Iterative Decoding Scheme for Cooperative Communications

Xiaoyan Xu

A thesis submitted for the degree of Doctor of Philosophy.
The University of Edinburgh.
September 2009

Abstract

Cooperative communication becomes a practical alternative to multiple-input multiple-output (MIMO) systems when MIMO cannot be implemented due to size, cost or hardware limitations. The diversity provided by cooperative communication techniques can effectively combat the deleterious effect of fading, thus cooperative diversity is believed as a powerful approach to facilitate high-speed and high-quality wireless communications.

Originally invented by Gallager in the early 1960s, Low-Density Parity-Check (LDPC) codes were largely ignored until Mackay and Neal rediscovered them in 1995. Nowadays owing to its sparse graph structure and low complexity iterative decoding methods, LDPC codes have not only been successful in the field of coding theory but are also being used in a variety of practical applications. LDPC codes have been adopted in several standards and they are a promising candidate for the channel coding scheme to be used in next generation mobile communication systems.

This thesis explores practical decoding schemes of LDPC codes for cooperative communication. We start with the Gaussian multiple access channel (MAC) since the MAC could be seen as one component in a multi-user cooperative communication system. In this part, we investigate the performance of LDPC codes with soft iterative successive interference cancellation (SIC) and decoding scheme using BPSK modulation. Next, we propose a cooperative diversity scheme for the communication model of two sources sharing a single relay under two scenarios. The scheme uses algebraic code superposition relaying in a frequency division mode or in the multiple access fading channel to create spatial diversity under the constraint of limited communication resources. We also describe in detail a novel, computationally efficient message passing algorithm at the destination's decoder which can extract the substantial spatial diversity contained in the code and signal superposition. The decoder is based on a sliding window structure where certain *a posteriori* soft values are retained to form *a priori* soft values for the next decoding. We show that despite the simplicity of the proposed scheme, diversity gains are efficiently leveraged by the simple combination of channel coding at the sources and network coding at the relay.

Acknowledgements

First of all, I would like to thank my supervisor Prof. Norbert Goertz for his continuous support, encouragement and invaluable advice throughout the years of my PhD.

I would like to thank Dr. Mark Flanagan for fruitful discussions with him and great help he provided in my research work.

Many thanks also go to Dr. John Thompson who took the time out of his busy schedule to offer me great help and suggestion in writing up this thesis, and Prof. Steve McLaughlin for his supervision.

My gratitude is also extended to my colleagues in IDCOM and all my dear friends in Edinburgh for providing much fun and laughter to take away the monotony of doing a PhD.

Last but most importantly, I would like to thank my parents for their emotional and financial support and their endless love. Without them, completion of this PhD will not have been possible. This thesis is dedicated to them.

Contents

List of figures

List of tables

Acronyms and abbreviations

APP	a posteriori probability
AWGN	additive white Gaussian noise
BCJR	Bahl-Cocke-Jelinek-Raviv
BER	bit error rate
BF	bit-flipping
BPSK	binary phase-shift keying modulation
BSC	binary symmetric channel
CDMA	code division multiple access
CRC	cyclic redundancy check
CSI	channel state information
DE	density evolution
FDMA	frequency division multiple access
FEC	forward error correction
i.i.d.	independent and identically distributed
ISI	intersymbol interference
LAN	local area network
LDGM	low-density generator matrix
LDPC	low-density parity check
LLR	log-likelihood ratio
MAC	multiple access channel
MAN	metropolitan area network
MAP	maximum a posteriori
MARC	multiple-access relay channel
MIMO	multiple-input and multiple-output
MLG	majority-logic
ML	maximum likelihood
MUD	multiuser detector
OFDM	orthogonal frequency-division multiplexing
pdf	probability density function

pmf	porbability mass function
PSD	power spectral density
RCPC	rate compatible punctured convolutional
RM	Reed-Muller
RS	Reed-Solomon
SIC	successive interference cancellation
SPA	sum-product algorithm
TDMA	time division multiple access
WBF	weighted bit-flipping

Nomenclature

a	Channel coefficient
$\mathbb{C}$	Complex numbers
$\mathcal{C}$	A binary linear block code
C	Channel capacity
d_{min}	Minimum Hamming distance
$\mathbf{H}$	Parity check matrix
$I(;)$	Mutual information
$\mathbf{G}$	Generator matrix
k	Number of information digits
$L(\cdot)$	Log-likelihood ratio
n	Number of encoded digits
N_{it}	Maximum global iteration number
$P(\cdot)$	Probability mass function
$p(\cdot)$	Probability density function
Ind $[\cdot]$	Indicator function: Takes the value 1 if the argument is true, and 0 if the argument is false
R	Code rate
$\mathbb{R}$	Real numbers
$\Re\{\cdot\}$	The real part of a complex number
$(\cdot)^T$	Transpose
$\mathbf{u}$	Information sequence
$\hat{\mathbf{u}}$	Estimated information sequence
$\mathbf{v}$	Encoded sequence
$\hat{\mathbf{v}}$	Estimated codeword
v_i	One variable
$(\cdot)^*$	Conjugate
$\oplus$	Addition in $GF(2)$
$\sum^{\oplus}$	Summation in GF(2)
$\sum_{\sim\{v_i\}}$	Summary for variable v_i

λ LLRs corresponding to message passed on the factor graph

γ Column weight in a parity check matrix

ρ Row weight in a parity check matrix

Chapter 1
Introduction

1.1 General Background

Around 1897, Marconi's successful demonstration of wireless telegraphy is seen as the beginning of wireless communication. Over more than a hundred years' development, especially the explosive growth in the past decade, wireless communication industry is the most vigorous areas in the communication field today.

The most successful wireless networking system so far has been the cellular system whose concept is to divide the service area into small cells for the purpose of frequency reuse. Each small cell is covered by a fixed base station which offers services for the users within this cell. Especially, the wireless link from a base station to a mobile user is called a downlink or forward channel, and the link from a mobile user to a base station is called an uplink or reverse channel.

The first wireless cellular system can be traced back to the 1960s, and used analog communications, such as Advanced Mobile Phone Service (AMPS) [Gol05]. Second-generation cellular systems, such as the global system for mobile (GSM) originated in Europe, time-division multiple access (TDMA) in US (IS-136) and CDMA (IS-95), are digital and were built mainly for voice data and slow data transmission [Gol05]. Although cellular systems were originally developed for telephony, the third-generation cellular systems (3G) are designed to handle data with specific characteristics and/or voice [TV05, DPSB07]. The fourth-generation (known as 4G or beyond 3G) is expected to provide high-quality voice, data and stream multimedia services based on an "Anytime, Anywhere" scale. The principle technologies in 4G include orthogonal frequency-division multiplexing (OFDM), multiple-input and multiple-output (MIMO), the Turbo principle, the cooperative relaying concept, etc. [STB09].

1.2 Cooperative Diversity

Wireless channels suffer from multipath fading which can cause poor transmission performance due to the loss of receive signal power with a fixed receiver noise level. Based on the fact

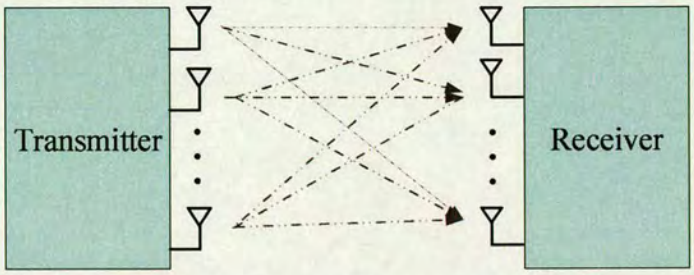

Figure 1.1: *A MIMO system.*

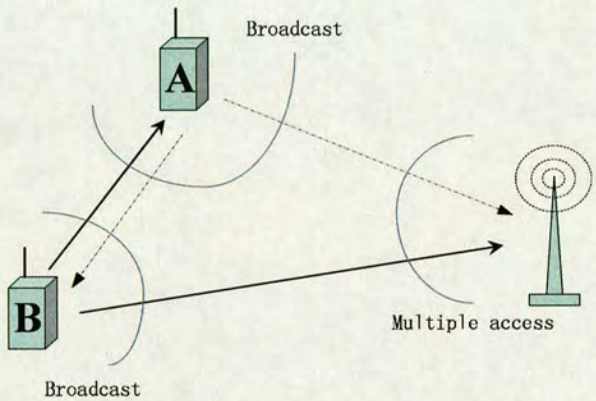

Figure 1.2: *An uplink cooperative diversity example with two mobile users (A and B) and one base station.*

that individual channels experience different levels of fading, the concept of mitigating the detrimental effects of multipath fading is to provide the receiver with multiple versions of the transmitted information through independent fading channels. Particularly spacial diversity is generated by transmitting the signal from different locations to the destination. The use of multiple antennas at both the transmitter and receiver, so called *multiple-input and multiple-output* (MIMO) technique depicted in Figure 1.1, can provide a spatial diversity gain, spatial multiplexing gain and array gain compared to single-antenna systems

However, MIMO techniques are not applicable to many wireless systems due to size, complexity or hardwire constraints. For example, in the uplink of cellular systems (as well as *ad hoc* and *sensor* networks), the size and power of a mobile handset is the limiting factor. Taking advantage of the broadcast nature of wireless channels, one user's transmitted signal can be captured by a relay station or another user in the network. As illustrated in Figure 1.2, when user A transmits its signal to the base station, there is actually a broadcasting process which means that user B nearby can also decode the information from user A. Hence, apart from the original transmission channel (user A to base station), user B can also help transmit user A's

information to the base station through its fading channel which leads to a virtual two-antenna transmitter system. Hence the basic idea behind *cooperative diversity* [SEA03a,SEA03b] is that one user's information, by help of other users or relaying stations in the network, can be transmitted through multiple independent fading channels to the destination which correspondingly forms a virtual distributed multiple antenna system. The cooperation model can be decomposed into two components: a broadcast channel (A transmitting, B and base sation receiving or B transmitting, A and base sation receiving) and a multiple access channel (A and B transmitting, base station receiving).

1.3 The Turbo Principle and Low Density Parity-Check (LDPC) Codes

In 1948, Shannon published a seminal paper [Sha48] which gave the first quantitative definition of information, thus creating the field of information theory. Shannon proved that reliable communication of information over unreliable channels is feasible by means of "channel codes". Since then, the ultimate goal of a communication system designer is to operate as close to the theoretical capacity as possible with acceptable transmitter and receiver complexity.

In the ensuing years after Shannon's publication, coding researchers were working on the good codes with a large number of structures, such as Reed-Solomon (RS) codes, Reed-Muller (RM) codes and BCH codes [LC04]. In 1993, [BGT93] proposed a coding technique, called *Turbo coding*, which provides exceptionally good performance. Turbo codes feature in a code design that produces pseudorandom properties and a decoder design that makes use of soft-output values and iterative decoding [LC04]. The expression "Turbo principle" is used to describe the key idea underlying the turbo-decoding algorithm, namely a "strategy exploiting the iterated exchange of soft information between different blocks in a communication receiver" [SDE+01b, SDE+01a, Hag02]. The turbo principle becomes a general principle which is applied in many mobile communications receivers to improve performance through the iterative processing of extrinsic information.

In 1963, Gallager [Gal63] invented LDPC codes by using the sum-product algorithm (SPA) as the decoding method. Unfortunately the decoding complexity was too high for practical implementation at that time. As a result, LDPC codes were largely neglected by coding community for decades except for some work [Tan81, Mar82]. In the wake of the discovery of turbo codes,

LDPC codes were rediscovered in the mid-1990s [MN95, SS96] as a class of Turbo-like codes and an explosion of research interest ensued due to their performance being comparable to and sometimes better than the original Turbo codes. The LDPC code constructed in [CFRU01] is capable of achieving reliable performance within 0.0045dB of the Shannon limit in the binary input additive white Gaussian noise (AWGN) channel. Owing to the low complexity iterative message-passing decoding method, LDPC codes are also attractive in practical applications: LDPC codes are already used as standards for the satellite transmission of digital television in DVB-S2 [DVB]; in 2008, LDPC beat convolutional Turbo codes to become the FEC scheme for the ITU-T G.hn standard [ITU] due to its lower decoding complexity and lower error floor compared to the proposed Turbo Codes at the desired range of operation.

1.4 Main Contributions of the Thesis

This work is concerned with the practical application of LDPC codes in cooperative communication environments.

As we can see from Figure 1.2, the multiple access channel (MAC) is one component in a multi-user cooperative communication system. Hence we begin with the investigation of LDPC codes in the Gaussian MAC. It is known from information theory [CT06] that successive interference cancellation (SIC) is a technique that can achieve the upper limits of the capacity region of the Gaussian MAC. However, the theory assumes that Gaussian signal alphabets are used which is not possible in practice. Hence, we go to the opposite extreme and investigate the performance of SIC schemes when a binary modulation signal alphabet is used. This is motivated by the idea to have very simple transmitters such as wireless network nodes communicating to a receiving "base station" where it is allowed for some higher complexity to perform more advanced signal processing such as SIC. A formulation of a multi-user detector is derived which lends itself to an efficient implementation by L-value algebra. Besides, as it is impossible to decode with no errors in practice, a soft iterative SIC and decoding scheme is presented by factor graph theory where the interference cancellation and decoding are carried out by exchanging the extrinsic L-values of bit-reliability information. A set of practical LDPC codes with moderate blocksize are used to investigate the practical performance of interference cancellation.

Although "coding" has been employed in several papers on cooperative communications [SE04, HN02b, HN06, JHHN04], diversity gain is obtained through channel coding gain by means of

single channel codes. Inspired by the fact that network coding is a technique well known for its capability to increase system throughput, joint channel and network coding is believed to be an efficient way to generate spatial diversity with limited power and bandwidth resources for uplink transmissions [XFKC06b, HD06]. In this thesis, a cooperative diversity scheme is proposed for the communication model of two sources sharing a single relay. Two scenarios and correspondingly two cooperative coding schemes are under consideration. The first scenario is in a frequency division mode where algebraic code superposition relaying is exploited to generate diversity gain. The second scenario is in a multiple access fading mode where the code superposition relay is accompanied by the interference introduced by the fading multiple access channel. A novel, computationally efficient message passing algorithm is described in detail for the destination's decoder which can extract substantial spatial diversity contained in the code and signal superposition. The decoder is based on a sliding window structure which attains a separation of the soft-input soft-output (SISO) decoder modules with respect to each received signal stream. Despite the simplicity of the proposed scheme, diversity gains are efficiently leveraged by the simple combination of channel coding at the sources and network coding at the relay.

1.5 Structure of the Thesis

The rest of the thesis is organized as follows. Chapter 2 provides the necessary background on channel coding, factor graph theory and cooperative communication. Chapter 3 investigates the performance of the soft iterative SIC and decoding scheme by LDPC codes under binary-input constraints. Chapter 4 first addresses the proposed cooperative coding scheme in the frequency division scenario, and then moves to the fading MAC scenario. Chapter 5 summarizes this thesis and discusses some related future work.

Chapter 2
Background

The contents in this chapter are mainly classified into three parts: channel coding theory, factor graph theory and cooperative communications. First the necessary background on channel coding theory is reviewed. Next, the factor graph theory is introduced in an individual section since it acts as a presentation tool throughout this thesis. Due to the strong relation between factor graph theory and graph-based decoding algorithms, Sum-product algorithm (SPA) decoding for LDPC codes and relative analysis methods for LDPC codes are presented at the end of the factor graph section. Last of all, this chapter reviews some basic concepts of cooperative communication which can effectively mitigate the detrimental effect of fading in wireless environments.

2.1 Basics of Channel Coding

Coding choice is strongly dependent on the channel model. In this section, we focus on channel coding under an additive white Gaussian noise (AWGN) channel as it shapes the coding discipline.

2.1.1 Shannon's channel coding theorem

Shannon's noisy channel coding theorem [Sha48] states that for a data transmission rate $R_T \leq C$, with C the capacity of a discrete-input memoryless channel (C and R_T are both in bits per transmission), there exists a coding scheme that achieves an arbitrary small probability of error; conversely, if $R_T > C$ any coding scheme will be unreliable. If M is the size of the discrete channel input alphabet, the code rate is given by $R = \dfrac{R_T}{\log_2(M)}$. In the case of binary channel inputs, we have R equals to R_T.

Suppose that E_s denotes the receive energy per channel symbol; E_b denotes the receive energy per information bit; N_0 denotes the one-sided noise power spectral density. Thus, we have the

following power correction equation:

$$\frac{E_s}{N_0} = \frac{E_b}{N_0} \cdot R_T \tag{2.1}$$

The error probability of a coded communication system is commonly expressed in terms of the ratio of energy per information bit E_b to the one-sided PSD N_0 of the channel noise.

The information capacity C of a communication channel is defined as the maximum mutual information between the input X and output Y over all distributions on the input symbols:

$$C = \max_{p(x)} I(X;\, Y) \tag{2.2}$$

where $p(x)$ is the PDF of the input variable X.

Consider a discrete-time AWGN channel with the following expression:

$$y_i = x_i + n_i \tag{2.3}$$

where x_i is the real-valued channel input at time i, y_i is the corresponding output and n_i is a white Gaussian noise random process.

The information capacity of a discrete-time, AWGN channel with input power constraint $E\{X^2\} \leq P$ and a noise variance of σ^2 is

$$C = \max_{p(x):E\{X^2\}\leq P} I(X;\, Y) = \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma^2}\right) \tag{2.4}$$

with C in information bits per transmission per dimension. In this thesis, we only consider discrete-time system since most continuous-time systems can be converted to discrete-time system via sampling. For a continuous time, B-band-limited AWGN channel with sampling at Nyquist rate $2B$, the capacity expression is

$$C = \frac{1}{2}\log_2\left(1 + \frac{2E_s}{N_0}\right) \leq \frac{1}{2}\log_2\left(1 + \frac{2 \cdot C \cdot E_b}{N_0}\right). \tag{2.5}$$

We can further obtain

$$\frac{E_b}{N_0} \geq \frac{2^{2C} - 1}{2C} \tag{2.6}$$

Equation (2.6) gives a theoretical limit on the minimum SNR required for a coded system with data transmission rate C to achieve error-free transmission.

Equations (2.4), (2.5) and (2.6) apply to the Gaussian-distributed inputs. For transmission over a binary-input, continuous-output AWGN channel with BPSK signalling, the Shannon limit (in terms of $\frac{E_b}{N_0}$) does not have a closed form but can be evaluated numerically.

The channel capacity for the binary-input discrete-time AWGN channel with equiprobable input bits $x \in \{0, 1\}$ and the channel output signal y is

$$C = \sum_{x \in \{0,1\}} \int_{-\infty}^{\infty} \frac{p(y|x)}{2} \log_2 \frac{2 \cdot p(y|x)}{p(y|x=1) + p(y|x=0)} dy \tag{2.7}$$

where $p(\cdot)$ represents the PDF of Gaussian noise and σ^2 is the variance of the noise:

$$p(y|x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-x)^2 \right). \tag{2.8}$$

Suppose one symbol is transmitted every T seconds, the symbol transmission rate (baud rate) is $1/T$. For a baseband bandwidth B, the symbol transmission rate is limited to $2B$. Hence we have $1/T = 2B$. For a coded system with code rate $R = k/n$ where k information bits correspond to the transmission of n symbols, the information transmission rate (data rate) is R/T bits per second (bps), where $R/T = 2RB$. In a binary coded system, $R < 1$. If the coded system aims to maintain the same information transmission rate as the uncoded system, its corresponding bandwidth shall become B/R which means the coded system requires bandwidth expansion by a factor of $1/R$. If the system does not permit bandwidth expansion, redundancy for coding can be covered by a combined coding and modulation technique, called coded modulation which allows significant coding gain without bandwidth expansion [IH77, Ung82].

2.1.2 Basics of channel coding

A block diagram of a typical data transmission system on an AWGN channel is shown in Figure 2.1. In the following, we focus on the channel encoder and channel decoder and we only consider binary linear codes.

Suppose each message block, denoted by $\mathbf{u}$, consists of k information digits. There are a total of 2^k distinct messages. The channel encoder transforms a k-bit information sequence $\mathbf{u}$ into

an n-bit codeword $\mathbf{v}$. The rule of a channel decoder is to transform the received sequence $\mathbf{r}$ into an estimated information sequence $\hat{\mathbf{u}}$ which is the estimated information sequence. The decoder can also produce $\hat{\mathbf{v}}$ which is the estimated codeword of $\mathbf{v}$ since there is a one-to-one correspondence between the information sequence $\mathbf{u}$ and the codeword $\mathbf{v}$.

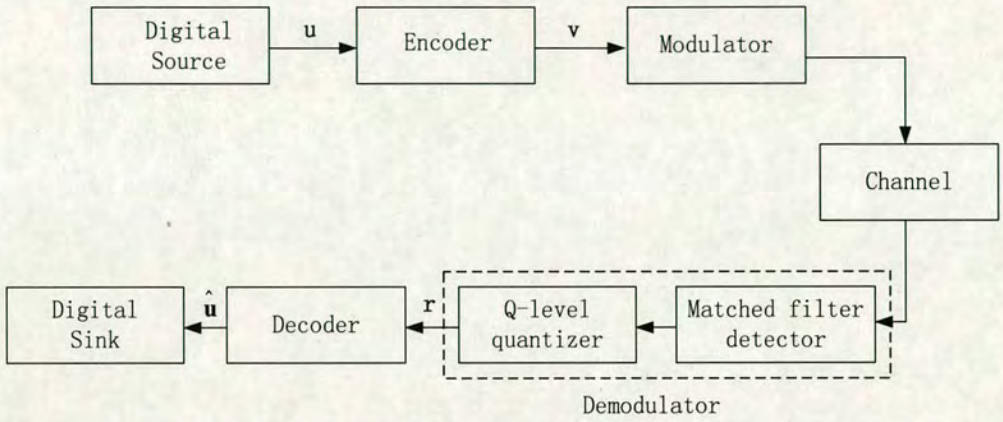

Figure 2.1: *A block diagram of a coded system on a discrete-time channel.*

ML and MAP decoding rules

A decoding error happens when $\hat{\mathbf{v}} \neq \mathbf{v}$. The error probability of the decoder is defined as

$$P(E) = \sum_{\mathbf{r}} P(E|\mathbf{r})P(\mathbf{r}) \tag{2.9}$$

where $P(\mathbf{r})$ is the probability of the received sequence $\mathbf{r}$ and independent of the decoding rule used. Thus, the decoding rule is to minimize $P(\hat{\mathbf{v}} \neq \mathbf{v}|\mathbf{r})$ for all $\mathbf{r}$, which is equivalent to maximizing

$$P(\mathbf{v}|\mathbf{r}) = \frac{P(\mathbf{r}|\mathbf{v})P(\mathbf{v})}{P(\mathbf{r})} \tag{2.10}$$

A decoding rule which aims to maximize equation (2.10) is called the *maximum a posteriori* (MAP) rule. If all information sequences are equally likely, $P(\mathbf{v})$ is the same for all $\mathbf{v}$. Thus, maximizing $P(\mathbf{v}|\mathbf{r})$ is equal to maximizing $P(\mathbf{r}|\mathbf{v})$. A decoder chosen to maximize $P(\mathbf{r}|\mathbf{v})$ is called a decoder of the *maximum likelihood* (ML) rule. We could see if $P(\mathbf{v})$ does not depend on $\mathbf{v}$, i.e., $P(\mathbf{v})$ is the same for all $\mathbf{v}$, the MAP and ML rules are equivalent.

Hamming weight and Hamming distance

The Hamming weight of a codeword is defined as the number of non-zero components. The

Hamming distance is defined as the number of positions in which two codewords $\mathbf{v}_1$ and $\mathbf{v}_2$ differ and equals to the Hamming weight of the mod-2 sum of $\mathbf{v}_1$ and $\mathbf{v}_2$. The minimum Hamming distance d_{min} of a binary linear code $\mathcal{C}$ is the minimum of all Hamming distances between pairs of distinct codewords in $\mathcal{C}$. The parameter d_{min} gives an important measure of code performance. A code with minimum distance d_{min} is capable of detecting all codewords with errors less than e_d and correcting all codewords with errors less than e_c, with $e_d + e_c < d_{min}$ and $e_c \leq e_d$ [Pro01].

Hard decision and soft decision decoding

At the output of the demodulator, if each component of the signal $\mathbf{r}$ is quantized in two levels, denoted as 0 and 1, the decoder processes this hard-decision binary received sequence based on a specific decoding method. This type of decoding is called *hard-decision* decoding. Hard decision decoding using algebraic structures of the code is called *algebraic decoding*. The objective of a hard-decision decoding is to choose the codeword nearest in Hamming distance to the received vector. The loss of information in a hard decision causes performance degration of the decoder.

If the outputs of the matched filter are unquantized or quantized in more than two levels, the demodulator makes soft decisions. Decoding by processing this soft-decision received sequence is called *soft-decision* decoding. Soft-decision decoding provides better error performance than hard-decision decoding due to the additional information contained in the unquantized or multilevel quantized signal. In general, maximum likelihood decoding with soft decision has a coding gain of 3dB over algebraic decoding [LC04]. The price to pay for better error performance is the computational complexity required by soft-decision decoding.

2.1.3 Linear block codes

Block codes and convolutional codes are two structurally different types of codes in common use today. In this thesis, we focus on LDPC codes which belong to linear block codes. Next we introduce the basic structure of binary linear block codes. More information about convolutional codes could be found in [LC04].

We still use $\mathbf{u}$ to denote each message block which consists of k information bits and a binary n-tuple $\mathbf{v}$ to denote the codeword of the message $\mathbf{u}$. For channel coding, $n > k$ and the code rate is $R = k/n$. The one-to-one correspondence between $\mathbf{u}$ and its $\mathbf{v}$ for linear block codes

can be written as

$$\mathbf{v} = \mathbf{u} \cdot \mathbf{G} = (u_0,\ u_1, \cdots,\ u_{k-1}) \cdot \begin{bmatrix} \mathbf{g_0} \\ \mathbf{g_1} \\ \vdots \\ \mathbf{g_{k-1}} \end{bmatrix}$$

where the $k \times n$ matrix $\mathbf{G}$ with k linearly independent rows is called a *generator matrix* for this linear code.

Definition 2.1.1. A block code of length n and 2^k codewords is called a linear $(n,\ k)$ code if and only if its 2^k codewords form a k-dimensional subspace of the vector space of all the n-tuples over the field of $GF(2)$ [LC04].

Remark 2.1.1. A binary block code is linear if and only if the modulo-2 sum of any two codewords is also a codeword.

Remark 2.1.2. In an $(n,\ k)$ linear code $\mathcal{C}$, every codeword $\mathbf{v}$ could be represented by a linear combination of k independent codewords $\mathbf{g_0},\ \mathbf{g_1}, \cdots,\ \mathbf{g_{k-1}}$:

$$\mathbf{v} = u_0\mathbf{g}_0 + u_1\mathbf{g}_1 + \cdots + u_{k-1}\mathbf{g}_{k-1}. \tag{2.11}$$

Remark 2.1.3. A desirable property for a linear block code to possess is systematic structure. Each systematic codeword could be divided into the message part, which is composed of k unaltered information (or message) digits, and the redundant checking part, which is composed of $(n - k)$ parity check digits made by the linear sums of the information digits. A linear systematic $(n,\ k)$ code is completely specified by a generator matrix $\mathbf{G}$ of the following form:

$$\mathbf{G} = \begin{bmatrix} \mathbf{g_0} \\ \mathbf{g_1} \\ \vdots \\ \mathbf{g_{k-1}} \end{bmatrix} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0,n-k-1} & 1 & 0 & \cdots & 0 \\ p_{10} & p_{11} & \cdots & p_{1,n-k-1} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{k-1,0} & p_{k-1,1} & \cdots & p_{k-1,n-k-1} & 0 & 0 & \cdots & 1 \end{bmatrix} = [\mathbf{P}\ \mathbf{I}_k]$$

where $\mathbf{I}_k$ denotes the $k \times k$ identity matrix. Any nonsystematic code generator matrix $\mathbf{G}$ can be converted to systematic form by means of row operations and column permutations.

The *parity check matrix* $\mathbf{H}$ is another matrix associated with each linear block code. $\mathbf{H}$ is an $m \times n$ $(m \geq n - k)$ matrix such that the rows of $\mathbf{H}$ span the null space of the linear block code

$\mathcal{C}$. A binary n-tuple $\mathbf{v}$ is a codeword in $\mathcal{C}$ if and only if

$$\mathbf{v} \cdot \mathbf{H}^T = \mathbf{0}_{1\times m} \tag{2.12}$$

Corresponding to the systematic form of $\mathbf{G}$, the systemtic form of $\mathbf{H}$ takes the following form (when $m = n - k$):

$$\mathbf{H} = [\mathbf{I}_{n-k} \ \mathbf{P}^T] = \begin{bmatrix} 1 & 0 & \cdots & 0 & p_{00} & p_{10} & \cdots & p_{k-1,0} \\ 0 & 1 & \cdots & 0 & p_{01} & p_{11} & \cdots & p_{k-1,1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & p_{0,n-k-1} & p_{1,n-k-1} & \cdots & p_{k-1,n-k-1} \end{bmatrix}$$

where $\mathbf{I}_{n-k}$ denote the $(n-k) \times (n-k)$ identity matrix.

2.1.4 LDPC codes

As we mentioned in the introduction part, LDPC codes were rediscovered in the mid-1990s following the invention of turbo codes which is another capacity-approaching code. Generally speaking, LDPC codes show the following advantage over turbo codes [LC04]: they do not require a long interleaver to achieve good performance since there is an inherent "interleaver" in the Gaussian elimination process in the transformation between the $\mathbf{G}$ matrix and the $\mathbf{H}$ matrix for linear block codes; they have better block error performance; the error floor of LDPC codes are lower than Turbo codes; the decoding algorithms are not trellis based and hence do not require as much memory as Turbo codes do. Next we shall present some basic concepts of LDPC codes. Because LDPC codes are codes defined on a graph, the graphical representation and corresponding decoding issue shall be in Section 2.2.4 after the introduction of factor graph theory.

Definition 2.1.2. A (γ, ρ) regular LDPC code is a class of linear block codes which is specified by the parity check matrix $\mathbf{H}$ with the following structural properties: (1) each row has ρ 1's; (2) each column has γ 1's; (3) the number of 1's in common between any two columns is no greater than 1; (4) both ρ and γ are small compared with the length of the code n and the number of rows ($m \geq n - k$) in $\mathbf{H}$ [Gal63].

The density of $\mathbf{H}$ is defined as $r = \rho/n$. The code rate can be computed as

$$R = \frac{k}{n} \geq \frac{n-m}{n} = 1 - \frac{\gamma}{\rho} \tag{2.13}$$

where the matrix $\mathbf{H}$ may not have full rank. The property of a sparse distribution of ones in the parity check matrix is amenable to operate iterative decoding algorithm on a Tanner graph as will be shown in Section 2.2. The LDPC codes in Definition 2.1.2 refers to Gallager's originally analysed regular codes which have parity check matrices of constant row weight and constant column weight. Normally, an ensemble (n, γ, ρ) of LDPC codes refers to those LDPC codes with length n, column weight γ and row weight ρ. According to [RSU01], the average behaviour of all codes of one ensemble (n, γ, ρ) concentrates around the expected behaviour when n is large enough. In [Gal63] Gallager pointed out that for a specific ensemble of regular LDPC codes, d_{min}/n for almost all codes in the ensemble is lower-bounded by a fixed function of the row and column weights. This phenomenon already indicates good performance; by contrast, the fraction d_{min}/n of binary BCH codes approaches zero with increasing n. For LDPC codes, the parameter d_{min} links to the error floor of the code while its iterative decoding performance is more decided by the girth (the length of the shortest cycle) in the Tanner graph (we shall introduce in Section 2.2) corresponding to the $\mathbf{H}$ structure. If all the columns or all the rows of $\mathbf{H}$ do not have the same weight, an LDPC code is said to be irregular. [LMSS01,RSU01] showed that optimized irregular LDPC codes could provide better performance than regular LDPC codes.

2.1.4.1 Construction of LDPC codes

Construction of parity check matrices for LDPC codes can be classified into two main categories: random and algebraic construction.

Random (or pseudorandom) constructions, also called computer-generated Gallager LDPC codes, are based on the generation of a parity check matrix (or parts of a parity check matrix) randomly filled with 0s and 1s. There are some constraints in the computer generation process due to the properties of LDPC codes, *e.g.*, making $\mathbf{H}$ fit a desired row weight and column weight distribution. Some computer randomly generated LDPC codes have good performance [Mac99b, Mac99a, RSU01, RU01], and even approach the random-coding bounds predicted by Shannon [RSU01], which is in accord with the description in [CG90] that a

capacity-approaching code with a practical decoding algorithm requires the code with large block length and a high degree of structure, however it should also appear random to a channel. There are some disadvantages of the random construction method: post-processing is necessary to eliminate cycles; lack of structure in **H** makes encoding complex and makes it difficult for performance analysis and hardware implementation.

In the algebraic constructions, the **H** matrices are produced based on certain mathematical foundations, such as finite geometries [KLF01] and circulant permutation matrices [TSS+04, CXDL04]. It was pointed out in [LC04] that for length up to 1057, finite-geometry codes outperform their corresponding randomly generated LDPC codes although for longer length randomly generated codes perform better with sum-product algorithm decoding. [LCZ+06] shows that quasi-cyclic (QC) LDPC have encoding advantage over other types of LDPC codes because QC LDPC can be encoded with simple shift registers with linear complexity based on generator matrices. Algebraic construction guarantees the designed code with certain properties and is easier for hardware implementation. However, due to the constrains of design structure the choice of rate value and length are limited.

2.1.4.2 LDPC decoding

LDPC codes can be decoded in several ways, such as majority-logic (MLG) decoding, bit-flipping (BF) decoding, weighted bit-flipping (WBF) decoding, a posteriori probability (APP) decoding and iterative decoding based on belief propagation (known as sum-product algorithm (SPA)) [LC04]. APP and SPA both deliver good error performance among the five types of decoding method of LDPC codes while APP decoding is computationally intractable [LC04]. In this thesis, we shall focus on SPA decoding since SPA not only gives the best error performance but also is practically implementable in terms of decoding complexity. Because SPA is highly linked to factor graph theory, the detailed decoding algorithm and relative analysis shall be presented in Section 2.2.

2.2 Factor Graph Theory and LDPC Decoding

In the work of [AM00], a "generalized distributive law" was described to facilitate the marginalization of an arbitrary function through its factorization into a number of simpler factors. The so-called "marginalize product-of-functions" (MPF) problem actually gets involved in many

algorithms of communications and signal processing. The theory of the factor graph was introduced in the work of [KFL01] which provides a visual approach to solve the MPF problem by the sum-product algorithm (SPA) which involves message passing on the corresponding factor graph.

This section is organized as follows. We present a concise overview of factor graph theory, including the factor graph presentation for linear block codes and *maximum a posteriori* (MAP) decoding. Then we focus on the LDPC decoding algorithm, since this forms the basic building blocks for this thesis. Last we also briefly introduce some analysis methods for LDPC codes.

2.2.1 Introduction

Algorithms in coding, signal processing and artificial intelligence often deal with complicated functions of many variables. The given complicated function could be regarded as a global function and thus be exploited as a product of several "local" functions with a subset of the variables. The visualization for such a factorization can be approached by a factor graph, a bipartite graph that expresses the relations among functions and the relation between the local function and its variables. The sum-product algorithm (SPA) is a generic message-passing algorithm following a simple computational rule which operates on its corresponding factor graph and aims to compute various marginal functions associated with the global function. A wide variety of applications which may be seen as specific instances of the SPA operating on a factor graph include: Pearl's "belief propagation" on a Bayesian network [Pea88]; the iterative decoding in a distributed fashion, called "code networks", for various forms of concatenated coding schemes [BDMP98]; the Viterbi algorithm for ML decoding of block and convolutional codes [For73, Wol78]; the BCJR algorithm for bit-by-bit MAP decoding applied to block or convolutional codes [BCJR74]; the iterative decoding of turbo codes and LDPC codes [FK96, Fre98, LS88, MN95, Mac99b, MMC98]; trellis-based schemes in turbo equalization [Fla05]; and electromyographic signal analysis [Koc07]. In the following overview of factor graph theory, we shall use the factor graph model introduced in [KFL01,FKLW97], as opposed to the Forney-style factor graphs presented in [For01, Loe04].

Let $v_1, v_2, \ldots, v_n$ represent a collection of variables. For each i, v_i lies in some finite domain A_i. Let $g(v_1, \ldots, v_n)$ represent a function of these variables with domain

$$S = A_1 \times A_2 \times \cdots \times A_n. \tag{2.14}$$

The domain S of g is called the configuration space for the given collection of variables and each element in S is a specific configuration of the variables. For the work in this thesis, we shall take domain $A_i = \mathbb{R}$. The 'sum' and 'product' mentioned later shall refer to the ordinary addition and multiplication of real numbers.

There are n marginal functions $g_i(v_i)$ associated with every function $g(v_1, \ldots, v_n)$ in which

$$g_i(v_i) = \sum_{\sim\{v_i\}} g(v_1, \ldots, v_n) \tag{2.15}$$

The notation $\sum_{\sim\{v_i\}}$, which is called the summary for v_i of g, denotes the summation over all the variables except v_i. Suppose that $g(v_1, \ldots, v_n)$ factors into a product of several local functions, each having a subset of $\{v_1, \ldots, v_n\}$ as arguments:

$$g(v_1, \ldots, v_n) = \prod_{j \in J} f_j(V_j) \tag{2.16}$$

where J is a discrete index set and each V_j is a subset of the variables $\{v_1, \ldots, v_n\}$.

Now we come to the definition of a factor graph [KFL01]: A factor graph is a bipartite graph that expresses the structure of the factorization (2.16). A factor graph has a variable node for each variable v_i and a factor node for each local function f_i, and an edge connecting variable node v_i to factor node f_i if and only if v_i is an argument of f_i.

Example 2.2.1. Suppose we have a factorization as follows

$$g(v_1, v_2, v_3, v_4, v_5) = f_A(v_1) f_B(v_2) f_C(v_1, v_2, v_3) f_D(v_3, v_4) f_E(v_3, v_5) \tag{2.17}$$

The factor graph for g is illustrated in Figure2.2.

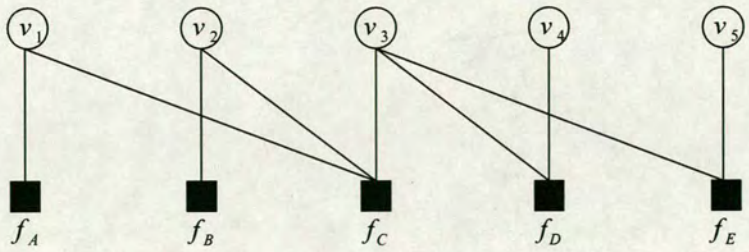

Figure 2.2: *Factor graph example.*

Sum-product Algorithm:

Message-passing refers to the algorithm in which messages are passing along the edges of a factor graph. In the case of the sum-product message-passing algorithm, *sum-product* refers to the rule used to calculate the updated messages. The resulting algorithm is also called the sum-product algorithm.

Genarally speaking, the first step in SPA is that all pendant (degree-1) nodes send out messages to their adjacent nodes in which each variable node sends out a unit message and each factor node sends out a description of itself (the function f_i). Then each node x (variable or factor) waits for messages to arrive from all but one of its neighbouring node (say y). When those messages arrive, x computes and sends the message to the remaining node y. Then x waits for a returning message from y. Once the message is received from y, x is able to send the computed messages to all of its neighbouring nodes but y in turn. The algorithm terminates once every edge in the graph has experienced the message passing schedule in both directions. Since there are various sums and products involved in the computation, the algorithm is referred as the sum-product algorithm.

The sum-product update rule for non-pendant nodes:

- **Variable to factor:** The message sent by a variable node v_i to a factor node f_i is the product of all the incoming messages from its neighbouring nodes except f_i.
- **Factor to variable:** The message sent by a factor node f_i to a variable node v_i is the product of the local function at f_i with all the incoming messages from its neighbouring nodes except f_i.

Example 2.2.2. Let q_{lj} denote the message sent from the variable node v_l to the factor node f_j; let σ_{jl} denote the message sent from the factor node f_j to the variable node v_l. Also, $M(j)$ denotes the set of indices of the variable nodes adjacent to the factor node f_j; $N(l)$ denotes the set of indices of the factor nodes adjacent to the variable node v_l. The messages computed by SPA could be described as follows:

- **variable to factor:**

$$q_{lj} = \prod_{t \in N(l) \setminus j} \sigma_{tl} \tag{2.18}$$

- **factor to variable:**

$$\sigma_{jl} = \sum_{\sim\{v_l\}} \left(f_j(V) \prod_{t \in M(j)\backslash l} q_{tj} \right) \tag{2.19}$$

in which $V = M(j)$ is the set of arguments of the function f. One example of the sum-product rule is shown in Figure 2.3. Equation (2.18) says that the operations at the variable node are simply the product of the incoming messages from the nearby function nodes since there is no local function involved. While equation (2.19) says that the operations at the check node relate to the function multiplications and are then followed by an application of the summary operator.

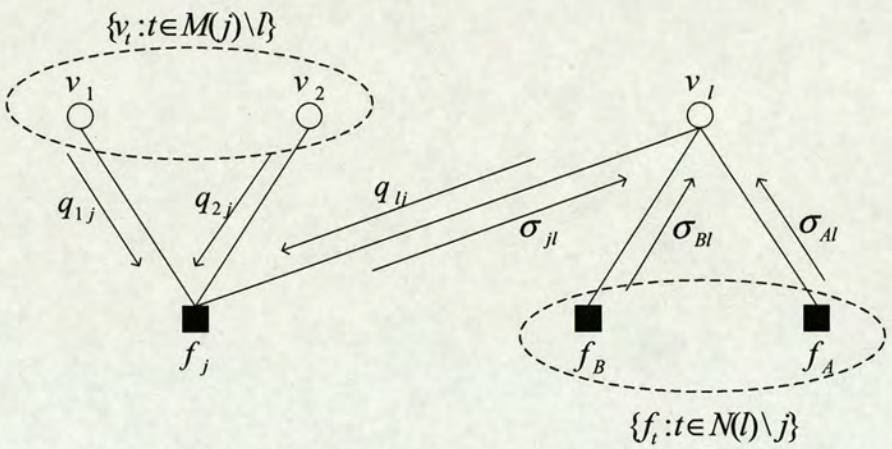

Figure 2.3: *Illustration for the sum-product rule.*

We now make two remarks upon the operation of the sum-product rule:

Remark 2.2.1. For the case of variable nodes of degree two, the message is simply passed through the node.

Remark 2.2.2. The prerequisite to obtain the correct computation of a marginal function is that the corresponding factor graph is cycle-free. A cycle-free factor graph can be equally converted to a rooted tree structure [KFL01] where the same information will not flow twice along one edge. On the contrary, for a factor graph with cycles, the same information may be passed multiple times on a given edge since there is no natural termination condition. In this way, the iterative algorithm may result in a failure of convergence and may not provide the correct marginal function computation. However, the SPA could still work locally in a factor graph with cycles. The corresponding computational rule is the same as before except that only the most recently passed message along any edge is used. A stop criterion is at the discretion of

the algorithm designer and estimations of the marginal functions are obtainable provided the girth (minimum cycle length) in the factor graph is not too small. For a code whose graph is not cycle-free, the presence of short cycles should be avoided; if the girth of its corresponding factor graph is very large, the loop-free approximation can be made [RU01]. This sub-optimal approach is known to give good performance in some applications such as Turbo codes and LDPC codes.

2.2.2 Factor graph for linear block codes

In the work of [Wil89], "behavioural" modeling is defined in set-theoretic terms by specifying the validity of the particular configurations of variables. This approach can be represented by a factor graph where the characteristic functions for the given behaviours are all indicated.

Let $v_0, v_1, \ldots, v_{n-1}$ be a collection of variables with configuration space $S = A_0 \times A_1 \times \cdots \times A_{n-1}$. A behaviour means any subset V of S. The elements of V are the valid configurations. If the domain of each variable is some finite alphabet A, the configuration space is the n-fold Cartesian product $S = A^n$. Thus, a block code of length n over A is a behaviour $V \subset S$ while the codewords refer to the valid configurations. The characteristic function for a binary linear block code V is defined as

$$\chi_V(\mathbf{v}) = \chi_V(v_0, v_1, \ldots, v_{n-1}) = \text{Ind}\,[(v_0, v_1, \ldots, v_{n-1}) \in V] \tag{2.20}$$

Example 2.2.3. We consider a (6, 3) linear block code defined by the parity check matrix

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \tag{2.21}$$

Here, V is the set of all binary 6-tuples $\mathbf{v} = (v_0, v_1, \ldots, v_5)$ that satisfies the parity-check constraints $\mathbf{v} \cdot \mathbf{H}^T = \mathbf{0}$. The indicator function Ind$[\cdot]$ takes the value 1 if the argument is true, and otherwise it takes the value 0. Therefore we have the characteristic function

$$\begin{aligned} \chi_V(v) &= \text{Ind}\,[(v_0, v_1, \ldots, v_5) \in V] \\ &= \text{Ind}\,[v_0 \oplus v_1 \oplus v_4 = 0] \cdot \text{Ind}\,[v_1 \oplus v_2 \oplus v_5 = 0] \\ &\quad \cdot \text{Ind}\,[v_0 \oplus v_2 \oplus v_3 = 0] \end{aligned} \tag{2.22}$$

2.2.3 Factor graph for MAP decoding

The probability distribution is an important class of functions that can be represented through a factorization of the variable's joint probability mass or density function. In probabilistic modelling of systems, statistical dependency among variables could be expressed by a factor graph.

Now we consider a transmission system in which $\mathbf{v} = (v_0, \ldots, v_{n-1})$ represents an n-bit codeword and $\mathbf{y} = (y_0, \ldots, y_{n-1})$ represents the receive sequence at the end of the transmission channel. (Here we ignore the modulation process which normally takes place in a standard transmission model.) The joint a posteriori probability (APP) for the transmitted codeword is

$$P(\mathbf{v}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{v})P(\mathbf{v})}{p(\mathbf{y})} \tag{2.23}$$

where we assume that the probability mass function (pmf) $P(\mathbf{v})$ is uniform over all valid codewords. Also we assume that the channel is memoryless:

$$p(\mathbf{y}|\mathbf{v}) = \prod_{i=0}^{n-1} p(y_i|v_i) \tag{2.24}$$

Since the observed sequence $\mathbf{y}$ is fixed for one decoding procedure, $P(\mathbf{v}|\mathbf{y})$ is regarded as a function of $\mathbf{v}$ only.

Next we combine the "behavioural" modeling for the linear block code presentation with the probabilistic modeling of systems for the APP of the codewords given the received output of a channel. The global function for the transmission system over the variable $\mathbf{v}$ can be described as the product of equation (2.20) and equation (2.24):

$$g(\mathbf{v}) = \chi_V(\mathbf{v}) \prod_{i=0}^{n-1} p(y_i|v_i) \tag{2.25}$$

Example 2.2.4. If V is the set of codewords as in the example 2.2.3, the global function is

$$g(\mathbf{v}) = \text{Ind}\,[v_0 \oplus v_1 \oplus v_4 = 0] \cdot \text{Ind}\,[v_1 \oplus v_2 \oplus v_5 = 0] \cdot \text{Ind}\,[v_0 \oplus v_2 \oplus v_3 = 0] \cdot \prod_{i=0}^{n-1} p(y_i|v_i) \tag{2.26}$$

The corresponding factor graph is shown in Figure 2.4, in which $p(y_i|v_i)$ is shown as pendant nodes.

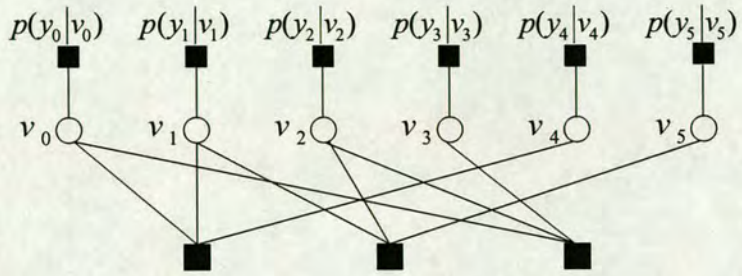

Figure 2.4: *Factor graph for the linear block codes together with the joint APP.*

2.2.4 Sum product algorithm for LDPC decoding

The sum-product algorithm is a decoding method for LDPC codes invented by Gallager [Gal63]. However, the full power of LDPC codes was not realized until Tanner introduced graphs [Tan81] (those factor graphs for LDPC codes are thus called Tanner graphs) to describe those codes. The term message-passing decoder comes from the message-passing nature of the SPA. One important feature for the message-passing decoder is that the decoding complexity grows linearly with the length of the code. It needs to be pointed out that the message-passing algorithm is sub-optimal if underlying cycles exist in a factor graph. However, SPA operating on Tanner graphs with cycles still gives good performance when the girth in the graph is large which necessarily requires that the block length of the code is large. This phenomenon is due to the fact that when the length of the cycle grows the independency of the messages also grow. In terms of the convergence behaviour of iterative decoding, an EXIT chart [tB01] is a useful tool for analysis and code design. Although the convergence of the iterative decoding algorithms is still not well understood, SPA based on belief propagation is believed to be an extremely powerful tool for decoding LDPC codes [WLK95, KF98, FMI99, LKF00, LK00, KLF00a, KLF00b, KLF00c, RSU01, RU01, CFRU01, KLF01, Ard04].

2.2.4.1 Probability domain

Conventionally, SPA is presented in probabilities. Suppose $p_l(0) = P(v_l = 0)$ and $p_l(1) = P(v_l = 1)$ are the prior probabilities of $v_l = 0$ and $v_l = 1$ respectively which are obtained by channel observation. Next we will present two update rules in SPA for LDPC decoding based

on (2.18) and (2.19).

- **bit node v_l to check node f_j:**

$$q_{lj}(v_l) = p_l(v_l) \prod_{t \in N(l) \setminus j} \sigma_{jl}(v_l) \tag{2.27}$$

- **check node f_j to bit node v_l:**

$$\sigma_{jl}(v_l) = \sum_{\sim\{v_l\}} \left(\prod_{t \in M(j) \setminus l} q_{tj}(v_t) \cdot \mathrm{Ind}\left\{ \sum_{t \in M(j)}^{\oplus} v_t = 0 \right\} \right) \tag{2.28}$$

Equation (2.27) says that the variable node processes the incoming messages from all but one of its neighbouring check nodes and then sends the updated information about itself back to that check node.

Equation (2.28) says that the function of a check node is to force its neighbours to satisfy an even parity. The outgoing message from a check node is the result of all the incoming messages except from the variable node which will receive the outgoing message. And the message sent from the check node indicates its belief towards the destination variable node. By presenting the message in the Log domain in the next subsection, we could easily see that the reliability of the outgoing messages is even less than that of the least reliable incoming messages.

The above two update rules lay the foundation for SPA decoding in the probability domain which are a straight derivation from the sum-product rule of (2.18) and (2.19) respectively. However from the computational point of view, messages in probability domain are not computationally convenient since there are lots of multiplications involved.

2.2.4.2 SPA for LDPC decoding in terms of log-likelihood ratios

The log-likelihood ratio (LLR) is the most preferable message type used in a practical application for its precise representation of the probability values, and also for its simplicity in update rules and computer implementations. This version of SPA is based on updating extrinsic information by each iteration. We will follow the presentation in [Fla05] to develop SPA decoding for LDPC codes in the log domain.

Let V be in $GF(2)$ with the elements $\{0,\ 1\}$. We shall refer each element v of V as a *bit*. The

LLR is defined as

$$L_V(v) = \ln\left(\frac{P_V(v=0)}{P_V(v=1)}\right) \tag{2.29}$$

The LLR is denoted as "soft" value, or the L-value of the variable V. The sign of $L_V(v)$ is regarded as the hard decision and the magnitude $|L_V(v)|$ is regarded as the reliability of the hard decision. For simplicity, we shall use $v(0)$, $v(1)$ and $L(v)$ to denote $P_V(v=0)$, $P_V(v=1)$ and $L_V(v)$ respectively.

Soft channel outputs

Suppose the following output is observed from a Rayleigh fading channel

$$y_i = a_i \cdot \widetilde{v_i} + n_i \tag{2.30}$$

where $\widetilde{v_i}$ is the BPSK modulated bit, *i.e.* $\widetilde{v_i} = 1 - 2 \cdot v_i$, $a_i \in \mathbb{C}$ represents the Rayleigh fading channel gain under current conditions and $n_i \in \mathbb{C}$ represents the additive white Gaussian noise with variance σ^2 in each of the real and imaginary direction. The pdf of the channel output y_i conditioned on $\widetilde{v_i}$ is

$$p(y_i|\widetilde{v_i}) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{|y_i - a_i \cdot \widetilde{v_i}|^2}{2\sigma^2}\right) \tag{2.31}$$

Thus, we can obtain the corresponding LLR with regard to the input bit v_i (suppose $p(v_i = 0) = p(v_i = 1) = 0.5$):

$$L(v_i) = L(v_i|y_i) = L(y_i|v_i) = \ln\frac{p(y_i|v_i=0)}{p(y_i|v_i=1)} = \ln\frac{p(y_i|\widetilde{v_i}=1)}{p(y_i|\widetilde{v_i}=-1)} = \frac{2\Re\{a_i^* \cdot y_i\}}{\sigma^2} \tag{2.32}$$

When $a_i = 1$ and $n \in \mathbb{R}$, this special condition corresponds to an AWGN channel. In this case,

$$L(v_i) = \frac{2y_i}{\sigma^2} \tag{2.33}$$

Bit node v_l to check node f_j

Suppose $q(v_l)$ ($v_l \in \{0,\ 1\}$) is the message to be passed from a bit node v_l to a check node f_j with the input messages σ_t from the rest of the check nodes and APP of the codewords $p(y_i|v_i)$

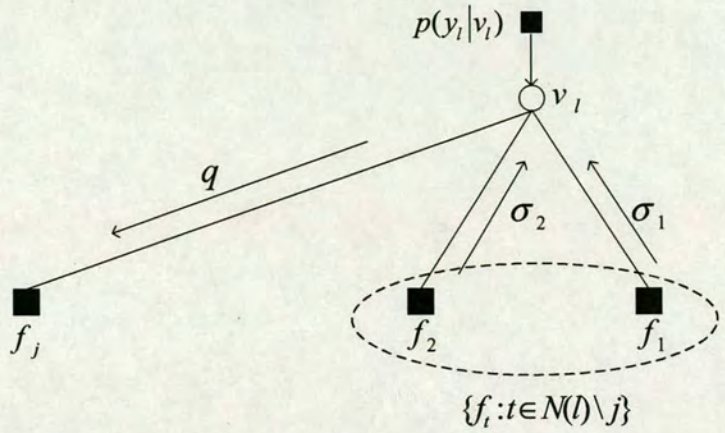

Figure 2.5: *Demonstration for the message passed from bit node v_l to check node f_j.*

as depicted in Figure 2.5. Then based on (2.27),

$$\begin{aligned} \ln\left(\frac{q(0)}{q(1)}\right) &= \ln\left[\left(\frac{p(y_i|v_i=0)}{p(y_i|v_i=1)}\right) \cdot \prod_{t\in N(l)\setminus j}\left(\frac{\sigma_t(0)}{\sigma_t(1)}\right)\right] \\ &= \ln\left(\frac{p(y_i|v_i=0)}{p(y_i|v_i=1)}\right) + \sum_{t\in N(l)\setminus j} \ln\left(\frac{\sigma_t(0)}{\sigma_t(1)}\right) \end{aligned}$$

Therefore

$$L(q) = L(v_i) + \sum_{t\in N(l)\setminus j} L(\sigma_t) \tag{2.34}$$

where $L(v_i)$ is the channel observed L-value which is also referred to as the *intrinsic* message for every variable node v_i and $\sum_{t\in N(l)\setminus j} L(\sigma_t)$ is referred as the *extrinsic* message which is due to the structure of the code and could be improved through exploiting redundancy in the code iteration by iteration. It can be seen that the reliability of a variable node improves as it receives positive (mostly correct) information about itself.

Check node f_j to bit node v_l

Suppose $\sigma(v_l)$ ($v_l \in \{0, 1\}$) is the message to be passed from a check node f_j to a bit node v_l with the input messages q_t from the rest of the code bits as depicted in Figure 2.6. The probability domain equation is

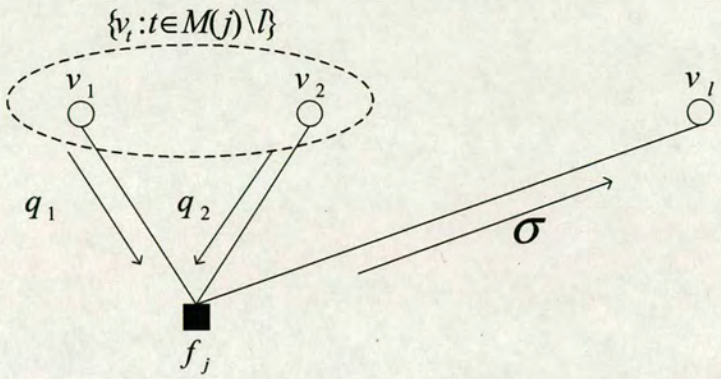

Figure 2.6: *Demonstration for the message passed from check node f_j to bit node v_l.*

$$\sigma(v_l) = \sum_{\sim\{v_l\}} \left(\prod_{t\in M(j)\backslash l} q_t(v_t) \cdot \mathrm{Ind}\left\{ \sum_{t\in M(j)}^{\oplus} v_t = 0 \right\} \right) \tag{2.35}$$

We have

$$\begin{aligned} \prod_{t\in M(j)\backslash l} [q_t(0) + q_t(1)] &= \sum_{\sum_{t\in M(j)\backslash l}^{\oplus} c_t = 0} \left[\prod_{t\in M(j)\backslash l} q_t(v_t) \right] \pm \sum_{\sum_{t\in M(j)\backslash l}^{\oplus} c_t = 1} \left[\prod_{t\in M(j)\backslash l} q_t(v_t) \right] \\ &= \sigma(0) + \sigma(1) \end{aligned} \tag{2.36}$$

Let $\delta'(v)$ denote the specific soft value of the binary bit v:

$$\delta'(v) = \frac{v(0) - v(1)}{v(0) + v(1)} = \tanh\left(\frac{L(v)}{2}\right) \tag{2.37}$$

Together with (2.36), we can obtain:

$$\delta'(\sigma) = \prod_{t\in M(j)\backslash l} \delta'(q_t) \tag{2.38}$$

And thus,

$$\tanh\left(\frac{L(\sigma)}{2}\right) = \prod_{t\in M(j)\backslash l} \tanh\left(\frac{L(q_t)}{2}\right) \tag{2.39}$$

which lead to the expression:

$$L(\sigma) = \ln \frac{1 + \prod_{t \in M(j) \backslash l} \tanh\left(\frac{L(q_t)}{2}\right)}{1 - \prod_{t \in M(j) \backslash l} \tanh\left(\frac{L(q_t)}{2}\right)} \tag{2.40}$$

It is worth mentioning that in [HOP96], the "box-plus" operation (denoted by $\boxplus$) was introduced for the special algebra for L-values. For two independent random variables v_1 and v_2

$$\begin{aligned} L(v_1) \boxplus L(v_2) &= L(v_1 \oplus v_2) = \ln \frac{1 + e^{L(v_1)} e^{L(v_2)}}{e^{L(v_1)} + e^{L(v_2)}} \\ &\approx \operatorname{sign}(L(v_1)) \cdot \operatorname{sign}(L(v_2)) \cdot \min(|L(v_1)|, |L(v_1)|) \end{aligned} \tag{2.41}$$

with the additional rules

$$L(v) \boxplus \infty = L(v), \;\; L(v) \boxplus -\infty = -L(v) \tag{2.42}$$

and

$$L(v) \boxplus 0 = 0 \tag{2.43}$$

Therefore, (2.40) could be expressed as

$$L(\sigma) = \sum_{t \in M(j) \backslash l}^{\boxplus} L(q_t) \approx \left(\prod_{t \in M(j) \backslash l} \operatorname{sign}(L(q_t)) \right) \cdot \min_{t \in M(j) \backslash l} |L(q_t)| \tag{2.44}$$

which indicates that the reliability of the outgoing extrinsic message from a check nodes is less than that of the least reliable incoming message. That is to say, the reliability of the extrinsic messages is decreased at the check nodes since the role of a check node is to "force" the adjacent variable nodes to satisfy an even parity check. From the perspective of the variable nodes, the *a posteriori* probabilities are still improved though.

SPA for decoding LDPC codes in terms of extrinsic information [LFKL00]

Let $\mathbf{H} = [h_{jl}]$ denote a $J \times L$ parity check matrix with each entry h_{jl} ($0 \leq j < J$ and $0 \leq l < L$) and N_{iter} denotes a maximun iteration number. The scalar z_{lj} denotes the LLR for the message passed from the bit node v_l to the check node f_j for each $h_{jl} = 1$; ε_{jl} denotes the LLR for the message passed from f_j to v_l for each $h_{jl} = 1$; $L(v_l)$ denotes the channel observed LLR for v_l; and z_l denotes the LLR for the summary of the global function for v_l.

Initialization:

- For each $l = 1,\ 2, \ldots,\ L$,

$$L(v_l) = \frac{2\Re\{a_l^* \cdot y_l\}}{\sigma^2} \tag{2.45}$$

- For each $l,\ j$ satisfying $h_{jl} = 1$

$$\varepsilon_{jl} = 0 \tag{2.46}$$

Main Loop: For $i_t = 1$ **to** ∞

- For each check node f_j where $h_{jl} = 1$

$$z_{lj} = L(v_l) + \sum_{t \in N(l) \backslash j} \varepsilon_{tl} \tag{2.47}$$

- For each code bit v_l where $h_{jl} = 1$

$$\varepsilon_{jl} = \ln \frac{1 + \prod_{t \in M(j) \backslash l} \tanh\left(\frac{z_{tj}}{2}\right)}{1 - \prod_{t \in M(j) \backslash l} \tanh\left(\frac{z_{tj}}{2}\right)} \tag{2.48}$$

- Compute for each code bit v_l

$$z_l = L(v_l) + \sum_{t \in N(l)} \varepsilon_{tl} \tag{2.49}$$

$$\hat{v}_l = \begin{cases} 0 & \text{if } z_l \geq 0 \\ 1 & \text{if } z_l < 0 \end{cases}$$

If $\hat{\mathbf{v}}\mathbf{H}^T = \mathbf{0}$ **or** $i_t = N_{it}$ **then** ***break;***

Endfor

2.2.4.3 The analysis of SPA decoding for LDPC codes

First, we shall have a look at the principle of iterative decoding. At the input of the iterative decoder, there are two types of knowledge about the transmitted codeword as shown in Figure 2.7: the intrinsic information (the information from the channel) which corresponds to $L(v_l)$

in equation (2.47) and the extrinsic information (the information from the previous decoding iteration) which corresponds to ε_{tl} in equation (2.47). Based on these two types of information, the decoder attempts to improve the knowledge about the transmitted codeword and at the same time, the corresponding extrinsic information will be used in the following iteration.

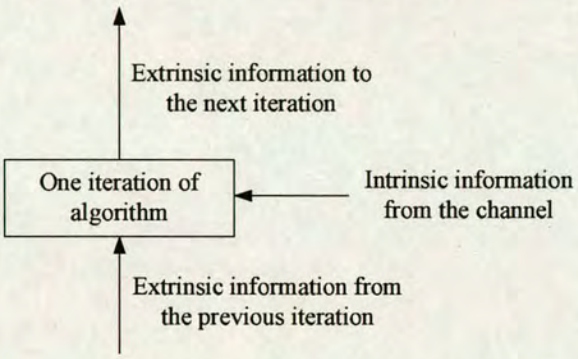

Figure 2.7: *The principle of iterative decoding.*

Decoding tree

The whole decoding process could be presented by means of a decoding tree. Figure 2.8 shows a depth-one decoding tree for a regular (3, 6) LDPC code. The outgoing message from a variable node v_l to the next iteration is comprised of the intrinsic information from the channel and the extrinsic information from the other two neighbour check nodes f_A and f_B. Consider the update at the check nodes. The outgoing message from f_A or f_B to v_l is comprised of five incoming message to f_A or f_B from their neighboring variable nodes except v_l. This completes one iteration process. Continuing in this way, the decoding tree of any depth is obtained. As long as the factor graph can be presented by a decoding tree, the messages in the decoding tree of any depth are independent. If the factor graph has girth of l, the messages in the decoding tree are independent up to depth $\frac{l}{2}$. Correspondingly the independent assumption is true up to $\frac{l}{2}$ iterations and will become an approximation for further iterations.

An intuition for the better performance of irregular LDPC codes is: variable nodes with more neighbours have more chances of getting the correct bits; check nodes with less neighbours have less chances of getting wrong; thus variable nodes with high degree attain faster convergence to correct bit values and then their neighbor check nodes get the correct bit value faster which result in that variable nodes with small degree get the correct values faster.

Density evolution

Density evolution was proposed by Richardson and Urbanke in 2001 for the analysis of LDPC

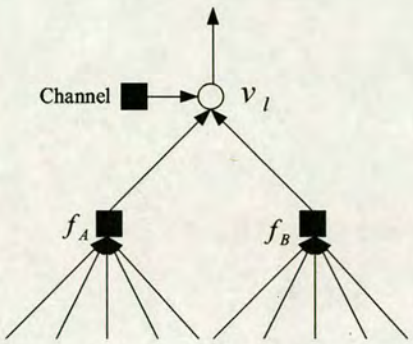

Figure 2.8: *The depth-one decoding tree for a regular (3, 6) LDPC code.*

codes [RU01]. The basic idea is that tracking the probability density during the iterations could give a good picture of the actual behaviour of the decoding algorithm. Density evolution is a two-phase computation algorithm which aims to track the average fraction of incorrect messages passed in each iteration. Together with Fourier transform techniques, density evolution can be used to obtain asymptotic thresholds below which belief propagation could decode the code successfully, and above which belief propagation can not provide successful decoding. The *threshold* of a LDPC code is defined as the worst channel condition for which the message error rate converges to zero as the number of iterations of this LDPC code approaches infinity. If the threshold of a code conforms to the Shannon limit, the code is said to be capacity-approaching. Capacity-approaching irregular LDPC codes could be designed by following the evolution of the density of messages in the decoder, that is, to set some irregularity in the code, check the performance with density evolution and vary the irregularity in the code until the desired performance is achieved. But performing density evolution accurately is a computationally demanding task. A relatively easier approach for designing irregular codes is by using EXIT charts [tB00, tB01].

Extrinsic information transfer (EXIT) charts

Another popular approach for analyzing iterative decoders, including codes with complicated constituent codes, is to use extrinsic information transfer (EXIT) charts [tB00, EGH01, tB01]. Instead of tracking the density of messages, an EXIT chart tracks the evolution of the mutual information between messages and decoded bits iteration by iteration [tB00]. Tracking the evolutions of other parameters like SNRs of the extrinsic messages [EGH01], error probabilities [AK04], etc., are seen as EXIT-like charts. Compared to density evolution which is computationally intensive, the EXIT chart is computationally simpler and applicable to many iterative decoders.

2.3 Cooperative Communications

In this section, we start with an introduction to wireless fading channels. Then we review some basic ideas on cooperative relaying strategies and user cooperation schemes. Last we briefly address three user cooperation schemes whose concepts relate to the problem studied in this thesis.

2.3.1 Wireless fading channel and capacity

As a medium for reliable communications, the wireless radio channel has several challenges due to noise, interference and other channel impediments caused by unpredictable user movements and environmental dynamics, which lead to the phenomenon of *fading*. In the following, we shall briefly review the basic characteristics of a wireless fading channel.

Observing the variations of the channel strength over time and frequency, there are two types of fading: large-scale fading and small-scale fading. Large-scale fading is caused by the path loss effect and shadowing effect which occur over a relatively large distance (a distance of the order of the cell size) [Gol05]. The large-scale fading is more relevant to issues such as cell-site planning [TV05], thus it is not the focus of this thesis. Small scale fading, which is also referred to as multipath fading, is caused by the constructive and destructive interference of the multiple signal path between the transmitter and receiver. Small scale fading is more important to the design of reliable and efficient communication systems. In this thesis, *fading* refers to small scale fading.

One reasonable model of the multipath environment is Rayleigh fading when there are many objects in the environment that scatter the radio signal before it arrives at the receiver. The envelope of the received signal follows a Rayleigh distribution. Specifically we consider a discrete-time complex baseband model where the channel coefficient can be expressed as

$$a = a_r + a_i j \tag{2.50}$$

where a_r and a_i are independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance $\frac{\sigma^2}{2}$. The magnitude $|a|$ satisfies a Rayleigh distribution density function given by

$$f(x) = \frac{x}{\sigma^2} \exp\left\{\frac{-x^2}{2\sigma^2}\right\}, \quad x \geq 0. \tag{2.51}$$

And the squared magnitude $|a|^2$ is exponentially distributed with density function

$$f(x) = \frac{1}{\sigma^2} \exp\left\{\frac{-x}{\sigma^2}\right\}, \quad x \geq 0. \tag{2.52}$$

In this thesis, the value of σ is assumed to be 1.

2.3.1.1 Flat fading and frequency-selective fading

The channel coherence bandwidth B_c is a statistical measure of the range of frequencies over which the channel passes all spectral components with approximately equal gain and linear phase [Rap01]. If the coherence bandwidth of the channel is larger than the bandwidth of the signal, this type of fading is called flat fading which means that all frequency components of the signal will experience the same magnitude of fading. If the coherence bandwidth of the channel is smaller than the bandwidth of the signal, this type of fading is called frequency-selective. Since different frequency components of the signal experience decorrelated fading, frequency-selective fading leads to intersymbol interference (ISI) of the received signal. The effect of frequency-selective fading could be mitigated by employing a diversity scheme such as orthogonal frequency-division multiplexing (OFDM).

2.3.1.2 Slow fading and fast fading

The channel coherence time T_c is a statistical measure of the minimum time duration for the magnitude change of the channel to become decorrelated from its previous value. In other words, the channel fading coefficient within T_c remains roughly the same. Coherence time T_c reflects the time-scale of the variation of the channel and is inversely proportional to Doppler spread due to the movement of transmitter or receiver.

Block fading is a fading model in which the fading coefficient a remains constant over some blocklength T and is i.i.d. across different coherence periods [Gol05]. Sometimes, we also use the term *quasi-static* to describe the fading channel that the fading coefficient remains constant for a duration of the codeword and changes independently from one codeword to next. In wireless communication, channels are often categorized into fast fading and slow fading. Fast fading occurs when one codeword is spread over multiple fading blocks. On the contrary, slow fading occurs when one codeword length is less than one fading block. In reality, the terms of fast or slow fading depend not only on the environment but also on the application with regard

to the practical delay requirements.

2.3.1.3 Capacity of the independent Rayleigh fading channel

A discrete-time complex baseband Rayleigh fading channel with stationary and ergodic time-varying fading coefficient a_i and AWGN n_i with $n \sim \mathcal{CN}(0,\ 1)$ is expressed by

$$y_i = \sqrt{\rho} a_i x_i + n_i \tag{2.53}$$

where x_i is the channel input at time i, y_i is the corresponding channel output and ρ is the transmit SNR. In this thesis, we assume that the receiver always has full channel state information (CSI) while a_i is not known to the transmitter.

Suppose a long codeword is transmitted and its symbols are likely to experience all states of the channel. Hence the channel is ergodic to the codeword. Under this fast fading scenario, Shannon capacity defines the maximum data rate that can be sent over the channel with an arbitrarily small number of errors. According to [MS84], the Shannon capacity of a fading channel with receiver CSI can be expressed as

$$C = \mathbb{E}_a\{\log_2(1 + \rho|a|^2)\} \tag{2.54}$$

where $\mathbb{E}_a$ denotes the expectation with respect to the fading coefficient and C in information bits per transmission per dimension.

The opposite situation of a channel is that the fading is so slow that the fading coefficient remains constant for the whole duration of a codeword which leads to a nonzero probability that the channel capacity is below the transmission rate R. As any codeword is unable to experience all the states of the channel, this channel is nonergodic. Consequently, Shannon capacity is not valid anymore. An alternative capacity definition for a fading channel with receiver CSI is capacity with outage. Outage is defined as a condition that the capacity for a channel realization is smaller than the transmission rate R. The outage probability for Rayleigh fading is defined as

$$P_{out} = P(R > C) = P(R > \log_2(1 + \rho|a|^2)) = P\left(|a|^2 < \frac{2^R - 1}{\rho}\right). \tag{2.55}$$

2.3.1.4 Coding for wireless channel

Codes designed for a fading channel are mainly based on employing an AWGN code combined with an interleaver to combat burst errors introduced by deep fades. But the criterion for coding changes to provide diversity. [Big05] shows that for the independent Rayleigh fading channel with high SNR, a sensible criterion for the selection of a code is the maximization of the minimum Hamming distance between any two codewords. Various diversity schemes in the wireless environments, *e.g.*, space, time or frequency diversity can be seen as an implementation of the simple repetition code. Coding across different channel realizations provides a certain amount of diversity, which counters the effects of multipath fading [KH00].

2.3.2 Cooperative communication strategy

This part gives a brief review of some basic concept in the area of cooperative communication. We shall begin with the classical relay channel then introduce three user cooperation schemes relating to the problems studied in this thesis.

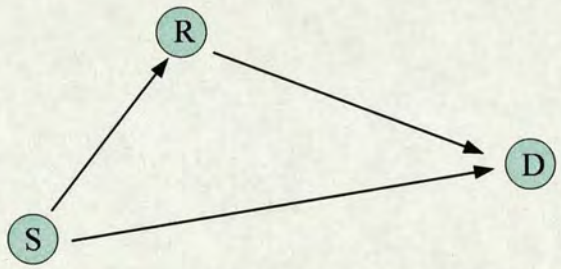

Figure 2.9: *A basic three-node relay model.*

2.3.2.1 Relay Model

The genesis of cooperative communication can be traced back to the work on the relay channel by Cover and El Gamal [CG79]. A relay model shown in Figure 2.9 consists of one source (S), one relay (R) and one destination (D). It was shown in [CG79] that the information theoretic capacity of this channel is bounded by the minimum rates of transmission of the constituent broadcast channel and MAC. In many cases, the overall capacity is better than the individual capacity between S and D.

According to operations at the relay, relaying schemes can be classified into three major categories: *amplify-and-forward* (AF) [LTW04], where the relay simply rescales and retransmits

the received signal, *decoded-and-forward* (DF) [LTW04], where the signal is fully decoded and then re-encoded at the relay before transmission to the destination, and *compress-and-forward* (CF) [XLC04], where the relay compresses the received signal with certain distortion.

For the AF scheme, the relay transmitted signal can be expressed as

$$x_r = \beta y_r = \beta(\sqrt{\rho} a_{sr} x_s + n_r) \tag{2.56}$$

where x_s denotes the source broadcasted signal; y_r denotes the relay received signal from source in the broadcast phase; a_{sr} denotes the fading coefficient in the source-relay link; n_r denotes the MAC receiver noise at the destination; β denotes the amplifying gain with the expression

$$\beta = \sqrt{\frac{1}{\rho |a_{sr}|^2 + 1}} \tag{2.57}$$

which is in order to make x_r a unit average power [LTW04]. Or β can be set as a fixed value [HA04]. Although the noisy signal is amplified and transmitted at the relay, the destination still receives two independently-faded versions of the original signal and is supposed to make better decisions for the transmitted symbols. Some hybrid protocols and power allocation issues in non-orthogonal AF scheme are discussed in [KTMG08]; relay selection issues for AF systems with interference are studied in [Kri09]. The drawback for the AF scheme is that sampling, amplifying and retransmitting analog values can be technologically non-trivial [LTW04].

For the DF scheme, the simplest scheme is to employ repetition code at the relay [LTW04], where

$$x_r = x_s \tag{2.58}$$

if the codeword is correctly decoded. Obviously, forwarding an erroneous codeword from relay to destination may lead to error propagation which is worse than non-cooperation. To avoid this error propagation phenomenon, Laneman in [LTW04] proposed an adaptive relaying scheme where only successfully decoded codewords (indicated by CRC check) will be forwarded to the destination.

For the CF scheme, the relay node may employ standard quantization or a source coding technique [XLC04, SHMX06] to process the received signal before transmission to the destination. CF, especially *Wyner-Ziv cooperation* [HL06] for optimal compression, is preferable if the noisy copy at the relay is highly correlated to the original source. The comparison of perfor-

mance for DF and CF in [KGG05] shows that the achievable rate of DF increases when the relay moves towards the source and reaches a maximum when the relay is near the source; as for CF, the achievable rate increases as the relay moves towards the destination and the rate reaches a maximum value when the relay is near the destination.

2.3.2.2 Cooperative Communication

Cooperative diversity is an active research topic in recent years. There is a staggering number of papers published under this theme. Here we only introduce the basic concepts of several cooperative schemes to assist our further discussion in Chapter 4.

User Cooperation Strategy

User cooperation refers to the cooperative diversity scheme where each user is responsible for transmitting not only its own information, but also the information of the other users. That is, the diversity gains in a cell are achieved via the cooperation of in-cell users. [SEA03a, SEA03b] first proposed a user cooperation model in a CDMA-based implementation of DF between a pair of users which could be represented by Figure 2.9 when the relay (R) is no longer a simple relay but another user. In the first and second interval, each user transmits its own bits. After detecting the other user's second bit, in the third interval both users transmit a linear combination of their own second bit and the estimate of the other user's second bit. [SEA03a, SEA03b] examined the achievable rate regions and outage capacity for the proposed scheme. The results in [SEA03a, SEA03b] show that even though the interuser channel is noisy, user cooperation leads not only to an increase in capacity for both users but also to a more robust system where users' achievable rates are less susceptible to channel variations.

Coded Cooperation

Coded cooperation was proposed which applies a practical channel coding scheme in the user cooperation strategy [HN02b, HN02a, HN06]. As depicted in Figure 2.10, the source and its partner encode a block of K source bits into N codeword bits respectively. The N bits are then partitioned into two segments after puncturing, with the first set being a punctured code with N_1 bits, and the second being the $(N - N_1)$ remaining parity bits. Initially, the user transmits N_1 bits and the partner attempts to decode the user's message. The N_1 bits are transmitted by the original user and the additional $(N - N_1)$ bits will be forwarded by the partner in the second phase if it managed to decode it successfully. In a word, the key idea behind coded cooperation is that each user tries to transmit redundant parity bits for its partner. By employing

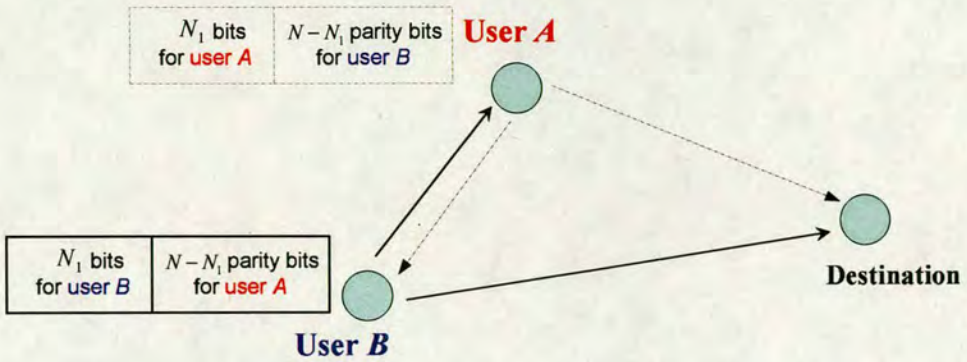

Figure 2.10: *In coded cooperation, user A and user B encode a block of K source bits into N codeword bits respectively. The N bits are then partitioned into two segments after puncturing, with the first set being a punctured code with N_1 bits, and the second being the $(N - N_1)$ remaining parity bits. Initially, user A/B transmits N_1 bits and then user B/A attempts to decode the user's message after receiving it. The N_1 bits are transmitted by the original user in the first frame and the additional $(N - N_1)$ bits will be forwarded by the other user in the second frame if successfully decoded.*

rate compatible punctured convolutional (RCPC) codes, a portion of each user's code bits arrive at the base station through a different, independent fading channel. The generated diversity gain provides a performance advantage over the non-cooperation transmission. Some extension work following this concept include [JHHN04], etc., which will be further discussed in Chapter 4.

Superposition Modulation and Code Superposition

In a traditional time division two-user cooperation scheme as shown in Figure 2.11 (case 2), each user transmits its own local information in the first half slot and then relays the information for its partner in the second half slot. The information transmitted in one user's relaying slot (*e.g.* "User A relays" in Figure 2.11) is already known to the other user (*e.g.* "User B" in Figure 2.11) as that is where the information originally comes from. From the perspective of the original user, the resources allocated to transmit the relayed bits by its partner is wasted. In order to make full use of the inter-user channel resource, *superposition modulation* [LV05] and *code superposition* [XFKC06a] are proposed to let one user simultaneously transmit its own data together with relaying information for the other user.

In *superposition modulation*, one user's local data and relayed data are allocated with different power level under total power constraint. At the destination, by using a soft MAP demodulator,

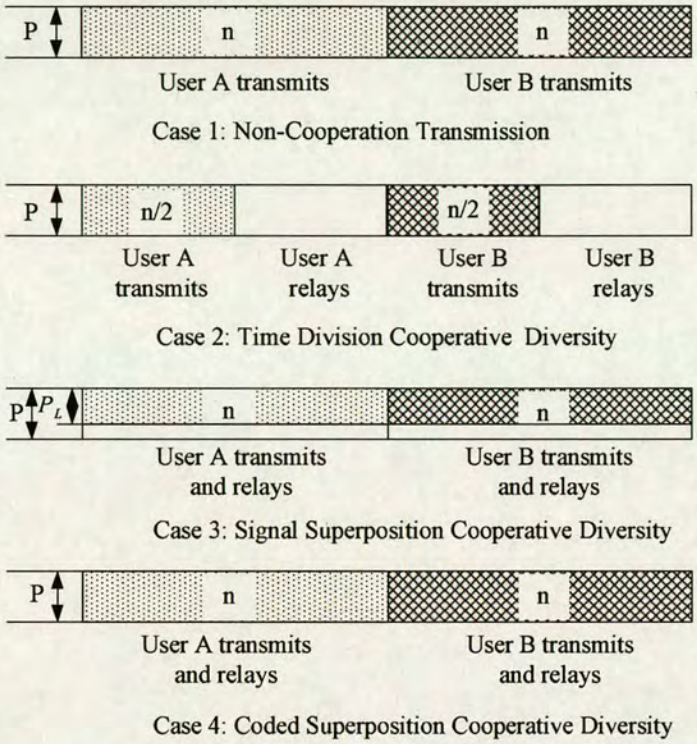

Figure 2.11: *Illustration of four different resource allocation schemes where the horizontal axis represents time while the vertical axis represents power. Two types of shaded portions correspond to the resources dedicated to the two users' local information transmission. In each half-slot, n symbols are transmitted with transmit power of P. In case 3, since part of the power has been used to modulate relayed codeword, local information is transmitted by partial power $P_L < P$ [XFKC07b].*

the *a posteriori* L-value for each information bit is obtained from two consecutive received signals. [LV05] demonstrate that signal superposition cooperative diversity scheme (case 3 in Figure 2.11) can outperform traditional DF time division cooperative diversity scheme (case 2 in Figure 2.11).

In *code superposition*, also called *network coding approach*, the local codeword and the relayed codeword are taken XOR before transmission. In the following, we only consider the case that each user successfully decodes its partner's codeword. In time slot t, the superimposed code $C(t)$ transmitted by each user is expressed by equation (2.59), where $i_L(t)$ represents the locally generated information, $i_R(t)$ represents the relayed information, $\mathbf{G}_L$ represents the

local generator matrix and $\mathbf{G}_R$ represents the generator matrix for the relayed bits.

$$C(t) = i_R(t)\mathbf{G}_R \bigoplus i_L(t)\mathbf{G}_L = \left[\begin{array}{cc} i_R(t) & i_L(t) \end{array} \right] \left[\begin{array}{c} \mathbf{G}_R \\ \mathbf{G}_L \end{array} \right] \tag{2.59}$$

For the decoder at the other user, the $i_R(t)\mathbf{G}_R$ is already known since that is the codeword generated by itself at last time slot. Thus the soft information needed to decode its partner's code is calculated by

$$\begin{aligned} L[i_L(t)\mathbf{G}_L](i) &= \log \frac{P_r\{[i_L(t)\mathbf{G}_L](i) = 0\}}{P_r\{[i_L(t)\mathbf{G}_L](i) = 1\}} \\ &= \begin{cases} L[C(t)](i) & \text{when} [i_R(t)\mathbf{G}_R](i) = 0 \\ -L[C(t)](i) & \text{when} [i_R(t)\mathbf{G}_R](i) = 1 \end{cases} \end{aligned} \tag{2.60}$$

The decoding at the destination is more complex. Next, we use the superscripts A and B to indicate the origins of information. In order to obtain $i_L^A(t)$, two codewords received at the destination are under consideration as shown in equation (2.61): $C^A(t)$ transmitted from user A and $C^B(t)$ transmitted from user B. As illustrated in equation (2.61), the information $i_L^A(t)$ is first protected by the local generator matrix $\mathbf{G}_L$ in $C^A(t)$ at user A and then protected by the relay generator matrix $\mathbf{G}_R$ in $C^B(t)$ at user B (in the case that user B successfully obtains $i_L^A(t)$). The decoding task is to "extract" the information of $i_L^A(t)$ in the two stacked codewords $C^A(t)$ and $C^B(t)$.

$$\begin{aligned} C^A(t) &= \left[\begin{array}{cc} i_L^B(t-1) & i_L^A(t) \end{array} \right] \left[\begin{array}{c} \mathbf{G}_R \\ \mathbf{G}_L \end{array} \right] \\ C^B(t) &= \left[\begin{array}{cc} i_L^A(t) & i_L^B(t) \end{array} \right] \left[\begin{array}{c} \mathbf{G}_R \\ \mathbf{G}_L \end{array} \right] \end{aligned} \tag{2.61}$$

In [XFKC07b], $\mathbf{G}_L$ and $\mathbf{G}_R$ are both 8-state convolutional codes. Hence a 64-state BCJR decoder [BCJR74] is required at the destination. The decoding is achieved by exchanging the corresponding extrinsic information between $C^A(t)$ and $C^B(t)$.

[XFKC07b] demonstrated that code superposition outperforms signal superposition under the same channel resource. In signal superposition scheme, part of the power has to be allocated to modulate the relayed codeword, while in code superposition, full power is obtainable for both local codeword and relayed codeword. It is indicated in [XFKC07b] that code superposition

(also referred to as *network coding* approach) is efficient to address the *cooperative dilemma* issue which concerns the resource allocation between the transmission of local information as well as relayed information. One challenge seems to be the decoder design which should be able to cope with the complicated decoding scenario and at the same time of acceptable decoding complexity. The above observations have inspired further discussions on this topic in Chapter 4.

2.4 Chapter Summary

This chapter presented a brief review of channel coding, factor graph theory and cooperative communications. LDPC codes are also introduced with specific emphasis on the SPA decoding method. This chapter lays the groundwork for the following two chapters. The next two chapters are concerned with the topic of exploiting LDPC codes in cooperative communication environments and the presentation is based on factor graph theory.

Chapter 3
Practical Successive Interference Cancellation in the Binary-Input Gaussian Multiple-Access Channel

3.1 Introduction

In the multiple access channel (MAC), all users share the common channel resources to transmit information to one destination. Time-Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA) systems are based on the orthogonal division of the resources "time" and "frequency", with only one user occupying a given frequency at a time. Contrary to that in CDMA, each user is assigned a distinct signature sequence that allows for orthogonal signal transmission with more than one user occupying a frequency at the same time [Pro01]; the price to pay is the increased bandwidth due to the "signature sequence" that has to be transmitted to send one information bit. Still, theoretical analysis shows that CDMA is a more efficient multiple access method [KA00] than TDMA and FDMA, mainly due to its greater flexibility.

In principle, the CDMA idea can also be used for channel-coded sequences without any explicit "spreading" code and in this case it is known from information theory [CT06] that successive interference cancellation (SIC) is a technique that can achieve the upper limits of the capacity region of the Gaussian multiple access channel. For the two-user MAC case, signal superposition together with SIC theoretically guarantees that one of the users can transmit at the capacity-approaching rate, as in the single-user AWGN channel, while the other user can still realise reliable transmission at a non-zero rate that is determined by the Gaussian MAC capacity region: essentially, the second user's signal is treated as "noise" by the first user, i.e., the first user suffers from a reduced rate due to interference. When the first user's rate is sufficiently low, the signal can be reliably decoded at the receiving end, and the re-encoded signal can be subtracted (cancelled) from the total received signal, leaving only the second signal and Gaussian receiver noise. This means the receiver "sees" a Gaussian channel without any interference

for the second user.

Although, at a first glance, this theoretical concept seems attractive for a practical application there are some potential drawbacks. First, the theory assumes that Gaussian signal alphabets are used, which is not possible in practice. Hence, we go to the opposite extreme and investigate the performance of SIC schemes when a binary modulation signal alphabet is used. This is motivated by the idea to have very simple transmitters such as wireless network nodes communicating to a receiving "base station" where we allow for some higher complexity to perform more advanced signal processing such as successive interference cancellation. Second, as it is impossible to decode with no errors in practice, the first user's transmit signal can not always be reconstructed correctly at the receiver, which means that the subtraction of the erroneous signal would lead to propagation of the first user's decoding errors into the decoding process for the second user. Hence, we investigate the soft multiple user demodulators combined with the decoding of channel codes in which the interference cancellation and the decoding are carried out in a "soft" iterative process by exchanging the extrinsic L-values carrying the bit-reliability information.

The major relative work in the field of soft iterative SIC, also called Turbo multiuser detector (MUD), is as follows: in [ARAS99, Moh98, RSAA98] the "turbo principle" is proposed in the realization of CDMA in combination with forward-error control coding, where a joint decoder is broken into the APP estimation of the coded symbols and parallel single-user FEC soft decoding; [WP99, Poo04] further develop Turbo MUD and demonstrate that the Turbo MUD method can provide good performance with relatively low complexity; [BC01] and [BC02] present a framework for iterative multiuser joint decoding of CDMA signals by factor graph theory, where DE is also used to obtain the asymptotic performance of the system in the limit of large code block lengths; in [SS06], the combination of forward-error control coding with CDMA using random spreading sequences is studied from the perspective of a large number of users.

In our work we use low-density parity-check (LDPC) codes which are very powerful error correcting codes. Well-designed codes with a very large blocksize can achieve very low bit-error-rates (BER) [RSU01]. In practical applications, however, delay constraints usually put an upper limit on the coding blocksize. Therefore, we use practical moderate-blocksize LDPC codes and investigate their performance in the Gaussian MAC with iterative SIC and decoding. We shall also use factor graph theory to present the LDPC coded MUD scheme. The major

contribution of this chapter is twofold. Firstly, we present a formulation of a multi-user detector under BPSK modulation which lends itself to an efficient implementation by L-value algebra; secondly, we employ a group of moderate-blocksize LDPC codes in the LDPC coded MUD in order to analyze the practical application of SIC under the BPSK modulation scenario and the corresponding simulation results are discussed.

This chapter is organized as follows. First this chapter will go through the capacity region for the additive white Gaussian noise (AWGN) Multiple Access Channel (MAC) based on information theory. Then a formulation of a multi-user detector is derived which lends itself to an efficient implementation by L-value algebra. Next is the presentation of the LDPC coded MUD by a factor graph. Finally this chapter investigates the practical potential of the iterative SIC and decoding scheme by employing a set of moderate-blocksize LDPC codes in the Binary-input AWGN MAC and provides some discussion with regards to the simulation results.

3.2 System Description and Channel Capacities

In this subsection, the theoretical background with regard to the two-user Gaussian multiple access channel will be reviewed.

3.2.1 System model (Two Users)

In a MAC, also called the uplink multiuser channel, all users send signals to one receiver. The discrete-time model for the two-user AWGN MAC is shown in Figure 3.1. The receive signal is given by

$$y[i] = x_a[i] + x_b[i] + n[i], \tag{3.1}$$

with i the time index, $n[i] \sim N(0, \sigma_w^2)$ the i.i.d. Gaussian receiver noise with power spectral density (PSD) $N_0/2$, the variance $\sigma_w^2 = N_0 \cdot B$ with B the baseband system bandwidth. For the sake of simplicity we assume real baseband signals; an extension to complex baseband signals is straightforward. The user signals are $x_a[i]$ and $x_b[i]$ respectively. We integrate any power scaling and any fixed path-loss into the average power constraints such that $\mathrm{E}\{(x_{a,b}[i])^2\} = P_{a,b}$.

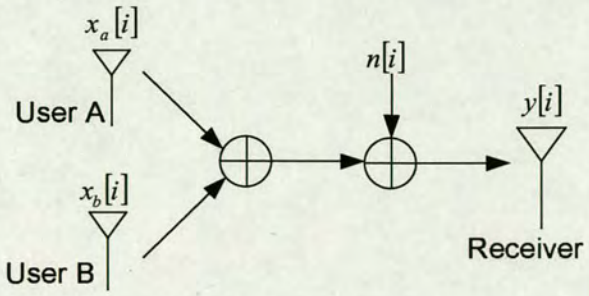

Figure 3.1: *A two-user multiple access channel.*

3.2.2 Capacity of the single-user discrete-time AWGN channel

The capacity for the ideal additive white Gaussian noise (AWGN) channel with average input power constraint P is achieved with a Gaussian distribution of the input symbols. The capacity equals

$$C = \frac{1}{2}\log_2(1 + \frac{P}{\sigma_w^2}) \text{ bits per transmission.} \tag{3.2}$$

For the band-limited AWGN channel with bandwidth B, the capacity of the channel can be rewritten as

$$C = B\log_2(1 + \frac{P}{N_0 B}) \text{ bits per second.} \tag{3.3}$$

The channel capacity for the binary-input discrete-time AWGN channel with equiprobable input bits $\xi \in \{0, 1\}$ and the channel output signal y is

$$C = \sum_{\xi \in \{0,1\}} \int_{-\infty}^{\infty} \frac{p(y|\xi)}{2} \log_2 \frac{2 \cdot p(y|\xi)}{p(y|\xi = 1) + p(y|\xi = 0)} dy \tag{3.4}$$

with C in bits per transmission (channel use). For a binary phase shift keying (BPSK) modulation (with coherent detection) we map the input bits $\xi \in \{0, 1\}$ to the modulation signal "constellation" $\{+1, -1\}$. However, as we have integrated any power scaling and path loss into the power constraints P_a, P_b we have for the received modulation signals

$$x_{a,b}(\xi) = \sqrt{P_{a,b}} \cdot (1 - 2 \cdot \xi) \tag{3.5}$$

Hence, the probability density function (pdf) of the Gaussian noise equals

$$p(y|\xi) \doteq \frac{1}{\sqrt{2\pi}\sigma_w} \exp(-\frac{1}{2\sigma_w^2}(y - x_{a,b}(\xi))^2) \,. \tag{3.6}$$

3.2.3 Capacity region for the two-user AWGN MAC

The two-user AWGN MAC capacity region is given by all rate vectors (R_a, R_b) satisfying the following constraints [CT06]:

$$\begin{aligned} R_a &\leq B\log_2\left(1+\frac{P_a}{N_0B}\right) \text{ bits / second} \\ R_b &\leq B\log_2\left(1+\frac{P_b}{N_0B}\right) \text{ bits / second} \\ R_a+R_b &\leq B\log_2\left(1+\frac{P_a+P_b}{N_0B}\right) \text{ bits / second} \end{aligned} \tag{3.7}$$

The interpretation for the capacity region is straightforward: The first two lines of (3.7) say that the transmission rates for the individual users can not exceed the Gaussian capacity limits in the single-user AWGN channel. The last line of (3.7) puts an upper limit on the sum rate of the users: this sum can not exceed the rate limit that we obtain on a Gaussian channel with the sum of the powers of both users.

The boundary of the feasible/infeasible rate region is shown in Figure 3.2 [Gol05], where C_k and C_k^* are given by

$$C_k = B\log_2\left(1+\frac{P_k}{N_0B}\right),\ k=a,b, \tag{3.8}$$

$$C_a^* = B\log_2\left(1+\frac{P_a}{N_0B+P_b}\right), \tag{3.9}$$

$$C_b^* = B\log_2\left(1+\frac{P_b}{N_0B+P_a}\right). \tag{3.10}$$

The points $(R_b, R_a) = (0, C_a)$ and $(C_b, 0)$ are achieved when one transmitter sends its data at its maximum rate in the single-user AWGN channel while the other user is silent. The rate pairs on the straight line between $(0, C_a)$ and $(C_b, 0)$ are achieved by a time-division strategy between two transmitters operating at their maximum rates with fixed power. The rate pairs on the line $(0, C_a) - (C_b^*, C_a)$ and the rate pairs on the line $(C_b, C_a^*) - (C_b, 0)$ are achieved by using signal superposition together with SIC. The rate pairs on the straight line connecting (C_b^*, C_a) and (C_b, C_a^*) are achieved by time-sharing between the two points or a rate-splitting technique [GRUW01]. For the latter, a user splits its data stream into multiple substreams and encodes these substreams as if they originated from different virtual users.

For comparison, Figure 3.2 also contains the rate region resulting from orthogonal signalling

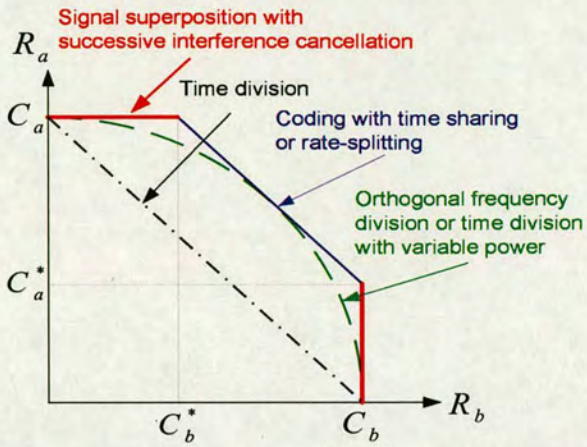

Figure 3.2: *Capacity region of the two-user multiple access channel [Gol05].*

(FDMA and TDMA) without SIC: the rate region achieved by those schemes touches the boundary of the capacity region in exactly one point, which corresponds to the rate at which we obtain the maximum sum rate. In all other points orthogonal signalling is strictly suboptimal. The problem is particularly significant when the users have very different receive power levels: in this case the weaker user will hardly get any rate at all in the capacity-achieving point [TV05, pp. 232–234].

3.2.4 Signal superposition and successive interference cancellation

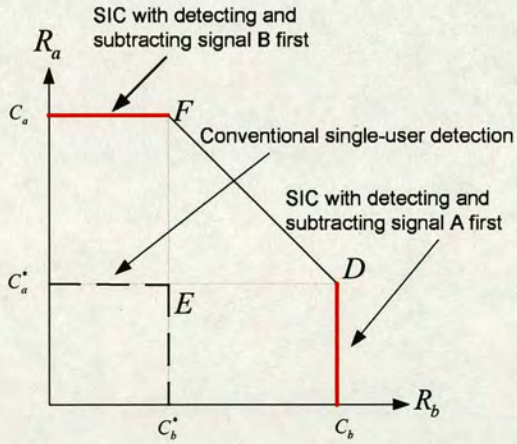

Figure 3.3: *Capacity boundary for SIC and conventional single-user detection in two-user multiple access channel.*

Successive Interference Cancellation (SIC) is a multiuser detection technique which can be used to achieve the boundary of the capacity region of a Multiple-Access Channel [TV05]. In

contrast to conventional single-user detection (e.g., [KA00]), in which other users' signals are treated as noise when decoding one user's signal, SIC is based on subtracting (cancelling) the already detected signals from the received signal before the detection of the next signal. The more disparate the users' powers are, the higher the potential gain from SIC.

The point D in Figure 3.3 indicates that user B could transmit at capacity-approaching rate in the point-to-point transmission model, while user A can still transmit at the maximum rate of C_a^* without errors. If signal A is the first one to be decoded, signal B is treated as Gaussian noise to user A in the first signal detection stage.

In Figure 3.3, the line $C_b - D$ is achieved by signal superposition together with SIC when signal A is first detected and subtracted. If the order of the signal detection is reversed and signal B is first detected, by SIC the capacity boundary $C_a - F$ could be achieved. The capacity boundary encompassed by C_a^*, E and C_b^* is for signal superposition without SIC which is the conventional single-user detection method. The theoretical capacity region shows that signal superposition with SIC is superior to the conventional single-user detection method.

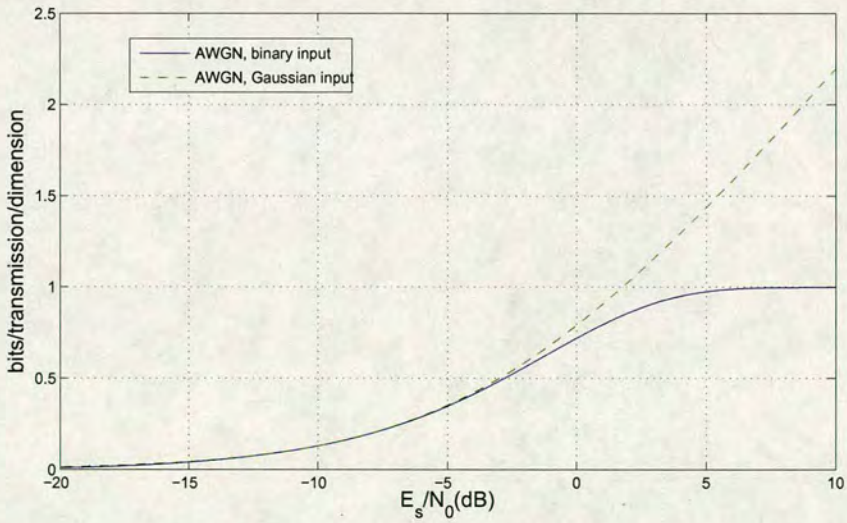

Figure 3.4: *Comparison of the capacities of single user AWGN channels with binary input and Gaussian input constraints.*

3.3 Multiuser APP-Demodulator for BPSK Modulation

In practice, reliable communication with channel codes of limited block size is impossible in a strict sense. Therefore, instead of "cancelling by signal subtraction" (which would cause error propagation in case of incorrect decoding), we resort to a "soft" *a posteriori* probability (APP) demodulator, which can deal with reliability information for bits rather than with hard decisions only. This concept (in the given framework) is known from *e.g.* [WP99] [CMT04] [LV05].

According to information theory, SIC is a technique which can achieve the upper limits of the capacity region of a multiple access channel. However, the theory is based on the assumption of Gaussian input symbols. We already know that in a single-user channel case, AWGN channels with binary inputs and with Gaussian inputs attain different capacities as illustrated in Figure 3.4. Since SIC is considered on the assumption of Gaussian input symbols in the information theory, the key challenge of this work here is to investigate the performance of SIC under binary input symbols by employing a group of moderate blocksize LDPC codes. Hence in what follows we consider the special case of coherently detected binary phase-shift keying modulation (BPSK) and we state a formulation of a multi-user detector that lends itself to an efficient implementation by L-value algebra.

The basic idea is to extract the *a posteriori* probability for the transmitted bits $Z_k[i]$ of all users $k = 1, 2, ..., K$ to be "one" or "zero" given the received channel value y (see Fig. 3.1 for the two-user case). The channel model for the general K-user case is given by

$$y[i] = \sum_{k=1}^{K} x_k[i] + n[i] \,, \tag{3.11}$$

where

$$x_k(z_k[i]) = \sqrt{P_k}(1 - 2z_k[i]) \tag{3.12}$$

and $z_k[i] \in \{0, 1\}$ is the bit sent by user k at time i; P_k is the receive power for the signal transmitted by user k. We now calculate the APP L-value [HOP96] of the bit $z_k[i]$:

$$L(Z_k[i]|y) = \ln \frac{\Pr(Z_k[i] = 0|y)}{\Pr(Z_k[i] = 1|y)} \tag{3.13}$$

where $Z_k[i]$ indicates the random variable of the bit and $z_k[i] \in \{0, 1\}$ its realisation. As we consider only one time-instant in the detector, we omit the time index i in what follows as long as there is no risk of confusion.

We expand

$$\Pr(Z_k = 0|y) = \sum_{\forall\{z_1,\ldots,z_K\}:z_k=0}^{K} \Pr(Z_1 = z_1, ..., Z_K = z_K|y) \tag{3.14}$$

and use the Bayes-rule which immediately leads to

$$L(Z_k|y) = \ln \frac{\sum_{\forall \mathbf{z}:z_k=0} p(y|\mathbf{z}) \cdot \Pr(\mathbf{z})}{\sum_{\forall \mathbf{z}:z_k=1} p(y|\mathbf{z}) \cdot \Pr(\mathbf{z})} \tag{3.15}$$

with the bit vector realisations $\mathbf{z} \doteq \{z_1, z_2, ..., z_K\}$ (we use the same definition for the bit vector random "variable" $\mathbf{Z}$). The notation $\forall \mathbf{z} : z_k = 0$ denotes all possible bit combinations of the other users, with the bit value of "zero" for the user k under consideration. For later use we further define $\mathbf{z}\backslash z_k \doteq \{z_1, z_2, ..., z_{k-1}, z_{k+1}, ..., z_K\}$. Moreover, we assume that the users' code bits are independent (realistic assumption for independent users), so that

$$\Pr(\mathbf{z}) = \prod_{k=1}^{K} \Pr(z_k) \, . \tag{3.16}$$

(above, we use the abbreviation $\Pr(Z_k = z_k) = \Pr(z_k)$).

If we assume a Gaussian channel, we can write the probability density function (PDF) of the channel as follows:

$$p(y|\mathbf{z}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2} \Big(y - \sum_{k=1}^{K} x_k(z_k) \Big)^2 \right) \tag{3.17}$$

Up to this stage, the equations are also valid for non-binary symbols, but now we start to exploit the special BPSK modulation scheme: we reformulate the exponent of the PDF in (3.17): with $x_0 \doteq -y$ and the abbreviation $x_k \doteq x_k(z_k), k = 1, 2, ...$ we obtain

$$\begin{aligned} \left(y - \sum_{k=1}^{K} x_k(z_k) \right)^2 &= \left(\sum_{k=0}^{K} x_k \right)^2 = \sum_{k=0}^{K} x_k^2 + 2 \sum_{k=0}^{K-1} x_k \sum_{l=k+1}^{K} x_l \\ &= y^2 + \sum_{k=1}^{K} x_k^2 - 2y \cdot \sum_{k=1}^{K} x_k + 2 \sum_{k=1}^{K-1} x_k \sum_{l=k+1}^{K} x_l \end{aligned} \tag{3.18}$$

As $x_k = \sqrt{P_k}(1 - 2z_k)$, with $z_k \in \{0, 1\}$, we have $x_k^2 = P_k$, which does *not* depend on the

choice of the bit z_k. Therefore, we can re-write (3.15) as follows:

$$\begin{aligned} L(Z_k|y) &= \ln \frac{\Pr(Z_k=0) \sum\limits_{\forall \mathbf{z}: z_k=0} p(y|\mathbf{z}) \cdot \Pr(\mathbf{z}\backslash z_k)}{\Pr(Z_k=1) \sum\limits_{\forall \mathbf{z}: z_k=1} p(y|\mathbf{z}) \cdot \Pr(\mathbf{z}\backslash z_k)} \\ &= L(Z_k) + \ln \frac{\sum\limits_{\forall \mathbf{z}: z_k=0} e^{f(y,z_1,\ldots,z_K)/\sigma^2} \cdot \Pr(\mathbf{z}\backslash z_k)}{\sum\limits_{\forall \mathbf{z}: z_k=1} e^{f(y,z_1,\ldots,z_K)/\sigma^2} \cdot \Pr(\mathbf{z}\backslash z_k)} \end{aligned} \tag{3.19}$$

with

$$f(y, z_1, ..., z_K) = y \cdot \sum_{\kappa=1}^{K} x_\kappa - \sum_{\xi=1}^{K-1} x_\xi \sum_{l=\xi+1}^{K} x_l \tag{3.20}$$

Note that, as above, $x_\kappa = \sqrt{P_\kappa}(1 - 2 \cdot z_\kappa)$ for BPSK modulation. We now separate out the bit z_k in (3.20), i.e., we isolate $x_k = x_k(z_k)$:

$$\begin{aligned} f(y, z_1, ..., z_K) &= x_k\Big(y - \sum_{\substack{\xi=1 \\ \xi\neq k}}^{K} x_\xi\Big) + y \sum_{\substack{l=1 \\ l\neq k}}^{K} x_l - \sum_{\substack{\xi=1 \\ \xi\neq k}}^{K-1} x_\xi \sum_{\substack{l=\xi+1 \\ l\neq k}}^{K} x_l \\ &= x_k \cdot y + \Big(y - x_k\Big) \sum_{\substack{\xi=1 \\ \xi\neq k}}^{K} x_\xi - \sum_{\substack{\xi=1 \\ \xi\neq k}}^{K-1} x_\xi \sum_{\substack{l=\xi+1 \\ l\neq k}}^{K} x_l \end{aligned} \tag{3.21}$$

When we use (3.21) in (3.19) with $x_k(z_k) = \sqrt{P_k}(1 - 2z_k)$ we find

$$\begin{aligned} L(Z_k|y) &= L(Z_k) + \log \frac{\exp(y \cdot x_k(0)/\sigma^2)}{\exp(y \cdot x_k(1)/\sigma^2)} + L_E(Z_k) \\ &= L(Z_k) + 2\frac{\sqrt{P_k}}{\sigma^2} \cdot y + L_E(Z_k) \end{aligned} \tag{3.22}$$

where

$$L_E(Z_k) = \log \frac{\sum\limits_{\forall \mathbf{z}\backslash z_k} e^{g_0(\mathbf{z}\backslash z_k)/\sigma^2} \cdot \Pr(\mathbf{z}\backslash z_k)}{\sum\limits_{\forall \mathbf{z}\backslash z_k} e^{g_1(\mathbf{z}\backslash z_k)/\sigma^2} \cdot \Pr(\mathbf{z}\backslash z_k)} \tag{3.23}$$

with

$$g_\rho(\mathbf{z}\backslash z_k) = \Big(y - x_k(\rho)\Big) \sum_{\substack{\xi=1 \\ \xi\neq k}}^{K} x_\xi - \sum_{\substack{\xi=1 \\ \xi\neq k}}^{K-1} x_\xi \sum_{\substack{l=\xi+1 \\ l\neq k}}^{K} x_l \,, \tag{3.24}$$

with $\rho \in \{0,1\}$ and $x_k = \sqrt{P_k}(1-2z_k)$. As long as we are using binary modulation, (3.23) can be evaluated for, e.g., up to $K = 10$ users without serious complexity problems. If more users are in the system and the received powers P_i for their signals are significantly different, we can approximate the sums in (3.24) by only considering a subset of the "strongest" users.

We can further develop (3.23) by using (3.16):

$$\Pr(\mathbf{z}\backslash z_k) = \prod_{\substack{l=1 \\ l\neq k}}^{K} \Pr(z_l)\,. \tag{3.25}$$

We now express the probabilities $\Pr(z_l)$ of the bits of user l by its corresponding "a-priori" L-value

$$L(Z_l) \doteq \ln \frac{\Pr(Z_l = 0)}{\Pr(Z_l = 1)}\,. \tag{3.26}$$

By inversion of (3.26) we obtain

$$\Pr(Z_l = z) = \frac{e^{-L(Z_l)\cdot z}}{1 + e^{-L(Z_l)}}, \quad z \in \{0,1\}\,. \tag{3.27}$$

We use (3.25) to find

$$\Pr(\mathbf{z}\backslash z_k) = \omega \cdot \exp\left(-\sum_{\substack{l=1 \\ l\neq k}}^{K} L(Z_l)\cdot z_l \right) \tag{3.28}$$

with the constant ω that will cancel out in (3.23). We combine (3.28) and (3.24) conveniently for use in (3.23) and obtain

$$\begin{aligned} h_\rho(\mathbf{z}\backslash z_k) &= \frac{1}{\sigma^2} g_\rho(\mathbf{z}\backslash z_k) - \sum_{\substack{l=1 \\ l\neq k}}^{K} L(Z_l)\cdot z_l \\ &= \frac{1}{\sigma^2}\left(\Big(y - x_k(\rho)\Big) \sum_{\substack{\xi=1 \\ \xi\neq k}}^{K} x_\xi - \sum_{\substack{\xi=1 \\ \xi\neq k}}^{K-1} x_\xi \sum_{\substack{l=\xi+1 \\ l\neq k}}^{K} x_l \right) - \sum_{\substack{l=1 \\ l\neq k}}^{K} L(Z_l)\cdot z_l \end{aligned} \tag{3.29}$$

with

$$L_E(Z_k) = \ln \frac{\sum_{\forall \mathbf{z}\backslash z_k} \exp(h_0(\mathbf{z}\backslash z_k))}{\sum_{\forall \mathbf{z}\backslash z_k} \exp(h_1(\mathbf{z}\backslash z_k))} \tag{3.30}$$

As only the factor $x_k(\rho)$ is different in the numerator and denominator of (3.30), most of the

values $h_\rho(\mathbf{z}\backslash z_k)$ need only to be computed once. Moreover, the probability-weighting by the L-values for the bits of other users appears as a simple sum of L-values (rightmost term in (3.29), which are scaled by the bit-values $z_l \in \{0,1\}$: hence, only the "one"-bits in the bit-vector sum-"index" $\mathbf{z}\backslash z_k$ need to be considered. The double-sum in (3.29) is most conveniently computed by starting with $\xi = K-1$: the result of the inner sum can then be recursively used to compute the next result. With a limited number of users in the system, within one detection all sums in (3.29), apart from the last term, can in principle be pre-computed and stored as they do not depend on y or on the L-values of the other users' bits. Moreover, we can ignore terms in the sums that contribute very little to the total result (which stem typically from users with very low channel SNR); this helps to cope with a larger number of users in the system.

In the following section, we have used (3.22), (3.29) and (3.30) for the special case of $K = 2$ users. It needs to be pointed out that the detection order by applying equation (3.22) leads to the different capacity regions as shown in Figure 3.3 for the two-user case.

3.4 Application of Low-Density Parity-Check Codes in Successive Interference Cancellation

The Turbo principle of iterative process has been applied among multiuser detection and error control coding where channel decoders provide intermediate soft decisions that can be exploited by a multiuser detector [Poo04]. This Turbo MUD method can provide good performance with relatively low complexity [WP99, Poo04]. The work of [WP99, KS01, KH97, KS03] focused on convolutionally encoded CDMA systems; the work of [YWN03] employs LDPC codes for turbo multiuser detection where density evolution (DE) is used to compute the theoretical threshold for LDPC codes. In principle, the CDMA idea can also be used for channel-coded sequences without any explicit "spreading" code. The receive structure for the LDPC coded Turbo MUD is shown in Figure 3.5 where the labels are corresponding to those in the factor graph presentation in Figure 3.6. In this section, we use practical moderate-blocksize LDPC codes and investigate their performance in the Gaussian MAC with iterative SIC and decoding. We shall first present some simulation curves to illustrate the performance of this LDPC coded Turbo MUD system; then we shall discuss some practically achievable points as compared with the theoretical description.

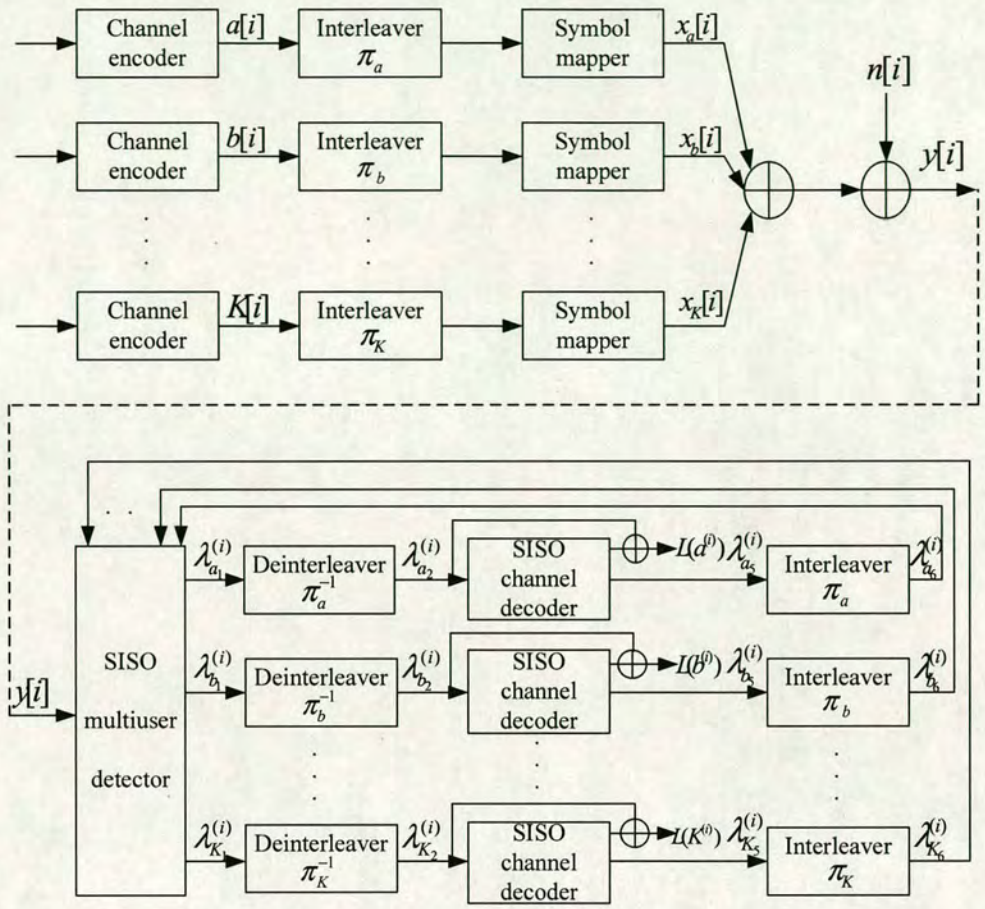

Figure 3.5: *The receiver structure for LDPC coded Turbo MUD where the labels are corresponding to those in the factor graph presentation in Figure 3.6.*

3.4.1 Iterative (LDPC) soft interference cancellation and decoding scheme

Based on the factor graph theory introduced in Section 2.2, the two-user iterative soft interference cancellation and decoding scheme by employing LDPC codes, or called LDPC coded Turbo MUD, is illustrated in Figure 3.6. In the following, we will refer to the signal transmitted by user A as signal A and refer to the signal transmitted by user B as signal B. For simplicity, the two soft-input soft-output (SISO) decoder modules for the constituent LDPC codes $\mathcal{C}_a$ (for user A) and $\mathcal{C}_b$ (for user B) are illustrated for the case $n = 3$, where circles depict bit nodes, squares depict check nodes and (i) depicts the ith bit. The factor node $\{F_i\}$ corresponds to the APP-demodulator and the variable node $y^{(i)}$ corresponds to the channel observed value $y[i]$ (or y when the time index i is omitted afterwards) in Section 3.3. In the following we define $\mathcal{I} = \{1, 2, \cdots, n\}$, and $x^{(i)}$ denotes the i-th bit of codeword $\mathbf{x}$. The notation "π"

and "π^{-1}" represent the interleaver and the de-interleaver respectively. $\mathcal{J}_a = \{1, 2, \cdots m_a\}$; $\mathcal{J}_b = \{1, 2, \cdots m_b\}$; $\mathcal{N}_a(i) = \{j \in \mathcal{J}_a \ : \ H_a(j,i) = 1\}$; $\mathcal{N}_b(i) = \{j \in \mathcal{J}_b \ : \ H_b(j,i) = 1\}$; $\mathcal{M}_a(j) = \{i \in \mathcal{I} \ : \ H_a(j,i) = 1\}$; $\mathcal{M}_b(j) = \{i \in \mathcal{I} \ : \ H_b(j,i) = 1\}$. The letter λ is used to denote LLRs corresponding to messages passed on the factor graph. The interleaving "π" is interpreted as

$$\mathbf{x} = \pi(\mathbf{y}) \quad \Longleftrightarrow \quad x^{(i)} = y^{(\pi(i))}$$

For the first iteration, assuming equally likely code bits, *i.e.*, no *a priori* information is available, we have $L(Z_k) = 0$ in (3.22) and $L(Z_l) = 0$ in (3.29). The message-passing schedule to start with $\lambda_{a_1}^{(i)}$ or $\lambda_{b_1}^{(i)}$ is decided by the detection order which leads to different capacity region for this pair of users as illustrated in Figure 3.3. Suppose that the signal A is the first to be detected, for the start the LLR $\lambda_{a_1}^{(i)}$ is calculated by setting $\lambda_{b_6}^{(i)} = 0$ based on (3.22), (3.29) and (3.30). Then $\lambda_{a_1}^{(i)}$ goes into the SISO decoder module $\mathcal{C}_a$ as the *a priori* information. After one local iteration within $\mathcal{C}_a$ (as presented in the section 2.2.4.2), the updated extrinsic value stream $\lambda_{a_6}^{(i)}$ expressing the constraints by the code structure is then fed back to the APP-demodulator $\{F_i\}$ as the *a priori* information (equivalent to $L(Z_l)$ in (3.29)) for $\lambda_{b_1}^{(i)}$ which initiates a similar local iterative process within $\mathcal{C}_b$. The message exchanging process between the two SISO decoder modules $\mathcal{C}_a$ and $\mathcal{C}_b$ then repeats itself. It needs to be pointed out that each time the APP-demodulator takes as input extrinsic LLRs $\lambda_{a_6}^{(i)}$ or $\lambda_{b_6}^{(i)}$ expressing the information of the transmitted bits in the absence of channel information and then computes the new extrinsic LLRs $\lambda_{b_1}^{(i)}$ or $\lambda_{a_1}^{(i)}$ expressing new information derived from the channel. Thus in the calculation of the extrinsic information $\lambda_{a_1}^{(i)}$ or $\lambda_{b_1}^{(i)}$, we have $L(Z_k) = 0$ in (3.22) while in the calculation of the *a posteriori* LLR of each bit (which is used to make decisions on the decoded bits at the last iteration) both *a priori* LLRs (those flowing into the decoder) and the extrinsic LLRs (those flowing out of the decoder) are taken into account. At the first iteration, the extrinsic LLRs $\lambda_{b_1}^{(i)}$ and $\lambda_{a_1}^{(i)}$ are statistically independent. However subsequently they will become more and more correlated since they use the same information indirectly. The global iteration process ceases when it reaches a certain maximum iteration number N_{it}, (*e.g.* 20 in the following simulations) or at least one component code satisfies the checksum condition. If only one component code satisfies the checksum condition, the other component code will perform in local iterations until it converges to a solution or reaches a total of N_{it} iterations.

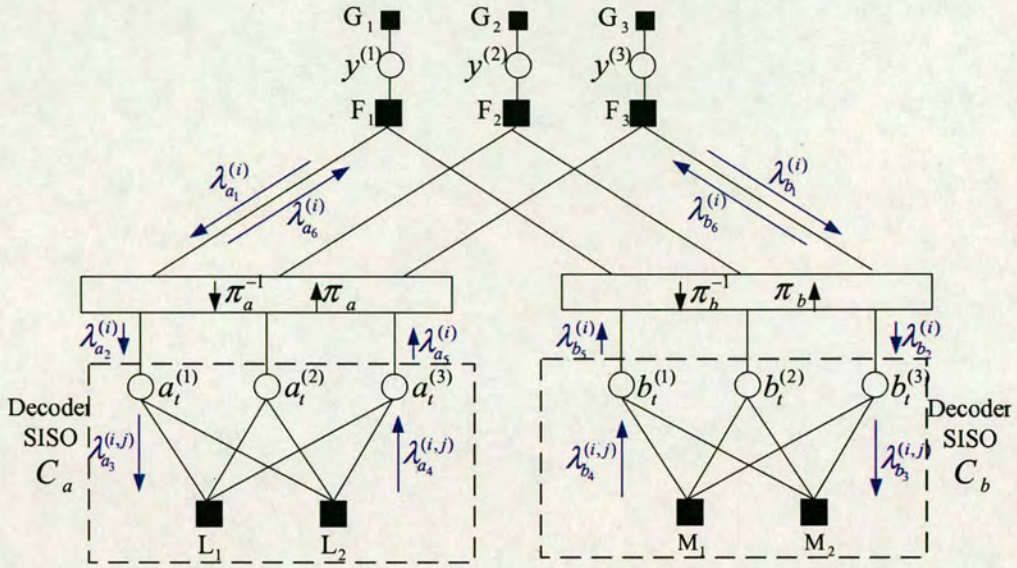

Figure 3.6: *Factor graph corresponding to the iterative LDPC soft SIC and decoding scheme. The message-passing schedule is such that extrinsic information is exchanged between the two decoder SISO modules for the constituent codes $\mathcal{C}_a$ and $\mathcal{C}_b$, via the factor nodes $\{F_i\}$ corresponding to the APP-demodulator. For ease of presentation, the factor graph is illustrated for the case $n = 3$ and trivial codes $\mathcal{C}_a$, $\mathcal{C}_b$.*

Factor Graph Based Decoding Algorithm

for LDPC coded Turbo MUD (two user case)

Initialization:

- For $i \in \mathcal{I}$,

$$\lambda_{b6}^{(i)} = 0 \tag{3.31}$$

- For $i \in \mathcal{I}, j \in \mathcal{N}_a(i)$

$$\lambda_{a4}^{(i,j)} = 0 \tag{3.32}$$

- For $i \in \mathcal{I}, j \in \mathcal{N}_b(i)$

$$\lambda_{b4}^{(i,j)} = 0 \tag{3.33}$$

Main Loop: For $i_t = 1$ **to** N_{it} **do**

- MUD by Equation (3.22)

$$\lambda_{a1}^{(i)} = L(Z_a|y) \tag{3.34}$$

- For $i \in \mathcal{I}$,

$$\lambda_{a_2}^{(\pi_a(i))} = \lambda_{a_1}^{(i)} \tag{3.35}$$

- SISO decoder $\mathcal{C}_a$: for $i \in \mathcal{I}, j \in \mathcal{N}_a(i)$

$$\lambda_{a_3}^{(i,j)} = \lambda_{a_2}^{(i)} + \sum_{l \in \mathcal{N}_a(i) \backslash \{j\}} \lambda_{a_4}^{(i,l)} \tag{3.36}$$

$$\lambda_{a_4}^{(i,j)} = \boxplus_{l \in \mathcal{M}_a(j) \backslash \{i\}} \lambda_{a_3}^{(l,j)} \tag{3.37}$$

- For $i \in \mathcal{I}$,

$$\lambda_{a_5}^{(i)} = \sum_{j \in \mathcal{N}_a(i)} \lambda_{a_4}^{(i,j)} \tag{3.38}$$

- Calculate *a posteriori* LLRs for codeword $\mathbf{a}_t$:

$$L(a_t^{(i)}) = \lambda_{a_2}^{(i)} + \lambda_{a_5}^{(i)} \tag{3.39}$$

- Make decisions on the code bits

$$\hat{a}_t^{(i)} = \begin{cases} 0 & \text{if } L(a_t^{(i)}) \geq 0 \\ 1 & \text{if } L(a_t^{(i)}) < 0 \end{cases}$$

- For $i \in \mathcal{I}$,

$$\lambda_{a_6}^{(i)} = \lambda_{a_5}^{(\pi_a(i))} \tag{3.40}$$

- MUD by Equation (3.22)

$$\lambda_{z_1}^{(i)} = L(Z_b|y) \tag{3.41}$$

- For $i \in \mathcal{I}$,

$$\lambda_{b_2}^{(\pi_b(i))} = \lambda_{b_1}^{(i)} \tag{3.42}$$

- SISO decoder $\mathcal{C}_b$: for $i \in \mathcal{I}, j \in \mathcal{N}_b(i)$

$$\lambda_{b_3}^{(i,j)} = \lambda_{b_2}^{(i)} + \sum_{l \in \mathcal{N}_b(i) \backslash \{j\}} \lambda_{b_4}^{(i,l)} \tag{3.43}$$

$$\lambda_{b_4}^{(i,j)} = \boxplus_{l \in \mathcal{M}_b(j) \backslash \{i\}} \lambda_{b_3}^{(l,j)} \tag{3.44}$$

- For $i \in \mathcal{I}$,

$$\lambda_{b_5}^{(i)} = \sum_{j \in \mathcal{N}_b(i)} \lambda_{b_4}^{(i,j)} \tag{3.45}$$

- Obtain the *a posteriori* LLR (prior for decoding in next half slot)

$$L_2(b_t^{(i)}) = \lambda_{b_2}^{(i)} + \lambda_{b_5}^{(i)} \tag{3.46}$$

- Make decisions on the code bits

$$\hat{b}_t^{(i)} = \begin{cases} 0 & \text{if } L(b_t^{(i)}) \geq 0 \\ 1 & \text{if } L(b_t^{(i)}) < 0 \end{cases}$$

- For $i \in \mathcal{I}$,

$$\lambda_{b_6}^{(i)} = \lambda_{b_5}^{(\pi_b(i))} \tag{3.47}$$

If $\hat{\mathbf{a}}_t \mathbf{H}_a^T = \mathbf{0}$ or $\hat{\mathbf{b}}_t \mathbf{H}_b^T = \mathbf{0}$ then *break;*

Endfor

3.4.2 Simulation results and analysis

The LDPC codes used here are randomly generated regular LDPC codes[1] of blocksize 1200 with column weight 3 for both users. For the theoretical capacity calculation, equation (3.2) is for Gaussian input symbols which could not be implemented in the real world, while the capacity equation (3.4) for binary input symbols could be used as a benchmark for BPSK modulation. In the following simulations, this set of LDPC codes is used in the two-user MAC with BPSK modulation in order to compare the achievable rates by iterative SIC and decoding with the capacity region. The theoretical signal-to-noise ratio (SNR) to achieve error free transmission and actual SNR for those middle blocksize LDPC codes to achieve BERs of 10^{-4} in the single-user AWGN channel are listed in Table 3.1.

In the multi-user framework below, the SNR next refers to a user's signal power P_k over the background Gaussian noise variance σ^2 in the channel, and the background Gaussian noise

[1] *The regular random LDPC matrices were constructed using the online software available at* http://www.cs.toronto.edu/~radford/ldpc.software.html. *This program generates random regular LDPC codes of any specified rate and block length and is capable of expurging four cycles in the LDPC constraint graph.*

Code rate	0.4	0.5	0.7	0.8	0.9
Theoretical SNR (E_s/N_0 in dB)	-4.22	-2.82	-0.28	1.07	2.74
Actual SNR (E_s/N_0 in dB)	-1.73	-0.61	1.45	3.01	4.3
Actual SNR (E_b/N_0 in dB)	2.25	2.4	3.0	3.98	4.75

Table 3.1: *Theoretical SNR and actual SNR for the 1200 blocksize LDPC codes by BPSK modulation.*

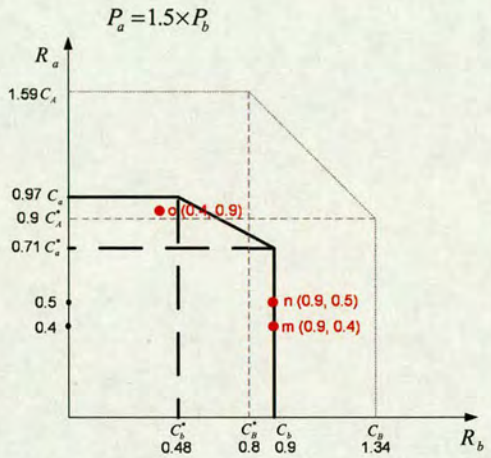

Figure 3.7: *Four corner points (C_a, C_a^*, C_b and C_b^*) and three rate pairs ($m(0.9, 0.4)$, $n(0.9, 0.5)$ and $o(0.4, 0.9)$) for examination under the scenario $P_a = 1.5 \times P_b$ for BPSK modulation. The region encompassed by $(0,\ C_A)$, $(C_B^*,\ C_A)$, $(C_B,\ C_A^*)$ and $(C_B,\ 0)$ is the capacity region defined by SIC of Gaussian input symbols under equivalent SNRs.*

Point (B, A)	SNR & BER	Signal A	Signal B
$M(0.9, 0.4)$	E_s/N_0 (dB)	6.06	4.3
	E_b/N_0 (dB)	10.03	4.75
	BER	1.2×10^{-5}	1.0×10^{-4}
$N(0.9, 0.5)$	E_s/N_0 (dB)	6.06	4.3
	E_b/N_0 (dB)	9.07	4.75
	BER	0.03	0.035
$O_1(0.4, 0.9)$	E_s/N_0 (dB)	6.06	4.3
	E_b/N_0 (dB)	6.5	8.3
	BER	9.0×10^{-4}	8.0×10^{-4}
$O_2(0.4, 0.9)$	E_s/N_0 (dB)	6.93	5.17
	E_b/N_0 (dB)	7.38	9.15
	BER	7.0×10^{-5}	1.0×10^{-4}

Table 3.2: *The BERs of four points $M(0.9, 0.4)$, $N(0.9, 0.5)$, $O_1(0.4, 0.9)$ and $O_2(0.4, 0.9)$ which are corresponding to the rate pairs of $m(0.9, 0.4)$, $n(0.9, 0.5)$ and $o(0.4, 0.9)$ with specific SNR values under the scenario $P_a = 1.5 \times P_b$.*

variance σ^2 is conveniently set to "1" (0dB) for the simulations.

3.4.2.1 First scenario: $P_a = 1.5 \times P_b$

In this scenario, the received power of signal A is 1.5 times that of the received power of signal B. C_b is chosen as 0.9. Thus according to power allocation between the two users, the binary input theoretical capacity values C_a (0.97) could be obtained by (3.4). Also C_a^* (0.71) and C_b^* (0.48) are obtained by (3.4) based on the assumption in the capacity calculation that the channel noise and the other user's signal appear as Gaussian interference. (This Gaussian interference assumption will be analysed later.) The theoretical binary-input capacity calculation shows that under the scenario of $P_a = 1.5 \times P_b$, as the rate of signal B (C_b) is fixed, the maximum rate of signal A could be C_a^* when signal A is first detected. In the same way, as the rate of signal A is fixed at C_a, the rate of signal B could be increased to as far as C_b^* with error free transmission when signal B is first detected. The resulted four rates (C_a, C_a^*, C_b and C_b^*) are demonstrated in Figure 3.7. In the following, we shall illustrate the performances of the LDPC coded Turbo MUD for three rate pairs: $m(0.9, 0.4)$, $n(0.9, 0.5)$ and $o(0.4, 0.9)$.

For each of the rate pairs, we shall focus on one specific pair of SNR condition which aims to examine the implementation of the SIC theory under the binary input condition. Because we use LDPC codes with a middle blocksize, strictly error free transmission is impossible in practice. Hence, instead of the theoretical E_s/N_0 (2.74dB) for the code rate of 0.9, E_s/N_0 at 4.3dB for signal B ensures that this 0.9 rate middle blocksize LDPC code provides the BER performance of 10^{-4}. Accordingly, the equivalent E_s/N_0 for the signal A is 6.06dB. Thus, the specific pair of SNRs is (4.3dB, 6.06dB) by E_s/N_0. Points $M(0.9, 0.4)$, $N(0.9, 0.5)$ and $O_1(0.4, 0.9)$ demonstrated in Table 3.2, Figure 3.8, Figure 3.9, Figure 3.12, Figure 3.13, Figure 3.15 and Figure 3.16 are of the specific pair of SNRs (4.3dB, 6.06dB) in the rate pair $m(0.9, 0.4)$, $n(0.9, 0.5)$ and $o(0.4, 0.9)$ respectively. The principle for the chosen SNR pair is that we aim to examine the effect of the SIC operation rather then the decoder itself. If the SIC theory works perfectly, this setting should guarantee the BERs for the both users fall in the range not above 10^{-4}. The examined points and their corresponding SNRs and BERs performance are listed in Table 3.2.

m(0.9, 0.4)

In the rate pair $m(0.9, 0.4)$, the user A with stronger power is the first to be detected. Then the message flows between the two decoder SISO modules via the APP-demodulator up to 20 iter-

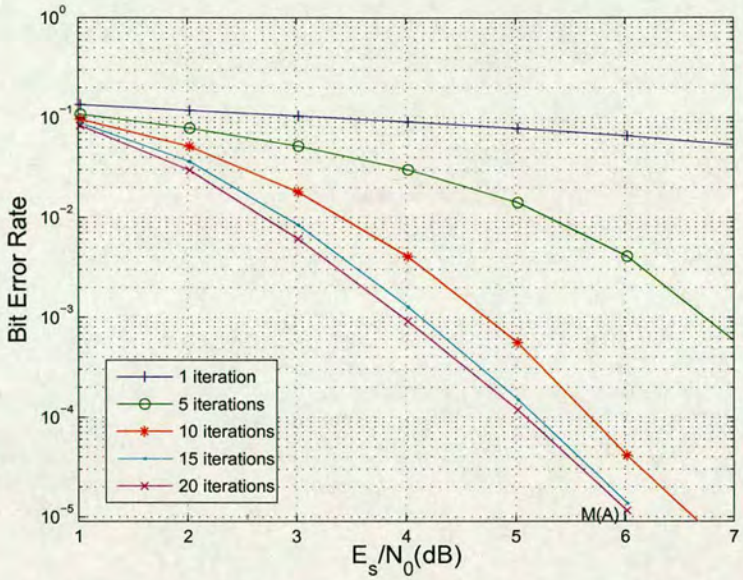

Figure 3.8: *Performance for user A with rate 0.4 in the rate pair* $m(0.9, 0.4)$ *under the iterative LDPC soft SIC and decoding scheme. The user A with stronger power is the first to be detected.* $M(A)$ *corresponds to signal A of point* $M(0.9, 0.4)$ *listed in Table 3.2.*

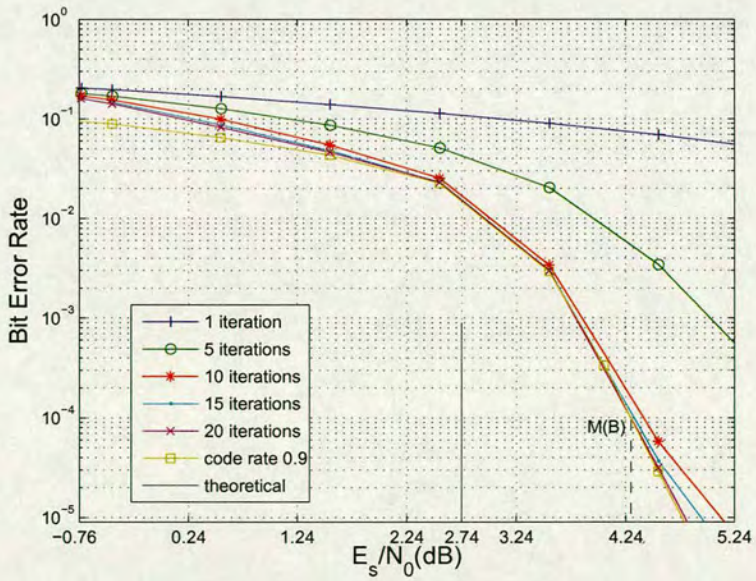

Figure 3.9: *Performance for user B with rate 0.9 in the rate pair* $m(0.9, 0.4)$ *under the iterative LDPC soft SIC and decoding scheme. The user B with weaker power is the last to be detected. The performance for the 0.9 code in the single user AWGN channel is provided for comparison. The theoretical SNR for code rate 0.9 in the binary-input channel is also indicated.* $M(B)$ *corresponds to signal B of point* $M(0.9, 0.4)$ *listed in Table 3.2.*

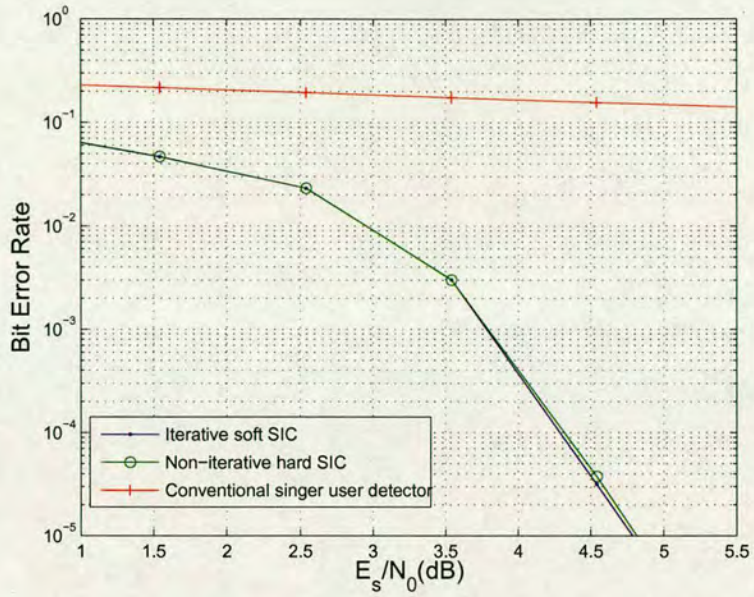

Figure 3.10: *Comparison between the performance of the second detected user B with rate 0.9 in the rate pair $m(0.9, 0.4)$ under the conventional single-user detection scheme and the iterative LDPC soft SIC and decoding scheme. The non-iterative SIC is also provided for comparison.*

ations. The performance for user A with rate 0.4 under the LDPC coded Turbo MUD scheme is shown in Figure 3.8; the performance for user B with rate 0.9 under the LDPC coded Turbo MUD scheme is shown in Figure 3.9 which indicates that as the SNR increases, the performance for the second detected user B of this rate pair goes towards the single-user performance. Figure 3.10 demonstrates the advantage of user B's performance under the LDPC coded Turbo MUD scheme over that under the conventional single-user detection scheme. Figure 3.10 also illustrates the comparison between iterative soft SIC and non-iterative hard SIC (where the iterative decoding process is only conducted within the local decoder structure after the multiuser detection operation). We can see that the iterative soft SIC provides slightly better performance compared to the non-iterative SIC in the waterfall region. The improvements are not as large as we expected which raises the issue regarding the tradeoff between global iterations (information exchange between different component codes) and local iterations (iterations within a single LDPC component). Since the SPA decoder itself is suboptimal, there is no fixed answer to the best ratio among global iterations and local iterations. Possible improvements may be achieved by the global optimizations, *e.g.* density evolution or EXIT charts for both component codes. The complexity to pay for the better performance of soft SIC is that the operations

in (3.22), (3.29) and (3.30) are carried out multiple times (N_{it} times). The way to reduce the complexity is as indicated at the end of Section 3.3, where the multiuser detector formulations derived for BPSK modulation could lend to an efficient implementation by L-value algebra.

Point $M(0.9, 0.4)$

The simulation result shows that by using the LDPC coded Turbo MUD scheme, the practical BER for considered SNR pair $M(0.9, 0.4)$ could fall within the targeted region which is not above 10^{-4} as shown in Figure 3.8 and Figure 3.9.

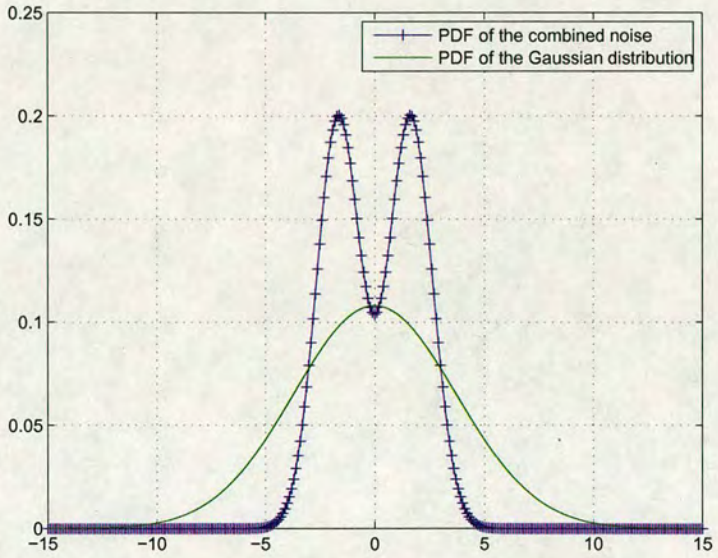

Figure 3.11: *PDF of the combined noise which is composed of Gaussian noise plus Bernoulli distributed signal from the other user. The PDF of the single Gaussian noise of equivalent power is also shown for comparison.*

n(0.9, 0.5)

In the rate pair $n(0.9, 0.5)$, the performance for user A with rate 0.5 under the LDPC coded Turbo MUD scheme is shown in Figure 3.12; the performance for user B with rate 0.9 under the LDPC coded Turbo MUD scheme is shown in Figure 3.13. Still, the user A with stronger power is the first to be detected and the iteration between the two decoder SISO modules via the APP-demodulator operates up to 20 times. We could also see from Figure 3.13 that the performance for the second detected user B under this power relation cannot converge to the single-user performance. Figure 3.14 demonstrates the advantage of user B's performance under the LDPC coded Turbo MUD scheme over that under the conventional single-user detection scheme.

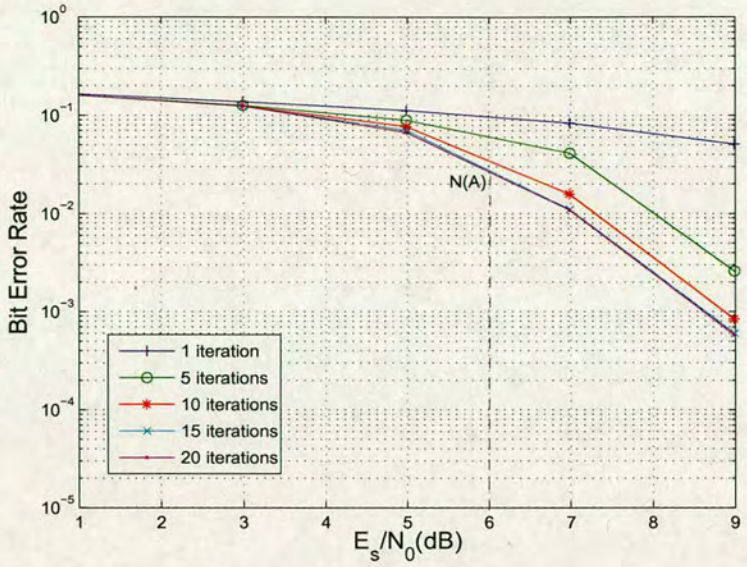

Figure 3.12: *Performance for user A with rate 0.5 in the rate pair $n(0.9, 0.5)$ under the iterative LDPC soft SIC and decoding scheme. The user A with stronger power is the first to be detected. $N(A)$ corresponds to signal A of point $N(0.9, 0.5)$ listed in Table 3.2.*

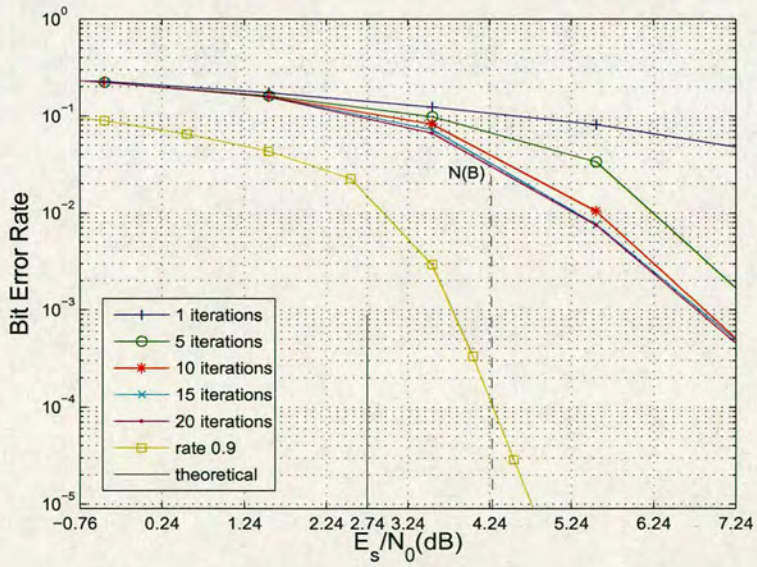

Figure 3.13: *Performance for user B with rate 0.9 in the rate pair $n(0.9, 0.5)$ under the iterative LDPC soft SIC and decoding scheme. The user B with weaker power is the last to be detected. The performance for the 0.9 code in the single user AWGN channel is also provided for comparison. The theoretical SNR for code rate 0.9 in the binary-input channel is indicated. $N(B)$ corresponds to signal B of point $N(0.9, 0.5)$ listed in Table 3.2.*

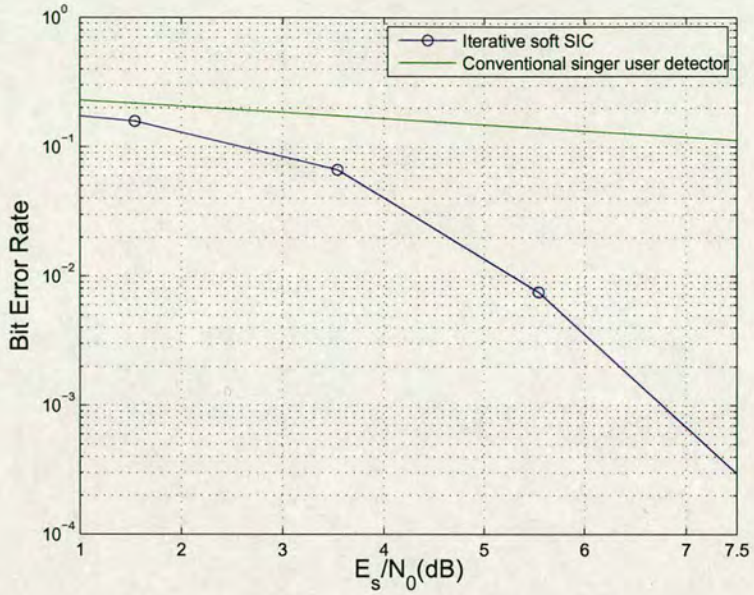

Figure 3.14: *Comparison between the performance of the second detected user B with rate 0.9 in the rate pair $n(0.9, 0.5)$ under the conventional single-user detection scheme and the iterative LDPC soft SIC and decoding scheme.*

Point $N(0.9, 0.5)$

According to the theoretical capacity calculation, the considered point $N(0.9, 0.5)$ with specific SNR values should be within the achievable region. However the practical BERs for the point $N(0.9, 0.5)$ are much higher than 10^{-4} as listed in Table 3.2. The equivalent SNR for the first decoded user A in the multiple access channel is E_s/N_0 of 0.39dB under the assumption that the background Gaussian channel noise (variance 0dB) and the signal B are Gaussian interference to the signal A. But for the middle blocksize rate 0.5 LDPC code, an E_s/N_0 of -0.61dB is enough for error free transmission in single-user AWGN channel. It should be noticed that for both users, the noise with power $(\sigma^2 + P_b)$ is the superposition of channel noise satisfying Gaussian distribution and the signal B that consists of random BPSK data which takes value $\sqrt{P_b}$ with probability p and $-\sqrt{P_b}$ with probability $(1-p)$ (usually p is 0.5 for random data). The pdf of this combined noise with power $(\sigma^2 + P_b)$ is

$$g(x) = p \cdot f(x + \sqrt{P_b}) + (1-p) \cdot f(x - \sqrt{P_b}), \tag{3.48}$$

in which $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x)^2}$. From equation (3.48), as shown in Figure 3.11, we could see that it is no longer Gaussian distributed noise to the detected signal. The distribution of

signal A is also random BPSK data taking value $\sqrt{P_a}$ with probability p and taking $-\sqrt{P_a}$ with probability $(1-p)$. The superposition of the symbols in signal A and the symbols in signal B can be constructive (of the same sign) or destructive (of the opposite sign) with the probability of 0.5 respectively. In the case of destructive addition, all the symbols in signal B cancel the symbols in signal A which results in an equivalent SNR for the detected signal A of -8.66dB (E_s/N_0). This point shows that when employing BPSK modulation in the two-user MAC, the weakness of the assumption with regard to the Gaussian input symbol in the theory is clear to see.

o(0.4, 0.9)

Now we reverse the order of the detection in which we will detect signal B with the weaker power first. According to the capacity calculation in this scenario, the maximum code rate for user A by the theoretical calculation should be 0.97. Here the code rate 0.9 is used instead in order to analyse the impact by reversing the order of detection. In the rate pair $o(0.4, 0.9)$, the user B with weaker power but lower rate is the first to be detected. Then the message flows between the two decoder SISO modules via the APP-demodulator up to 20 iterations. The performance for user B with rate 0.4 under the LDPC coded Turbo MUD scheme is shown in Figure 3.15; the performance for the stronger user A with rate 0.9 under the LDPC coded Turbo MUD scheme is shown in Figure 3.16 which indicates that the performance for the user A which is stronger but second detected cannot converge with the single-user performance. Figure 3.17 demonstrates the advantage of user B's performance under the LDPC coded Turbo MUD scheme over that under the conventional single-user detection scheme.

Point $O_1(0.4, 0.9)$ & *Point* $O_2(0.4, 0.9)$

The BERs for the considered SNR point $O_1(0.4, 0.9)$ are above 10^{-4}. This could still be explained following the imperfect Gaussian input assumption for the two-user BPSK case. When we increase SNR pair to E_s/N_0 of (6.93dB, 5.17dB) which corresponds to the point $O_2(0.4, 0.9)$ as shown in Table 3.2, the BERs can fall within the target region which is not above 10^{-4}. The SNR value of the point $O_2(0.4, 0.9)$ is higher than the SNR values for the point $M(0.9, 0.4)$ to achieve the BER of the same order. This actually shows that the achievable rate regions under different detection orders are not symmetric. The "gap" of the two users' achievable rates are much bigger when the weaker user is first detected compared to that when the stronger user is first detected. Here, the power of one user is only 1.76dB difference. If the two signals are of much disparate power received at the destination, *e.g.*,

20dB, the weaker user can only achieve very low rate if it is the first to be detected [TV05].

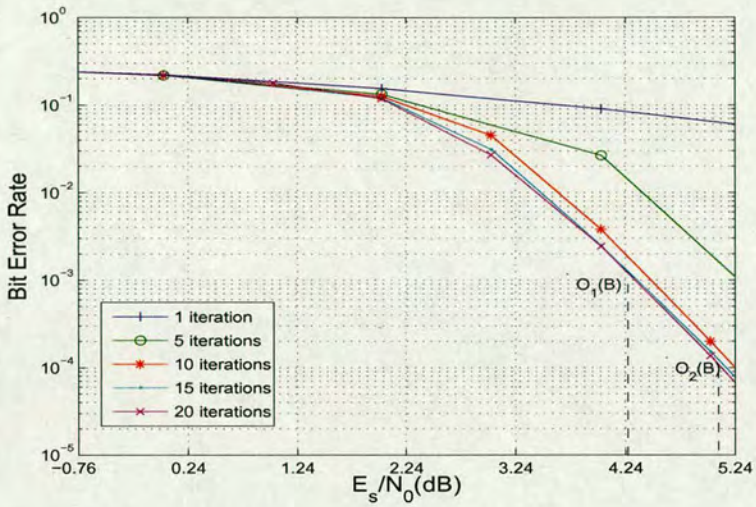

Figure 3.15: *Performance for user B with rate 0.4 in the rate pair $o(0.4, 0.9)$ under the iterative LDPC soft SIC and decoding scheme. The user B with weaker power is the first to be detected. $O_1(B)$ and $O_2(B)$ respectively correspond to signal B of point $O_1(0.4, 0.9)$ and $O_2(0.4, 0.9)$ listed in Table 3.2.*

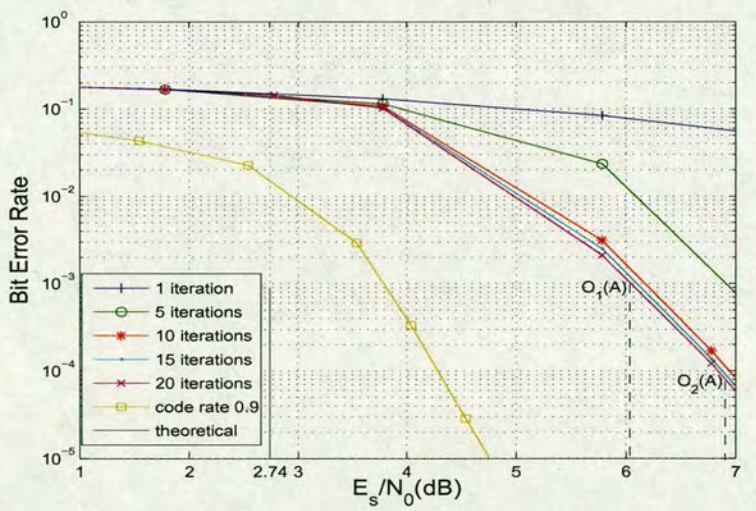

Figure 3.16: *Performance for user A with rate 0.9 in the rate pair $o(0.4, 0.9)$ under the iterative LDPC soft SIC and decoding scheme. The user A with stronger power is the last to be detected. The performance for the 0.9 code in the single user AWGN channel is provided for comparison. The theoretical SNR for code rate 0.9 in the binary-input channel is also indicated. $O_1(A)$ and $O_2(A)$ respectively correspond to signal A of point $O_1(0.4, 0.9)$ and $O_2(0.4, 0.9)$ listed in Table 3.2.*

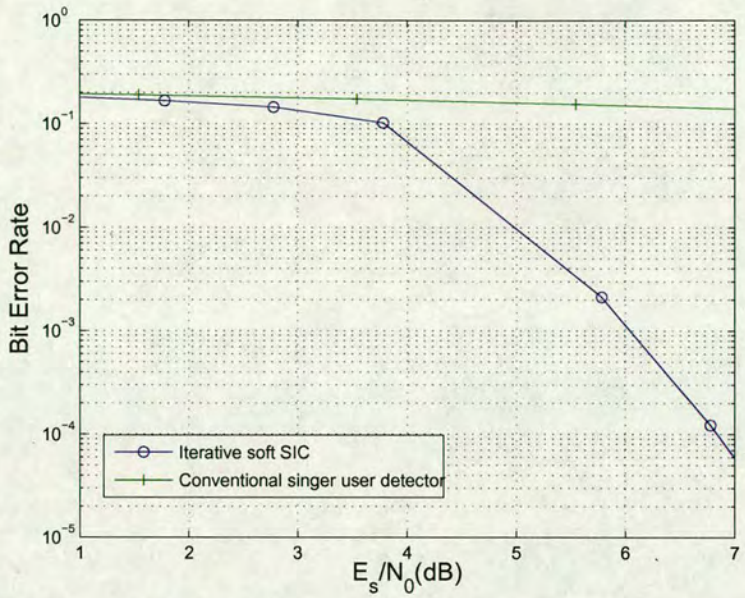

Figure 3.17: *Comparison between the performance of the user A with rate 0.9 in the rate pair* $o(0.4, 0.9)$ *under the conventional single-user detection scheme and the iterative LDPC soft SIC and decoding scheme.*

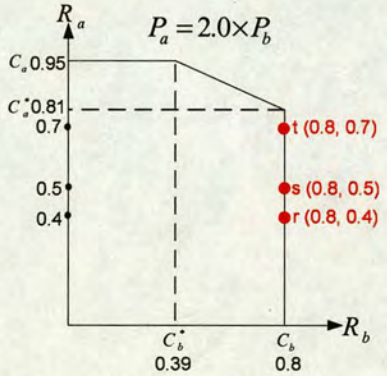

Figure 3.18: *Illustration of the four boundary rates* (C_a, C_a^*, C_b, C_b^*) *and three rate pairs* ($r(0.8, 0.4)$, $s(0.8, 0.5)$, $t(0.8, 0.7)$) *under the scenario* $P_a = 2.0 \times P_b$.

3.4.2.2 Second scenario: $P_a = 2.0 \times P_b$

In this scenario, the received power of signal A is twice that of the received power of signal B. The four boundary rates (C_a, C_a^*, C_b, C_b^*) and three rate pairs ($r(0.8, 0.4)$, $s(0.8, 0.5)$, $t(0.8, 0.7)$) are marked in Figure 3.18. The SNRs and BERs performance of the examined points $R(0.8, 0.4)$, $S(0.8, 0.5)$ and $T(0.8, 0.7)$ corresponding to the rate pairs $r(0.8, 0.4)$, $s(0.8, 0.5)$ and $t(0.8, 0.7)$ respectively are listed in Table 3.3. By detecting the stronger signal

Point (B, A)	SNR & BER	Signal A	Signal B
$R(0.8, 0.4)$	E_s/N_0 (dB)	6.02	3.01
	E_b/N_0 (dB)	9.99	3.98
	BER	8.01×10^{-6}	9.66×10^{-6}
$S(0.8, 0.5)$	E_s/N_0 (dB)	6.02	3.01
	E_b/N_0 (dB)	9.03	3.98
	BER	8.85×10^{-6}	1.27×10^{-5}
$T(0.8, 0.7)$	E_s/N_0 (dB)	6.02	3.01
	E_b/N_0 (dB)	7.57	3.98
	BER	0.082	0.128

Table 3.3: *The BERs of three points* $R(0.8, 0.4)$, $S(0.8, 0.5)$ *and* $T(0.8, 0.7)$ *corresponding to the rate pairs of* $r(0.8, 0.4)$, $s(0.8, 0.5)$ *and* $t(0.8, 0.7)$ *with specific SNR values under the scenario* $P_a = 2.0 \times P_b$.

A first and then applying iterative SIC and decoding, the BER for points $R(0.8, 0.4)$ and $S(0.8, 0.5)$ are within the target region, while the unideal performance at the point (0.8, 0.7) is due to the imperfect Gaussian input assumption for the two-user BPSK case.

3.5 Conclusion

Signal superposition together with SIC is a technique which can achieve the capacity region for the Gaussian MAC. In this chapter, we investigated the performance of SIC when binary modulated signal alphabets are sent by the users and derived a formulation of a multi-user detector which lends itself to an efficient implementation by L-value algebra. We also present a soft iterative SIC and decoding scheme, or called LDPC coded MUD, by factor graph theory where the interference cancellation and decoding are carried out by exchanging the extrinsic L-values of bit-reliability information. A set of practical LDPC codes with moderate blocksize are used to investigate the practical performance of interference cancellation. From the simulations, we can see that the order of the signal detection has an impact on the required signal power to achieve the desired transmission rate. We also see when BPSK modulation is used in the two-user MAC, the assumption of Gaussian input symbols is violated which may lead to degrading performance; this is worth further investigations in future research work.

Chapter 4
Joint Channel and Network Coding for Cooperative Diversity in Shared-Relay Environments

4.1 Introduction

While wireless channels suffer from fading, at the same time the broadcast nature of wireless channels provides the possibility of a third party other than the destination "overhearing" the information that the source transmits. Thus apart from the original transmission channel, the same information could be transmitted to the destination through another independently fading channel. This generated spatial diversity can effectively combat the deleterious effect of fading [SEA03a, SEA03b, NHH04]. Although "coding" has been employed in several papers on cooperative communications [SE04, HN02b, HN06, JHHN04], diversity gain is obtained through channel coding gain by means of single channel codes. In recent years, there has been increasing interest in applying the idea of algebraic code superposition, also called "network coding" to the cooperative communications scenarios [XFKC06b, XFKC06a, XFKC07a, XFKC07b, HSOB05, HD06, HH06, HHK08]. The network coding approach provides an efficient way to generate spatial diversity under the constraint of limited resources. One challenge is the problem of decoder design which should be able to cope with the complex received signal at the destination [XFKC06b, XFKC06a, XFKC07a, XFKC07b, HSOB05, HD06, HH06, HHK08].

The multiple-access relay channel (MARC) (as shown in Figure 4.1) is a model for network topologies where multiple sources communicate with a single destination in the presence of a relay node. In [KW00], MARC was introduced and the capacity outer and inner bounds for it were derived. Examples of such networks topologies include cellular networks, hybrid wireless LAN/WAN networks, sensor and ad hoc networks where cooperation between the nodes is either undesirable or not possible, but an intermediate relay node can be used to aid communication between the sources and the destination [SKM04]. In [HD06, HSOB05, ZZLL07],

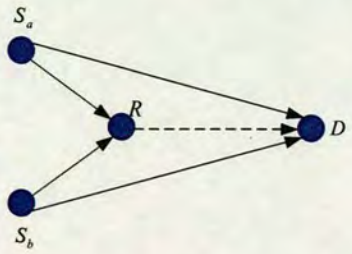

Figure 4.1: *Four-node communications network. Sources S_a and S_b share a common relay R as well as having direct links to the destination D.*

a practical decoding scheme was considered under network-coding-based MARC where the packets from the two sources are linearly combined at the relay. In [CKL06], it was shown that theoretically network-coding-based MARC provides a performance improvement over similar systems without network coding, with less hardware and a lower bandwidth cost. In our work we also focus on this cooperative transmission model for the situation where it is impractical for one mobile user to "capture" the other user's signal during its uplink transmission in the cellular network. Moreover, relay-based cooperative processing provides greater security than direct user cooperation in which user information must be shared.

[XFKC06a] proposed the convolutional code method to decode algebraically superposed codeword, also called a "nested code" [XFKC06b], where a 64-state convolutional code is required for the XORed codeword composed of two 8-state encoders at the destination node. In [XFKC07a], a code superposition scheme employing low-density generator matrix (LDGM) codes is proposed to reduce the decoding complexity at the destination. But in order to do the graph-based decoding, the systematic bits must be retained without superposition which means that the potential superposition diversity is lower than that obtainable from fully superposed codewords. In [HSOB05], a combined low-density parity-check (LDPC) code construction scheme including two channel code components and one network code component is produced by random parity-check matrix generation under certain constraints. The network codes are actually the parity checks for two channel codewords; this necessitates more complicated relay operations than simple superposition. In [BL05] and [BL06], the authors proposed an adaptive network coded cooperation (ANCC) scheme, which couples the instantaneous network topology with codes on graphs, such as LDPC codes. Since the number of terminals engaged in operation decides the number of variable nodes in the final parity check matrix which employs the message-passing algorithm for decoding, this method aims to target cooperation among a large number of terminals under the wireless ad hoc network scenario. The decoding method

of ANCC cannot deal with the challenge of cooperation among a small number of users, *e.g.* two users.

In this chapter, we will present two pieces of work corresponding to two scenarios under the two-source one-relay network-coded MARC transmission model as shown in Figure 4.1. In the first scenario, we propose a cooperative coding scheme which, in contrast to previous work where TDMA-only or FDMA-only relaying is assumed, allows continuous transmission of superposed codewords by the relay, thus making efficient use of communication resources to leverage spatial diversity gains. The proposed scheme also arises from the observation that idle frequency channels do exist and could be exploited in current TDMA systems. The second scenario is where the source-destination transmission and relay-destination transmission use a fading multiple access channel (MAC). In this scenario we propose a cooperative coding scheme which is different from the previous work of [XFKC06a, XFKC07b, XFKC07a, HD06, HSOB05] where superposed codewords experience a channel orthogonal to that of the original transmission, and also different from the previous work of [KAA04] where simple codeword retransmission is employed in the multiple access Gaussian relay channel. Our scheme allows continuous transmission of superposed codewords by the relay and at the same time targets the challenge of coping with the interference introduced by the multiple access channel, thus making efficient use of communication resources to leverage spatial diversity gains.

The major contributions of this chapter come from the factor graph based decoding algorithm at the destination node for the above two scenarios. We shall start with the simpler first scenario. In this part of the chapter, we are going to describe a novel and efficient decoding algorithm based on message passing on the destination node's factor graph for the purpose of exploiting the spatial diversity contained in the algebraically superposed codewords. The algorithm attains a separation of the two soft-input soft-output (SISO) decoder modules corresponding to the codes employed by the two sources; for convolutional codes, this separation affords a complexity advantage over decoding of the "nested code" [XFKC06a, XFKC07b]; for LDPC codes, it affords a more efficient Tanner graph schedule than fully parallel decoding [HSOB05]. In the second scenario, we shall make some further development based on the previous decoding scenario. We also detail a novel computationally efficient decoding algorithm based on message passing on the destination node's factor graph for the purpose of exploiting the spatial diversity contained in the algebraically superposed codewords together with the signal superposition introduced by the fading multiple access channel. The corresponding sliding-window

factor graph based decoding algorithm attains a separation of the three soft-input soft-output (SISO) decoder modules with respect to each received signal stream; we believe this work is the first practical decoding scheme dealing with the extraction of the substantial spatial diversity contained in the code superposition together with the signal superposition.

The main body of this chapter is divided into two sections with respect to the proposed cooperative coding schemes under two scenarios. Within each section, first the relative backgrounds will be introduced; then we present the proposed cooperative coding scheme; next we shall address in detail the corresponding decoding algorithm at the destination node; a theoretical lower bound on the frame (codeword) error rate (FER) of the system will also be derived; finally a simulation-based comparison of the proposed scheme and the reference schemes will be given.

4.2 Proposed Cooperative Coding Scheme in a Shared Relay Environment with an Extra Available Frequency Band

4.2.1 Background

In this subsection, we will first briefly review the relations between error control coding and cooperative communications, and also the relation between network coding and cooperative communications. Then we shall describe the motivation behind our work in this section.

4.2.1.1 Error control coding vs. Cooperative communications

It has long been known that path diversity is able to improve the effective SNR of a fading wireless channel [Jak74]. [Ala98, TSC98, TJC99] showed that transmitters who possess multiple antennas can use block and trellis space-time codes to reduce the bit and frame error probability and improve the end-to-end system performance. However, when mobiles fail to support multiple antennas due to size or other constraints, the above methods cannot be used to provide uplink transmit diversity. The idea of user cooperation was born due to the observation that the broadcast nature of wireless channels provides the possibility of a third party other than the destination receiving the information that the source transmits. Apart from the original transmission channel, the same information could be transmitted to the destination through another independently fading channel. This generates a virtual MIMO system which can effectively combat the deleterious effect of fading. The fundamental ideas behind coop-

eration can date back to the work of [CG79] on the relay channel and the work of [Wil83] on the multiple access channel. However, the earliest work specifically on user cooperation is [SEA98, SEA03a, SEA03b] where the achievable rate regions and outage probabilities are examined under an information-theoretic model and a CDMA implementation is presented.

The first framework where channel coding was used in the user cooperation strategy is called *coded cooperation* [HN02b, HN02a, HN06] which we already introduced in Section 2.3.2.2. In a word, the key idea behind coded cooperation is that each user tries to transmit redundant parity bits for its partner. Later, an extension of coded cooperation was proposed in [JHHN04] where space-time cooperation is applied in the second segment transmission and Turbo codes are implemented for this scheme. In [ZV03, LVWD06, CV05, CdBSA05] distributed channel codes are applied to exploit the single channel coding gain by "combining" the redundancy introduced in different signal segments by transmissions from source and relay separately. Thus, the distributed channel code method for the relay channel could be seen as a joint routing-channel coding scheme [HD06].

4.2.1.2 Network coding vs. Cooperative communications

Forward error control coding, which allows the receiver to detect and correct errors (within some bound) without the need to ask the sender for additional data, was largely limited to point-to-point communication links [LC04]. The traditional layered approach in wireless network applications limited error control coding techniques to protect the data in a number of point-to-point links under the "decomposition" of the whole network. However, the throughput of a network could be improved by developing error control methodologies that not only concern the point-to-point link but a larger network setting as well.

[ACLY00] proposed the the concept termed *network coding* which aims to increase the achievable throughput in a network. The basic idea is that intermediate nodes in a network are allowed not only to forward the data they received but also to mix their received data from different users before passing it on. [ACLY00] proved that in a multicast transmission with one source it is possible to achieve the min-cut capacity[1] between the source and the sinks with network coding. [LYC03] proved that linear coding, which regards a block of data as a vector over a certain base field and allows a node to apply a linear transformation to a vector before passing

[1] Max-flow min-cut theorem states that the maximum amount of flow is equal to the capacity of a minimum cut [Bol79] in a network.

it on, suffices to achieve the max-flow min-cut throughput for multicast transmissions. [KM03] extended the network coding framework from multicast networks to arbitrary networks and robust networking. [KKH+05] showed that by using opportunistic network coding scheme COPE, where data packets are XORed at the network layer, the network throughput can be significantly improved compared to the current 802.11 mesh network. [ZLL06] addressed electromagnetic-wave-based network coding at the physical layer and demonstrates its ability to enhance the throughput performance of multi-hop wireless networks. [DEH+05,RSW03,KRH+08] studied the application of network coding in wireless networks. It is worth mentioning that network coding, whose key feature is to encourage the relay to forward the mixture of its observations, is naturally suited to wireless communications environments to generate cooperative diversity. [DLGT07] studied the combination of the network coding and relay selection to combat fading and exploit the dynamic nature of the wireless environment. In [CKL06,WK07], the outage probability results of an network-coding-based adaptive decode-and-forward schemes are presented.

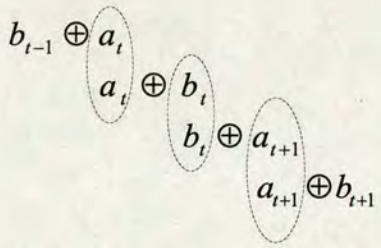

Figure 4.2: *Illustration for a typical decoding scenario at the destination node introduced by XOR operation. Multiple codewords received at the destination can be viewed as a chain.* $\mathbf{a}_t$ *and* $\mathbf{b}_t$ *represent codewords generated in the time slot t from source a and source b respectively.*

4.2.1.3 Motivation

The above work with regard to network coding is either from an information theoretical perspective or focused on a single layer of the network. The potential benefit arising from the interaction of physical layer channel coding and network layer network coding was identified by [WCK05] where the exchange of independent information between two nodes in a wireless network can be efficiently performed by exploiting network coding and the physical layer broadcast property offered by the wireless medium. Also joint network-channel coding [HSOB05] and nested codes [XFKC06b] were proposed with practical decoding methods. It was shown from [HSOB05, XFKC06b] that the redundancy in the network code could be

used to support the channel code for better protection as well as carrying new information which helps to generate spatial diversity. The joint channel and network coding approach provides an efficient way to generate spatial diversity with limited power resources. However one problem is the decoder design at the destination which should be able to cope with the complicated decoding scenario where typically, the pattern of "stacked" codewords appears as depicted in Figure 4.2. The convolutional code method to decode this algebraic superposition was proposed in [XFKC06a] where there is a price to pay in complexity at the destination node since a 64-state convolutional code is required for the XORed codeword composed of two 8-state encoders where the complexity for the BCJR algorithm [BCJR74] at the destination node is squared compared with the decoding of the single component code. In order to reduce the complexity, [XFKC07a] proposed an algebraic superposition scheme by means of LDGM codes [GFZ03] where the decoding complexity for XORed codeword grows linearly rather than quadratically with the component code complexity. By the definition for the generator matrix of the LDGM codes, the information bits can be naturally separated from the superposition and only the parity bits are superposed as done in [XFKC07a]. The authors in [XFKC06a,XFKC07a] tried to tackle this decoding problem from the perspective of the generator matrix corresponding to the XORed codewords. Following this process, it will be difficult to apply LDPC codes since knowing the $\mathbf{G}$ matrix first and then recovering the structure of the $\mathbf{H}$ matrix is theoretically achievable but computationally intractable. Inspired by this, our decoding scheme comes from the perspective of the resulting factor graph presentation for the $\mathbf{H}$ matrix where LDPC codes naturally fit.

4.2.2 System model

We consider the four-node communications network depicted in Figure 4.1, with two sources S_a and S_b, one relay R, and one destination D common to the two sources. The assumption for the system model is that all the point-to-point links are flat-fading.

4.2.2.1 System overview

The communication period is divided into $L+1$ *time slots* $t = 0, 1, \cdots, L$; each time slot $t \in \{0, 1, \cdots, L\}$ is further subdivided into 2 *half slots* ($2t$, $2t+1$) as presented in Table 4.1. Source S_a has L messages to transmit, which it encodes into L n-bit codewords $\{\mathbf{a}_t : t = 0, 1, \cdots, L-1\}$. The code used at source S_a is $\mathcal{C}_a$ and is defined by the $m_a \times n$

Time Slot		0		1		...	t		t+1		...	L-1		L	
Half Slot		0	1	2	3	...	2t	2t+1	2t+2	2t+3	...	2L-2	2L-1	2L	2L+1
Relay Receives	From Source (f1)	$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_1$	$\mathbf{b}_1$	...	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_{t+1}$	$\mathbf{b}_{t+1}$	...	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$		
Destination Receives	From Source (f1)	$\mathbf{c}_0=$ $\mathbf{a}_0$	$\mathbf{c}_1=$ $\mathbf{b}_0$	$\mathbf{c}_2=$ $\mathbf{a}_1$	$\mathbf{c}_3=$ $\mathbf{b}_1$	...	$\mathbf{c}_{2t}=$ $\mathbf{a}_t$	$\mathbf{c}_{2t+1}=$ $\mathbf{b}_t$	$\mathbf{c}_{2t+2}=$ $\mathbf{a}_{t+1}$	$\mathbf{c}_{2t+3}=$ $\mathbf{b}_{t+1}$	...	$\mathbf{c}_{2L-2}=$ $\mathbf{a}_{L-1}$	$\mathbf{c}_{2L-1}=$ $\mathbf{b}_{L-1}$		
	From Relay (f2)		$\mathbf{d}_1=$ $\pi(\mathbf{a}_0)$	$\mathbf{d}_2=$ $\pi(\mathbf{b}_0)$ $\oplus\mathbf{a}_0$	$\mathbf{d}_3=$ $\pi(\mathbf{a}_1)$ $\oplus\mathbf{b}_0$	...	$\mathbf{d}_{2t}=$ $\pi(\mathbf{b}_{t-1})$ $\oplus\mathbf{a}_{t-1}$	$\mathbf{d}_{2t+1}=$ $\pi(\mathbf{a}_t)$ $\oplus\mathbf{b}_{t-1}$	$\mathbf{d}_{2t+2}=$ $\pi(\mathbf{b}_t)$ $\oplus\mathbf{a}_t$	$\mathbf{d}_{2t+3}=$ $\pi(\mathbf{a}_{t+1})$ $\oplus\mathbf{b}_t$	...	$\mathbf{d}_{2L-2}=$ $\pi(\mathbf{b}_{L-2})$ $\oplus\mathbf{a}_{L-2}$	$\mathbf{d}_{2L-1}=$ $\pi(\mathbf{a}_{L-1})$ $\oplus\mathbf{b}_{L-2}$	$\mathbf{d}_{2L}=$ $\pi(\mathbf{b}_{L-1})$ $\oplus\mathbf{a}_{L-1}$	$\mathbf{d}_{2L+1}=$ $\mathbf{b}_{L-1}$
Destination Decodes				$\mathbf{a}_0$	$\mathbf{b}_0$	...	$\mathbf{a}_{t-1}$	$\mathbf{b}_{t-1}$	$\mathbf{a}_t$	$\mathbf{b}_t$	...	$\mathbf{a}_{L-2}$	$\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$

Table 4.1: *Transmission schedule of proposed cooperative coding scheme.*

parity-check matrix $\mathbf{H}_a = (H_a(j,i))$. Similarly, source S_b has L messages to transmit, which it encodes into L n-bit codewords $\{\mathbf{b}_t : t = 0, 1, \cdots, L-1\}$. The code used at source S_b is $\mathcal{C}_b$ and is defined by the $m_b \times n$ parity-check matrix $\mathbf{H}_b = (H_b(j,i))$. Thus, the codes $\mathcal{C}_a$ and $\mathcal{C}_b$ have the same length but not necessarily the same rate. In general, the codes used at the two sources can be LDPC or convolutional; in this chapter we concentrate on LDPC codes. S_a and S_b broadcast their modulated codewords to the relay and destination nodes using TDMA in frequency band f_1 and there is no cooperation between the two sources (neither S_a nor S_b could capture the signal broadcast from the other source). For each $t \in \{0, 1, \cdots, 2L-1\}$, let $\mathbf{c}_t$ denote the codeword broadcast by the source in half slot t; thus $\mathbf{c}_{2t} = \mathbf{a}_t$ and $\mathbf{c}_{2t+1} = \mathbf{b}_t$ for $t \in \{0, 1, \cdots, L-1\}$.

4.2.2.2 Relaying protocol

The relay decodes and then re-encodes each codeword received from the source (the cooperative scheme is based on the scenario where the source is quite close to the relay). The relay also has a buffer in which it stores the codewords it obtained in the previous two half slots. At each half slot t ($t = 2, 3, \cdots, 2L$), the relay interleaves the codeword received in half slot $t-1$ and superposes it (XOR operation) with the codeword received in half slot $t-2$; it then transmits the resulting codeword to the destination in frequency band f_2. Special cases arise at half slots 1 and $2L+1$ in which only a single codeword is stored at the relay and no XOR operation is performed. The allocation of different frequency bands to the transmission channel of the relay and the broadcast channels of the sources allows for simultaneous transmission and reception by the relay node, thus allowing efficient leveraging of communication resources for spatial diversity. Let $\mathbf{d}_t$ denote the codeword transmitted from the relay to the destination in

half slot $t \in \{1, 2, \cdots, 2L+1\}$; thus $\mathbf{d}_t = \pi(\mathbf{c}_{t-1}) \oplus \mathbf{c}_{t-2}$, for $t = 2, 3, \cdots 2L$. For each $t = 0, 1, \cdots L-1$, source S_a's codeword $\mathbf{a}_t$ is decoded at the end of half slot $2t+2$ and source S_b's codeword $\mathbf{b}_t$ is decoded at the end of half slot $2t+3$. The transmission schedule for this cooperative coding scheme is illustrated in Table 4.1. It may be seen from Table 4.1 that spatial diversity for each message is contained in three codewords received at the destination.

Note that decoding and re-encoding of received codewords using a different code (as in the scheme of [XFKC06a]) is not performed by the relay in this scheme; this is in order to keep the relay operation as simple as possible. The interleaver π actually provides the "interleaver gain" for decoding at the destination. The interleaver is not in general necessary in the case of LDPC coding; however it may be used to avoid a large multiplicity of 8-cycles in the Tanner graph for the case where $\mathbf{H}_a \doteq \mathbf{H}_b$.

4.2.2.3 Reference systems

In order to demonstrate the benefit provided by network coding in the cooperative communications, we provide two reference systems. The first reference system is the consecutive relaying (see Table 4.2), in which the time and frequency allocations are the same as in the proposed cooperative scheme, except that the relay simply retransmits the previously received codeword rather than a codeword superposition. The second reference system is the simple TDMA/FDMA based code superposition relaying scheme depicted in Tables 4.3 and 4.4 respectively. It is easily seen that in these schemes, spatial diversity for each message is provided by two codewords received at the destination. A simulation-based comparison of all schemes described in this section under a transmit power constraint will be given in Section 4.2.5.

Time Slot		0		1		...	t		t+1		...	L-1		
Half Slot		0	1	2	3	...	2t	2t+1	2t+2	2t+3	...	2L-2	2L-1	2L
Destination Receives	From Source (f1)	$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_1$	$\mathbf{b}_1$	...	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_{t+1}$	$\mathbf{b}_{t+1}$	...	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$	
	From Relay (f2)		$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_1$	...	$\mathbf{b}_{t-1}$	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_{t+1}$	...	$\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$

Table 4.2: *Transmission schedule of consecutive relaying scheme.*

Time Slot										
0	1	2	...	3t	3t+1	3t+2	...	3L-3	3L-2	3L-1
$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_0 \oplus \pi(\mathbf{b}_0)$	...	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_t \oplus \pi(\mathbf{b}_t)$	...	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$	$\mathbf{a}_{L-1} \oplus \pi(\mathbf{b}_{L-1})$

Table 4.3: *Section of transmission schedule for the TDMA relaying scheme.*

	0	1	...	t	t+1	...	L-1	
Frequency Band f1	$\mathbf{a}_0$	$\mathbf{a}_1$	...	$\mathbf{a}_t$	$\mathbf{a}_{t+1}$	...	$\mathbf{a}_{L-1}$	
Frequency Band f2	$\mathbf{b}_0$	$\mathbf{b}_1$	...	$\mathbf{b}_t$	$\mathbf{b}_{t+1}$	...	$\mathbf{b}_{L-1}$	
Frequency Band f3		$\mathbf{a}_0 \oplus \pi(\mathbf{b}_0)$	...	$\mathbf{a}_{t-1} \oplus \pi(\mathbf{b}_{t-1})$	$\mathbf{a}_t \oplus \pi(\mathbf{b}_t)$	...	$\mathbf{a}_{L-2} \oplus \pi(\mathbf{b}_{L-2})$	$\mathbf{a}_{L-1} \oplus \pi(\mathbf{b}_{L-1})$

Table 4.4: *Section of transmission schedule for the FDMA relaying scheme.*

4.2.3 Decoding algorithm at destination node

Iterative decoding is adopted to exploit the diversity contained in the joint channel and network codes. The whole process is based on a sliding window structure for the relay transmitted superposed codewords where the output L-values of the decoder in the current frame will be passed on to assist the decoding in the next frame. An evolution of the knowledge on each source transmitted codeword is obtained by the iterative exchanges of the extrinsic information between two component codes of the superposed codeword as well as the sliding window structure spanning on multiple frames.

In the following, we will first describe the decoding principle by the example of codeword $\mathbf{a}_t$ and then present the detailed decoding algorithm based on factor graph theory.

4.2.3.1 Decoding overview

Half Slot		$2t$	$2t+1$	$2t+2$	
Source Transmits (f1)	...	$\mathbf{c}_{2t} = \mathbf{a}_t$	$\mathbf{c}_{2t+1} = \mathbf{b}_t$	$\mathbf{c}_{2t+2} = \mathbf{a}_{t+1}$	...
Relay Transmits (f2)	...	$\mathbf{d}_{2t} = \pi(\mathbf{b}_{t-1}) \oplus \mathbf{a}_{t-1}$	$\mathbf{d}_{2t+1} = \pi(\mathbf{a}_t) \oplus \mathbf{b}_{t-1}$	$\mathbf{d}_{2t+2} = \pi(\mathbf{b}_t) \oplus \mathbf{a}_t$	...
Decoding Evolution Process For $\mathbf{a}_t$		$L_1(c_{2t}^{(i)})$	$L_2(c_{2t}^{(i)})$	$\hat{a}_t^{(i)}$	
		Step 1	Step 2	Step 3	

Table 4.5: *The demonstration for the three-step decoding evolution. Here we take the decoding of* $\mathbf{a}_t$ *as an example. The decoding of* $\mathbf{a}_t$ *spans three transmission frames (half slots)* $2t \rightarrow 2t+1 \rightarrow 2t+2$*, resulting in the three-step evolution* $L_1(c_{2t}^{(i)}) \rightarrow L_2(c_{2t}^{(i)}) \rightarrow \hat{a}_t^{(i)}$.

Without loss of generality, we consider the decoding of codeword $\mathbf{a}_t$ at the end of half slot $2t+2$, for $t \in \{1, 2, \cdots, L-1\}$. As we can see from Table 4.5, there are three copies of $\mathbf{a}_t$ contained in the codewords $\mathbf{c}_{2t}$, $\mathbf{d}_{2t+1}$ and $\mathbf{d}_{2t+2}$ respectively. In the following, we shall use $x^{(i)}$ to denote the ith bit of the codeword $\mathbf{x}$. We also use $L_1(c_t^{(i)})$ and $L_1(d_t^{(i)})$ to denote the respective *a priori* LLRs derived from the corresponding received streams. Also, $L_2(\cdot)$ denotes the (updated) *a posteriori* LLRs which will be used as the *a priori* LLRs in the next decoding. The decoding of $\mathbf{a}_t$ spans three transmission frames (half slots) $2t \rightarrow 2t+1 \rightarrow 2t+2$, resulting in the three-step evolution $L_1(c_{2t}^{(i)}) \rightarrow L_2(c_{2t}^{(i)}) \rightarrow \hat{a}_t^{(i)}$ as illustrated in Table 4.5. In the first step, $L_1(c_{2t}^{(i)})$ is derived from the received stream at the end of the source-destination link. In the second step, the *a priori* $L_1(c_{2t}^{(i)})$ is fed into the current decoder structure for half slot $2t+1$ and the *a posteriori* LLR $\{L_2(c_{2t}^{(i)})\}$ is produced consequently. In the third step, $\{L_2(c_{2t}^{(i)})\}$ will be used as *a priori* LLR in the decoding structure for the current frame.

Now we focus on the decoding structure at the third step (half slot $2t+2$) which is based on the codeword $\mathbf{d}_{2t+2} = \pi(\mathbf{b}_t) \oplus \mathbf{a}_t$. We assume that *a priori* LLRs on $\mathbf{c}_{2t+1}$ and $\mathbf{c}_{2t}$, denoted by $\{L_1(c_{2t+1}^{(i)})\}$ and $\{L_2(c_{2t}^{(i)})\}$ respectively, are available from the channel observation value and previous decoding in the second step. In addition to the decoding of $\mathbf{a}_t$, the decoder will produce *a posteriori* LLRs $\{L_2(c_{2t+1}^{(i)})\}$ which will be used as *a priori* LLRs in the next decoding step.

4.2.3.2 Detailed algorithm

In this section, we provide a concise description of the factor graph based decoding algorithm [KFL01] at the destination decoder. The factor graph for the decoding under the half time slot $2t+2$ is illustrated in Figure 4.3, where circles depict variable nodes and squares depict factor nodes. Extrinsic information is exchanged between the two soft-input soft-output (SISO) decoder modules for the constituent codes $\mathcal{C}_a$ and $\mathcal{C}_b$, via the factor nodes $\{F_i\}$ which correspond to the network coding operation at the relay. For simplicity, the graph is illustrated for the case $n = 3$, and where $\mathcal{C}_a$ and $\mathcal{C}_b$ are (trivial) LDPC codes. For convolutional constituent codes, the two SISO modules execute BCJR algorithms. Note that in the convolutional case, the separation of the two (e.g. M-state) decoder SISO modules gives a complexity advantage over schemes which use a larger (e.g. M^2-state) decoder to decode the "nested" code generated at the relay (see e.g. [XFKC06a]). In the LDPC case this separation of SISO modules yields a more efficient message-passing schedule than implements fully parallel decoding on the Tanner graph of the nested code (see e.g. [HSOB05]).

Next we introduce some notational conventions pertaining to the following algorithm description. In all cases, the letter λ is used to denote LLRs corresponding to messages passed on the factor graph. The interleaving "π" is interpreted as

$$\mathbf{x} = \pi(\mathbf{y}) \quad \Longleftrightarrow \quad x^{(i)} = y^{(\pi(i))}$$

Some index sets are defined as follows. $\mathcal{J}_a = \{1, 2, \cdots m_a\}$; $\mathcal{J}_b = \{1, 2, \cdots m_b\}$; $\mathcal{N}_a(i) = \{j \in \mathcal{J}_a \ : \ H_a(j,i) = 1\}$; $\mathcal{N}_b(i) = \{j \in \mathcal{J}_b \ : \ H_b(j,i) = 1\}$; $\mathcal{M}_a(j) = \{i \in \mathcal{I} \ : \ H_a(j,i) = 1\}$; $\mathcal{M}_b(j) = \{i \in \mathcal{I} \ : \ H_b(j,i) = 1\}$. Also, $\boxplus$ denotes the (commutative and associative) "box-plus" operation [HOP96], i.e.

$$\boxplus_{s \in \mathcal{S}} \lambda_s = \log \left(\frac{1 + \prod_{s \in \mathcal{S}} \tanh(\lambda_s/2)}{1 - \prod_{s \in \mathcal{S}} \tanh(\lambda_s/2)} \right)$$

N denotes the maximum number of decoding iterations.

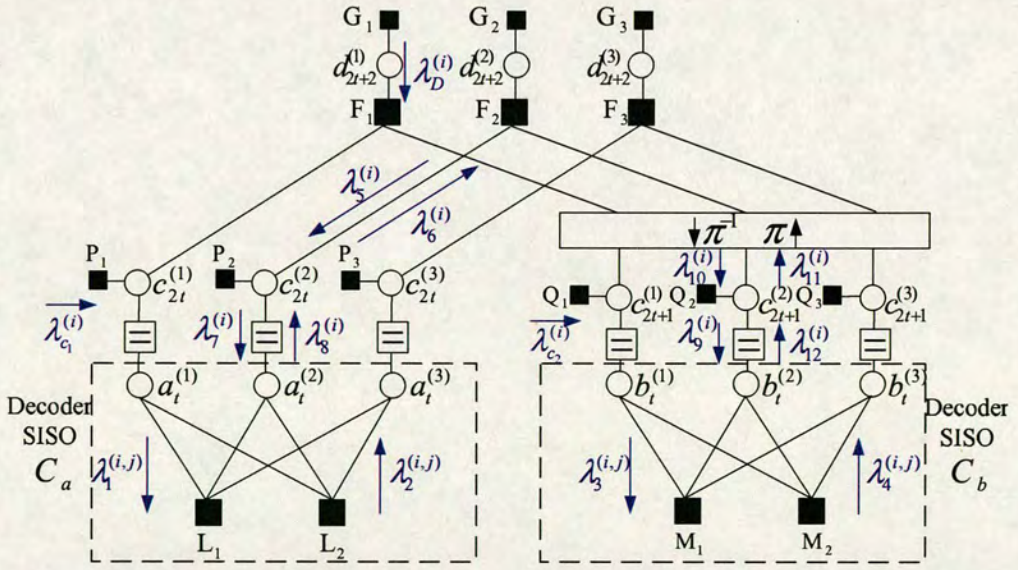

Figure 4.3: *Factor graph corresponding to the destination's decoding of codeword* $\mathbf{a}_t$. *The message-passing schedule is such that extrinsic information is exchanged between the two decoder SISO modules for the constituent codes* C_a *and* C_b, *via the factor nodes* $\{F_i\}$ *corresponding to the network coding operation at the relay. For ease of presentation, the factor graph is illustrated for the case* $n = 3$ *and trivial codes* C_a, C_b.

Factor Graph Based Decoding Algorithm at Destination Node – Decoding of Codeword $\mathbf{a}_t$

Initialization:

- For $i \in \mathcal{I}$,

$$\lambda_{c_1}^{(i)} = L_2(c_{2t}^{(i)}) \tag{4.1}$$
$$\lambda_{c_2}^{(i)} = L_1(c_{2t+1}^{(i)}) \tag{4.2}$$
$$\lambda_{8}^{(i)} = 0 \tag{4.3}$$
$$\lambda_{D}^{(i)} = L_1(d_{2t+2}^{(i)}) \tag{4.4}$$

- For $i \in \mathcal{I}, j \in \mathcal{N}_a(i)$

$$\lambda_2^{(i,j)} = 0 \tag{4.5}$$

- For $i \in \mathcal{I}, j \in \mathcal{N}_b(i)$

$$\lambda_4^{(i,j)} = 0 \tag{4.6}$$

Main Loop: For $k = 1$ to N do

- For $i \in \mathcal{I}$,

$$\lambda_6^{(i)} = \lambda_{c_1}^{(i)} + \lambda_8^{(i)} \tag{4.7}$$

- Network coding constraints: for $i \in \mathcal{I}$,

$$\lambda_{10}^{(\pi(i))} = \lambda_6^{(i)} \boxplus \lambda_D^{(i)} \tag{4.8}$$

- For $i \in \mathcal{I}$,

$$\lambda_9^{(i)} = \lambda_{c_2}^{(i)} + \lambda_{10}^{(i)} \tag{4.9}$$

- SISO decoder $\mathcal{C}_b$: for $i \in \mathcal{I}, j \in \mathcal{N}_b(i)$

$$\lambda_3^{(i,j)} = \lambda_9^{(i)} + \sum_{l \in \mathcal{N}_b(i) \setminus \{j\}} \lambda_4^{(i,l)} \tag{4.10}$$
$$\lambda_4^{(i,j)} = \boxplus_{l \in \mathcal{M}_b(j) \setminus \{i\}} \lambda_3^{(l,j)} \tag{4.11}$$

For $i \in \mathcal{I}$,

$$\lambda_{12}^{(i)} = \sum_{j \in \mathcal{N}_b(i)} \lambda_4^{(i,j)} \tag{4.12}$$

- Obtain the *a posteriori* LLR (prior for decoding in next half slot)

$$L_2(c_{2t+1}^{(i)}) = \lambda_9^{(i)} + \lambda_{12}^{(i)} \tag{4.13}$$

- For $i \in \mathcal{I}$,

$$\lambda_{11}^{(i)} = \lambda_{c_2}^{(i)} + \lambda_{12}^{(i)} \tag{4.14}$$

$$\lambda_5^{(i)} = \lambda_D^{(i)} \boxplus \lambda_{11}^{(\pi(i))} \tag{4.15}$$

- For $i \in \mathcal{I}$,

$$\lambda_7^{(i)} = \lambda_{c_1}^{(i)} + \lambda_5^{(i)} \tag{4.16}$$

- SISO decoder $\mathcal{C}_a$: for $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_1^{(i,j)} = \lambda_7^{(i)} + \sum_{l \in \mathcal{N}_a(i) \setminus \{j\}} \lambda_2^{(i,l)} \tag{4.17}$$

$$\lambda_2^{(i,j)} = \boxplus_{l \in \mathcal{M}_a(j) \setminus \{i\}} \lambda_1^{(l,j)} \tag{4.18}$$

For $i \in \mathcal{I}$,

$$\lambda_8^{(i)} = \sum_{j \in \mathcal{N}_a(i)} \lambda_2^{(i,j)} \tag{4.19}$$

- Calculate *a posteriori* LLRs for codeword $\mathbf{a}_t$:

$$L(a_t^{(i)}) = \lambda_7^{(i)} + \lambda_8^{(i)} \tag{4.20}$$

- Make decisions on the code bits

$$\hat{a}_t^{(i)} = \begin{cases} 0 & \text{if } L(a_t^{(i)}) \geq 0 \\ 1 & \text{if } L(a_t^{(i)}) < 0 \end{cases}$$

If $\hat{\mathbf{a}}_t \mathbf{H}_a^T = \mathbf{0}$ then *break;*

Endfor

For the special case of decoding the first codeword $\mathbf{a}_0$, we set

$$L_2(c_0^{(\pi(i))}) = L_1(c_0^{(\pi(i))}) + L_1(d_1^{(i)}) \tag{4.21}$$

for all $i \in \mathcal{I}$ in Equation (4.1). Also, for the special case of decoding the final codeword $\mathbf{b}_{L-1}$, further modifications are made to the algorithm as follows:

- Equation (4.2), (4.6), (4.7)-(4.14), (4.15) are deleted
- Equation (4.16) is replaced by

$$\lambda_7^{(i)} = \lambda_{c_1}^{(i)} \tag{4.22}$$

4.2.4 A lower bound on the system frame error rate

In this section a theoretical lower bound on the frame (codeword) error rate (FER) of the proposed scheme in Table 4.1 is derived; the analysis follows the lines of [XFKC07b].

Half Slot		$2t$	$2t+1$	$2t+2$	$2t+3$	
Source-Destination Link (f1)	...	$\mathbf{c}_{2t} = \mathbf{a}_t$	$\mathbf{c}_{2t+1} = \mathbf{b}_t$	...	...	...
Relay-Destination Link (f2)	...	...	$\mathbf{d}_{2t+1} = \pi(\mathbf{a}_t) \oplus \mathbf{b}_{t-1}$	$\mathbf{d}_{2t+2} = \pi(\mathbf{b}_t) \oplus \mathbf{a}_t$	$\mathbf{d}_{2t+3} = \pi(\mathbf{a}_{t+1}) \oplus \mathbf{b}_t$	...

Table 4.6: *Illustration for the five received codewords at the destination which contain the information of* $\mathbf{a}_t$ *and* $\mathbf{b}_t$.

In the case where an extra frequency band is used, let us consider the decoding of $\mathbf{a}_t$ and $\mathbf{b}_t$, transmitted from the sources in half slots $2t$ and $2t+1$ respectively. As illustrated in Table 4.6, the codewords received by the destination from the sources in half slots $2t$ and $2t+1$ contain information only relating to $\mathbf{a}_t$ and $\mathbf{b}_t$. Due to latency constraints embodied in the relay operation, the message of $\mathbf{a}_t$ is contained in $\mathbf{d}_{2t+1}$ and $\mathbf{d}_{2t+2}$; in the same way, the message of $\mathbf{b}_t$ is also contained in $\mathbf{d}_{2t+2}$ and $\mathbf{d}_{2t+3}$. The received information which may be used in decoding $\mathbf{a}_t$ and $\mathbf{b}_t$ is contained in five received codewords at the destination – transmissions from sources to destination in half slots $2t$ and $2t+1$, and transmissions from relay to destination in half slots $2t+1$, $2t+2$ and $2t+3$. Recall that the codeword length is n, and let $r = k/n$ denote the code rate of each encoder; also let $C(\gamma)$ denote the capacity of a binary-input point-to-point link with temporal SNR γ:

$$C(\gamma) = \sum_{\xi \in \{0,1\}} \int_{-\infty}^{\infty} \frac{p(y|\xi,\gamma)}{2} \log_2 \frac{2 \cdot p(y|\xi,\gamma)}{p(y|\xi=1,\gamma) + p(y|\xi=0,\gamma)} dy \tag{4.23}$$

where $p(y|\xi,\gamma)$ is the pdf of Gaussian noise

$$p(y|\xi,\gamma) = \sqrt{\frac{\gamma}{\pi}} \exp(-\gamma \cdot (y - x(\xi))^2) \tag{4.24}$$

with BPSK modulation $x(\xi) = 1 - 2 \cdot \xi$.

Then the number of information bits at the destination which may be used for joint decoding of $\mathbf{a}_t$ and $\mathbf{b}_t$ is not greater than $n[C(\gamma_{AD}) + C(\gamma_{BD}) + C(\gamma_{RD1}) + C(\gamma_{RD2}) + C(\gamma_{RD3})]$, where γ_{RD1}, γ_{RD2} and γ_{RD3} (the relay-destination SNRs in consecutive half slots) are independent and identically distributed random variables, and γ_{AD} and γ_{BD} (the source-destination SNRs) are independent and identically distributed random variables if the two source-to-destination link average SNRs are equal. There is no possibility of jointly decoding both packets, and an outage event is inevitable, if

$$n[C(\gamma_{AD}) + C(\gamma_{BD}) + C(\gamma_{RD1}) + C(\gamma_{RD2}) + C(\gamma_{RD3})] < 2k \tag{4.25}$$

Thus, a lower bound for the FER is given by

$$\text{FER} \geq \frac{1}{2} \cdot \Pr[C(\gamma_{AD}) + C(\gamma_{BD}) + C(\gamma_{RD1}) + C(\gamma_{RD2}) + C(\gamma_{RD3}) < 2r] \tag{4.26}$$

The reason for the factor $1/2$ in equation (4.26) is that the outage probability is a good indicator of frame error rate, but that depends on the definition of the "frame". In this case, there are two codewords superimposed in one frame in the relay-destination transmission, so an outage will affect two codewords. The factor $1/2$ is a conservative estimate to assert how many of these two codewords will be erroneous if an outage event described in equation (4.25) happens.

The Monte Carlo simulated lower bound is included in the simulation results of Section 4.2.5.

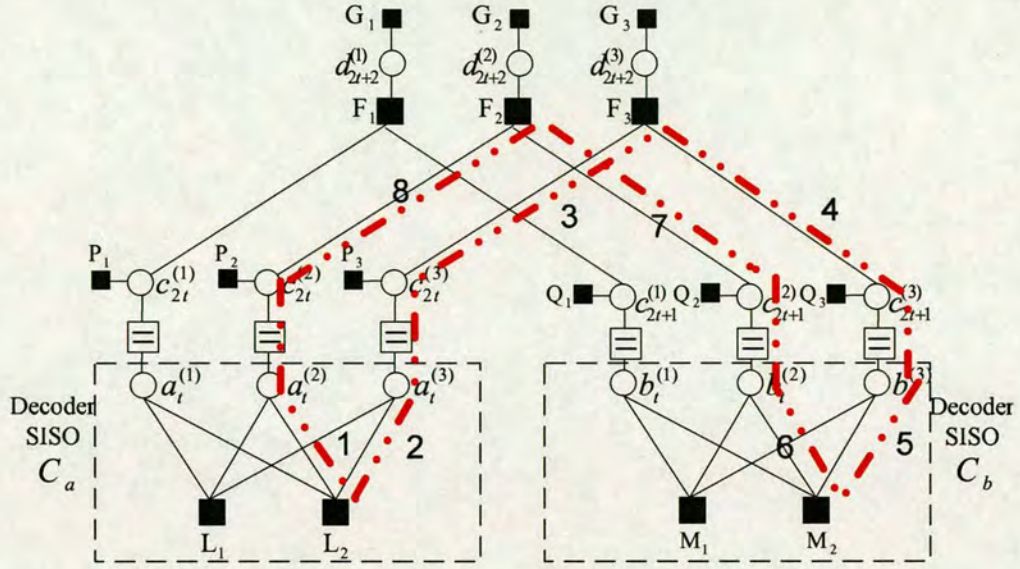

Figure 4.4: *Demonstration for the large multiplicity of 8-cycles in the absence of the interleaver π when $\mathbf{H}_a = \mathbf{H}_b$, such as $\mathbf{a}_t^{(2)} \Rightarrow L_2 \Rightarrow \mathbf{a}_t^{(3)} \Rightarrow F_3 \Rightarrow \mathbf{b}_t^{(3)} \Rightarrow M_2 \Rightarrow \mathbf{b}_t^{(2)} \Rightarrow F_2 \Rightarrow \mathbf{a}_t^{(2)}$ where edges like $\mathbf{C}_{2t}^{(2)} \Rightarrow \mathbf{a}_t^{(2)}$ are not counted.*

4.2.5 Simulation results

In this section, we provide a comparison of the proposed cooperative coding scheme with the two reference cooperative schemes. The first reference scheme is the consecutive relaying scheme of Table 4.2. Here a single LDPC decoding per codeword is sufficient for reception, where the LLRs for decoder initialization are found by adding the LLRs corresponding to the broadcast and relayed versions of the pertinent codeword. The second reference scheme is the simple TDMA/ FDMA code superposition relaying scheme of Table 4.3 and 4.4. For this scheme, joint decoding of $\mathbf{a}_t$ and $\mathbf{b}_t$ is performed using a decoding algorithm similar to that of Section 4.2.3. The decoder also uses message passing between the two constituent decoders via network coding constraints; full details are omitted due to space limitations. The fair comparison of the three cooperative schemes is based on the constraint that in simulations, each scheme uses the same codes $\mathcal{C}_a$ and $\mathcal{C}_b$, and the same total energy E for transmission of the $2L$ source messages.

The codes used for simulations are randomly generated rate $1/2$ regular LDPC codes of block length $n = 1200$ with column weight 3 and no 4-cycles in the Tanner graph. In simulations we choose $\mathbf{H}_a = \mathbf{H}_b$, and assume BPSK modulation for all systems. If no interleaver is used and the two codes $\mathcal{C}_a$ and $\mathcal{C}_b$ are identical, there will be a large multiplicity of 8-cycles in the full

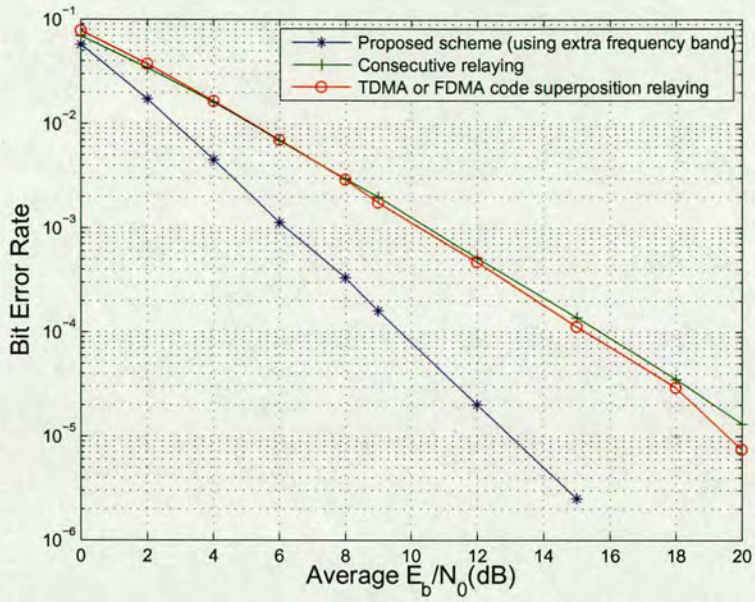

Figure 4.5: *Comparative BER performance for the proposed cooperative scheme. The performance is shown with respect to two reference schemes: simple TDMA/FDMA code superposition relaying; and consecutive relaying using the extra frequency band.*

graph as depicted in Figure 4.4. Putting in a random interleaver π can avoid 8-cycle multiplicity in the Tanner graph. We consider a quasi-static Rayleigh fading channel, for which the fading coefficients are constant within each half slot (one codeword) and change independently from one half slot to the next. We assume equal average signal-to-noise ratio (SNR) on the two source-destination links and the relay-destination link, and we assume that the destination has perfect knowledge of the channel fading coefficients and noise variances. As for the two source-relay links, which play a key role in the performance of the system since poor link quality may lead to catastrophic error propagation at the destination decoder, the simulation setup is that the source-relay links for the comparative performance of the proposed scheme with other reference systems are both ideally error-free. We also provide a bundle of curves indicating the degration of the overall performance as the signal-to-noise ratios of the source-relay links worsen.

Simulated performance results in terms of BER and FER are shown in Figures 4.5 and 4.6 respectively. The curve corresponding to the proposed cooperative coding scheme exhibits an increased diversity gain with respect to the reference schemes in the SNR region of interest. The proposed scheme attains approximately an order of magnitude decrease in both BER and FER over the other schemes at an E_b/N_0 of 8 dB. From the another point of view, to achieve

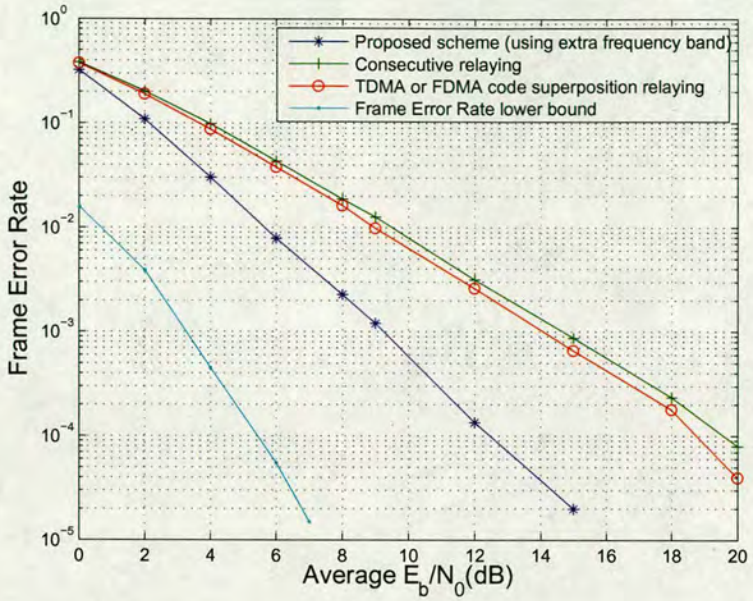

Figure 4.6: *Comparative FER performance for the proposed cooperative scheme. The performance is shown with respect to two reference schemes: simple TDMA/FDMA code superposition relaying; and consecutive relaying using the extra frequency band. Also plotted is the theoretical lower bound on the FER given by (4.26).*

the performance of BER of 10^{-3} and FER of 10^{-2}, the proposed scheme acquires an advantage of 4dB in E_b/N_0 over the second reference scheme of simple TDMA/FDMA based code superposition relaying. We could see from Table 4.1 that in the proposed cooperative coding scheme every source transmitted codeword enjoys a diversity order of three, while in the first reference scheme of consecutive relaying (Table 4.2) and the second reference scheme of simple TDMA/FDMA based code superposition relaying (Table 4.3 and 4.4) every source transmitted codeword only enjoys a diversity order of two. The second reference scheme of simple TDMA/FDMA based code superposition relaying also attains a performance advantage over the first reference scheme of consecutive relaying due to the fact that network coding is a more energy efficient method to produce the same diversity order under the same total power constraint. We also notice that in Figure 4.6 there is an obvious gap between the proposed scheme and the theoretical FER lower bound derived in section 4.2.4 which could be explained as follows: in the capacity calculation, the information embodied in the XOR codewords is assumed obtainable while in practice "splitting" information contained in the XOR codeword is quite a challenging task for the practical coding scheme; in the proposed scheme, the decoding error of the previous codeword may be passed on to the next decoding operation because of

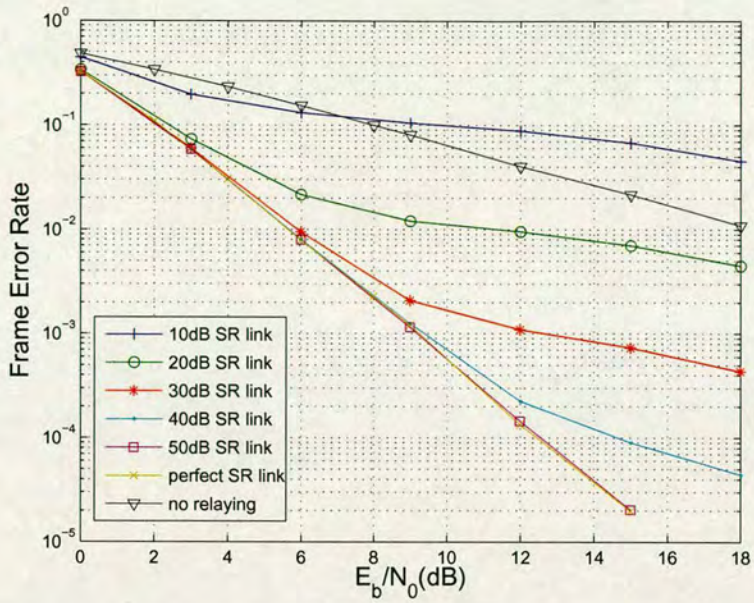

Figure 4.7: *Comparative FER performance for the proposed cooperative scheme in the case of extra frequency band. The performances are shown with various scenarios of the source-relay links: the signal-to-noise ratio ranges from 10dB to 50dB. Performance for the simple TDMA scheme without cooperation or relaying is provided for comparison.*

the "chain" formed by the multiple codewords; the sum-product algorithm for LDPC decoding itself is suboptimal; also only middle blocksize LDPC codes are used here while in the theory the code blocksize is unlimited.

In Figure 4.7, a bundle of curves corresponding to various source-relay link scenarios are shown. It shows that the overall performance of the cooperative scheme improves as the SNR of the source-relay links increases. When the SNRs of the source-relay links are 50dB, at which the outage probabilities are 10^{-5}, the overall performance approaches that of the ideal case. The phenomenon indicates that in the network coded multiple access relay channel the source-relay link plays a key role in the overall performance since the re-encoding operation at the relay may lead to error propagation. The aim of Figure 4.7 is to show the impact of the error propagation effect when the condition of the source-relay links are not perfect. It is also worth noting that adaptive relaying (where the relay discards any unsuccessfully decoded codewords indicated by CRC rather than relaying them) [LTW04] will provide an improved performance without error floor phenomenon. This improvement could be continued as a future work.

4.2.6 Conclusion

In this section, we have proposed a simple but effective cooperative coding scheme for the shared-relay scenario. The scheme uses algebraic code superposition relaying in a frequency division mode to create spatial diversity under the constraint of limited communications resources. The decoding algorithm at the destination node is based on message passing on a factor graph corresponding to multiple received frames at the destination, and extracts spatial diversity gains in a computationally efficient manner. We show that despite the simplicity of the proposed scheme, diversity gains are efficiently leveraged by the simple combination of channel coding at the sources and network coding at the relay. Simulation results demonstrate that the proposed scheme outperforms competitive schemes based on consecutive relaying and TDMA/FDMA based code superposition relaying.

4.3 Proposed Cooperative Coding Scheme in a Shared Relay Environment in the Case of Fading Multiple Access Channel

4.3.1 Background

In the work of [HSOB05] and [HD06], the cooperation model was considered as a specific case for the network-coding-based MARC [KW00] where the source-to-destination channel and the relay-to-destination channel are assumed to be orthogonal to each other. In [LYK+08], the interference between the source-to-destination channel and the relay-to-destination channel are under consideration while the relay simply plays a "forwarding" role without the network coding operation. In this section, we will extend our work of the last section to a cooperative coding scheme which allows continuous transmission of superposed codewords by the relay and at the same time tackling with the interference introduced by the multiple access fading channel, thus making efficient use of communication resources to leverage spatial diversity gains. The corresponding decoder at the destination node is based on a sliding window structure where certain *a posteriori* LLRs are retained as *a priori* LLRs for the next decoding stage which aims to efficiently extract the substantial spatial diversity contained in the code superposition and signal superposition.

Before we start with the system description, there is a piece of related work that needs to be introduced. [YWYHM08] proposed the design of a cooperation coding scheme for a two-user

MAC where two block Markov coding schemes, namely, the multiplexed coding and the superposition coding are under consideration. Although the configuration is different from our work here, the approach in [YWYHM08] is similar in philosophy to our proposed scheme. However, the partially multiplexed (PMP) coding scheme introduced in [YWYHM08] requires a separation of two-user's information bits from code superposition in the **G** matrix or **H** matrix construction which means it cannot yield the full network coding gain provided by the construction in our work.

4.3.2 System model

We still consider the four-node communications network depicted in Figure 4.1, with two sources S_a and S_b, one relay R, and one destination D common to the two sources. Also, flat-fading channels are assumed for all point-to-point links.

Time Slot	0		1		...	t		t+1		...	L-1		L	
Half Slot	0	1	2	3	...	2t	2t+1	2t+2	2t+3	...	2L-2	2L-1	2L	2L+1
Source Transmits	$c_0=$ a_0	$c_1=$ b_0	$c_2=$ a_1	$c_3=$ b_1	...	$c_{2t}=$ a_t	$c_{2t+1}=$ b_t	$c_{2t+2}=$ a_{t+1}	$c_{2t+3}=$ b_{t+1}	...	$c_{2L-2}=$ a_{L-1}	$c_{2L-1}=$ b_{L-1}		
Relay Transmits		$d_1=$ $\pi(a_0)$	$d_2=$ $\pi(b_0)$ $\oplus a_0$	$d_3=$ $\pi(a_1)$ $\oplus b_0$	...	$d_{2t}=$ $\pi(b_{t-1})$ $\oplus a_{t-1}$	$d_{2t+1}=$ $\pi(a_t)$ $\oplus b_{t-1}$	$d_{2t+2}=$ $\pi(b_t)$ $\oplus a_t$	$d_{2t+3}=$ $\pi(a_{t+1})$ $\oplus b_t$	...	$d_{2L-2}=$ $\pi(b_{L-2})$ $\oplus a_{L-2}$	$d_{2L-1}=$ $\pi(a_{L-1})$ $\oplus b_{L-2}$	$d_{2L}=$ $\pi(b_{L-1})$ $\oplus a_{L-1}$	$d_{2L+1}=$ b_{L-1}
Destination Receives	e_0	e_1	e_2	e_3	...	e_{2t}	e_{2t+1}	e_{2t+2}	e_{2t+3}	...	e_{2L-2}	e_{2L-1}	e_{2L}	e_{2L+1}
Destination Decodes			a_0	b_0	...	a_{t-1}	b_{t-1}	a_t	b_t	...	a_{L-2}	b_{L-2}	a_{L-1}	b_{L-1}

Table 4.7: *Transmission schedule of proposed cooperative coding scheme for the case of multiple access channel.*

4.3.2.1 Overview

The transmission schedule for the new scheme in this section is illustrated in Table 4.7. The major difference between the system model in Table 4.7 and the previous scenario in Table 4.1 is that in the new model the source-destination links and the relay-destination link are in the same frequency band and therefore a two-user fading MAC is formed from the perspective of the destination. The details will be addressed in the following relaying protocol.

4.3.2.2 Relaying protocol

The relay decodes and then re-encodes each codeword received from the source. The relay has a buffer in which it stores the codewords it obtained in the previous two half slots. At

each half slot ($t = 2, 3, \cdots, 2L$), the relay interleaves the codeword obtained in half slot $t-1$ and superposes it (XOR operation) with the codeword obtained in half slot $t-2$; it then transmits the resulting codeword to the destination in the fading multiple access channel whose channel resources are also shared with the source transmission. Special cases arise at half slots 1 and $2L+1$ in which only a single codeword is stored at the relay and no XOR operation is performed. Let $\mathbf{d}_t$ denote the codeword transmitted from the relay to the destination in half slot $t \in \{1, 2, \cdots, 2L+1\}$; thus $\mathbf{d}_t = \pi(\mathbf{c}_{t-1}) \oplus \mathbf{c}_{t-2}$, for $t = 2, 3, \cdots 2L$. Let $\mathbf{e}_t$ denote the signal stream received by the destination in half slot $t \in \{0, 1, \cdots, 2L+1\}$. The use of the MAC requires no extra resources for the proposed cooperative transmission scheme as compared to the non-cooperative scheme. As we shall see, the spatial diversity gain generated by the superposition relaying outweighs the signal interference degradation inherent in using the MAC. To simplify the analysis, we assume that directional antennas are employed at the relay, so that interference between the source-relay link and the relay-destination link may be neglected. For each $t = 0, 1, \cdots L-1$, source S_a's codeword $\mathbf{a}_t$ is decoded at the end of half slot $2t+2$ and source S_b's codeword $\mathbf{b}_t$ is decoded at the end of half slot $2t+3$. The transmission schedule for this cooperative coding scheme is illustrated in Table 4.7. It can be seen that spatial diversity for each message is contained in three separate transmissions spanning three half slots.

To avoid the full-duplex requirement at the relay in a practical communication system, a second relay R' which employs a simple amplify-and-forward (AF) scheme could be used between R and D in the current cell frequency $f1$. The transmission from R to R' could employ the neighboring cell frequency $f2$, and thus the uplink MAC still use $f1$ for this cell. In the counterpart neighboring cell with frequency $f2$, $f1$ is used in the R to R' link as depicted in Figure 4.8. In the overall communications system, there will be no extra frequency band occupied and only negligible inter-cell interference. The decoding procedure at the destination node is then identical to that presented in this thesis.

Time Slot	0		1		...	t		t+1		...	L-1		
Half Slot	0	1	2	3	...	2t	2t+1	2t+2	2t+3	...	2L-2	2L-1	2L
Source Transmits	$\mathbf{c}_0=\mathbf{a}_0$	$\mathbf{c}_1=\mathbf{b}_0$	$\mathbf{c}_2=\mathbf{a}_1$	$\mathbf{c}_3=\mathbf{b}_1$	...	$\mathbf{c}_{2t}=\mathbf{a}_t$	$\mathbf{c}_{2t+1}=\mathbf{b}_t$	$\mathbf{c}_{2t+2}=\mathbf{a}_{t+1}$	$\mathbf{c}_{2t+3}=\mathbf{b}_{t+1}$	...	$\mathbf{c}_{2L-2}=\mathbf{a}_{L-1}$	$\mathbf{c}_{2L-1}=\mathbf{b}_{L-1}$	
Relay Transmits		$\mathbf{d}_1=\pi(\mathbf{a}_0)$	$\mathbf{d}_2=\pi(\mathbf{b}_0)$	$\mathbf{d}_3=\pi(\mathbf{a}_1)$	...	$\mathbf{d}_{2t}=\pi(\mathbf{b}_{t-1})$	$\mathbf{d}_{2t+1}=\pi(\mathbf{a}_t)$	$\mathbf{d}_{2t+2}=\pi(\mathbf{b}_t)$	$\mathbf{d}_{2t+3}=\pi(\mathbf{a}_{t+1})$	...	$\mathbf{d}_{2L-2}=\pi(\mathbf{b}_{L-2})$	$\mathbf{d}_{2L-1}=\pi(\mathbf{a}_{L-1})$	$\mathbf{d}_{2L}=\pi(\mathbf{b}_{L-1})$
Destination Receives	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	...	$\mathbf{e}_{2t}$	$\mathbf{e}_{2t+1}$	$\mathbf{e}_{2t+2}$	$\mathbf{e}_{2t+3}$	...	$\mathbf{e}_{2L-2}$	$\mathbf{e}_{2L-1}$	$\mathbf{e}_{2L}$
Destination Decodes		$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_1$	...	$\mathbf{b}_{t-1}$	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_{t+1}$	...	$\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$

Table 4.8: *Transmission schedule of consecutive relaying scheme for the case of multiple access channel.*

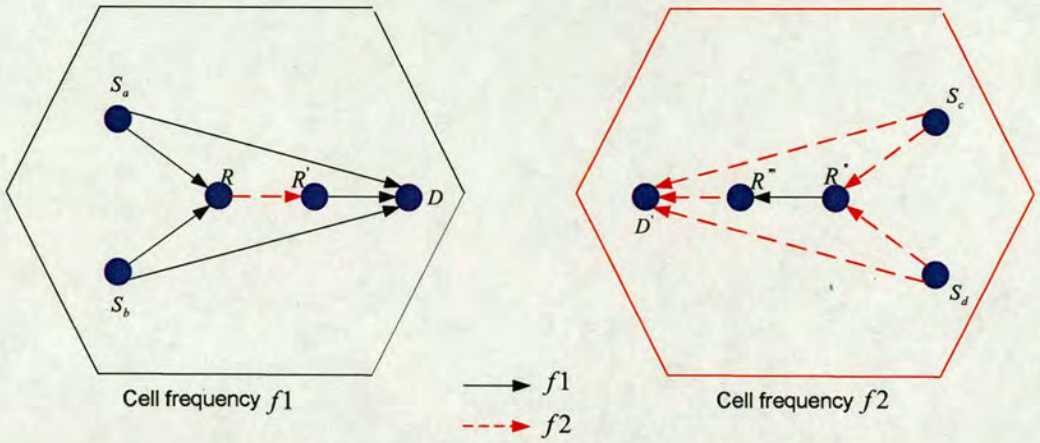

Figure 4.8: *One example model to avoid the full-duplex requirement at the relay.*

4.3.2.3 Reference schemes

The first reference scheme is the consecutive relaying in the fading MAC depicted in Table 4.8; here the relay simply re-transmits the (interleaved) previously received codeword rather than a codeword superposition. It is easily seen that in this reference scheme, spatial diversity for each message is contained in two separate transmissions spanning two half slots. The second reference scheme is a simple two-user TDMA transmission scheme without any relaying or cooperation among the users. A simulation-based comparison of the three schemes described in this section under a transmit power constraint will be given in Section 4.3.6.

4.3.3 Soft demodulator for BPSK modulation in the multiple access channel

In the case where the MAC is used, a SISO signal processing module is required to update the extrinsic information regarding the coded bits involved in the signal superposition, using corresponding input extrinsic information together with the received channel values.

Recall that $\mathbf{c}_t$ denotes the codeword broadcast by the source in half slot t and $\mathbf{d}_t$ denotes the codeword transmitted by the relay in half slot t. In the following we define $\mathcal{I} = \{1, 2, \cdots, n\}$, and $x^{(i)}$ denotes the i-th bit of codeword $\mathbf{x}$ while $e^{(i)}$ denotes the i-th received channel value of signal stream $\mathbf{e}$. Considering a single half slot $t \in \{1, \cdots, 2L-1\}$, the channel model is given by

$$e_t^{(i)} = \phi_t^{(i)} \cdot \alpha(c_t^{(i)}) + \psi_t^{(i)} \cdot \beta(d_t^{(i)}) + n_t^{(i)}, \quad i \in \mathcal{I}\,, \tag{4.27}$$

where $\phi_t^{(i)}$ and $\psi_t^{(i)}$ represent the fading processes on the source-destination and relay-destination

links, respectively, and $n_t^{(i)}$ is complex AWGN with variance σ^2 per dimension. Also

$$\alpha(c_t^{(i)}) = \sqrt{P_S}(1 - 2c_t^{(i)}) \tag{4.28}$$

$$\beta(d_t^{(i)}) = \sqrt{P_R}(1 - 2d_t^{(i)}) \tag{4.29}$$

holds for BPSK modulation, with P_S and P_R representing the received power for the symbols transmitted by the source and relay respectively.

The function of the soft demodulator is to take as input extrinsic log-likelihood ratios (LLRs) on the transmitted bits (LLRs in the absence of channel information); without loss of generality we consider the bit $c_t^{(i)}$,

$$L^E(c_t^{(i)}) = \ln\left(\frac{\Pr\left(c_t^{(i)} = 0\right)}{\Pr\left(c_t^{(i)} = 1\right)}\right) \tag{4.30}$$

and compute the new extrinsic LLRs (incremental LLRs expressing new information derived from the channel)

$$L^O(c_t^{(i)}) = L(c_t^{(i)}) - L^E(c_t^{(i)}) \tag{4.31}$$

where the *a posteriori* LLRs $L(c_t^{(i)})$ (incorporating channel information) are given by

$$L(c_t^{(i)}) = \ln\left(\frac{\Pr\left(c_t^{(i)} = 0|e_t^{(i)}\right)}{\Pr\left(c_t^{(i)} = 1|e_t^{(i)}\right)}\right) \tag{4.32}$$

$$= \ln\left(\frac{\sum_{\{c_t^{(i)}, d_t^{(i)}\}: c_t^{(i)}=0} \Pr\left(c_t^{(i)}, d_t^{(i)}|e_t^{(i)}\right)}{\sum_{\{c_t^{(i)}, d_t^{(i)}\}: c_t^{(i)}=1} \Pr\left(c_t^{(i)}, d_t^{(i)}|e_t^{(i)}\right)}\right) \tag{4.33}$$

$$= \ln\left(\frac{\sum_{\{c_t^{(i)}, d_t^{(i)}\}: c_t^{(i)}=0} p\left(e_t^{(i)}|c_t^{(i)}, d_t^{(i)}\right) \cdot \Pr(c_t^{(i)}, d_t^{(i)})}{\sum_{\{c_t^{(i)}, d_t^{(i)}\}: c_t^{(i)}=1} p\left(e_t^{(i)}|c_t^{(i)}, d_t^{(i)}\right) \cdot \Pr(c_t^{(i)}, d_t^{(i)})}\right) \tag{4.34}$$

We assume that the users' code bits are independent (a realistic assumption for independent users), so that

$$\Pr(c_t^{(i)}, d_t^{(i)}) = \Pr(c_t^{(i)}) \cdot \Pr(d_t^{(i)}) \tag{4.35}$$

Also, the probability density function (PDF) of the received channel value conditioned on the

transmitted bits may be written as

$$p(e_t^{(i)}|c_t^{(i)}, d_t^{(i)}) = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)} - \phi_t^{(i)} \cdot \alpha(c_t^{(i)}) - \psi_t^{(i)} \cdot \beta(d_t^{(i)})\right|^2\right) \tag{4.36}$$

Therefore, (4.31) may be re-written as (4.37), for which we use the shorthand $L^O(c_t^{(i)}) = f_1(e_t^{(i)}, L^E(d_t^{(i)}))$. In the same way, the new extrinsic LLR of the code bit $d_t^{(i)}$ may be obtained from (4.38), for which we use the shorthand $L^O(d_t^{(i)}) = f_2(e_t^{(i)}, L^E(c_t^{(i)}))$.

$$L^O(c_t^{(i)}) = \ln\left(\frac{\exp\left(\frac{\left|e_t^{(i)}-\sqrt{P_S}\phi_t^{(i)}-\sqrt{P_R}\psi_t^{(i)}\right|^2}{-2\sigma^2} + L^E(d_t^{(i)})\right) + \exp\left(\frac{\left|e_t^{(i)}-\sqrt{P_S}\phi_t^{(i)}+\sqrt{P_R}\psi_t^{(i)}\right|^2}{-2\sigma^2}\right)}{\exp\left(\frac{\left|e_t^{(i)}+\sqrt{P_S}\phi_t^{(i)}-\sqrt{P_R}\psi_t^{(i)}\right|^2}{-2\sigma^2} + L^E(d_t^{(i)})\right) + \exp\left(\frac{\left|e_t^{(i)}+\sqrt{P_S}\phi_t^{(i)}+\sqrt{P_R}\psi_t^{(i)}\right|^2}{-2\sigma^2}\right)}\right) \tag{4.37}$$

$$L^O(d_t^{(i)}) = \ln\left(\frac{\exp\left(\frac{\left|e_t^{(i)}-\sqrt{P_R}\psi_t^{(i)}-\sqrt{P_S}\phi_t^{(i)}\right|^2}{-2\sigma^2} + L^E(c_t^{(i)})\right) + \exp\left(\frac{\left|e_t^{(i)}-\sqrt{P_R}\psi_t^{(i)}+\sqrt{P_S}\phi_t^{(i)}\right|^2}{-2\sigma^2}\right)}{\exp\left(\frac{\left|e_t^{(i)}+\sqrt{P_R}\psi_t^{(i)}-\sqrt{P_S}\phi_t^{(i)}\right|^2}{-2\sigma^2} + L^E(c_t^{(i)})\right) + \exp\left(\frac{\left|e_t^{(i)}+\sqrt{P_R}\psi_t^{(i)}+\sqrt{P_S}\phi_t^{(i)}\right|^2}{-2\sigma^2}\right)}\right) \tag{4.38}$$

4.3.4 Decoding algorithm at destination node

The principle of the decoding operation is to extract the diversity gain contained in the joint channel and network codes plus the signal superposition introduced by the nature of the fading MAC. The information for each codeword is contained in three consecutive streams received at the destination. Thus the factor graph for the decoding at the destination node is based on a sliding window structure for each of the received streams $\mathbf{e}_t$, $t \in \{1, \cdots, 2L+1\}$.

4.3.4.1 Decoding overview

Half Slot		$2t$	$2t+1$	$2t+2$	
Source Transmits	...	$\mathbf{c}_{2t} = \mathbf{a}_t$	$\mathbf{c}_{2t+1} = \mathbf{b}_t$	$\mathbf{c}_{2t+2} = \mathbf{a}_{t+1}$	...
Relay Transmits	...	$\mathbf{d}_{2t} = \pi(\mathbf{b}_{t-1}) \oplus \mathbf{a}_{t-1}$	$\mathbf{d}_{2t+1} = \pi(\mathbf{a}_t) \oplus \mathbf{b}_{t-1}$	$\mathbf{d}_{2t+2} = \pi(\mathbf{b}_t) \oplus \mathbf{a}_t$	...
Destination Receives	...	$\mathbf{e}_{2t}$	$\mathbf{e}_{2t+1}$	$\mathbf{e}_{2t+2}$	...
Decoding Evolution Process For $\mathbf{a}_t$		$L_1(c_{2t}^{(i)})$	$L_2(c_{2t}^{(i)})$	$\hat{a}_t^{(i)}$	
		Step 1	Step 2	Step 3	

Table 4.9: *The demonstration for the three-step decoding evolution in the case of fading MAC. Here we still take the decoding of* $\mathbf{a}_t$ *as an example.* $\mathbf{a}_t$ *is contained in three consecutive received streams at the destination:* $\mathbf{e}_{2t}$, $\mathbf{e}_{2t+1}$ *and* $\mathbf{e}_{2t+2}$. *We use* $L_1(c_{2t}^{(i)})$ *and* $L_2(c_{2t}^{(i)})$ *to demonstrate the L-value obtained at the corresponding half time slot.*

Without loss of generality, we still consider the decoding of codeword $\mathbf{a}_t$ at the end of half slot $2t+2$, for $t \in \{0, 1, \cdots, L-2\}$. As demonstrated in Table 4.9, $\mathbf{a}_t$ is contained in three received streams: $\mathbf{e}_{2t}$, $\mathbf{e}_{2t+1}$ and $\mathbf{e}_{2t+2}$. We still use $L_1(c_{2t}^{(i)}) \rightarrow L_2(c_{2t}^{(i)}) \rightarrow \hat{a}_t^{(i)}$ to demonstrate the evolution of the decoding process for $\mathbf{a}_t$, in which $L_1(c_{2t}^{(i)})$ and $L_2(c_{2t}^{(i)})$ represent the L-value obtained in the corresponding step. Next we shall use a factor graph to illustrate this evolution process which corresponds to the different positions in the graph.

The factor graphs corresponding to the half slot $2t$, $2t+1$ and $2t+2$ are illustrated in Figure 4.9, Figure 4.10 and Figure 4.11 respectively where circles depict variable nodes and squares depict factor nodes. In each factor graph, the extrinsic information is exchanged between three soft-input soft-output (SISO) decoder modules for the constituent codes, via the factor nodes $\{F_i\}$ which correspond to the network coding operation at the relay and the factor nodes $\{G_i\}$ which correspond to the soft demodulator for the MAC as given by equation (4.37) and equation (4.38). For simplicity, the graphs are all illustrated for the case $n = 3$, and where $\mathcal{C}_a$ and $\mathcal{C}_b$ are (trivial) LDPC codes.

The whole decoding process could be seen as a sliding window structure where certain *a posteriori* LLRs are obtained and are used as *a priori* LLRs in the next decoding structure. The

SISO decoder for a specific codeword, *e.g.* $\mathbf{a}_t$, goes through three position in the whole factor graph as illustrated in the Figure 4.9 for half slot $2t$, Figure 4.10 for half slot $2t+1$ and Figure 4.11 for half slot $2t+2$. Accordingly, the LLRs for $\mathbf{a}_t$ are evolved by the exchange of the extrinsic L-values under the constraints of parity checks in channel coding, the network coding operation and the MAC demodulator until the final stage of the third step where decision of $\hat{a}_t^{(i)}$ is made.

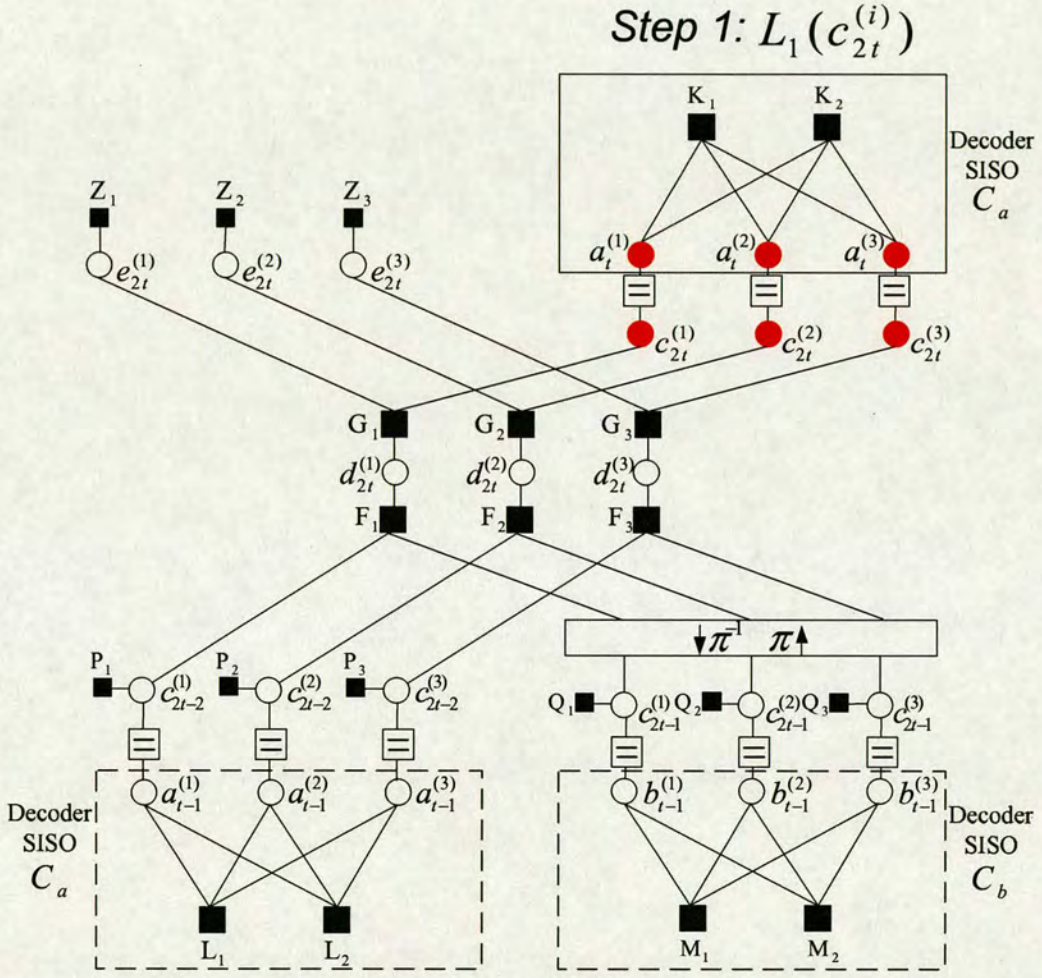

Figure 4.9: *Factor graph corresponding to the half slot 2t which is the first step of the decoding evolution for codeword* $\mathbf{a}_t$. *At this stage, the code structure corresponding to* $\mathbf{a}_t$ *takes up the position of the upper right decoder SISO as indicated in the factor graph. Accordingly, the* a posteriori *LLR* $L_1(c_{2t}^{(i)})$ *is obtained which will be used as* a priori *LLR in the next step.*

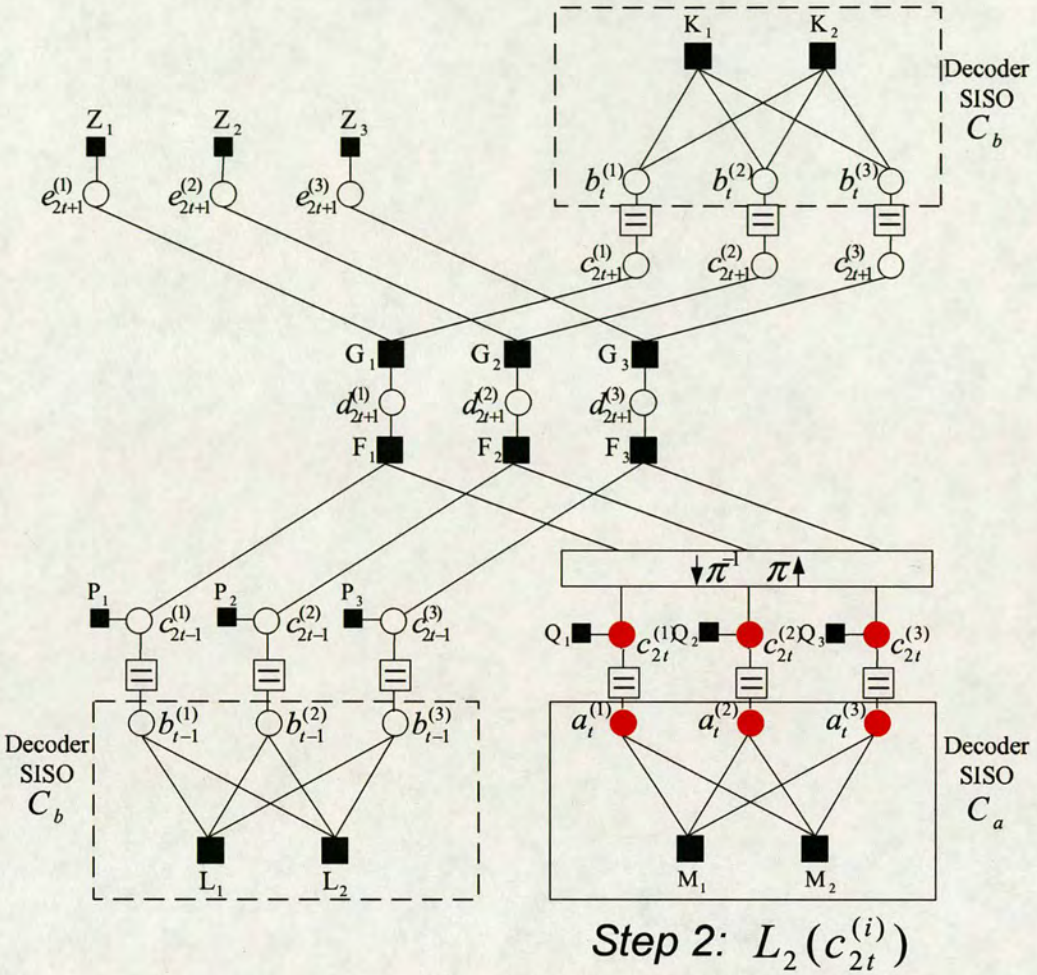

Figure 4.10: *Factor graph corresponding to the half slot $2t+1$ which is the second step of the decoding evolution for codeword $\mathbf{a}_t$. At this stage, the code structure corresponding to $\mathbf{a}_t$ takes up the position of the lower right decoder SISO as indicated in the factor graph. $L_1(c^{(i)}_{2t})$ which was obtained in the first step is used as* a priori *LLR in the current step and accordingly* a posteriori *LLR $L_2(c^{(i)}_{2t})$ is produced which will be used as* a priori *LLR in the next step. Also note that in the current factor graph structure, $L_1(c^{(i)}_{2t+1})$ is obtained which will also be used in the next step.*

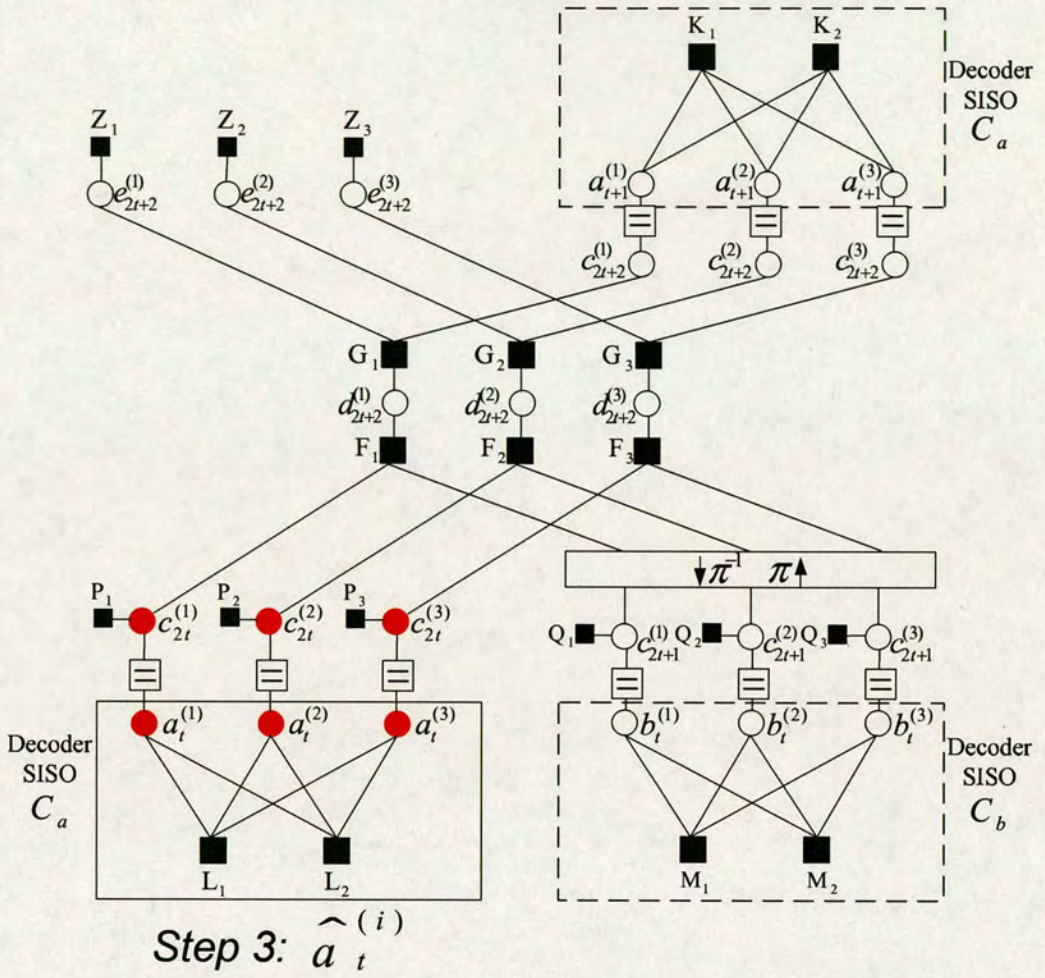

Figure 4.11: *Factor graph corresponding to the half slot $2t+2$ which is the final step of the decoding evolution for codeword $\mathbf{a}_t$. At this stage, the code structure corresponding to $\mathbf{a}_t$ takes up the position of the lower left decoder SISO as indicated in the factor graph. $L_2(c_{2t}^{(i)})$ which was obtained in the second step is used as* a priori *LLR in the current step and the final decision $\hat{a}_t^{(i)}$ is made. Also note that in the current factor graph structure, $L_1(c_{2t+2}^{(i)})$ and $L_2(c_{2t+1}^{(i)})$ are obtained which will be used in the succeeding decoding window.*

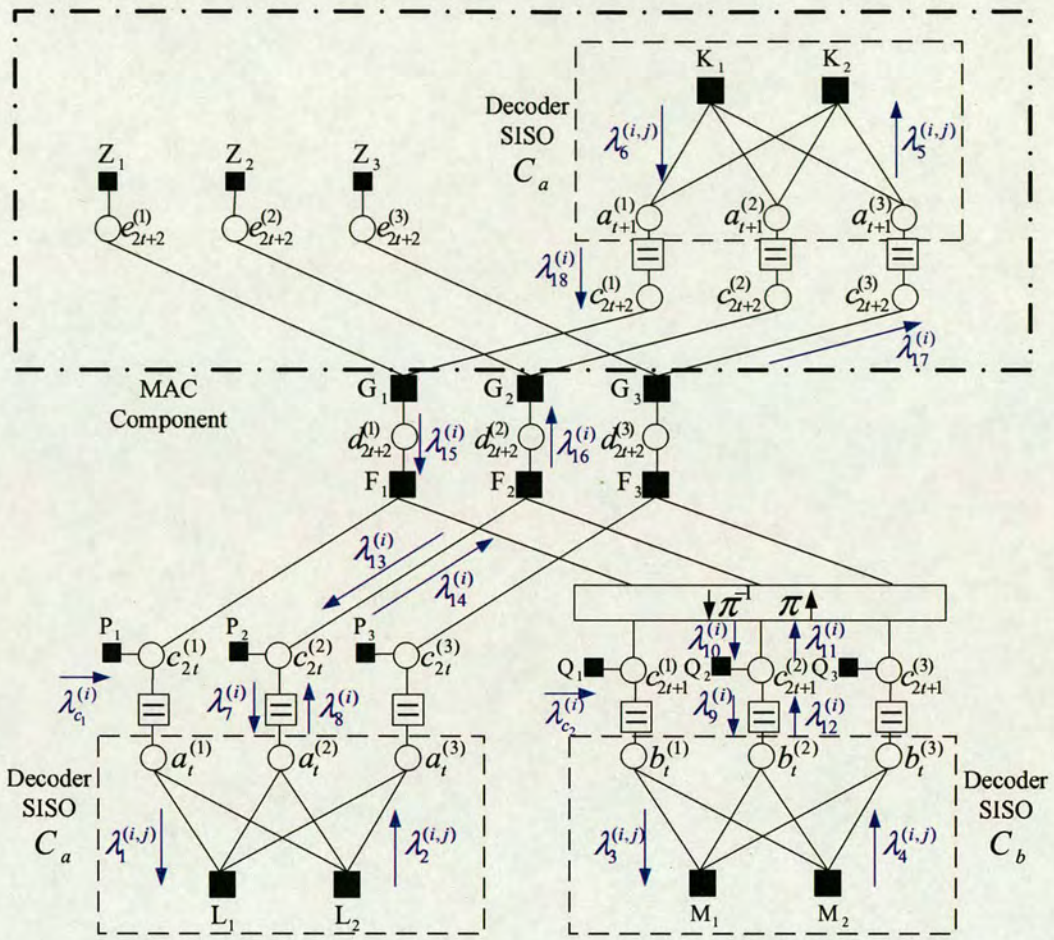

Figure 4.12: *Factor graph corresponding to the destination's decoding of codeword* $\mathbf{a}_t$. *For ease of presentation, the factor graph is illustrated for the case* $n = 3$ *and trivial codes* $\mathcal{C}_a$, $\mathcal{C}_b$. *For the case of the fading multiple access channel, the three decoder SISO modules exchange extrinsic information via the factor nodes* $\{F_i\}$ *which correspond to the network coding operation at the relay and the factor nodes* $\{G_i\}$ *which correspond to the soft demodulator for the MAC.*

4.3.4.2 Detailed algorithm

Next, we provide a concise description of the factor graph based decoding algorithm [KFL01] at the destination decoder for the half slot $2t+2$ where the final decision $\hat{a}_t^{(i)}$ is made. As described before, *a priori* LLRs on $\mathbf{c}_{2t+1}$ and $\mathbf{c}_{2t}$, denoted $\{L_1(c_{2t+1}^{(i)})\}$ and $\{L_2(c_{2t}^{(i)})\}$ respectively, are available from the previous decoding. In addition to decoding of $\mathbf{a}_t$, the decoder will produce *a posteriori* LLRs $\{L_1(c_{2t+2}^{(i)})\}$ and (updated) *a posteriori* LLRs $\{L_2(c_{2t+1}^{(i)})\}$; these will be used as *a priori* LLRs in the next decoding step.

The two codewords contained in the received signal stream $\mathbf{e}_{2t+2}$ are:

$$\begin{aligned} \mathbf{c}_{2t+2} &= \mathbf{a}_{t+1} \\ \mathbf{d}_{2t+2} &= \pi(\mathbf{c}_{2t+1}) \oplus \mathbf{c}_{2t} \\ &= \pi(\mathbf{b}_t) \oplus \mathbf{a}_t \end{aligned}$$

In the following, we shall use the same notational conventions pertaining to the following algorithm description as we did in Section 4.2.3.

Factor Graph Based Decoding Algorithm at Destination Node – Decoding of Codeword $\mathbf{a}_t$

Initialization:

- For $i \in \mathcal{I}$,

$$\lambda_{c_1}^{(i)} = L_2(c_{2t}^{(i)}) \tag{4.39}$$

$$\lambda_{c_2}^{(i)} = L_1(c_{2t+1}^{(i)}) \tag{4.40}$$

$$\lambda_8^{(i)} = 0 \tag{4.41}$$

$$\lambda_{15}^{(i)} = 0 \tag{4.42}$$

- For $i \in \mathcal{I}, j \in \mathcal{N}_a(i)$

$$\lambda_2^{(i,j)} = 0 \tag{4.43}$$

- For $i \in \mathcal{I}, j \in \mathcal{N}_b(i)$

$$\lambda_4^{(i,j)} = 0 \tag{4.44}$$

- For $i \in \mathcal{I}, j \in \mathcal{N}_a(i)$

$$\lambda_6^{(i,j)} = 0 \tag{4.45}$$

Main Loop: For $k = 1$ to N do

- For $i \in \mathcal{I}$,

$$\lambda_{14}^{(i)} = \lambda_{c_1}^{(i)} + \lambda_8^{(i)} \tag{4.46}$$

- Network coding constraints: for $i \in \mathcal{I}$,

$$\lambda_{10}^{(\pi(i))} = \lambda_{14}^{(i)} \boxplus \lambda_{15}^{(i)} \tag{4.47}$$

- For $i \in \mathcal{I}$,

$$\lambda_9^{(i)} = \lambda_{c_2}^{(i)} + \lambda_{10}^{(i)} \tag{4.48}$$

- SISO decoder $\mathcal{C}_b$: for $i \in \mathcal{I}, j \in \mathcal{N}_b(i)$

$$\lambda_3^{(i,j)} = \lambda_9^{(i)} + \sum_{l \in \mathcal{N}_b(i) \setminus \{j\}} \lambda_4^{(i,l)} \tag{4.49}$$

$$\lambda_4^{(i,j)} = \boxplus_{l \in \mathcal{M}_b(j) \setminus \{i\}} \lambda_3^{(l,j)} \tag{4.50}$$

For $i \in \mathcal{I}$,

$$\lambda_{12}^{(i)} = \sum_{j \in \mathcal{N}_b(i)} \lambda_4^{(i,j)} \tag{4.51}$$

- Obtain the *a posteriori* LLR (prior for decoding in next half slot)

$$L_2(c_{2t+1}^{(i)}) = \lambda_9^{(i)} + \lambda_{12}^{(i)} \tag{4.52}$$

- For $i \in \mathcal{I}$,

$$\lambda_{11}^{(i)} = \lambda_{c_2}^{(i)} + \lambda_{12}^{(i)} \tag{4.53}$$

$$\lambda_{16}^{(i)} = \lambda_{14}^{(i)} \boxplus \lambda_{11}^{(\pi(i))} \tag{4.54}$$

- Soft demodulation: for $i \in \mathcal{I}$,

$$\lambda_{17}^{(i)} = f_1(e_{2t+2}^{(i)}, \lambda_{16}^{(i)}) \tag{4.55}$$

- SISO decoder $\mathcal{C}_a$: for $i \in \mathcal{I}, j \in \mathcal{N}_a(i)$

$$\lambda_5^{(i,j)} = \lambda_{17}^{(i)} + \sum_{l \in \mathcal{N}_a(i) \setminus \{j\}} \lambda_6^{(i,l)} \tag{4.56}$$

$$\lambda_6^{(i,j)} = \boxplus_{l \in \mathcal{M}_a(j) \setminus \{i\}} \lambda_5^{(l,j)} \tag{4.57}$$

For $i \in \mathcal{I}$,

$$\lambda_{18}^{(i)} = \sum_{j \in \mathcal{N}_a(i)} \lambda_6^{(i,j)} \tag{4.58}$$

- Obtain the *a posteriori* LLR (prior for decoding in next half slot)

$$L_1(c_{2t+2}^{(i)}) = \lambda_{17}^{(i)} + \lambda_{18}^{(i)} \tag{4.59}$$

- Soft demodulation: for $i \in \mathcal{I}$,

$$\lambda_{15}^{(i)} = f_2(e_{2t+2}^{(i)}, \lambda_{18}^{(i)}) \tag{4.60}$$

$$\lambda_{13}^{(i)} = \lambda_{15}^{(i)} \boxplus \lambda_{11}^{(\pi(i))} \tag{4.61}$$

- For $i \in \mathcal{I}$,

$$\lambda_7^{(i)} = \lambda_{c_1}^{(i)} + \lambda_{13}^{(i)} \tag{4.62}$$

- SISO decoder $\mathcal{C}_a$: for $i \in \mathcal{I}, j \in \mathcal{N}_a(i)$

$$\lambda_1^{(i,j)} = \lambda_7^{(i)} + \sum_{l \in \mathcal{N}_a(i) \setminus \{j\}} \lambda_2^{(i,l)} \tag{4.63}$$

$$\lambda_2^{(i,j)} = \boxplus_{l \in \mathcal{M}_a(j) \setminus \{i\}} \lambda_1^{(l,j)} \tag{4.64}$$

For $i \in \mathcal{I}$,

$$\lambda_8^{(i)} = \sum_{j \in \mathcal{N}_a(i)} \lambda_2^{(i,j)} \tag{4.65}$$

- Calculate *a posteriori* LLRs for codeword $\mathbf{a}_t$:

$$L(a_t^{(i)}) = \lambda_7^{(i)} + \lambda_8^{(i)} \tag{4.66}$$

- Make decisions on the code bits

$$\hat{a}_t^{(i)} = \begin{cases} 0 & \text{if } L(a_t^{(i)}) \geq 0 \\ 1 & \text{if } L(a_t^{(i)}) < 0 \end{cases}$$

If $\hat{\mathbf{a}}_t \mathbf{H}_a^T = \mathbf{0}$ **then** ***break;***

Endfor

The preceding presentation was for the general case of decoding (half slots $t = 2, 3, \cdots, 2L-1\}$). Special cases are handled in a straightforward manner as follows. In half slot $t = 0$, a single LDPC decoding of $\mathbf{a}_0$ based on LLRs derived from the received stream $\mathbf{e}_0$ produces $\{L_1(c_0^{(i)})\}$. In half slot $t = 1$, decoding proceeds as in the general case except that $\lambda_{14}^{(i)}$ is set to $+\infty$ for all $i \in \mathcal{I}$; $\{L_1(c_1^{(i)})\}$ and $\{L_2(c_0^{(i)})\}$ are produced (note that this decoder has the same structure as the decoder for the reference scheme of consecutive relaying in the fading MAC as shown in Table 4.8). In half slot $t = 2L$, the decoder structure used is that for superposition decoding as described in Section 4.2.3; also only $\{L_2(c_{2L-1}^{(i)})\}$ are produced. In time slot $t = 2L + 1$, a single LDPC decoding of $\mathbf{b}_{L-1}$ is performed based on the sum of the LLRs derived from the received stream $\mathbf{e}_{2L+1}$ and the LLRs $\{L_2(c_{2L-1}^{(i)})\}$ obtained from the previous decoding.

4.3.5 A lower bound on the system frame error rate

In the case where relaying is in the MAC, the information of each relay-transmitted signal is already contained in the source-transmitted signal from previous half slots. Therefore, assuming correct decoding in previous half slots, the relay transmitted signal could be perfectly cancelled from the destination's received signal, and hence the receiver will "see" an interference-free fading channel for the source-transmitted signal. Therefore, the theoretical lower bound on the FER for the case of additional frequency band in the section 4.2.4 also applies to the case of MAC relaying. The Monte Carlo simulated lower bound is included in the simulation results in Section 4.3.6.

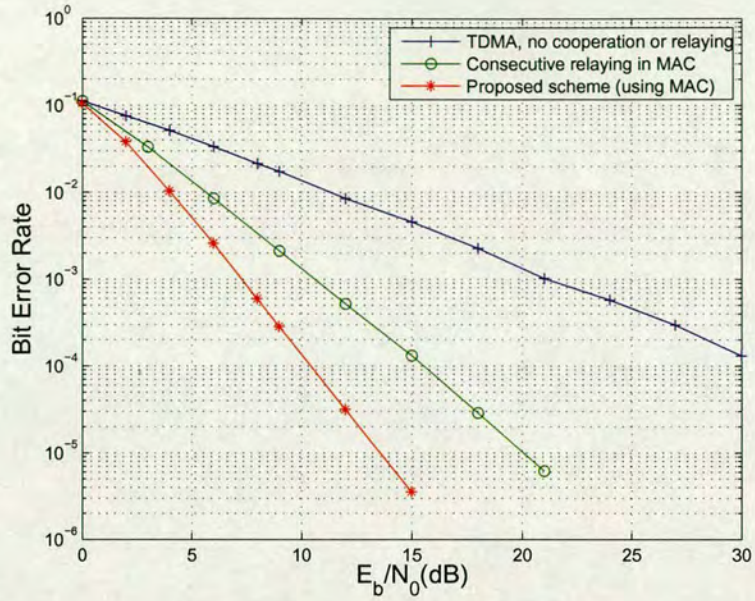

Figure 4.13: *Comparative BER performance for the proposed cooperative scheme in the case of multiple access channel. The performance is shown with respect to two reference schemes: simple TDMA without cooperation or relaying; and consecutive relaying in the MAC.*

4.3.6 Simulation results

The codes used for simulations are randomly generated rate $1/2$ regular LDPC codes of block length $n = 1200$ with column weight 3 and no 4-cycles in the Tanner graph. In simulations we choose $\mathbf{H}_a = \mathbf{H}_b$, and assume BPSK modulation for all systems. A random interleaver π is used to avoid 8-cycle multiplicity in the Tanner graph as described before. We still consider a quasi-static Rayleigh fading channel, for which the fading coefficients are constant within each half slot (one codeword) and change independently from one half slot to the next. We assume equal average signal-to-noise ratio (SNR) on the two source-destination links and the relay-destination link, and we assume that the destination has perfect knowledge of the channel fading coefficients and noise variances. As for the two source-relay links, which play a key role in the performance of the system since poor link quality may lead to catastrophic error propagation at the destination decoder, the simulation setup is that the source-relay links for the comparative performance of the proposed scheme with other reference systems are both ideally error-free. We also provide a bundle of curves indicating the degration of the overall performance as the signal-to-noise ratios of the source-relay links worsen.

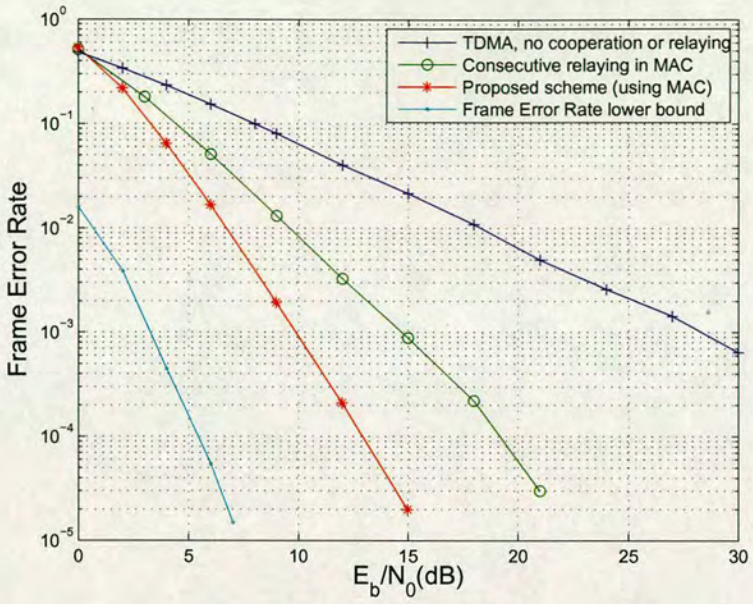

Figure 4.14: *Comparative FER performance for the proposed cooperative scheme in the case of multiple access channel. The performance is shown with respect to two reference schemes: simple TDMA without cooperation or relaying; and consecutive relaying in the MAC. Also plotted is the theoretical lower bound on the FER given by (4.26).*

Fair comparison of the proposed cooperative schemes with reference schemes is based on the constraint that in simulations, each scheme uses the same codes $\mathcal{C}_a$ and $\mathcal{C}_b$, and the same total energy E for transmission of the $2L$ source messages. We provide a comparison of the proposed cooperative scheme with two reference schemes introduced in Section 4.3.2.3. The first is consecutive relaying in the fading MAC as shown in Table 4.8. The second is a simple TDMA transmission scheme without cooperation or relaying.

Simulated performance results for bit error rate (BER) and frame error rate (FER) are shown in Figures 4.13 and 4.14 respectively. The curve corresponding to the proposed cooperative scheme exhibits the steepest slope (highest diversity gain) of the three, outperforming the others in the SNR region of interest. The scheme attains approximately an order of magnitude decrease in both BER and FER over consecutive relaying in the fading MAC at an E_b/N_0 of 10 dB. To achieve the performance of BER of 10^{-3} and FER of 10^{-2}, the proposed scheme acquires an advantage of 3dB in E_b/N_0 over the first reference scheme of consecutive relaying in the fading MAC and an advantage of nearly 12dB in E_b/N_0 over the second reference scheme of the simple TDMA transmission scheme without cooperation or relaying. The simulation results

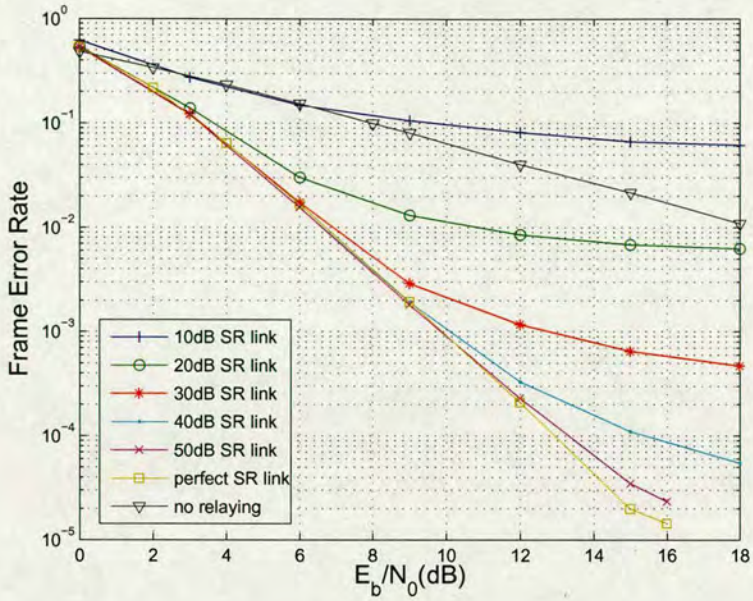

Figure 4.15: *Comparative FER performance for the proposed cooperative scheme in the case of multiple access channel. The performances are shown with various scenarios of the source-relay links: the signal-to-noise ratio ranges from 10dB to 50dB. Performance for the simple TDMA scheme without cooperation or relaying is provided for comparison.*

are consistent with the fact that the proposed scheme enjoys a diversity order of three, while the first reference scheme of consecutive relaying in the fading MAC attains a diversity order of two and the second reference scheme of the simple TDMA transmission scheme without cooperation or relaying only attains a diversity order of one. Under the same resources, the network coding approach provides an efficient way to generate diversity gain. The explanation for the gap between the theoretical FER lower bound and performance of the actual proposed scheme in Figure 4.14 is similar to that in section 4.2.5. The main reason for this gap is due to the imperfect codes and sub-optimal decoding algorithm in practical applications.

In Figure 4.15, a bundle of curves corresponding to various source-relay link scenarios are shown. It can be seen that the overall performance of the cooperative scheme is improved as the SNR of the source-relay links increases. When the SNR of the source-relay links is 50dB, at which the outage probabilities are of 10^{-5}, the overall performance approaches that of the ideal case. The intention of Figure 4.15 is to show the impact of the error propagation effect. A more advanced version of the proposed cooperative coding scheme is to apply adaptive relaying operations [LTW04] at the relay where the relay simply discards any unsuccessfully decoded

codewords rather than relaying them. This will lead to an improved performance without error floor phenomenon.

4.3.7 Conclusion

In this section we propose a cooperative diversity scheme where algebraic code superposition relaying is employed in the multiple access fading channel to create spatial diversity under the constraint of limited communications resources. We also describe in detail a novel computationally efficient message passing algorithm at the destination's decoder which extracts the substantial spatial diversity contained in the code superposition and signal superposition. The decoder is based on a sliding window structure where certain *a posteriori* LLRs are retained as *a priori* LLRs for the next decoding. We show that despite the simplicity of the proposed scheme, diversity gains are efficiently leveraged by the simple combination of channel coding at the sources and network coding at the relay.

4.4 Chapter Summary

Cooperative diversity has been recognized as an effective and low-cost technique to combat fading and enhance transmission reliability. Motivated by the fact that network coding is a technique well known for its capability to increase system throughput, joint channel and network coding is proposed in this chapter as an efficient way to generate spatial diversity under limited resource for uplink transmissions. We propose a cooperative diversity scheme for the communication model of two sources sharing a single relay under two scenarios. The scheme uses algebraic code superposition relaying in a frequency division mode or in the multiple access fading channel. In both cases a corresponding sliding-window factor graph based decoding algorithm is used at the destination node to extract available spatial diversity gains. Computational efficiency of the decoding algorithm is achieved in both cases through separation of the relevant SISO decoder modules and efficient connectivity via factor nodes corresponding to code and signal superposition operations. Simulation results demonstrate that the proposed schemes outperform competitive reference schemes.

Chapter 5
Conclusion

Cooperative diversity and LDPC codes are two principle state-of-the-art techniques to realize high-speed and high-quality wireless communication. In this thesis, we focus on practical coding schemes in cooperative communication environments, especially decoding schemes for LDPC codes. This chapter summarizes the thesis and provides suggestions for further research.

5.1 Summary

In Chapter 1, we briefly reviewed the history of wireless communications. The cooperative relaying concept and the Turbo principle, as the main topics of this thesis, are two key techniques to be employed for next generation mobile communication systems. We introduced cooperative diversity, which can be viewed as a virtual MIMO system when multiple-antenna structures cannot be employed in real mobile devices due to size, cost or hardware limitations. We also introduced LDPC codes, which were rediscovered after the success of Turbo codes and exhibit great potential for application in future wireless systems.

In Chapter 2, we offered detailed background with respect to error control coding (LDPC codes), factor graph theory and cooperative communication. We began with the basic concepts of channel codes under the AWGN channel as it shapes the coding discipline. LDPC codes, which is a class of linear block codes, have a sparse parity check matrix structure which facilitates low complexity message-passing decoding. Although the sum-product algorithm for decoding LDPC codes is suboptimum, in conjunction with soft demodulation, LDPC codes with sum-product algorithm decoding can achieve throughput results very near the capacity of the AWGN channel. As for cooperative communication, we started with the nature of the wireless fading channel. Particularly, we considered the Rayleigh flat fading channel model in this thesis. To combat the detrimental effect of fading, cooperative communication provides the receiver with multiple versions of the transmitted information through independent fading channels. As for user cooperation schemes, we focus on coded cooperation, signal superposition and coded superposition which combine channel coding and modulation techniques with

cooperative communication and inspired the work in this thesis. Also in this chapter, factor graph theory is introduced as it acts as an important graphical tool for the presentation of our results.

In Chapter 3, we concentrated on the application of LDPC codes in the multiple access channel since it can be seen as an important component of the whole cooperative communication system. Successive interference cancellation is a technique that can achieve the upper limits of the capacity region of the Gaussian multiple access channel indicated by information theory. Our concerns in this chapter are the practical application issues of successive interference cancellation. The first concern is that the Gaussian input assumption in theory is not possible in practice. Hence we go to the opposite extreme and investigate BPSK modulation. The second concern is that the subtraction of the erroneous signal would lead to propagation of the first user's decoding errors into the ensuing decoding process for the second user. Accordingly, we investigate a soft multiple user demodulator combined with the decoding of channel codes in which the interference cancellation and the decoding are carried out in a "soft" iterative process under the "Turbo principle". The major contributions of this chapter are twofold. Firstly, we present a formulation of a multi-user detector for BPSK modulation which lends itself to an efficient implementation by L-value algebra; secondly, we employ a group of moderate-blocksize LDPC codes in the LDPC coded multiple user demodulator in order to analyze practical application of successive interference cancellation under BPSK modulation scenario. The corresponding simulation results are discussed.

In Chapter 4, we proposed a cooperative diversity scheme for the communication model of a two-user multiple-access relay channel. The scheme uses algebraic code superposition relaying under two scenarios. In the first scenario, the source-destination channels and relay-destination channel are orthogonal to each other thus resulting in the destination received codewords containing only algebraically superposed codewords which is similar to the decoding scenario in [XFKC07b, XFKC07a, HD06, HSOB05]. Correspondingly, we proposed a novel and efficient decoding algorithm based on message passing on the destination node's factor graph for the purpose of exploiting the spatial diversity contained in the algebraically superposed codewords. The algorithm attains a separation of the two soft-input soft-output decoder modules corresponding to the codes employed by the two sources; for convolutional codes, this separation affords a complexity advantage over decoding of the "nested code" [XFKC06a,XFKC07b]; for LDPC codes, it affords a more efficient Tanner graph schedule than fully parallel decod-

ing [HSOB05]. In the second scenario, we consider that in a multiple-access relay channel the source-destination channels and the relay-destination channel are not orthogonal to each other and hence forms a fading multiple access channel. The corresponding decoding scenario at the destination is more complex compared to the first scenario since there is not only code superposition but also signal superposition. Developed from the previous algorithm in the first scenario, a factor graph based decoding algorithm is used at the destination node to extract the substantial spatial diversity contained in the code superposition and signal superposition. The computational efficiency of the decoding algorithm is achieved in both scenarios through separation of the relevant soft-input soft-output decoder modules and connectivity via factor nodes corresponding to code and signal superposition operations. Simulation results demonstrate that the proposed schemes outperform competitive reference schemes under a fair comparison of power constraints.

5.2 Future Work

In Chapter 3, the simulation results indicated that the assumption of Gaussian inputs in theory is not in accord with the practical modulation situation which results in a gap between theoretical evaluation and real performance. An interesting issue in the practical application is that, to one user, signals from the rest of users and the background Gaussian noise do not always appear Gaussian distributed. In some cases, *e.g.* a two user multiple access channel case using BPSK modulation, the gap is quite clear. Hence, investigations of channel code design not only in consideration of the Gaussian noise type but also the specific distribution of the interference signals could be an interesting direction to improve the overall performance of an LDPC coded multiple user demodulator.

In Chapter 4, as we already mentioned, a further development of the proposed scheme is to add a cyclic redundancy check at the relay to have an adaptive relaying schedule where the relay simply discards any unsuccessfully decoded codewords rather than forwarding them. For the joint channel and network coding scheme, the work we did in this thesis only focuses on the decoding scenarios at the destination and their corresponding algorithms. Another interesting development for this scheme is the encoding process at the relay. Although some codewords at the relay fail to pass cyclic redundancy check, it is possible that most code bits are still of high reliability and thus contain valuable information to be exploited. Based on the L-values of the decoded bits at the relay, a soft "encoding" scheme [SV05] could take into account the

reliability of each code bits and provide useful information for the decoding at the destination. We conjecture this can improve the overall performance of the schemes. In our work, we only considered a two-user case. To extend the proposed joint channel and network coding scheme to a scenario comprised of a large amount of users is an interesting option. Possible research directions to extend this work also include the optimization issue (by using density evolution method or EXIT chart method) for joint channel and network codes.

References

[ACLY00] R. Ahlswede, Ning Cai, S.-Y.R. Li, and R.W. Yeung, *Network information flow*, IEEE Trans. Inform. Theory **46** (2000), no. 4, 1204–1216.

[AK04] M. Ardakani and F. R. Kschischang, *A more accurate one-dimensional analysis and design of irregular LDPC codes*, IEEE Transactions on Communications **52** (2004), no. 12, 2106–2114.

[Ala98] S.M. Alamouti, *A simple transmit diversity technique for wireless communications*, IEEE Journal on Selecte Areas in Communications **16** (1998), no. 8, 1451–1458.

[AM00] S.M. Aji and R.J. McEliece, *The generalized distributive law*, IEEE Trans. on Information Theory **46** (2000), no. 2, 325–343.

[ARAS99] P.D. Alexander, M.C. Reed, J.A. Asenstorfer, and C.B. Schlegel, *Iterative multiuser interference reduction: turbo cdma*, IEEE Trans. on Communications **47** (1999), no. 7, 1008–1014.

[Ard04] M. Ardakani, *Efficient analysis, design and decoding of low-density parity-check codes*, Ph.D. thesis, University of Toronto, 2004.

[BC01] J. Boutros and G. Caire, *Iterative multiuser joint decoding: unified framework and asymptotic analysis*, Proc. IEEE International Symposium on Information Theory, 2001, pp. 317–.

[BC02] ———, *Iterative multiuser joint decoding: unified framework and asymptotic analysis*, IEEE Trans. on Information Theory **48** (2002), no. 7, 1772–1793.

[BCJR74] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, *Optimal decoding of linear codes for minimizing symbol error rate (corresp.)*, IEEE Trans. on Information Theory **20** (1974), no. 2, 284–287.

[BDMP98] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, *Soft-input soft-output modules for the construction and distributed iterative decoding of code networks*, European Transactions on Telecommunications **9(2)** (March-April, 1998).

[BGT93] C. Berrou, A. Glavieux, and P. Thitimajshima, *Near Shannon limit error-correcting coding and decoding: Turbo codes*, Proc. IEEE Intl. Conf. Commun. (ICC 93), 1993.

[Big05] Ezio Biglieri, *Coding for wireless channels*, Springer Science+Business Media, Inc., 2005.

[BL05] X. Bao and J. Li, *Matching code-on-graph with network-on-graph: Adaptive network coding for wireless relay networks*, Proceeding of 43rd Annual Allerton Conference on Communication, Control and Computing (Allerton) , Champaign, IL, (Sept. 2005).

[BL06] ———, *A unified channel-network coding treatment for wireless ad-hoc networks*, Proceeding of IEEE International Symposium on Information Theory (ISIT) (2006).

[Bol79] B Bollobas, *Graph theory, an introductory course*, New York: Springer-Verlag, 1979.

[CdBSA05] A. Chakrabarti, A. de Baynast, A. Sabharwal, and B. Aazhang, *LDPC code design for half-duplex decode-and-forward relaying*, 43rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, USA (2005).

[CFRU01] S.-Y. Chung, G.D. Forney, T.J Richardson, and R.L. Urbanke, *On the design of low-density parity-check codes within 0.0045 dB of the shannon limit*, IEEE Communications Letters **5** (2001), 58–60.

[CG79] T. Cover and A.E. Gamal, *Capacity theorems for the relay channel*, IEEE Trans. on Information Theory **25** (1979), no. 5, 572–584.

[CG90] J. T. Coffey and R. M. Goodman, *Any code of which we cannot think is good*, IEEE Trans. on Information Theory **36** (1990), no. 6, 1453–1461.

[CKL06] Y. Chen, S. Kishore, and J. Li, *Wireless diversity through network coding*, Proceeding of IEEE Wireless Communications and Networking Conference (WCNC), Las Vegas, NV (2006).

[CMT04] G. Caire, R. Muller, and T. Tanaka, *Iterative multiuser joint decoding: optimal power allocation and low-complexity implementation*, IEEE Trans. on Information Theory **50** (2004), no. 9, 1950–1973.

[CT06] T.M. Cover and J.A. Thomas, *Elements of information theory*, second ed., John Wiley & Sons, Inc., 2006.

[CV05] Y Cao and B. Vojcic, *Cooperative coding using serial concatenated convolutional codes*, Proc. IEEE Wireless Communications and Networking Conference, vol. 2, 2005, pp. 1001–1006 Vol. 2.

[CXDL04] L. Chen, J. Xu, I. Diurdjevic, and S. Lin, *Near-shannon-limit quasi-cyclic low-density parity-check codes*, IEEE Trans. on communications **52** (2004), no. 7, 1038–1042.

[DEH+05] S. Deb, M. Effros, T. Ho, D.R. Karger, R. Koetter, D.S. Lun, M. Medard, and N. Ratnakar, *Network coding for wireless applications: A brief tutorial*, IWWAN (2005).

[DLGT07] Z. Ding, K. K. Leung, D. L. Goeckel, and D. Towsley, *On the study of network coding with wireless diversity*, Proceeding of 1st Annual Conference of ITA (2007).

[DPSB07] E. Dahlman, S. Parkvall, J. Skoeld, and P. Beming, *3G evolution: HSPA and LTE for mobile broadband*, Academic press, 2007.

[DVB] *Digital video broadcasting (dvb)*.

[EGH01] H. El Gamal and Jr. Hammons, A. R., *Analyzing the turbo decoder using the gaussian approximation*, IEEE Transactions on Information Theory **47** (2001), no. 2, 671–686.

[FK96] B. J. Frey and F. R. Kshischang, *Probability propagation and iterative decoding*, Proc. 34th Allerton Conf. Communications, Control, and Computing (1996).

[FKLW97] B. J. Frey, F. R. Kshischang, H.-A. Loeliger, and N. Wiberg, *Factor graphs and algorithms*, Proc. 2003 35th Allerton Conf. Communications, Control, and Computing (Sept. 29-Oct. 1, 1997), 666–680.

[Fla05] M. F. Flanagan, *Combined coding and equalization for dispersive communications channels*, Ph.D. thesis, University College Dublin, 2005.

[FMI99] M. Fossorier, M. Mihaljevic, and H. Imai, *Reduced complexity iterative decoding of low density parity check codes*, IEEE Trans. on Information Theory **47** (1999), 673–680.

[For73] G.D. Forney, *The Viterbi algorithm*, Proc. IEEE **61** (1973), 268–278.

[For01] Jr. Forney, G.D., *Codes on graphs: normal realizations*, IEEE Trans. on Information Theory **47** (2001), no. 2, 520–548.

[Fre98] B. J. Frey, *Graphical models for machine learning and digital communication*, Cambridge, MA: MIT Press, 1998.

[Gal63] R.G. Gallager, *Low density parity check codes*, Cambridge:MA:MIT Press, 1963.

[GFZ03] J. Garcia-Frias and W. Zhong, *Approaching Shannon performance by iterative decoding of linear codes with low-density generator matrix*, IEEE Communications Letters **7** (2003), 266– 268.

[Gol05] A. Goldsmith, *Wireless communications*, Cambridge University Press, 2005.

[GRUW01] A.J. Grant, B. Rimoldi, R.L. Urbanke, and P.A. Whiting, *Rate-splitting multiple access for discrete memoryless channels*, IEEE Trans. on Information Theory **47** (2001), no. 3, 873–890.

[HA04] M. O. Hasna and M. S. Alouini, *A performance study of dual-hop transmissions with fixed gain relays*, IEEE Transactions on Wireless Communications **3** (2004), no. 6, 1963–1968.

[Hag02] J. Hagenauer, *The turbo principle in mobile communications*, International Symposium on Information Theory and Its Applications (ISITA), 2002.

[HD06] C. Hausl and P. Dupraz, *Joint network-channel coding for the multiple-access relay channel*, Sensor and Ad Hoc Communications and Networks, 2006. SECON '06. 2006 3rd Annual IEEE Communications Society on, vol. 3, 28-28 Sept. 2006, pp. 817–822.

[HH06] C. Hausl and J. Hagenauer, *Iterative network and channel decoding for the two-way relay channel*, Proc. IEEE International Conference on Communications ICC '06, vol. 4, 2006, pp. 1568–1573.

[HHK08] J. Hou, C. Hausl, and R. Koetter, *Distributed turbo coding schemes for asymmetric two-way relay communication*, Proc. International Symposium on Turbo Codes & Related Topics, 2008.

[HL06] Ruiyuan Hu and Jing Li, *Practical compress-forward in user cooperation: Wyner-ziv cooperation*, IEEE International Symposium on Information Theory, July 2006, pp. 489–493.

[HN02a] T.E. Hunter and A. Nosratinia, *Coded cooperation under slow fading, fast fading, and power control*, Proceeding of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, vol. 1, 2002, pp. 118–122 vol.1.

[HN02b] ———, *Cooperation diversity through coding*, Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on, 2002, p. 220.

[HN06] ———, *Diversity through coded cooperation*, IEEE Trans. on Wireless Communications **5** (2006), no. 2, 283–289.

[HOP96] J. Hagenauer, E. Offer, and L. Papke, *Iterative decoding of binary block and convolutional codes*, IEEE Trans. on Information Theory **42** (1996), no. 2, 429–445.

[HSOB05] C. Hausl, F. Schreckenbach, I. Oikonomidis, and G. Bauch, *Iterative network and channel decoding on a tanner graph*, 43rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, USA (2005).

[IH77] H. Imai and S. Hirakawa, *A new multilevel coding method using error-correcting codes*, IEEE Trans. on Information Theory **23** (1977), no. 3, 371–76.

[ITU] *Next generation home network technology standard (itu -t g.hn).*

[Jak74] W. C. Jakes, *Microwave mobile communications*, Ed., Wiley, New York, 1974.

[JHHN04] M. Janani, A. Hedayat, T.E. Hunter, and A. Nosratinia, *Coded cooperation in wireless communications: space-time transmission and iterative decoding*, Signal Processing, IEEE Transactions on **52** (2004), no. 2, 362–371.

[KA00] D. Koulakiotis and A.H. Aghvami, *Data detection techniques for DS/CDMA mobile systems: a review*, IEEE [see also IEEE Wireless Communications] Personal Communications **7** (2000), no. 3, 24–34.

[KAA04] M.A. Khojastepour, N. Ahmed, and B. Aazhang, *Code design for the relay channel and factor graph decoding*, Thirty-Eighth Asilomar Conference on Signal, Systems and Computers **2** (2004), 2000–2004.

[KF98] F.R. Kschischang and B.J. Frey, *Iterative decoding of compound codes by probability propagation in graphical models*, IEEE Journal on Selected Areas in Communications **16** (1998), 219–230.

[KFL01] F. R. Kshischang, B. J. Frey, and H.-A Loeliger, *Factor graphs and the sum-product algorithm*, IEEE Trans. on Information Theory **47** (2001), no. 2, 498–519.

[KGG05] G. Kramer, M. Gastpar, and P. Gupta, *Cooperative strategies and capacity theorems for relay networks*, Information Theory, IEEE Transactions on **51** (2005), no. 9, 3037–3063.

[KH97] S. Kaiser and J. Hagenauer, *Multi-carrier cdma with iterative decoding and soft-interference cancellation*, Proc. IEEE Global Telecommunications Conference GLOBECOM '97, vol. 1, 1997, pp. 6–10 vol.1.

[KH00] R. Knopp and P. A. Humblet, *On coding for block fading channels*, IEEE Transactions on Information Theory **46** (2000), no. 1, 189–205.

[KKH+05] S. Katti, D. Katabi, W. Hu, H. Rahul, and M. Medard, *The importance of being opportunistic: practical network coding for wireless environments*, Allerton (2005).

[KLF00a] Y. Kou, S. Lin, and M. P. C. Fossorier, *Low density parity check codes based on finite geometries: A rediscovery and new results*, IEEE International symposiums of information theory, Italy (2000).

[KLF00b] ———, *Low density parity check codes: construction based on finite geometries*, Proc. IEEE Global Telecommunications Conference GLOBECOM '00, vol. 2, 2000, pp. 825–829 vol.2.

[KLF00c] Y. Kou, S. Lin, and M.P.C Fossorier, *Construction of low density parity che codes-a geometric approach*, Proc. of 2nd International symposiums on Turbo codes and related topics, Brest, France, 2000.

[KLF01] ———, *Low density parity check codes based on finite geometries: A rediscovery and new results*, IEEE Trans. on Information Theory **47** (2001), 2711–2736.

[KM03] R. Koetter and M. Medard, *An algebraic approach to network coding*, IEEE/ACM Trans. on networking **11** (2003), no. 5, 782–795.

[Koc07] V. M. Koch, *A factor graph approach to model-based signal separation*, Ph.D. thesis, ETH Zurich, 2007.

[KRH+08] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, *Xors in the air: Practical wireless network coding*, IEEE/ACM Transactions on Networking **16** (2008), 497–510.

[KS01] P.L. Kafle and A.B. Sesay, *Performance of turbo coded multicarrier CDMA with iterative multiuser detection and decoding*, Proc. Canadian Conference on Electrical and Computer Engineering, vol. 1, 2001, pp. 105–110 vol.1.

[KS03] ———, *Iterative semi-blind multiuser detection for coded MC-CDMA uplink system*, IEEE Trans. on Communications **51** (2003), no. 7, 1034–1039.

[KTMG08] I. Krikidis, J. Thompson, S. McLaughlin, and N. Goertz, *Optimization issues for cooperative amplify-and-forward systems over block-fading channels*, IEEE Transactions on Vehicular Technology **57** (2008), no. 5, 2868–2884.

[KTMG09] ———, *Max-min relay selection for legacy amplify-and-forward systems with interference*, IEEE Transactions on Wireless Communications **8** (2009), 3016–3027.

[KW00] G. Kramer and A.J. Wijngaarden, *On the white gaussian multiple-access relay channel*, IEEE International Symposium on Information Theory (ISIT) (2000), 40.

[LC04] S. Lin and Daniel J. Costello, *Error control coding: Fundamentals and applications*, 2nd ed., Pearson Prentice Hall, 2004.

[LCZ+06] Z Li, L. Chen, L. Zeng, S. Lin, and W. H. Fong, *Efficient encoding of quasi-cyclic low-density parity-check codes*, IEEE Transactions on Communications **54** (2006), no. 1, 71–81.

[LFKL00] R. Lucas, M.P.C. Fossorier, Y. Kou, and S. Lin, *Iterative decoding of one-step majority logic decodable codes based on belief propagation*, IEEE Trans. on communications **48** (2000), no. 6, 931–937.

[LK00] S. Lin and Y. Kou, *A geometric approach to the construction of low density parity check codes*, presented at IEEE 29th Communication Theory Workshop, Haines City, Fla., 2000.

[LKF00] S. Lin, Y. Kou, and M.P.C Fossorier, *Finite geometry low density parity check codes: construction, structure and decoding*, Proceedings of the ForneyFest Kluwer academic, Boston, Mass., 2000.

[LMSS01] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, *Improved low-density parity check codes using irregular graphs*, IEEE Trans. on Information Theory **47** (2001), no. 2, 585–598.

[Loe04] H.-A. Loeliger, *An introduction to factor graphs*, IEEE Signal Processing Magazine **21** (2004), no. 1, 28–41.

[LS88] S. L. Lauritzen and D. J. Spiegelhalter, *Local computations with probabilities on graphical structures and their application to expert systems*, J. Roy. Statist. Soc. **50** (1988), 157–224.

[LTW04] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, *Cooperative diversity in wireless networks: Efficient protocols and outage behavior*, IEEE Transactions on Information Theory **50** (2004), no. 12, 3062–3080.

[LV05] E.G. Larsson and B.R. Vojcic, *Cooperative transmit diversity based on superposition modulation*, IEEE Communications Letters **9** (2005), no. 9, 778–780.

[LVWD06] Y Li, B. Vucetic, T.F. Wong, and M. Dohler, *Distributed turbo coding with soft information relaying in multihop relay networks*, IEEE Journal on Selected Areas in Communications **24** (2006), no. 11, 2040–2050.

[LYC03] S.-Y.R. Li, R.W. Yeung, and Ning Cai, *Linear network coding*, IEEE Trans. on Information Theory **49** (2003), no. 2, 371–381.

[LYK+08] C. Li, G. Yue, M.A. Khojastepour, X. Wang, and M. Madihian, *LDPC-coded cooperative relay systems: performance analysis and code design*, IEEE Trans. on Communications **56** (2008), 485–496.

[Mac99a] D. J. C MacKay, *Sparse graph codes*, 5th Intl. Symp. Commun. Theory and Applications, 1999, pp. 11–16.

[Mac99b] D.J.C. MacKay, *Good error-correcting codes based on very sparse matrices*, IEEE Trans. on Information Theory **45** (Mar. 1999), 399–431.

[Mar82] G. A. Margulis, *Explicit constructions of graphs without short cycles and low-density codes*, Combinatorica 2 **1** (1982), 71–78.

[MMC98] R. J. McEliece, D.J.C. MacKay, and J.-F Cheng, *Turbo decoding as an instance of pearl's 'belief propagation' algorithm*, IEEE J. Select. Areas Commun. **16** (1998), 140–152.

[MN95] D.J.C. MacKay and R.M. Neal, *Good codes based on very sparse matrices*, 5th IMA conference in cryptography and coding **1025** (1995), 100–111.

[Moh98] M. Moher, *An iterative multiuser decoder for near-capacity communications*, IEEE Trans. on Communications **46** (1998), 870–880.

[MS84] R. McEliece and W. Stark, *Channels with block interference*, IEEE Transactions on Information Theory **30** (1984), no. 1, 44–53.

[NHH04] A. Nosratinia, T.E. Hunter, and A. Hedayat, *Cooperative communication in wireless networks*, IEEE Communications Magazine **42** (2004), no. 10, 74–80.

[Pea88] J. Pearl, *Probabilistic reasoning intelligent systems, 2nd ed.*, San Francisco, CA: Kaufmann, 1988.

[Poo04] H.V. Poor, *Iterative multiuser detection*, IEEE Signal Processing Magazine **21** (2004), 81–88.

[Pro01] J.G. Proakis, *Digital communications*, The McGraw-Hill Companies,Inc., 2001.

[Rap01] T.S. Rappaport, *Wireless communications: Principles and practice*, Prentice Hall PTR, 2001.

[RSAA98] M.C. Reed, C. B. Schlegel, P.D. Alexander, and J.A. Asenstorfer, *Iterative multiuser detection for cdma with fec: near-single-user performance*, IEEE Trans. on Communications **46** (1998), 1693–1699.

[RSU01] T.J. Richardson, M.A. Shokrollahi, and R.L. Urbanke, *Design of capacity-approaching irregular low-density parity-check codes*, IEEE Trans. on Information Theory **47** (2001), no. 2, 619–637.

[RSW03] A. Ramamoorthy, J. Shi, and R. Wesel, *On the capacity of network coding for wireless networks*, 41st Annual Allerton Conference on Communication, Control and Computing (2003).

[RU01] T.J. Richardson and R.L. Urbanke, *The capacity of low-density parity-check codes under message-passing decoding*, IEEE Trans. on Information Theory **47** (2001), no. 2, 599–618.

[SDE+01a] P. H. Siegel, D. Divsalar, E. Eleftheriou, J. Hagenauer, D. Rowitch, and W. H. Tranter, *Guest editorial the turbo principle: from theory to practice*, IEEE Journal on Selected Areas in Communications **19** (2001), no. 5, 793–799.

[SDE+01b] P. H. Siegel, D. Divsalar, E. Elepftheriou, J. Hagenauer, and D. Rowitch, *Guest editorial - the turbo principle: from theory to practice ii*, IEEE Journal on Selected Areas in Communications **19** (2001), no. 9, 1657–1661.

[SE04] A. Stefanov and E. Erkip, *Cooperative coding for wireless networks*, IEEE Transactions on communications **52** (2004), 273–277.

[SEA98] A. Sendonaris, E. Erkip, and B. Aazhang, *Increasing uplink capacity via user cooperation diversity*, Proc. IEEE International Symposium on Information Theory, 1998, pp. 156–.

[SEA03a] ———, *User cooperation diversity-Part I: System description*, IEEE Transactions on Communications **51** (2003), 1927 – 1938.

[SEA03b] ———, *User cooperation diversity-Part II: Implementation aspects and performance analysis*, IEEE Transactions on Communications **51** (2003), 1939 – 1948.

[Sha48] C. E. Shannon, *A mathematical theory of communication*, Bell System Technical Journal **27** (1948), 379–423 (Part 1); 623–56 (Part 2).

[SHMX06] V. Stankovic, A. Host-Madsen, and Zixiang Xiong, *Cooperative diversity for wireless ad hoc networks*, IEEE Signal Processing Magazine **23** (2006), no. 5, 37–49.

[SKM04] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, *Capacity theorems for the multiple-access relay channel*, 42nd Annual Allerton Conference on Communication, Control, and Computing, 2004.

[SS96] M. Sipser and D. A. Spielman, *Expander codes*, IEEE Trans. on Information Theory **42** (1996), no. 6, 1710–1722.

[SS06] Zhenning Shi and C. Schlegel, *Iterative multiuser detection and error control code decoding in random cdma*, IEEE Trans. on signal processing **54** (2006), no. 5, 1886–1895.

[STB09] S. Sesia, I. Toufik, and M. Baker, *LTE: The UMTS long term evolution*, WILEY, 2009.

[SV05] H. H. Sneesens and L. Vandendorpe, *Soft decode and forward improves cooperative communications*, Proc. 1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 13–13 Dec. 2005, pp. 157–160.

[Tan81] R. Tanner, *A recursive approach to low complexity codes*, IEEE Trans. on Information Theory **27** (1981), no. 5, 533–547.

[tB00] S. ten Brink, *Iterative decoding trajectories of paralle concatenated codes*, Proc. 3rd IEE/ITG Conference on Source and Channel Coding (2000), 75–80.

[tB01] S. ten Brink, *Convergence behavior of iteratively decoded parallel concatenated codes*, IEEE Transactions on Communications **49** (2001), no. 10, 1727–1737.

[TJC99] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, *Space-time block codes from orthogonal designs*, IEEE Trans. on Information Theory **45** (1999), no. 5, 1456–1467.

[TSC98] V. Tarokh, N. Seshadri, and A.R. Calderbank, *Space-time codes for high data rate wireless communication: performance criterion and code construction*, IEEE Trans. on Information Theory **44** (1998), no. 2, 744–765.

[TSS+04] R.M. Tanner, D. Sridhara, A. Sridharan, T.E. Fuja, and Jr. Costello, D.J., *LDPC block and convolutional codes based on circulant matrices*, IEEE Trans. on Information Theory **50** (2004), no. 12, 2966–2984.

[TV05] D. Tse and P. Viswanath, *Fundamentals of wireless communication*, Cambridge University Press, 2005.

[Ung82] G. Ungerboeck, *Channel coding with multilevel/phase signals*, IEEE Trans. on Information Theory **28** (1982), no. 1, 55–67.

[WCK05] Y. Wu, P. A. Chou, and S.-Y. Kung, *Information exchange in wireless networks with network coding and physical-layer broadcast*, Proc. Conference on Information Sciences and Systems, 2005.

[Wil83] F. Willems, *The discrete memoryless multiple access channel with partially cooperating encoders (corresp.)*, IEEE Trans. on Information Theory **29** (1983), no. 3, 441–445.

[Wil89] J. C. Willems, *Models for dynamics*, Dynamics Reported, Volume 2, U. Kirchgraber and H. O.Walther, Eds. New York:Wiley (1989), 171–269.

[WK07] D. Woldegebreal and H. Karl, *Network-coding-based adaptive decode and forward cooperative transmission in wireless networks: outage analysis*, Proceeding of 13th European Wireless Conference (2007).

[WLK95] N. Wiberg, H.-A Loeliger, and R. Koetter, *Codes and iterative decoding on general graphs*, European Transactions on Telecommunications **6** (1995), 513–526.

[Wol78] J. Wolf, *Efficient maximum likelihood decoding of linear block codes using a trellis*, IEEE Trans. on Information Theory **24** (1978), no. 1, 76–80.

[WP99] X. Wang and H.V. Poor, *Iterative (turbo) soft interference cancellation and decoding for coded cdma*, IEEE Trans. on Communications **47** (1999), no. 7, 1046–1061.

[XFKC06a] L. Xiao, T.E. Fuja, J. Kliewer, and D.J. Costello, *Cooperative diversity based on code superposition*, IEEE International Symposium on Information Theory (ISIT) (2006).

[XFKC06b] L. Xiao, T.E. Fuja, J. Kliewer, and D.J.Jr. Costello, *Nested codes with multiple interpretations*, Proc. 40th Annual Conference on Information Sciences and Systems, 2006, pp. 851–856.

[XFKC07a] L. Xiao, T.E. Fuja, J. Kliewer, and D.J. Costello, *Algebraic superposition of LDGM codes for cooperative diversity*, IEEE International Symposium on Information Theory (ISIT) (2007).

[XFKC07b] ———, *A network coding approach to cooperative diversity*, IEEE Trans. Information Theory **53** (2007), no. 10, 3714–3722.

[XLC04] Zixiang Xiong, A. D. Liveris, and S. Cheng, *Distributed source coding for sensor networks*, IEEE Signal Processing Magazine **21** (2004), no. 5, 80–94.

[YWN03] Guosen Yue, Xiaodong Wang, and K.R. Narayanan, *Design of low density parity check codes for turbo multiuser detection*, Proc. IEEE International Conference on Communications ICC '03, vol. 4, 2003, pp. 2703–2707 vol.4.

[YWYHM08] Guosen Yue, Xiaodong Wang, Zigui Yang, and A. Host-Madsen, *Coding schemes for user cooperation in low-power regimes*, IEEE Trans. on Signal Processing **56** (2008), no. 5, 2035–2049.

[ZLL06] S. Zhang, S. Liew, and P. Lam, *Physical layer network coding*, Proc. 12th Annual International Conference on Mobile Computing and Networking (ACM MobiCom) (2006).

[ZV03] B. Zhao and M.C. Valenti, *Distributed turbo coded diversity for relay channel*, Electronics Letters **39** (2003), 786– 787.

[ZZLL07] Shengli Zhang, Yu Zhu, Soung-Chang Liew, and Khaled Ben Letaief, *Joint design of network coding and channel decoding for wireless networks*, Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE, 11-15 March 2007, pp. 779–784.

Appendix A
Original Publications

The author of this thesis has the following accepted or submitted publications during the course of her Ph.D.

A.1 Journal Paper

- X. Xu, M. Flanagan, N. Goertz; "Joint Channel and Network Coding for Cooperative Diversity in a Shared-Relay Environment", submitted to *IEEE Transactions on Wireless Communications*, August 2008.

A.2 Conference Papers

- X. Xu, N. Goertz; "Practical Successive Interference Cancellation in the Binary-Input Gaussian Multiple Access Channel", *International ITG Conference on Source and Channel Coding*, Ulm, Germany, January 2008, ISSN 0932-6022.

- X. Xu, M. Flanagan, N. Goertz; "A Shared-Relay Cooperative Diversity Scheme Based on Joint Channel and Network Coding in the Multiple Access Channel", *International Symposium on Turbo Codes and Related Topics (Turbo-Coding-2008)*, 243-248, Lausanne, Switzerland, September 2008.

- X. Xu, M. Flanagan, C. Koller, N. Goertz; "A Shared-Relay Cooperative Diversity Scheme Based on Joint Channel and Network Coding", *IEEE International Symposium on Information Theory and its Applications (ISITA 2008)*, 417-422, Auckland, New Zealand, December 2008.

The original publications are included in the following pages.

Joint Channel and Network Coding for Cooperative Diversity in a Shared-Relay Environment

Xiaoyan Xu, Mark F. Flanagan, *Member, IEEE,*

and Norbert Goertz, *Senior Member, IEEE*

Abstract

In this paper we propose a cooperative diversity scheme for the communication model of two sources sharing a single relay under two scenarios. The scheme uses algebraic code superposition relaying in the multiple access fading channel or in a frequency division mode to create spatial diversity under the constraint of limited communications resources. We also describe in detail a novel computationally efficient message passing algorithm at the destination's decoder which extracts the substantial spatial diversity contained in the code superposition and signal superposition. The decoder is based on a sliding window structure where certain *a posteriori* LLRs are retained to form *a priori* LLRs for the next decoding. We show that despite the simplicity of the proposed scheme, diversity gains are efficiently leveraged by the simple combination of channel coding at the sources and network coding at the relay.

Index Terms

Cooperative diversity, low-density parity-check codes, network coding, factor graphs, iterative decoding, multiple access channel.

I. Introduction

WHILE wireless channels suffer from fading, at the same time the broadcast nature of wireless channels provides the possibility of a third party other than the destination "overhearing" the information that the source transmits. Thus apart from the original transmission channel, the same

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), grant number EP/D 002184/1, and by the Scottish Funding Council for the Joint Research Institute with Heriot-Watt University which is a part of the Edinburgh Research Partnership. The material in this paper is to be presented in part at the 5th International Symposium on Turbo Codes and Related Topics, Turbo-Coding-2008, Lausanne, Switzerland, September 2008.

X. Xu and N. Goertz are with the Institute for Digital Communications, Joint Research Institute for Signal & Image Processing, The University of Edinburgh, Scotland, UK (e-mail: X.Xu@ed.ac.uk, Norbert.Goertz@ieee.org).

Mark F. Flanagan is with Department of Electronics, Computer Science and Systems (DEIS), The University of Bologna, Italy (e-mail: mark.flanagan@ieee.org).

information could be transmitted to the destination through another independently fading channel. This generated spatial diversity can effectively combat the deleterious effect of fading [1]–[3]. Although "coding" has been employed in several literature on cooperative communications [4]–[7], the diversity gain are obtained through channel coding gain by means of single channel codes. In recent years, there has been increasing interest in applying the idea of algebraic code superposition, also called "network coding" [8]–[12] to the cooperative communications scenario. The network coding approach provides an efficient way to generate spatial diversity under the constraint of limited resources. One challenge is the problem of decoder design which should be able to cope with the complicated decoding situation at the destination [8]–[12].

In [11]–[13], the model of a typical network coding unit is considered in which the packets from the two sources are linearly combined at the relay. In our work we also focus on this cooperative transmission model for the situation where it is impractical for one mobile user to "capture" the other user's signal during its uplink transmission in the cellular network. Moreover, relay-based cooperative processing provides greater security than direct user cooperation in which user information must be shared. In [14], it is shown theoretically that network-coded distributed antenna systems provide a performance improvement over similar systems without network coding, with less hardware and a lower bandwidth cost.

In [10], a code superposition scheme employing low-density generator matrix (LDGM) codes is proposed to reduce the decoding complexity at the destination. But in order to do the graph-based decoding, the systematic bits must be retained without superposition which means that the potential superposition diversity is lower than that obtainable from fully superposed codewords. In [12], a combined low-density parity-check (LDPC) code construction scheme including two channel code components and one network code component is produced by random parity-check matrix generation under certain constraints. The network codes are actually the parity checks for two channel codewords; this necessitates more complicated relay operations than simple superposition. In [15] and [16], the authors proposed an adaptive network coded cooperation (ANCC) scheme, which couples the instantaneous network topology with codes on graphs, such as LDPC codes. Since the number of terminals engaged in operation decides the number of variable nodes in the final parity check matrix which employs the message-passing algorithm for decoding, this method aims to target cooperation among a large number of terminals under the wireless ad hoc network scenario. The decoding method of ANCC cannot deal with the challenge of cooperation among a small number of users, *e.g.* two users.

In this work, we consider the two-source one-relay cooperative model under two scenarios. The first scenario is where the source-destination transmission and relay-destination transmission use a fading multiple access channel (MAC). In this scenario we propose a cooperative coding scheme which is different from the previous work of [8]–[12] where superposed codewords experience a channel orthogonal to that of the original transmission, and also different from the previous work of [17] where simple codeword retransmission is employed in the multiple access Gaussian relay channel. Our scheme allows continuous transmission of superposed codewords by the relay and at the same time targets the challenge of coping with the interference introduced by the multiple access channel, thus making efficient use of communication resources to leverage spatial diversity gains. The second scenario is where an extra frequency band is available for the relay-destination transmission. The cooperative coding scheme in this scenario, in contrast to previous work where TDMA-only or FDMA-only relaying is assumed, again allows continuous transmission of superposed codewords by the relay, making efficient use of communication resources to leverage available diversity gains. This second scenario arises from the observation that idle frequency channels do exist and could be exploited in current TDMA systems. We also detail a novel efficient decoding algorithm based on message passing on the destination node's factor graph for the purpose of exploiting the spatial diversity contained in the algebraically superposed codewords together with (in the case of no extra available frequency band) the signal superposition introduced by the MAC. For the case of multiple access channel, the algorithm attains a separation of the three soft-input soft-output (SISO) decoder modules corresponding to each received signal stream; for the case of extra available frequency band, the algorithm attains a separation of the two SISO decoder modules corresponding to the codes employed by the two sources. In both cases, for convolutional codes this separation affords a complexity advantage over decoding of the "nested code" [8], [9]; for LDPC codes it affords a more efficient Tanner graph schedule than fully parallel decoding [12].

II. Proposed Cooperative Coding Scheme

A. Cooperative Coding Scheme: Case of Fading Multiple Access Channel

We consider the four-node communications network depicted in Figure 1, with two sources S_a and S_b, one relay R, and one destination D common to the two sources. The communication period is divided into $L+1$ *time slots* $t = 0, 1, \cdots, L$; each time slot $t \in \{0, 1, \cdots, L\}$ is further subdivided into 2 *half slots* $(2t,\ 2t+1)$. Source S_a has L messages to transmit, which it encodes into L n-bit codewords

$\{\mathbf{a}_t : t = 0, 1, \cdots, L-1\}$. The code used at source S_a is $\mathcal{C}_a$ and is defined by the $m_a \times n$ parity-check matrix $\mathbf{H}_a = (H_a(j,i))$. Similarly, source S_b has L messages to transmit, which it encodes into L n-bit codewords $\{\mathbf{b}_t : t = 0, 1, \cdots, L-1\}$. The code used at source S_b is $\mathcal{C}_b$ and is defined by the $m_b \times n$ parity-check matrix $\mathbf{H}_b = (H_b(j,i))$. Thus, the codes $\mathcal{C}_a$ and $\mathcal{C}_b$ have the same length but not necessarily the same rate. In general, the codes used at the two sources can be LDPC or convolutional; in this paper we concentrate on LDPC codes. S_a and S_b broadcast their modulated codewords to the relay and destination nodes using TDMA and there is no cooperation between the two sources. For each $t \in \{0, 1, \cdots, 2L-1\}$, let $\mathbf{c}_t$ denote the codeword broadcast by the source in half slot t; thus $\mathbf{c}_{2t} = \mathbf{a}_t$ and $\mathbf{c}_{2t+1} = \mathbf{b}_t$ for $t \in \{0, 1, \cdots, L-1\}$.

The relay decodes and then re-encodes each codeword received from the source (the cooperative scheme is based on the scenario where the source is quite close to the relay). The relay also has a buffer in which it stores the codewords it obtained in the previous two half slots. At each half slot ($t = 2, 3, \cdots, 2L$), the relay interleaves the codeword obtained in half slot $t-1$ and superposes it (XOR operation) with the codeword obtained in half slot $t-2$; it then transmits the resulting codeword to the destination in the fading multiple access channel whose channel resources are also shared with the source transmission. Special cases arise at half slots 1 and $2L+1$ in which only a single codeword is stored at the relay and no XOR operation is performed. Let $\mathbf{d}_t$ denote the codeword transmitted from the relay to the destination in half slot $t \in \{1, 2, \cdots, 2L+1\}$; thus $\mathbf{d}_t = \pi(\mathbf{c}_{t-1}) \oplus \mathbf{c}_{t-2}$, for $t = 2, 3, \cdots 2L$. Let $\mathbf{e}_t$ denote the signal stream received by the destination in half slot $t \in \{0, 1, \cdots, 2L+1\}$. The use of the MAC allows for no extra expense in terms of channel resources for the proposed cooperative transmission scheme as compared to the non-cooperative scheme. As we shall see, the spatial diversity gain generated by the superposition relaying outweighs the signal interference degradation inherent in using the MAC. To simplify the analysis, we assume that directed antennas are employed at the relay, so that interference between the source-relay link and the relay-destination link may be neglected.[1] For each $t = 0, 1, \cdots L-1$, source S_a's codeword $\mathbf{a}_t$ is decoded at the end of half slot $2t+2$ and source S_b's codeword $\mathbf{b}_t$ is decoded at the end of half slot $2t+3$. The transmission schedule for this cooperative coding scheme is illustrated in Figure 2. It can be seen that spatial diversity for each message is contained in three separate transmissions spanning three

[1]To avoid full-duplex at the relay in a practical communication system, a second relay R' which employs a simple amplify-and-forward (AF) scheme could be used between R and D in the current cell frequency $f1$. The transmission from R to R' could employ the neighboring cell frequency $f2$, and thus the uplink MAC still maintain $f1$ for this cell. In the counterpart neighboring cell with frequency $f2$, $f1$ is us ed in the R to R' link. In the overall communications system, there will be no extra frequency band occupied and only negligible inter-cell interference. The decoding procedure at the destination node is then identical to that presented in this paper.

half slots.

As will be seen in Section IV, the interleaver π provides the "interleaver gain" for decoding at the destination. The interleaver is not in general necessary in the case of LDPC coding; however it may be used to avoid a large multiplicity of 8-cycles in the Tanner graph for the case where $\mathbf{H}_a = \mathbf{H}_b$.

As one reference system, the transmission schedule for consecutive relaying in the MAC is depicted in Figure 3; here the relay simply re-transmits the (interleaved) previously received codeword rather than a codeword superposition. It is easily seen that in this scheme, spatial diversity for each message is contained in two separate transmissions spanning two half slots. A simulation-based comparison of the two schemes described in this section under a transmit power constraint will be given in Section V-B1.

B. Cooperative Coding Scheme: Case of Extra Available Frequency Band

Under some TDMA protocols, idle frequency bands do exist and may be employed in the above cooperation scheme to avoid the signal interference inherent in using the multiple access channel. In the cooperative coding scheme with an extra available frequency band, S_a and S_b in turn broadcast their modulated codewords to the relay and destination nodes using TDMA in frequency band f_1, while the relay transmits the corresponding superposed codeword to the destination in frequency band $f2$. Apart from this, the transmission schedule and relay operation are the same as in the fading multiple access channel case described in Section II-A. This transmission schedule is also represented by Figure 2 with the difference that the destination no longer receives a superposed signal $\mathbf{e}_t$ but signals transmitted from sources and relay separately.

As a reference scheme, the corresponding transmission schedule for consecutive relaying uses the same time and frequency allocations, but the relay simply retransmits the previously received codeword rather than a codeword superposition. This transmission schedule is represented by Figure 3 with the modification that the destination no longer receives a superposed signal $\mathbf{e}_t$ but separate transmissions from sources and relay. Also, the transmission schedules for simple TDMA and FDMA based relaying schemes are depicted in Figures 4 and 5 respectively. It is easily seen that in all three of these reference schemes, spatial diversity for each message is contained in two codewords received at the destination. A simulation-based comparison of all schemes described in this section under a transmit power constraint will be given in Section V-B2.

III. Soft Demodulator for BPSK Modulation in the Multiple Access Channel

In the case where the MAC is used, a SISO signal processing module is required to update the extrinsic information regarding the coded bits involved in the signal superposition, using corresponding input extrinsic information together with the received channel values. This section describes how this may be achieved.

In the following we define $\mathcal{I} = \{1, 2, \cdots, n\}$, and $x^{(i)}$ denotes the i-th bit of codeword $\mathbf{x}$ while $e^{(i)}$ denotes the i-th received channel value of signal stream $\mathbf{e}$. Considering a single half slot $t \in \{1, \cdots, 2L-1\}$, the channel model is given by

$$e_t^{(i)} = \phi_t^{(i)} \cdot \alpha(c_t^{(i)}) + \psi_t^{(i)} \cdot \beta(d_t^{(i)}) + n_t^{(i)}, \quad i \in \mathcal{I}\,, \tag{1}$$

where $\phi_t^{(i)}$ and $\psi_t^{(i)}$ represent the fading processes on the source-destination and relay-destination links, respectively, and $n_t^{(i)}$ is complex AWGN with variance σ^2 per dimension. Also

$$\alpha(c_t^{(i)}) = \sqrt{P_S}(1 - 2c_t^{(i)}) \tag{2}$$

$$\beta(d_t^{(i)}) = \sqrt{P_R}(1 - 2d_t^{(i)}) \tag{3}$$

holds for BPSK modulation, with P_S and P_R representing the received power for the symbols transmitted by the source and relay respectively.

The function of the soft demodulator is to take as input extrinsic log-likelihood ratios (LLRs) on the transmitted bits (LLRs in the absence of channel information); without loss of generality we consider the bit $c_t^{(i)}$,

$$L^E(c_t^{(i)}) = \ln\left(\frac{\Pr\left(c_t^{(i)} = 0\right)}{\Pr\left(c_t^{(i)} = 1\right)}\right) \tag{4}$$

and compute the new extrinsic LLRs (incremental LLRs expressing new information derived from the channel)

$$L^O(c_t^{(i)}) = L(c_t^{(i)}) - L^E(c_t^{(i)}) \tag{5}$$

where the *a posteriori* LLRs $L(c_t^{(i)})$ (incorporating channel information) are given by

$$L(c_t^{(i)}) = \ln\left(\frac{\Pr\left(c_t^{(i)}=0|e_t^{(i)}\right)}{\Pr\left(c_t^{(i)}=1|e_t^{(i)}\right)}\right) \tag{6}$$

$$= \ln\left(\frac{\sum_{\{c_t^{(i)},d_t^{(i)}\}:c_t^{(i)}=0} \Pr\left(c_t^{(i)},d_t^{(i)}|e_t^{(i)}\right)}{\sum_{\{c_t^{(i)},d_t^{(i)}\}:c_t^{(i)}=1} \Pr\left(c_t^{(i)},d_t^{(i)}|e_t^{(i)}\right)}\right) \tag{7}$$

$$= \ln\left(\frac{\sum_{\{c_t^{(i)},d_t^{(i)}\}:c_t^{(i)}=0} p\left(e_t^{(i)}|c_t^{(i)},d_t^{(i)}\right)\cdot \Pr(c_t^{(i)},d_t^{(i)})}{\sum_{\{c_t^{(i)},d_t^{(i)}\}:c_t^{(i)}=1} p\left(e_t^{(i)}|c_t^{(i)},d_t^{(i)}\right)\cdot \Pr(c_t^{(i)},d_t^{(i)})}\right) \tag{8}$$

We assume that the users' code bits are independent (a realistic assumption for independent users), so that

$$\Pr(c_t^{(i)},d_t^{(i)}) = \Pr(c_t^{(i)})\cdot\Pr(d_t^{(i)}) \tag{9}$$

Also, the probability density function (PDF) of the received channel value conditioned on the transmitted bits may be written as

$$p(e_t^{(i)}|c_t^{(i)},d_t^{(i)}) = \frac{1}{2\pi\sigma^2}\cdot\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)}-\phi_t^{(i)}\cdot\alpha(c_t^{(i)})-\psi_t^{(i)}\cdot\beta(d_t^{(i)})\right|^2\right) \tag{10}$$

Therefore, (5) may be re-written as (11), for which we use the shorthand $L^O(c_t^{(i)}) = f_1(e_t^{(i)}, L^E(d_t^{(i)}))$. In the same way, the new extrinsic LLR of the code bit $d_t^{(i)}$ may be obtained from (12), for which we use the shorthand $L^O(d_t^{(i)}) = f_2(e_t^{(i)}, L^E(c_t^{(i)}))$.

$$L^O(c_t^{(i)}) = \ln\left(\frac{\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)}-\sqrt{P_S}\phi_t^{(i)}-\sqrt{P_R}\psi_t^{(i)}\right|^2+L^E(d_t^{(i)})\right)+\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)}-\sqrt{P_S}\phi_t^{(i)}+\sqrt{P_R}\psi_t^{(i)}\right|^2\right)}{\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)}+\sqrt{P_S}\phi_t^{(i)}-\sqrt{P_R}\psi_t^{(i)}\right|^2+L^E(d_t^{(i)})\right)+\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)}+\sqrt{P_S}\phi_t^{(i)}+\sqrt{P_R}\psi_t^{(i)}\right|^2\right)}\right) \tag{11}$$

$$L^O(d_t^{(i)}) = \ln\left(\frac{\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)}-\sqrt{P_R}\psi_t^{(i)}-\sqrt{P_S}\phi_t^{(i)}\right|^2+L^E(c_t^{(i)})\right)+\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)}-\sqrt{P_R}\psi_t^{(i)}+\sqrt{P_S}\phi_t^{(i)}\right|^2\right)}{\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)}+\sqrt{P_R}\psi_t^{(i)}-\sqrt{P_S}\phi_t^{(i)}\right|^2+L^E(c_t^{(i)})\right)+\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)}+\sqrt{P_R}\psi_t^{(i)}+\sqrt{P_S}\phi_t^{(i)}\right|^2\right)}\right) \tag{12}$$

IV. Decoding Algorithm at Destination Node

A. Decoding Algorithm for the Cooperative Coding Scheme: Case of Fading Multiple Access Channel

Without loss of generality, we consider the decoding of codeword $\mathbf{a}_t$ at the end of half slot $2t+2$, for $t \in \{0, 1, \cdots, L-2\}$. The two codewords contained in the received signal stream $\mathbf{e}_{2t+2}$ are:

$$\begin{aligned} \mathbf{c}_{2t+2} &= \mathbf{a}_{t+1} \\ \mathbf{d}_{2t+2} &= \pi(\mathbf{c}_{2t+1}) \oplus \mathbf{c}_{2t} \\ &= \pi(\mathbf{b}_t) \oplus \mathbf{a}_t \end{aligned}$$

We assume that *a priori* LLRs on $\mathbf{c}_{2t+1}$ and $\mathbf{c}_{2t}$, denoted $\{L_1(c_{2t+1}^{(i)})\}$ and $\{L_2(c_{2t}^{(i)})\}$ respectively, are available from the previous decoding. In addition to decoding of $\mathbf{a}_t$, the decoder will produce *a posteriori* LLRs $\{L_1(c_{2t+2}^{(i)})\}$ and (updated) *a posteriori* LLRs $\{L_2(c_{2t+1}^{(i)})\}$; these will be used as *a priori* LLRs in the next decoding.

Next, we provide a concise description of the factor graph based decoding algorithm [18] at the destination decoder. The factor graph for the decoding is illustrated in Figure 6, where circles depict variable nodes and squares depict factor nodes. Extrinsic information is exchanged between three soft-input soft-output (SISO) decoder modules for the constituent codes (two codes $\mathcal{C}_a$ and one code $\mathcal{C}_b$), via the factor nodes $\{F_i\}$ which correspond to the network coding operation at the relay and the factor nodes $\{G_i\}$ which correspond to the soft demodulator for the MAC as given by (11) and (12). For simplicity, the graph is illustrated for the case $n = 3$, and where $\mathcal{C}_a$ and $\mathcal{C}_b$ are (trivial) LDPC codes. For convolutional constituent codes, the SISO modules execute BCJR algorithms. Note that in the convolutional case, the separation of the two (e.g. M-state) decoder SISO modules gives a complexity advantage over schemes which use a larger (e.g. M^2-state) decoder to decode the "nested" code generated at the relay (see e.g. [8], [9]). In the LDPC case this separation of SISO modules effects a more efficient message-passing schedule than does fully parallel decoding on the Tanner graph of the nested code (see e.g. [12]).

Next we introduce some notational conventions pertaining to the following algorithm description. In all cases, the letter λ is used to denote LLRs corresponding to messages passed on the factor graph (i.e. extrinsic LLRs). The interleaving "π" is interpreted as

$$\mathbf{x} = \pi(\mathbf{y}) \iff x^{(i)} = y^{(\pi(i))} \ \forall i \in \mathcal{I}$$

Some index sets are defined as follows. $\mathcal{J}_a = \{1, 2, \cdots m_a\}$; $\mathcal{J}_b = \{1, 2, \cdots m_b\}$; $\mathcal{N}_a(i) = \{j \in \mathcal{J}_a \ : \ H_a(j,i) = 1\}$; $\mathcal{N}_b(i) = \{j \in \mathcal{J}_b \ : \ H_b(j,i) = 1\}$; $\mathcal{M}_a(j) = \{i \in \mathcal{I} \ : \ H_a(j,i) = 1\}$; $\mathcal{M}_b(j) = \{i \in \mathcal{I} \ : \ H_b(j,i) = 1\}$. Also, $\boxplus$ denotes the (commutative and associative) "box-plus" operation [19], i.e.

$$\boxplus_{s \in \mathcal{S}} \lambda_s = \log \left(\frac{1 + \prod_{s \in \mathcal{S}} \tanh(\lambda_s/2)}{1 - \prod_{s \in \mathcal{S}} \tanh(\lambda_s/2)} \right)$$

N denotes the maximum number of decoding iterations.

Factor Graph Based Decoding Algorithm at Destination Node – Decoding of Codeword $\mathbf{a}_t$

Initialization:

- For $i \in \mathcal{I}$,

$$\lambda_{c_1}^{(i)} = L_2(c_{2t}^{(i)}) \tag{13}$$

$$\lambda_{c_2}^{(i)} = L_1(c_{2t+1}^{(i)}) \tag{14}$$

$$\lambda_8^{(i)} = 0 \tag{15}$$

$$\lambda_{15}^{(i)} = 0 \tag{16}$$

- For $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_2^{(i,j)} = 0 \tag{17}$$

- For $i \in \mathcal{I}$, $j \in \mathcal{N}_b(i)$

$$\lambda_4^{(i,j)} = 0 \tag{18}$$

- For $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_6^{(i,j)} = 0 \tag{19}$$

Main Loop: For $k = 1$ to N do

- For $i \in \mathcal{I}$,

$$\lambda_{14}^{(i)} = \lambda_{c_1}^{(i)} + \lambda_8^{(i)} \tag{20}$$

- Network coding constraints: for $i \in \mathcal{I}$,

$$\lambda_{10}^{(\pi(i))} = \lambda_{14}^{(i)} \boxplus \lambda_{15}^{(i)} \tag{21}$$

- For $i \in \mathcal{I}$,

$$\lambda_{9}^{(i)} = \lambda_{c_2}^{(i)} + \lambda_{10}^{(i)} \tag{22}$$

- SISO decoder $\mathcal{C}_b$: for $i \in \mathcal{I}$, $j \in \mathcal{N}_b(i)$

$$\lambda_{3}^{(i,j)} = \lambda_{9}^{(i)} + \sum_{l \in \mathcal{N}_b(i) \backslash \{j\}} \lambda_{4}^{(i,l)} \tag{23}$$

$$\lambda_{4}^{(i,j)} = \boxplus_{l \in \mathcal{M}_b(j) \backslash \{i\}} \lambda_{3}^{(l,j)} \tag{24}$$

For $i \in \mathcal{I}$,

$$\lambda_{12}^{(i)} = \sum_{j \in \mathcal{N}_b(i)} \lambda_{4}^{(i,j)} \tag{25}$$

- Obtain the *a posteriori* LLR (prior for decoding in next half slot)

$$L_2(c_{2t+1}^{(i)}) = \lambda_{9}^{(i)} + \lambda_{12}^{(i)} \tag{26}$$

- For $i \in \mathcal{I}$,

$$\lambda_{11}^{(i)} = \lambda_{c_2}^{(i)} + \lambda_{12}^{(i)} \tag{27}$$

$$\lambda_{16}^{(i)} = \lambda_{14}^{(i)} \boxplus \lambda_{11}^{(\pi(i))} \tag{28}$$

- Soft demodulation: for $i \in \mathcal{I}$,

$$\lambda_{17}^{(i)} = f_1(e_{2t+2}^{(i)}, \lambda_{16}^{(i)}) \tag{29}$$

- SISO decoder $\mathcal{C}_a$: for $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_{5}^{(i,j)} = \lambda_{17}^{(i)} + \sum_{l \in \mathcal{N}_a(i) \backslash \{j\}} \lambda_{6}^{(i,l)} \tag{30}$$

$$\lambda_{6}^{(i,j)} = \boxplus_{l \in \mathcal{M}_a(j) \backslash \{i\}} \lambda_{5}^{(l,j)} \tag{31}$$

For $i \in \mathcal{I}$,

$$\lambda_{18}^{(i)} = \sum_{j \in \mathcal{N}_a(i)} \lambda_{6}^{(i,j)} \tag{32}$$

- Obtain the *a posteriori* LLR (prior for decoding in next half slot)

$$L_1(c_{2t+2}^{(i)}) = \lambda_{17}^{(i)} + \lambda_{18}^{(i)} \tag{33}$$

- Soft demodulation: for $i \in \mathcal{I}$,

$$\lambda_{15}^{(i)} = f_2(e_{2t+2}^{(i)}, \lambda_{18}^{(i)}) \tag{34}$$

$$\lambda_{13}^{(i)} = \lambda_{15}^{(i)} \boxplus \lambda_{11}^{(\pi(i))} \tag{35}$$

- For $i \in \mathcal{I}$,

$$\lambda_7^{(i)} = \lambda_{c_1}^{(i)} + \lambda_{13}^{(i)} \tag{36}$$

- SISO decoder $\mathcal{C}_a$: for $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_1^{(i,j)} = \lambda_7^{(i)} + \sum_{l \in \mathcal{N}_a(i) \backslash \{j\}} \lambda_2^{(i,l)} \tag{37}$$

$$\lambda_2^{(i,j)} = \boxplus_{l \in \mathcal{M}_a(j) \backslash \{i\}} \lambda_1^{(l,j)} \tag{38}$$

For $i \in \mathcal{I}$,

$$\lambda_8^{(i)} = \sum_{j \in \mathcal{N}_a(i)} \lambda_2^{(i,j)} \tag{39}$$

- Calculate *a posteriori* LLRs for codeword $\mathbf{a}_t$:

$$L(a_t^{(i)}) = \lambda_7^{(i)} + \lambda_8^{(i)} \tag{40}$$

- Make decisions on the code bits

$$\hat{a}_t^{(i)} = \begin{cases} 0 & \text{if } L(a_t^{(i)}) \geq 0 \\ 1 & \text{if } L(a_t^{(i)}) < 0 \end{cases}$$

If $\hat{\mathbf{a}}_t \mathbf{H}_a^T = 0$ **then** ***break;***

Endfor

It may be seen that the decoding of $\mathbf{a}_t$ spans three transmission frames (half slots) $2t \rightarrow 2t+1 \rightarrow 2t+2$, during which the decoding follows the three-step evolution $L_1(c_{2t}^{(i)}) \rightarrow L_2(c_{2t}^{(i)}) \rightarrow \hat{a}_t^{(i)}$.

The preceding presentation was for the general case of decoding (half slots $t = 2, 3, \cdots, 2L-1\}$). Special cases are handled in a straightforward manner as follows. In half slot $t = 0$, a single LDPC

decoding of $\mathbf{a}_0$ based on LLRs derived from the received stream $\mathbf{e}_0$ produces $\{L_1(c_0^{(i)})\}$. In half slot $t = 1$, decoding proceeds as in the general case except that $\lambda_{14}^{(i)}$ is set to $+\infty$ for all $i \in \mathcal{I}$; $\{L_1(c_1^{(i)})\}$ and $\{L_2(c_0^{(i)})\}$ are produced (note that this decoder has the same structure as the decoder for the reference scheme of consecutive relaying in the fading MAC as shown in Figure 3). In half slot $t = 2L$, the decoder structure used is that for superposition decoding as described in [8], [9]; also only $\{L_2(c_{2L-1}^{(i)})\}$ are produced. In time slot $t = 2L+1$, a single LDPC decoding of $\mathbf{b}_{L-1}$ is performed based on the sum of the LLRs derived from the received stream $\mathbf{e}_{2L+1}$ and the LLRs $\{L_2(c_{2L-1}^{(i)})\}$ obtained from the previous decoding.

B. Decoding Algorithm for the Cooperative Coding Scheme: Case of Extra Available Frequency Band

In the case where an extra frequency band is used for relaying, the destination obtains separate coded transmissions from the sources and the relay. Here $L_1(c_t^{(i)})$ and $L_1(d_t^{(i)})$ denote the respective *a priori* LLRs derived from the corresponding received streams. Also, $L_2(\cdot)$ denotes the (updated) *a posteriori* LLRs which will be used as the *a priori* LLRs in the next decoding.

Concerning the factor graph based decoding algorithm at the destination node, in this case the "MAC Component" section of Figure 6 is no longer necessary and the factor nodes $\{G_i\}$ are simply pendant (degree-1) nodes in this case. These factor nodes now correspond not to the MAC operation but to the conditional PDFs of the channel values received from the relay.

We first consider without loss of generality the decoding of codeword $\mathbf{a}_t$ at the end of half slot $2t+2$, for $t \in \{1, 2, \cdots, L-1\}$. The algorithm described in Section IV-A is used for decoding, with the following modifications.

- Equation (19), (28)-(34) are deleted
- Equation (16) is replaced by

$$\lambda_{15}^{(i)} = L_1(d_{2t+2}^{(i)}) \tag{41}$$

For the special case of decoding the first codeword $\mathbf{a}_0$, we set $L_2(c_0^{(\pi(i))}) = L_1(c_0^{(\pi(i))}) + L_1(d_1^{(i)})$ for all $i \in \mathcal{I}$. Also, for the special case of decoding the final codeword $\mathbf{b}_{L-1}$, further modifications are made to the algorithm as follows:

- Equation (14), (18), (20)-(27), (35) are deleted
- Equation (36) is replaced by

$$\lambda_7^{(i)} = \lambda_{c_1}^{(i)} \tag{42}$$

In this case also, the decoding of $\mathbf{a}_t$ spans three transmission frames (half slots) $2t \rightarrow 2t+1 \rightarrow 2t+2$, resulting in the three-step decoding evolution $L_1(c_{2t}^{(i)}) \rightarrow L_2(c_{2t}^{(i)}) \rightarrow \hat{a}_t^{(i)}$.

V. Simulation Results and Discussion

A. A Lower Bound on the System Frame Error Rate

In this section a theoretical lower bound on the frame (codeword) error rate (FER) of the system is derived; the analysis follows the lines of [9].

In the case where an extra frequency band is used, let us consider the decoding of $\mathbf{a}_t$ and $\mathbf{b}_t$, transmitted in half slots $2t$ and $2t+1$ respectively. The codewords received by the destination from the sources in half slots $2t$ and $2t+1$ contain information only relating to $\mathbf{a}_t$ and $\mathbf{b}_t$. Due to latency constraints embodied in the relay operation, the message of $\mathbf{a}_t$ is contained in $\mathbf{d}_{2t+1}$ and $\mathbf{d}_{2t+2}$; in the same way, the message of $\mathbf{b}_t$ is also contained in $\mathbf{d}_{2t+2}$ and $\mathbf{d}_{2t+3}$. The received information which may be used in decoding $\mathbf{a}_t$ and $\mathbf{b}_t$ is contained in five received codewords at the destination – transmissions from sources to destination in half slots $2t$ and $2t+1$, and transmissions from relay to destination in half slots $2t+1$, $2t+2$ and $2t+3$. Recall that the codeword length is n, and let $r = k/n$ denote the code rate of each encoder; also let $C(\gamma)$ denote the capacity of a binary-input point-to-point link with temporal SNR γ. Then the number of information bits at the destination which may be used for joint decoding of $\mathbf{a}_t$ and $\mathbf{b}_t$ is not greater than $n[C(\gamma_{AD})+C(\gamma_{BD})+C(\gamma_{RD1})+C(\gamma_{RD2})+C(\gamma_{RD3})]$, where γ_{RD1}, γ_{RD2} and γ_{RD3} (the relay-destination SNRs in consecutive half slots) are independent and identically distributed random variables, and γ_{AD} and γ_{BD} (the source-destination SNRs) are independent and identically distributed random variables if the two source-to-destination link average SNRs are equal. When this value is lower than $2k$, there is no possibility of jointly decoding both packets, and an outage event is inevitable. Thus, a lower bound for the FER is given by

$$\text{FER} \geq \frac{1}{2} \cdot \Pr[C(\gamma_{AD}) + C(\gamma_{BD}) + C(\gamma_{RD1}) + C(\gamma_{RD2}) + C(\gamma_{RD3}) < 2r] \tag{43}$$

In the case where relaying is in the MAC, the information of each relay-transmitted signal is already contained in the source-transmitted signal of previous half slots. Therefore, assuming correct decoding in previous half slots, the relay transmitted signal could be perfectly cancelled from the destination received signal, and hence the receiver will "see" an interference-free fading channel for the source-transmitted signal. Therefore, the theoretical lower bound on the FER for the case of additional frequency band applies

also to the case of MAC relaying. The Monte Carlo simulated lower bound is included in the simulation results of Section V-B.

B. Simulation Results

The codes used for simulations are randomly generated rate $1/2$ regular LDPC codes of block length $n = 1200$ with column weight 3 and no 4-cycles in the Tanner graph. In simulations we choose $\mathbf{H}_a = \mathbf{H}_b$, and assume BPSK modulation for all systems. A random interleaver π is used to avoid 8-cycle multiplicity in the Tanner graph. We consider a quasi-static Rayleigh fading channel, for which the fading coefficients are constant within each half slot (one codeword) and change independently from one half slot to the next. We assume equal average signal-to-noise ratio (SNR) on the two source-destination links and the relay-destination link, and we assume that the destination has perfect knowledge of the channel fading coefficients and noise variances. As for the two source-relay links, which play a key role in the performance of the system since poor link quality may lead to catastrophic error propagation at the destination decoder, the simulation setup is that the source-relay links for the comparative performance of the proposed scheme with other reference systems are both ideally error-free. We also provide a bundle of curves indicating the degration of the overall performance as the signal-to-noise ratios of the source-relay links worsen.

Fair comparison of the proposed cooperative schemes with reference schemes is based on the constraint that in simulations, each scheme uses the same codes $\mathcal{C}_a$ and $\mathcal{C}_b$, and the same total energy E for transmission of the $2L$ source messages.

1) Case of Fading Multiple Access Channel: We provide a comparison of the proposed cooperative scheme with two reference schemes. The first is consecutive relaying in the fading MAC as shown in Figure 3. The second is a simple TDMA transmission scheme without cooperation or relaying.

Simulated performance results for bit error rate (BER) and frame error rate (FER) are shown in Figures 7 and 8 respectively. The curve corresponding to the proposed cooperative scheme exhibits the steepest slope (highest diversity gain) of the three, outperforming the others in the SNR region of interest. The scheme attains approximately an order of magnitude decrease in both BER and FER over consecutive relaying in the fading MAC at an E_b/N_0 of 10 dB. In Figure 9, a bundle of curves corresponding to various source-relay link scenarios are shown. It may be seen that the overall performance of the cooperative scheme is improved as the SNR of the source-relay links increases. When the signal-to-noise ratio of the source-relay links is 50dB, at which the outage probabilities are 10^{-5}, the overall performance approaches that of the ideal case.

2) Case of Extra Frequency Band: The first reference scheme is the consecutive relaying scheme of Figure 3. Here a single LDPC decoding per codeword is sufficient for reception, where the LLRs for decoder initialization are found by adding the LLRs corresponding to the broadcast and relayed versions of the pertinent codeword. The second reference scheme is the simple TDMA/FDMA code superposition relaying scheme of Figures 4 and 5. For this scheme, joint decoding of $\mathbf{a}_t$ and $\mathbf{b}_t$ is performed using a decoding algorithm similar to that of Section IV-B. The decoder also uses message passing between the two constituent decoders via network coding constraints in a straightforward manner.

Simulated performance results in terms of BER and FER are shown in Figures 10 and 11 respectively. The curve corresponding to the proposed cooperative coding scheme again exhibits an increased diversity gain with respect to the reference schemes in the SNR region of interest. The proposed scheme attains approximately an order of magnitude decrease in both BER and FER over the other schemes at an E_b/N_0 of 8 dB. In Figure 12, a bundle of curves corresponding to various source-relay link scenarios are shown. Again it is seen that the overall performance of the cooperative scheme improves as the SNR of the source-relay links increases. When the signal-to-noise ratio of the source-relay links is 50dB, at which the outage probabilities are 10^{-5}, the overall performance approaches that of the ideal case.

VI. Conclusion

In this paper, we have proposed a simple but effective cooperative coding scheme based on the shared-relay cooperative model. The scheme operates in two possible modes, depending on whether an extra frequency band is available for relaying. In both cases a corresponding sliding-window factor graph based decoding algorithm is used at the destination node to extract available spatial diversity gains. Computationally efficiency of the decoding algorithm is achieved in both cases through separation of the relevant SISO decoder modules and efficient connectivity via factor nodes corresponding to code and signal superposition operations. Simulation results demonstrate that the proposed schemes outperform competitive reference schemes.

REFERENCES

[1] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-Part I: System description," *Communications, IEEE Transactions on*, vol. 51, pp. 1927 – 1938, 2003.

[2] ——, "User cooperation diversity-Part II: Implementation aspects and performance analysis," *Communications, IEEE Transactions on*, vol. 51, pp. 1939 – 1948, 2003.

[3] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine*, pp. 74–80, Octorber 2004.

[4] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Transactions on communications*, vol. 52, pp. 273–277, 9-11 Sept. 2004.

[5] T. Hunter and A. Nosratinia, "Cooperation diversity through coding," in *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, 2002, p. 220.

[6] ——, "Diversity through coded cooperation," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 2, pp. 283–289, Feb. 2006.

[7] M. Janani, A. Hedayat, T. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: space-time transmission and iterative decoding," *Signal Processing, IEEE Transactions on*, vol. 52, no. 2, pp. 362–371, Feb. 2004.

[8] L. Xiao, T. Fuja, J. Kliewer, and D. Costello, "Cooperative diversity based on code superposition," *IEEE International Symposium on Information Theory (ISIT)*, 2006.

[9] L. Xiao, T. E. Fuja, J. Kliewer, and D. J. Costello, "A network coding approach to cooperative diversity," *IEEE Trans. Information Theory*, vol. 53, no. 10, pp. 3714–3722, Oct. 2007.

[10] L. Xiao, T. Fuja, J. Kliewer, and D. Costello, "Algebraic superposition of LDGM codes for cooperative diversity," *IEEE International Symposium on Information Theory (ISIT)*, 2007.

[11] C. Hausl and P. Dupraz, "Joint network-channel coding for the multiple-access relay channel," in *Sensor and Ad Hoc Communications and Networks, 2006. SECON '06. 2006 3rd Annual IEEE Communications Society on*, vol. 3, 28-28 Sept. 2006, pp. 817–822.

[12] C. Hausl, F. Schreckenbach, I. Oikonomidis, and G. Bauch, "Iterative network and channel decoding on a tanner graph," *43rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, USA*, 2005.

[13] S. Zhang, Y. Zhu, S.-C. Liew, and K. B. Letaief, "Joint design of network coding and channel decoding for wireless networks," in *Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE*, 11-15 March 2007, pp. 779–784.

[14] Y. Chen, S. Kishore, and J. Li, "Wireless diversity through network coding," *Proceeding of IEEE Wireless Communications and Networking Conference (WCNC), Las Vegas, NV*, 2006.

[15] X. Bao and J. Li, ""matching code-on-graph with network-on-graph: Adaptive network coding for wireless relay networks," *Proceeding of 43rd Annual Allerton Conference on Communication, Control and Computing (Allerton) , Champaign, IL,*, Sept. 2005.

[16] ——, "A unified channel-network coding treatment for wireless ad-hoc networks," *Proceeding of IEEE International Symposium on Information Theory (ISIT)*, 2006.

[17] M. A. Khojastepour, N. Ahmed, and B. Aazhang, "Code design for the relay channel and factor graph decoding," *Thirty-Eighth Asilomar Conference on Signal, Systems and Computers*, vol. 2, pp. 2000–2004, 2004.

[18] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Information Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[19] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *Information Theory, IEEE Transactions on*, vol. 42, no. 2, pp. 429–445, March 1996.

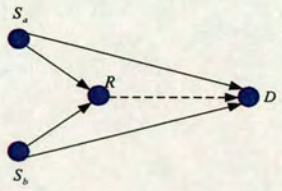

Fig. 1. Four-node communications network. Sources S_a and S_b share a common relay R as well as having direct links to the destination D.

Time Slot	0		1		...	t		t+1		...	L-1		L	
Half Slot	0	1	2	3	...	2t	2t+1	2t+2	2t+3	...	2L-2	2L-1	2L	2L+1
Source Transmits	$\mathbf{c}_0=$ $\mathbf{a}_0$	$\mathbf{c}_1=$ $\mathbf{b}_0$	$\mathbf{c}_2=$ $\mathbf{a}_1$	$\mathbf{c}_3=$ $\mathbf{b}_1$	...	$\mathbf{c}_{2t}=$ $\mathbf{a}_t$	$\mathbf{c}_{2t+1}=$ $\mathbf{b}_t$	$\mathbf{c}_{2t+2}=$ $\mathbf{a}_{t+1}$	$\mathbf{c}_{2t+3}=$ $\mathbf{b}_{t+1}$	...	$\mathbf{c}_{2L-2}=$ $\mathbf{a}_{L-1}$	$\mathbf{c}_{2L-1}=$ $\mathbf{b}_{L-1}$		
Relay Transmits		$\mathbf{d}_1=$ $\pi(\mathbf{a}_0)$	$\mathbf{d}_2=$ $\pi(\mathbf{b}_0)$ $\oplus\mathbf{a}_0$	$\mathbf{d}_3=$ $\pi(\mathbf{a}_1)$ $\oplus\mathbf{b}_0$	...	$\mathbf{d}_{2t}=$ $\pi(\mathbf{b}_{t-1})$ $\oplus\mathbf{a}_{t-1}$	$\mathbf{d}_{2t+1}=$ $\pi(\mathbf{a}_t)$ $\oplus\mathbf{b}_{t-1}$	$\mathbf{d}_{2t+2}=$ $\pi(\mathbf{b}_t)$ $\oplus\mathbf{a}_t$	$\mathbf{d}_{2t+3}=$ $\pi(\mathbf{a}_{t+1})$ $\oplus\mathbf{b}_t$	...	$\mathbf{d}_{2L-2}=$ $\pi(\mathbf{b}_{L-2})$ $\oplus\mathbf{a}_{L-2}$	$\mathbf{d}_{2L-1}=$ $\pi(\mathbf{a}_{L-1})$ $\oplus\mathbf{b}_{L-2}$	$\mathbf{d}_{2L}=$ $\pi(\mathbf{b}_{L-1})$ $\oplus\mathbf{a}_{L-1}$	$\mathbf{d}_{2L+1}=$ $\mathbf{b}_{L-1}$
Destination Receives	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	...	$\mathbf{e}_{2t}$	$\mathbf{e}_{2t+1}$	$\mathbf{e}_{2t+2}$	$\mathbf{e}_{2t+3}$	...	$\mathbf{e}_{2L-2}$	$\mathbf{e}_{2L-1}$	$\mathbf{e}_{2L}$	$\mathbf{e}_{2L+1}$
Destination Decodes			$\mathbf{a}_0$	$\mathbf{b}_0$	...	$\mathbf{a}_{t-1}$	$\mathbf{b}_{t-1}$	$\mathbf{a}_t$	$\mathbf{b}_t$	...	$\mathbf{a}_{L-2}$	$\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$

Fig. 2. Transmission schedule of proposed cooperative coding scheme for the case of multiple access channel. For the case of extra available frequency band, the destination no longer receives a superposed signal $\mathbf{e}_t$ but signals transmitted from sources and relay separately.

Time Slot	0		1		...	t		t+1		...	L-1		
Half Slot	0	1	2	3	...	2t	2t+1	2t+2	2t+3	...	2L-2	2L-1	2L
Source Transmits	$\mathbf{c}_0=\mathbf{a}_0$	$\mathbf{c}_1=\mathbf{b}_0$	$\mathbf{c}_2=\mathbf{a}_1$	$\mathbf{c}_3=\mathbf{b}_1$	...	$\mathbf{c}_{2t}=\mathbf{a}_t$	$\mathbf{c}_{2t+1}=\mathbf{b}_t$	$\mathbf{c}_{2t+2}=\mathbf{a}_{t+1}$	$\mathbf{c}_{2t+3}=\mathbf{b}_{t+1}$	...	$\mathbf{c}_{2L-2}=\mathbf{a}_{L-1}$	$\mathbf{c}_{2L-1}=\mathbf{b}_{L-1}$	
Relay Transmits		$\mathbf{d}_1=\pi(\mathbf{a}_0)$	$\mathbf{d}_2=\pi(\mathbf{b}_0)$	$\mathbf{d}_3=\pi(\mathbf{a}_1)$	...	$\mathbf{d}_{2t}=\pi(\mathbf{b}_{t-1})$	$\mathbf{d}_{2t+1}=\pi(\mathbf{a}_t)$	$\mathbf{d}_{2t+2}=\pi(\mathbf{b}_t)$	$\mathbf{d}_{2t+3}=\pi(\mathbf{a}_{t+1})$	...	$\mathbf{d}_{2L-2}=\pi(\mathbf{b}_{L-2})$	$\mathbf{d}_{2L-1}=\pi(\mathbf{a}_{L-1})$	$\mathbf{d}_{2L}=\pi(\mathbf{b}_{L-1})$
Destination Receives	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	...	$\mathbf{e}_{2t}$	$\mathbf{e}_{2t+1}$	$\mathbf{e}_{2t+2}$	$\mathbf{e}_{2t+3}$	...	$\mathbf{e}_{2L-2}$	$\mathbf{e}_{2L-1}$	$\mathbf{e}_{2L}$
Destination Decodes		$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_1$	...	$\mathbf{b}_{t-1}$	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_{t+1}$	...	$\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$

Fig. 3. Transmission schedule of consecutive relaying scheme for the case of multiple access channel. For the case of extra available frequency band, the destination no longer receives a superposed signal $\mathbf{e}_t$ but signals transmitted from sources and relay separately.

Time Slot										
0	1	2	...	3t	3t+1	3t+2	...	3L-3	3L-2	3L-1
$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_0\oplus\mathbf{b}_0$	...	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_t\oplus\mathbf{b}_t$	...	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$	$\mathbf{a}_{L-1}\oplus\mathbf{b}_{L-1}$

Fig. 4. Transmission schedule for the TDMA relaying scheme.

	0	1	...	t	t+1	...	L-1	
Frequency Band f1	$\mathbf{a}_0$	$\mathbf{a}_1$	...	$\mathbf{a}_t$	$\mathbf{a}_{t+1}$	...	$\mathbf{a}_{L-1}$	
Frequency Band f2	$\mathbf{b}_0$	$\mathbf{b}_1$	...	$\mathbf{b}_t$	$\mathbf{b}_{t+1}$	...	$\mathbf{b}_{L-1}$	
Frequency Band f3		$\mathbf{a}_0\oplus$ $\mathbf{b}_0$	...	$\mathbf{a}_{t-1}\oplus$ $\mathbf{b}_{t-1}$	$\mathbf{a}_t\oplus$ $\mathbf{b}_t$	...	$\mathbf{a}_{L-2}\oplus$ $\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}\oplus$ $\mathbf{b}_{L-1}$

Fig. 5. Transmission schedule for the FDMA relaying scheme.

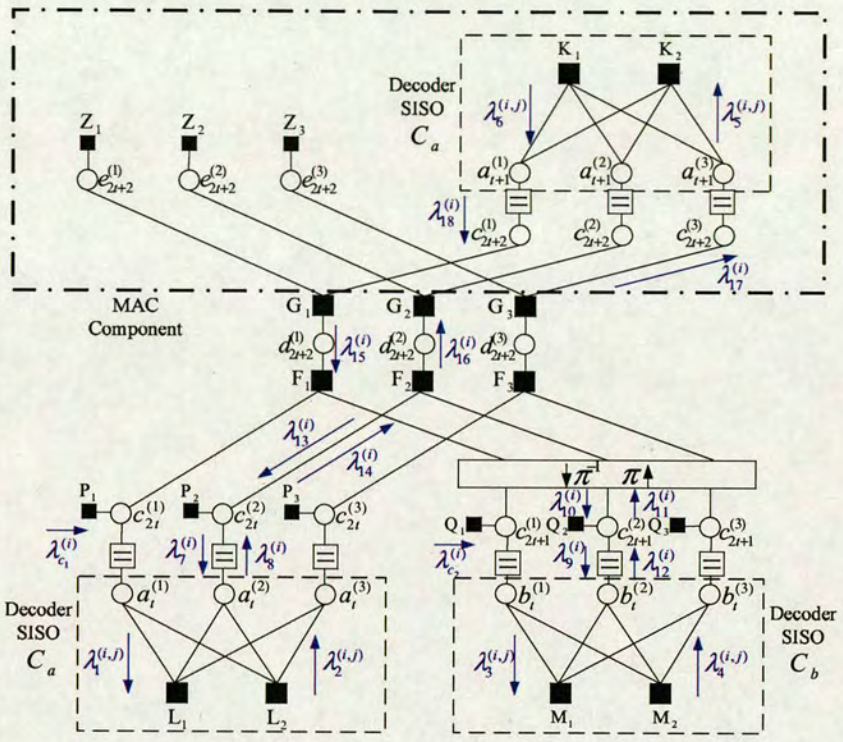

Fig. 6. Factor graph corresponding to the destination's decoding of codeword $\mathbf{a}_t$. For ease of presentation, the factor graph is illustrated for the case $n = 3$ and trivial codes $\mathcal{C}_a$, $\mathcal{C}_b$. For the case of the fading multiple access channel, the three decoder SISO modules exchange extrinsic information via the factor nodes $\{F_i\}$ which correspond to the network coding operation at the relay and the factor nodes $\{G_i\}$ which correspond to the soft demodulator for the MAC. For the case with an extra available frequency band, the "MAC component" subgraph is omitted (the factor nodes $\{G_i\}$ are pendant nodes in this case), and the message-passing schedule simply exchanges extrinsic information between the two decoder SISO modules for the constituent codes $\mathcal{C}_a$ and $\mathcal{C}_b$, via the factor nodes $\{F_i\}$.

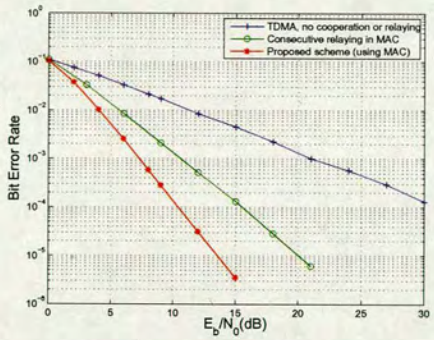

Fig. 7. Comparative BER performance for the proposed cooperative scheme in the case of multiple access channel. The performance is shown with respect to two reference schemes: simple TDMA without cooperation or relaying; and consecutive relaying in the MAC.

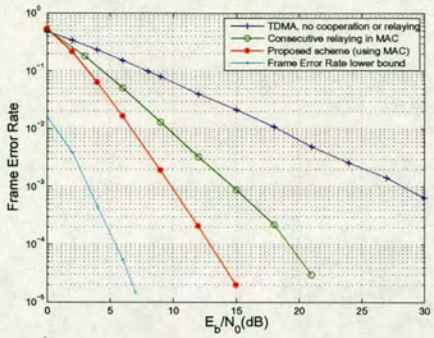

Fig. 8. Comparative FER performance for the proposed cooperative scheme in the case of multiple access channel. The performance is shown with respect to two reference schemes: simple TDMA without cooperation or relaying; and consecutive relaying in the MAC. Also plotted is the theoretical lower bound on the FER given by (43).

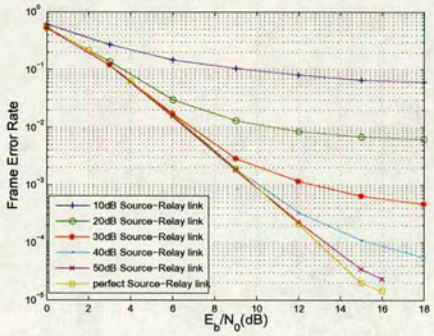

Fig. 9. Comparative FER performance for the proposed cooperative scheme in the case of multiple access channel. The performances are shown with various scenarios of the source-relay links: the signal-to-noise ratio ranges from 10dB to 50dB.

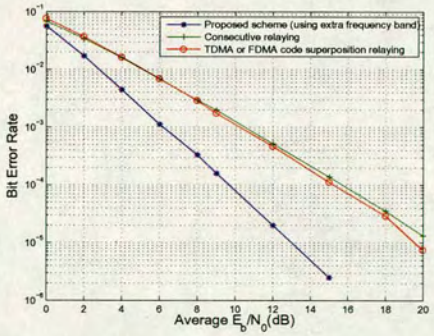

Fig. 10. Comparative BER performance for the proposed cooperative scheme in the case of extra frequency band. The performance is shown with respect to two reference schemes: simple TDMA/FDMA code superposition relaying; and consecutive relaying using the extra frequency band.

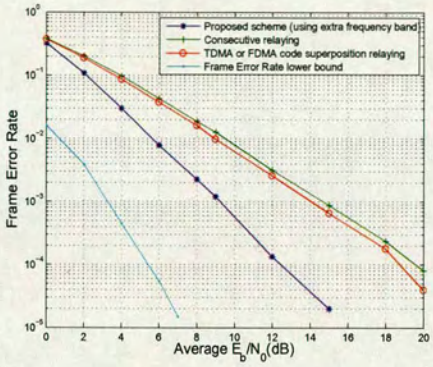

Fig. 11. Comparative FER performance for the proposed cooperative scheme in the case of extra frequency band. The performance is shown with respect to two reference schemes: simple TDMA/FDMA code superposition relaying; and consecutive relaying using the extra frequency band. Also plotted is the theoretical lower bound on the FER given by (43).

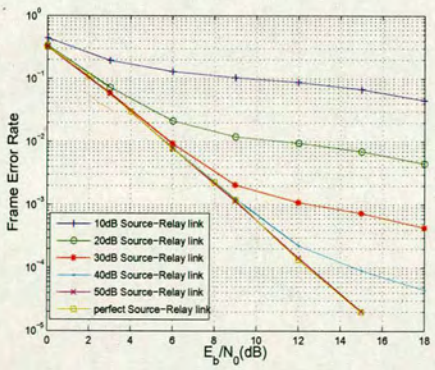

Fig. 12. Comparative FER performance for the proposed cooperative scheme in the case of extra frequency band. The performances are shown with various scenarios of the source-relay links: the signal-to-noise ratio ranges from 10dB to 50dB.

Practical Successive Interference Cancellation in the Binary-Input Gaussian Multiple-Access Channel

Xiaoyan Xu and Norbert Goertz
Institute for Digital Communications
Joint Research Institute for Signal & Image Processing
The University of Edinburgh, Scotland
Email: {X.Xu, Norbert.Goertz}@ed.ac.uk

***Abstract*—In the multiple access channel, successive interference cancellation (SIC) can be used to achieve the boundary points of the capacity region. In this paper, we investigate the practical potential of SIC by employing a set of practical moderate-blocksize Low-Density Parity-Check codes in the Multiple Access additive white Gaussian noise channel; in particular we consider binary modulation. The theoretically achievable points in the capacity region are compared with the practically achievable operating points at a low bit-error rate. It is shown that SIC makes available a higher transmission rate compared with simple separate detection that treats the signal of the other user as noise. This statement is also true for the use of binary modulation and a practical efficient implementation with channel codes of moderate blocksize.**

I. Introduction

In the multiple access channel (MAC), all users share the common channel resources to transmit information to a destination. Time-Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA) systems are based on the orthogonal division of the resources "time" and "frequency", with only one user occupying a frequency at a time. Contrary to that in CDMA each user is assigned a distinct signature sequence that allows for orthogonal signal transmission with more than one user occupying a frequency at the same time [1]; the price to pay is increased bandwidth due to the "signature sequence" that has to be transmitted to send one information bit. Still, theoretical analysis shows that CDMA is a more efficient multiple access method [2] than TDMA and FDMA, mainly due to its greater flexibility.

In principle, the CDMA idea can also be used for channel-coded sequences without any explicit "spreading" code and in this case it is known from Information Theory [3] that Successive Interference Cancellation (SIC) is a technique that can achieve the upper limits of the capacity region of the Gaussian Multiple Access Channel. Signal superposition together with SIC theoretically guarantee that one of the users can transmit at the capacity-approaching rate, as in the single-user AWGN channel, while the other user can still realize reliable transmission at a non-zero rate that is determined by the Gaussian MAC capacity region: essentially, the second user's signal is treated as "noise" by the first user, i.e., the first user suffers from a reduced rate due to interference. When the first user's rate is sufficiently low, the signal can be reliably decoded at the receiving end, and the re-encoded signal can be subtracted (cancelled) from the total received signal, leaving only the Gaussian receiver noise. This means the receiver "sees" a Gaussian channel without any interference for the second user.

Although, at a first glance, this theoretical concept seems also attractive for a practical application there are some potential drawbacks. First, the theory assumes that Gaussian signal alphabets are used, which is not possible in practice. Hence, we go to the opposite extreme and investigate the performance of SIC schemes when a binary modulation signal aphabet is used. This is motivated by the idea to have very simple transmitters such as wireless network nodes communicating to a receiving "base station" where we allow for some higher complexity to perform more advanced signal processing such as successive interference cancellation. Second, as it is impossible to transmit strictly reliable in practice, the first user's transmit signal can not always be reconstructed correctly at the reciever, which means the subtraction of the erroneous signal would lead to propagation of the first user's decoding errors into the decoding process for the second user. Hence, it is unclear what the performance benefits of SIC will be in practice: this is exactly the topic of the presented work in which we investigate soft multiple user demodulators in which the interference cancellation in carried out in a "soft" sense using bit-reliability information.

In our work we use low-density parity-check (LDPC) codes which are very powerful error correcting codes. Well-designed and with very large blocksize these codes can achieve very low bit-error-rates (BER) [4]. In practical applications, however, delay constraints usually put an upper limit on the coding blocksize. Therefore, we use pratical moderate-blocksize LDPC codes and investigate their performance in the Gaussian MAC with soft successive interference cancellation.

II. System description and Channel Capacities

A. System Model (Two Users)

In a Multiple Access Channel (MAC), also called uplink multiuser channel, all users send signals to one receiver. The discrete-time model for the two-user AWGN MAC is shown in Figure 1. The receive signal is given by

$$y[i] = x_a[i] + x_b[i] + n[i], \tag{1}$$

with i the time index, $n[i] \sim N(0, \sigma_w^2)$ the i.i.d. Gaussian receiver noise with power spectral density (PSD) $N_0/2$, the variance $\sigma_w^2 = N_0 \cdot B$ with B the base-band system bandwidth. For sake of simplicity we assume real baseband signals; an extension to complex baseband signals is straighforward. The received user signals are $x_a[i]$ and $x_b[i]$. We integrate any power scaling and any fixed path-loss into the average power constraints such that $\mathrm{E}\{(x_{a,b}[i])^2\} = P_{a,b}$.

B. Capacity of the single-user discrete-time AWGN Channel

The capacity for the ideal additive white Gaussian noise (AWGN) channel with average input power constraint P is achieved with a Gaussian distribution of the input symbols. The capacity equals

$$C = \frac{1}{2}\log_2(1 + \frac{P}{\sigma_w^2}) \text{ bits per transmission.} \tag{2}$$

For the band-limited AWGN channel with bandwidth B, the capacity of the channel can be rewritten as

$$C = B\log_2(1 + \frac{P}{N_0 B}) \text{ bits per second.} \tag{3}$$

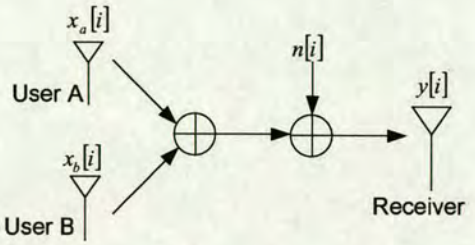

Fig. 1. Multiple access channel.

The channel capacity for the binary-input discrete-time AWGN channel with equiprobable input bits $\xi \in \{0,1\}$ and the channel output signal y is

$$C = \sum_{\xi\in\{0,1\}} \int_{-\infty}^{\infty} \frac{p(y|\xi)}{2} \log_2 \frac{2 \cdot p(y|\xi)}{p(y|\xi=1) + p(y|\xi=0)} dy \quad (4)$$

with C in bits per transmission (channel use). For binary phase shift keying modulation (with coherent detection) we map the input bits $\xi \in \{0,1\}$ to the modulation signal "constellation" $\{+1,-1\}$. However, as we have integrated any power scaling and path loss into the power constraints P_a, P_b we have for the received modulation signals

$$x_{a,b}(\xi) = \sqrt{P_{a,b}} \cdot (1 - 2 \cdot \xi) \quad (5)$$

Hence, the probability density function (pdf) of the Gaussain noise equals

$$p(y|\xi) \doteq \frac{1}{\sqrt{2\pi\sigma_w}} \exp(-\frac{1}{2\sigma_w^2}(y - x_{a,b}(\xi))^2) \,. \quad (6)$$

C. Capacity Region for the two-user AWGN MAC

The two-user capacity region is given by all rate vectors (R_a, R_b) satisfying the following constrains [3]:

$$\begin{aligned} R_a &\leq B \log_2 \left(1 + \frac{P_a}{N_0 B}\right) \text{ bits / second} \\ R_b &\leq B \log_2 \left(1 + \frac{P_b}{N_0 B}\right) \text{ bits / second} \\ R_a + R_b &\leq B \log_2 \left(1 + \frac{P_a + P_b}{N_0 B}\right) \text{ bits / second} \end{aligned} \quad (7)$$

The interpretation for the capacity region is straightforward: The first two lines of (7) say that the transmission rates for the individual users can not exceed the Gaussian capacity limits in the single-user AWGN channel. The last line of (7) puts an upper limit on the sum rate of the users: this sum can not exceed the rate limit that we obtain on a Gaussian channel with the sum of the powers of both users.

The capacity region for the two-user Gaussian MAC is shown in Figure 2 [5], where C_k and C_k^* are given by

$$C_k = B \log_2 \left(1 + \frac{P_k}{N_0 B}\right), \; k = a, b, \quad (8)$$

$$C_a^* = B \log_2 \left(1 + \frac{P_a}{N_0 B + P_b}\right), \quad (9)$$

$$C_b^* = B \log_2 \left(1 + \frac{P_b}{N_0 B + P_a}\right). \quad (10)$$

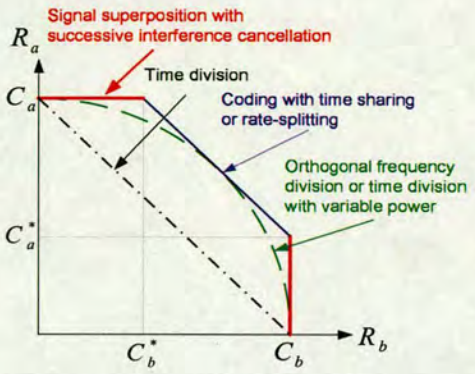

Fig. 2. Capacity region of the two-user multiple access channel [5].

The points $(R_b, R_a) = (0, C_a)$ and $(C_b, 0)$ are achieved when one transmitter sends its data at its maximum rate in the single-user AWGN channel while the other user is silent. The rate pairs on the straight line between $(0, C_a)$ and $(C_b, 0)$ are achieved by a time-division strategy between two transmitters operating at their maximum rates with fixed power. The rate pairs on the line $(0, C_a) - (C_b^*, C_a)$ and the rate pairs on the line $(C_b, C_a^*) - (C_b, 0)$ are achieved by using signal superposition together with SIC. The rate pairs on the straight line connecting (C_b^*, C_a) and (C_b, C_a^*) are achieved by time-sharing between the two points or a rate-splitting technique [6]. For the latter a user splits its data stream into multiple substreams and encodes these substreams as if they originated from different virtual users.

For comparison, Figure 2 also contains the rate region resulting from orthogonal signalling (TDMA, FDMA) without SIC: the rate region achieved by those schemes touches the boundary of the capacity region in exactly one point, which corresponds to the rate at which we obtain the maximum sum rate. In all other points orthogonal signaling is strictly suboptimal. The problem is particularly significant, when the users have strongly different receive power: in this case the weaker user will hardly get any rate at all in the capacity-achieving point [7, pp. 232–234].

III. Application of Low-Density Parity-Check Codes in Successive Interference Cancellation

A. Codes

LDPC codes were invented by Gallager [8], [9] in the early 1960s and largely ignored until they were re-discovered [10] almost 40 years later. LDPC codes are linear block codes with a particular structure for the parity check matrix $\mathbf{H}$ in which the fraction of nonzero entries is small. This allows graph-based decoders (e.g. Sum Product Algorithm SPA [11]) to be applied efficiently with remarkable performance. A key feature of the Sum Product Decoding algorithm is that soft reliability information from the channel output can be exploited for decoding.

We don't deal with details of LDPC codes and the SPA decoder as such in this paper (full details can be found, e.g., in [11]). We rather provide a scheme with which log-likelihhod ratios (L-values) [12] can be obtained for the SIC scheme that we apply in the MAC: L-values are the key to a successful application of soft-in put channel decoding by the sum-product algorithm.

B. Signal Superposition and Successive Interference Cancellation

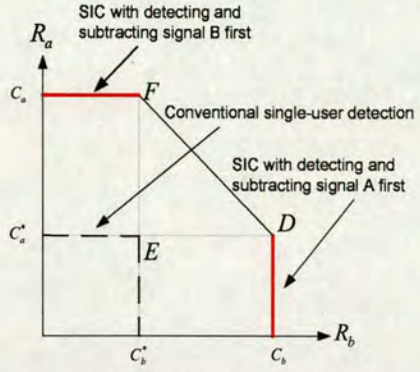

Fig. 3. Capacity boundary for SIC and conventional single-user detection in two-user multiple access channel.

Successive Interference Cancellation (SIC) is a multiuser detection technique which can be used to achieve the boundary of the capacity region of a Multiple-Access Channel [7]. In contrast to conventional single-user detection (e.g., [2]), in which other users' signals are treated as noise when decoding one user's signal, SIC is based on subtracting (cancelling) the already detected signals from the received signal before the detection of the next signal. The more disparate the users' powers are the higher the potential gain from SIC.

The point D in Figure 3 indicates that user B could transmit at capacity-approaching rate in the point-to-point transmission model, while user A can still transmit at the maximun rate of C_a^* without errors. If signal A is the first one to be decoded, signal B is treated as Gaussian noise to user A in the first signal detection stage.

In Figure 3, the line $C_b - D$ is achieved by signal superposition together with SIC when signal A is first detected and subtracted. If the order of the signal detection is reversed and signal B is first detected, by SIC the capacity boundary $C_a - F$ could be achieved. The capacity boundary encompassed by C_a^*, E and C_b^* is for signal superposition without SIC which is the conventional single-user detection method. The theoretical capacity region shows that signal superposition with SIC is superior to the conventional single-user detection method.

In practice, however, reliable communication with channel codes of limited block size is impossible in a strict sense. Therefore, instead of "cancelling by signal subtraction" (which would cause error propagation in case of incorrect decoding), we resort to a "soft" *a posteriori* probability (APP) demodulator, which can deal with reliability information for bits rather than with hard decisions only. This concept (in the given framework) is known from e.g. [13], [14]. In what follows we consider the special case of coherently detected binary phase-shift keying modulation (BPSK) and we state a formulation of a multi-user detector that lends itself to an efficient implementation by L-value algebra. We also compare the simulation results with the unconstrained Gaussian MAC capacity region.

C. Multiuser APP-Demodulator for BPSK Modulation

The basic idea is to extract the *a posteriori* probability for the transmitted bits $B_k[i]$ of all users $k = 1, 2, ..., K$ to be "one" or "zero" given the received channel vallue y (see Fig. 1 for the two-user case). The channel model for the general K-user case is given by

$$y[i] = \sum_{k=1}^{K} x_k[i] + n[i] \ , \tag{11}$$

where

$$x_k(b_k[i]) = \sqrt{P_k}(1 - 2b_k[i]) \tag{12}$$

and $b_k[i] \in \{0, 1\}$ is the bit sent by user k at time i; P_k is the receive power for the signal transmitted by user k. We now calculate the APP L-value [12] of the bit $b_k[i]$:

$$L(B_k[i]|y) = \log \frac{\Pr(B_k[i] = 0|y)}{\Pr(B_k[i] = 1|y)} \tag{13}$$

where $B_k[i]$ indicates the random variable of the bit and $b_k[i] \in \{0, 1\}$ its realisation. As we consider only one time-instant in the detector, we omit the time index i in what follows as long as there is no risk of confusion.

We expand

$$\Pr(B_k = 0|y) = \sum_{\forall \{b_1,...,b_K\}: b_k = 0}^{K} \Pr(B_1 = b_1, ..., B_K = b_K|y) \tag{14}$$

and use the Bayes-rule which immediately leads to

$$L(B_k|y) = \log \frac{\sum_{\forall \mathbf{b}: b_k = 0} p(y|\mathbf{b}) \cdot \Pr(\mathbf{b})}{\sum_{\forall \mathbf{b}: b_k = 1} p(y|\mathbf{b}) \cdot \Pr(\mathbf{b})} \tag{15}$$

with the bit vector realisations $\mathbf{b} \doteq \{b_1, b_2, ..., b_K\}$ (we use the same definition for the bit vector random "variable" $\mathbf{B}$). The notation $\forall \mathbf{b}: b_k = 0$ denotes all possible bit combinations of the other users, with the bit value of "zero" for the user k under consideration. For later use we further define $\mathbf{b} \backslash b_k \doteq \{b_1, b_2, ..., b_{k-1}, b_{k+1}, ..., b_K\}$. Moreover, we assume that the users' code bits are independent (rather realistic assumption for independent users), so that

$$\Pr(\mathbf{b}) = \prod_{k=1}^{K} \Pr(b_k) \ . \tag{16}$$

(above, we use the abbreviation $\Pr(B_k = b_k) = \Pr(b_k)$).

If we assume a Gaussian channel, we can write the probability density function (PDF) of the channel as follows:

$$p(y|\mathbf{b}) = \frac{1}{\sqrt{2\pi}\sigma_w} \exp\left(-\frac{1}{2\sigma_w^2} \left(y - \sum_{k=1}^{K} x_k(b_k) \right)^2 \right) \tag{17}$$

Up to this stage, the eqation are also valid for non-binary symbols, but now we start to exploit the special BSPK modulation scheme: we reformulate the exponent of the PDF in (17): with $x_0 \doteq -y$ and the abbreviation $x_k \doteq x_k(b_k), k = 1, 2, ...$ we obtain

$$\begin{aligned} \left(y - \sum_{k=1}^{K} x_k(b_k) \right)^2 &= \left(\sum_{k=0}^{K} x_k \right)^2 \\ &= \sum_{k=0}^{K} x_k^2 + 2 \sum_{k=0}^{K-1} x_k \sum_{l=k+1}^{K} x_l \\ &= y^2 + \sum_{k=1}^{K} x_k^2 - 2y \cdot \sum_{k=1}^{K} x_k + 2 \sum_{k=1}^{K-1} x_k \sum_{l=k+1}^{K} x_l \end{aligned} \tag{18}$$

As $x_k = \sqrt{P_k}(1-2b_k)$, with $b_k \in \{0,1\}$, we have $x_k^2 = P_k$, which does *not* depend on the choice of the bit b_k. Therefore, we can re-write (15) as follows:

$$L(B_k|y) = \log \frac{\Pr(B_k=0) \sum_{\forall \mathbf{b}: b_k=0} p(y|\mathbf{b}) \cdot \Pr(\mathbf{b}\backslash b_k)}{\Pr(B_k=1) \sum_{\forall \mathbf{b}: b_k=1} p(y|\mathbf{b}) \cdot \Pr(\mathbf{b}\backslash b_k)}$$
$$= L(B_k) + \log \frac{\sum_{\forall \mathbf{b}: b_k=0} e^{f(y,b_1,\ldots,b_K)/\sigma_w^2} \cdot \Pr(\mathbf{b}\backslash b_k)}{\sum_{\forall \mathbf{b}: b_k=1} e^{f(y,b_1,\ldots,b_K)/\sigma_w^2} \cdot \Pr(\mathbf{b}\backslash b_k)} \quad (19)$$

with

$$f(y,b_1,\ldots,b_K) = y \cdot \sum_{\kappa=1}^{K} x_\kappa - \sum_{\xi=1}^{K-1} x_\xi \sum_{l=\xi+1}^{K} x_l \quad (20)$$

Note that, as above, $x_\kappa = \sqrt{P_\kappa}(1-2\cdot b_\kappa)$ for BPSK modulation. We now separate out the bit b_k in (20), i.e., we isolate $x_k = x_k(b_k)$:

$$f(y,b_1,\ldots,b_K) = x_k\Big(y - \sum_{\substack{\xi=1\\ \xi\neq k}}^{K} x_\xi\Big) + y\sum_{\substack{l=1\\ l\neq k}}^{K} x_l - \sum_{\substack{\xi=1\\ \xi\neq k}}^{K-1} x_\xi \sum_{\substack{l=\xi+1\\ l\neq k}}^{K} x_l$$
$$= x_k \cdot y + \Big(y - x_k\Big) \sum_{\substack{\xi=1\\ \xi\neq k}}^{K} x_\xi - \sum_{\substack{\xi=1\\ \xi\neq k}}^{K-1} x_\xi \sum_{\substack{l=\xi+1\\ l\neq k}}^{K} x_l \quad (21)$$

When we use (21) in (19) with $x_k(b_k) = \sqrt{P_k}(1-2b_k)$ we find

$$L(B_k|y) = L(B_k) + \log \frac{\exp(y\cdot x_k(0)/\sigma_w^2)}{\exp(y\cdot x_k(1)/\sigma_w^2)} + L_E(B_k)$$
$$= L(B_k) + 2\frac{\sqrt{P_k}}{\sigma_w^2}\cdot y + L_E(B_k) \quad (22)$$

where

$$L_E(B_k) = \log \frac{\sum_{\forall \mathbf{b}\backslash b_k} e^{g_0(\mathbf{b}\backslash b_k)/\sigma_w^2} \cdot \Pr(\mathbf{b}\backslash b_k)}{\sum_{\forall \mathbf{b}\backslash b_k} e^{g_1(\mathbf{b}\backslash b_k)/\sigma_w^2} \cdot \Pr(\mathbf{b}\backslash b_k)} \quad (23)$$

with

$$g_\rho(\mathbf{b}\backslash b_k) = \Big(y - x_k(\rho)\Big) \sum_{\substack{\xi=1\\ \xi\neq k}}^{K} x_\xi - \sum_{\substack{\xi=1\\ \xi\neq k}}^{K-1} x_\xi \sum_{\substack{l=\xi+1\\ l\neq k}}^{K} x_l\,, \quad (24)$$

with $\rho \in \{0,1\}$ and $x_k = \sqrt{P_k}(1-2b_k)$. As long as we are using binary modulation, (23) can be evaluated for, e.g., up to $K = 10$ users without serious complexity problems. If more users are in the system and the received powers P_i for their signals are significantly different, we can approximate the sums in (24) by only considering a subset of the "strongest" users.

We can further develop (23) by using (16):

$$\Pr(\mathbf{b}\backslash b_k) = \prod_{\substack{l=1\\ l\neq k}}^{K} \Pr(b_l)\,. \quad (25)$$

We now express the probabilities $\Pr(b_l)$ of the bits of user l by its corresponding "a-priori" L-value

$$L(B_l) \doteq \log \frac{\Pr(B_l=0)}{\Pr(B_l=1)}\,. \quad (26)$$

By inversion of (26) we obtain

$$\Pr(B_l = b) = \frac{e^{-L(B_l)\cdot b}}{1+e^{-L(B_l)}}, \quad b \in \{0,1\}\,. \quad (27)$$

We use (25) to find

$$\Pr(\mathbf{b}\backslash b_k) = A \cdot \exp\left(-\sum_{\substack{l=1\\ l\neq k}}^{K} L(B_l)\cdot b_l\right) \quad (28)$$

with the constant A that will cancel out in (23). We combine (28) and (24) coveniently for use in (23) and obtain

$$h_\rho(\mathbf{b}\backslash b_k) = \frac{1}{\sigma_w^2} g_\rho(\mathbf{b}\backslash b_k) - \sum_{\substack{l=1\\ l\neq k}}^{K} L(B_l)\cdot b_l$$
$$= \frac{1}{\sigma_w^2}\left(\Big(y - x_k(\rho)\Big)\sum_{\substack{\xi=1\\ \xi\neq k}}^{K} x_\xi - \sum_{\substack{\xi=1\\ \xi\neq k}}^{K-1} x_\xi \sum_{\substack{l=\xi+1\\ l\neq k}}^{K} x_l\right) - \sum_{\substack{l=1\\ l\neq k}}^{K} L(B_l)\cdot b_l \quad (29)$$

with

$$L_E(B_k) = \log \frac{\sum_{\forall \mathbf{b}\backslash b_k} \exp(h_0(\mathbf{b}\backslash b_k))}{\sum_{\forall \mathbf{b}\backslash b_k} \exp(h_1(\mathbf{b}\backslash b_k))} \quad (30)$$

As only the factor $x_k(\rho)$ is different in the numerator and denominator of (30), most of $h_\rho(\mathbf{b}\backslash b_k)$ needs only to be computed once. Moreover, the probability-weighting by the L-values for the bits of other users appears as a simple sum of L-values (rightmost term in (30), which are scaled by the bit-values $b_l \in \{0,1\}$: hence, only the "one"-bits in the bit-vector sum-"index" $\mathbf{b}\backslash b_k$ need to be considered. The double-sum in (29) is most conveniently computed by starting with $\xi = K-1$: the result of the inner sum can then be recursively used to compute the next result. As long as the channel-SNRs don't change and with a limited number of users in the system, all sums in (29), apart from the last term, can in principle be pre-computed and stored as they don't depend on y or on the L-values of the other users' bits. Moreover, we can ignore terms in the sums that contribute very little to the total result (which stem typically from users with very low channel SNR); this helps to cope with a larger number of users in the system. In the simulations below we have used (22), (29) and (30) for the special case of $K = 2$ users.

IV. Simulation Results and Analysis

The LDPC codes used here are randomly generated regular LDPC codes[1] of blocksize 1200 and column weight 3. For the theoretical capacity calculation, (2) is for Gaussian input symbols which could not be implemented in the real world, while the capacity equation (4) for binary input symbols could be used as a benchmark for BPSK modulation. In the following simulations, this set of LDPC codes is used in the two-user MAC with BPSK modulation in order to compare the achievable rates by signal superposition and SIC with the capacity region. The theoretical signal-to-noise ratio (SNR) to achieve error free transmission and actual SNR for those middle blocksize LDPC codes to achieve BER below 10^{-5} in the single-user AWGN

[1]The regular random LDPC matrices were constructed using the online software available at http://www.cs.toronto.edu/~radford/ldpc.software.html This program generates random regular LDPC codes of any specified rate and block length and is capable of expurging four cycles in the LDPC constraint graph.

TABLE I
THEORETICAL SNR AND ACTUAL SNR FOR THE 1200 BLOCKSIZE LDPC CODES BY BPSK MODULATION.

Code rate	0.4	0.5	0.7	0.8	0.9
Theoretical SNR (E_s/N_0 in dB)	-4.2182	-2.8233	-0.277	1.0704	2.7393
Actual SNR (E_s/N_0 in dB)	-1.9791	-0.8103	1.4510	2.5310	4.2424

TABLE II
SIMULATION RESULTS FOR $P_a = 1.5 * P_b$

Point	SNR & BER	Signal A	Signal B
(0.9, 0.4)	SNR (dB)	6.0033	4.2424
	BER	0	3.5×10^{-5}
(0.9, 0.5)	SNR (dB)	6.0033	4.2424
	BER	0.0295	0.0374
(0.4, 0.9)	SNR (dB)	6.3849	4.6240
	BER	8.25×10^{-5}	8.5×10^{-5}

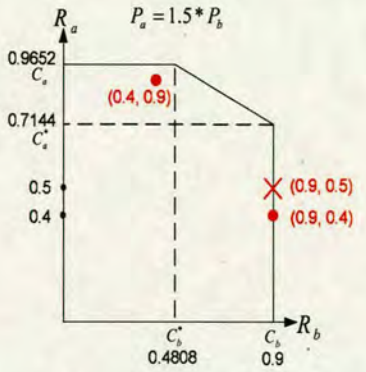

Fig. 4. Practical achievable point by using middle blocksize LDPC codes. $P_a = 1.5 * P_b$

channel are listed in Table I. The SNR is represented by E_s/N_0 which is the ratio of energy per transmitted code symbol and the one-sided power spectral density N_0 of the Gaussian channel noise.

In the multi-user framework below, the SNR next refers to a user's signal power P_k over the background Gaussian noise variance σ_w^2 in the channel, and the background Gaussian noise variance σ_w^2 is conveniently set to "1" (0dB) for the simulations.

A. First scenario: $P_a = 1.5 * P_b$

In this scenario, the received power of signal A is 1.5 times that of the received power of signal B. The parameters used to set up the simulation are SNR 6.00dB for signal A and SNR 4.24dB for signal B. The theoretical SNR for 0.9 information bits per channel use should be 2.74dB shown in Table I, but for this middle blocksize LDPC code SNR of 4.24dB (about 1.5dB more than the theoretical value) is needed to achieve BER below 10^{-4}. The theoretical values C_a and C_b are obtained by (4). C_a^* is obtained by (4) according to the assumption in the capacity calculation that the channel noise and the signal B both represent Gaussian interference to signal A. C_b^* is obtained in the same way by assuming signal A and channel noise are both "Gaussian interference" to B. By signal superposition and soft SIC, the first signal is detected as described in Section III-C and the channel code is decoded by SPA algorithm [11]. Then, the second signal B is obtained and decoded as described above. The simulation result is shown in table II and the theoretical capacity value and the practical achievable region for this scenario is shown in Figure 4.

1) point (0.9, 0.4): The simulation result shows that by using signal superposition and soft SIC, error free transmission (or very low BER for user A) could be achieved when the signal A with stronger received power transmits at the rate of 0.4 and the signal B with weaker received power transmits at the rate of 0.9. The stronger signal A is detected first.

2) point (0.9, 0.5): According to the theoretical capacity calculation, (0.9, 0.5) should be within the achievable region. But the results show that this rate pair is not achievable. The equivalent SNR for the first decoded user A in the multiple access channel is -0.09dB. But for the middle blocksize rate 0.5 LDPC code, SNR of -0.81dB is enough for error free transmission in single-user AWGN channel. It should be noticed that to the first signal A, the noise with power $(\sigma_w^2 + P_b)$ is the superposition of channel noise satisfying Gaussian distribution and the signal B satisfying the Bernoulli distribution (which takes value -1 with probability p and +1 with probability $1-p$ in the BPSK modulation). The pdf of this combined noise with power $(\sigma_w^2 + P_b)$ is

$$g(x) = p \cdot f(x+a) + (1-p) \cdot f(x-a), \tag{31}$$

in which $f(x) = \frac{1}{\sqrt{2\pi}\sigma_w} e^{-\frac{1}{2\sigma_w^2}(x)^2}$ and a represents the symbol power after BPSK. From (31) we could see that it is no longer Gaussian distributed noise to the first user. So the decoding of the first signal is not as good as in the single-user AWGN channel even though higher SNR is provided.

3) point (0.4, 0.9): According to the setup SNR in this scenario, the maximum code rate for user A by the theoretical calculation should be 0.9652. Here the code rate 0.9 is used instead in order to analyse the impact by reversing the order of detection. The decoding order for the point (0.4, 0.9) is to detect signal B with weaker power first, then after operating SIC the data for signal A with stronger signal is decoded. BER below 10^{-4} is also achievable for this point but the corresponding SNR needs to be (6.38dB, 4.62dB) which is higher than the SNR needed for point (0.9, 0.4). It means that if the decoding order is reversed to decode the weaker signal first, more signal power is needed to achieve the error free transmission in the MAC. Here, the power of one user is only 1.5 times of the other. If the two users have quite disparate power received at the destination, the decoding order will have an obvious influence on the total signal power to meet given target transmission rates [7].

B. Second scenario: $P_a = 2.0 * P_b$

In this scenario, the received power of signal A is twice of the received power of signal B. The parameters used to set up the simulation are listed in Table III. By detecting the stronger signal A first and operating SIC, point (0.8, 0.4) and point (0.8, 0.5) are achievable while the point (0.8, 0.7) could not be reached. The simulation results are shown in Table III and Figure 5.

V. CONCLUSION

Signal superposition together with SIC is the technique which achieves the boundary of the capacity region for the Gaussian Multiple Access Channel. A set of practical LDPC codes with moderate blocksize are used to investigate the practical performance of interference cancellation. In order to match the BER performance under the BPSK modulation, the binary-input channel capacity equation is used to obtain the theoretical values as the benchmark for the simulations. In practice SIC provides a higher transmission rate for the second

TABLE III
SIMULATION RESULTS FOR $P_a = 2.0 * P_b$

Point	SNR & BER	Signal A	Signal B
(0.8, 0.4)	SNR (dB)	5.5413	2.5310
	BER	0	0
(0.8, 0.5)	SNR (dB)	5.5413	2.5310
	BER	0	0
(0.8, 0.7)	SNR (dB)	5.5413	2.5310
	BER	0.092	0.1414

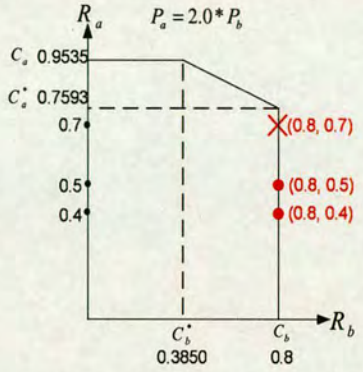

Fig. 5. Practical achievable point by using middle blocksize LDPC codes. $P_a = 2.0 * P_b$.

detected user compared with single-user detection method. The order of the signal detection and subtraction also has an impact on the required signal power to achieve the desired transmission rate. If the previously detected signal is just subtracted, error propagation will occur when the signal was not decoded correctly. For that reason we used a "soft" interference cancellation scheme that takes into account the reliablities with which bit decisions after channel decoding can be taken. The soft interference canceller was developed into a new form that lends itself for an efficient implementation that will also allow for iterative interference cancellation; this will be a topic of further work. We plan to present further results in the final version of the paper.

VI. Acknowledgement

This work was supported by Engineering and Physical Sciences Research Council (EPSRC), grant number EP/D 002184/1, and by the Scottish Funding Council for the Joint Research Institute with the Heriot-Watt University which is a part of the Edinburgh Research Partnership.

References

[1] John G. Proakis, *Digital Communications*, The McGraw-Hill Companies,Inc., 2001.
[2] D. Koulakiotis and A.H. Aghvami, "Data detection techniques for ds/cdma mobile systems: a review," *Personal Communications. IEEE [see also IEEE Wireless Communications]*, vol. 7, no. 3, pp. 24–34, June 2000.
[3] Joy A. Thomas Thomas M. Cover, *Elements of Information Theory*, John Wiley & Sons, Inc., second edition, 2006.
[4] T.J. Richardson and R.L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *Information Theory. IEEE Transactions on*, vol. 47, no. 2, pp. 599–618, Feb 2001.
[5] Andrea Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.
[6] A.J. Grant, B. Rimoldi, R.L. Urbanke, and P.A. Whiting, "Rate-splitting multiple access for discrete memoryless channels," *Information Theory. IEEE Transactions on*, vol. 47, no. 3, pp. 873–890, March 2001.
[7] David Tse and Pramod Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.
[8] R.G. Gallager, *Low Density Parity Check Code*, Cambridge:MA:MIT Press, 1963.
[9] R. Gallager, "Low-density parity-check codes," *Information Theory. IEEE Transactions on*, vol. 8, no. 1, pp. 21–28, Jan 1962.
[10] D.J.C. MacKay and R.M. Neal, "Near shannon limit performance of low density parity check codes," in *Electronics Letters*, 29 Aug. 1996, vol. 32, p. 1645.
[11] S. Lin and Daniel J. Costello, *Error Control Coding*, Pearson Prentice Hall, 2004.
[12] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *Information Theory. IEEE Transactions on*, vol. 42, no. 2, pp. 429–445, March 1996.
[13] G. Caire, R. Müller, and T. Tanaka, "Iterative multiuser joint decoding: Optimal power allocation and low-complexity implementation," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1950–1973, Sep. 2004.
[14] E.G. Larsson and B.R. Vojcic, "Cooperative transmit diversity based on superposition modulation," *Communications Letters. IEEE*, vol. 9, no. 9, pp. 778–780, Sep 2005.

A Shared-Relay Cooperative Diversity Scheme Based on Joint Channel and Network Coding in the Multiple Access Channel

Xiaoyan Xu, Mark F. Flanagan, Norbert Goertz
Institute for Digital Communications, Joint Research Institute for Signal & Image Processing
The University of Edinburgh, Scotland, UK
Email: X.Xu@ed.ac.uk, mark.flanagan@ieee.org, Norbert.Goertz@ieee.org

***Abstract*—In this paper we propose a cooperative diversity scheme for the scenario of two sources sharing a single relay. The scheme uses algebraic code superposition relaying in the multiple access fading channel to create spatial diversity under the constraint of limited communications resources. We also describe in detail a novel computationally efficient message passing algorithm at the destination's decoder which extracts the substantial spatial diversity contained in the code superposition and signal superposition. The decoder is based on a sliding window structure where certain *a posteriori* LLRs are retained as *a priori* LLRs for the next decoding. We show that despite the simplicity of the proposed scheme, diversity gains are efficiently leveraged by the simple combination of channel coding at the sources and network coding at the relay.**

I. Introduction

While wireless channels suffer from fading, at the same time the broadcast nature of wireless channels provides the possibility of a third party other than the destination "overhearing" the information that the source transmits. Thus apart from the original transmission channel, the same information could be transmitted to the destination through another independently fading channel. This generated spatial diversity can effectively combat the deleterious effect of fading [1]. In recent years, there has been increasing interest in applying the idea of algebraic code superposition, also called "network coding" [2]–[6] to the cooperative communications scenario. The network coding approach provides an efficient way to generate spatial diversity under the constraint of limited resources. One challenge is the problem of decoder design which should be able to cope with the complicated decoding situation at the destination [2]–[6].

In [5]–[7], the model of a typical network coding unit is considered in which the packets from the two sources are linearly combined at the relay. In our work we also focus on this cooperative transmission model for the situation where it is impractical for one mobile user to "capture" the other user's signal during its uplink transmission in the cellular network. Moreover, relay-based cooperative processing provides greater security than direct user cooperation in which user information must be shared.

In [4], a code superposition scheme employing low-density generator matrix (LDGM) codes is proposed to reduce the decoding complexity at the destination. But in order to do the graph-based decoding, the systematic bits must be retained without superposition which means that the potential superposition diversity is lower than that obtainable from fully superposed codewords. In [6], a combined LDPC code construction scheme including two channel code components and one network code component is produced by random parity-check matrix generation under certain constraints. The network codes are actually the parity checks for two channel codewords; this necessitates more complicated relay operations than simple superposition.

In this work we propose a cooperative coding scheme which is different from the previous work of [2]–[6] where superposed codewords experience a channel orthogonal to that of the original transmission, and also different from the previous work of [8] where simple codeword retransmission is employed in the multiple access Gaussian relay channel. Our scheme allows continuous transmission of superposed codewords by the relay and at the same time targets the challenge of coping with the interference introduced by the multiple access channel, thus making efficient use of communication resources to leverage spatial diversity gains. We also detail a novel efficient decoding algorithm based on message passing on the destination node's factor graph for the purpose of exploiting the spatial diversity contained in the algebraically superposed codewords together with the signal superposition introduced by the multiple access channel. The algorithm attains a separation of the three soft-input soft-output (SISO) decoder modules corresponding to each received signal stream; for convolutional codes, this separation affords a complexity advantage over decoding of the "nested code" [2], [3]; for LDPC codes, it affords a more efficient Tanner graph schedule than fully parallel decoding [6].

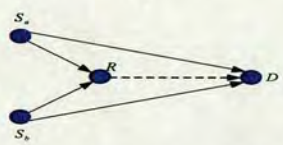

Fig. 1. Four-node communications network. Sources S_a and S_b share a common relay R as well as having direct links to the destination D.

Time Slot	0		1		...	t		t+1		...	L-1		L	
Half Slot	0	1	2	3	...	2t	2t+1	2t+2	2t+3	...	2L-2	2L-1	2L	2L+1
Source Transmits	$\mathbf{c}_0=$ $\mathbf{a}_0$	$\mathbf{c}_1=$ $\mathbf{b}_0$	$\mathbf{c}_2=$ $\mathbf{a}_1$	$\mathbf{c}_3=$ $\mathbf{b}_1$	...	$\mathbf{c}_{2t}=$ $\mathbf{a}_t$	$\mathbf{c}_{2t+1}=$ $\mathbf{b}_t$	$\mathbf{c}_{2t+2}=$ $\mathbf{a}_{t+1}$	$\mathbf{c}_{2t+3}=$ $\mathbf{b}_{t+1}$	...	$\mathbf{c}_{2L-2}=$ $\mathbf{a}_{L-1}$	$\mathbf{c}_{2L-1}=$ $\mathbf{b}_{L-1}$		
Relay Transmits		$\mathbf{d}_1=$ $\pi(\mathbf{a}_0)$	$\mathbf{d}_2=$ $\pi(\mathbf{b}_0)$ $\oplus\mathbf{a}_0$	$\mathbf{d}_3=$ $\pi(\mathbf{a}_1)$ $\oplus\mathbf{b}_0$	...	$\mathbf{d}_{2t}=$ $\pi(\mathbf{b}_{t-1})$ $\oplus\mathbf{a}_{t-1}$	$\mathbf{d}_{2t+1}=$ $\pi(\mathbf{a}_t)$ $\oplus\mathbf{b}_{t-1}$	$\mathbf{d}_{2t+2}=$ $\pi(\mathbf{b}_t)$ $\oplus\mathbf{a}_t$	$\mathbf{d}_{2t+3}=$ $\pi(\mathbf{a}_{t+1})$ $\oplus\mathbf{b}_t$	...	$\mathbf{d}_{2L-2}=$ $\pi(\mathbf{b}_{L-2})$ $\oplus\mathbf{a}_{L-2}$	$\mathbf{d}_{2L-1}=$ $\pi(\mathbf{a}_{L-1})$ $\oplus\mathbf{b}_{L-2}$	$\mathbf{d}_{2L}=$ $\pi(\mathbf{b}_{L-1})$ $\oplus\mathbf{a}_{L-1}$	$\mathbf{d}_{2L+1}=$ $\mathbf{b}_{L-1}$
Destination Receives	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	...	$\mathbf{e}_{2t}$	$\mathbf{e}_{2t+1}$	$\mathbf{e}_{2t+2}$	$\mathbf{e}_{2t+3}$	...	$\mathbf{e}_{2L-2}$	$\mathbf{e}_{2L-1}$	$\mathbf{e}_{2L}$	$\mathbf{e}_{2L+1}$
Destination Decodes			$\mathbf{a}_0$	$\mathbf{b}_0$	...	$\mathbf{a}_{t-1}$	$\mathbf{b}_{t-1}$	$\mathbf{a}_t$	$\mathbf{b}_t$	...	$\mathbf{a}_{L-2}$	$\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$

Fig. 2. Transmission schedule of proposed cooperative coding scheme.

Time Slot	0		1		...	t		t+1		...	L-1		
Half Slot	0	1	2	3	...	2t	2t+1	2t+2	2t+3	...	2L-2	2L-1	2L
Source Transmits	$\mathbf{c}_0=\mathbf{a}_0$	$\mathbf{c}_1=\mathbf{b}_0$	$\mathbf{c}_2=\mathbf{a}_1$	$\mathbf{c}_3=\mathbf{b}_1$	...	$\mathbf{c}_{2t}=\mathbf{a}_t$	$\mathbf{c}_{2t+1}=\mathbf{b}_t$	$\mathbf{c}_{2t+2}=\mathbf{a}_{t+1}$	$\mathbf{c}_{2t+3}=\mathbf{b}_{t+1}$	...	$\mathbf{c}_{2L-2}=\mathbf{a}_{L-1}$	$\mathbf{c}_{2L-1}=\mathbf{b}_{L-1}$	
Relay Transmits		$\mathbf{d}_1=\pi(\mathbf{a}_0)$	$\mathbf{d}_2=\pi(\mathbf{b}_0)$	$\mathbf{d}_3=\pi(\mathbf{a}_1)$	...	$\mathbf{d}_{2t}=\pi(\mathbf{b}_{t-1})$	$\mathbf{d}_{2t+1}=\pi(\mathbf{a}_t)$	$\mathbf{d}_{2t+2}=\pi(\mathbf{b}_t)$	$\mathbf{d}_{2t+3}=\pi(\mathbf{a}_{t+1})$	...	$\mathbf{d}_{2L-2}=\pi(\mathbf{b}_{L-2})$	$\mathbf{d}_{2L-1}=\pi(\mathbf{a}_{L-1})$	$\mathbf{d}_{2L}=\pi(\mathbf{b}_{L-1})$
Destination Receives	$\mathbf{e}_0$	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	...	$\mathbf{e}_{2t}$	$\mathbf{e}_{2t+1}$	$\mathbf{e}_{2t+2}$	$\mathbf{e}_{2t+3}$	...	$\mathbf{e}_{2L-2}$	$\mathbf{e}_{2L-1}$	$\mathbf{e}_{2L}$
Destination Decodes		$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_1$	...	$\mathbf{b}_{t-1}$	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_{t+1}$	...	$\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$

Fig. 3. Transmission schedule of consecutive relaying scheme in the multiple access channel.

II. Proposed Cooperative Coding Scheme

We consider the four-node communications network depicted in Figure 1, with two sources S_a and S_b, one relay R, and one destination D common to the two sources. The communication period is divided into $L+1$ *time slots* $t = 0, 1, \cdots, L$; each time slot $t \in \{0, 1, \cdots, L\}$ is further subdivided into 2 *half slots* ($2t$, $2t+1$). Source S_a has L messages to transmit, which it encodes into L n-bit codewords $\{\mathbf{a}_t : t = 0, 1, \cdots, L-1\}$. The code used at source S_a is $\mathcal{C}_a$ and is defined by the $m_a \times n$ parity-check matrix $\mathbf{H}_a = (H_a(j,i))$. Similarly, source S_b has L messages to transmit, which it encodes into L n-bit codewords $\{\mathbf{b}_t : t = 0, 1, \cdots, L-1\}$. The code used at source S_b is $\mathcal{C}_b$ and is defined by the $m_b \times n$ parity-check matrix $\mathbf{H}_b = (H_b(j,i))$. Thus, the codes $\mathcal{C}_a$ and $\mathcal{C}_b$ have the same length but not necessarily the same rate. In general, the codes used at the two sources can be LDPC or convolutional; in this paper we concentrate on LDPC codes. S_a and S_b broadcast their modulated codewords to the relay and destination nodes using TDMA and there is no cooperation between the two sources. For each $t \in \{0, 1, \cdots, 2L-1\}$, let $\mathbf{c}_t$ denote the codeword broadcast by the source in half slot t; thus $\mathbf{c}_{2t} = \mathbf{a}_t$ and $\mathbf{c}_{2t+1} = \mathbf{b}_t$ for $t \in \{0, 1, \cdots, L-1\}$.

The relay decodes and then re-encodes each codeword received from the source (the cooperative scheme is based on the scenario where the source is quite close to the relay). The relay also has a buffer in which it stores the codewords it obtained in the previous two half slots. At each half slot ($t = 2, 3, \cdots, 2L$), the relay interleaves the codeword obtained in half slot $t-1$ and superposes it (XOR operation) with the codeword obtained in half slot $t-2$; it then transmits the resulting codeword to the destination in the fading multiple access channel (MAC) whose channel resources are also shared with the source transmission. Special cases arise at half slots 1 and $2L+1$ in which only a single codeword is stored at the relay and no XOR operation is performed. Let $\mathbf{d}_t$ denote the codeword transmitted from the relay to the destination in half slot $t \in \{1, 2, \cdots, 2L+1\}$; thus $\mathbf{d}_t = \pi(\mathbf{c}_{t-1}) \oplus \mathbf{c}_{t-2}$, for $t = 2, 3, \cdots 2L$. Let $\mathbf{e}_t$ denote the signal stream received by the destination in half slot $t \in \{0, 1, \cdots, 2L+1\}$. The use of the MAC allows for no extra expense in terms of channel resources for the proposed cooperative transmission scheme as compared to the non-cooperative scheme. As we shall see, the spatial diversity gain generated by the superposition relaying outweighs the signal interference degradation inherent in using the MAC. To simplify the analysis, we assume that directed antennas are employed at the relay, so that interference between the source-relay link and the relay-destination link may be neglected.[1] For each $t = 0, 1, \cdots L-1$, source S_a's codeword $\mathbf{a}_t$ is decoded at the end of half slot $2t+2$ and source S_b's codeword $\mathbf{b}_t$ is decoded at the end of half slot $2t+3$. The transmission schedule for this cooperative coding scheme is illustrated in Figure 2. It can be seen that spatial diversity for each message is contained in three separate transmissions spanning three half slots.

As will be seen in Section IV, the interleaver π provides the "interleaver gain" for decoding at the destination. The interleaver is not in general necessary in the case of LDPC coding; however it may be used to avoid a large multiplicity of 8-cycles in the Tanner graph for the case where $\mathbf{H}_a = \mathbf{H}_b$.

[1] To avoid full-duplex at the relay in a practical communication system, a second relay R' which employs a simple amplify-and-forward (AF) scheme could be used between R and D in the current cell frequency $f1$. The transmission from R to R' could employ the neighboring cell frequency $f2$, and thus the uplink MAC still maintain $f1$ for this cell. In the counterpart neighboring cell with frequency $f2$, $f1$ is used in the R to R' link. In the overall communications system, there will be no extra frequency band occupied and only negligible inter-cell interference. The decoding procedure at the destination node is then identical to that presented in this paper.

As one reference system, the transmission schedule for consecutive relaying in the MAC is depicted in Figure 3; here the relay simply re-transmits the (interleaved) previously received codeword rather than a codeword superposition. It is easily seen that in this scheme, spatial diversity for each message is contained in two separate transmissions spanning two half slots. A simulation-based comparison of the two schemes described in this section under a transmit power constraint will be given in Section V.

III. Soft Demodulator for BPSK Modulation in the Multiple Access Channel

In the following we define $\mathcal{I} = \{1, 2, \cdots, n\}$, and $x^{(i)}$ denotes the i-th bit of codeword $\mathbf{x}$ while $e^{(i)}$ denotes the i-th received channel value of signal stream $\mathbf{e}$. Considering a single half slot $t \in \{1, \cdots, 2L-1\}$, the channel model is given by

$$e_t^{(i)} = \phi_t^{(i)} \cdot \alpha(c_t^{(i)}) + \psi_t^{(i)} \cdot \beta(d_t^{(i)}) + n_t^{(i)}, \quad i \in \mathcal{I}, \quad (1)$$

where $\phi_t^{(i)}$ and $\psi_t^{(i)}$ represent the fading processes on the source-destination and relay-destination links, respectively, and $n_t^{(i)}$ is complex AWGN with variance σ^2 per dimension. Also

$$\alpha(c_t^{(i)}) = \sqrt{P_S}(1 - 2c_t^{(i)}) \quad (2)$$

$$\beta(d_t^{(i)}) = \sqrt{P_R}(1 - 2d_t^{(i)}) \quad (3)$$

holds for BPSK modulation, with P_S and P_R representing the receive power for the symbols transmitted by the source and relay respectively.

The function of the soft demodulator is to take as input extrinsic LLRs on the transmitted bits (LLRs in the absence of channel information); without loss of generality we consider the bit $c_t^{(i)}$,

$$L^E(c_t^{(i)}) = \ln\left(\frac{\Pr(c_t^{(i)} = 0)}{\Pr(c_t^{(i)} = 1)}\right) \quad (4)$$

and compute the new extrinsic LLRs (incremental LLRs expressing new information derived from the channel)

$$L^O(c_t^{(i)}) = L(c_t^{(i)}) - L^E(c_t^{(i)}) \quad (5)$$

where the *a posteriori* LLRs $L(c_t^{(i)})$ (incorporating channel information) are given by

$$\begin{aligned} L(c_t^{(i)}) &= \ln\left(\frac{\Pr(c_t^{(i)} = 0 | e_t^{(i)})}{\Pr(c_t^{(i)} = 1 | e_t^{(i)})}\right) \\ &= \ln\left(\frac{\sum_{\{c_t^{(i)}, d_t^{(i)}\}: c_t^{(i)} = 0} \Pr(c_t^{(i)}, d_t^{(i)} | e_t^{(i)})}{\sum_{\{c_t^{(i)}, d_t^{(i)}\}: c_t^{(i)} = 1} \Pr(c_t^{(i)}, d_t^{(i)} | e_t^{(i)})}\right) \\ &= \ln\left(\frac{\sum_{\{c_t^{(i)}, d_t^{(i)}\}: c_t^{(i)} = 0} p(e_t^{(i)} | c_t^{(i)}, d_t^{(i)}) \cdot \Pr(c_t^{(i)}, d_t^{(i)})}{\sum_{\{c_t^{(i)}, d_t^{(i)}\}: c_t^{(i)} = 1} p(e_t^{(i)} | c_t^{(i)}, d_t^{(i)}) \cdot \Pr(c_t^{(i)}, d_t^{(i)})}\right) \end{aligned} \quad (6)$$

We assume that the users' code bits are independent (a realistic assumption for independent users), so that

$$\Pr(c_t^{(i)}, d_t^{(i)}) = \Pr(c_t^{(i)}) \cdot \Pr(d_t^{(i)}) \quad (7)$$

Also, the probability density function (PDF) of the receive channel value conditioned on the transmitted bits may be written as

$$p(e_t^{(i)} | c_t^{(i)}, d_t^{(i)}) = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)} - \phi_t^{(i)} \cdot \alpha(c_t^{(i)}) - \psi_t^{(i)} \cdot \beta(d_t^{(i)})\right|^2\right) \quad (8)$$

Therefore, (5) may be re-written as (9), for which we use the shorthand $L^O(c_t^{(i)}) = f_1(e_t^{(i)}, L^E(d_t^{(i)}))$. In the same way, the new extrinsic LLR of the code bit $d_t^{(i)}$ may be obtained from (10), for which we use the shorthand $L^O(d_t^{(i)}) = f_2(e_t^{(i)}, L^E(c_t^{(i)}))$.

IV. Decoding Algorithm at Destination Node

Without loss of generality, we consider the decoding of codeword $\mathbf{a}_t$ at the end of half slot $2t+2$, for $t \in \{0, 1, \cdots, L-2\}$. The two codewords contained in the received signal stream $\mathbf{e}_{2t+2}$ are:

$$\begin{aligned} \mathbf{c}_{2t+2} &= \mathbf{a}_{t+1} \\ \mathbf{d}_{2t+2} &= \pi(\mathbf{c}_{2t+1}) \oplus \mathbf{c}_{2t} \\ &= \pi(\mathbf{b}_t) \oplus \mathbf{a}_t \end{aligned}$$

We assume that *a priori* LLRs on $\mathbf{c}_{2t+1}$ and $\mathbf{c}_{2t}$, denoted $\{L_1(c_{2t+1}^{(i)})\}$ and $\{L_2(c_{2t}^{(i)})\}$ respectively, are available from the previous decoding. In addition to decoding of $\mathbf{a}_t$, the decoder will produce *a posteriori* LLRs $\{L_1(c_{2t+2}^{(i)})\}$ and (updated) *a posteriori* LLRs $\{L_2(c_{2t+1}^{(i)})\}$; these will be used as *a priori* LLRs in the next decoding.

Next, we provide a concise description of the factor graph based decoding algorithm [9] at the destination decoder. The factor graph for the decoding is illustrated in Figure 4, where circles depict variable nodes and squares depict factor nodes. Extrinsic information is exchanged between three soft-input soft-output (SISO) decoder modules for the constituent codes (two codes $\mathcal{C}_a$ and one code $\mathcal{C}_b$), via the factor nodes $\{F_i\}$ which correspond to the network coding operation at the relay and the factor nodes $\{G_i\}$ which correspond to the soft demodulator for the MAC as given by (9) and (10). For simplicity, the graph is illustrated for the case $n = 3$, and where $\mathcal{C}_a$ and $\mathcal{C}_b$ are (trivial) LDPC codes. For convolutional constituent codes, the SISO modules execute BCJR algorithms. Note that in the convolutional case, the separation of the two (e.g. M-state) decoder SISO modules gives a complexity advantage over schemes which use a larger (e.g. M^2-state) decoder to decode the "nested" code generated at the relay (see e.g. [2], [3]). In the LDPC case this separation of SISO modules effects a more efficient message-passing schedule than does fully parallel decoding on the Tanner graph of the nested code (see e.g. [6]).

$$L^O(c_t^{(i)}) = \ln\left(\frac{\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)} - \sqrt{P_S}\phi_t^{(i)} - \sqrt{P_R}\psi_t^{(i)}\right|^2 + L^E(d_t^{(i)})\right) + \exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)} - \sqrt{P_S}\phi_t^{(i)} + \sqrt{P_R}\psi_t^{(i)}\right|^2\right)}{\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)} + \sqrt{P_S}\phi_t^{(i)} - \sqrt{P_R}\psi_t^{(i)}\right|^2 + L^E(d_t^{(i)})\right) + \exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)} + \sqrt{P_S}\phi_t^{(i)} + \sqrt{P_R}\psi_t^{(i)}\right|^2\right)}\right) \tag{9}$$

$$L^O(d_t^{(i)}) = \ln\left(\frac{\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)} - \sqrt{P_R}\psi_t^{(i)} - \sqrt{P_S}\phi_t^{(i)}\right|^2 + L^E(c_t^{(i)})\right) + \exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)} - \sqrt{P_R}\psi_t^{(i)} + \sqrt{P_S}\phi_t^{(i)}\right|^2\right)}{\exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)} + \sqrt{P_R}\psi_t^{(i)} - \sqrt{P_S}\phi_t^{(i)}\right|^2 + L^E(c_t^{(i)})\right) + \exp\left(-\frac{1}{2\sigma^2}\left|e_t^{(i)} + \sqrt{P_R}\psi_t^{(i)} + \sqrt{P_S}\phi_t^{(i)}\right|^2\right)}\right) \tag{10}$$

Next we introduce some notational conventions pertaining to the following algorithm description. In all cases, the letter λ is used to denote LLRs corresponding to messages passed on the factor graph (i.e. extrinsic LLRs). The interleaving "π" is interpreted as

$$\mathbf{x} = \pi(\mathbf{y}) \iff x^{(i)} = y^{(\pi(i))} \ \forall i \in \mathcal{I}$$

Some index sets are defined as follows. $\mathcal{J}_a = \{1, 2, \cdots m_a\}$; $\mathcal{J}_b = \{1, 2, \cdots m_b\}$; $\mathcal{N}_a(i) = \{j \in \mathcal{J}_a : H_a(j,i) = 1\}$; $\mathcal{N}_b(i) = \{j \in \mathcal{J}_b : H_b(j,i) = 1\}$; $\mathcal{M}_a(j) = \{i \in \mathcal{I} : H_a(j,i) = 1\}$; $\mathcal{M}_b(j) = \{i \in \mathcal{I} : H_b(j,i) = 1\}$. Also, $\boxplus$ denotes the (commutative and associative) "box-plus" operation [10], i.e.

$$\boxplus_{s\in\mathcal{S}} \lambda_s = \log\left(\frac{1 + \prod_{s\in\mathcal{S}} \tanh(\lambda_s/2)}{1 - \prod_{s\in\mathcal{S}} \tanh(\lambda_s/2)}\right)$$

N denotes the maximum number of decoding iterations.

Factor Graph Based Decoding Algorithm at Destination Node – Decoding of Codeword $\mathbf{a}_t$

Initialization:

- For $i \in \mathcal{I}$,

$$\lambda_{c_1}^{(i)} = L_2(c_{2t}^{(i)}) \tag{11}$$
$$\lambda_{c_2}^{(i)} = L_1(c_{2t+1}^{(i)}) \tag{12}$$
$$\lambda_8^{(i)} = 0 \tag{13}$$
$$\lambda_{15}^{(i)} = 0 \tag{14}$$

- For $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_2^{(i,j)} = 0 \tag{15}$$

- For $i \in \mathcal{I}$, $j \in \mathcal{N}_b(i)$

$$\lambda_4^{(i,j)} = 0 \tag{16}$$

- For $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_6^{(i,j)} = 0 \tag{17}$$

Main Loop: For $k = 1$ to N do

- For $i \in \mathcal{I}$,

$$\lambda_{14}^{(i)} = \lambda_{c_1}^{(i)} + \lambda_8^{(i)} \tag{18}$$

- Network coding constraints: for $i \in \mathcal{I}$,

$$\lambda_{10}^{(\pi(i))} = \lambda_{14}^{(i)} \boxplus \lambda_{15}^{(i)} \tag{19}$$

- For $i \in \mathcal{I}$,

$$\lambda_9^{(i)} = \lambda_{c_2}^{(i)} + \lambda_{10}^{(i)} \tag{20}$$

- SISO decoder $\mathcal{C}_b$: for $i \in \mathcal{I}$, $j \in \mathcal{N}_b(i)$

$$\lambda_3^{(i,j)} = \lambda_9^{(i)} + \sum_{l\in\mathcal{N}_b(i)\setminus\{j\}} \lambda_4^{(i,l)} \tag{21}$$
$$\lambda_4^{(i,j)} = \boxplus_{l\in\mathcal{M}_b(j)\setminus\{i\}} \lambda_3^{(l,j)} \tag{22}$$

For $i \in \mathcal{I}$,

$$\lambda_{12}^{(i)} = \sum_{j\in\mathcal{N}_b(i)} \lambda_4^{(i,j)} \tag{23}$$

- Obtain the *a posteriori* LLR (prior for decoding in next half slot)

$$L_2(c_{2t+1}^{(i)}) = \lambda_9^{(i)} + \lambda_{12}^{(i)} \tag{24}$$

- For $i \in \mathcal{I}$,

$$\lambda_{11}^{(i)} = \lambda_{c_2}^{(i)} + \lambda_{12}^{(i)} \tag{25}$$
$$\lambda_{16}^{(i)} = \lambda_{14}^{(i)} \boxplus \lambda_{11}^{(\pi(i))} \tag{26}$$

- Soft demodulation: for $i \in \mathcal{I}$,

$$\lambda_{17}^{(i)} = f_1(e_{2t+2}^{(i)}, \lambda_{16}^{(i)}) \tag{27}$$

- SISO decoder $\mathcal{C}_a$: for $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_5^{(i,j)} = \lambda_{17}^{(i)} + \sum_{l\in\mathcal{N}_a(i)\setminus\{j\}} \lambda_6^{(i,l)} \tag{28}$$
$$\lambda_6^{(i,j)} = \boxplus_{l\in\mathcal{M}_a(j)\setminus\{i\}} \lambda_5^{(l,j)} \tag{29}$$

For $i \in \mathcal{I}$,

$$\lambda_{18}^{(i)} = \sum_{j\in\mathcal{N}_a(i)} \lambda_6^{(i,j)} \tag{30}$$

- Obtain the *a posteriori* LLR (prior for decoding in next half slot)

$$L_1(c_{2t+2}^{(i)}) = \lambda_{17}^{(i)} + \lambda_{18}^{(i)} \tag{31}$$

- Soft demodulation: for $i \in \mathcal{I}$,

$$\lambda_{15}^{(i)} = f_2(e_{2t+2}^{(i)}, \lambda_{18}^{(i)}) \tag{32}$$
$$\lambda_{13}^{(i)} = \lambda_{15}^{(i)} \boxplus \lambda_{11}^{(\pi(i))} \tag{33}$$

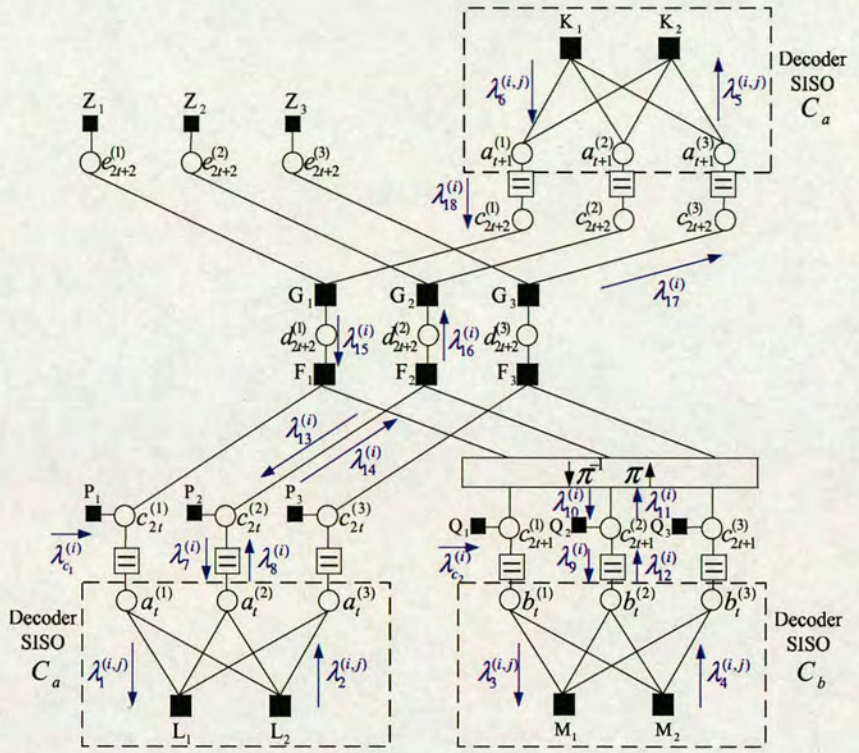

Fig. 4. Factor graph corresponding to the destination's decoding of codeword $\mathbf{a}_t$. The three decoder SISO modules exchange extrinsic information via the factor nodes $\{F_i\}$ which correspond to the network coding operation at the relay and the factor nodes $\{G_i\}$ which correspond to the soft demodulator for the MAC. For ease of presentation, the factor graph is illustrated for the case $n = 3$ and trivial codes $\mathcal{C}_a$, $\mathcal{C}_b$.

- For $i \in \mathcal{I}$,

$$\lambda_7^{(i)} = \lambda_{c_1}^{(i)} + \lambda_{13}^{(i)} \tag{34}$$

- SISO decoder $\mathcal{C}_a$: for $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_1^{(i,j)} = \lambda_7^{(i)} + \sum_{l \in \mathcal{N}_a(i) \setminus \{j\}} \lambda_2^{(i,l)} \tag{35}$$

$$\lambda_2^{(i,j)} = \boxplus_{l \in \mathcal{M}_a(j) \setminus \{i\}} \lambda_1^{(l,j)} \tag{36}$$

For $i \in \mathcal{I}$,

$$\lambda_8^{(i)} = \sum_{j \in \mathcal{N}_a(i)} \lambda_2^{(i,j)} \tag{37}$$

- Calculate *a posteriori* LLRs for codeword $\mathbf{a}_t$:

$$L(a_t^{(i)}) = \lambda_7^{(i)} + \lambda_8^{(i)} \tag{38}$$

- Make decisions on the code bits

$$\hat{a}_t^{(i)} = \begin{cases} 0 & \text{if } L(a_t^{(i)}) \geq 0 \\ 1 & \text{if } L(a_t^{(i)}) < 0 \end{cases}$$

If $\hat{\mathbf{a}}_t \mathbf{H}_a^T = \mathbf{0}$ **then** ***break;***

Endfor

It may be seen that the decoding of $\mathbf{a}_t$ spans three transmission frames (half slots) $2t \rightarrow 2t+1 \rightarrow 2t+2$, during which the decoding follows the three-step evolution $L_1(c_{2t}^{(i)}) \rightarrow L_2(c_{2t}^{(i)}) \rightarrow \hat{a}_t^{(i)}$.

The preceding presentation was for the general case of decoding (half slots $t = 2, 3, \cdots, 2L-1$). Special cases are handled in a straightforward manner as follows. In half slot $t = 0$, a single LDPC decoding of $\mathbf{a}_0$ based on LLRs derived from the receive stream $\mathbf{e}_0$ produces $\{L_1(c_0^{(i)})\}$. In half slot $t = 1$, decoding proceeds as in the general case except that $\lambda_{14}^{(i)}$ is set to $+\infty$ for all $i \in \mathcal{I}$; $\{L_1(c_1^{(i)})\}$ and $\{L_2(c_0^{(i)})\}$ are produced (note that this decoder has the same structure as the decoder for the reference scheme of consecutive relaying in the fading MAC as shown in Figure 3). In half slot $t = 2L$, the decoder structure used is that for superposition decoding as described in [2], [3]; also only $\{L_2(c_{2L-1}^{(i)})\}$ are produced. In time slot $t = 2L+1$, a single LDPC decoding of $\mathbf{b}_{L-1}$ is performed based on the sum of the LLRs derived from the receive stream $\mathbf{e}_{2L+1}$ and the LLRs $\{L_2(c_{2L-1}^{(i)})\}$ obtained from the previous decoding.

V. Simulation Results

In this section, we provide a comparison of the proposed cooperative scheme with two reference schemes. The first is consecutive relaying in the fading MAC as shown in Figure 3. The second is a simple TDMA transmission scheme without cooperation or relaying. Fair comparison of the three cooperative schemes is based on the constraint that in simulations, each scheme uses the same codes $\mathcal{C}_a$ and $\mathcal{C}_b$, and the same total energy E for transmission of the $2L$ source messages.

The LDPC codes used for the simulations are randomly generated rate $1/2$ regular LDPC codes of block length $n = 1200$ with column weight 3 and no 4-cycles in the Tanner graph. In simulations we choose $\mathbf{H}_a = \mathbf{H}_b$, and assume BPSK modulation for all three systems. A random interleaver π is used to avoid 8-cycle multiplicity in the Tanner graph. We consider a quasi-static Rayleigh fading channel, for which the fading coefficients are constant within each half slot (one codeword) and change independently from one half slot to the next. We assume equal signal-to-noise ratio (SNR) on the two source-destination links and the relay-destination link, and we assume that the destination has perfect knowledge of the channel fading coefficients and noise variances. As for the two source-relay links, which play a key role in the performance of the system since poor link quality may lead to catastrophic error propagation at the destination decoder, the simulation setup is such that the outage probabilities of the source-relay links are both around 10^{-5} – this is a suitable value for the operation of cooperative schemes. The overall performance of both cooperative schemes will degrade as the outage probability of the source-relay links increases; results under varying outage probability on the source-relay links are omitted due to the space limitations.

Simulated performance results for bit error rate (BER) and frame (codeword) error rate (FER) are shown in Figures 5 and 6 respectively. The curve corresponding to the proposed cooperative scheme exhibits the steepest slope (due to increased diversity gain) of the three, outperforming the others in the SNR region of interest. The scheme attains approximately an order of magnitude decrease in both BER and FER over consecutive relaying in the fading MAC at an E_b/N_0 of 10 dB.

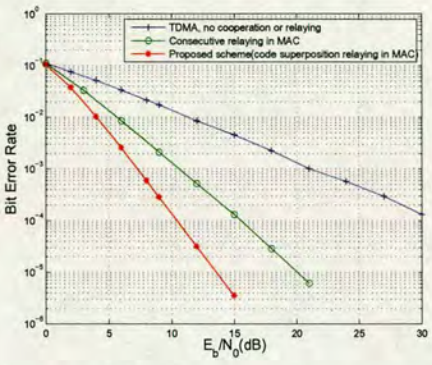

Fig. 5. Comparative BER performance for the three communication schemes.

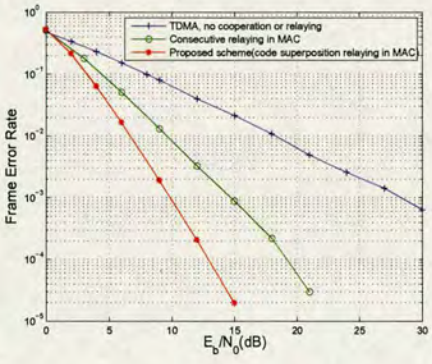

Fig. 6. Comparative FER performance for the three communication schemes.

VI. Acknowledgment

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), grant number EP/D 002184/1, and by the Scottish Funding Council for the Joint Research Institute with Heriot-Watt University which is a part of the Edinburgh Research Partnership.

References

[1] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine*, pp. 74–80, Octorber 2004.

[2] L. Xiao, T. Fuja, J. Kliewer, and D. Costello, "Cooperative diversity based on code superposition," *IEEE International Symposium on Information Theory (ISIT)*, 2006.

[3] L. Xiao, T. E. Fuja, J. Kliewer, and D. J. Costello, "A network coding approach to cooperative diversity," *IEEE Trans. Information Theory*, vol. 53, no. 10, pp. 3714–3722, Oct. 2007.

[4] L. Xiao, T. Fuja, J. Kliewer, and D. Costello, "Algebraic superposition of LDGM codes for cooperative diversity," *IEEE International Symposium on Information Theory (ISIT)*, 2007.

[5] C. Hausl and P. Dupraz, "Joint network-channel coding for the multiple-access relay channel," in *Sensor and Ad Hoc Communications and Networks, 2006. SECON '06. 2006 3rd Annual IEEE Communications Society on*, vol. 3, 28-28 Sept. 2006, pp. 817–822.

[6] C. Hausl, F. Schreckenbach, I. Oikonomidis, and G. Bauch, "Iterative network and channel decoding on a tanner graph," *43rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, USA*, 2005.

[7] S. Zhang, Y. Zhu, S.-C. Liew, and K. B. Letaief, "Joint design of network coding and channel decoding for wireless networks," in *Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE*, 11-15 March 2007, pp. 779–784.

[8] M. A. Khojastepour, N. Ahmed, and B. Aazhang, "Code design for the relay channel and factor graph decoding," *Thirty-Eighth Asilomar Conference on Signal, Systems and Computers*, vol. 2, pp. 2000–2004, 2004.

[9] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Information Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[10] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *Information Theory, IEEE Transactions on*, vol. 42, no. 2, pp. 429–445, March 1996.

International Symposium on Information Theory and its Applications, ISITA2008
Auckland, New Zealand, 7-10, December, 2008

A Shared-Relay Cooperative Diversity Scheme Based on Joint Channel and Network Coding

Xiaoyan Xu[†], Mark F. Flanagan[‡], Christian Koller[††] and Norbert Goertz[†]

[†] Institute for Digital Communications,
Joint Research Institute for Signal & Image Processing
The University of Edinburgh, Scotland, UK
E-mail: X.Xu@ed.ac.uk, Norbert.Goertz@ieee.org

[‡] Department of Electronics,
Computer Science and Systems (DEIS)
The University of Bologna, Italy
E-mail: mark.flanagan@ieee.org

[††] Dept. of Electrical Engineering,
275 Fitzpatrick Hall, University of Notre Dame
Notre Dame, IN 46556, US
E-mail: Christian.Koller.3@nd.edu

Abstract
In cooperative communications, the technique of algebraic code superposition (or *network coding*) may efficiently utilize limited communications resources to produce spatial diversity. In this paper, we propose a simple cooperative diversity scheme for the scenario of two sources sharing a single relay. No cooperative coding is required at the sources, and only very simple interleave-and-superpose operations are performed at the relay. We propose a novel computationally efficient message passing algorithm at the destination's decoder which utilizes spatial diversity gathered over multiple transmission frames. We show that despite the simplicity of the proposed scheme, diversity gains are efficiently leveraged by the simple combination of channel coding at the sources and network coding at the relay.

1. INTRODUCTION

While wireless channels suffer from fading, at the same time the broadcast nature of wireless channels provides the possibility of a third party other than the destination "overhearing" the information that the source transmits. Thus apart from the original transmission channel, the same information could be transmitted to the destination through another independently fading channel. This generated spatial diversity can effectively combat the deleterious effect of fading [1]. In recent years, there has been increasing interest in applying the idea of algebraic code superposition, also called "network coding" [2, 3, 4, 5, 6] to the cooperative communications scenario. The network coding approach provides an efficient way to generate spatial diversity under the constraint of limited resources. One challenge is the problem of decoder design which should be able to cope with the complicated decoding situation at the destination [2, 3, 4, 5, 6].

In [5, 6], the model of a typical network coding unit is considered in which the packets from the two sources are linearly combined at the relay. In our work we also focus on this cooperative transmission model for the situation where it is impractical for one mobile user to "capture" the other user's signal during its uplink transmission in the cellular network. Moreover, relay-based cooperative processing provides greater security than direct user cooperation in which user information must be shared.

In [4], a code superposition scheme employing low-density generator matrix (LDGM) codes is proposed to reduce the decoding complexity at the destination. But in order to do the graph-based decoding, the systematic bits must be retained without superposition which means that the potential superposition diversity is lower than that obtainable from fully superposed codewords. In [6], a combined LDPC code construction scheme including two channel code components and one network code component is produced by random parity-check matrix generation under certain constraints. The network codes are actually the parity checks for two channel codewords; this necessitates more complicated relay operations than simple superposition.

In this work we propose a cooperative coding scheme which, in contrast to previous work where TDMA-only or FDMA-only relaying is assumed, allows continuous transmission of superposed codewords by the relay, thus making efficient use of communication resources to leverage spatial diversity gains. The proposed scheme also arises from the observation that idle

frequency channels do exist and could be exploited in current TDMA systems. We shall detail a novel efficient decoding algorithm based on message passing on the destination node's factor graph for the purpose of exploiting the spatial diversity contained in the algebraically superposed codewords. The algorithm attains a separation of the two soft-input soft-output (SISO) decoder modules corresponding to the codes employed by the two sources; for convolutional codes, this separation affords a complexity advantage over decoding of the "nested code" [2, 3]; for LDPC codes, it affords a more efficient Tanner graph schedule than fully parallel decoding [6].

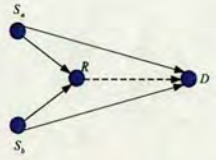

Figure 1: Four-node communications network. Sources S_a and S_b share a common relay R as well as having direct links to the destination D.

2. PROPOSED COOPERATIVE CODING SCHEME

We consider the four-node communications network depicted in Figure 1, with two sources S_a and S_b, one relay R, and one destination D common to the two sources. The communication period is divided into $L+1$ *time slots* $t = 0, 1, \cdots, L$; each time slot $t \in \{0, 1, \cdots, L\}$ is further subdivided into 2 *half slots* $(2t, 2t+1)$. Source S_a has L messages to transmit, which it encodes into L n-bit codewords $\{\mathbf{a}_t : t = 0, 1, \cdots, L-1\}$. The code used at source S_a is $\mathcal{C}_a$ and is defined by the $m_a \times n$ parity-check matrix $\mathbf{H}_a = (H_a(j,i))$. Similarly, source S_b has L messages to transmit, which it encodes into L n-bit codewords $\{\mathbf{b}_t : t = 0, 1, \cdots, L-1\}$. The code used at source S_b is $\mathcal{C}_b$ and is defined by the $m_b \times n$ parity-check matrix $\mathbf{H}_b = (H_b(j,i))$. Thus, the codes $\mathcal{C}_a$ and $\mathcal{C}_b$ have the same length but not necessarily the same rate. In general, the codes used at the two sources can be LDPC or convolutional; in this paper we concentrate on LDPC codes. S_a and S_b broadcast their modulated codewords to the relay and destination nodes using TDMA in frequency band f_1 and there is no cooperation between the two sources. For each $t \in \{0, 1, \cdots, 2L-1\}$, let $\mathbf{c}_t$ denote the codeword broadcast by the source in half slot t; thus $\mathbf{c}_{2t} = \mathbf{a}_t$ and $\mathbf{c}_{2t+1} = \mathbf{b}_t$ for $t \in \{0, 1, \cdots, L-1\}$.

The relay decodes and then re-encodes each codeword received from the source (the cooperative scheme is based on the scenario where the source is quite close to the relay). The relay also has a buffer in which it stores the codewords it obtained in the previous two half slots. At each half slot $(t = 2, 3, \cdots, 2L)$, the relay interleaves the codeword received in half slot $t-1$ and superposes it (XOR operation) with the codeword received in half slot $t-2$; it then transmits the resulting codeword to the destination in frequency band f_2. Special cases arise at half slots 1 and $2L+1$ in which only a single codeword is stored at the relay and no XOR operation is performed. The allocation of different frequency bands to the transmission channel of the relay and the broadcast channels of the sources allows for simultaneous transmission and reception by the relay node, thus allowing efficient leveraging of communication resources for spatial diversity. Let $\mathbf{d}_t$ denote the codeword transmitted from the relay to the destination in half slot $t \in \{1, 2, \cdots, 2L+1\}$; thus $\mathbf{d}_t = \pi(\mathbf{c}_{t-1}) \oplus \mathbf{c}_{t-2}$, for $t = 2, 3, \cdots 2L$. For each $t = 0, 1, \cdots L-1$, source S_a's codeword $\mathbf{a}_t$ is decoded at the end of half slot $2t+2$ and source S_b's codeword $\mathbf{b}_t$ is decoded at the end of half slot $2t+3$. The transmission schedule for this cooperative coding scheme is illustrated in Figure 2. It may be seen from the Figure that spatial diversity for each message is contained in three codewords received at the destination.

Note that decoding and re-encoding of received codewords using a different code (as in the scheme of [2]) is not performed by the relay in this scheme; this is in order to keep the relay operation as simple as possible. The interleaver π actually provides the "interleaver gain" for decoding at the destination. The interleaver is not in general necessary in the case of LDPC coding; however it may be used to avoid a large multiplicity of 8-cycles in the Tanner graph for the case where $\mathbf{H}_a = \mathbf{H}_b$.

The transmission schedule for consecutive relaying (see Figure 3), in which the time and frequency allocations are the same as in the proposed cooperative scheme, except that the relay simply retransmits the previously received codeword rather than a codeword superposition. Also, the transmission schedules for simple TDMA and FDMA based relaying schemes are depicted in Figures 4 and 5 respectively. It is easily seen that in all three of these schemes, spatial diversity for each message is contained in two codewords received at the destination. A simulation-based comparison of all schemes described in this section under a transmit power constraint will be given in Section 5.

3. DECODING ALGORITHM AT DESTINATION NODE

Time Slot		0		1		...	t		t+1		...	L-1		L	
Half Slot		0	1	2	3	...	2t	2t+1	2t+2	2t+3	...	2L-2	2L-1	2L	2L+1
Relay Receives	From Source (f1)	$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_1$	$\mathbf{b}_1$	...	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_{t+1}$	$\mathbf{b}_{t+1}$	...	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$		
Destination Receives	From Source (f1)	$\mathbf{c}_0=$ $\mathbf{a}_0$	$\mathbf{c}_1=$ $\mathbf{b}_0$	$\mathbf{c}_2=$ $\mathbf{a}_1$	$\mathbf{c}_3=$ $\mathbf{b}_1$	...	$\mathbf{c}_{2t}=$ $\mathbf{a}_t$	$\mathbf{c}_{2t+1}=$ $\mathbf{b}_t$	$\mathbf{c}_{2t+2}=$ $\mathbf{a}_{t+1}$	$\mathbf{c}_{2t+3}=$ $\mathbf{b}_{t+1}$	...	$\mathbf{c}_{2L-2}=$ $\mathbf{a}_{L-1}$	$\mathbf{c}_{2L-1}=$ $\mathbf{b}_{L-1}$		
	From Relay (f2)		$\mathbf{d}_1=$ $\pi(\mathbf{a}_0)$	$\mathbf{d}_2=$ $\pi(\mathbf{b}_0)$ $\oplus\mathbf{a}_0$	$\mathbf{d}_3=$ $\pi(\mathbf{a}_1)$ $\oplus\mathbf{b}_0$	...	$\mathbf{d}_{2t}=$ $\pi(\mathbf{b}_{t-1})$ $\oplus\mathbf{a}_{t-1}$	$\mathbf{d}_{2t+1}=$ $\pi(\mathbf{a}_t)$ $\oplus\mathbf{b}_{t-1}$	$\mathbf{d}_{2t+2}=$ $\pi(\mathbf{b}_t)$ $\oplus\mathbf{a}_t$	$\mathbf{d}_{2t+3}=$ $\pi(\mathbf{a}_{t+1})$ $\oplus\mathbf{b}_t$	...	$\mathbf{d}_{2L-2}=$ $\pi(\mathbf{b}_{L-2})$ $\oplus\mathbf{a}_{L-2}$	$\mathbf{d}_{2L-1}=$ $\pi(\mathbf{a}_{L-1})$ $\oplus\mathbf{b}_{L-2}$	$\mathbf{d}_{2L}=$ $\pi(\mathbf{b}_{L-1})$ $\oplus\mathbf{a}_{L-1}$	$\mathbf{d}_{2L+1}=$ $\mathbf{b}_{L-1}$
Destination Decodes				$\mathbf{a}_0$	$\mathbf{b}_0$	...	$\mathbf{a}_{t-1}$	$\mathbf{b}_{t-1}$	$\mathbf{a}_t$	$\mathbf{b}_t$	...	$\mathbf{a}_{L-2}$	$\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$

Figure 2: Transmission schedule of proposed cooperative coding scheme.

Time Slot		0		1		...	t		t+1		...	L-1		
Half Slot		0	1	2	3	...	2t	2t+1	2t+2	2t+3	...	2L-2	2L-1	2L
Destination Receives	From Source (f1)	$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_1$	$\mathbf{b}_1$	...	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_{t+1}$	$\mathbf{b}_{t+1}$	...	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$	
	From Relay (f2)		$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_1$	...	$\mathbf{b}_{t-1}$	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_{t+1}$	...	$\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$

Figure 3: Transmission schedule of consecutive relaying scheme.

Time Slot									
0	1	2	...	3t	3t+1	3t+2	...	3L-3	3L-2	3L-1
$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{a}_0\oplus\mathbf{b}_0$	...	$\mathbf{a}_t$	$\mathbf{b}_t$	$\mathbf{a}_t\oplus\mathbf{b}_t$	...	$\mathbf{a}_{L-1}$	$\mathbf{b}_{L-1}$	$\mathbf{a}_{L-1}\oplus\mathbf{b}_{L-1}$

Figure 4: Section of transmission schedule for the TDMA relaying scheme.

	0	1	...	t	t+1	...	L-1	
Frequency Band f1	$\mathbf{a}_0$	$\mathbf{a}_1$	...	$\mathbf{a}_t$	$\mathbf{a}_{t+1}$	...	$\mathbf{a}_{L-1}$	
Frequency Band f2	$\mathbf{b}_0$	$\mathbf{b}_1$	...	$\mathbf{b}_t$	$\mathbf{b}_{t+1}$	...	$\mathbf{b}_{L-1}$	
Frequency Band f3		$\mathbf{a}_0\oplus$ $\mathbf{b}_0$	...	$\mathbf{a}_{t-1}\oplus$ $\mathbf{b}_{t-1}$	$\mathbf{a}_t\oplus$ $\mathbf{b}_t$	...	$\mathbf{a}_{L-2}\oplus$ $\mathbf{b}_{L-2}$	$\mathbf{a}_{L-1}\oplus$ $\mathbf{b}_{L-1}$

Figure 5: Section of transmission schedule for the FDMA relaying scheme.

Without loss of generality, we consider the decoding of codeword $\mathbf{a}_t$ at the end of half slot $2t+2$, for $t \in \{1, 2, \cdots, L-1\}$. We focus on the codeword $\mathbf{d}_{2t+2} = \pi(\mathbf{b}_t) \oplus \mathbf{a}_t$ for this decoding. We use $L_1(c_t^{(i)})$ and $L_1(d_t^{(i)})$ to denote the respective *a priori* LLRs derived from the corresponding received streams. Also, $L_2(\cdot)$ denotes the (updated) *a posteriori* LLRs which will be used as the *a priori* LLRs in the next decoding. We assume that *a priori* LLRs on $\mathbf{c}_{2t+1}$ and $\mathbf{c}_{2t}$, denoted $\{L_1(c_{2t+1}^{(i)})\}$ and $\{L_2(c_{2t}^{(i)})\}$ respectively, are available from the channel observation value and previous decoding respectively. In addition to the decoding of $\mathbf{a}_t$, the decoder will produce *a posteriori* LLRs $\{L_2(c_{2t+1}^{(i)})\}$ which will be used as *a priori* LLRs in the next decoding.

Next, we provide a concise description of the factor graph based decoding algorithm [7] at the destination decoder. The factor graph for the decoding is illustrated in Figure 6, where circles depict variable nodes and squares depict factor nodes. Extrinsic information is exchanged between the two soft-input soft-output (SISO) decoder modules for the constituent codes $\mathcal{C}_a$ and $\mathcal{C}_b$, via the factor nodes $\{F_i\}$ which correspond to the network coding operation at the relay. For simplicity, the graph is illustrated for the case $n = 3$, and where $\mathcal{C}_a$ and $\mathcal{C}_b$ are (trivial) LDPC codes. For convolutional constituent codes, the two SISO modules execute BCJR algorithms. Note that in the convolutional case, the separation of the two (e.g. M-state) decoder SISO modules gives a complexity advantage over schemes which use a larger (e.g. M^2-state) decoder to decode the "nested" code generated at the relay (see e.g. [2]). In the LDPC case this separation of SISO modules effects a more efficient message-passing schedule than does fully parallel decoding on the Tanner graph of the nested code (see e.g. [6]).

Next we introduce some notational conventions pertaining to the following algorithm description. In all cases, the letter λ is used to denote LLRs corresponding to messages passed on the factor graph. The interleaving "π" is interpreted as

$$\mathbf{x} = \pi(\mathbf{y}) \iff x^{(i)} = y^{(\pi(i))}$$

Some index sets are defined as follows. $\mathcal{J}_a = \{1, 2, \cdots m_a\}$; $\mathcal{J}_b = \{1, 2, \cdots m_b\}$; $\mathcal{N}_a(i) = \{j \in \mathcal{J}_a : H_a(j,i) = 1\}$; $\mathcal{N}_b(i) = \{j \in \mathcal{J}_b : H_b(j,i) = 1\}$; $\mathcal{M}_a(j) = \{i \in \mathcal{I} : H_a(j,i) = 1\}$; $\mathcal{M}_b(j) = \{i \in \mathcal{I} : H_b(j,i) = 1\}$. Also, $\boxplus$ denotes the (commutative and associative) "box-plus" operation [8], i.e.

$$\boxplus_{s\in\mathcal{S}}\, \lambda_s = \log\left(\frac{1+\prod_{s\in\mathcal{S}}\tanh(\lambda_s/2)}{1-\prod_{s\in\mathcal{S}}\tanh(\lambda_s/2)}\right)$$

N denotes the maximum number of decoding iterations.

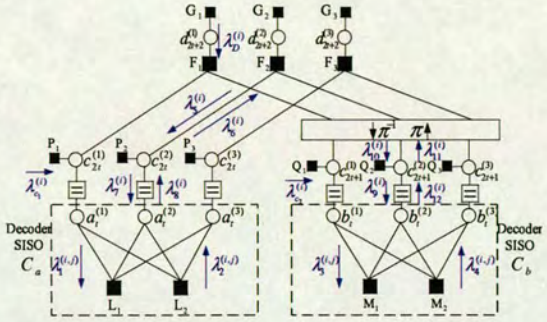

Figure 6: Factor graph corresponding to the destination's decoding of codeword $\mathbf{a}_t$. The message-passing schedule is such that extrinsic information is exchanged between the two decoder SISO modules for the constituent codes $\mathcal{C}_a$ and $\mathcal{C}_b$, via the factor nodes $\{F_i\}$ corresponding to the network coding operation at the relay. For ease of presentation, the factor graph is illustrated for the case $n = 3$ and trivial codes $\mathcal{C}_a$, $\mathcal{C}_b$.

Factor Graph Based Decoding Algorithm at Destination Node – Decoding of Codeword $\mathbf{a}_t$

Initialization:

- For $i \in \mathcal{I}$,

$$\lambda_{c_1}^{(i)} = L_2(c_{2t}^{(i)}) \quad (1)$$
$$\lambda_{c_2}^{(i)} = L_1(c_{2t+1}^{(i)}) \quad (2)$$
$$\lambda_8^{(i)} = 0 \quad (3)$$
$$\lambda_D^{(i)} = L_1(d_{2t+2}^{(i)}) \quad (4)$$

- For $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_2^{(i,j)} = 0 \quad (5)$$

- For $i \in \mathcal{I}$, $j \in \mathcal{N}_b(i)$

$$\lambda_4^{(i,j)} = 0 \quad (6)$$

Main Loop: For $k = 1$ to N do

- For $i \in \mathcal{I}$,

$$\lambda_6^{(i)} = \lambda_{c_1}^{(i)} + \lambda_8^{(i)} \quad (7)$$

- Network coding constraints: for $i \in \mathcal{I}$,

$$\lambda_{10}^{(\pi(i))} = \lambda_6^{(i)} \boxplus \lambda_D^{(i)} \quad (8)$$

- For $i \in \mathcal{I}$,

$$\lambda_9^{(i)} = \lambda_{c_2}^{(i)} + \lambda_{10}^{(i)} \quad (9)$$

- SISO decoder $\mathcal{C}_b$: for $i \in \mathcal{I}$, $j \in \mathcal{N}_b(i)$

$$\lambda_3^{(i,j)} = \lambda_9^{(i)} + \sum_{l \in \mathcal{N}_b(i) \setminus \{j\}} \lambda_4^{(i,l)} \quad (10)$$
$$\lambda_4^{(i,j)} = \boxplus_{l \in \mathcal{M}_b(j) \setminus \{i\}} \lambda_3^{(l,j)} \quad (11)$$

For $i \in \mathcal{I}$,

$$\lambda_{12}^{(i)} = \sum_{j \in \mathcal{N}_b(i)} \lambda_4^{(i,j)} \quad (12)$$

- Obtain the *a posteriori* LLR (prior for decoding in next half slot)

$$L_2(c_{2t+1}^{(i)}) = \lambda_9^{(i)} + \lambda_{12}^{(i)} \quad (13)$$

- For $i \in \mathcal{I}$,

$$\lambda_{11}^{(i)} = \lambda_{c_2}^{(i)} + \lambda_{12}^{(i)} \quad (14)$$
$$\lambda_5^{(i)} = \lambda_D^{(i)} \boxplus \lambda_{11}^{(\pi(i))} \quad (15)$$

- For $i \in \mathcal{I}$,

$$\lambda_7^{(i)} = \lambda_{c_1}^{(i)} + \lambda_5^{(i)} \quad (16)$$

- SISO decoder $\mathcal{C}_a$: for $i \in \mathcal{I}$, $j \in \mathcal{N}_a(i)$

$$\lambda_1^{(i,j)} = \lambda_7^{(i)} + \sum_{l \in \mathcal{N}_a(i) \setminus \{j\}} \lambda_2^{(i,l)} \quad (17)$$
$$\lambda_2^{(i,j)} = \boxplus_{l \in \mathcal{M}_a(j) \setminus \{i\}} \lambda_1^{(l,j)} \quad (18)$$

For $i \in \mathcal{I}$,

$$\lambda_8^{(i)} = \sum_{j \in \mathcal{N}_a(i)} \lambda_2^{(i,j)} \quad (19)$$

- Calculate *a posteriori* LLRs for codeword $\mathbf{a}_t$:

$$L(a_t^{(i)}) = \lambda_7^{(i)} + \lambda_8^{(i)} \quad (20)$$

- Make decisions on the code bits

$$\hat{a}_t^{(i)} = \begin{cases} 0 & \text{if } L(a_t^{(i)}) \geq 0 \\ 1 & \text{if } L(a_t^{(i)}) < 0 \end{cases}$$

If $\hat{\mathbf{a}}_t \mathbf{H}_a^T = \mathbf{0}$ then *break;*

Endfor

For the special case of decoding the first codeword $\mathbf{a}_0$, we set $L_2(c_0^{(\pi(i))}) = L_1(c_0^{(\pi(i))}) + L_1(d_1^{(i)})$ for all $i \in \mathcal{I}$. Also, for the special case of decoding the final codeword $\mathbf{b}_{L-1}$, further modifications are made to the algorithm as follows:

- Equation (2), (6), (7)-(14), (15) are deleted
- Equation (16) is replaced by

$$\lambda_7^{(i)} = \lambda_{c_1}^{(i)} \tag{21}$$

In this case also, the decoding of $\mathbf{a}_t$ spans three transmission frames (half slots) $2t \rightarrow 2t+1 \rightarrow 2t+2$, resulting in the three-step decoding evolution $L_1(c_{2t}^{(i)}) \rightarrow L_2(c_{2t}^{(i)}) \rightarrow \hat{a}_t^{(i)}$.

4. A LOWER BOUND ON THE OUTAGE PROBABILITY

In this section a theoretical lower bound on the frame (codeword) error rate (FER) of the system is derived; the analysis follows the lines of [3].

In the case where an extra frequency band is used, let us consider the decoding of $\mathbf{a}_t$ and $\mathbf{b}_t$, transmitted in half slots $2t$ and $2t+1$ respectively. The codewords received by the destination from the sources in half slots $2t$ and $2t+1$ contain information only relating to $\mathbf{a}_t$ and $\mathbf{b}_t$. Due to latency constraints embodied in the relay operation, the message of $\mathbf{a}_t$ is contained in $\mathbf{d}_{2t+1}$ and $\mathbf{d}_{2t+2}$; in the same way, the message of $\mathbf{b}_t$ is also contained in $\mathbf{d}_{2t+2}$ and $\mathbf{d}_{2t+3}$. The received information which may be used in decoding $\mathbf{a}_t$ and $\mathbf{b}_t$ is contained in five received codewords at the destination – transmissions from sources to destination in half slots $2t$ and $2t+1$, and transmissions from relay to destination in half slots $2t+1$, $2t+2$ and $2t+3$. Recall that the codeword length is n, and let $r = k/n$ denote the code rate of each encoder; also let $C(\gamma)$ denote the capacity of a binary-input point-to-point link with temporal SNR γ. Then the number of information bits at the destination which may be used for joint decoding of $\mathbf{a}_t$ and $\mathbf{b}_t$ is not greater than $n[C(\gamma_{AD}) + C(\gamma_{BD}) + C(\gamma_{RD1}) + C(\gamma_{RD2}) + C(\gamma_{RD3})]$, where γ_{RD1}, γ_{RD2} and γ_{RD3} (the relay-destination SNRs in consecutive half slots) are independent and identically distributed random variables, and γ_{AD} and γ_{BD} (the source-destination SNRs) are independent and identically distributed random variables if the two source-to-destination link average SNRs are equal. When this value is lower than $2k$, there is no possibility of jointly decoding both packets, and an outage event is inevitable. Thus, a lower bound for the FER is given by

$$\text{FER} \geq \frac{1}{2} \cdot \Pr[C(\gamma_{AD}) + C(\gamma_{BD}) + C(\gamma_{RD1}) + C(\gamma_{RD2}) + C(\gamma_{RD3}) < 2r] \tag{22}$$

The Monte Carlo simulated lower bound is included in the simulation results of Section 5.

5. SIMULATION RESULTS

In this section, we provide a comparison of the proposed cooperative coding scheme with two reference cooperative schemes. The first is the consecutive relaying scheme of Figure 3. Here a single LDPC decoding per codeword is sufficient for reception, where the LLRs for decoder initialization are found by adding the LLRs corresponding to the broadcast and relayed versions of the pertinent codeword. The second reference scheme is the simple TDMA/ FDMA code superposition relaying scheme of Figures 4 and 5. For this scheme, joint decoding of $\mathbf{a}_t$ and $\mathbf{b}_t$ is performed using a decoding algorithm similar to that of Section 3. The decoder also uses message passing between the two constituent decoders via network coding constraints; full details are omitted due to space limitations. The fair comparison of the three cooperative schemes is based on the constraint that in simulations, each scheme uses the same codes $\mathcal{C}_a$ and $\mathcal{C}_b$, and the same total energy E for transmission of the $2L$ source messages.

The codes used for simulations are randomly generated rate $1/2$ regular LDPC codes of block length $n = 1200$ with column weight 3 and no 4-cycles in the Tanner graph. In simulations we choose $\mathbf{H}_a = \mathbf{H}_b$, and assume BPSK modulation for all systems. A random interleaver π is used to avoid 8-cycle multiplicity in the Tanner graph. We consider a quasi-static Rayleigh fading channel, for which the fading coefficients are constant within each half slot (one codeword) and change independently from one half slot to the next. We assume equal average signal-to-noise ratio (SNR) on the two source-destination links and the relay-destination link, and we assume that the destination has perfect knowledge of the channel fading coefficients and noise variances. As for the two source-relay links, which play a key role in the performance of the system since poor link quality may lead to catastrophic error propagation at the destination decoder, the simulation setup is that the source-relay links are both ideally error-free.

Simulated performance results in terms of BER and FER are shown in Figures 7 and 8 respectively. The curve corresponding to the proposed cooperative coding scheme again exhibits an increased diversity gain with respect to the reference schemes in the SNR region of interest. The proposed scheme attains approximately an order of magnitude decrease in both BER and FER over the other schemes at an E_b/N_0 of 8 dB.

6. CONCLUSION

In this paper, we have proposed a simple but effective cooperative coding scheme for the shared-relay scenario. The scheme requires only channel coding at

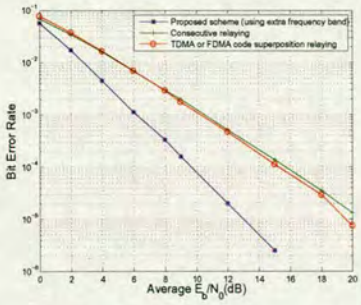

Figure 7: Comparative BER performance for the proposed cooperative scheme. The performance is shown with respect to two reference schemes: simple TDMA/FDMA code superposition relaying; and consecutive relaying using the extra frequency band.

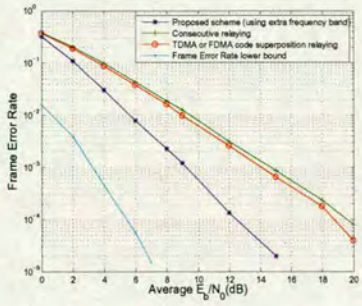

Figure 8: Comparative FER performance for the proposed cooperative scheme. The performance is shown with respect to two reference schemes: simple TDMA/FDMA code superposition relaying; and consecutive relaying using the extra frequency band. Also plotted is the theoretical lower bound on the FER given by (22).

the sources and interleaving and network coding at the relay, thus relegating complexity to the base station for uplink transmission. The decoding algorithm at the destination node is based on message passing on a factor graph corresponding to multiple received frames at the destination, and extracts spatial diversity gains in a computationally efficient manner. Simulation results demonstrate that the proposed scheme outperforms competitive schemes based on consecutive relaying and TDMA/FDMA based code superposition relaying.

7. ACKNOWLEDGEMENT

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC), grant number EP/D 002184/1, and by the Scottish Funding Council for the Joint Research Institute with Heriot-Watt University which is a part of the Edinburgh Research Partnership. The authors also would like to thank Dr. John Thompson for helpful discussions.

References

[1] A. Nosratinia, T.E. Hunter, and A. Hedayat: "Cooperative communication in wireless networks," *IEEE Communications Magazine*, pp.74-80, Oct., 2004.

[2] L. Xiao, T.E. Fuja, J. Kliewer, and D.J. Costello: "Cooperative Diversity Based on Code Superposition," *IEEE International Symposium on Information Theory (ISIT)*, 2006.

[3] L. Xiao, T.E. Fuja, J. Kliewer, and D.J. Costello: "A network coding approach to cooperative diversity," *IEEE Trans. Information Theory*, vol.53, pp.3714-3722, 2007.

[4] L. Xiao, T.E. Fuja, J. Kliewer, and D.J. Costello: "Algebraic superposition of LDGM codes for cooperative diversity," *IEEE International Symposium on Information Theory (ISIT)*, 2007.

[5] C. Hausl and P. Dupraz: "Joint Network-Channel Coding for the Multiple-Access Relay Channel," *3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks*, pp.817-822, 2006.

[6] C. Hausl, F. Schreckenbach, I. Oikonomidis and G. Bauch: "Iterative Network and Channel Decoding on a Tanner Graph," *43rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, USA*, 2005.

[7] F. R. Kschischang, B. J. Frey, and H. A. Loeliger: "Factor graphs and the sum-product algorithm," *IEEE Trans. Information Theory*, vol.47, pp.498-519, 2001.

[8] J. Hagenauer, E, Offer and L. Papke:"Iterative decoding of binary block and convolutional codes," *IEEE Trans. Information Theory*, vol.42, pp.429–445, 1996.