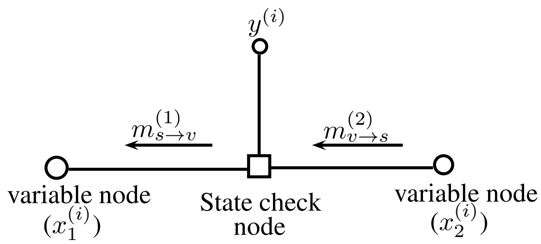

Figure 5.3: Message flow through the state check node.

that

$$m^{(1)}_{s\to v} = \log \frac{q_{00}\beta^{(1)}_{00} e^{m^{(2)}_{v\to s}} + q_{01}\beta^{(1)}_{01}}{q_{10}\beta^{(1)}_{10} e^{m^{(2)}_{v\to s}} + q_{11}\beta^{(1)}_{11}}, \tag{5.13}$$

where $\beta^{(1)}_{ij} = Pr(X_1 = (-1)^i | X_2 = (-1)^j)$. Similarly, we have,

$$m^{(2)}_{s\to v} = \log \frac{q_{00}\beta^{(2)}_{00} e^{m^{(1)}_{v\to s}} + q_{10}\beta^{(2)}_{10}}{q_{01}\beta^{(1)}_{10} e^{m^{(1)}_{v\to s}} + q_{11}\beta^{(2)}_{11}}, \tag{5.14}$$

where $\beta^{(2)}_{ij} = Pr(X_2 = (-1)^j | X_1 = (-1)^i)$. The detailed derivation is given in Appendix 6.2. Note that the parity bits of each source are independent due to the dense nature of the generator matrices in LDPC codes. Therefore $\beta^{(1)}_{ij}$ and $\beta^{(2)}_{ij}$ for a state check node which is connected to the parity bits of two sources are calculated accordingly.

5.4 EXIT Analysis For Two-user Code Design

In this section, we develop the EXIT analysis for joint design of LDPC codes two dependent users (sources), by exploiting the decoder structure and the message maps defined in Sec. 5.3. The EXIT analysis based code design procedure tracks the evolution of mutual information between a message (*i.e.* LLRs) passed along an edge and the code-bit represented by the variable node attached to that edge, during decoding iterations of the BP algorithm. For successful decoding of the code being optimized, we need this mutual information to increase after each iteration. We therefore need to calculate how this mutual information evolves as a function of the number of decoding iterations.

The computation of mutual information requires that the incoming and outgoing LLRs of a node be Gaussian random variables with variance equal to twice the mean [34]. The information updates through variable nodes and parity check nodes is well known [35]. Specifically, the information update through a variable nodes stems from the message update through that variable node and the central limit theorem. An outgoing message from a node has the mean equal to the sum of incoming means and the outgoing message is Gaussian given a reasonably high node degree. The information update for a parity-check node relies on a deterministic relationship with the variable nodes connected to it, as defined by the parity check equations. The information update for a parity-check node can therefore be calculated by simply using the duality relationship that its has with a variable node [75]. The aforementioned Gaussian assumption is exact for communication of a single binary source through a Bi-AWGN channel. For other types of channels, the assumption is approximate but is known to yield good results due to the universality of LDPC codes (*i.e.* code designed for one type of channel can perform well on a channel of different type) [84]. It should also be noted that analytical computation of aforementioned mutual information is possible if the pdf of the outgoing LLRs from the nodes can be modeled as Gaussian mixtures [34].

Since $\mathbf{H}_1$ and $\mathbf{H}_2$ correspond to two LDPC codes, the mutual information for outgoing messages from variable and parity check nodes in the joint factor-graph of the two LDPC codes can be computed in a similar manner to single-user LDPC codes. These are summarized in Sec. 5.4.1 and 5.4.2 for completeness. In the next section, a new approach is proposed for information updates through state-check nodes in two-user case. As no duality relation exists between state-check nodes and variable nodes, the aforementioned deterministic approach cannot be used to compute the mean and variance of output LLRs. This is because, the messages through the state-check nodes depend on the source statistics and not on the code structure, which makes it difficult to extend the standard EXIT analysis-based LDPC code design procedure as in [35] to jointly optimize LDPC codes for GMAC. We therefore develop a different method to approximate the outgoing message from the state-check node in Sec 5.5.

Before we proceed, let us define the following notation to simplify the presentation. The average extrinsic (*a prior*) information passed along an edge of a node with degree *deg* is defined as $I^{(k)}_{node\rightarrow node,deg}$, where $k \in \{1,2\}$ represents the source and *node* can be v, p or s representing variable, parity check or state check nodes respectively (for state check nodes, the subscript *deg* is not required since the degree is always 2). The average information is represented in the same way, with the subscript for the degree removed. The relationship between the mean μ of the message density and the extrinsic information I is given by

$$I = J(\mu), \tag{5.15}$$

where the function $J(\cdot)$ is given in [72]. While $J(\cdot)$ and $J^{-1}(\cdot)$ cannot be calculated in close-form, approximations can be used as in [75].

5.4.1 Variable Node Information Update

Let the mean of a message be denoted by $\mu^{(k)}_{node\rightarrow node,deg}$. Since the mean of the sum of independent random variables is the sum of their means, by (5.10), the mean of an outgoing message on an edge from a degree j variable node to a parity check node can be given by

$$\mu^{(k)}_{v\rightarrow p,j} = (j-2)\mu^{(k)}_{p\rightarrow v} + \mu^{(k)}_{s\rightarrow v}, \tag{5.16}$$

for $k \in \{1,2\}$. Making use of the function $J(\cdot)$ given by (5.15) in (5.16), we get for the average information update,

$$I^{(k)}_{v\rightarrow p,j} = J\Big((j-2)J^{-1}\big(I^{(k)}_{p\rightarrow v}\big) + J^{-1}\big(I^{(k)}_{s\rightarrow v}\big)\Big). \tag{5.17}$$

By averaging over the variable node degree distribution $\lambda^{(k)}(x)$, we find the average mutual information from a variable node to a parity check node (given by $I^{(k)}_{v\to p}$) as

$$\begin{aligned} I^{(k)}_{v\to p} &= \sum_{j=2}^{d_{vmax}} \lambda^{(k)}_j I^{(k)}_{v\to p,j} \\ &= \sum_{j=2}^{d_{vmax}} \lambda^{(k)}_j J\Big((j-2)J^{-1}\big(I^{(k)}_{p\to v}\big) + J^{-1}\big(I^{(k)}_{s\to v}\big)\Big). \end{aligned} \tag{5.18}$$

Similarly, from (5.11), the average information transferred from a variable node to a state check node is given by

$$I^{(k)}_{v\to s} = \sum_{j=2}^{d_{vmax}} \lambda_j J\Big((j-1)J^{-1}(I^{(k)}_{p\to v,i})\Big). \tag{5.19}$$

5.4.2 Parity Check Node Information Update

From (5.12), we see that the mean of the outgoing message of a degree d_c check-node is given by

$$\mathbb{E}\Big\{\tanh\frac{m^{(k)}_{p\to v}}{2}\Big\} = \prod_{\substack{i=1\\ i\neq j}}^{d_c} \mathbb{E}\Big\{\tanh\frac{m^{(k)}_{v\to p,i}}{2}\Big\}, \tag{5.20}$$

where the expectation is taken with respect to the density of $m^{(k)}_{p\to v}$. While computing the expectation of the above equation is not analytically feasible, two different approximations have been used in the literature. First is based on finding an approximation for the expectation (5.20) (*e.g.,* in [79]), while the other is based on the duality approximation [80], [32, pp. 236]. In our experiments, we have observed that the codes designed through both approaches are similar in terms of their degree distributions and BER performance. We use the latter approach as given the message updates for the parity check nodes , we can simply exploit the duality between a parity check node and a variable node by taking the Fourier transform. In the case of extrinsic information, we note that the EXIT function i of a code $\mathbb{C}$ is related to the EXIT function $i^\perp$ of its dual code $\mathbb{C}^\perp$ by [32, pp. 236],

$$i^\perp = 1 - i. \tag{5.21}$$

In other words, for the information update over a parity check node with degree j, we first compute the dual parameter $1 - I$, then the corresponding mean by computing $J^{-1}(1 - I)$. Since the means at a variable (the dual of the parity-check code) add, multiply the result by $j - 1$. Finally, bring the resulting mean back to mutual information by applying $J(\dot{)}$ and compute the dual parameter by using 5.21. Therefore, for a degree j parity check node, we get

$$I^{(k)}_{p\to v,j} = 1 - J\Big((j-1)J^{-1}\big(1 - I^{(k)}_{v\to p}\big)\Big). \tag{5.22}$$

By averaging over the check-node degree distribution, we get the average information as

$$\begin{aligned} I^{(k)}_{p\to v} &= \sum_{j=2}^{d_{cmax}} \rho_j^{(k)} I^{(k)}_{p\to v,j} \\ &= 1 - \sum_{j=2}^{d_{cmax}} \rho_j^{(k)} J\Big((j-1)J^{-1}\big(1 - I^{(k)}_{v\to p}\big)\Big). \end{aligned} \tag{5.23}$$

5.5 Information Update In State-Check Nodes

Recall that, EXIT analysis requires that outgoing messages of nodes be approximately Gaussian with mean μ and variance 2μ [34]. However, as discussed earlier, the mean $\mu_{s\to v}$ of the outgoing messages of a state-check node cannot be computed analytically as in the case of variable and parity-check nodes. In the following we propose an effective (leads to good two-user code designs) and computationally efficient method to approximate the density of outgoing messages from state-check nodes by a Gaussian and to estimate its mean and variance.

Consider a message $m^{(2)}_{v\to s}$ passed from a variable node of X_2 to a state-check node s

as shown in Fig. 5.3 (calculation for message from X_1 to X_2 is similar). Since $m^{(2)}_{v\to s}$ is an outgoing message from a variable node it is Gaussian distributed with mean $\mu_{v\to s}$ and variance $2\mu_{v\to s}$. Hence, we can estimate the density of the outgoing message $m^{(1)}_{s\to v}$ from the state-check node using Monte-Carlo simulations. These density functions are shown in Fig. 5.4 for several values of the inter-source correlation parameter α and $\mu_{s\to v}$. It can be seen that these densities are skewed and the Gaussian approximation is not straight forward. Note also that, these densities in general can be considered bimodal for any values of α and μ.

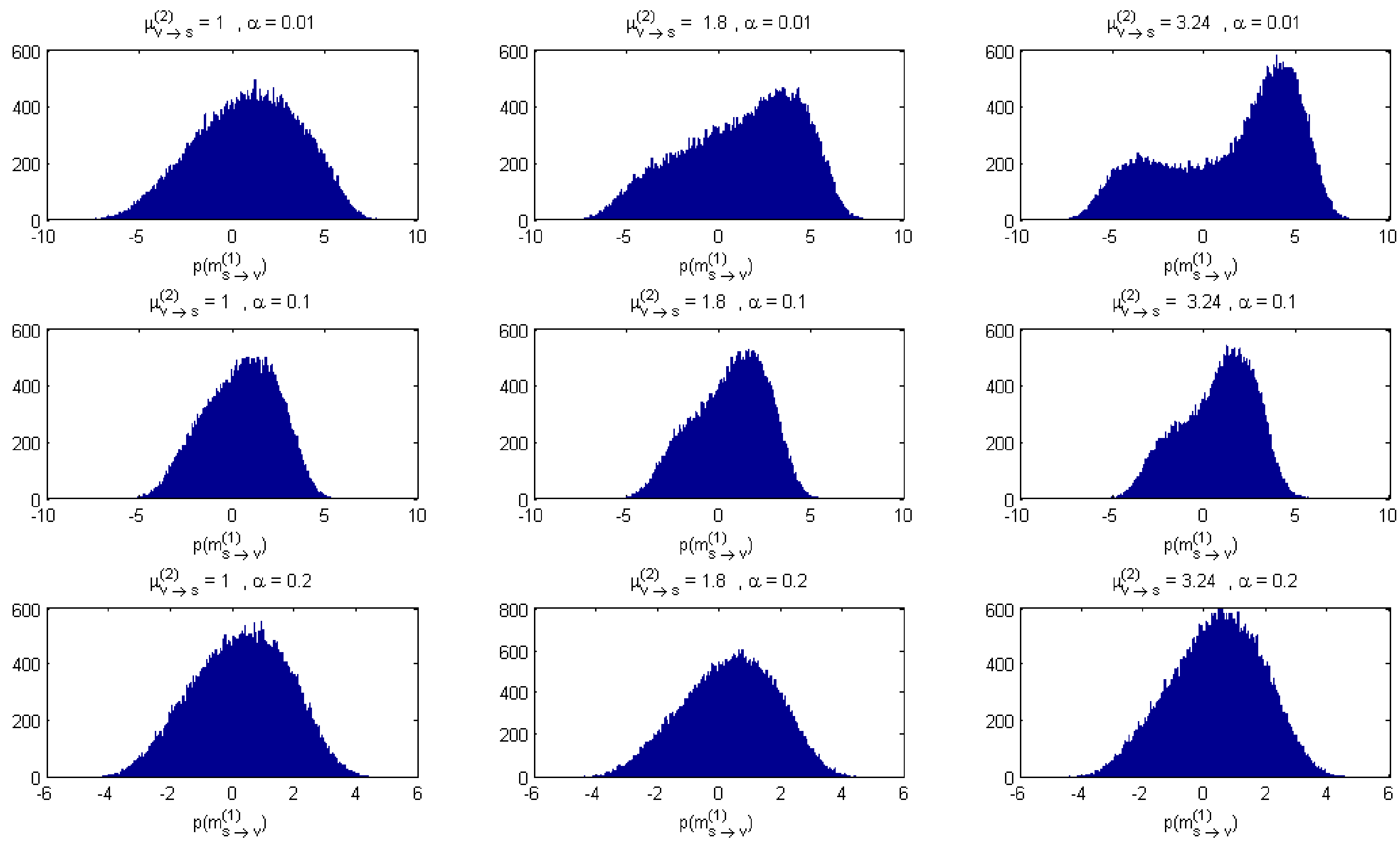

Figure 5.4: Histograms for outgoing message from a state check node for different values of $\mu^{(2)}_{v\to s}$ and α, $\sigma^2 = 5$.

5.5.1 *Gaussian Approximation of Outgoing Messages*

The problem of approximating an arbitrary distribution through a Gaussian has been well studied. The most widely used approaches are the transformation-based methods, see [85–87]. The interested reader is referred to [86] for a discussion about a family of transformations along with their merits and demerits. These methods are essentially parametric where the parameter estimation is usually done through methods such as ML or Bayesian inference. They also require performing the *inverse transform* operation once the required processing is done on the Gaussian density.

For our problem, both parameter estimation and inverse operation can make the code optimization algorithm intractable. Since our optimization algorithm is based upon mutual information transfer, we need an approximation which can model the output distribution well enough, does not entail any extra overhead for the algorithm, and leads to the maximum mutual information for the outgoing message for a given input mean. Based on these considerations we focus on the following approaches to compute the mean μ of a Gaussian pdf $\mathcal{N}(\mu, 2\mu)$ of the outgoing messages from a state-check node.

- *Mean-matched Gaussian approximation* - The mean is estimated from observations.
- *Mode-matched Gaussian approximation*- The mode m of the pdf is estimated from observations and we set $\mu = m$. The fitting of a Gaussian distribution at the mode of an arbitrary distribution is a well known method in the literature known as the Laplace approximation [88]. The only difference here is that we use the variance as twice the mean as required by EXIT analysis.
- *Two-component Gaussian mixture approximation*- The density is approximated by fitting a two component Gaussian mixture $a_1\mathcal{N}(\mu_1, \sigma_1^2) + a_2\mathcal{N}(\mu_2, \sigma_2^2)$. For simplicity, $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, a_1,$ and a_2 are estimated from the frequency of occurrence in observations.

The rationale for using these approximations can be seen from Fig. 5.4. Note that for some values of α, $\mu_{v\to s}$ and σ^2, the output density displays two dominant modes (although not as pronounced as in the case of orthogonal channels shown in Fig. 4.4). For some values, the output density does resembles a Gaussian, for which the output mean and mode are equal, while for some values, the density is unimodal and does not resemble a Gaussian.

In order to compare the performance of these approaches on the basis of mutual information, we compute the output mutual information $I^{(1)}_{s\to v} = I(m^{(1)}_{s\to v}; x^{(i)}_1)$ as a function of input mutual information $I^{(2)}_{v\to s} = I(m^{(1)}_{v\to s}; x^{(2)}_2)$ in Fig. 5.3. Fig. 5.5 and Fig. 5.6 show this comparison for two cases selected from Fig. 5.4. In Fig. 5.5 we compare the mean-matched and mode-matched approximations for a case where the output density is unimodal and using a mixture based method does not make sense here. In Fig. 5.6, we select a case where the output histogram does show a bimodal density and hence we include a case of a two-component Gaussian mixture approximation. In both cases we see that the highest output mutual information is achieved by fitting a Gaussian on the actual mode of the output density. The mode-matched method was found to yield the maximum output mutual information for various other values of α, $\mu^{(2)}_{v\to s}$ and σ^2 as well. Therefore, the mode-matched Gaussian approximation has been adopted for code design in this chapter. As will be shown (see Fig. 5.10), the joint codes designed by using this approximation also yield the lowest decoding bit-error probability, compared to the other two approaches.

5.5.2 Computation of Mutual Information

In this section, we propose a method to compute the mutual information for state-check nodes during EXIT analysis, using the mode-matched Gaussian approximation. The first step in this procedure is to compute the mean values of the outgoing messages of each state-check node. To this end consider the outgoing message from a state-check node to a variable node, as shown in Fig. 5.3).

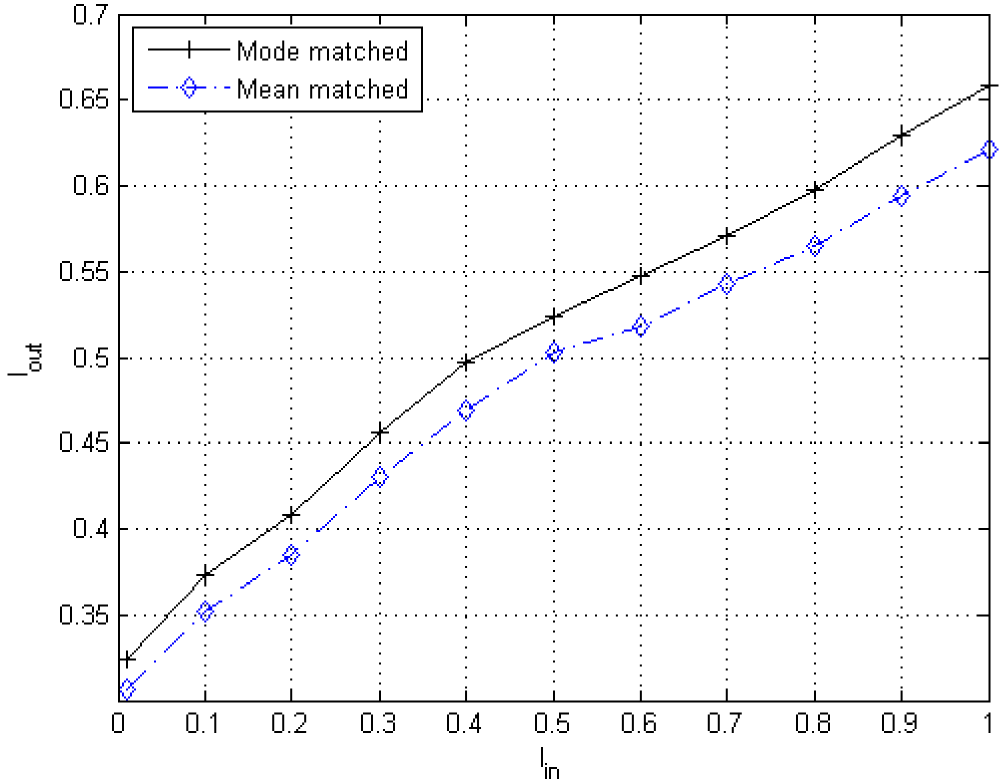

Figure 5.5: Mutual information update through the state check node for $\mu^{(2)}_{v\to s} = 1.8$ and $\alpha = 0.01$. Note in the corresponding histogram in Fig. 5.4 the unimodal (and non-Gaussian) nature of the output density.

1. Given the mean value of the (Gaussian) outgoing message $m^{(2)}_{v\to s}$ of the variable nodes $x^{(2)}_i$ (representing the i^{th} bit of source X_2), generate a sufficiently large number of N samples of $m^{(2)}_{v\to s}$ from $\mathcal{N}(\mu^{(2)}_{v\to s}, 2\mu^{(2)}_{v\to s})$.

2. Generate N sample of the channel output y_i (representing the i^{th} symbol of channel output Y) using the pdf given in (5.2).

3. Use the message update given by (5.13) to compute the corresponding N samples of the outgoing message $m^{(1)}_{s\to v}$ from the state-check node, and estimate the mode m of the pdf of $m^{(1)}_{s\to v}$. Set $\mu^{(1)}_{s\to v} = m$.

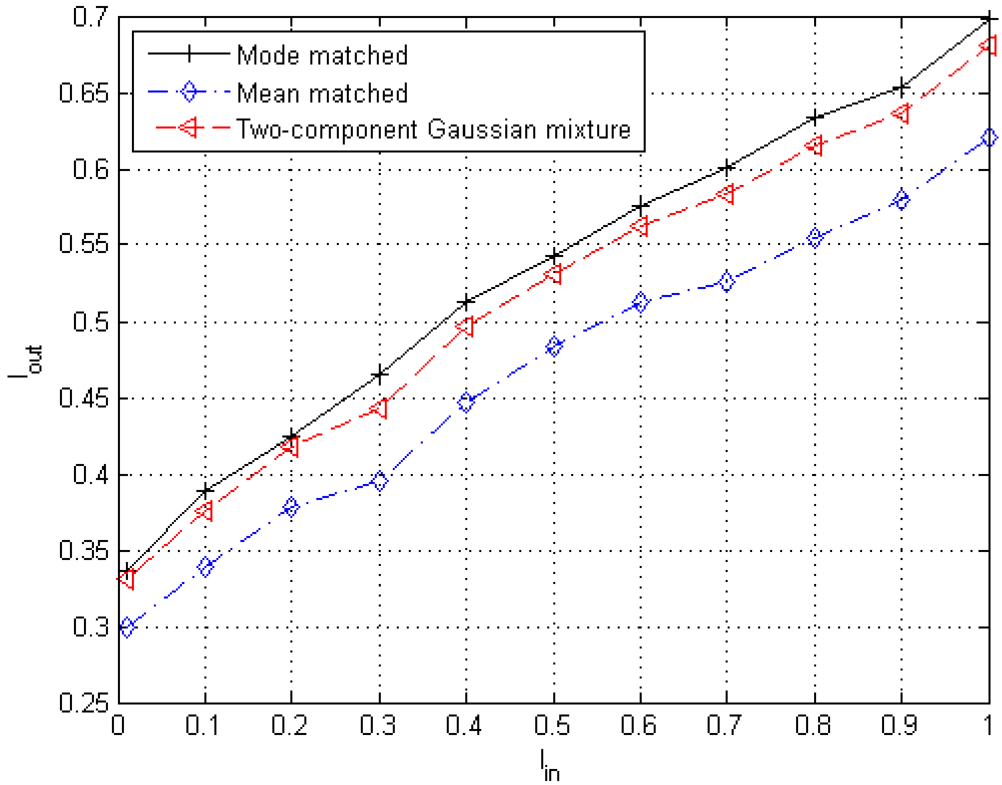

Figure 5.6: Mutual information update through the state check node for $\mu^{(2)}_{v\to s} = 3.24$ and $\alpha = 0.01$. Note in the corresponding histogram in Fig. 5.4 the bimodal nature of the output density.

Since $\mu^{(1)}_{s\to v}$ computed above is some function ψ of $\mu^{(2)}_{v\to s}$ and the channel noise variance σ^2, we can write

$$\mu^{(1)}_{s\to v} = \psi(\mu^{(2)}_{v\to s}, \sigma^2). \tag{5.24}$$

From (5.15) it follows that the mutual information update is given by

$$I^{(1)}_{s\to v} = J\Big(\psi(J^{-1}(I^{(2)}_{v\to s}), \sigma^2)\Big). \tag{5.25}$$

By plugging (5.25) into (5.18) we then obtain the overall mutual information update at the l^{th} iteration of the EXIT analysis as

$$I^{(1)}_{v\to p}(l) = \sum_{j=2}^{d_{vmax}} \lambda^{(1)}_j J\Big((j-2)J^{-1}\big(I^{(1)}_{p\to v}(l)\big) + \psi(J^{-1}(I^{(2)}_{v\to s}(l)), \sigma^2)\Big). \tag{5.26}$$

The overall mutual information update for $I^{(2)}_{v\to p}(l)$ can be obtained by interchanging the superscripts $^{(1)}$ and $^{(2)}$ in (5.26).

5.6 Code Optimization Algorithm

In this section, we summarize the two-user code optimization algorithm, which is an extension of the single-user LDPC code optimization procedure in [34] to the joint factor-graph with state-check nodes. The joint design of two LDPC codes for DJSC coding involves finding the degree distributions $\lambda^{(k)}(x)$ for $k \in \{1, 2\}$ which maximize the code rate R, or equivalently (for fixed $\rho^{(1)}(x)$ and $\rho^{(2)}(x)$) maximize,

$$\sum_i \frac{\lambda^{(1)}_i}{i} + \frac{\lambda^{(2)}_i}{i}. \tag{5.27}$$

As used for orthogonal channels, we use a concentrated degree polynomial of the form $\rho(x) = \rho x^{s-2} + (1-\rho)x^s$ for some $s \geq 2$ and $0 < \rho \leq 1$ (which is sufficient for achieving near optimal performance [79, Theorem 2]). The objective function given in (5.27) is maximized subject to the following set of constraints.

1. *Probability constraint*

   Since the coefficients of the polynomials $\lambda(x)$ represent the probabilities, we should have for $k \in \{1, 2\}$,

$$\sum_{i=2}^{d_{vmax}} \left(\lambda_i^{(k)}\right) = 1, \tag{5.28}$$

$$\lambda_i^{(k)} \in [0, 1]. \tag{5.29}$$

2. *Information constraint*

   The extrinsic information has to be non-decreasing for each iteration, so we need

$$I_{v \to p}^{(k)}(l+1) > I_{v \to p}^{(k)}(l),\ \forall l \in \{1, 2, ..., \infty\},\ k \in \{1, 2\}. \tag{5.30}$$

   where $I_{v \to p}^{(1)}(l)$ is given in (5.26).

3. *Stability constraint*

   It is known that the stability condition for the MAC under consideration is the same as that of two single user binary channels with Gaussian noise [32]. Thus, we have for $k \in \{1, 2\}$,

$$\lambda_2^{(k)} < \exp \frac{1}{2\sigma^2} / \sum_{j=2}^{d_{cmax}} (j-1)\rho_j^{(k)}. \tag{5.31}$$

The objective function (5.27) and the constraints (5.29) to (5.31) are all linear in the code parameters $\lambda_i^{(k)}$. Therefore, we can use a linear program to find the optimized code.

5.7 Simulation Results

In this section, we present simulation results obtained by designing DJSC codes for a GMAC using the code design method developed in this chapter. These results show that the proposed approach yield DJSC codes which can outperform traditional single-user codes as well previously proposed DJSC codes for GMAC.

Table 5.1: Degree profiles for joint LDPC codes generated by the design algorithm, for different values of channel noise variance σ^2. Inter-source correlation parameter $\alpha = 0.1$.

	σ^2			
	0.3	0.4	0.5	0.6
$\rho(x)$	x^9	x^8	x^7	x^6
$\lambda_1(x) = \lambda_2(x)$	$0.284x + 0.3124x^2$ $+0.0222x^4 + 0.1344x^7$ $+0.0977x^8 + 0.1277x^{19}$ $+0.08x^{99}$	$0.3012x + 0.3321x^2$ $+0.0982x^3 + 0.1322x^{10}$ $+0.1363x^{99}$	$0.3111x + 0.3544x^2$ $+0.1655x^3 + 0.0786x^{11}$ $+0.0321x^{16} + 0.0583x^{99}$	$0.4411x + 0.4235x^2$ $0.0233x^{17} + 0.0321x^{17}$ $+0.0682x^{99}$
R	0.5614	0.5226	0.4857	0.4528
$\frac{1}{2}I(X_1, X_2; Y\|\alpha = 0.5)$	0.5014	0.4731	0.4393	0.4113
$\frac{1}{2}I(X_1, X_2; Y\|\alpha = 0.1)$	0.6328	0.6017	0.5744	0.5509

Given the inter-source correlation parameter α and the channel noise variance σ^2, the goal the DJSC code design is to find two LDPC codes whose rate R is maximized. For a given pair of (α, σ^2) and ρ^s for some fixed s, the proposed code design algorithm yields coefficients of the polynomials λ_1, λ_2 which also implicitly define the actual rates of the two LDPC codes. In general, the maximum sum-rate achievable over GMAC for two dependent sources is not known. However, as discussed in Sec. 3.3 when the source are independent ($\alpha = 0.5$) the maximum sum-rate achievable by any code design is $I(X_1, X_2, Y)$. Table 5.1 presents the degree profiles and the rate R of the DJSC code designs for a pair of source with $\alpha = 0.1$ at different channel noise levels. The table also shows the maximum code rate achievable (channel capacity) with independent sources $I(X_1, X_2, Y|\alpha = 0.5)$ as well as the actual value of the joint mutual information $I(X_1, X_2; Y|\alpha = 0.1)$ as given by (5.3). These results show that DJSC codes design can actually achieve code rates higher than the channel capacity for independent sources. This clearly shows the advantage of JSC coding of dependent sources over a GMAC.

In order to further demonstrate this, we next compare there different coding schemes as follows.

1. *Scheme 1*: Regardless of the actual inter-source correlation α, we simply assume that the two sources are independent ($\alpha = 0.5$) in code design as well as in decoding. Essentially, these code at best can only achieve a channel capacity of $I(X_1, X_2; Y|\alpha =$

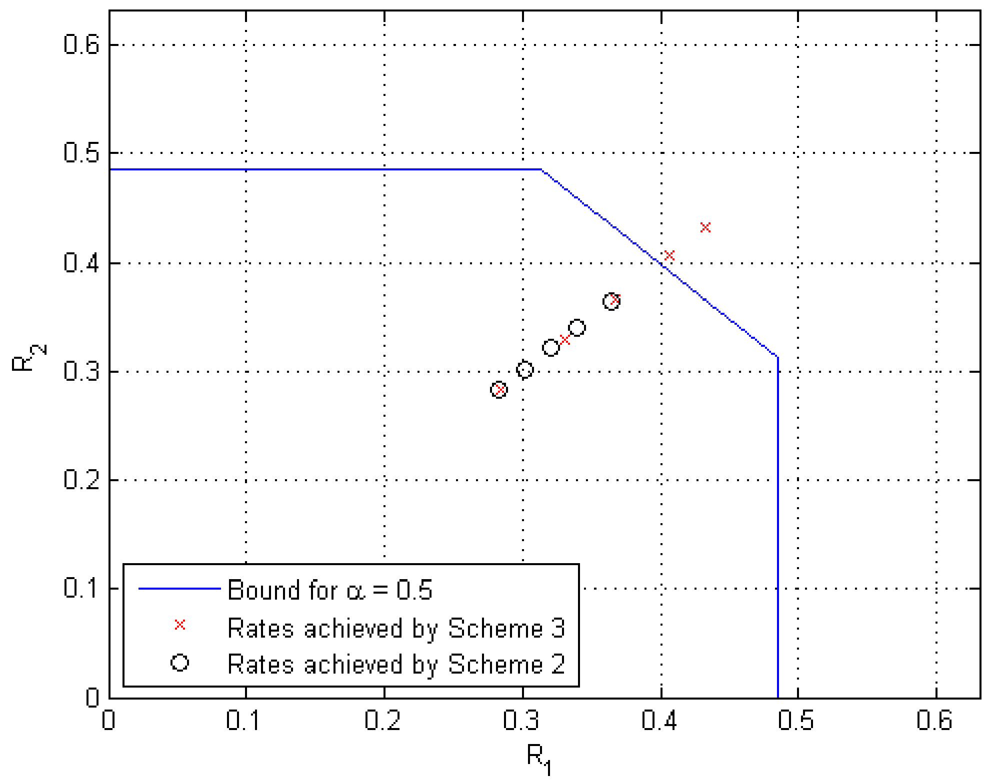

Figure 5.7: Achievable channel coding rates (in bits per channel use) achieved by Scheme 2 ($\alpha_{design} = 0.5, \alpha_{decode} = \alpha_{actual}$) and Scheme 3($\alpha_{design} = \alpha_{actual}, \alpha_{decode} = \alpha_{actual}$) at P_e of 10^{-6}. $\sigma^2 = 1$. The points correspond to $\alpha = 0.5$ (lowest rate points), $0.4, 0.3, 0.2$ and 0.1(highest rate points).

0.5). We denote the scheme by ($\alpha_{design} = 0.5, \alpha_{decode} = 0.5$).

2. *Scheme 2*: We assume independent sources for code design, but use the actual value of α in the decoder. We denote the scheme by ($\alpha_{design} = 0.5, \alpha_{decode} = \alpha_{actual}$).

3. *Scheme 3*: We use the actual value of α for both code design and in decoding. We denote the scheme by ($\alpha_{design} = \alpha_{actual}, \alpha_{decode} = \alpha_{actual}$).

Fig. 5.7 shows the achievable channel coding rates (measured in bits per channel use) for Schemes 2 and 3 for different values of correlation parameter α for $\sigma^2 = 1$ at a decoding error probability of 10^{-6}. As the correlation increases, the both schemes can achieve higher

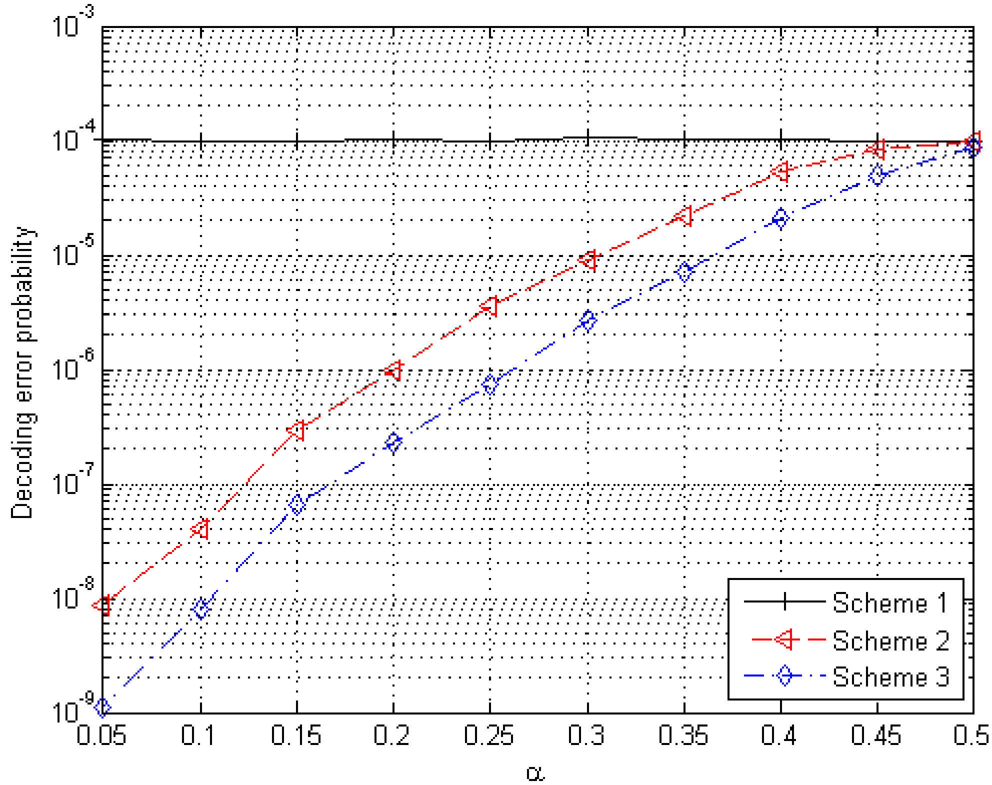

Figure 5.8: Comparison of different coding schemes for independent with correlated sources, codeword length = 10^6 bits.

rates. However note that for $\alpha = 0.2$ and 0.1, Scheme 3 can actually achieve rates which are higher than the theoretical upper bound for independent sources, which demonstrates the advantage of optimizing the two codes for the joint distribution of the two dependent sources.

Fig. 5.8 shows the bit error rate (BER) of the three schemes described above as a function of the correlation parameter α. As expected, we see that the codes optimized for the joint distribution of the sources yields the best performance. It can also be seen that, the codes designed for independent sources can achieve a significant performance improvement if the actual value of joint source probabilities are used for joint decoding. Note that with $(\alpha_{design} = 0.5, \alpha_{decode} = 0.5)$ and $(\alpha_{design} = 0.5, \alpha_{decode} = \alpha_{actual})$, the same pair of codes have been used for all values of α. We see that as correlation increases (*i.e.*, α decreases), the improvement achieved by incorporating the joint source statistics for code optimization and for joint decoding becomes more pronounced.

In Fig. 5.9), the probability of decoding error of the three schemes described above are as a function of codeword length n. At a given value of σ^2 , the JSC rates (channel uses per source bit) are kept the same for all three schemes. We see that, with the increase in block length, while the BER decreases for all three schemes but *Scheme 3* yields the best performance.

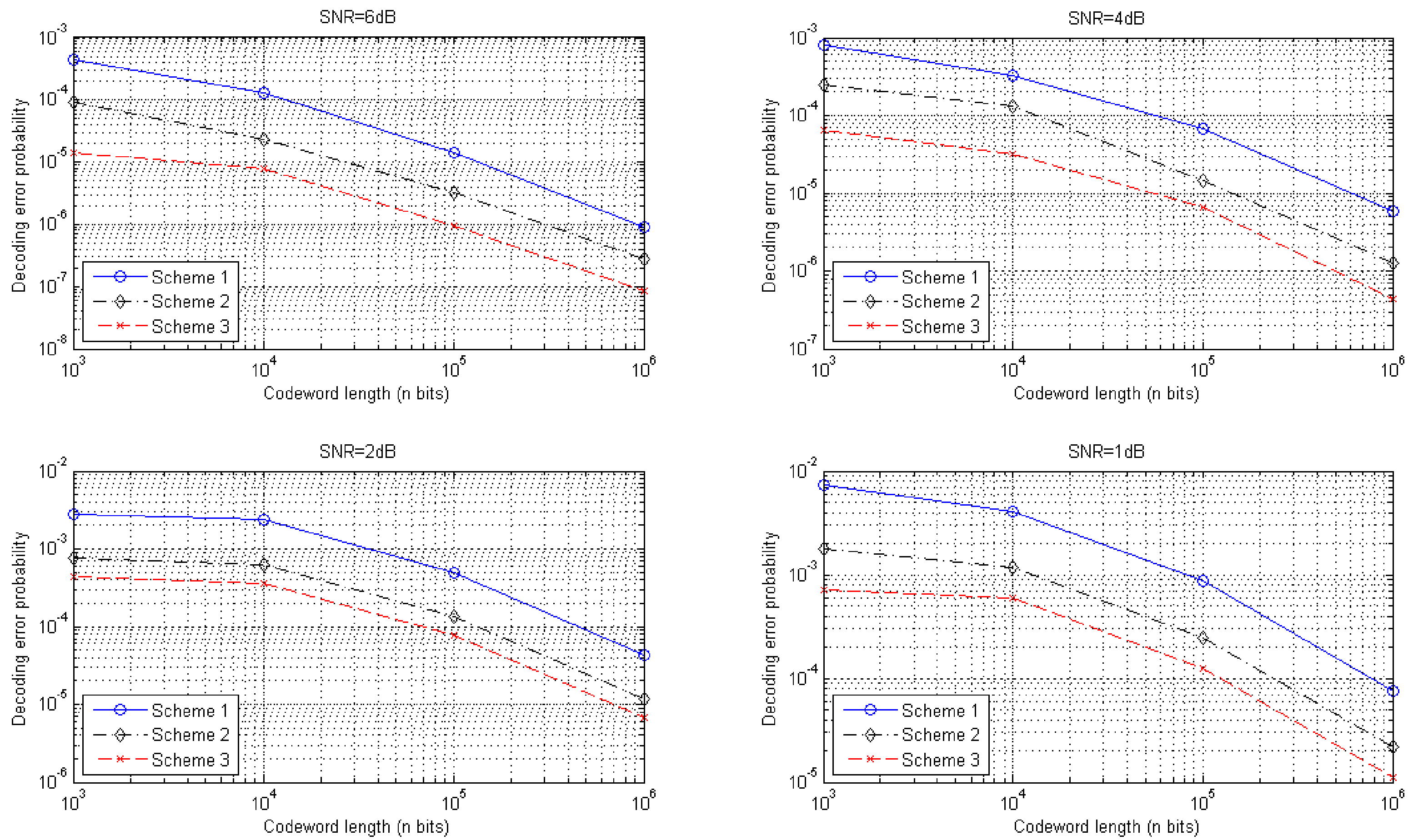

Figure 5.9: Comparison of three schemes at different codeword lengths at various channel SNR's. $\alpha = 0.2$.

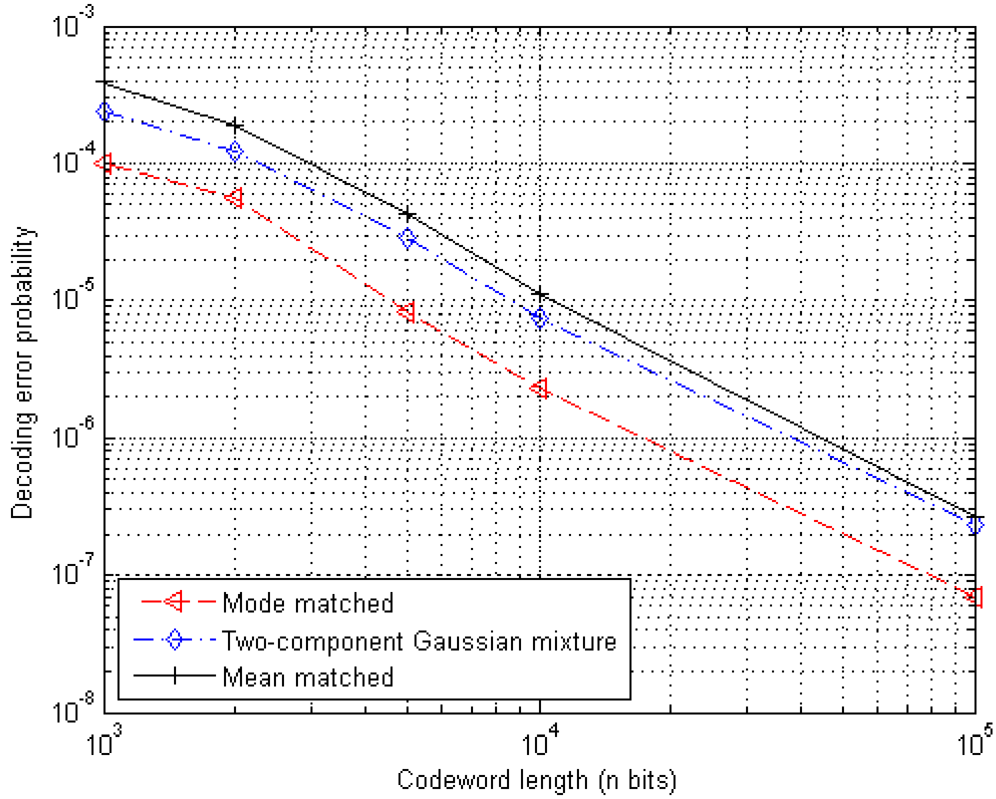

Figure 5.10: BER performance of codes designed through the three methods described in Fig. 5.6 as a function of codeword length. $\sigma^2 = 5$ and $\alpha = 0.01$

Recall that Fig. 5.5 and Fig. 5.6 indicate that codes designed by using the mode-matched approximation for the state-check node information update gives the maximum output mutual information as compared to the mean-matched or mixture-based approximations. In Fig. 5.10 we compare the BER performance for codes designed using these approximations. Here, the correlation parameter α and the channel noise variance σ^2 are kept the same as in Fig. 5.6 (for which the outgoing message density displays a bimodal behavior). We see that in terms of BER, the mode-matched approximations yields the best code designs. For example at BER of 10^{-6}, block length required for codes based on mode-matched approximation is approximately 1.7×10^4 bits, while with mean-matched approximation, we need a block length of approximately 3.8×10^4 bits.

In Fig. 5.11, we compare the JSC rates in channel uses/source bit, achieved by the

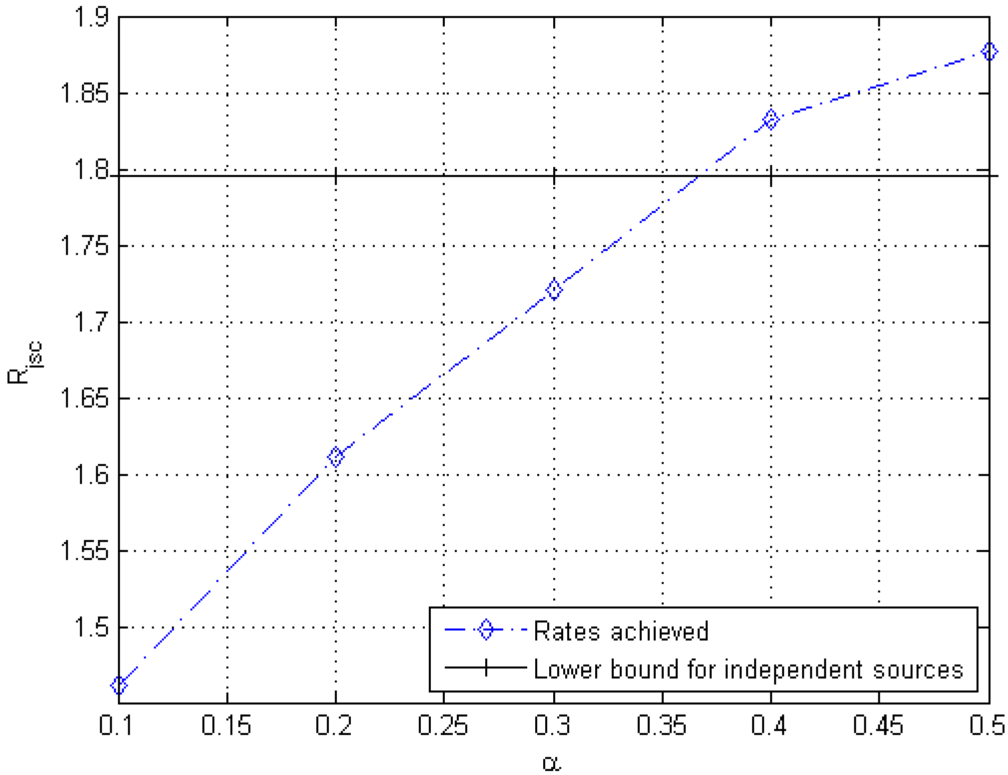

Figure 5.11: JSC rates (channel uses per source bit) as a function of the correlation parameter for $\sigma^2 = 0.5$. The theoretical limit for communication of independent sources over the same GMAC is also shown.

proposed DJSC codes as a function of the correlation parameter α at a decoding error probability of 10^{-6}. The rate upper-bound given by (5.3) for independent sources over the same channel is also shown. Note that for $\alpha \leq 0.36$, the JSC rates achieved by the DJSC codes, even beat the theoretical limit for the independent sources.

While the proposed DJSC codes are clearly aimed at coding of dependent sources, the same approach can also be used to design JSC codes for communicating independent sources over a GMAC by simply setting $\alpha = 0.5$. Previously, in [4], the authors proposed a joint code design method for communication of independent sources over a GMAC. Fig. 5.12 compares JSC code for two independent sources designed with approach presented in this chapter with the code designs in [4, Fig. 3], which shows that codes designed with the proposed approach

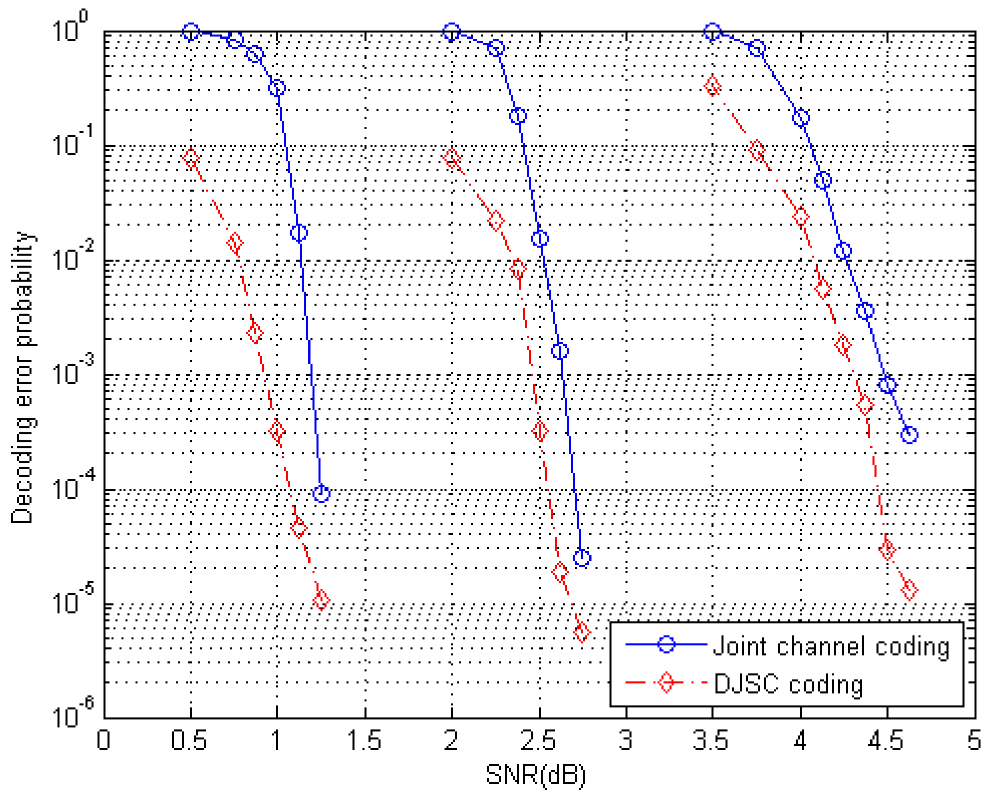

Figure 5.12: BER comparison of the proposed DJSC coding scheme with the joint channel coding scheme for independent sources, reported in [4, Fig. 3]. The rate R is measured in bits per channel use.

outperforms those in [4]. The main difference is due to the better approximation of outgoing message pdf of state-check nodes as discussed in Sec 5.5. In particular we observe a coding gain of approximately 0.2 dB at a decoding error probability of 10^{-3} for all coding rates considered here.

Next we compare performance of the DJSC codes proposed in this dissertation against those reported in [3] where the authors present a different DJSC code design method based on the use of LDGM codes. The use of LDGM codes has the advantage that the correlation among the sources is roughly preserved in both information bits and the parity bits in transmitted codewords, where as with systematic LDPC codes the parity bits transmitted for two sources are not strongly dependent, as the sparse parity check matrices leads to

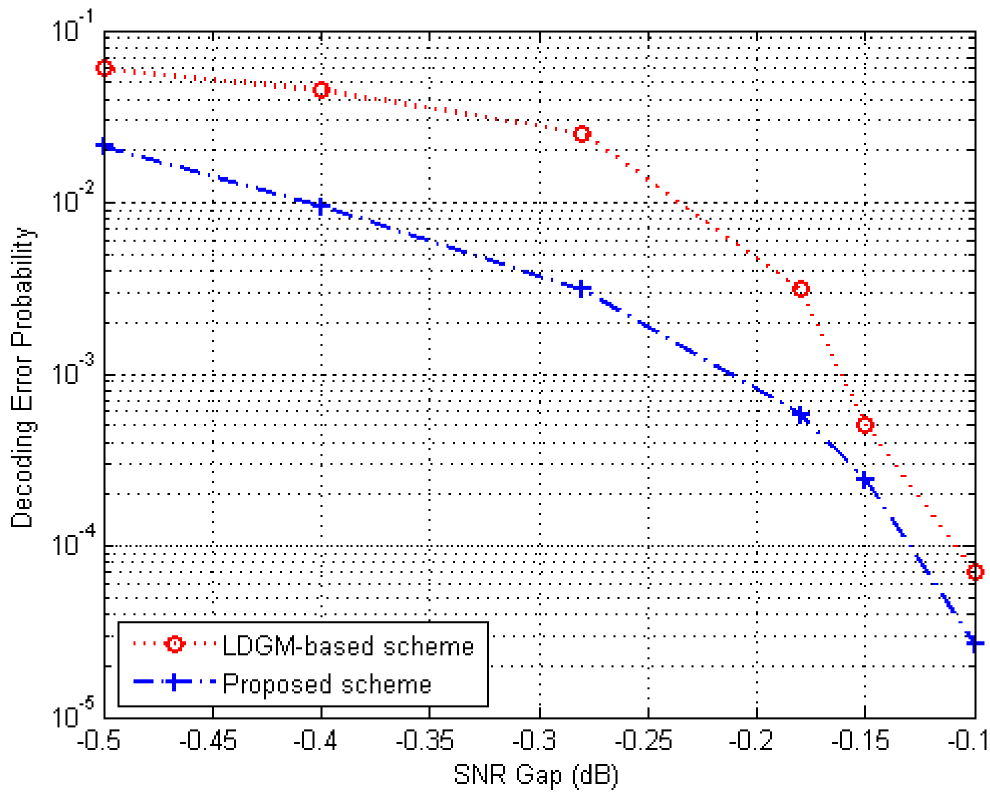

Figure 5.13: Comparison of the proposed coding scheme with previously reported scheme based on LDGM codes (Fig. 8 in [3]). ($\alpha = 0.1$)

dense generator matrices. However, LDGM codes have an inherent disadvantage that they typically show high error floors [32]. To get around this problem, the authors in [3] have used serial, parallel and hybrid concatenation of LDGM codes which removes the error floor but increases the encoding complexity due to the use of two or more codes and an interleaver. This problem is not present in the proposed LDPC codes. In Fig. 5.13 we compare the BER of DJSC codes based on LDPC codes with those reported in [3, Fig. 8]. The SNR Gap (dB) in this figure gives the gap between the SNR of the actual channel for which the code is designed and the SNR which represents the theoretical limit for independent sources for the same channel.

5.8 Conclusions

Motivated by the fact that for communicating dependent sources over a GMAC, DJSC coding can be used to achieve rates higher than those that can be achieved with separate source and channel coding. In this chapter, a new approach to designing a DJSC code with symmetric rates for a pair of correlated binary sources transmitted over a GMAC has been presented. In the proposed approach, linear programming is used to find the optimal degree polynomials of a pair of irregular LDPC codes. For codes designs for dependent sources and GMAC, a new method is identified for modeling the outgoing message densities of state-check nodes in the joint decoder factor-graph. Experimental results show that codes designed with the proposed approximation leads to codes which can outperform other known codes. The experimental results show that incorporating joint distribution of the sources in the code optimization algorithm yields improved performance as compared to a design which assumes independent sources . It has been demonstrated that for sufficiently long codeword lengths DJSC codes for dependent sources and GMAC can achieve rates higher than the theoretical rate upper bound for independent sources.

Chapter 6

Conclusions and Future Work

6.1 Summary and Conclusions

In this dissertation, the problem of designing DJSC codes for correlated sources over multi-user communication systems was investigated. The main contributions can be summarized as follows:

- A new approach to designing DJSC codes for a pair of correlated binary sources transmitted over an orthogonal MAC has been developed. Although the source-channel separation principle holds for this case, for short to moderate codeword lengths (on the order of few thousand bits), new DJSC codes were shown to significantly outperform separate source and channel coding. The proposed design approach is an extension of the well known EXIT analysis-based LDPC code optimization for single-user communication to multi-user communication of correlated sources. It was shown that the proposed DJSC codes can approach the theoretical rate bounds when the codeword length is increased.These codes have also been shown to outperform the other DJSC codes previously reported in the literature.
- It is known that the source-channel separation principle does not hold for GMAC,

and for optimal performance DJSC coding is required This problem has not been well addressed in previous work. In order to address this issue a new approach to designing DJSC codes for transmission of two correlated binary sources over a GMAC has been developed. The proposed approach generalizes the EXIT analysis-based single-user LDPC code designs to coding of two correlated sources over a GMAC. To this end, a new method has been proposed for modeling the density functions of the messages passed in BP decoding of two LDPC codes using a single factor graph. Simulation results show that new DJSC code designs can achieve rates higher than the theoretical upper bound of separate source and channel coding for sufficiently large codeword lengths. To our knowledge the only other work that considers DJSC codes design for GMAC is [4]. The simulation results presented here indicated that the proposed LDPC code designs can outperform LDGM codes in [3].

6.2 Possible Future Work

- An interesting avenue of future work is the extension of the proposed DJSC code design approaches to more than two sources. For orthogonal MAC, this involves designing good codes for several sources with the property that multiple UEP levels are provided for different bits within a codeword as determined by the pairwise correlation between the sources. For a GMAC, the message density modeling for EXIT analysis becomes more involved and it may require a different approximation method than that considered in Chapter 5.

- In this dissertation it has been assumed that sources and the channel are stationary. Another important extension include designing adaptive DJSC codes which are robust against variations in source correlation as well as the channel noise level. This requires estimating the inter-source correlation during the decoding process. This may require introducing a two stage decoder in which the first stage estimates the inter-source correlation and the second stage produces (soft) estimates which are fed back to the first stage. Similar methodology is applied in [89] for correlation estimation for the asymmetric case.

- In this dissertation it has been assumed that the sources have uniform distributions. In this case, the redundancy is due to the dependence between the sources and no redundancy exist within each source. In future work the code design for non-uniform sources can be can also be considered. In this case, the redundancy within each source can be exploited to further improve the efficiency/reliability of communication. For a single non-uniform source, the JSC code design has been considered in [90] which may be extended to the multi-user case as well. However, the extension is not straightforward in the sense that, codes designed are not necessarily systematic and the preservation of the inter-source correlation over the channel appears non-trivial.

Appendix A

Derivation of (5.13) and (5.14)

Since all the messages passed along the edges are in LLR form, we have the message passed from a state check node to a variable node of X_1 as,

$$m^{(1)}_{s\to v} = \log \frac{p(x_1 = +1|y)}{p(x_1 = -1|y)}. \tag{A.1}$$

The numerator in the logarithm can be expanded as

$$\begin{aligned} p(x_1 = +1|y) &= \frac{p(x_1 = +1, y)}{p(y)} && \text{(A.2)} \\ &= \frac{p(y, x_1 = +1, x_2 = +1) + p(y, x_1 = +1, x_2 = -1)}{p(y)} && \text{(A.3)} \\ &= \frac{p(y|x_1 = +1, x_2 = +1)p_{00} + p(y|x_1 = +1, x_2 = -1)p_{01}}{p(y)} && \text{(A.4)} \\ &= \frac{q_{00}p_{00} + q_{01}p_{01}}{p(y)}. && \text{(A.5)} \end{aligned}$$

Similarly, we can expand the denominator as

$$
\begin{aligned}
p(x_1 = +1|y) &= \frac{p(x_1 = -1, y)}{p(y)} && \text{(A.6)} \\
&= \frac{p(y, x_1 = -1, x_2 = +1) + p(y, x_1 = -1, x_2 = -1)}{p(y)} && \text{(A.7)} \\
&= \frac{p(y|x_1 = -1, x_2 = +1)p_{00} + p(y|x_1 = -1, x_2 = -1)p_{01}}{p(y)} && \text{(A.8)} \\
&= \frac{q_{10}p_{10} + q_{11}p_{11}}{p(y)}. && \text{(A.9)}
\end{aligned}
$$

Now we write,

$$
\begin{aligned}
p_{00} &= Pr(X_1 = +1, X_2 = +1) && \text{(A.10)} \\
&= Pr(X_2 = +1)Pr(X_1 = +1|X_2 = +1) \\
&= Pr(X_2 = +1)\beta_{00}^{(1)}.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
p_{01} &= Pr(X_2 = -1)\beta_{01}^{(1)}, && \text{(A.11)} \\
p_{10} &= Pr(X_2 = +1)\beta_{10}^{(1)}, \\
p_{11} &= Pr(X_2 = -1)\beta_{11}^{(1)}.
\end{aligned}
$$

Plugging (A.11) and (A.12) into (A.5) and (A.9), dividing by $Pr(X_2 = -1)$, and using $m_{s\to v}^{(2)} = \log \frac{Pr(X_2=+1)}{Pr(X_2=-1)}$, we get (5.13).The message update for $m_{s\to v}^{(2)}$ can be calculated in a similar fashion.

Bibliography

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

[2] A. Aaron and B. Girod, "Compression with side information using turbo codes," in *Data Compression Conference*, pp. 252–261, April 2002.

[3] J. Garcia-Frias, Y. Zhao, and W. Zhong, "Turbo-like codes for transmission of correlated sources over noisy channels," *IEEE Signal Processing Magazine*, pp. 58–66, Sept. 2007.

[4] A. Roumy and D. Declercq, "Characterization and optimization of LDPC codes for the 2-user Gaussian multiple access channel," *EURASIP Journal on Wireless Communications and Networking*, June 2007. Article ID 74890.

[5] T. Cover and J. Thomas, *Elements of Information Theory.* Wiley-Interscience, 2 ed., 2006.

[6] J. G. Proakis and M. Salehi, *Digital Communications.* McGraw Hill International Editions, 2008.

[7] C. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE National Conventional Board, Part 4*, pp. 142–163, 1959.

[8] D. A. Huffman, "A method for the construction of minimum redundancy codes," *Proceedings of the IRE*, vol. 40, pp. 1098–1101, 1952.

[9] D. Ornstein and B. Weiss, "Entropy and data compression schemes," *IEEE Transactions on Information Theory*, vol. vol. 39, pp. 78–83, Jan. 1993.

[10] T. Berger and J. D. Gibson, "Lossy source coding," *IEEE Transactions on Information Theory*, vol. 44, pp. 2693–2723, October 1998.

[11] R. Hamming, "Error detecting and error correcting codes," *The Bell System Technical Journal*, vol. vol. 26, pp. 147–160, April 1950.

[12] I. M. Jacobs, "Practical applications of coding," *IEEE Transactions on Information Theory*, vol. vol. 20, pp. 305–310, May 1974.

[13] J. D. J. Costello and J. G. D. Forney, "Channel coding: the road to channel capacity," *Proceedings of the IEEE*, vol. 95, pp. 1150–1177, June 2007.

[14] K. Sayood and J. C. Borkenhagen, "Use of residual redundancy in the design of joint source-channel coders," *IEEE Transactions on Communications*, vol. 39, no. 6, pp. 838–846, 1991.

[15] J. L. Massey, "Joint source channel coding," *Communication Systems and Random Process Theory*, pp. 279–293, 1978.

[16] Q. Xu, V. Stankovic, and Z. Xiong, "Distributed joint source-channel coding of video using raptor codes," *IEEE Journal on Selected Areas in Communications*, vol. vol. 25, pp. 851–861, May 2007.

[17] M. Fresia, F. Perez-Cruz, H. V. Poor, and S. Verdú, "Joint source and channel coding," *IEEE Signal Processing Magazine*, pp. 104–113, November 2010.

[18] I. F. Akyildiz, S. Weilian, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communication Magazine*, vol. 40, no. 8, 2002.

[19] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low cost outdoor localization for very small devices," *IEEE Personal Communications Magazine*, October 2000.

[20] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. vol 19, pp. 471–480, July 1973.

[21] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. vol. 22, pp. 1–10, January 1976.

[22] T. Berger and S. Tung, "Encoding of correlated analog sources," *IEEE USSR Joint Workshop on Information Theory*, pp. 7–10, 1975.

[23] A. Wyner, "Recent results in shannon theory," *IEEE Transactions on Information Theory*, vol. 20, pp. 2–10, 1974.

[24] Z. Xiong, A. D. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, September 2004.

[25] X. Zhu, Y. Liu, and L. Zhang, "Distributed joint source channel coding in wireless sensor networks," *Sensors*, pp. 4901–4917, June 2009.

[26] R. Ahlswede, "Multi-way communication channels," *in proceedings of the IEEE Int. Symposium on Information Theory, Tsahkadsor, Armenian SSSR.*, pp. 23–52, 1971.

[27] H. Liao, *Multiple access channels*. PhD thesis, University of Hawaii, Honolulu, HI, USA, 1972.

[28] D. Gündüz and E. Erkip, "Source and channel coding for correlated sources over multiuser channels," *IEEE Transactions on Information Theory*, vol. vol. 55, September 2009.

[29] S. Ray, M. Medard, M. Effros, and R. Koetter, "On separation for multiple access channels," *Information Theory Workshop*, 2006.

[30] T. Cover, A. E. Gamal, and M. Salehi, "Multiple-access channel with arbitrarily correlated sources," *IEEE Transactions on Information Theory*, vol. vol. IT-26, pp. 648–657, Nov. 1980.

[31] S. S. Pradhan and K. Ramachandran, "Distributed source coding: Symmetric rates and applications to sensor network," in *Data Compression Conference*, pp. 363–372, March 2000.

[32] T. Richardson and R. Urbanke, *Modern Coding Theory.* Cambridge Univ. Press, 2008.

[33] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Joint source-channel coding of binary sources with side information at the decoder using IRA codes," in *IEEE Workshop on Multimedia Signal Processing*, pp. 53–56, Dec. 2002.

[34] T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density-parity-check codes," *IEEE Transactions on Information Theory*, vol. 47, pp. 619–637, Feb. 2001.

[35] T. Richardson and R. Urbanke, *Modern Coding Theory.* Cambridge University Press, 2008.

[36] I. Shahid and P. Yahampath, "Distributed joint source-channel coding of correlated binary sources in wireless sensor networks," *In Proceedings of the International Symposium Wireless Comm. Systems*, Nov. 2011.

[37] I. Shahid and P. Yahampath, "Design of low-delay distributed joint source-channel codes using irregular LDPC codes," *Proceedings of the IEEE Vehicular Technology Conference*, Sept. 2012.

[38] I. Shahid and P. Yahampath, "Distributed joint source-channel coding using unequal error protection LDPC codes," *IEEE Transactions on Communications, accepted for publication*, 2013.

[39] S. Vembu, S. Verdú, and Y. Steinberg, "The source channel separation theorem revisited," *IEEE Transactions on Information Theory*, vol. 41, pp. 44–54, Jan 1995.

[40] T. Hindelang, J. Hagenauer, and S. Heinen, "Source-controlled channel decoding: Estimation of correlated parameters," *Proceedings of International ITG Conference on Source and Channel Coding (ITGSCC)*, pp. 259–266, 2000.

[41] M. Grangetto, P. Cosman, and G. Olmo, "Joint source/chaninel coding and map decoding of arithmetic codes," *IEEE Transactions on Communications*, vol. 53, no. 6, pp. 1007–1016, 2005.

[42] E. Ordentlich, G. Seroussi, S. Verdú, and K. Viswanathan, "Universal algorithms for channel decoding of uncompressed sources," *Information and Quantum Systems Lab*, February 2008.

[43] C. Poulliat, D. Declercq, C. Lamy-Bergot, and I. Fijalkow, "Analysis and optimization of irregular LDPC codes for joint source-channel decoding," *IEEE Communcation Letters*, vol. 9, pp. 1064–1066, December 2005.

[44] L. Pu, Z. Wuand, A. Bilgin, M. W. Marcellin, and B. Vasic, "LDPC-based iterative joint source-channel decoding for JPEG2000," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 577–581, 2007.

[45] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," *Lecture Notes in Computer Science*, vol. 2634, pp. 162–177, Apr 2003.

[46] J. Kusuma, L. Doherty, and K. Ramchandran, “Distributed compression for sensor networks,” *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, vol. 1, pp. 82–85, Oct. 2001.

[47] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” *IEEE Transactions on Information Theory*, vol. 19, pp. 471–480, Jul. 1973.

[48] J. J. Xiao and Z. Q. Luo, “Multiterminal source-channel communication over an orthogonal multiple-access channel,” *IEEE Transactions on Information Theory*, vol. 53, no. 9, pp. 3255–3264, 2007.

[49] R. Palanki, A. K. Khandekar, and R. J. McEliece, “Graph-based codes for synchronous multiple access channels,” *Proceedings of the Allerton Conference on Communications, Control and Computing*, October 2001.

[50] R. Peterson and D. J. Costello, “Binary convolutional codes for a multiple access channel,” *IEEE Transactions on Information Theory*, vol. IT-25, January 1979.

[51] F. Ali and B. Honary, “Collaborative coding and decoding techniques for multiple access channel,” *Proceedings of IEE Communications*, vol. 141, pp. 56–62, 1994.

[52] N. Ibrahim and G. Kaleh, “Iterative decoding and soft interference cancellation for the Gaussian multiple access channel,” *URSI International Symposium on Signals, Systems and Electronics*, pp. 156–161, 1998.

[53] S. Lin and D. J. C. Jr., *Error Control Coding*. Prentice Hall, 2004.

[54] S. S. Pradhan and K. Ramachandran, “Distributed source coding using syndromes (DISCUS): Design and construction,” in *Data Compression Conference*, March 1999.

[55] A. Roumy, K. Lajnef, and C. Guillemot, “Rate-adaptive turbo-syndrome scheme for Slepian-Wolf coding,” *IEEE Asilomar Conference on Signals, Systems and Computers*, Nov 2007.

[56] Z. Tu, J. Li, and R. S. Blum, “An efficient SF-ISF approach for the Slepian-Wolf source coding problem,” *EURASIP Journal on Applied Signal Processing*, pp. 961–971, May 2005.

[57] J. Bajcsy and P. Mitran, “Coding for the Slepian-Wolf problem with turbo codes,” *Proceeding of IEEE Global Communications Conference (GLOBECOM)*, pp. 1400–1404, Dec. 2001.

[58] F. Cabarcas and J. Garcia-Frias, “Approaching the Slepian-Wolf boundary using practical channel codes,” *IEEE International Symposium on Information Theory*, p. 330, June 2004.

[59] J. Garcia-Frias and Y. Zhao, “Compression of correlated binary sources using turbo codes,” *IEEE Communication Letters*, vol. 5, pp. 417–419, Oct. 2001.

[60] M. Sartipi and F. Fekri, “Distributed source coding using short to moderate length rate-compatible LDPC codes: The entire Slepian-Wolf region.,” *IEEE Transactions on Communications*, vol. 56, no. 3, pp. 400–411, 2008.

[61] N. Gehrig and P. L. Dragotti, “Slepian-Wolf codes with systematic and non-systematic linear codes,” *IEEE Communication Letters*, pp. 61–63, Jan 2005.

[62] J. Garcia-Frias, “Joint source-channel decoding of correlated sources over noisy channels,” *Data Compression Conference*, pp. 283–292, 2001.

[63] J. Garcia-Frias and Y. Zhao, “Near-Shannon/Slepian-Wolf performance for unknown correlated sources over AWGN channels,” *IEEE Transactions on Communications*, vol. 53, pp. 555–559, Apr. 2005.

[64] Q. Xu, V. Stankovic, and Z. Xiong, "Distributed joint source-channel coding of video using Raptor codes," *IEEE Journal on Selected Areas in Communication*, vol. 25, pp. 851–861, May 2007.

[65] F. Daneshgaran, M. Laddomada, and M. Mondin, "Iterative joint channel decoding of correlated sources employing serially concatenated convolutional codes," *IEEE Transactions on Information Theory*, vol. 51, pp. 2721–2763, July 2005.

[66] F. Daneshgaran, M. Laddomada, and M. Mondin, "LDPC-based channel coding of correlated sources with iterative decoding," *IEEE Transactions in Communication*, vol. 54, pp. 577–582, Apr 2006.

[67] F. Daneshgaran, M. Laddomada, and M. Mondin, "Iterative joint decoding of correlated sources," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 2659–2663, Oct. 2006.

[68] J. Barros and S. D. Servetto, "Network information flow with correlated sources," *IEEE Transactions on Information Theory*, vol. 52, pp. 155–170, Jan 2006.

[69] P. L. Dragotti and M. Gastpar, *Distributed Source Coding*. Academic Press, 2009.

[70] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Compression of binary sources with side information at the decoder," *IEEE Communication Letters*, vol. 6, pp. 440–442, Oct 2002.

[71] P. Tan and J. Li, "Enhancing the robustness of distributed compression using ideas from channel coding," *Proceeding of IEEE Global Communications Conference (GLOBECOM)*, pp. 2385–2389, 2005.

[72] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, pp. 1727–1737, Oct. 2001.

[73] V. Kumar and O. Milenkovic, “On unequal error protection of LDPC codes based on Plotkin-type constructions,” *IEEE Transactions on Communications*, vol. 54, no. 6, pp. 994–1005, 2006.

[74] N. Rahnavard, H. Pishro-Nik, and F. Fekri, “Unequal error protection using partially regular LDPC codes,” *IEEE Transactions on Communications*, vol. 55, pp. 387–391, Mar. 2007.

[75] S. ten Brink, G. Kramer, and A. Ashikhmin, “Design of low-density-parity-check codes for modulation and detection,” *IEEE Transactions on Communications*, vol. 52, pp. 670–678, Apr. 2004.

[76] J. Garcia-Frias and F. Cabarcas, “Approaching the Slepian-Wolf boundary using practical channel codes,” *Signal Processing*, vol. 86, pp. 3096–3101, 2006.

[77] D. Schonberg, K. Ramachandran, and S. S. Pradhan, “Distributed code construction for the entire Slepian-Wolf rate region for arbitrarily correlated sources,” in *37. Asilomar Conference on Signals, Systems, and Computers*, pp. 835–839, 2003.

[78] A. W. Eckford and W. Yu, “Density evolution for the simultaneous decoding of LDPC-based Slepian-wolf source codes,” in *International Symposium on Information Theory (ISIT)*, pp. 1401–1405, 2005.

[79] S.-Y. Chung, T. Richardson, and R. Urbanke, “Analysis of sumproduct decoding of LDPC codes using a gaussian approximation,” *IEEE Transactions on Information Theory*, vol. 47, pp. 657–670, Feb. 2001.

[80] E. Sharon, A. Ashikhmin, and S. Litsyn, “EXIT functions for binary input memoryless symmetric channels,” *IEEE Transactions on Communications*, vol. 54, pp. 1207–1214, July 2006.

[81] M. Luby, G. Mitzenmacher, A. Shokrollahi, and D. Spielman, "Improved low-density parity-check codes using irregular graphs," *IEEE Transactions on Information Theory*, vol. 47, pp. 585–598, Feb. 2001.

[82] J. Li, Z. Tu, and R. S. Blum, "Slepian-Wolf coding for nonuniform sources using turbo codes," in *Data Compression Conference*, pp. 312–321, 2004.

[83] G. F. Jr., "Codes on graphs: Normal realizations," *IEEE Transactions on Information Theory*, vol. 47, pp. 520–548, Feb. 2001.

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

[85] E. John, H. Ahn, L. Prichep, M. Trepetin, D. Brown, and H. Kaye, "Developmental equations for the electroencephalogram," *Science, 210*, pp. 1255–1258, 2013.

[86] G. Box and D. Cox, "An analysis of transformations," *J R Stat Soc B, 26*, pp. 211–252, 1964.

[87] T. Gasser, P. Bcher, and J. Mcks, "Transformations towards the normal distribution of broad band spectral parameters of the EEG," *Electroencephalogr Clin Neurophysiol, 53*, pp. 119–124, 1982.

[88] A. Azevedo-Filho and R. D. Shachter, "Laplace's method approximations for probabilistic interference in belief networks with continuous variables," *Uncertainty in Artificial Intelligence*, 1994.

[89] S. Wang, L. Cui, L. Stankovic, V. Stankovic, and S. Cheng, "Adaptive correlation estimation with particle filtering for distributed video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 5, pp. 649–658, 2012.

[90] M. Fresia, F. Prez-Cruz, H. V. Poor, and S. Verdú, “Joint source and channel coding,” *IEEE Signal Processing Magazine*, pp. 104–113, Nov 2010.