Detection Techniques for Alternate-Relaying Cooperative Communications Systems

By

©Hala Mostafa, M.Sc., Eng.

A Thesis

submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Memorial University of Newfoundland

May 2014

Doctor of Philosophy (2014) (Electrical and Computer Engineering)

Memorial University of Newfoundland, St. John's, NL, Canada

Title: Detection Techniques for Alternate-Relaying Cooperative Communications Systems

Author: Hala Mostafa

Supervisory Committee

Prof. Mohamed Hossam Ahmed

Prof. Octavia A. Dobre

Prof. Cheng Li

Abstract

Cooperative technology constitutes a breakthrough in the design of wireless communication systems. This is due to its relatively simple implementation and its significant performance gains in terms of link reliability, system capacity, and transmission range. In cooperative communications, multiple terminals in a network cooperate by relaying each other's information, forming a virtual antenna array, and, thus realizing spatial diversity in a distributed fashion. It is not surprising that cooperative communications have become a strong candidate for many wireless applications, such as cellular networks, wireless local area network, mobile ad-hoc networks, and wireless sensor networks. However, cooperative technology is not without challenges. A major problem in this technology is the reduction in spectral efficiency, which results from the half-duplex constraint at the relays and orthogonal relay transmission. This has spurred researchers to investigate cooperative strategies to recover the spectral efficiency loss. Such strategies can be classified into three main categories.

One category supposes that each source transmits a 'superimposed' signal, which consists of its own data and relaying information. This superposition can be performed in code or in modulation domain. Obviously, if the relay does not have its own data, a full-rate transmission can not be achieved.

The second category is to employ adaptive modulation techniques where the spectral efficiency is improved by changing modulation size with fixed symbol rate. However, the transmitter needs to know the channel signal-to-noise (SNR) such that the best suitable modulation is chosen and the receiver must be informed on the used modulation in order to decode the information.

This leads to an increased overhead in the system as compared with a fixed modulation system, and will increase the complexity of the receiver too.

The third one utilizes two-relay, which alternatively transmit and receive. A key feature of this category is that the source continues to transmit data, while the two relays take turns in receiving and transmitting the data from the source. Due to the simultaneous transmission of the data streams through both direct and one of relay channels, harmful interference occurs at the relays and destination. The interference occurred at the relays and destination represents a drawback in this case, though.

According to our best knowledge, no previous research was done to develop the optimal detectors for alternate-relaying cooperative (ARC) systems. Further, all the previous works for ARC systems have in common that they do not exploit any properties of the underlying error correcting codes. It is therefore necessary to propose optimal detection techniques for uncoded and coded two-relay systems. This motivated us to do this research. In this thesis, we proposed optimal and suboptimal detectors to mitigate the influence of the interference signal for the uncoded and coded decode-and-forward (DF) ARC systems.

Acknowledgments

During my time as a Ph.D. researcher, I had the opportunity to get to know many interesting and clever people. I would like to thank all who, in some way or other, have contributed to this work or to my great social environment.

I wish to express my sincere appreciation to my supervisors Prof. Mohamed Hossam Ahmed and Prof. Octavia A. Dobre for their invaluable guidance, encouragement, enthusiasm and friendship. Also, they greatly and sincerely have contributed a great deal to the development and production of this thesis in the best possible way. I hope that this work reflects a small portion of their abilities as excellent researchers.

I would like to thank all friends and colleagues in the wireless research group of the Faculty of Engineering and Applied Science at Memorial University for the help I have received during more than four years in the group and for a pleasant working atmosphere.

I wish to convey warmest thanks to my parents who have given me endless support, and provided me with the opportunity to reach this far with my studies.

I am indebted to my husband and my children for their continuous love, encouragement, and patience all the way.

Finally, the financial support provided to me by the Arab Republic of Egypt is gratefully acknowledged.

Thank you all!

Hala Mostafa

Contents

List of Figures

Abbreviations and Acronyms

AF	Amplify and Forward
APP	A Posteriori Probability
ARC	Alternate-Relaying Cooperative
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BCJR	Bahl, Cocke, Jelinek, and Raviv
BICM-ID	Bit Interleaved Coded Modulation - Iterative Decoding
BW	Bandwidth
CDMA	Code Division Multiple Access
DF	Decode and Forward
FDMA	Frequency Division Multiple Access
IEEE	Institute of Electrical and Electronics Engineers
LAN	Local Area Network
L-value	Likelihood value
MAP	Maximum A Posteriori
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
MRC	Maximum Ratio Combining

MS	Memory Size
SISO	Soft Input Soft Output
SNR	Signal to Noise Ratio
TDMA	Time Division Multiple Access
TCM	Trellis Coded Modulation
WLAN	Wireless Local Area Networks
WPAN	Wireless Personal Area Networks

Notations

x^*	the complex conjugate of x
$\widehat{x}$	the detected value of x
$\Re(x)$	the real part of x
$\Im(x)$	the imaginary part of x
$\mathrm{E}[\mathrm{x}]$	the expected value of random variable x
$\lvert x \rvert$	the absolute value of x
σ_x^2	the variance of random variable x
$\Pr(x \mid y)$	the probability density function of occurrence of x given y
$\arg\max_x f(x)$	the value of x that maximizes $f(x)$
$\mathbf{x}$	the vector $\mathbf{x}$
$\mathbf{x}^{\mathrm{T}}$	the transpose of $\mathbf{x}$
$\mathbf{x}^{\mathrm{H}}$	the complex conjugate transpose of $\mathbf{x}$ (Hermitian transpose)
$\lVert\mathbf{x}\rVert$	the norm of the vector $\mathbf{x} = \sqrt{\sum_k x_k^2}$, x_k is the kth element in the $\mathbf{x}$
j	$\sqrt{-1}$
Ω	the signal constellation

Publication

The research described in this thesis published in 7 refereed scholarly journal and conference papers.

Journals

1. H. Mostafa et al., "Simplified Maximum Likelihood Detection Schemes For Full-Rate Cooperative Systems", IET Communications, vol. 7, no. 17, pp. 1899-1906, Nov. 2013.

2. H. Mostafa et al., "Decoding Techniques for Alternate-Relaying BICM Cooperative Systems" EURASIP Journal on Wireless Communications and Networking, pp. 1-13, Sept. 2013.

3. H. Mostafa et al., "BICM Alternate-Relaying Cooperative Systems with Multiple Antennas at the Destination" Submitted to IEEE Transactions on Vehicular Technology.

Conferences

1. H. Mostafa et al., "Maximum-Likelihood Detectors for Full-Rate Cooperative Communication Systems", IEEE ICC, Kyoto, Japan, June 2011.

2. H. Mostafa et al., "Detection Techniques for Two-Relays Decode and Forward Cooperative Systems", IEEE GLOBECOM, Houston, Texas, USA, Dec., 2011.

3. H. Mostafa et al., "Decoding Techniques for Coded Full-Rate Decode and Forward Cooperative Transmission ", IEEE ICC, Ottawa, Canada, June 2012.

4. H. Mostafa et al., “Receiver Design For Alternate-Relaying Cooperative Systems with Multiple Antennas at the Destination”, IEEE GLOBECOM, Atlanta, GA., USA, Dec., 2013.

5. H. Mostafa et al., “BICM Alternate-Relaying Cooperative Systems with Multiple Antennas at the Destination” Accepted in IEEE ICC, Sydney, Australia, 2014.

Chapter 1

Introduction

Wireless communications have seen a remarkably fast technological evolution. Although separated by only a few years, each new generation of wireless devices has brought significant improvements in terms of link communication speed, device size, battery life, applications, etc. One of the major challenges for reliable transmission in wireless communications is the random and unpredictable nature of the channel. This nature results in a phenomenon known as fading. Fading is mainly determined by its multipath nature due to the reflections of the signals on natural and man-made objects. Thus, at the receiver antenna, multiple of replicas of the transmitted signal arrive over different paths from different directions with different amplitudes and phases. The superposition of these replicas results in amplitude and phase variations of the composite received signal. In addition, as the transmitter, the receiver or the objects may move at a relative speed with respect to each other, the paths change with time. This turns a good reception into a signal breakdown due to a sudden extreme attenuation and vice versa. Compensating for the adverse effects of fading is therefore critical to the design of wireless communications systems with guaranteed performance.

Diversity is an effective way to overcome the problem of fading in wireless channels. This improves transmission performance by making use of more than one independently faded version of the transmitted signal. If several replicas of the signals are transmitted over multiple channels that exhibit independent fading, the probability that all signals experience deep fading simultaneously is significantly reduced. Diversity can be realized in different ways, including time, frequency, and spatial diversity [1–5].

Time diversity is a diversity technique where identical signals are transmitted during different time slots. These time slots are uncorrelated, i.e., the temporal separation between those slots is greater than the coherence time of the wireless channel[1]. The main shortcoming of this technique is that the redundancy is provided in the time domain with a penalty of a loss in spectral efficiency.

Frequency diversity uses several carrier frequencies, that are separated by at least the coherence bandwidth of the channel, to transmit the same signal [6, 7].[2] Consequently, the carrier frequencies are uncorrelated, i.e., they do not experience the same fades. Similar to time diversity, in frequency diversity, the redundancy is provided in the frequency domain with the penalty of a loss in the spectral efficiency. Additionally, the structure of the receiver is complicated, as it must be able to work with a number of frequencies.

Spatial diversity can be achieved with multiple-input multiple-output (MIMO) systems by employing multiple antennas at the transmitter and/or receiver side. MIMO systems have been widely acknowledged and adopted in various wireless standards [3, 4, 8, 9].

[1]The coherence time is the period of time after which the correlation function of two samples of the channel response taken at the same frequency drops below a certain predetermined threshold [6, 7].

[2]The coherence bandwidth is a statistical measure of the range of frequencies over which the channel passes all the spectral components with approximately equal gain and linear phase [6, 7].

Although MIMO systems are an attractive solution to obtain diversity without incurring any loss of the scarce resources, the use of them on compact terminals, such as cell phones and wireless sensor nodes, is impractical due to size, power, and price limitations. In order to overcome these limitations, a new form of realizing spatial diversity has been recently introduced under the name of cooperative diversity [5, 10–14]. The basic idea behind cooperative diversity is that the signal transmitted by the source node is overheard by other nodes, which can be defined as relays. The source and relays can jointly process and transmit their information, creating a virtual antenna array, although each of them is equipped with only one antenna. Similar to physical antenna arrays, this virtual antenna array combats fading in wireless channels by providing the receiver with essentially redundant signals over independent channels that can be combined to average individual channel effects. There is no doubt that cooperative diversity has found applications in various networks such as cellular, ad-hoc, and sensor networks (see [15–17] and the references therein). For this wide applicability of cooperative relaying, standardization bodies have also started taking interest into it. Several task groups within the IEEE, especially, IEEE 802.11s [18], IEEE 802.15.3 [19], IEEE 802.15.4 [20], IEEE 802.15.5 [21], and IEEE 802.16 [22] along with [23] are contributing toward the integration of cooperative techniques in various wireless networks.

1.1 Overview of Cooperative Communications

In a typical wireless communication network, such as a cellular radio network, the number of active nodes is much smaller than the idle ones. Thus, it is of interest to invite idle nodes to cooperate as relays, and to exploit them as an extra dimension to further enhance the communication reliability. Generally, there are two kinds of

modes: full-duplex and half-duplex.

In a full-duplex mode, a relay transmits and receives simultaneously in the same band; however, the transmitted signal interferes at the relay with the received signal. In theory, it is possible for the relay to cancel out the interference because it knows the transmitted signal. However, in practice, a small error in interference cancellation can be fatal because the transmitted signal is typically 100-150 dB stronger than the received signal, as indicated in [11]. This error results from inaccurate knowledge of the device characteristics or from the effects of quantization and finite-precision processing. Therefore, the full-duplex mode is not commonly used.

In a half-duplex mode, a relay cannot simultaneously transmit and receive. In other words, the source and relay transmissions are orthogonal in order to eliminate any potential interference. Orthogonality can be in time-domain, in frequency domain, or using any set of signals that are orthogonal over the time-frequency plane. A major problem of the half-duplex relaying mode is the reduction in spectral efficiency [11].

To clarify the basic principles of half-duplex cooperative transmission, consider a communication system as shown in Figure 1.1, in which there are N relays that cooperate in the communication between source node S and destination node D. The communication process between the source and destination is split into two phases.[3] In the first phase, the source transmits the signal to the destination and the relays. In the second phase, N relays deal with the received signal and transmits it to the destination. The second phase of transmission requires N time slots to guarantee orthogonal transmission between relays, and then time division multi-

[3]It is worthy to mention that there are many other cooperative strategies in the literature, such as cooperative space-time coding and cooperative beamforming [5, 10–14]. In this thesis, we focus on cooperative diversity.

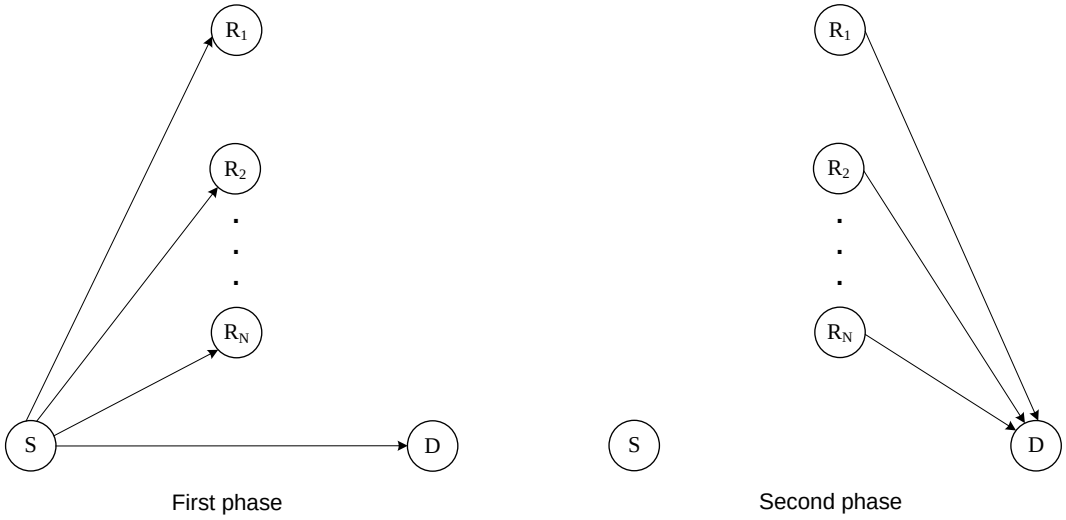

Figure 1.1: Cooperative transmission.

ple access (TDMA) is required. This is because of the half-duplex constraint, which is the practical perspective that the relays can not simultaneously transmit and receive. Furthermore, frequency division multiple access (FDMA) or code division multiple access (CDMA) could also be used so that the relays can transmit simultaneously to the destination, however at the cost of bandwidth expansion. This implies that relays introduce a delay, i.e., the source uses only one time slot from N+1 time slots when TDMA is used or uses one time slot from 2 time slots when FDMA or CDMA is employed at the second phase. This leads to an decrease in the instantaneous spectral efficiency.

The processing at the relay differs according to the employed relaying strategies. Basically, the two relaying strategies analyzed in the literature are: amplify and forward and decode and forward.

Amplify-and-Forward (AF) is one of the simplest relay strategies. In AF, each relay simply forwards to the destination a scaled version of the received signals, including both information and noise. By properly combining received signals from the source and relays, the destination node makes a final decision. Since the des-

tination receives multiple copies of signals transmitted from the source and relays through multiple independent paths, spatial diversity can always be achieved by the AF strategy at high signal-to-noise ratios (SNRs). Obviously the major drawback of AF protocols is noise amplification at the relays.

Decode-and-Forward (DF) is another commonly used strategy for eliminating the noise effect, especially for coded systems. The relay decodes the received signals and re-encodes them before forwarding to the destination. When the channel quality in the link between the source and relay is good, the process of decoding and re-encoding provides more powerful error correcting capabilities than AF. However, when the link from the source to the relay suffers from deep fading, decoding errors may occur at the relay. In this case, if the relay re-encodes these incorrect bits, error propagation will occur and lead to even worse performance.

As indicated in Figure 1.1 and the corresponding discussion, we recall that the source remains idle during the second phase of the transmission where relays deal with the received signal and transmits it to the destination. This leads to a a significant reduction of the spectral efficiency. Recently, adaptive modulation [24–27], superposition modulation [28–31], non-orthogonal relaying [32,33], and alternate-relaying [34–36] are considered to be promising cooperative strategies to improve spectral efficiency of cooperative systems.[4]

In adaptive modulation, the spectral efficiency is improved by changing modulation size with fixed symbol rate, e.g., [24–27]. The use of adaptive modulation in the context of AF cooperative has been studied in [24,25,37], where adaptive modulation is performed only at the source. Alternatively, DF allows the source and relays to employ different modulation schemes, and, as a consequence, increases

[4]The best relay scheme also improves the spectral efficiency through using only two channels instead of $N+1$ channels. The best relay is the relay that achieves the highest instantaneous channel gain to the destination.

the degree of freedom in multi-link adaptation [26,27]. Generally, the major disadvantage of adaptive modulation techniques is that the transmitter needs to know the channel signal-to-noise (SNR); such that the best suitable modulation is chosen and the receiver must be informed on the used modulation in order to decode the information. This leads to an increased overhead in the system as compared with a fixed modulation system, and will also increase the complexity of the receiver.

The superposition modulation is applicable to the scenario where two nodes are willing to transmit to a common destination and wish to get help from each other to provide better performance [28–31]. The main idea of superposition modulation is that each node in the network transmits a superimposed signal containing its own data and an overheard version of the other source data, which can improve spectral efficiency of cooperative system.

In the non-orthogonal relaying (e.g., [32,33]), the source is active all the time. In the first half of the transmission interval, the source sends data to a relay and destination. However, since the relay is assumed to be half-duplex, the relay does not receive what the source transmits in the second half of the transmission interval. This results in a reduction in the diversity order of the system. Furthermore, an additional processing is required at the destination in order to separate the signals received simultaneously in the second half of the transmission interval.

Alternate-relaying cooperative (ARC) transmission protocols are introduced (e.g., [34–36, 38–43]), in which the source communicates with the destination via two relays. The basic idea behind these protocols is to use two successive relays to mimic a full-duplex relay. More specifically, at any time slot, the source sends its information to one of the relays and the destination, while the other relay forwards the information received from the source in the previous time slot to the destination. In this way, the source can continuously transmit data and hence the

spectral efficiency loss is recovered. Recently, alternate-relaying transmission protocols have been further extended to multiple relay and multiple user scenarios (e.g., [44–48]). As being related to the direction of the thesis, next Section focuses on the literature review related to ARC systems.

1.2 Overview of BICM-ID

Bandwidth-efficient bit-interleaved coded modulation (BICM) has been considered as a state-of-the art coded modulation technique to extract time diversity over wireless fading channels [49]. The basic idea behind BICM is to concatenate a channel encoder, such as a linear block code or a convolutional code, and a random bit interleaver spanning over several coherence time intervals. By doing so, a diversity gain that is equal to the minimum binary Hamming distance of the channel code can be obtained. In BICM, the channel encoder and the modulator can be designed independently, offering more flexibility in system design. The performance of BICM can be further improved by applying the iterative decoding (ID) approach, which results in BICM with iterative decoding (BICM-ID) systems [50–53]. At a comparable decoding complexity, it is well known that BICM-ID systems can outperform trellis coded modulation (TCM) systems [54–57] over both faded and non-faded channels. It is worth noting that the BICM-ID technique has already been adopted by several wireless communications standards such as HIPERLAN/2, IEEE 802.11a/g (WLAN), IEEE 802.16e (mobile WiMAX), 3GPP/HSDPA, DVB-S2, and DVBT [58].

Typically, to further improve the performance of cooperative systems, joint signal design and error coding codes are performed at the source and relays. One form of coded cooperative transmission can be implemented by transmitting the

overall codeword in a distributed manner where different parts of the codeword are transmitted by relays through independent wireless links (e.g. [14]). This creates additional degrees of freedom. However, this poses challenges in code construction. How to combine ARC systems with BICM transmission in an effective manner is the objective of this research.

1.3 Related Work

The vast majority of research in ARC transmission protocols focuses on information-theoretic analysis to assess achievable rates, capacity bounds, and diversity-multiplexing tradeoff (e.g., [38, 40, 42, 43, 59–63]). However, the major issue associated with these protocols is how to cancel or utilize the interference, which is caused by simultaneous transmission of the source and one of the relays at the same time, in a simple way. Basically, the method used to treat interference at the relays and destination depends on whether the relays employ AF or DF relaying strategies.

For AF alternate-relaying protocols, the authors in [38] propose successive decoding at the destination with partial cancellation of inter-relay interference. The authors in [40] introduce an inter-relay self interference cancellation detector, where the cancellation is performed at one of the relays; however the detection process at the destination requires high computational complexity. Recently, the authors in [35, 39] propose a full inter-relay interference detector, where the cancellation is performed at the destination in a simple manner; however, noise accumulation associated with the detector limits the overall system performance. Generally, the deployment of AF relaying strategy in alternate-relaying cooperative systems is challenging, since the interference and noise accumulation which

results from the inter-relay link degrade the overall performance. Accordingly, the detection process at the destination has to be associated with sophisticated interference cancellation detectors. This increases the computational complexity significantly, as shown in [40]. Due to its symbol-by-symbol decision base, DF is an attractive relaying strategy to avoid interference accumulation at the destination for alternate-relaying cooperative systems[5].

For DF alternate-relaying protocols, it is usually assumed that the inter-relay link is either sufficiently weak and then the interference can be treated as extra noise, or sufficiently strong and the interference can be canceled through successive interference cancellation at the relay[6] [38, 59]. However, these two extreme scenarios may not always occur in practical systems. In [60,62], dirty paper coding based on interference pre-subtraction at the source is proposed to cancel the inter-relay interference. However, this requires high computational complexity and full knowledge of channel state information of all links at the source, which is not easy to achieve in practice. Beam-forming/smart antennas [34] and code-division multiple access (CDMA) techniques [41] are also proposed to eliminate the interference at the relays and destination. The former technique comes at the cost of complexity, whereas the latter comes at the cost of bandwidth. In [42], the authors propose employing multiple antenna at the relays to cancel out the inter-relay interference. However, implementing multiple antennas at the relays is impractical for some wireless applications due to size, power, and cost constraints. The authors in [43] rotate the signal constellation, such that there are no two symbols having the same real or imaginary value. Real and imaginary dimensions are then assigned for the direct and relaying paths, respectively to achieve full interference cancellation at

[5]DF relaying strategy can be seen simply as a particular non-linearity version of AF.

[6]In the successive interference cancellation detection, the strongest signal is detected first, and then its contribution is subtracted from the received signal before detecting the other signal.

the relays. However, the direct link is not considered and the rotated constellation causes a dramatic decrease in the bit-error-rate (BER) performance when compared with that of the original constellation.

1.4 Motivation and Objective of the Thesis

To the best of our knowledge, all the reported works for ARC systems have the following in common:

- they do not provide the optimal data detection solution;
- they restrict themselves to uncoded transmission and do not exploit any properties of the underlying error correcting codes;

These works motivated us to do this research. We can summarize the objective of the thesis in the following main point:

- To propose novel data detection algorithms for uncoded and coded DF-ARC systems, with a destination node supporting single and multiple antenna.

1.5 Organization of the Thesis

The thesis is organized into five chapters.

Chapter 2 describes the basic principles that are used throughout the thesis. It begins with the fundamentals of detection theory. In addition, a brief overview of convolutional codes with particular focus on Viterbi and BCJR algorithms as powerful tools for decoding, is presented. Finally, the bit interleaved coded modulation iterative decoding (BICM-ID) scheme is also highlighted.

Chapter 3 develops novel detection algorithms for uncoded ARC-DF systems. It begins with the transmission model for uncoded ARC-DF systems. Then, the maximum-likelihood (ML) detector at the relays is proposed. The optimal detector at the destination supporting one antenna is also developed. Furthermore, a sub-optimal one is introduced. In addition, this chapter discusses the receiver design of a destination node supporting multiple antenna for uncoded ARC-DF systems. Finally, the performance of the proposed detectors is verified through computer simulations.

Chapter 4 develops novel detection algorithms for BICM ARC-DF systems. It begins with the transmission model for coded ARC-DF systems. Then, a novel detector at the relays is proposed.The optimal detector at the destination supporting one antenna is developed. Furthermore, a sub-optimal one is also introduced. In addition, this chapter discusses the receiver design of a destination node supporting multiple antenna for BICM ARC-DF systems. Finally, the performance of the proposed detectors is verified through computer simulations.

The thesis is concluded in Chapter 5 and some topics for future research are addressed.

Chapter 2

Basic Principles

As mentioned in Chapter one, the thesis focuses on proposing new detection algorithms for uncoded and coded alternate-relaying cooperative (ARC) systems. As a prerequisite the reader should be familiar with the fundamentals of the wireless digital communications, such as signal detection theory and error-correcting concepts. This chapter provides the basic principles that are used throughout the thesis. The fundamentals of signal detection theory are presented in Section 2.1. Comprehensive coverage of convolutional codes, including encoders, the structural properties of convolutional codes, and decoders is provided in Section 2.2. Bit-interleaved coded modulation iterative-decoding (BICM-ID) schemes are reviewed in Section 2.3. Finally, the chapter is summarized in Section 2.4.

2.1 Signal Detection Theory

Signal detection theory is concerned with the analysis of received signals to determine the presence or absence of the signals of interest, to classify the signals present, and to extract information included in these signals. Generally, we are

given a finite set of possible hypotheses about an experiment, along with observations related statistically to the previous hypotheses, and the theory provides rules for making best decisions (according to some performance criterion) about which hypothesis is likely to be true. The field of signal detection covers a broader range of applications namely: radar, image processing, digital communications, etc.

In digital communications, the hypotheses are the possible messages, and the observables are the outputs of a probabilistic channel. Suppose we have M hypotheses (signals), labeled by S_i, $i = 0, 1, \cdots, M-1$, associated with a probabilistic experiment. We also adopt prior probabilities on the hypotheses, denoted $\Pr(S_i)$. Based on the observation R, the decision maker produces $\hat{S}_i$. We are interested in the best decision-making algorithm in sense of minimizing the probability of decision error, $\Pr\left(\hat{S}_i \neq S_i\right)$. There are basically two criteria that are widely applied to minimize $\Pr\left(\hat{S}_i \neq S_i\right)$: the maximum likelihood (ML) criterion and the maximum a posteriori (MAP) criterion [64, 65]. MAP criterion exploits priori information on the hypotheses. ML criterion is suited to problems where a priori information is missing. Since some prior information about the hypotheses is incorporated in the MAP criterion, the MAP criterion can improve the detection accuracy as compared to the ML criterion.

MAP criterion

MAP criterion maximizes the a posteriori probability:

$$\hat{S}_i = \arg\max_{S_i} \Pr(S_i|R). \tag{2.1}$$

Using Bayes' rule, we can write

$$\Pr(S_i|R) = \frac{\Pr(R|S_i)\Pr(S_i)}{\Pr(R)}. \tag{2.2}$$

However, the denominator on the right-side in (2.2) does not involve i, and maximizing (2.2) is equivalent to maximizing $\Pr(R|S_i)\Pr(S_i)$.

ML criterion

When $\Pr(S_i)$ is unknown, MAP criterion is not possible and ML is then performed. In ML, we maximize the likelihood function

$$\hat{S}_i = \arg\max_{S_i} \Pr(R|S_i). \tag{2.3}$$

Observe that when S_i has a uniform a priori distribution over its entire domain, the MAP criterion reduces to the ML.

Because it is the maximizing value (i.e. the argument) that is important in ML/MAP criterion, and not the value of the maximum (the function) itself, it is common to ignore or suppress constants in the likelihood/a posteriori function that do not depend upon the parameter set. Also, it is more convenient to consider the logarithm of the likelihood/a posteriori function, called log-likelihood/ log-a posteriori function. Since the logarithm is monotonically increasing, the maximizing the log-likelihood/log-a posteriori function is equivalent to maximizing the likelihood/a posteriori function.

2.2 Convolutional Codes

Convolutional codes have been widely used in applications such as space and satellite communications, cellular mobile, digital video broadcasting, etc [66, 67]. Further, convolutional codes are the building blocks for other important classes of error correcting codes such as turbo codes and trellis coded modulation [54–57, 68, 69]. Their popularity stems from their simple structure and availability of easily implementable maximum likelihood decoding.

Convolutional codes have a long history. Elias first introduced them in 1955 [70] as an alternative to block codes. He showed that redundancy could be introduced into a data stream through the use of a linear shift register. In 1967 Viterbi discovered an approach for decoding convolutional codes which he showed to be "asymptotically optimal" [71, 72]. In 1973 Forney [73] showed that what is now known as the Viterbi algorithm is actually a maximum-likelihood decoding algorithm for convolutional codes.

In the following sub-sections we consider a basic structure of convolutional codes. An emphasis is placed on decoding algorithms including the Viterbi and BCJR algorithms. For more details, the interested reader is encouraged to consult texts such as [67, 74].

2.2.1 Convolutional Encoders

We will use the rate 1/2 linear convolutional encoder of Figure 2.1 as a running example to illustrate the insights into the fundamental structures of these encoders. The rate of this encoder is established by the fact that the encoder outputs two bits

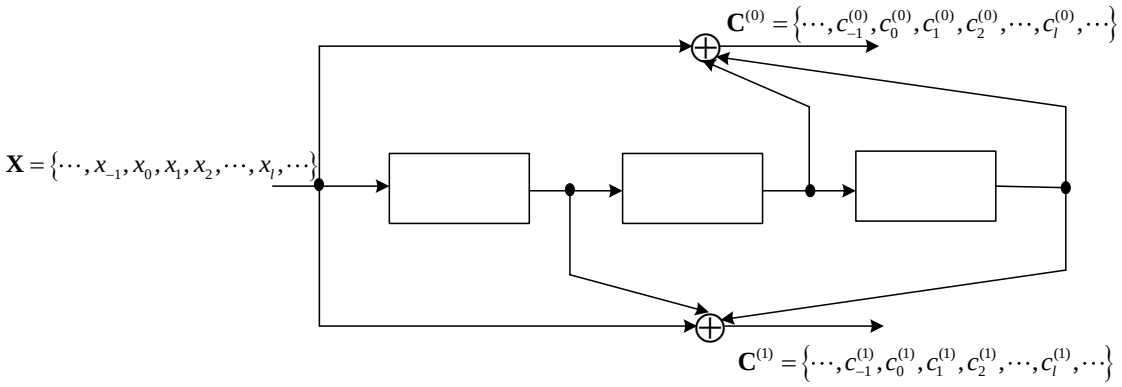

Figure 2.1: A rate-1/2 convolutional encoder.

for every input bit. In Figure 2.1 the binary data stream

$$\mathbf{X} = \{\ldots, x_{-1}, x_0, x_1, x_2, \ldots, x_l, \ldots\}, \tag{2.4}$$

is fed into a shift register circuit consisting of a series of memory elements, where l is a time instant. With each successive input to the shift register, the values of the memory elements are tapped off and added according to a fixed pattern, creating a pair of output coded data streams

$$\mathbf{C}^{(0)} = \left\{\ldots, c_{-1}^{(0)}, c_0^{(0)}, c_1^{(0)}, c_2^{(0)}, \ldots, c_l^{(0)}, \ldots\right\}, \tag{2.5}$$

and

$$\mathbf{C}^{(1)} = \left\{\ldots, c_{-1}^{(1)}, c_0^{(1)}, c_1^{(1)}, c_2^{(1)}, \ldots, c_l^{(1)}, \ldots\right\}, \tag{2.6}$$

where we assume that $c_l^{(i)}$, $i = 0, 1$ are binary symbols. These output streams can be multiplexed by a commutator to create a single coded data stream

$$\mathbf{C} = \left\{\ldots, c_0^{(0)} c_0^{(1)}, c_1^{(0)} c_1^{(1)}, c_2^{(0)} c_2^{(1)}, \ldots, c_l^{(0)} c_l^{(1)}, \ldots\right\}, \tag{2.7}$$

which is the convolutional codeword.

Each element in the output streams $\mathbf{C}^{(0)}$ and $\mathbf{C}^{(1)}$ is a linear combination of the elements in the input stream $\mathbf{X}$. At time l the input to the encoder is x_l and the output is $\mathbf{c}_l = \left\{c_l^{(0)} c_l^{(1)}\right\}$, where

$$c_l^{(0)} = x_i + x_{i-2} + x_{i-3}, \tag{2.8}$$

and

$$c_l^{(1)} = x_i + x_{i-1} + x_{i-3}, \tag{2.9}$$

where $+$ denotes modulo-2 addition.

For example, the encoder in Figure 2.1 is used to encode the information sequence $\mathbf{X} = \{10110\}$. We obtain $\mathbf{C}^{(0)} = \{10001010\}$ and $\mathbf{C}^{(1)} = \{11111110\}$, and the convolutional codeword is then $\{11, 01, 01, 01, 11, 01, 11, 00\}$. Here we assume that the shift-register contents are initialized to zero before the encoding process begins.

The connections between the shift register elements and the modulo-2 adders can be conveniently described by the following two generator sequences,

$$\begin{aligned} \mathbf{g}^{(0)} &= \left\{g_0^{(0)} g_1^{(0)} g_2^{(0)} g_3^{(0)}\right\} = \{1011\}, \\ \mathbf{g}^{(1)} &= \left\{g_0^{(0)} g_1^{(0)} g_2^{(0)} g_3^{(0)}\right\} = \{1101\}, \end{aligned} \tag{2.10}$$

where $\mathbf{g}^{(0)}$ and $\mathbf{g}^{(1)}$ represent the lower connections, respectively, with the leftmost entry being the connection to the leftmost stage. The term convolutional codes comes from the observation that the ith output sequence $\mathbf{C}^{(i)}$, $i = 1, 2$, represents

the convolution of the input sequence and the ith generator sequence,

$$\mathbf{C}^{(i)} = \mathbf{X} * g^{(i)}, \tag{2.11}$$

where * denotes the convolution operator.

As shown in Figure 2.1, the output, $c_l^{(i)}$, is affected by x_l and the the past three information bits. In general, the constraint length of a convolutional code is defined as the number of shifts over which a single information bit can influence the encoder output. In other way, the constraint length is usually taken to be the length of the input shift register plus one [67,74].

2.2.2 The Structural Properties of Convolutional Codes

Traditionally, the structure properties of convolutional codes are portrayed in graphical form by using state and trellis diagrams.

State Diagram

Since the convolutional encoder is a sequential circuit, its operation can be described by a state diagram. The state of the encoder is defined as the shift register contents. The encoder shown in Figure 2.1 contains three binary memory elements that can collectively assume any one of eight possible states. Designate these states $\{S_0, S_1, \ldots, S_7\}$ and associate them with the contents of the memory elements as shown below

$$\begin{aligned} S_0 &\longleftrightarrow (000) & \qquad S_4 &\longleftrightarrow (001) \\ S_1 &\longleftrightarrow (100) & \qquad S_5 &\longleftrightarrow (101) \\ S_2 &\longleftrightarrow (010) & \qquad S_6 &\longleftrightarrow (011) \\ S_3 &\longleftrightarrow (110) & \qquad S_7 &\longleftrightarrow (111). \end{aligned}$$

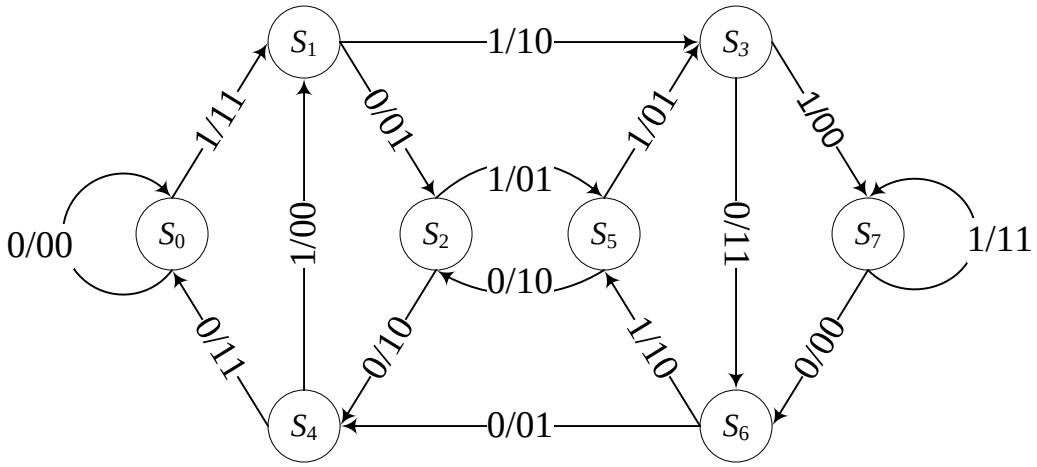

Figure 2.2: State diagram of the encoder shown in Figure 2.1.

The encoder moves between states in a deterministic manner. Given a current state $(\lambda\mu\nu)$, the next state can be either $(0\lambda\mu)$ corresponding to a zero input or $(1\lambda\mu)$ corresponding to a one input, where $\lambda, \mu, \nu \in \{0, 1\}$. The state diagram for the encoder in Figure 2.1 is shown in Figure 2.2. Each branch in the state diagram has a label of the form $x/c^{(0)}c^{(1)}$, where x is the input bit that causes the state transition and $c^{(0)}c^{(1)}$ is the corresponding pair of output bits.

For example: The information sequence $\mathbf{X} = \{101011\}$ is encoded using the encoder shown in Figure 2.1 and its state diagram shown in Figure 2.2. This information sequence designates the path $S_0, S_1, S_2, S_5, S_2, S_5, S_3$ through the encoder state diagram in Figure 2.2. The corresponding code word is thus $\{11, 01, 01, 10, 01, 01\}$. We start encoding with a zero state encoder and the end state must be known, so additional m zero bits are added to the end of the information to drive the encoder to state zero. These additional zero bits are called the tail.

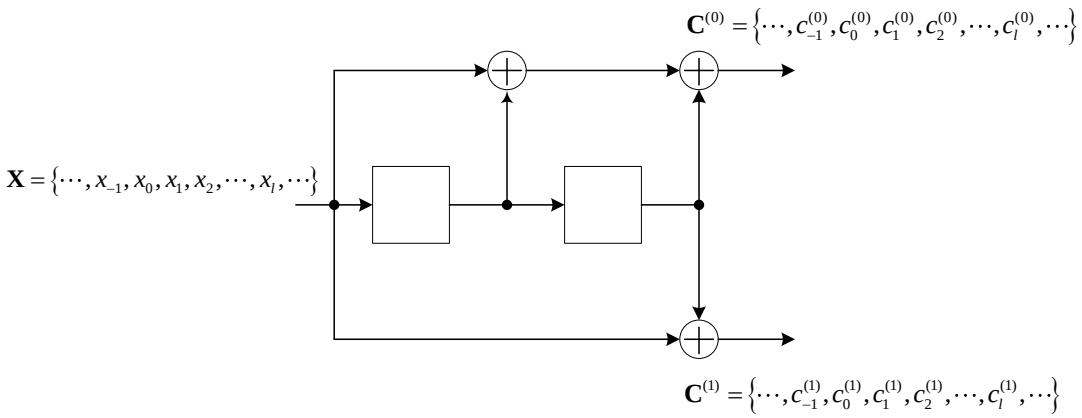

Figure 2.3: A rate-1/2 convolutional encoder with $m = 2$.

Trellis Diagram

A trellis diagram is an extension of a convolutional code's state diagram that explicitly shows the passage of time. Consider the rate-1/2 convolutional encoder with two shift registers, m = 2, and generator sequences $g^{(0)} = \{111\}$ and $g^{(1)} = \{101\}$ shown in Figure 2.3. The state diagram of this encoder is shown in Figure 2.4. In Figure 2.5 the state diagram is extended in time to form a trellis diagram. The branches of the trellis diagram are labeled with the output bits corresponding to the associated state transitions, with the lower branch corresponding to the zero-input bit, and the upper branch corresponding to the one-input bit.

Every codeword in a convolutional code is associated with a unique path, starting and stopping at state S_0, through the associated trellis diagram. Consider a rate $1/n$ binary convolutional encoder with total memory m. The associated trellis diagram has 2^m states at each stage, or time increment t. There are 2 branches leaving each node, one branch for each possible combination of input values.

It is assumed that after the input sequence has been entered into the encoder, m state transitions are necessary to return the encoder to state S_0. Given an input sequence of L' bits, the trellis diagram must have $(L' + m)$ stages, the first and last

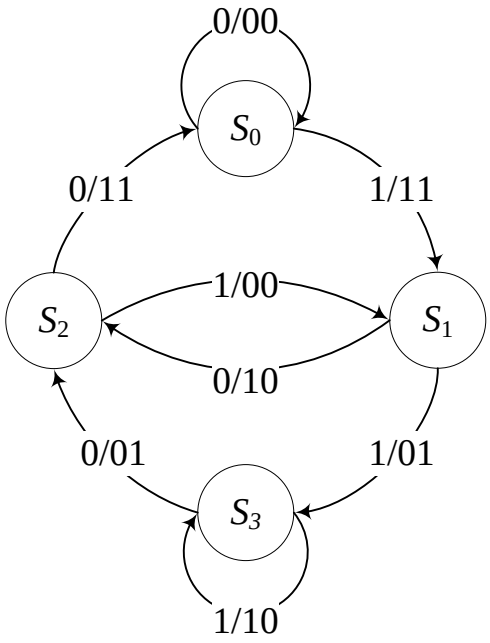

Figure 2.4: State diagram of the encoder shown in Figure 2.3.

stages starting and stopping, respectively, in state S_0. There are thus $2^{L'}$ distinct paths through the general trellis, each corresponding to a convolutional codeword of length $n(L' + m)$.

For example: An input sequence of length $L' = 3$, $\mathbf{X} = \{101\}$ is shown in Figure 2.6 to correspond to a five-branch path associated with the $2(3 + 2) = 10$ bits convolutional codeword $\mathbf{C} = \{11, 10, 00, 10, 11\}$.

2.2.3 Viterbi Decoding

Let $\mathbf{X} = \{x_0, x_1, \cdots, x_L\}$ denote a message vector with length L' including tail bits, $\mathbf{C} = \left\{c_0^{(0)}c_0^{(1)}, c_1^{(0)}c_1^{(1)}, \cdots, c_{L'}^{(0)}c_{L'}^{(1)}\right\}$ denote the corresponding codeword and $\mathbf{R} = \left\{r_0^{(0)}r_0^{(1)}, r_1^{(0)}r_1^{(1)}, \cdots, r_{L'}^{(0)}r_{L'}^{(1)}\right\}$ is the discrete received vector. Here we consider a rate $1/2$ convolutional code, generalization to any other rate is straightforward. Given $\mathbf{R}$, the decoder is required to make an estimate $\hat{\mathbf{X}}$ of the message

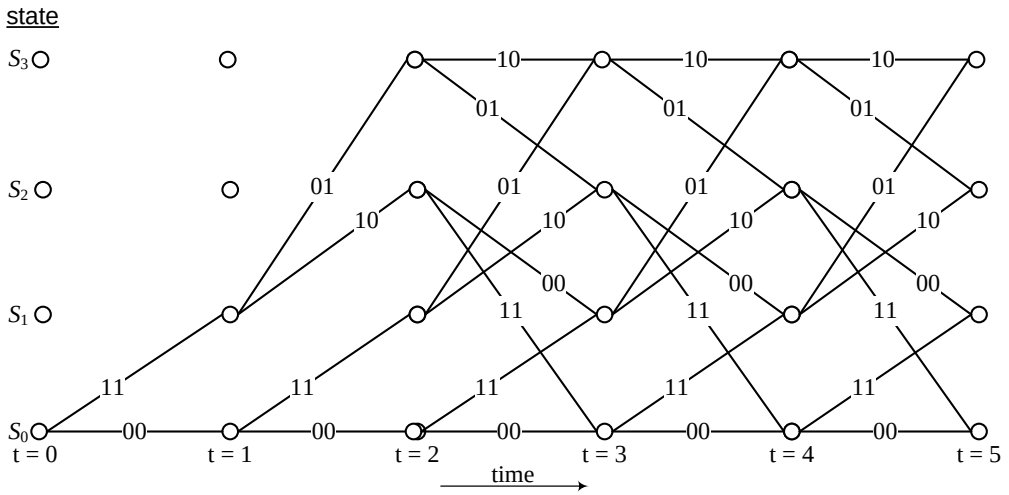

Figure 2.5: Trellis diagram for the encoder in Figure 2.3.

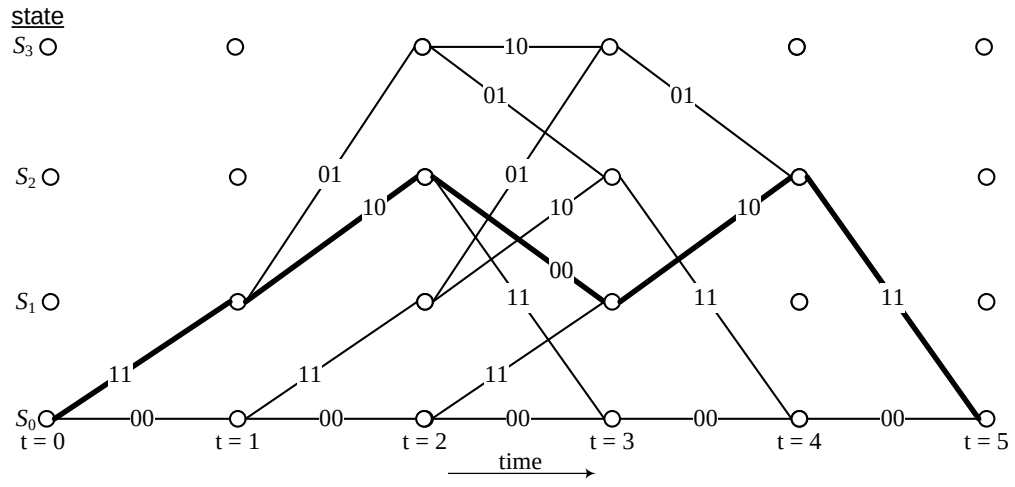

Figure 2.6: Trellis diagram for inputs of length 3 to the encoder in Figure 2.3.

vector. Since there is one-to-one correspondence between the message vector $\mathbf{X}$ and the codeword $\mathbf{C}$, the decoder may equivalently produce an estimate $\hat{\mathbf{C}}$ of the codeword. We put $\hat{\mathbf{X}} = \mathbf{X}$ if and only if $\hat{\mathbf{C}} = \mathbf{C}$. Otherwise, a decoding error is committed in the receiver. The decoding rule for choosing the estimate $\mathbf{C}$, given the received vector, is said to be optimum when the probability of decoding error is minimized. Assuming, equiprobable messages, the probability of decoding error is minimized if the estimate $\hat{\mathbf{C}}$ is chosen to maximize the log-likelihood function, $\log \Pr(\mathbf{R}|\mathbf{C})$. Here $\Pr(\mathbf{R}|\mathbf{C})$ denotes the conditional probability of receiving $\mathbf{R}$, given $\mathbf{C}$ was sent.

The maximum likelihood decoder is described as follows: choose the estimate $\hat{\mathbf{C}}$ for which the $\log \Pr(\mathbf{R}|\mathbf{C})$ is maximum. Mathematically, we can write

$$\hat{\mathbf{C}} = \underset{\mathbf{C}}{\operatorname{argmax}} \log \Pr(\mathbf{R}|\mathbf{C}). \tag{2.12}$$

We assume a flat fading channel and the noise process affecting a given bit in the received vector $\mathbf{R}$ is independent of the noise process affecting all of the other received bits. Since the probability of joint, independent events is simply the product of probabilities of the individual events [75], it follows that

$$\Pr(\mathbf{R}|\mathbf{C}) = \prod_{l=0}^{L'-1} \prod_{i=0}^{1} \Pr\left(r_l^{(i)}|c_l^{(i)}\right), \tag{2.13}$$

where

$$\Pr\left(r_l^{(i)}|c_l^{(i)}\right) = \frac{1}{\pi\sigma^2} e^{-\left|r_l^{(i)} - h_l^{(i)} c_l^{(i)}\right|^2/\sigma^2}, \tag{2.14}$$

where $h_l^{(i)}$ is the channel coefficient and σ^2 is the noise variance. The expression

$$\prod_{i=0}^{1} \Pr\left(r_l^{(i)}|c_l^{(i)}\right),$$

is called the branch metric because it indicates the component of (2.13) associated with the lth stage of the trellis. By taking the logarithm of each side of (2.13) we obtain the log likelihood function

$$\log\Pr(\mathbf{R}|\mathbf{C}) = \sum_{l=0}^{L'-1} \sum_{i=0}^{1} \log\Pr\left(r_l^{(i)}|c_l^{(i)}\right). \tag{2.15}$$

The log-likelihood function $\log\Pr(\mathbf{R}|\mathbf{C})$, which we denote by $\mathrm{G}(\mathbf{R}|\mathbf{C})$, represents the metric associated with the sequence $\mathbf{C}$. We can now define a partial path metric for the first t stages of a path as

$$\mathrm{G}(\mathbf{R}|\mathbf{C})_t = \sum_{l=0}^{t-1} \sum_{i=0}^{1} \log\Pr\left(r_l^{(i)}|c_l^{(i)}\right). \tag{2.16}$$

We have labeled the branches of our trellis with the output bits corresponding to a particular input to the encoder and the encoder's current state.

In the Viterbi algorithm, each node in the trellis is assigned a number. This number is the partial path metric of the path that starts at state S_0 at time $t = 0$ and ends at that node. The assignment of the numbers to trellis nodes is routine until we reach the point in the trellis in which more than one path enters each node. In this case we choose as the node label the best partial path metric among the metrics for all of the entering paths. The path with the best metric is the survivor, while the other entering paths are non-survivors. If two or more paths share the best metric, the survivor is randomly selected from among the best paths.

The selection of the survivors lies at the heart of the Viterbi algorithm. The algorithm terminates when all of the nodes in the trellis have been labeled and their entering survivors have been determined. We then go to the last node in the trellis (state S_0 at time L') and trace back through the trellis. At any given node, we can only continue backward on a path that survived upon entry into that node.

Algorithm 2.1

Viterbi Algorithm

Step 1. Set the metric stored at state S_0 at time $t = 0$ equal 0.

Step 2. At time t compute the partial path metrics for all paths entering a state by adding the branch metric entering that state to the metric of the connecting survivor at the preceding time unit.

Step 3. For each state, store the survivor together with its metric and eliminate all other paths.

Step 4. If $t < L'$, increase t and return to step 2. Otherwise, stop.

Since each node has only one entering survivor, our trace-back operation always yields a unique path. This path is the ML path. The following algorithm, when applied to the received sequence $\mathbf{R}$, finds the ML path through the trellis. The algorithm processes $\mathbf{R}$ in an iterative manner. At each step, it compares the metrics of all paths entering each state, and stores the survivor together with its metric. Following the surviving branches backward through the trellis gives a path. The path thus defined is unique and corresponding to the ML codeword. Algorithm 2.1 summarizes the Viterbi algorithm.

2.2.4 BCJR Decoding

Given the received sequence $\mathbf{R}$, the Viterbi algorithm finds the codeword $\mathbf{C}$ that maximizes the log-likelihood function. Once the ML codeword $\mathbf{C}$ is determined, its corresponding information $\mathbf{X}$ becomes the decoded output. Because the Viterbi algorithm finds the most likely codeword, it minimizes the word error rate, $\Pr\left(\hat{\mathbf{C}} \neq \mathbf{C}|\mathbf{R}\right)$. In many cases, however, we are interested in minimizing the bit-error-rate (BER), $\Pr\left(\hat{x}_l \neq x_l|\mathbf{R}\right)$ rather than the word error rate, $\Pr\left(\hat{\mathbf{C}} \neq \mathbf{C}|\mathbf{R}\right)$. To minimize the BER, the a posteriori probability $\Pr\left(\hat{x}_l = x_l|\mathbf{R}\right)$ must be maximized. An algorithm that maximizes $\Pr\left(\hat{x}_l = x_l|\mathbf{R}\right)$ is called a maximum a posteriori probability (MAP) decoder.

In 1974 Bahl, Cocke, Jelinek, and Raviv [76] introduced a maximum a posteriori (MAP) decoder, called the BCJR (or MAP) algorithm that can be applied any linear code, block or convolutional code. The BCJR algorithm shares many similarities with the Viterbi algorithm. Both are based on the same trellis, both assign branch metric to transitions, and both progress through the trellis recursively. However, rather than making one pass through the trellis from start to end, as is done in the Viterbi algorithm, the BCJR algorithm makes two phases: one forward pass from start to end, and then a second backward pass from end to start. Therefore, roughly speaking, the complexity of the BCJR algorithm is twice that of the Viterbi algorithm, and thus the Viterbi decoding is preferred in the case of equally likely information bits. When the information bits are not equally likely, however, better performance is achieved with a MAP decoder. On the other hand, Viterbi and BCJR algorithms differ in a fundamental way because they compute very different quantities. The Viterbi algorithm computes hard decisions by performing sequence detection, while the BCJR algorithm computes soft information about the data symbols in the form of a posteriori probabilities for each of the transmitted symbol. Recently, the BCJR algorithm is critical to modern iteratively-decoded error-correcting codes including turbo codes and low-density parity-check codes as the a priori probabilities of the information bits change from iteration to iteration, a MAP decoder gives the best performance [77–79].

In the following discussion we describe the BCJR algorithm for the case of convolutional codes used on binary-input, continuous-output AWGN channel and flat fading. We do not assume that the information bits are equally likely. The algorithm calculates the a posteriori log-likelihood ratios (L-values) as

$$L\left(x_l\right) = \log\left[\frac{\Pr\left(x_l = 1|\mathbf{R}\right)}{\Pr\left(x_l = 0|\mathbf{R}\right)}\right], \tag{2.17}$$

called the APP L-values, of each information bit, and the decoder output is given by

$$\hat{x}_l = \begin{cases} 1, & L(x_l) > 0 \\ 0, & L(x_l) < 0 \end{cases}, l = 0, 1, \cdots, L' - 1. \tag{2.18}$$

In iterative decoding, the APP L-values can be taken as the decoder outputs, resulting in a soft-input soft-output (SISO) decoding algorithm. The magnitude of an APP L-value is a measure of the confidence we have in the preceding decision: the more the $L(x_l)$ magnitude is far away from the zero threshold decision, the more we trust in the bit estimation we have made. We write the APP value $\Pr(x_l = 1|\mathbf{R})$ as follows:

$$\Pr(x_l = 1|\mathbf{R}) = \frac{\Pr(x_l = 1, \mathbf{R})}{\Pr(\mathbf{R})} = \frac{\sum_{\mathbf{X} \in \mathbf{X}^+} \Pr(\mathbf{R}|\mathbf{C}) \Pr(\mathbf{X})}{\sum_{\mathbf{X}} \Pr(\mathbf{R}|\mathbf{C}) \Pr(\mathbf{X})}, \tag{2.19}$$

where $\mathbf{X}^+$ is the set of all information sequences $\mathbf{X}$ such that $x_l = 1$. Rewriting $\Pr(x_l = 0|\mathbf{R})$ in the same way, we can write the expression (2.17) for the APP L-value as

$$L(x_l) = \ln \frac{\sum_{\mathbf{X} \in \mathbf{X}^+} \Pr(\mathbf{R}|\mathbf{C}) \Pr(\mathbf{X})}{\sum_{\mathbf{X} \in \mathbf{X}^-} \Pr(\mathbf{R}|\mathbf{C}) \Pr(\mathbf{X})}, \tag{2.20}$$

where $\mathbf{X}^-$ is the set of all information sequences $\mathbf{X}$ such that $x_l = 0$. MAP decoding can be achieved by computing the APP L-values $L(x_l)$, $l = 0, 1, 2, \cdots, L' - 1$ directly from (2.20) and then applying (2.18); however, except for very short block trellis L', the amount of computation required is prohibitive. For codes with trellis structure and a reasonable number of states, such as short convolutional codes, employing a recursive computational procedure based on the trellis structure of the code considerably simplifies the process.

First, making use of the trellis structure of the code, we can reformulate (2.19) as follows;

$$\Pr\left(x_l = 1|\mathbf{R}\right) = \frac{\Pr\left(x_l = 1, \mathbf{R}\right)}{\Pr\left(\mathbf{R}\right)} = \frac{\sum\limits_{(s',s)\in \mathrm{Y}_l^+} \Pr\left(s_l = s', s_{l+1} = s, \mathbf{R}\right)}{\Pr\left(\mathbf{R}\right)}, \tag{2.21}$$

where Y_l^+ is the set of all state pairs $s_l = s'$ and $s_{l+1} = s$ that correspond to the input bit $x_l = 1$ at time l. Reformulating the expression $\Pr\left(x_l = 0|\mathbf{R}\right)$ in the same way, we can write (2.17) for the APP L-value as

$$L\left(x_l\right) = \log \frac{\sum\limits_{(s',s)\in \mathrm{Y}_l^+} \Pr\left(s_l = s', s_{l+1} = s, \mathbf{R}\right)}{\sum\limits_{(s',s)\in \mathrm{Y}_l^-} \Pr\left(s_l = s', s_{l+1} = s, \mathbf{R}\right)}, \tag{2.22}$$

where Y_l^- is the set of all state pairs $s_l = s'$ and $s_{l+1} = s$ that correspond to the input bit $x_l = 0$ at time l. Equations (2.20) and (2.22) are equivalent expressions for the APP L-value $L\left(x_l\right)$, but whereas the summations in (2.20) extend only over a set of state pairs. Hence for a large block length L', (2.22) is considerably simpler to evaluate.

Now, we show how the joint probability density function (pdf), $\Pr\left(s', s, \mathbf{R}\right)$, in (2.22) can be evaluated recursively. The key to the BCJR algorithm is a decomposition of $\Pr\left(s', s, \mathbf{R}\right)$ for a transition at time l into three factors: the first depending on only the "past" observations, $\mathbf{r}_{t<l} = \{\mathbf{r}_t : t < l\}$, the second depending only on the "present" $\mathbf{r}_l$, and the third depending only on the "future" observations, $\mathbf{r}_{t>l} = \{\mathbf{r}_t : t > l\}$. We begin by writing

$$\Pr\left(s', s, \mathbf{R}\right) = \Pr\left(s', s, \mathbf{r}_{t<l}, \mathbf{r}_l, \mathbf{r}_{t>l}\right). \tag{2.23}$$

Application of Bayes' rule yields

$$\begin{aligned}
\Pr\left(s', s, \mathbf{R}\right) &= \Pr\left(\mathbf{r}_{t>l} \left| s', s, \mathbf{r}_{t<l}, \mathbf{r}_l \right.\right) \Pr\left(s', s, \mathbf{r}_{t<l}, \mathbf{r}_l\right) \\
&= \Pr\left(\mathbf{r}_{t>l} \left| s', s, \mathbf{r}_{t<l}, \mathbf{r}_l \right.\right) \Pr\left(s, \mathbf{r}_l \left| s', \mathbf{r}_{t<l} \right.\right) \mathrm{P}\left(s', \mathbf{r}_{t<l}\right) \qquad (2.24) \\
&= \Pr\left(\mathbf{r}_{t>l} \left| s \right.\right) \Pr\left(s, \mathbf{r}_l \left| s' \right.\right) \Pr\left(s', \mathbf{r}_{t<l}\right),
\end{aligned}$$

where the last equality follows from the fact that the probability of the received branch at time l depends only on the state and the input symbol at time l. Defining

$$\alpha_l(s') = \Pr\left(s', \mathbf{r}_{t<l}\right), \qquad (2.25)$$

$$\gamma_l(s', s) = \Pr\left(s, \mathbf{r}_l \left| s' \right.\right), \qquad (2.26)$$

$$\beta_{l+1}(s) = \Pr\left(\mathbf{r}_{t>l} \left| s \right.\right), \qquad (2.27)$$

we can rewrite (2.24) as

$$\Pr\left(s', s, \mathbf{R}\right) = \beta_{l+1}(s)\gamma_l(s', s)\alpha_l(s'). \qquad (2.28)$$

The expression for the probability $\alpha_{l+1}(s)$ can be rewritten as

$$\begin{aligned}
\alpha_{l+1}(s) &= \Pr\left(s, \mathbf{r}_{t<l+1}\right) = \sum_{s' \in \sigma_p} \Pr\left(s', s, \mathbf{r}_{t<l+1}\right) \\
&= \sum_{s' \in \sigma_l} \Pr\left(s, \mathbf{r}_l \left| s', \mathbf{r}_{t<l} \right.\right) \Pr\left(s', \mathbf{r}_{t<l}\right) \\
&= \sum_{s' \in \sigma_l} \Pr\left(s, \mathbf{r}_l \left| s' \right.\right) \Pr\left(s', \mathbf{r}_{t<l}\right) \\
&= \sum_{s' \in \sigma_l} \gamma_l(s', s)\alpha_l(s'), \qquad (2.29)
\end{aligned}$$

where σ_l is the set of all states at time l. Thus, we can compute a forward metric $\alpha_{l+1}(s)$ for each state s at time $l+1$ using the forward recursion (2.29). Similarly,

we can write the expression for the probability $\beta_l(s')$ as

$$\beta_l(s') = \sum_{s' \in \sigma_{l+1}} \gamma_l(s', s)\beta_{l+1}(s), \tag{2.30}$$

where σ_{l+1} is the set of all states at time $l+1$, and we can compute a backward metric $\beta_l(s')$ for each state s' at time l using the backward metric (2.30). The forward recursion begins at time $l = 0$ with initial condition

$$\alpha_0(s) = \begin{cases} 1, & s = \mathbf{0} \\ 0, & s \neq \mathbf{0}. \end{cases} \tag{2.31}$$

Similarly, the backward recursion begins at time $l = L'$ with the initial condition

$$\beta_L(s) = \begin{cases} 1, & s = \mathbf{0} \\ 0, & s \neq \mathbf{0}. \end{cases} \tag{2.32}$$

We can write the branch metric $\gamma_l(s', s)$ as

$$\begin{aligned} \gamma_l(s', s) &= \Pr\left(s, \mathbf{r}_l \,|s'\right) = \frac{\Pr\left(s', s, \mathbf{r}_l\right)}{\Pr\left(s'\right)} \\ &= \frac{\Pr\left(s', s\right)}{\Pr\left(s'\right)} \frac{\Pr\left(s', s, \mathbf{r}_l\right)}{\Pr\left(s', s\right)} \\ &= \Pr\left(s \,|s'\right) \Pr\left(\mathbf{r}_l \,|s', s\right) \\ &= \Pr\left(x_l\right) \Pr\left(\mathbf{r}_l \,|\mathbf{c}_l\right), \end{aligned} \tag{2.33}$$

where $\mathbf{c}_l$ is the output corresponding to the state transition $s' \to s$ at time l and

$$\Pr\left(\mathbf{r}_l \,|\mathbf{c}_l\right) = \frac{1}{\pi\sigma^2} \exp\left(-\left|\mathbf{r}_l - \mathbf{h}_l\mathbf{c}_l\right|^2 / \sigma^2\right) \tag{2.34}$$

where $\mathbf{h}_l$ is the corresponding channel coefficients.

Log MAP BCJR

The BCJR algorithm is in general of high complexity, and, on the other hand, sums and products involved in its calculation can lead to underflow and overflow problems in practical implementations. These calculations also require considerable amount of memory to store all the values, until a decoding decision is taken. A logarithmic version of this algorithm appears to be a solution to the many calculation problems that the original version of the BCJR algorithm faces. The basic idea is that by converting calculations into their logarithmic form, products convert into sums. The logarithm of a sum of two or more terms seems to be a new complication, but this operation is solved by using the following equation,

$$\max^*(x,y) \equiv \log\left(e^x + e^y\right) = \max(x,y) + \log\left(1 + e^{-|x-y|}\right), \tag{2.35}$$

where the term $\log\left(1 + e^{-|x-y|}\right)$ can be either exactly calculated, or in practical implementations of this algorithm, obtained from a look up table.

The logarithmic version of the MAP BCJR algorithm greatly reduces the overflow and underflow effects in its applications. The version is known as the LOG MAP BCJR. Another and even simpler version of the LOG MAP BCJR is the so-called MAX LOG MAP BCJR algorithm, in which the term $\log\left(1 + e^{-|x-y|}\right)$ is omitted in the calculations, and (2.35) is used by simply evaluating the MAX value of the involved quantities [80].

Defining $\alpha^*_{l+1}(s) = \log\left(\alpha_{l+1}(s)\right)$, $\beta^*_l(s') = \log\left(\beta_l(s')\right)$, and

$$\begin{aligned}
\gamma^*_l(s',s) &= \log\left(\gamma_l(s',s)\right), \\
&= \log\left(\Pr\left(x_l\right)\right) + \log\left(\Pr\left(\mathbf{r}_l \,|\mathbf{c}_l\right)\right), \\
&= \log\left(\Pr\left(x_l\right)\right) - \left|\mathbf{r}_l - \mathbf{h}_l\mathbf{c}_l\right|^2 / \sigma^2.
\end{aligned} \tag{2.36}$$

We exploit (2.33) and (2.34) to obtain the second and third lines of the previous equation, respectively, and we drop $\frac{1}{\pi\sigma^2}$ of (2.34) as this is a common term in all branches. Using (2.28) - (2.35), we can write

$$\begin{aligned}\alpha_{l+1}^*(s) &= \log \sum_{s' \in \sigma_l} \gamma_l(s', s)\alpha_l(s'), \\ &= \log \sum_{s' \in \sigma_l} e^{\left[\gamma_l^*(s', s)\alpha_l^*(s')\right]}, \\ &= \max^*_{s' \in \sigma_l} \left[\gamma_l^*(s', s) + \alpha_l^*(s')\right], \quad l = 0, 1, \cdots, L' - 1, \end{aligned} \tag{2.37}$$

and

$$\alpha_0^*(s) \equiv \log\left(\alpha_0(s)\right) = \begin{cases} 0, & s = \mathbf{0} \\ -\infty, & s \neq \mathbf{0} \end{cases}. \tag{2.38}$$

Similarly,

$$\begin{aligned}\beta_l^*(s') &= \log \sum_{s \in \sigma_{l+1}} \gamma_l(s', s)\beta_{l+1}(s), \\ &= \log \sum_{s \in \sigma_l} e^{\left[\gamma_l^*(s', s)\beta_{l+1}^*(s)\right]}, \\ &= \max^*_{s \in \sigma_{l+1}} \left[\gamma_l^*(s', s) + \beta_{l+1}^*(s)\right], \quad l = L' - 1, L' - 2, \cdots, 0, \end{aligned} \tag{2.39}$$

and

$$\beta_{L'}^*(s) \equiv \log\left(\beta_{L'}(s)\right) = \begin{cases} 0, & s = \mathbf{0} \\ -\infty, & s \neq \mathbf{0} \end{cases}. \tag{2.40}$$

Further, we can now write the expressions for the pdf $\Pr\left(s, s', \mathbf{r}\right)$ and the APP L-value $L\left(x_l\right)$ as

$$\Pr\left(s, s', \mathbf{r}\right) = e^{\beta_{l+1}^*(s) + \gamma_l^*(s', s) + \alpha_l^*(s')} \tag{2.41}$$

Algorithm 2.2

The Log-domain BCJR algorithm

Step 1. Initialize the forward and backward metrics $\alpha_0^*(s)$ and $\beta_{L'}^*(s)$ using (2.38) and (2.40).
Step 2. Compute the branch metrics $\gamma_l^*(s',s)$, $l = 0, 1, \cdots, L'-1$, using (2.36).
Step 3. Compute the forward metrics $\alpha_{l+1}^*(s)$, $l = 0, 1, \cdots, L'-1$, using (2.37).
Step 4. Compute the backward metrics $\beta_l^*(s')$, $l = L'-1, L'-2, \cdots, 0$, using (2.39).
Step 5. Compute the APP L-values $L(x_l)$ $l = 0, 1, \cdots, L'-1$, using (2.42).
Step 6. (Optional) Compute the hard decisions $\hat{x}_l$, $l = 0, 1, \cdots, L'-1$, using (2.18).

and

$$\begin{aligned} L(x_l) &= \log\left\{\sum_{(s,s')\in Y_l^+} e^{\beta_{l+1}^*(s)+\gamma_l^*(s',s)+\alpha_l^*(s')}\right\} \\ &- \log\left\{\sum_{(s,s')\in Y_l^-} e^{\beta_{l+1}^*(s)+\gamma_l^*(s',s)+\alpha_l^*(s')}\right\}, \\ &= \max^*_{(s,s')\in Y_l^+}\left[\beta_{l+1}^*(s)+\gamma_l^*(s',s)+\alpha_l^*(s')\right] \\ &- \max^*_{(s,s')\in Y_l^-}\left[\beta_{l+1}^*(s)+\gamma_l^*(s',s)+\alpha_l^*(s')\right]. \end{aligned} \tag{2.42}$$

Here, we use the fact that three-variable max*function can be obtained by applying the 2-variable max*function twice as

$$\max^*(x,y,z) \equiv \log\left(e^x+e^y+e^z\right) = \max^*\left(\max^*(x,y),z\right), \tag{2.43}$$

Because using max-star function, the LOG MAP BCJR, so called the Log-domain BCJR is considerably simpler to implement and provides greater numerical stability than the original BCJR algorithm. The Log-domain BCJR algorithm is summarized in Algorithm 2.2.

2.3 BICM-ID

For a long time, error correcting codes and modulation have been treated as distinct tasks in communication systems. By integrating error correcting codes and modulation, Ungerboeck in 1987 [68, 69] proposed Trellis Coded Modulation (TCM), which is capable of achieving significant coding gains[1] over different wireless channels. Among many adaptions [54–57] suggested to improve the BER performance of TCM schemes, bit interleaved coded modulation (BICM) proposed by Zehavi [49] gives the largest improvement for fading channels. BICM is based on the combination of coding, bit interleaving[2], and mapping (modulation). A significant improvement for BICM is achieved by using iterative decoding (ID) as studied in [50–53, 81]. A soft input soft output (SISO) [82] decoder is used in the BICM-ID scheme where the decoder's output is fed back to the input of the demapper. BICM-ID is flexible in allowing any type of code to be used in combination with virtually any modulation type. BICM-ID can be used as a bandwidth efficient approach to provide excellent performance over flat Rayleigh channels, provided that SISO channel decoding is applied.

In the following section, we present a short overview of BICM-ID. The details about iterative probabilistic decoding can be found in many tutorials such as [58, 83, 84].

2.3.1 Encoder

Figure 2.7 shows the block diagram of a BICM-ID system. The code that is used in this work is a convolutional code. The transmitter consists of the serial concate-

[1] The coding gain is the amount of additional signal to noise ratio that would be required to provide the same bit-error-rate (BER) performance for an uncoded signal.

[2] The interleaver is a one-to-one mapping that rearranges the ordering of a sequence of symbols in a deterministic manner.

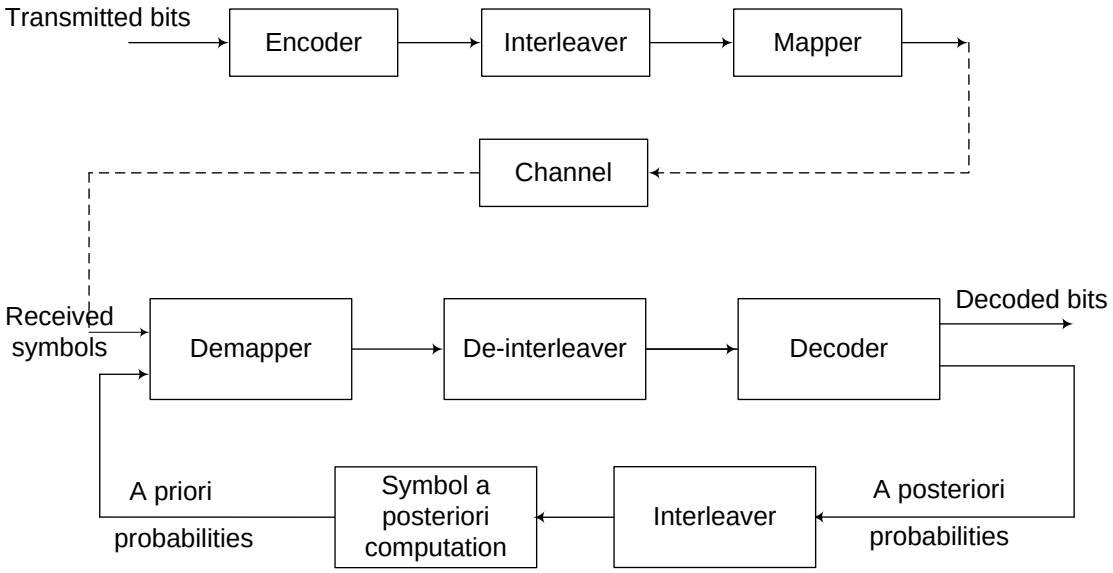

Figure 2.7: Block diagram of BICM-ID with soft feedback.

nation of the convolutional encoder, bit interleaver and the mapper. For convolutional codes with rate $1/n$, each input data bit at time t is encoded to the coded output bit sequence $\mathbf{c}_t = \left[c_t^0, c_t^1, \cdots, c_t^{n-1}\right]$. We represent the output of the convolutional encoder by

$$\mathbf{C} = \left[c_0^0, c_0^1, \cdots, c_0^{n-1}, \cdots, c_t^0, c_t^1, \cdots, c_t^{n-1}, \cdots\right] = \left[\mathbf{c}_0, \cdots, \mathbf{c}_t, \cdots\right]. \tag{2.44}$$

The bit interleaver permutes the coded bits in a deterministic way to make sure that adjacent bits are separated by several bits after interleaving. At the de-interleaver, the permutation is inverted. The idea behind using the interleaver and de-interleaver is to spread out bursts of channel errors in time, such that the decoder can handle them as if they were random errors. This improves the error correction capabilities of coding schemes over bursty channels. The output of the interleaver is represented by

$$\mathbf{V} = \left[v_0^0, v_0^1, \cdots, v_0^{m-1}, \cdots, v_t^0, v_t^1, \cdots, v_t^{m-1}, \cdots\right] = [\mathbf{v}_0, \cdots, \mathbf{v}_t, \cdots] \tag{2.45}$$

where v_t^i is the ith output bit at time position t. Each m-tuple of the interleaver output is grouped together, $\mathbf{v}_t = \left[v_t^0, v_t^1, \cdots, v_t^{m-1}\right]$, and is mapped to a complex symbol x_t chosen from an M-ary constellation Ω by a signal mapping μ

$$d_t = \mu\left(\mathbf{v}_t\right) \quad x_t \in \Omega. \tag{2.46}$$

2.3.2 Decoder

For transmission over flat fading channels, the decision variables can be written as

$$y_t = h_t d_t + n_t \tag{2.47}$$

where h_t is the channel coefficient and n_t is complex valued AWGN. In principle, MAP decoding of the received sequence $\mathbf{y} = [y_0, \cdots y_t, \cdots]$ can be carried out as

$$\widehat{\mathbf{u}}_{\mathrm{MAP}} = \arg\max_{\mathbf{u}} \ \mathrm{P}\left(\mathbf{u}\,|\mathbf{y}, \mathbf{h}\right)$$

where $\mathbf{h} = [h_0, \cdots h_t, \cdots]$ are the channel coefficients and $\mathbf{u} = [u_0, \cdots u_t, \cdots]$ is the trial value of the transmitted information sequence. Unfortunately, due to the presence of the bit-based interleaver, the true MAP decoding of BICM requires joint demapping and decoding and is therefore too complex to implement in practice [85]. In [49], Zehavi suggested a suboptimal method using two separate steps: bit metric generation and Viterbi decoding. To improve the BER performance of a

BICM scheme, the Viterbi decoding is replaced by a SISO[3] decoder and the output of the SISO decoder is fed back to the demapper for bit metric regeneration [53,81]. The MAP bit metrics are calculated as [85,86]

$$\begin{aligned}\lambda\left(v_t^i=b\right) &= \Pr\left(v_t^i=b|y_t,h_t\right) \quad i=0,\cdots,m-1;\ b=0,1\\ &= \sum_{d\in\psi(i,b)}\Pr\left(d|y_t,h_t\right) \\ &\propto \sum_{d\in\psi(i,b)}\Pr\left(y_t|d,h_t\right)\mathrm{P}\left(d\right)\end{aligned} \tag{2.48}$$

where the signal subset $\psi\left(i,b\right) = \left\{\mu\left(\left[v_t^0,v_t^1,\cdots,v_t^{m-1}\right]\right)|v_t^i=b\right\}$. The size of each subset $\psi\left(i,b\right)$ is 2^{m-1}. The a priori probability $\Pr\left(d\right)$ is unavailable on the first iteration of the demapping. Therefore, in the initialization phase it is assumed that all d are equally probable. (2.48) is used as the input to the SISO decoder, which then generates the a posteriori probabilities for the coded bits. From the a posteriori probabilities for the coded bits, the extrinsic a posteriori probabilities are extracted. This extrinsic information depends only on the redundant information introduced by the used code.

On the second iteration, the extrinsic a posteriori probabilities[4] $\Pr\left(c_t^i;E\right)$ coming out from the SISO decoder are interleaved and fed back as a priori probabilities $\Pr\left(v_t^i;A\right)$ to the demapper as shown in Figure 2.7, where A and E denote a priori and extrinsic a posteriori respectively. Assuming that the probabilities $\Pr\left(v_t^0;A\right)$,..., $\Pr\left(v_t^{m-1};A\right)$ are independent by using a good interleaver, the symbol a priori

[3]The SISO decoder of the BICM-ID scheme is actually a maximum a posterior probability (MAP) decoder that computes the a posterior probabilities of the coded bits.

[4]The a priori and extrinsic a posteriori probabilities are well explained in the literature of turbo codes [67,87].

probability $\Pr(x)$ can be computed as

$$\begin{aligned}\Pr(d) &= \Pr\left(\mu\left(\mathbf{v}_t(d); A\right)\right) \\ &= \prod_{j=0}^{m-1} \Pr\left(v_t^j(d) = b; A\right). \end{aligned} \tag{2.49}$$

Using (2.48) and (2.49), the extrinsic a posteriori bit probabilities for the second pass demodulation can be written as

$$\begin{aligned}\Pr\left(v_t^i = b; E\right) &= \frac{\Pr\left(v_t^i = b|\mathrm{y}_t\right)}{\Pr\left(v_t^i = b; A\right)} \\ &= \frac{\sum_{d\in\psi(i,b)} \Pr\left(y_t|d, h_t\right) \prod_{j=0}^{m-1} \Pr\left(v_t^j(d) = b; A\right)}{\Pr\left(v_t^i(d) = b; A\right)} \\ &= \sum_{d\in\psi(i,b)} \Pr\left(y_t|d, h_t\right) \prod_{j=0, j\neq i}^{m-1} \Pr\left(v_t^j(d) = b; A\right). \end{aligned} \tag{2.50}$$

Equation (2.50) shows that we only need the a priori probabilities $\Pr\left(v_t^j = b; A\right)$ of the other bits than the considered one $(i \neq j)$ from the same channel symbol $\mathbf{v}_t$ when recalculating the bit metrics. The receiver then uses (2.50) to regenerate the bit metrics and iterates between the demapper and the decoder. After the last iteration, the final decoded outputs are the hard decisions based on the a posteriori probabilities.

2.4 Summary

In this chapter, a general overview of signal detection theory and convolutional codes has been presented. First, we have briefly described the signal detection theory with focusing on ML and MAP algorithms. Next, we have explained ba-

sic ideas of convolutional codes. Viterbi and BCJR algorithms, as powerful tools for decoding, have been also addressed. Finally, the chapter has covered the bit interleaved coded modulation iterative decoding (BICM-ID) schemes.

Chapter 3

Data Detection for Uncoded ARC Systems

A key issue in the alternate-relaying cooperative (ARC) communication systems is the interference which is caused by the simultaneous transmission of the source and one of the relays at the same time. In this chapter, we propose maximum likelihood (ML) detectors to mitigate the interference in uncoded decode-and-forward (DF) ARC systems. At the relays, the proposed ML detector is employed by averaging out the interference signal. Furthermore, unlike previous work in which interference cancellation is required at the destination, we exploit the interference signal as a beneficial resource to develop the optimal detectors at the destination equipped with single and multiple antennas. The major drawback of the proposed optimal detectors is the delay because the destination has to receive and store the entire frame before performing data detection. Due to the inevitable delay restriction, sub-optimal detectors are developed. In contrast with the optimal detector, the sub-optimal detectors exploit two consecutive received packets to decode one packet. Extensive simulation results have been presented to demonstrate the effec-

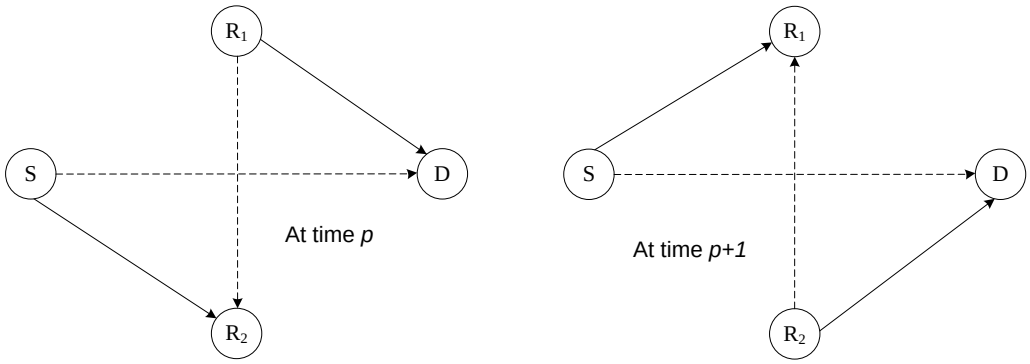

Figure 3.1: System model of the alternate-relaying cooperative system.

tiveness of the proposed detectors.

The chapter is organized as follows. In Section 3.1, the system model and problem formulation are presented. Then, the ML detector at the relays is introduced in Section 3.2. The optimal and sub-optimal detectors at the destination equipped with single and multiple antenna are proposed in Sections 3.3 and 3.4, respectively. The performance of the proposed detectors is evaluated through computer simulations in Section 3.5. Finally, we conclude the chapter in Section 3.6.

3.1 System Model and Problem Formulation

A four-node network with one source (S), one destination (D), and two relays (R_1, R_2) is considered. We assume that each node is equipped only with one antenna and the relays operate in half-duplex mode. Each link is assumed to be frequency-flat and the fading coefficient between two nodes A and B is denoted as h_{AB}, where $A, B \in \{S, R_1, R_2, D\}$, $A \neq B$. This assumption is valid as long as the signal bandwidth is less than the channel coherence bandwidth. All nodes have equal additive white Gaussian noise (AWGN) variance of σ^2. The network supports frame-based transmissions, in which each frame is composed of P packets. The pth packet $\mathbf{d}^{(p)} = \left[d^{(p)}(1),\, d^{(p)}(2),\, \cdots,\, d^{(p)}(N_d)\right]$ consists of N_d data symbols, which are ran-

domly and independently drawn from a constellation Ω, $p = 1, 2, \cdots, P$. Each frame is transmitted continuously, one packet per time slot, over wireless channels to the destination and relays. The transmission schedule for the P time slots for each frame is described as follows (see Figure 3.1):

- At even time slot p: S sends $\mathbf{d}^{(p)}$ and R_1 (named the forwarding relay) sends $\mathbf{d}^{(p-1)}\psi_{R_1}(p-1)$, where $\psi_{R_1}(p-1)$ is an indicator function which is unity when the relay R_1 was able to correctly detect the $(p-1)$th data packet and zero otherwise[1]. R_2 (named the listening relay) listens to $\mathbf{d}^{(p)}$ from S while being interfered by $\mathbf{d}^{(p-1)}\psi_{R_1}(p-1)$ from R_1. D receives $\mathbf{d}^{(p-1)}\psi_{R_1}(p-1)$ from R_1 while being interfered by $\mathbf{d}^{(p)}$ from S.

- At odd time slot $p+1$: S sends $\mathbf{d}^{(p+1)}$ and R_2 (named the forwarding relay) sends $\mathbf{d}^{(p)}\psi_{R_2}(p)$. R_1 (named the listening relay) listens to $\mathbf{d}^{(p+1)}$ from S while being interfered by $\mathbf{d}^{(p)}\psi_{R_2}(p)$ from R_2. D receives $\mathbf{d}^{(p)}\psi_{R_2}(p)$ from R_2 while being interfered by $\mathbf{d}^{(p+1)}$ from S.

These transmission steps are then continuously repeated until P packets are transmitted by the source. In practice, in order to achieve diversity for the last packet, one additional time slot is required at the end of the transmission; S remains silent while R_1 (or R_2) sends $\mathbf{d}^{(P)}$. This slight loss in the rate, $\frac{1}{P}$, is asymptotically zero for large values of P. As one can observe from Figure 3.2, the destination receives two copies of each packet, one from the source-destination link and the other from the forwarding relay-destination link, if this packet was correctly detected at this relay; otherwise, the destination receives only one copy through the source-destination link. This implies that diversity gain can still be achieved by this protocol, while

[1]In order to check whether their decoded packets are correct, relays can apply different techniques, such as cyclic redundancy check (CRC) code or signal-to-noise (SNR) ratio threshold. In this work, we assume that CRC error detection is already embedded into information packets, and is able to detect all the packet errors, as assumed in many other studies, e.g., [88, 89].

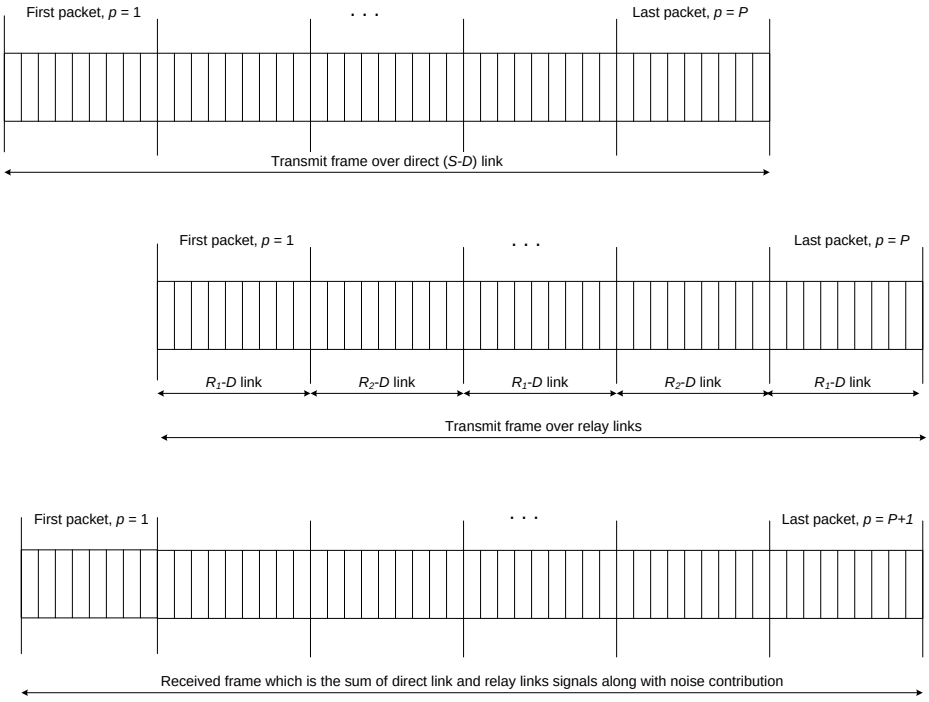

Figure 3.2: The detailed signal structure.

the source transmits continuously. As a result, the bandwidth efficiency is not sacrificed, and full-rate transmission is retained.

Based on Figure 3.1, at time slot p, the nth received symbol at the destination $y_D^{(p)}(n)$, and relay R_2, $y_{R_2}^{(p)}(n)$, can be respectively expressed as

$$y_D^{(p)}(n) = d^{(p)}(n)h_{SD}^{(p)}(n) + d^{(p-1)}(n)\psi_{R_1}(p-1)h_{R_1D}^{(p)}(n) + w_D^{(p)}(n), \tag{3.1}$$

$$y_{R_2}^{(p)}(n) = d^{(p)}(n)h_{SR_2}^{(p)}(n) + d^{(p-1)}(n)\psi_{R_1}(p-1)h_{R_1R_2}^{(p)}(n) + w_{R_2}^{(p)}(n), \tag{3.2}$$

where $w_D^{(p)}(n)$ and $w_{R_2}^{(p)}(n)$ are AWGN contributions at destination and relay R_2, respectively. Further, at time slot $p+1$, the received signal at the destination and relay R_1 can be respectively written as

$$y_D^{(p+1)}(n) = d^{(p+1)}(n)h_{SD}^{(p+1)}(n) + d^{(p)}(n)\psi_{R_2}(p)h_{R_2D}^{(p+1)}(n) + w_D^{(p+1)}(n), \tag{3.3}$$

$$y_{R_1}^{(p+1)}(n) = d^{(p+1)}(n)h_{SR_1}^{(p+1)}(n) + d^{(p)}(n)\psi_{R_2}(p)h_{R_1R_2}^{(p+1)}(n) + w_{R_1}^{(p+1)}(n), \quad (3.4)$$

where $w_{R_1}^{(p+1)}(n)$ is the AWGN contribution at relay R_1. Our goal is to develop the ML detectors at the relays and destination.

3.2 ML Detector at the Relays

From (3.2) and (3.4), one can observe that data received at a listening relay from the source is interfered by data sent from the forwarding relay. This is because at any time slot, there is always a relay transmitting data simultaneously with the source. In this section, we propose an ML detector at the listening relay based on averaging out the interference signal. Without loss of generality, we can write the received signal at relay R_x ($x \in \{1,2\}$) as

$$y_{R_x}^{(p)}(n) = d^{(p)}(n)h_{SR_x}^{(p)}(n) + d^{(p-1)}(n)\psi_{R_y}(p-1)h_{R_xR_y}^{(p)}(n) + w_{R_x}^{(p)}(n), \quad (3.5)$$

where $y \in \{1,2\}$ and $x \neq y$. Based on the ML principle, the detected symbol, $d^{(p)}(n)$, at relay R_x can be found as

$$\hat{d}^{(p)}(n) = \arg \max_{d^{(p)}(n)} \log \Pr\left(y_{R_x}^{(p)}(n), d^{(p-1)}(n) \,\middle|\, d^{(p)}(n), h_{SR_x}^{(p)}(n), h_{R_xR_y}^{(p)}(n)\right). \quad (3.6)$$

Since the symbol $d^{(p-1)}(n)$ is not known at relay R_x, we remove its contribution by averaging out as

$$\begin{aligned} &\Pr\left(y_{R_x}^{(p)}(n) \,\middle|\, d^{(p)}(n), h_{SR_x}^{(p)}(n), h_{R_xR_y}^{(p)}(n)\right) \\ &= \sum_{d^{(p-1)}(n)\in\Omega} \Pr\left(y_{R_x}^{(p)}(n) \,\middle|\, d^{(p)}(n), d^{(p-1)}(n), h_{SR_x}^{(p)}(n), h_{R_xR_y}^{(p)}(n)\right) \times \Pr\left(d^{(p-1)}(n)\right), \end{aligned} \quad (3.7)$$

and

$$\Pr\left(y_{R_x}^{(p)}(n)\,\middle|\,d^{(p)}(n), d^{(p-1)}(n), h_{SR_x}^{(p)}(n), h_{R_xR_y}^{(p)}(n)\right) = \frac{1}{\pi\sigma^2}\exp\left(-\left\|y_{R_x}^{(p)}(n) - d^{(p)}(n)h_{SR_x}^{(p)}(n) - d^{(p-1)}(n)\psi_{R_y}(p-1)h_{R_xR_y}^{(p)}(n)\right\|^2/\sigma^2\right). \tag{3.8}$$

We assume that the transmitted symbols are equally probable; thus, $\Pr\left(d^{(p-1)}(n)\right)$ does not affect the optimization.

3.3 Data Detection at the Destination

In the following we provide optimal and sub-optimal detectors at the destination node.

3.3.1 Optimal Detector

From (3.1) and (3.3), one can observe that the source and relays send their messages to the destination in a sequential form. For illustration, let us consider that the source transmits the frame $\left\{\cdots, \mathbf{d}^{(p-1)}, \mathbf{d}^{(p)}, \mathbf{d}^{(p+1)}, \cdots\right\}$. Accordingly, the source and relays send their data packets as $\left\{\cdots, \left[\mathbf{d}^{(p)}, \mathbf{d}^{(p-1)}\psi_{R_x}(p-1)\right], \left[\mathbf{d}^{(p+1)}, \mathbf{d}^{(p)}\psi_{R_y}(p)\right], \left[\mathbf{d}^{(p+2)}, \mathbf{d}^{(p+1)}\psi_{R_x}(p+1)\right], \cdots\right\}$. Consequently, in contrast to the relays, the destination receives each symbol in each packet two times, through the source-destination link and the relay-destination link, only if a relay was able to correctly detect the packet that contains this symbol; otherwise, the transmitted symbol is received only one time through the direct link. The equivalent block diagram of the DF alternate-relaying cooperative system can be represented as shown in Figure 3.3. From this Figure it is obvious that the equivalent model is analogous

to a convolutional code with a constraint length of two. Hence, one can describe the transmitter of the equivalent system through a trellis diagram. The trellis consists of M states, which is equal to the modulation order. There are M branches leaving from each state corresponding to M different input patterns. For example, the trellis diagram of a transmitter using 8-PSK modulation is shown in Figure 3.3, where $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8\}$ denote the 8-PSK symbols. Each branch is labeled by $\alpha_i/\alpha_i\alpha_x$, where α_i is the input symbol, and $\alpha_i\alpha_x$ represent the two symbols transmitted through the direct and relay links, respectively[2]. Each row indicates the branch labels for transitions from states corresponding to the inputs $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8\}$, respectively. Accordingly, by using this equivalent model, we can apply the Viterbi algorithm [66, 90] with a minor modification to derive the optimal detector[3].

The key idea of the proposed optimal detector is to employ N_d parallel Viterbi algorithms, as shown in Figure 3.4. The input of the nth Viterbi algorithm is the nth symbols from each received packet, $\mathbf{y}(n) = \left[y_D^{(1)}(n), y_D^{(2)}(n), \cdots, y_D^{(P)}(n)\right]$, with $y_D^{(p)}(n)$ and $y_D^{(p+1)}(n)$ defined in (3.1) and (3.3), respectively[4]. The nth Viterbi algorithm chooses the sequence $\mathbb{D}(n) = \left[d^{(1)}(n), d^{(2)}(n), \cdots, d^{(P)}(n)\right]$ that maximizes the log-likelihood function $\log \Pr\left(\mathbb{Y}(n) \,|\mathbb{D}(n), \mathbb{H}(n)\right)$

$$\widehat{\mathbb{D}(n)} = \arg \max_{\mathbb{D}(n)} \log \Pr\left(\mathbb{Y}(n) \,|\mathbb{D}(n), \mathbb{H}(n)\right), \tag{3.9}$$

[2]Strictly speaking, since the indicator function takes one or zero, the equivalent model is analogous to a punctured convolutional code.

[3]Note that when the information symbols are equally likely, the (ML) Viterbi algorithm provides the optimal BER performance [66, 67, 90].

[4]From (3.1) and (3.3), one can observe that each transmitted symbol $d^{(p)}(n)$ is included in $y_D^{(p)}(n)$ and $y_D^{(p+1)}(n)$. Accordingly, in order to provide an optimal detector for the nth symbol in each packet we have to rely on the received sequence $\left[y^{(1)}(n), y^{(2)}(n), \cdots, y^{(P)}(n)\right]$. That is why we mix the received packets, as shown in Figure 3.4.

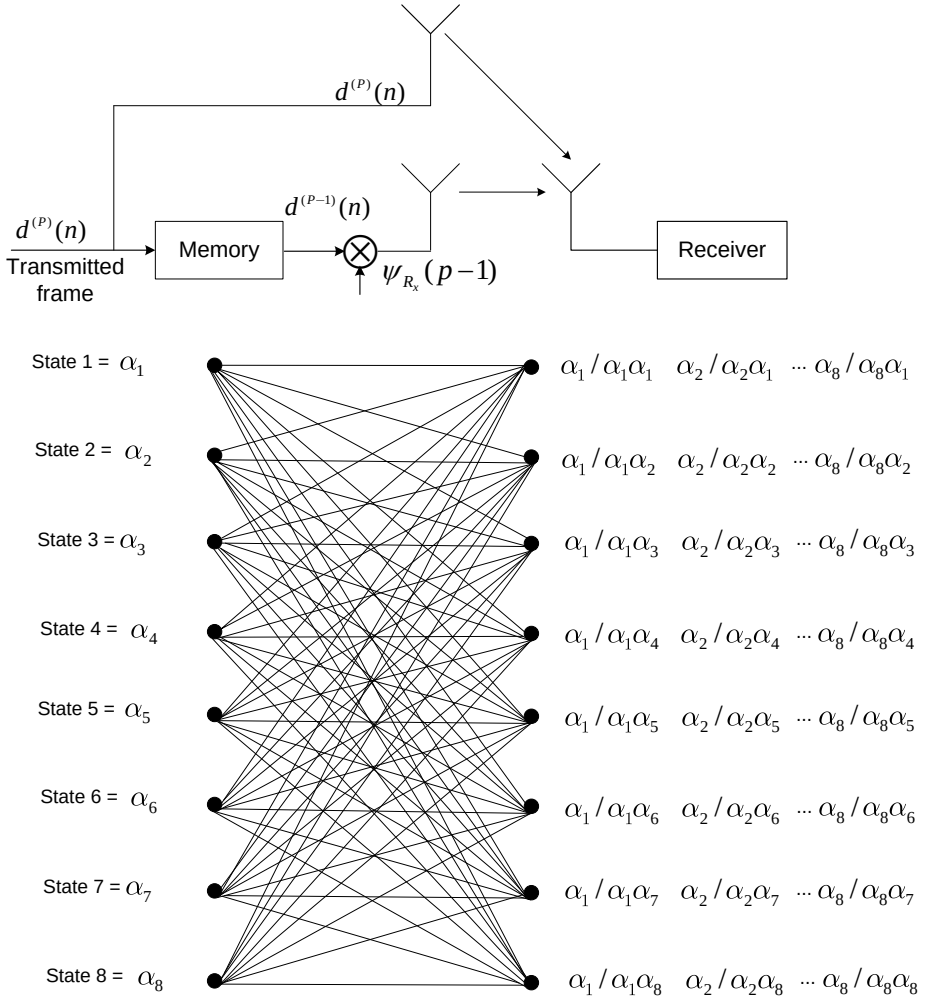

Figure 3.3: The equivalent system of the DF-ARC system, with trellis diagram for 8-PSK modulation as an example.

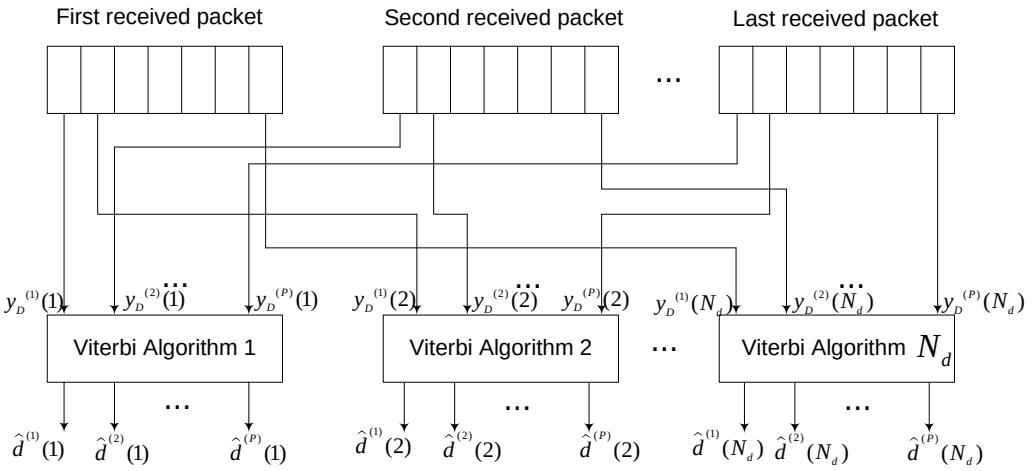

Figure 3.4: The optimal receiver for the full-rate DF-ARC system.

where $\mathbb{H}(n) = \left[h_{SD}^{(1)}(n), h_{R_xD}^{(1)}(n), h_{SD}^{(2)}(n), h_{R_yD}^{(2)}(n), \cdots, h_{SD}^{(P)}(n), h_{R_xD}^{(P)}(n)\right]$ represents the channel coefficients. Since the probability of joint independent events is simply the product of probabilities of the individual events, it follows that

$$\Pr\left(\mathbb{Y}(n)\,|\mathbb{D}(n), \mathbb{H}(n)\right) = \prod_{p=1}^{P} \Pr\left(y_D^{(p)}(n) \,\middle|\, \mathbf{d}^{(p)}(n), \mathbf{h}^{(p)}(n)\right), \tag{3.10}$$

where $\mathbf{d}^{(p)}(n) = \left[d^{(p)}(n),\, d_{R_x}^{(p-1)}(n)\right]$ and $\mathbf{h}^{(p)}(n) = \left[h_{SD}^{(p)}(n),\, h_{R_xD}^{(p)}(n)\right]$ and the transition probability $\Pr\left(y_D^{(p)}(n) \,\middle|\, \mathbf{d}^{(p)}(n), \mathbf{h}^{(p)}(n)\right)$ is defined as

$$\begin{aligned}&\Pr\left(y_D^{(p)}(n) \,\middle|\, \mathbf{d}^{(p)}(n), \mathbf{h}^{(p)}(n)\right)\\ &= \tfrac{1}{\pi\sigma^2}\exp\left(-\left|y_D^{(p)}(n) - d^{(p)}(n)h_{SD}^{(p)}(n) - d^{(p-1)}(n)\psi_{R_x}(p-1)h_{R_xD}^{(p)}(n)\right|^2/\sigma^2\right).\end{aligned} \tag{3.11}$$

This yields

$$\log\Pr\left(\mathbb{Y}(n)\,|\mathbb{D}(n), \mathbb{H}(n)\right) = \sum_{p=1}^{P} \log\Pr\left(y_D^{(p)}(n) \,\middle|\, \mathbf{d}^{(p)}(n), \mathbf{h}^{(p)}(n)\right). \tag{3.12}$$

The log-likelihood function $\log\Pr\left(\mathbb{Y}(n)\,|\mathbb{D}(n), \mathbb{H}(n)\right)$, which we denote by

$\mathrm{G}\left(\mathbb{Y}(n)\,|\mathbb{D}(n),\mathbb{H}(n)\right)$, represents the metric associated with the sequence $\mathbb{D}(n)$. The probability $\Pr\left(y_D^{(p)}(n)\,\middle|\,\mathbf{d}^{(p)}(n),\mathbf{h}^{(p)}(n)\right)$ is referred to as the branch metric. We can now express a partial path metric for the first t branches of a path as[5]

$$\mathrm{G}\left([\mathbb{Y}(n)\,|\mathbb{D}(n),\mathbb{H}(n)]_t\right)=\sum_{p=1}^{t}\log\Pr\left(y_D^{(p)}(n)\,\middle|\,\mathbf{d}^{(p)}(n),\mathbf{h}^{(p)}(n)\right). \tag{3.13}$$

Implementation Aspects

- Since the Viterbi algorithm is a trace back algorithm, the end state must be known. An additional N_d zero symbols are added at the end of the transmitted frame to drive the "virtual" encoder to state zero. In the literature, these additional zeros are referred to as the tail [66, 90]. This will result in a slight bandwidth loss compared with the direct transmission. This loss is given by $\frac{1}{P}$, which approaches zero for a large number of packets per frame.

- We use adaptive DF relaying where the relays forward only when they correctly detect the received packets; otherwise, relays remain idle. Relays can apply different strategies, such as cyclic redundancy check (CRC) code or signal-to-noise (SNR) ratio threshold, in order to check whether their decoded packets are correct. In this work, we assume that CRC error detection is already embedded into information packets, and is able to detect all the packet errors, as assumed in many other studies, e.g., [88, 89]. As such, the destination has to receive acknowledgment from the relays about the status of each packet. For illustration, if the destination receives a negative acknowledgment for the pth packet, this implies that the relay R_x ($x = 1$ or 2) is unable to decode and forward this packet.

- Unfortunately, the proposed optimal detector requires a large memory and

[5]For details of the Viterbi algorithm, the reader is referred to Section 2.2.3.

delay. The destination has to receive and store the entire frame before starting data detection. These limitations are an increasing function on the modulation order (M) and the number of packets per frame (P). Accordingly, these constraints may prevent the optimal detector from a practical implementation. This motivated us to propose the following sub-optimal detector.

3.3.2 Sub-Optimal Detector

In this section, we use the successive interference cancellation principle to tackle the problems associated with the optimal detector, while maintaining the diversity gain. The key principle of the sub-optimal detector is to employ two consecutive received packets, $\left\{y_D^{(p)}(n)\right\}_{n=1}^{N_d}$ and $\left\{y_D^{(p+1)}(n)\right\}_{n=1}^{N_d}$, given in (3.1) and (3.3), respectively to detect the packet $\left\{d^{(p)}(n)\right\}_{n=1}^{N_d}$. More specifically, let us remove the contribution of $\left\{d^{(p-1)}(n)\right\}_{n=1}^{N_d}$ from $\left\{y_D^{(p)}(n)\right\}_{n=1}^{N_d}$ and form

$$z_D^{(p)}(n) = y_D^{(p)}(n) - d^{(p-1)}(n)\psi_{R_x}(p-1)h_{R_xD}^{(p)}(n). \tag{3.14}$$

Using (3.1), we can write

$$z_D^{(p)}(n) = d^{(p)}(n)h_{SD}^{(p)}(n) + w_D^{(p)}(n). \tag{3.15}$$

Here we follow the common assumption associated with most of the successive detection algorithms found in the literature (e.g., [91–93]) that the previous symbol $\hat{d}^{(p-1)}(n)$ is perfectly detected, when we perform detection for $d^{(p)}(n)$. This assumption is necessary for the theoretical development of the sub-optimal detector;

Table 3.1: A comparison of optimal and sub-optimal detectors.

	complexity	delay	MS	BW loss
ML detector	$22M^2N_dP$	N_dP	N_dP	$1/P$
Sub-optimal detector	$33M^2N_dP + 8N_d(P-1)$	N_d	N_d	zero

however, in the simulations, we relax this assumption. With (3.3) rewritten as

$$y_D^{(p+1)}(n) = d^{(p+1)}(n)h_{SD}^{(p+1)}(n) + d^{(p)}(n)\psi_{R_y}(p)h_{R_yD}^{(p+1)}(n) + w_D^{(p+1)}(n), \quad (3.16)$$

our aim is to detect the symbol $d^{(p)}(n)$ from $z_D^{(p)}(n)$ and $y_D^{(p+1)}(n+1)$. Based on the ML principle, the detected symbol $\hat{d}^{(p)}(n)$ can be obtained by averaging over the symbol $d^{(p+1)}(n)$. Therefore, the ML detector of $\hat{d}^{(p)}(n)$ is

$$\hat{d}^{(p)}(n) = \arg \max_{d^{(p)}(n)} \sum_{d^{(p+1)}(n)\in\Omega} \log \Pr\left(\mathbf{y}'^{(p)}(n) \left| \mathbf{d}'^{(p)}(n), \mathbf{h}'^{(p)}(n)\right.\right) \Pr\left(d^{(p+1)}(n)\right), \quad (3.17)$$

where $\mathbf{y}'^{(p)}(n) = \left[z_D^{(p)}(n)\, y_D^{(p+1)}(n)\right]^T$, $\mathbf{d}'^{(p)}(n) = \left[d^{(p)}(n)\, d^{(p+1)}(n)\right]^T$, and

$$\mathbf{h}'^{(p)}(n) = \begin{bmatrix} h_{SD}^{(p)}(n) & 0 \\ h_{R_yD}^{(p+1)}(n)\psi_{R_y}(p) & h_{SD}^{(p+1)}(n) \end{bmatrix}, \quad (3.18)$$

and

$$\Pr\left(\mathbf{y}'^{(p)}(n) \left| \mathbf{d}'^{(p)}(n), \mathbf{h}'^{(p)}(n)\right.\right) = \tfrac{1}{\pi\sigma^2}\exp\left(-\left\|\mathbf{y}'^{(p)}(n) - \mathbf{h}'^{(p)}(n)\mathbf{d}'^{(p)}(n)\right\|^2/\sigma^2\right). \quad (3.19)$$

After detecting $\left\{d^{(p)}(n)\right\}_{n=1}^{N_d}$, we can remove its contribution from $\left\{y_D^{(p+1)}(n)\right\}_{n=1}^{N_d}$ forming $z_D^{(p)}(n+1) = y_D^{(p+1)}(n) - \hat{d}_{R_y}^{(p)}(n)\psi_{R_y}(p)h_{R_xD}^{(p+1)}(n)$. Similarly, from $\left\{z_D^{(p+1)}(n)\right\}_{n=1}^{N_d}$ and $\left\{y_D^{(p+2)}(n)\right\}_{n=1}^{N_d}$ and by averaging over $\left\{d^{(p+2)}(n)\right\}_{n=1}^{N_d}$, we can detect $\left\{d^{(p+1)}(n)\right\}_{n=1}^{N_d}$, and so on.

A comparison between the proposed ML and sub-optimal detectors in terms of complexity, delay, memory size (MS), and bandwidth loss (BW loss) is presented in Table 3.1. We evaluate the computational complexity of the proposed detectors in terms of the number of required floating point operations (see the Appendix). One can notice that the complexity of both detectors is proportional to the number of symbols per packet N_d, number of packets per frame P and square of the modulation order, M^2. We also notice that the sub-optimal detector outperforms the ML detectors in terms of the required delay, MS, and BW loss at the expense of a slightly computational complexity.

3.4 Data Detection for Multiple Antenna Destination

In this section, we assume that the destination is a base station that has the capability to support $A_D \geqslant 1$ antennas, while the source and relays are small nodes with limited power and few resources, thus, each of them is equipped only with one antenna. Our aim is to perform data detection at the destination.

3.4.1 Optimal Receiver

Based on Figure 3.1, the nth received symbol of the pth and $(p+1)$th received packets at the fth antenna of the destination can be respectively expressed as

$$y_D^{(f,p)}(n) = d^{(p)}(n)h_{SD}^{(f,p)}(n) + d^{(p-1)}(n)\psi_{R_x}(p-1)h_{R_xD}^{(f,p)}(n) + w_D^{(f,p)}(n), \quad (3.20)$$

$$y_D^{(f,p+1)}(n) = d^{(p+1)}(n)h_{SD}^{(f,p+1)}(n) + d^{(p)}(n)\psi_{R_y}(p)h_{R_yD}^{(f,p+1)}(n) + w_D^{(f,p+1)}(n), \quad (3.21)$$

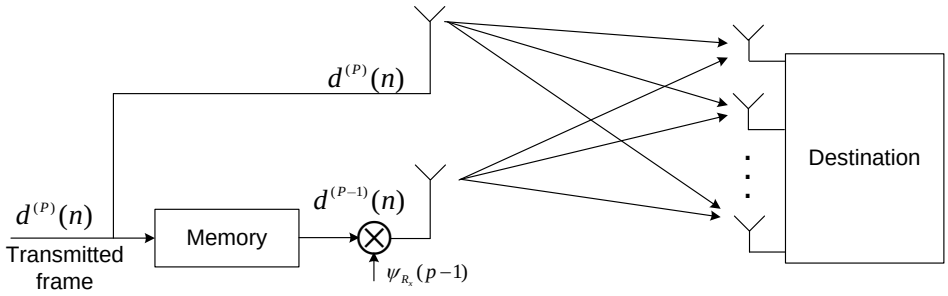

Figure 3.5: System model of the DF-ARC system, with a destination supporting multiple antenna.

where $h_{GD}^{(f,p)}(n)$ is the channel coefficient between the antenna of node G and the fth receive antenna of the destination, $G \in \{S, R_x, R_y\}$, $x, y \in \{1, 2\}$ and $x \neq y$, $w_D^{(f,p)}(n)$ and $w_D^{(f,p+1)}(n)$ are the corresponding additive white Gaussian noise (AWGN) at the fth antenna of the destination. Using the equivalent model as shown in Figure 3.5, we can apply the BCJR algorithm [76] to develop the optimal detector.

The key idea of the proposed optimal receiver is to employ A_D serial BCJR-based detectors; each is associated with one of the A_D antennas. Each detector delivers a posteriori probabilities (APPs) about the transmitted data which is used as a priori information for the next detector, as shown in Figure 3.6. For the first detector, we assume no prior information available. Upon completion, the APPs provided by the last detector are used to make final decisions on the transmitted data symbols.

Each detector employs N_d parallel BCJR algorithms, as shown in Figure 3.6. The input of the nth BCJR algorithm is the nth symbols from each packet received on the fth antenna, $\mathbf{y}^{(f)}(n) = \left[y_D^{(f,1)}(n), y_D^{(f,2)}(n), \cdots, y_D^{(f,P)}(n)\right]$, with $y_D^{(f,p)}(n)$ and $y_D^{(f,p+1)}(n)$ defined in (3.20) and (3.21), respectively. From (3.20) and (3.21), one can notice that each transmitted symbol $d^{(p)}(n)$ is included in $y_D^{(f,p)}(n)$ and $y_D^{(f,p+1)}(n)$. Accordingly, we mix the received packets on the fth antenna, as shown in Figure

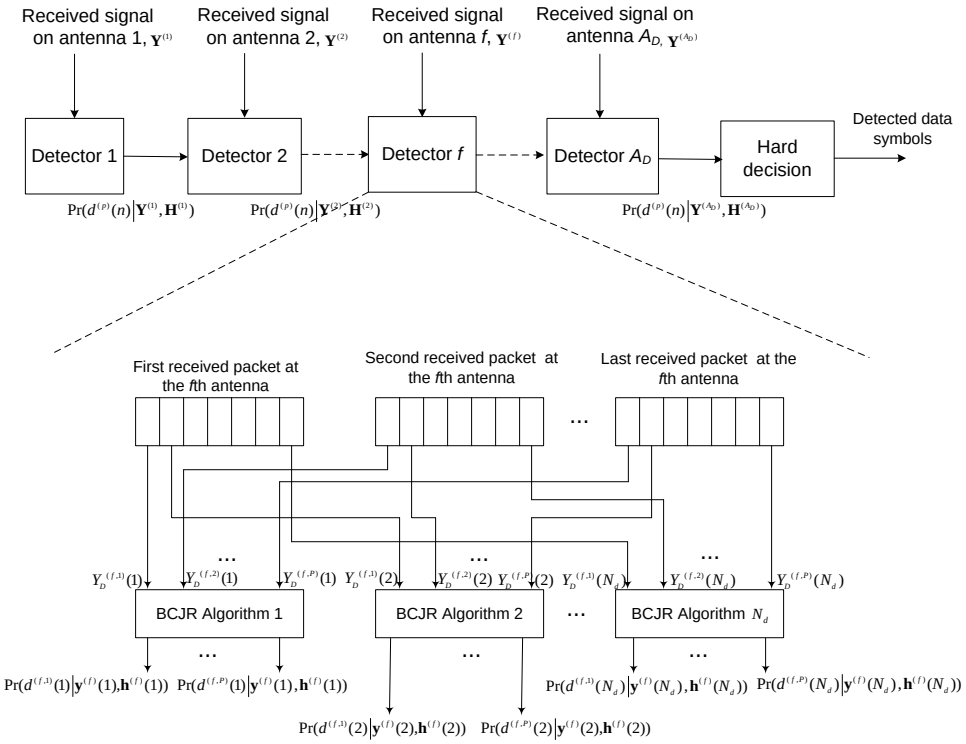

Figure 3.6: Optimal receiver for multiple receive antennas alternate-relaying DF cooperative system. $\mathbf{H}^{(f)}$ is the vector that includes the channel coefficients between the source and relays and the fth antenna of the destination.

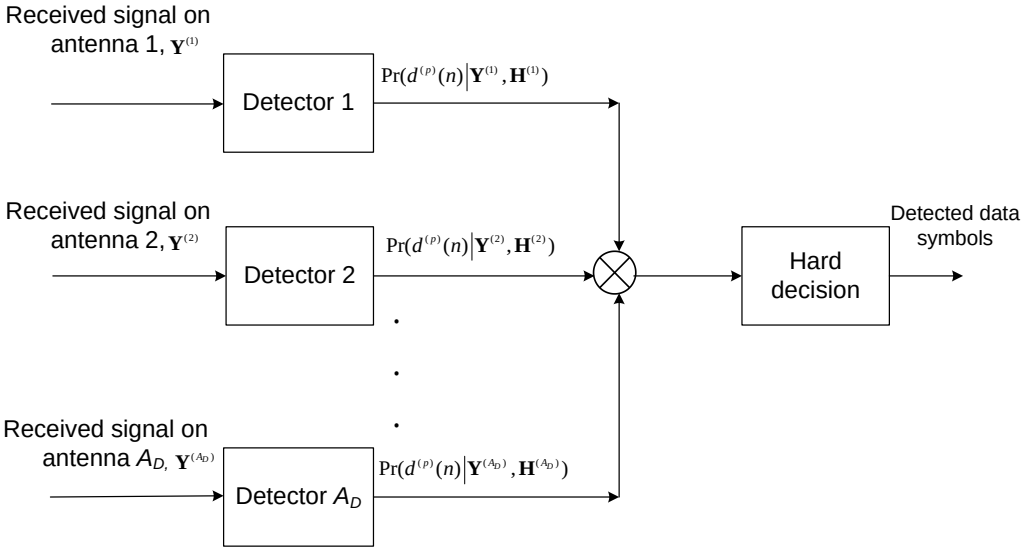

Figure 3.7: Parallel representation of the optimal detector. $\mathbf{H}^{(f)}$ is the vector that includes the channel coefficients between the source and relays and the fth antenna of the destination.

3.6, in order to exploit the repetition of the transmitted symbols. The nth BCJR algorithm of the fth detector computes the APPs of the nth symbol of the pth received packet on the fth antenna, $\Pr\left(d^{(p)}(n)\left|\mathbf{y}^{(f)}(n),\mathbf{h}^{(f)}(n)\right.\right)$, where $\mathbf{h}^{(f)}(n) = \left[h_{SD}^{(f,1)}(n), h_{R_xD}^{(f,1)}(n), h_{SD}^{(f,2)}(n), h_{R_yD}^{(f,2)}(n), \cdots, h_{SD}^{(f,P)}(n), h_{R_xD}^{(f,P)}(n)\right]$ represents the channel coefficients, and $p = 1, \cdots, P$. Using (3.20) and following the principles of the BCJR algorithm which is explained in Section 2.2.4, one can write the branch metric between the states s' and s for the nth symbol in the pth packet given the received symbol $y_D^{(p)}(n)$ as

$$\gamma_n^{(f,p)}(s',s) = \frac{1}{\sqrt{\pi\sigma^2}}\Pr\left(d^{(p)}(n)\right) \times \exp\left(-\left|y_D^{(f,p)}(n) - d^{(p)}(n)h_{SD}^{(f,p)}(n) - d^{(p-1)}(n)\psi_{R_x}(p-1)h_{R_xD}^{(f,p)}(n)\right|^2/\sigma^2\right), \tag{3.22}$$

where $\left[d^{(p)}(n)\, d^{(p-1)}(n)\right]$ represents the output associated with this branch and $\Pr\left(d^{(p)}(n)\right)$ is the a priori probability.

Implementation Aspects

- As the proposed optimal receiver is based on sequential detection, the required processing time for data detection is A_D times of the processing time of one detector. In order to reduce the processing time by the factor A_D, the optimal receiver can be performed in a parallel fashion, as shown in Figure 3.7. Here we exploit the fact that the APPs of data symbols are the multiplication of the BCJR output in the case of no a priori probabilities available and the a priori probabilities of these data symbols (e.g., [77]).

- When the destination node has only one receive antenna, $A_D = 1$, the optimal receiver can be implemented by using Viterbi algorithms instead of the BCJR algorithms. The reason behind this is that the BCJR and Viterbi algorithms have the same bit-error-rate performance when there is no a priori

information available about data symbols, and the computational complexity of the Viterbi algorithms is approximately one-third of that of the BCJR algorithms [67,76].

- Since the BCJR algorithm is a backward recursion, the end state must be known. An additional N zero symbols are included at the end of the transmitted frame to drive the "virtual" encoder to state zero. In the literature, these additional zeros are referred to as the tail [67,90,94]. This will lead to a slight bandwidth loss compared with the direct transmission. This loss is $\frac{1}{P}$, which approaches zero for a large number of packets per frame.

- The proposed optimal receiver requires a large memory size and delay. The destination has to receive and store the entire frame before starting data detection. These limitations are an increasing function of the modulation order (M) and the number of packets per frame (P). Accordingly, these constraints may prevent the optimal detector from a practical implementation. This motivates us to propose the following sub-optimal detector.

3.4.2 Sub-Optimal Detector

In this section, we employ the successive interference cancellation principle to avoid the delay problem associated with the optimal receiver, while maintaining the diversity gain. The basic idea behind the sub-optimal receiver is to use A_D sub-optimal detectors, and as shown in Figure 3.8, each detector is associated with one of the A_D receive antennas. The sub-optimal receiver exploits two consecutive received packets on each receive antenna, $\left\{y_D^{(f,p)}(n)\right\}_{n=1}^{N_d}$ and $\left\{y_D^{(f,p+1)}(n)\right\}_{n=1}^{N_d}$, $f = 1, \cdots, A_D$, given in (3.20) and (3.21), respectively to detect the packet $\left\{d^{(p)}(n)\right\}_{n=1}^{N_d}$. The fth detector operates as follows. We remove the

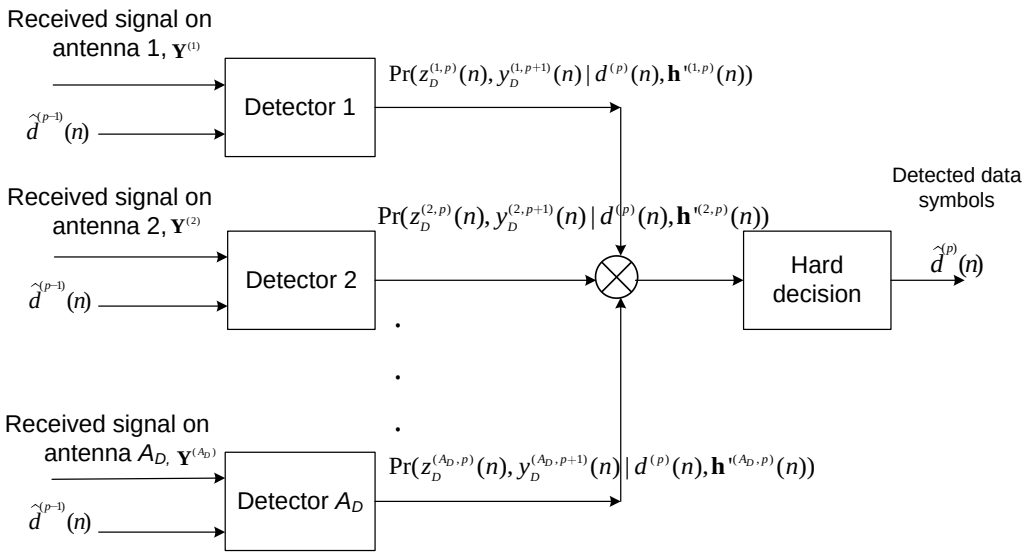

Figure 3.8: Sub-optimal receiver for DF alternate-relaying cooperative systems.

contribution of $\left\{\hat{d}^{(p-1)}(n)\right\}_{n=1}^{N_d}$ from $\left\{y_D^{(f,p)}(n)\right\}_{n=1}^{N_d}$ and form

$$z_D^{(f,p)}(n) = y_D^{(f,p)}(n) - \hat{d}^{(p-1)}(n)\psi_{R_x}(p-1)h_{R_xD}^{(f,p)}(n). \tag{3.23}$$

Using (3.20), we can write

$$z_D^{(f,p)}(n) = d^{(p)}(n)h_{SD}^{(f,p)}(n) + w_D^{(f,p)}(n). \tag{3.24}$$

Here we follow the common assumption associated with most of the successive detection algorithms found in the literature (e.g., [91,92]) that the previous symbol $\hat{d}^{(p-1)}(n)$ is perfectly detected, when we perform detection for $d^{(p)}(n)$. This assumption is necessary for the theoretical development of the sub-optimal receiver; however, in the simulations we relax this assumption. With (3.21) rewritten as

$$y_D^{(f,p+1)}(n) = d^{(p+1)}(n)h_{SD}^{(f,p+1)}(n) + d^{(p)}(n)\psi_{R_y}(p)h_{R_yD}^{(f,p+1)}(n) + w_D^{(f,p+1)}(n). \tag{3.25}$$

Therefore, we can write

$$\Pr\left(z_D^{(f,p)}(n), y_D^{(f,p+1)}(n) \left| d^{(p)}(n), \mathbf{h}'^{(f,p)}(n)\right.\right) = \sum_{d^{(p+1)}(n)\in\Omega} \Pr\left(z_D^{(f,p)}(n), y_D^{(f,p+1)}(n) \left| d^{(p)}(n), d^{(p+1)}(n), \mathbf{h}'^{(f,p)}(n)\right.\right) \Pr\left(d^{(p+1)}(n)\right), \tag{3.26}$$

where

$$\mathbf{h}'^{(f,p)}(n) = \begin{bmatrix} h_{SD}^{(f,p)}(n) & 0 \\ h_{R_yD}^{(f,p+1)}(n)\psi_{R_y}(p) & h_{SD}^{(f,p+1)}(n) \end{bmatrix}, \tag{3.27}$$

and

$$\Pr\left(z_D^{(f,p)}(n), y_D^{(f,p+1)}(n) \left| d^{(p)}(n), d^{(p+1)}(n), \mathbf{h}'^{(f,p)}(n)\right.\right) = \frac{1}{\pi\sigma^2}\exp\left(-\left\|\left[z_D^{(f,p)}(n), y_D^{(f,p+1)}(n)\right]^T - \mathbf{h}'^{(f,p)}(n)\left[d^{(p)}(n), d^{(p+1)}(n)\right]^T\right\|^2/\sigma^2\right), \tag{3.28}$$

where the superscript T refers to the vector transpose. Based on the channel information provided by each sub-optimal detector and using the maximum-likelihood criterion, the detected value of $d^{(p)}(n)$ can be estimated as

$$\hat{d}^{(p)}(n) = \arg\max_{d^{(p)}(n)} \prod_{f=1}^{A_D} \Pr\left(z_D^{(f,p)}(n), y_D^{(f,p+1)}(n) \left| d^{(p)}(n), \mathbf{h}'^{(p)}(n)\right.\right).$$

After detecting $\left\{d^{(p)}(n)\right\}_{n=1}^{N_d}$, its contribution can be removed from $\left\{y_D^{(f,p+1)}(n)\right\}_{n=1}^{N_d}$ forming $z_D^{(f,p)}(n+1) = y_D^{(f,p+1)}(n) - d^{(p)}(n)\psi_{R_y}(p)h_{R_yD}^{(f,p+1)}(n)$. Similarly, from $\left\{z_D^{(p+1)}(n)\right\}_{n=1}^{N_d}$ and $\left\{y_D^{(p+2)}(n)\right\}_{n=1}^{N_d}$ and by averaging over $\left\{d^{(p+2)}(n)\right\}_{n=1}^{N_d}$, we can detect $\left\{d^{(p+1)}(n)\right\}_{n=1}^{N_d}$, and so on.

A comparison between the proposed optimal and sub-optimal detectors in terms of complexity, delay, memory size, and bandwidth loss is shown in Table 3.2. The abbreviation MS stands for the required memory size (MS) while the BW

Table 3.2: A comparison of optimal and sub-optimal detectors, with a destination supporting multiple antenna.

	Complexity	Delay	MS	BW loss
Optimal	$70M^2N_dPA_D$	N_dP	N_dPA_D	$1/P$
Sub-optimal	$\left(33M^2N_dP + 8N_d(P-1)\right)A_D$	N_d	N_dA_D	zero

loss refers to the bandwidth loss due to the all-zero packet sent at the end and the delay is expressed by the number of data symbols required to store. Similar to the analysis in the Appendix, the computational complexity of the proposed detectors is evaluated in terms of the number of required floating point operations. One can notice that the complexity of both detectors is proportional to the number of symbols per packet N_d, number of packets per frame P and square of the modulation order, M^2. We also notice that the sub-optimal detector outperforms the ML detectors in terms of the required delay, MS, and BW loss at the expense of a slightly computational complexity.

3.5 Simulation Results

In this section, we present simulation results to illustrate the performance of the proposed detectors. For each link, the channel is assumed to be frequency non-selective and modeled as a zero-mean independent complex Gaussian random variable. To capture the effect of the path loss on the performance, we consider $\mathrm{E}\left[|h_{AB}|^2\right] = (d_{SD}/d_{AB})^{\epsilon}$ [11], where h_{AB}, d_{SD}, and d_{AB} are the channel coefficient, the distance between source and destination, and the distance between nodes A and B, respectively, and ϵ is the path loss exponent. The channel coefficients remain constant over the packet length. A frame based transmission is assumed. The values given here for the number of packets, the distances, and the path-loss

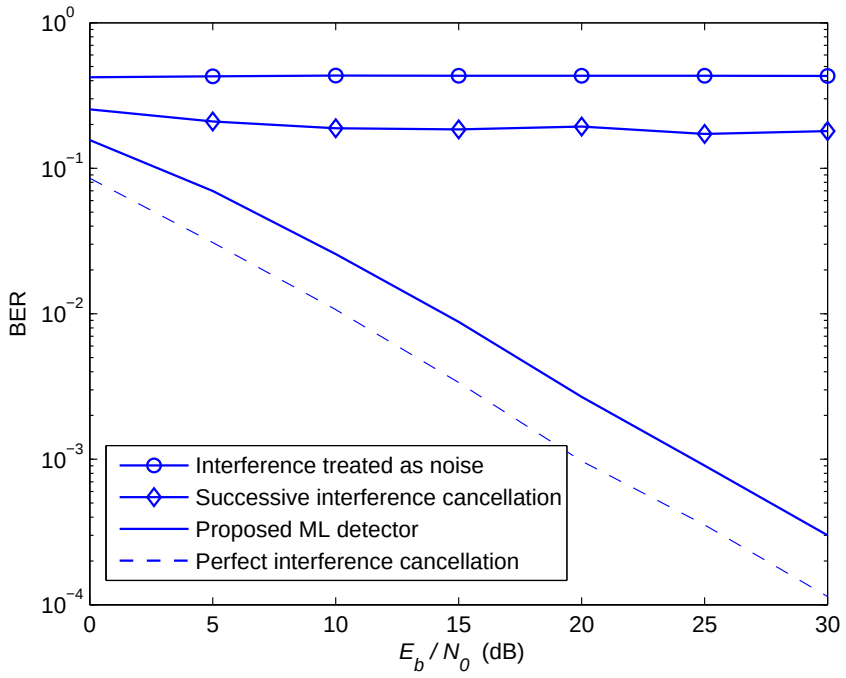

Figure 3.9: BER of the proposed ML detector at the relays.

exponents are the default ones and they are the used values unless otherwise mentioned. Each frame has P =20 packets. The packet length is 150 information bits, after modulation, this leads to N_d = 150/m symbols, with m as the number of bits per symbol. The distance between the source and the two relays equals 0.4, the distance between the two relays and the destination equals 0.6208, and the distance between the two relays is 0.2. All these distances are normalized to the source to destination distance. Unless mentioned otherwise, 8-PSK modulation is used and $\epsilon = 2$.

At Relays

Figure 3.9 shows the BER performance of the proposed ML detector at the relays. For the sake of comparison, we also show the performance when the interference signal could perfectly be removed. In the sequel, this is referred to as the perfect

interference cancellation. In addition, we show the BER performance of the successive interference cancellation detector proposed in [38, 59], which is based on detecting the strongest signal first, then subtracting its contributions from the received interfered signal before detecting the other signal. As one can observe, the detector that treats the interference as additional noise leads to unacceptable performance degradation when compared with the case of perfect interference cancellation. This is because the inter-relay link is not weak enough when compared with the source to relays links. Furthermore, the performance of the sequential detector, which is not acceptable as well, can be justified as follows. Bearing in mind that the channel coefficients are random and all transmissions use the same transmit power level, we can not guarantee that the instantaneous value of one of the two terms $\left|d^{(p)}(n)h^{(p)}_{SR_x}(n)\right|^2$ and $\left|d^{(p-1)}(n)\psi_{R_y}(p-1)h^{(p)}_{R_xR_y}(n)\right|^2$ is much greater than the other term, although $\mathrm{E}\left[\left|h^{(p)}_{R_1R_2}(n)\right|^2\right] = 4\mathrm{E}\left[\left|h^{(p)}_{SR_1}(n)\right|^2\right] = 4\mathrm{E}\left[\left|h^{(p)}_{SR_2}(n)\right|^2\right]$ as we considered in the simulations. A significant improvement in performance is achieved by applying the proposed ML-based detector, for which the BER performance degradation reduces to about 5 dB.

Figure 3.10 shows the BER performance of the proposed detector at the relays for different modulation orders. As one can observe, the difference between the performance of the ML proposed detector and the case of perfect interference cancellation increases with the modulation order. This occurs because increasing the modulation order yields to averaging the interference signal over a large number of possible transmitted symbols, and thus disturbing the detector.

Figure 3.11 illustrates the BER performance of the proposed detector at the relays for different values of the path-loss exponent, ϵ, at the relays. It turns out that the difference between the performance of the ML proposed detector and the case of perfect interference cancellation decreases with the path-loss exponent.

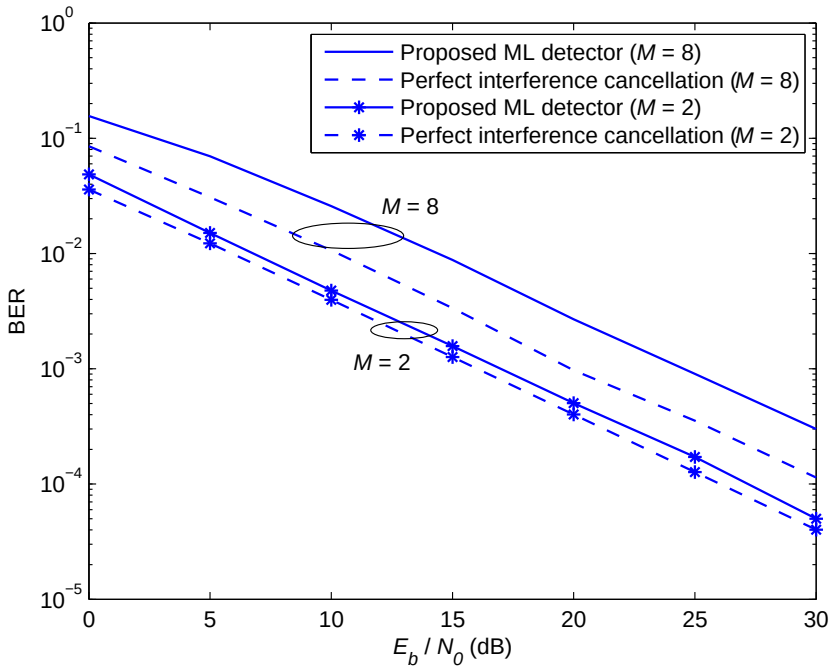

Figure 3.10: Influence of the modulation order, M, on the BER of the proposed ML detector at the relays.

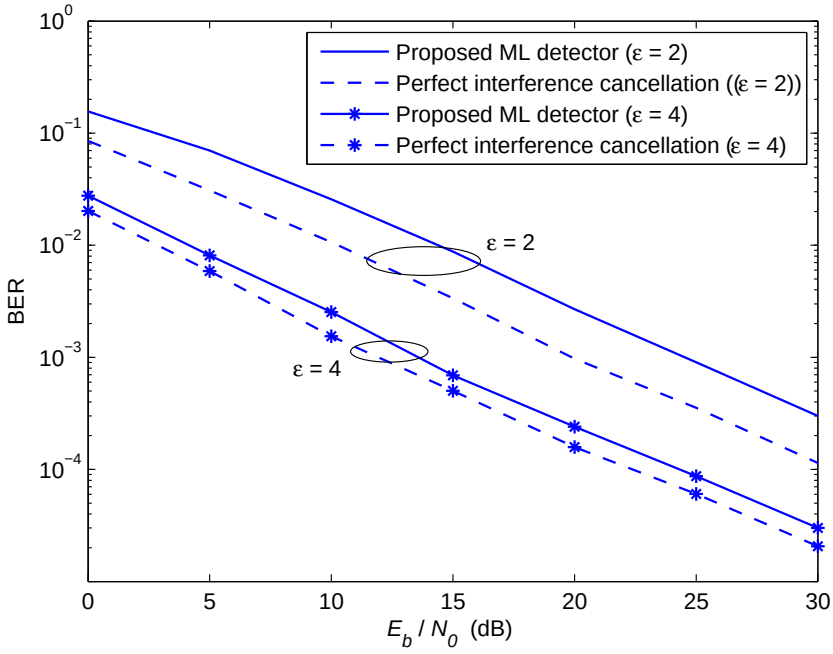

Figure 3.11: Influence of the path-loss exponent, ϵ, on the BER of the proposed ML detector at the relays.

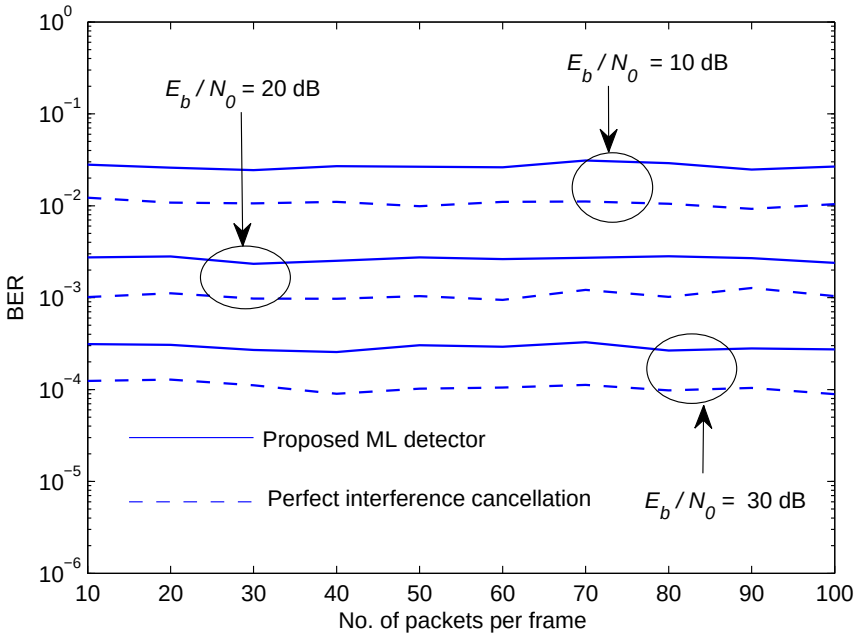

Figure 3.12: Influence of the number of packets per frame on the BER of the proposed detector at the relays.

Figure 3.12 shows the BER performance of the proposed detector as a function of the number of packets per frame at the relays. It is obvious that the BER performance is independent of the number of packets per frame. This can be explained as follows. At the relays, the proposed ML receiver is based on a symbol-by-symbol detection procedure. Accordingly, the influence of the frame length on the BER performance can be negligible provided that the number of bits per packet is kept constant.

Figure 3.13 depicts the BER performance of the proposed detectors as a function of the number of bits per packet at the relays. The results indicate that, at the relay, the BER performance is independent of the number of bits per packet. This is because the detection process is performed symbol-by-symbol as we explained in the proceeding paragraph.

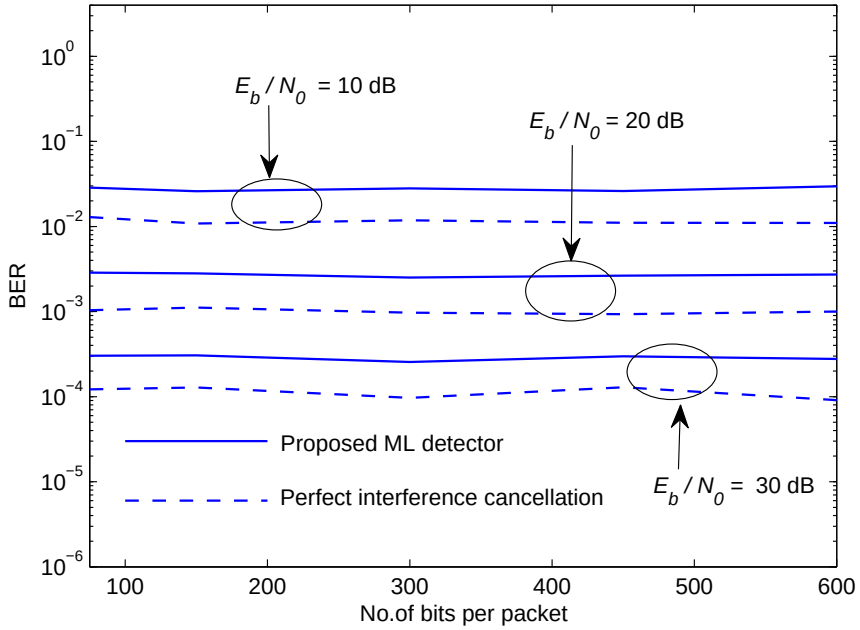

Figure 3.13: Influence of the number of bits per packet on the BER of the proposed detector at the relays.

Single-Antenna Destination

Figure 3.14 depicts the BER performance of the proposed detectors at the destination. The BER performance of the half-rate one relay and best-relay selection from a set of two available relays are also shown. For a fair comparison between the half-rate and full-rate systems, we set the same transmitted power. Furthermore, we apply 64-PSK modulation for the half-rate systems and 8-PSK for the full-rate systems in order to maintain the same data rate. We also show the performance when the interference could perfectly be canceled. In other words, we assume that the contribution of the packets $(p-1)$th and $(p+1)$th is perfectly removed during detection of the pth packet. As a consequence, the pth packet can be detected by employing the classical maximum ratio combining technique as seen in Figure 3.15. In addition, we show the BER performance of the DF alternate-relaying proposed

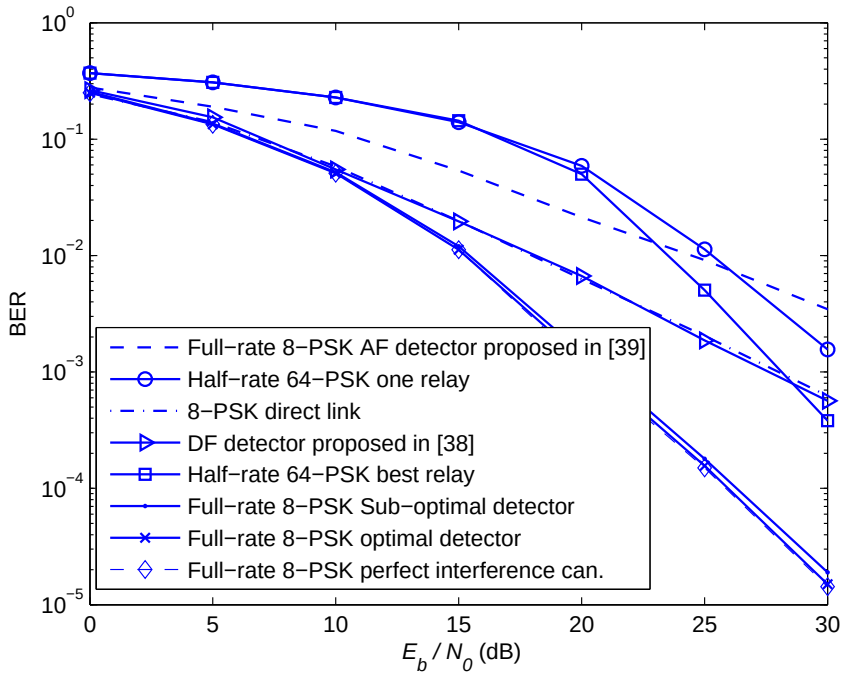

Figure 3.14: BER of the proposed ML detectors at the destination.

in [38] and the latest work on AF alternate-relaying [39]. As one can observe, the performance of the proposed full-rate cooperative systems outperforms that of the half-rate cooperative systems and full-rate cooperative systems proposed in [38] and [39]. For example, at BER = 10^{-3}, the proposed optimal detector outperforms the half-rate best relay, DF detector [38], and AF detector [39] by 8 dB, 7 dB, and 15 dB, respectively[6]. Moreover, the optimal detector outperforms the sub-optimal detector by around 0.25 dB. In addition, we notice that the performance of the optimal detector coincides with the case of perfect interference cancellation.

Figures 3.16 and 3.17 show the BER performance at the destination of the proposed detectors for different modulation orders and different path-loss exponent values. As one can observe, the performance of the sub-optimal and optimal de-

[6]It is worthy to mention that the work in [38] does not take the direct link into account, and as such, no diversity gain is obtained at the destination. Furthermore, the full interference cancellation process proposed in [39] results in extra noise components that limit the BER performance.

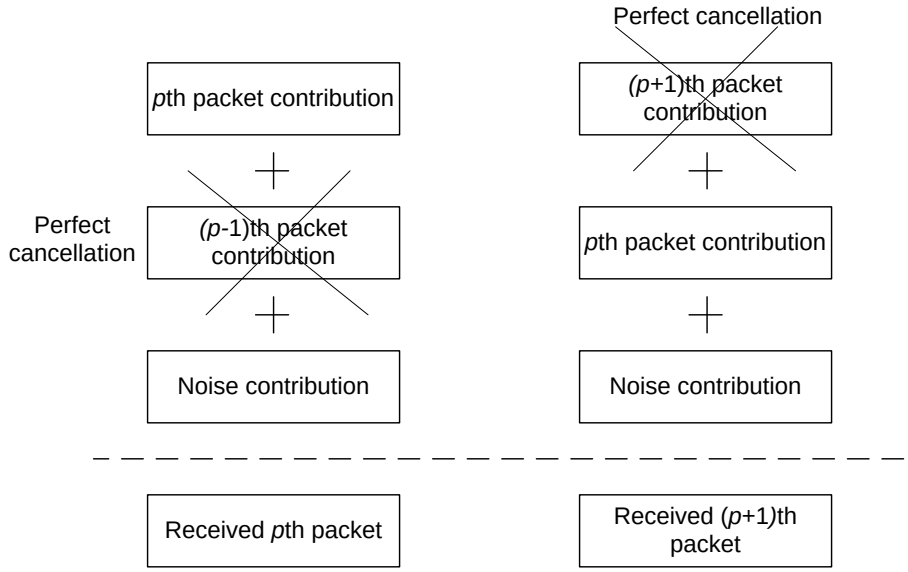

Figure 3.15: Perfect interference cancellation at the destination.

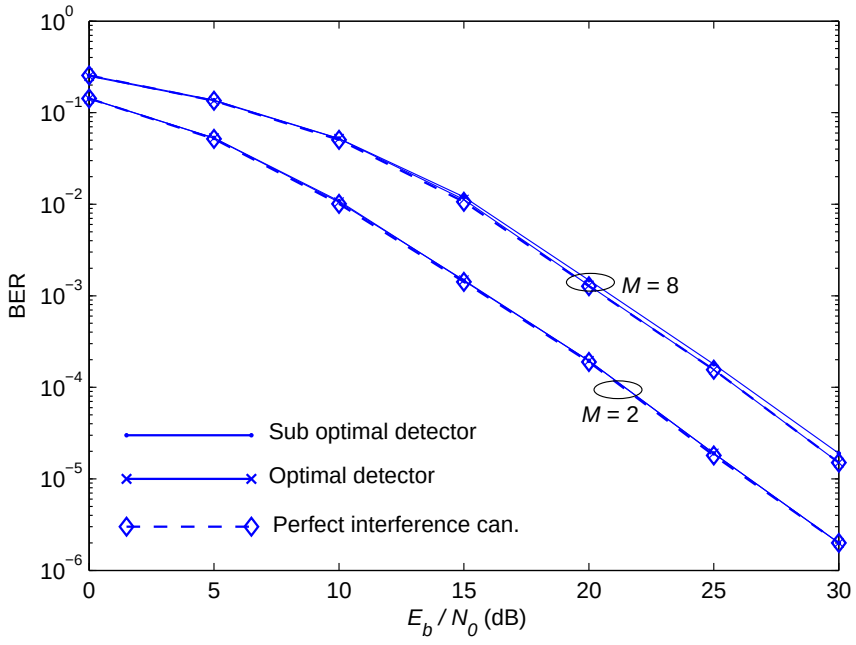

Figure 3.16: Influence of the modulation order, M, on the BER of the proposed detectors at the destination.

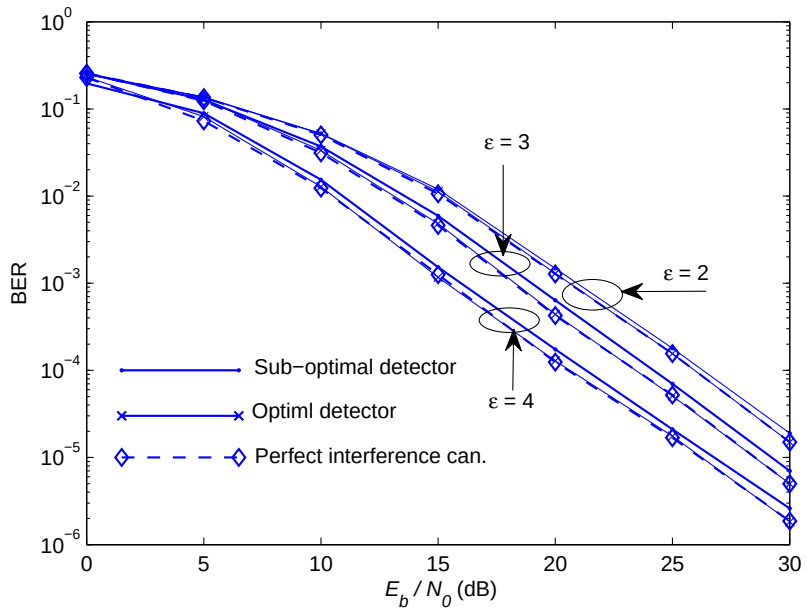

Figure 3.17: Influence of the path-loss exponent, ϵ, on the BER of the proposed detectors at the destination.

tectors almost coincides with the case of perfect interference cancellation.

Figures 3.18 and 3.19 show the BER performance of the proposed detectors at the destination for different numbers of packets per frame and number of bits per packet, respectively. It is obvious that the BER performance of the proposed detectors is independent of the number of packets per frame. Furthermore, one can observe that the dependency of the BER performance on the number of bits per packet increases when E_b/N_o increases, although the dependency is small, as shown in Figure 3.19. This is explained as at the high E_b/N_0 the effect of noise diminishes and this clarifies the influence of the bits per packets.

In the following discussion, we show the effect of the strength of the inter-relay link on the overall BER performance. This can not be seen purely in a two-dimensional geometry because if the distance between the two relays changes, the distance between the source and the two relays and the distance between two re-

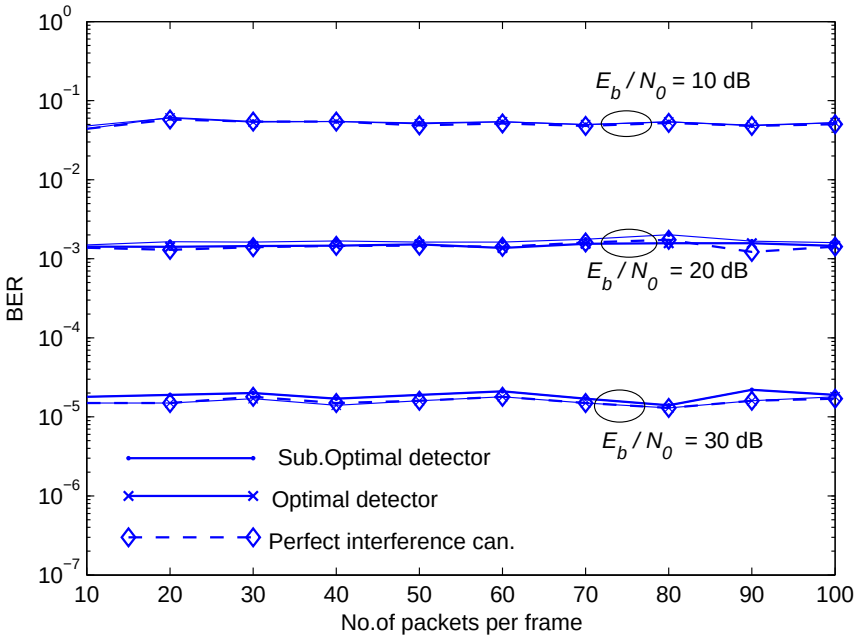

Figure 3.18: Influence of the number of packets per frame, P, on the BER of the proposed detectors at the destination.

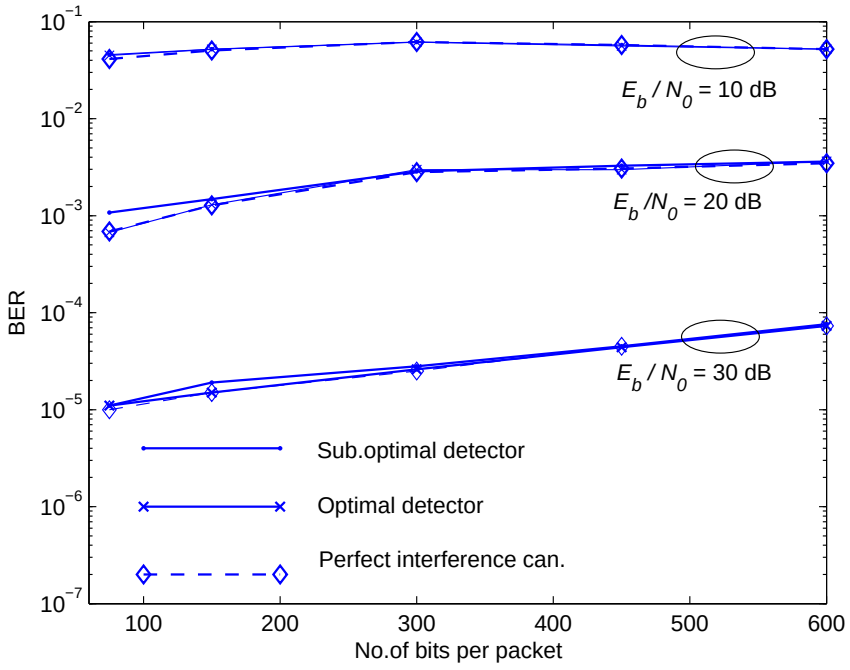

Figure 3.19: Influence of the number of bits per packet on the BER of the proposed detector at the destination.

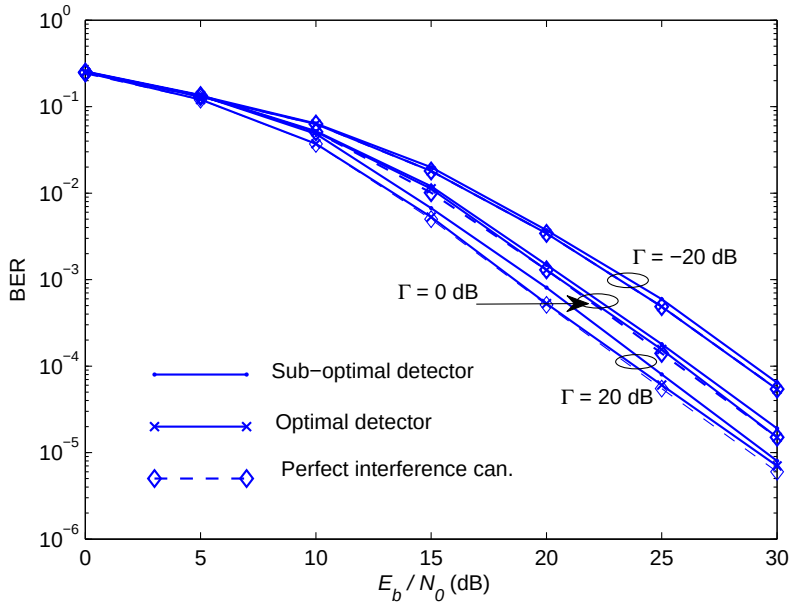

Figure 3.20: BER performance of the proposed detectors at the destination, in the presence of shadowing for the inter-relay link.

lays and destination change, as well. To overcome this problem, we consider shadowing in the two-relay link, while the other links are not be affected by shadowing. For simulation purpose, we consider $\mathrm{E}\left[\left|h_{R_1R_2}\right|^2\right] = (d_{SD}/d_{R_1R_2})^{\epsilon}\, 10^{\Gamma/10}$, with Γ as the shadowing term expressed by dB[7]. Figure 3.20 shows the effect of Γ on the overall BER performance at the destination, for 8-PSK modulation. One can observe that the proposed optimal and sub-optimal detectors provide a good BER performance for a wide range of Γ. In addition, the BER performance of the proposed detectors improves with the increase of Γ. This is because increasing Γ enhances the reliability of the inter-relay link, which in turn decreases the BER performance at the relays, thus improving the overall BER performance at the destination.

[7]In wireless literature, the shadowing term Γ is normally modeled as a random variable drawn from a Gaussian distribution with a mean of 0 dB. However, here we consider Γ as a constant to clearly show the effect of the inter-relay link on the BER performance.

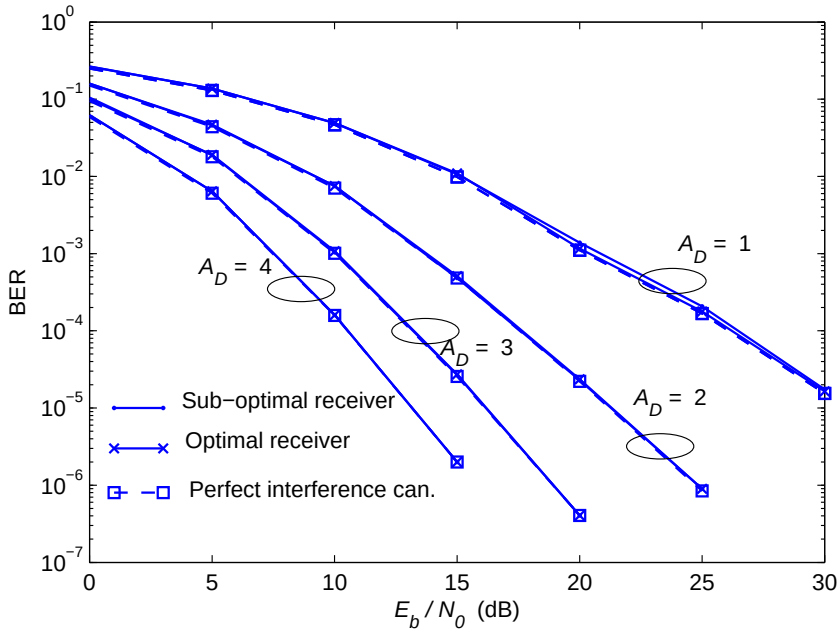

Figure 3.21: BER performance of the proposed receivers.

Multiple-Antenna Destination

Figure 3.21 depicts the BER performance of the proposed receivers at the destination for $A_D = 1$, 2, 3, and 4. We also show the performance when the interference could perfectly be canceled, i.e, the contribution of the packets $(p-1)$th and $(p+1)$th is perfectly removed during detection of the pth packet. As a consequence, the pth packet can be detected by employing the classical maximum ratio combining technique. In the sequel, this is referred to as the perfect interference cancellation scenario. This can be considered as a lower bound of the BER performance. As one can observe, the proposed optimal and sub-optimal receivers have approximately the same BER permanence, which is very close to the lower bound of the BER performance.

Previous results have been restricted to the idealistic case of independent and identically distributed (i.i.d.) channels, i.e., uncorrelated fading. However, in prac-

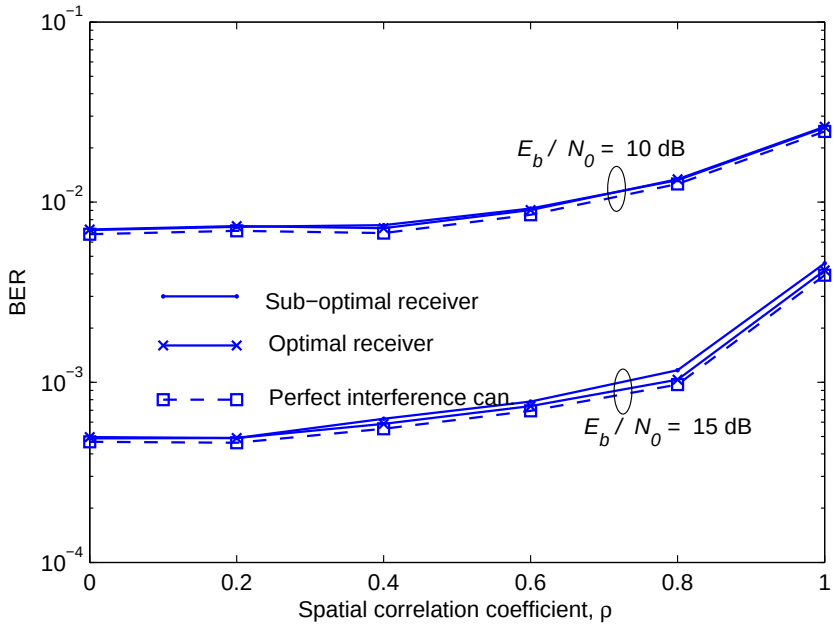

Figure 3.22: Effect of the spatial correlation on the BER performance of the proposed receivers for $A_D = 2$.

tice, the channels may be dependent due to, for example, insufficient spacing between antenna elements [95,96]. The spatial correlation is known to be detrimental for the performance of multi antenna systems as reported in many other studies (e.g., [95, 96]). Figure 3.22 investigates the effect of spatial correlation on the proposed receivers for a destination node supporting two receive antennas. For simplicity, it is assumed that the following channels have the same spatial correlation coefficient, ρ: source-destination (antenna one), source-destination (antenna two), relays (R_1 and R_2)-destination (antenna one), and relays (R_1 and R_2)-destination (antenna two). As one observes from the figure, the proposed receivers are robust to spatial fading correlation, up to ρ =0.6.

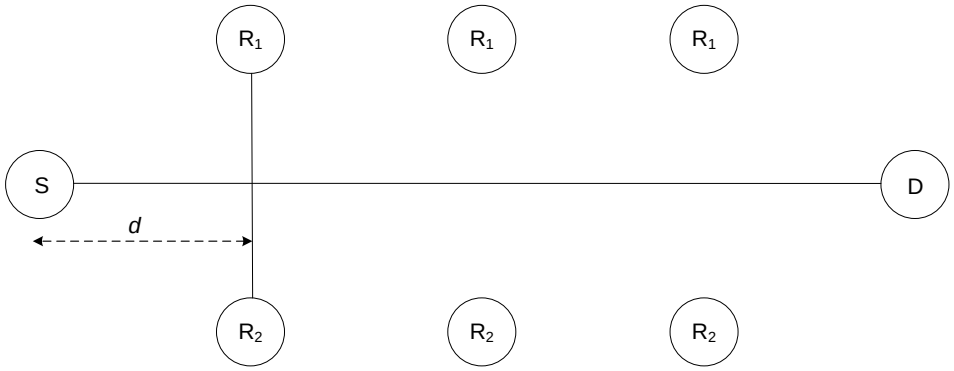

Figure 3.23: Relays positions for scenario 1, $d_{R_1R_2} = 0.2$.

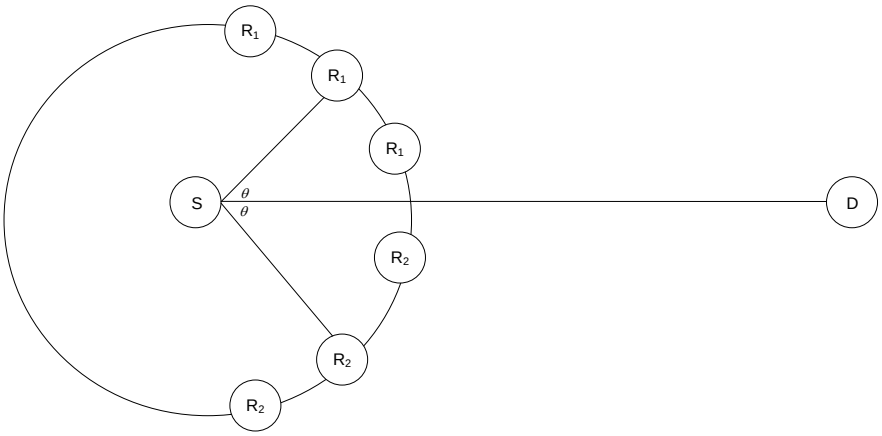

Figure 3.24: Relays positions for scenario 2, $d_{SR_1} = d_{SR_2} = 0.4$.

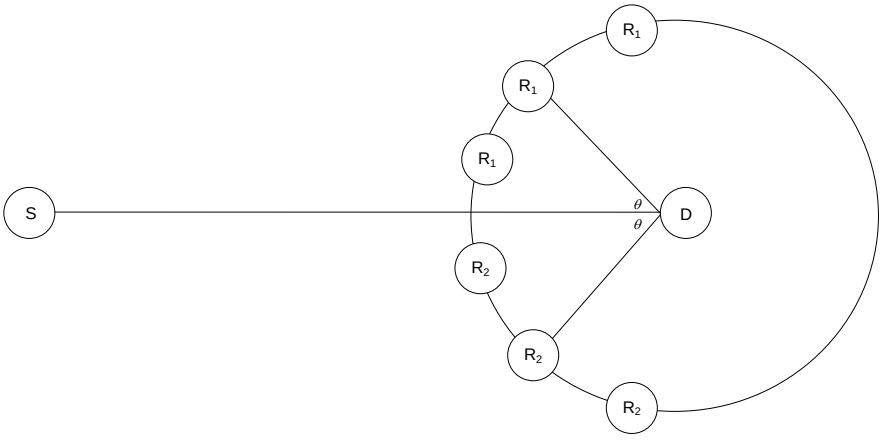

Figure 3.25: Relays positions for scenario 3, $d_{R_1D} = d_{R_2D} = 0.4$.

Relays positions

In the following, we investigate the effect of relays position on the BER performance of the DF-ARC system with a single-antenna destination. We consider three scenarios as shown in Figures 3.23, 3.24, and 3.25. In the first scenario, the distance between the two relays is kept constant of 0.2. In the second scenario, the distance between the source and two relays is kept constant of 0.4. In the third scenario, the distance between the destination and two relays is kept constant of 0.4. All these distances are normalized to the source to destination distance. The corresponding BER performance of the three scenarios is shown in Figures 3.26-3.31, where d is defined as in Figure 3.23 and θ is indicated as in Figures 3.24 and 3.25. Generally, increasing the distance between the source and relays leads to source-relays channel degradation which in turn increases the BER at the relays. This is clearly observed in Figures 3.26 and 3.28. Figure 3.27 shows that the difference between the BER of the proposed detector in the presence of interference and the BER when the interference could be removed (perfect interference cancellation) increases with θ, and the maximum of the difference occurs at $\theta = 90$. This can easily be explained as at this particular point the useful signal and interference components of the received signal have the same power, thus it is difficult for the proposed detector to extract the source information from the received signal. This observation can be also seen in the first scenario, when $d = 0$ as in Figure 3.26. Figures 3.29, 3.30, and 3.31 indicate that the corresponding BER at the destination for scenarios 1, 2, and 3 does not relatively depend on the relays positions.

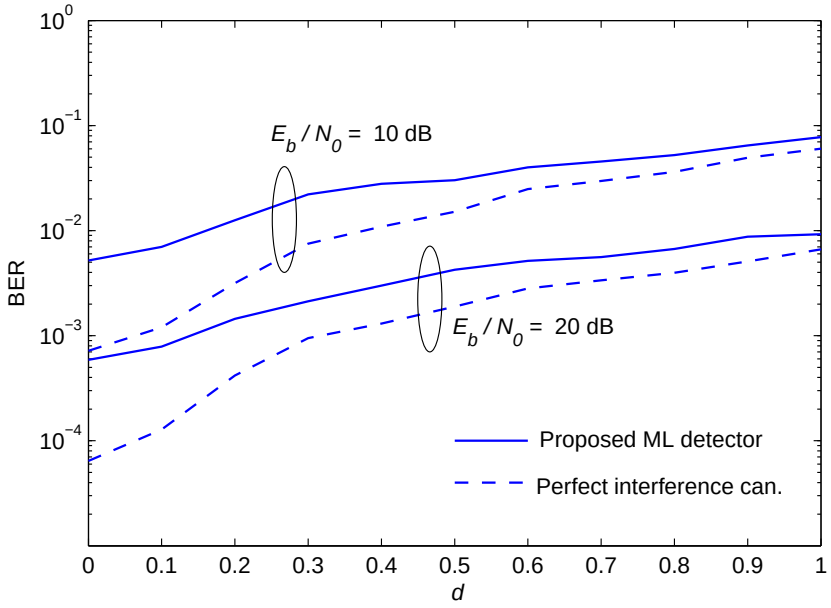

Figure 3.26: BER performance at relays of scenario 1.

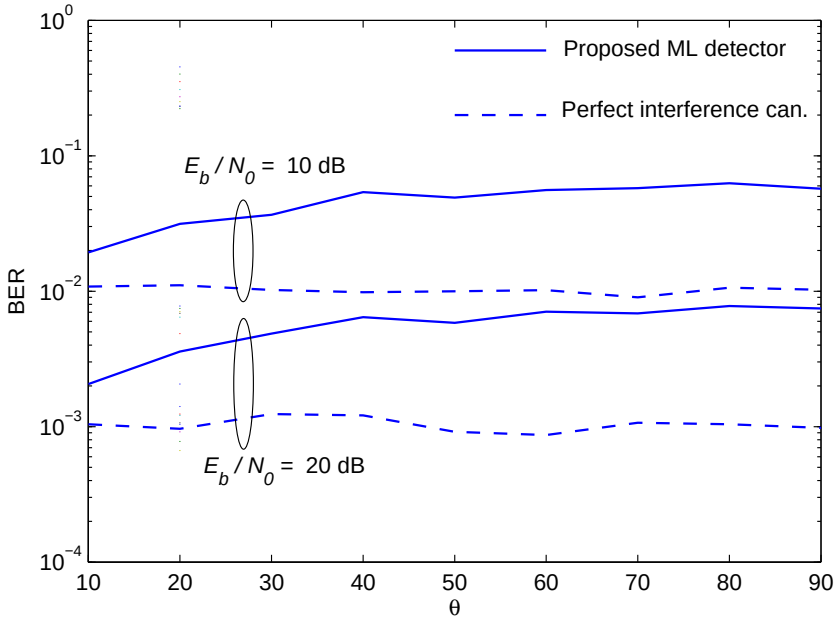

Figure 3.27: BER performance at the relays of scenario 2.

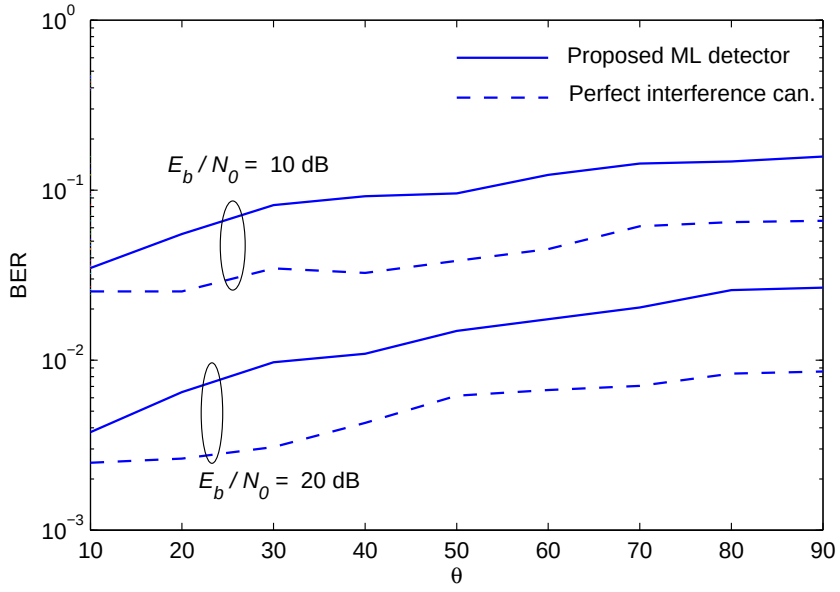

Figure 3.28: BER performance at the relays of scenario 3.

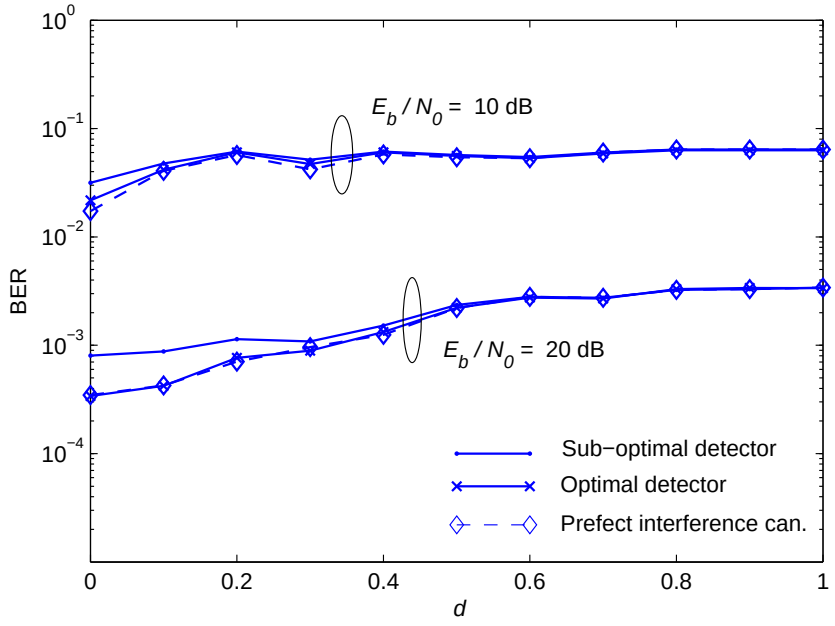

Figure 3.29: BER performance at the destination of scenario 1.

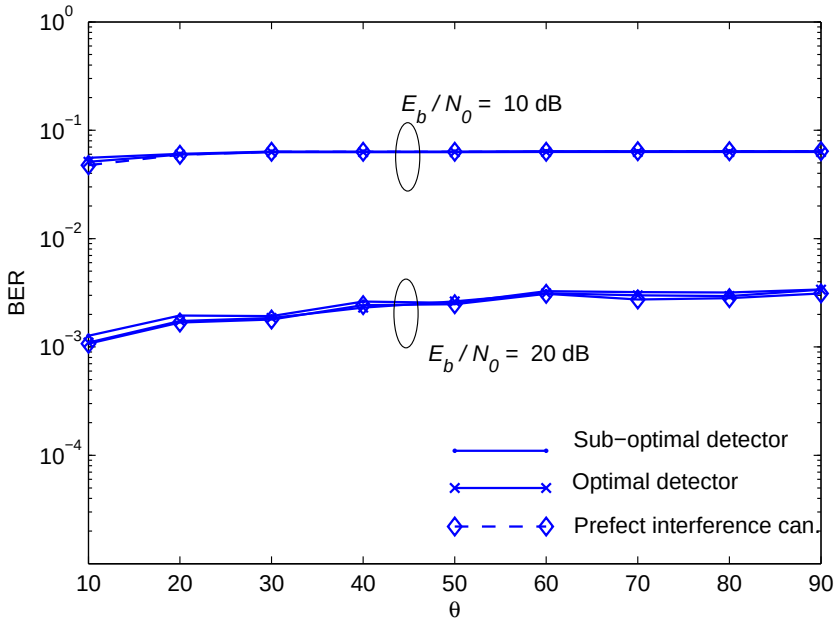

Figure 3.30: BER performance at the destination of scenario 2.

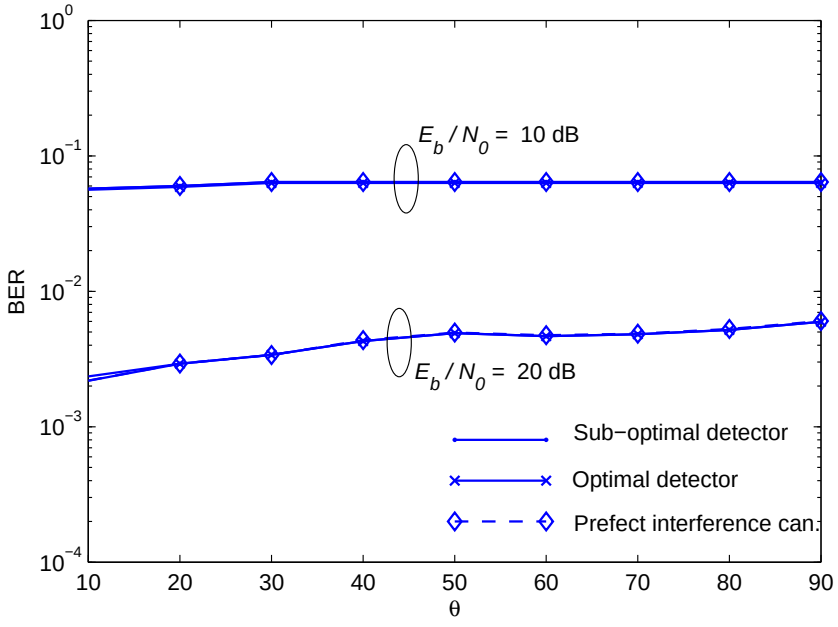

Figure 3.31: BER performance at the destination of scenario 3.

3.6 Summary

In this chapter, we have proposed a maximum-likelihood detector at the relays. This is based on averaging out the interference signal. Furthermore, unlike previous work in which interference cancellation is required at the destination, we exploit the interference signal as a beneficial resource to develop the optimal detector at the destination. It is shown that the optimal detector can be implemented by parallel Viterbi algorithms. The major drawback of the proposed optimal detector is the delay because the destination has to receive and store the entire frame before performing data detection. Due to the inevitable delay restriction, a sub-optimal detector is developed. In contrast with the optimal detector, the sub-optimal detector exploits two consecutive received packets to decode one packet.

In addition, the proposed algorithms have been extended for a new development of a multiple antenna destination. We demonstrated that the optimal detector can be performed by parallel detectors; each is based on a family of BCJR algorithms. A sub-optimal detector has been also proposed to avoid the delay limitation that is associated with the optimal one.

Generally, simulation results indicated that the performance of the optimal and sub-optimal detectors is very close to that of the ideal case with perfect separation of direct and relaying links. In addition, the sub-optimal detector outperforms the optimal one in terms of the required delay, memory size, bandwidth loss, and computational complexity.

Appendix

In this appendix, we compute the number of floating point operations (flops) for the optimal and the sub-optimal detectors at the destination[8]. We take into account that multiplication and addition of two complex numbers require 6 and 2 flops, respectively, and the square of the absolute value of a complex number takes 3 flops [97]. For illustration, $|a_1 + b_1|^2 = a_1 \times a_1 + b_1 \times b_1$, where $a_1 \times a_1$ is the first flop, $b_1 \times b_2$ is the second flop, and the addition is the third flop, where a_1 and b_1 are real-valued. Following [67], we assume that the number of flops required for look-up table functions, such as e^{a_1} and $\log a_1$, is zero as the value of the input can be directly mapped to the memory address that includes the output value. As such, there are no flops required for such functions. Bearing this in mind, it is easy to compute the required flops for any algorithm.

For the optimal detector, it is easy to verify that the number of flops that are required to compute the branch metric in the log domain, in (3.11) given by

$$-\left|y_D^p(n) - d^{(p)}(n)h_{SD}^{(p)}(n) - d^{(p-1)}(n)\psi_{R_x}(p-1)h_{R_xD}^{(p)}(n)\right|^2/\sigma^2$$

is 20 flops. For each trellis unit, there are M branches entering each state. With each branch, the branch metrics are added with the corresponding previous state metrics and we choose the path that corresponds to the largest metric. It follows that we need $22M^2 - M \simeq 22M^2$ flops for each trellis unit, where M flops are required for choosing the largest metric [98]. As such, for the entire frame we need $22M^2N_dP$.

In the case of the sub-optimal detector, it turns out that the number of flops required to compute (3.19), in the log domain, is 32. Accordingly, the computation

[8]Floating-point operations include any operations that involve fractional numbers.

of (3.17) can be performed as

$$\log \Pr\left(\mathbf{y}'^{(p)}(n) \,\middle|\, \mathbf{d}'^{(p)}(n), \mathbf{h}'^{(p)}(n)\right) \Pr\left(d^{(p+1)}(n)\right) \quad \Longrightarrow$$

$$32 \text{ flops},$$

$$\sum_{d^{(p+1)}(n)\in\Omega} \log \Pr\left(\mathbf{y}'^{(p)}(n) \,\middle|\, \mathbf{d}'^{(p)}(n), \mathbf{h}'^{(p)}(n)\right) \Pr\left(d^{(p+1)}(n)\right) \quad \Longrightarrow$$

$$32M + M - 1 \text{ flops},$$

$$\arg \max_{\tilde{d}^{(p)}(n)} \sum_{d^{(p+1)}(n)\in\Omega} \log \Pr\left(\mathbf{y}'^{(p)}(n) \,\middle|\, \mathbf{d}'^{(p)}(n), \mathbf{h}'^{(p)}(n)\right) \Pr\left(d^{(p+1)}(n)\right), \quad \Longrightarrow$$

$$(32M + M - 1)\, M + M - 1 \approx 33M^2 \text{ flops}.$$

In the previous equations, we assume equally-likely transmitted, $\Pr\left(d^{(p+1)}(n)\right) = \frac{1}{M}$. As such, we drop the multiplication factor, $\Pr\left(d^{(p+1)}(n)\right)$ as this does not affect the maximization step in (3.17). By a closer look at the proposed sub-optimal detector, one can notice that the previous computation is repeated $N_d P$ times. In addition, the interference cancellation step (in (3.14) can be accomplished with $8N_d$ flops per packet. Accordingly, the proposed sub-optimal detector requires $33M^2 N_d P + 8N_d(P-1)$ flops.

Chapter 4

Data Detection for Coded ARC Systems

In this chapter, we propose the use of bit-interleaved coded modulation (BICM) in decode-and-forward (DF) alternate-relaying cooperative (ARC) communication systems. A modified BICM-ID decoder is proposed at the relay to mitigate the influence of interference resulting from simultaneous transmission of data streams through both direct and one of relay channels. Furthermore, at the destination, we exploit the interference signal, which results from the simultaneous transmission of data streams through both direct and one of the relay channels, to develop an optimal detector. The optimal detector exchanges soft information between decoders and MAP algorithms in an iterative way for performance improvement. The major drawback of the proposed optimal detector is the delay, i.e., the destination has to receive and store all received packets before performing data detection. Due to the inevitable delay restriction, a sub-optimal detector is developed. In contrast with the optimal one, the sub-optimal detector exploits two consecutive received packets to decode one packet. In addition, we develop optimal and suboptimal de-

Figure 4.1: Block diagram of the transmitter.

tectors for a destination node supporting multiple antenna. Extensive simulation results are presented to demonstrate the effectiveness of the proposed detectors.

The chapter is organized as follows. In Section 4.1, the system model and problem formulation are presented. Then, the ML detector at the relays is introduced in Section 4.2. The optimal and sub-optimal detectors at the destination equipped with single and multiple antenna are proposed in Sections 4.3 and 4.4, respectively. The performance of the proposed detectors is evaluated through computer simulations in Section 4.5. Finally, we conclude the chapter in Section 4.6.

4.1 System Model and Problem Formulation

We consider a simple cooperative communication network composed of four nodes: source node S, relay nodes R_1 and R_2, and destination node D. All the terminals are equipped with a single antenna and relays operate in half-duplex mode. The source transmission is divided into frames, each consisting of P packets. Each frame is generated as follows (see Figure 4.1). The pth packet $\mathbf{b}^{(p)} = \left[b^{(p)}(1), b^{(p)}(2), \cdots, b^{(p)}(N_b)\right]$ of N_b information bits is encoded, resulting in N_c coded bits, $\mathbf{c}^{(p)} = \left[c^{(p)}(1), c^{(p)}(2), \cdots, c^{(p)}(N_c)\right]$. The code that is used in this work is a convolutional code[1]. In order to break the sequential fading correlation and increase the diversity order, the bit interleaver permutes the coded

[1]For the sake of simplicity, we use a convolutional code. However, the proposed detectors in this chapter are valid for other coding schemes.

bits in a deterministic way. Each m consecutive bits of the interleaved sequence are grouped to form the nth vector $\mathbf{v}_n^{(p)} = \left[v_n^{(p)}(1), v_n^{(p)}(2), \cdots, v_n^{(p)}(m)\right]$, where $v_n^{(p)}(f)$ is the fth bit of the nth vector of the pth packet. The output of the interleaver can be represented by $\mathbf{v}^{(p)} = \left[\mathbf{v}_1^{(p)}, \mathbf{v}_2^{(p)}, \cdots, \mathbf{v}_{N_d}^{(p)}\right]$, where $N_d = N_c/m$. The modulator maps each $\mathbf{v}_n^{(p)}$ to a complex transmitted symbol $d^{(p)}(n) = \mu(\mathbf{v}_n^{(p)})$ chosen from an M-point signal constellation Ω, where μ is the labeling map and $M = 2^m$. Accordingly, one can describe the output of the mapper for the pth packet as $\mathbf{d}^{(p)} = \left[d^{(p)}(1),\, d^{(p)}(2),\, \cdots,\, d^{(p)}(N_d)\right]$.

The frame is transmitted continuously, one packet per time slot, over a wireless channel to the destination. The transmission schedule for the P time slots for each frame is as described in Section 3.1. As assumed in the previous chapter, the channels are assumed to be frequency non-selective and modeled as a zero-mean independent complex Gaussian random variable. In addition, we assume that all nodes have equal additive white Gaussian noise (AWGN) variance of σ^2. Finally, we assume that perfect channel state information is available at the destination and relays. For reading convenience, we rewrite the nth received symbol of the pth and $(p+1)$th packets at the relays and destination as (see Figures 3.1 and 3.2)

$$y_D^{(p)}(n) = d^{(p)}(n)h_{SD}^{(p)}(n) + d^{(p-1)}(n)\psi_{R_1}(p-1)h_{R_1D}^{(p)}(n) + w_D^{(p)}(n), \tag{4.1}$$

$$y_{R_2}^{(p)}(n) = d^{(p)}(n)h_{SR_2}^{(p)}(n) + d^{(p-1)}(n)\psi_{R_1}(p-1)h_{R_1R_2}^{(p)}(n) + w_{R_2}^{(p)}(n), \tag{4.2}$$

and

$$y_D^{(p+1)}(n) = d^{(p+1)}(n)h_{SD}^{(p+1)}(n) + d^{(p)}(n)\psi_{R_2}(p)h_{R_2D}^{(p+1)}(n) + w_D^{(p+1)}(n), \tag{4.3}$$

$$y_{R_1}^{(p+1)}(n) = d^{(p+1)}(n)h_{SR_1}^{(p+1)}(n) + d^{(p)}(n)\psi_{R_2}(p)h_{R_1R_2}^{(p+1)}(n) + w_{R_1}^{(p+1)}(n). \quad (4.4)$$

Here $h_{AB}^{(p)}(n)$ is the channel coefficient between nodes A and B, $w_A^{(p)}(n)$ is the noise contribution, $A, B \in \{S, D, R_1, R_2\}$, and $\psi_x(\diamond)$ is an indicator function which is unity when the relay x, $x \in \{R_1, R_2\}$, was able to correctly detect the $(\diamond)$th data packet and zero otherwise. Our goal is to develop BICM detectors at relays and destination.

4.2 Decoding Technique at the Relays

In DF ARC systems, it is usually assumed that successive interference cancellation, where the strongest signal is detected first, and then its contribution is subtracted from the received signal before detecting the other signal, can be employed at the relays [38, 59]. In order to provide a reliable BER performance, this requires that the inter-relay link is either sufficiently weak or sufficiently strong when compared with the source-relays links. However, these two extreme scenarios may not always occur in practical systems. In this section, at the relays, we modify the bit metric generation to have a better performance in the presence of the interference resulting from the forwarding relay transmission.

From (4.2) and (4.4), one can observe that data received at a listening relay from the source is interfered by the data sent from the forwarding relay. This is because at any time slot, there is always a relay transmitting data simultaneously with the source. Without loss of generality, we can write the received signal at relay R_x $(x \in \{1, 2\})$ as

$$y_{R_x}^{(p)}(n) = d^{(p)}(n)h_{SR_x}^{(p)}(n) + d^{(p-1)}(n)\psi_{R_y}(p-1)h_{R_xR_y}^{(p)}(n) + w_{R_x}^{(p)}(n), \quad (4.5)$$

where $y \in \{1,2\}$ and $x \neq y$. The demapper takes the received symbol $y_{R_x}^{(p)}(n)$ and the channel coefficients $\mathbf{h}^{(p)}(n) = \left[h_{SR_x}^{(p)}(n), \psi_{R_y}(p-1)h_{R_xR_y}^{(p)}(n)\right]$ as its inputs to compute the bit metric, $\lambda\left(v_n^{(p)}(f) = b\right)$, $\forall f = 1, \cdots, m$ and $b = 0, 1$, using the maximum a posteriori criterion as [53]

$$\lambda\left(v_n^{(p)}(f) = b\right) = \Pr\left(v_n^{(p)}(f), d^{(p-1)}(n) \mid y_{R_x}^{(p)}(n), \mathbf{h}^{(p)}(n)\right). \tag{4.6}$$

Since the symbol $d^{(p-1)}(n)$ is not known at the relay R_x, we remove its contribution by averaging out as

$$\begin{aligned}\lambda\left(v_n^{(p)}(f) = b\right) = &\sum_{d^{(p-1)}(n)\in\Omega} \Pr\left(d^{(p-1)}(n)\right) \\ &\Pr\left(v_n^{(p)}(f) = b \mid y_{R_x}^{(p)}(n), d^{(p-1)}(n), \mathbf{h}^{(p)}(n)\right).\end{aligned} \tag{4.7}$$

We assume that the transmitted symbols are equally probable; thus, $\Pr\left(d^{(p-1)}(n)\right) = M^{-1}$. Further, the probability $\Pr\left(v_n^{(p)}(f) = b \mid y_{R_x}^{(p)}(n), d^{(p-1)}(n), \mathbf{h}^{(p)}(n)\right)$ can be computed as

$$\begin{aligned}&\Pr\left(v_n^{(p)}(f) = b \mid y_{R_x}^{(p)}(n), d^{(p-1)}(n), \mathbf{h}^{(p)}(n)\right) = \\ &\quad\sum_{a\in\Psi(f,b)} \Pr\left(y_{R_x}^{(p)}(n) \mid a, d^{(p-1)}(n), \mathbf{h}^{(p)}(n)\right) \times \Pr(a),\end{aligned} \tag{4.8}$$

where the subset $\Psi(f,b) = \left\{\mu\left(\left[v_n^{(p)}(1), v_n^{(p)}(2), \cdots, v_n^{(p)}(m)\right] \mid v_n^{(p)}(f) = b\right)\right\}$. The a priori probability $\Pr(a)$ is unavailable on the first iteration of the demapping. Therefore, in the initialization phase it is assumed that all a are equally probable. Equation (4.8) is used as the input to the SISO decoder, which then generates the a posteriori probabilities for the coded bits. On the second iteration, these probabili-

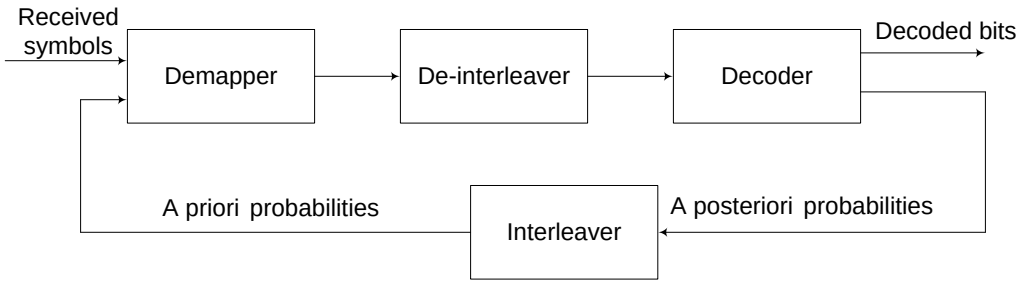

Figure 4.2: Block diagram of BICM with iterative decoding for single-input single-output systems.

ties are interleaved and fed back as a priori probabilities to the demapper as shown in Figure 4.2.

4.3 Decoding Techniques at the Destination

For single-input single-output systems, Zehavi suggested a detection method using two separate steps: bit metric generation and SISO decoding, as shown in Figure 4.2 [49]. The demapper generates $2m$ bit metrics for each received symbol, associated with the m positions, each having binary values 0 and 1. These bit metrics are then de-interleaved and fed to the SISO decoder. The soft information provided by the decoder is used to enhance the bit metrics in a recursive manner.

In this section, we show how to exploit the interference signal at the destination, to develop an optimal detector for DF alternate-relaying BICM cooperative systems.

4.3.1 Optimal Detection

The proposed optimal detector consists of three major steps, as depicted in Figure 4.3. In the first step, we apply the MAP algorithms to compute the a posteriori probability of each transmitted symbol. In the second step, we forward these

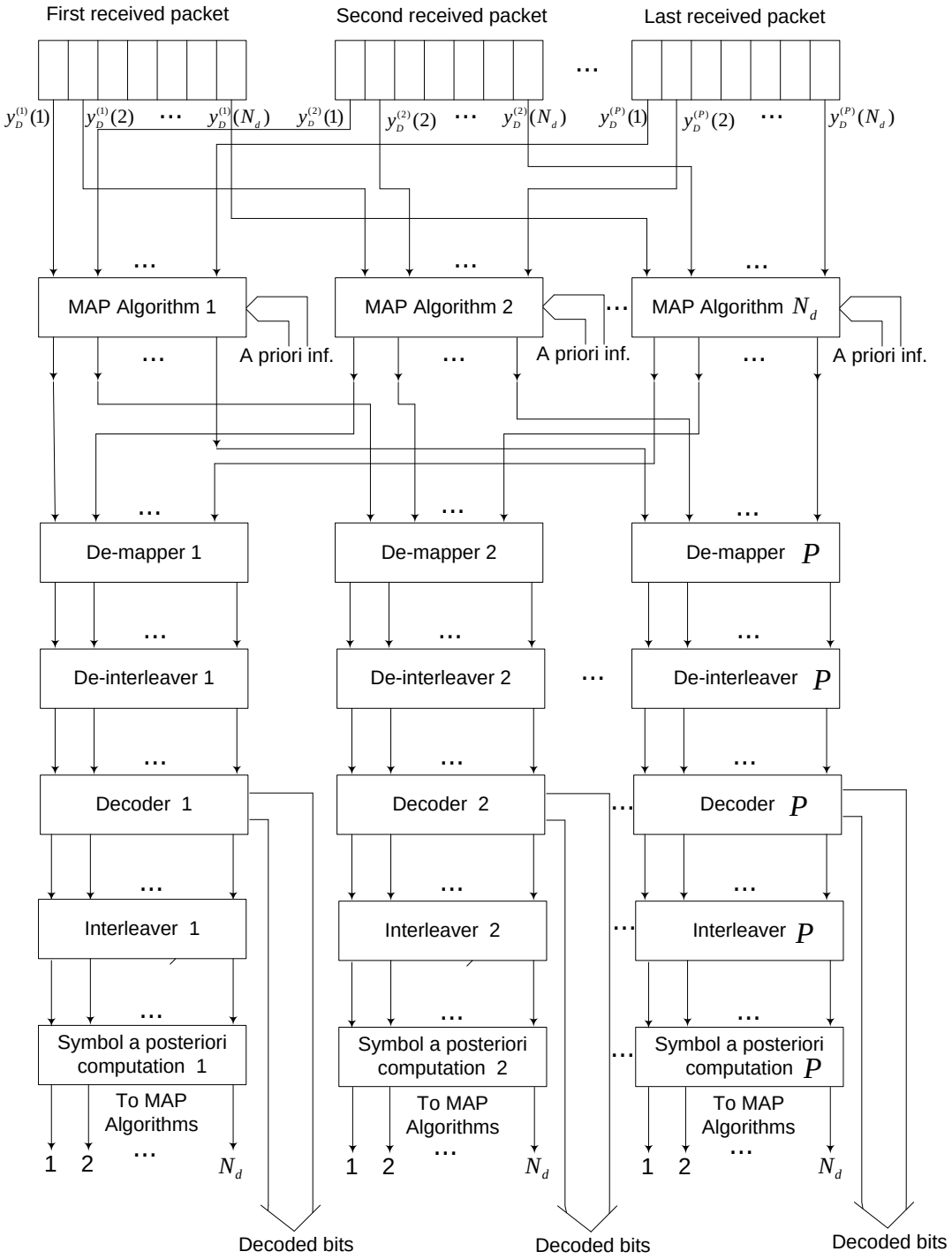

Figure 4.3: Optimal proposed detector at destination.

probabilities to demappers to compute the bit metrics. Finally, the SISO decoders receive these metrics and extract the a posteriori information of the transmitted bits. The output of the decoder is fed back to the MAP algorithms for performance improvement. As one observes, the proposed detector iterates between MAP algorithms and SISO decoders.

4.3.1.1 MAP Algorithms

The key idea of the proposed optimal detector is to employ N_d parallel MAP algorithms, as shown in Figure 4.3. The input of the nth MAP algorithm is the nth symbols from each received packet, $\mathbf{Y}(n) = \left[y_D^{(1)}(n), y_D^{(2)}(n), \cdots, y_D^{(P)}(n)\right]$, with $y_D^{(p)}(n)$ and $y_D^{(p+1)}(n)$ defined in (4.1) and (4.3) respectively[2]. The nth MAP algorithm calculates the a posteriori probabilities $\Pr\left(d^{(p)}(n) = \vartheta \mid \mathbf{Y}(n), \mathbf{H}(n)\right)$ for $p = 1, \cdots, P$, and for all θ belonging to the constellation Ω, and $\mathbf{H}(n) = \left[h_{SD}^{(1)}(n), h_{R_xD}^{(1)}(n), h_{SD}^{(2)}(n), h_{R_yD}^{(2)}(n), \cdots, h_{SD}^{(P)}(n), h_{R_xD}^{(P)}(n)\right]$. Here $\Pr(x_1|, x_2, x_3)$ is defined as the probability density function of occurring the event x_1 given the events x_2 and x_3. The MAP algorithm can be implemented based on the BCJR algorithm [76], with the transition probability[3], $\gamma_n^{(p)}(s', s)$, as

$$\gamma_n^{(p)}(s', s) = \frac{1}{\sqrt{\pi\sigma^2}} \Pr\left(d^{(p)}(n)\right) \exp\left(-\left|y_D^p(n) - d^{(p)}(n) h_{SD}^{(p)}(n) - d^{(p-1)}(n)\psi_{R_x}(p-1) h_{R_xD}^{(p)}(n)\right|^2 / \sigma^2\right), \quad (4.9)$$

[2]From (3.1) and (3.3), one can observe that the symbol $d^{(p)}(n)$ is included in $y_D^{(p)}(n)$ and $y_D^{(p+1)}(n)$. In addition, the symbol $d^{(p+1)}(n)$ is included in $y_D^{(p+1)}(n)$ and $y_D^{(p+2)}(n)$. Accordingly, in order to provide an optimal detector for nth symbol in each packet we have to rely on the received sequence $\left[y^{(1)}(n), y^{(2)}(n), \cdots, y^{(P)}(n)\right]$. That is why we mix received packets, as shown in Figure 4.3.

[3]$\gamma_n^{(p)}(s', s)$ represents the probability to transit from the state s' to the state s for the nth symbol in the pth packet given the received symbol $y_D^{(p)}(n)$.

where $\left[d^{(p)}(n)\; d_{R_x}^{(p-1)}(n)\right]$ represents the output associated with this transition. The a priori probability $\Pr\left(d^{(p)}(n)\right)$ is unavailable at the first iteration. Therefore, in the initialization phase it is assumed that all $d^{(p)}(n)$ are equally probable. Equation (4.9) is used as the input to the demapper, which then generates the bit metrics.

4.3.1.2 Demapper

The a posteriori probabilities provided by N_d MAP algorithms can be exploited to compute the bit metrics. Referring to Figure 4.3, after the MAP algorithms, the decoding process can be divided into P parallel branches, where the branch pth decodes the pth packet. Each branch contains a demapper, de-interleaver, decoder, interleaver, and symbol a posteriori computation unit. In order to forward each branch its corresponding a posteriori probabilities, the outputs of MAP algorithms are de-permuted; the pth output of each MAP algorithm is forwarded to the pth branch. Mathematically, the a posteriori probabilities $\left\{\Pr\left(d^{p}(n)\mid y_D^{(p)}(n),\mathbf{H}(n)\right)\right\}_{n=1}^{N_d}$ are passed to the pth branch. Accordingly, the bit metric, $\lambda\left(v_n^{(p)}(f)=b\right)$, can be computed as[4]

$$\lambda\left(v_n^{(p)}(f)=b\right)=\sum_{d^{(p)}(n)\in\Psi(f,b)}\Pr\left(d^{(p)}(n)\mid y_D^{(p)}(n),\mathbf{H}(n)\right), \tag{4.10}$$

where the subset $\Psi(f,b)=\left\{\mu\left(\left[v_n^{(p)}(1),v_n^{(p)}(2),\cdots,v_n^{(p)}(m)\right]\mid v_n^{(p)}(f)=b\right)\right\}$.

4.3.1.3 Decoder

After de-interleaving, the bit metrics is passed to the decoder to provide the a posteriori probabilities of coded bits. These probabilities are then interleaved and forwarded to the symbol a posteriori computation unit. Assuming that the probabili-

[4]For more details on the bit metric concept, the reader is referred to [49, 51–53, 81, 85].

ties $\Pr\left(v_n^{(p)}(1)\right), \Pr\left(v_n^{(p)}(2)\right), \cdots, \Pr\left(v_n^{(p)}(m)\right)$ are independent by using a good interleaver, the symbol a posteriori probability $\Pr\left(d^{(p)}(n)\right)$ can be computed as

$$\Pr\left(d^{(p)}(n)\right) = \prod_{f=1}^{m} \Pr\left(v_n^{(p)}(f)\right). \tag{4.11}$$

These a posteriori symbol probabilities are provided to the MAP algorithms as a priori information, as seen in (4.9). At the last iteration, the final decoded outputs are the hard decisions based on the a posteriori probabilities.

Implementation Aspects

- Since the MAP algorithm is a trace back algorithm, the end state must be known. An additional N_d zero symbols are added to the end of the transmitted frame to drive the encoder to state zero. In the literature, these additional zeros are called the tail [49, 51–53, 81, 85]. This will result in a sightly bandwidth loss compared with the direct transmission. This loss is given by $1/P$, which approaches zero for a large number of packets per frame.

- We follow the common assumption used in the cooperative literature that the relays forward the received packets only when they have been correctly decoded; otherwise, relays remain idle. As such, the destination has to receive an acknowledgment from relays about the status of each packet. For illustration, if the destination receives negative acknowledgment for the pth packet, this implies that the relay R_x ($x = 1$ or 2) is unable to decode and forward this packet.

- Unfortunately, the proposed optimal detector requires a large memory and delay. The destination has to receive and store the entire frame before starting data detection. These limitations are an increasing function on the modula-

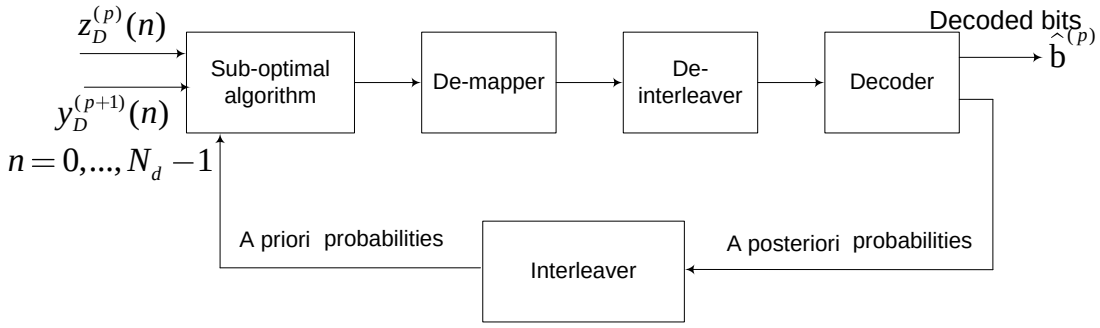

Figure 4.4: Sub-optimal detector at the destination.

tion order and number of packets per frame. Accordingly, these constraints may prevent it from practical implementation. This motivates us to propose a sub-optimal detector.

4.3.2 Sub-Optimal Detection

Instead of computing the exact values of symbol a posteriori probabilities $\Pr\left(d^{(p)}(n) = \vartheta \mid \mathbf{Y}(n), \mathbf{H}(n)\right)$ provided by the MAP algorithms, in the proposed sub-optimal detector, these probabilities are calculated in an approximated way. We employ P sub-optimal algorithms in the sub-optimal detector instead of the P MAP algorithms in the optimal detector, as shown in Figure 4.4. Note that these P sub-optimal algorithms are applied in a serial fashion, while in the optimal one the P MAP algorithms are applied in a parallel fashion. The key principle of the sub-optimal detector is to employ two consecutive received packets, $\left\{y_D^{(p)}(n)\right\}_{n=1}^{N_d}$ and $\left\{y_D^{(p+1)}(n)\right\}_{n=1}^{N_d}$, given in (4.1) and (4.3) respectively, to detect the packet $\left\{d^{(p)}(n)\right\}_{n=1}^{N_d}$. More specifically, let us remove the contribution of $\left\{d^{(p-1)}(n)\right\}_{n=1}^{N_d}$ from $\left\{y_D^{(p)}(n)\right\}_{n=1}^{N_d}$ forming

$$z_D^{(p)}(n) = y_D^{(p)}(n) - d^{(p-1)}(n)\psi_{R_x}(p-1)h_{R_xD}^{(p)}(n). \tag{4.12}$$

Here we follow the common assumption associated with the most of successive detection algorithms found in the literature (e.g., [91, 92]) that the previous symbols $\left\{d^{(p-1)}(n)\right\}_{n=1}^{N_d}$ are perfectly detected, when we perform detection for $\left\{d^{(p)}(n)\right\}_{n=1}^{N_d}$. This assumption is necessary for the theoretical development of the sub-optimal detector; however, in simulations, we relax this assumption. Using (4.1), we can write

$$z_D^{(p)}(n) = d^{(p)}(n)h_{SD}^{(p)}(n) + w_D^{(p)}(n). \tag{4.13}$$

With (3.3) rewritten as

$$y_D^{(p+1)}(n) = d^{(p+1)}(n)h_{SD}^{(p+1)}(n) + d^{(p)}(n)\psi_{R_y}(p)h_{R_yD}^{(p+1)}(n) + w_D^{(p+1)}(n), \tag{4.14}$$

and based on (4.13) and (4.14), one can write

$$\begin{aligned} &\mathrm{P}\left(d^{(p)}(n)\left|\mathbf{y}'^{(p)}(n), \mathbf{h}'^{(p)}(n)\right.\right) = \sum_{d^{(p+1)}(n)\in\Omega} \mathrm{P}\left(\mathbf{d}'^{(p)}(n)\left|\mathbf{y}'^{(p)}(n), \mathbf{h}'^{(p)}(n)\right.\right) \\ &= \sum_{d^{(p+1)}(n)\in\Omega} \mathrm{P}\left(\mathbf{y}'^{(p)}(n)\left|\mathbf{d}'^{(p)}(n), \mathbf{h}'^{(p)}(n)\right.\right) \mathrm{P}\left(d^{(p+1)}(n)\right) \mathrm{P}\left(d^{(p)}(n)\right). \end{aligned} \tag{4.15}$$

Here $\mathbf{y}'^{(p)}(n) = \left[z_D^{(p)}(n)\, y_D^{(p+1)}(n)\right]^T$, $\mathbf{d}'^{(p)}(n) = \left[d^{(p)}(n)\, d^{(p+1)}(n)\right]^T$,

$$\mathbf{h}'^{(p)}(n) = \begin{bmatrix} h_{SD}^{(p)}(n) & 0 \\ h_{R_yD}^{(p+1)}(n)\psi_{R_y}(p) & h_{SD}^{(p+1)}(n) \end{bmatrix}, \tag{4.16}$$

Table 4.1: A comparison of the optimal and sub-optimal detectors. The abbreviation MS stands for the required memory size and BW loss for the bandwidth loss. Delay is expressed by the number of data symbols required to store.

	Complexity	Delay	MS	BW loss
Optimal	$65M^2N_dP + \Upsilon$	N_dP	N_dP	$1/P$
Sub-optimal	$\left(34M^2 - M\right)N_dP + 8N_d(P-1) + \Upsilon$	N_d	N_d	zero

and

$$\mathrm{P}\left(\mathbf{y}'^{(p)}(n)\,\middle|\,\mathbf{d}'^{(p)}(n), \mathbf{h}'^{(p)}(n)\right) = \tfrac{1}{\pi\sigma^2}\exp\left(-\left\|\mathbf{y}'^{(p)}(n) - \mathbf{h}'^{(p)}(n)\mathbf{d}'^{(p)}(n)\right\|^2/\sigma^2\right). \tag{4.17}$$

After computing the a posteriori probabilities of the data symbols as in (4.15), (4.10) can be applied to produce the bits metrics. As for the optimal detector, these bit metrics are then de-interleaved and passed to the SISO detector. The soft information provided by the decoder is fed back to (4.15) to refine the computation.

After detecting $\left\{d^{(p)}(n)\right\}_{n=1}^{N_d}$, their contribution can be removed from $\left\{y_D^{(p+1)}(n)\right\}_{n=1}^{N_d}$, forming $z_D^{(p)}(n+1) = y_D^{(p+1)}(n) - d^{(p)}(n)\psi_{R_y}(p)h_{R_yD}^{(p+1)}(n)$. Similarly, from $\left\{z_D^{(p+1)}(n)\right\}_{n=1}^{N_d}$ and $\left\{y_D^{(p+2)}(n)\right\}_{n=1}^{N_d}$ and by averaging over $\left\{d^{(p+2)}(n)\right\}_{n=1}^{N_d}$, the $(p+1)$th packet, $\left\{d^{(p+1)}(n)\right\}_{n=1}^{N_d}$, can be detected, and so on.

A comparison of the proposed optimal and sub-optimal detectors in terms of complexity, delay, MS, and BW loss is shown in Table 4.1. We evaluate the computational complexity of the proposed detectors in terms of the required flops per iteration. In the table, $\Upsilon = P\left(MN_dm + (m-1)N_d + \varpi\right)$ refers to the term that is common in both detectors, where ϖ is the number of flops that requires for one decoder; ϖ depends on encoder parameters such as constraint length, code rate and generator polynomial. One can notice that the complexity of both detectors is directly proportional to the number of symbols per packet N_d and square of the

modulation order, M^2. We also notice that the sub-optimal detector outperforms the optimal detector in terms of the required delay, MS, and BW loss, with a lower computational complexity.

4.4 Data Detection for Multiple Antenna Destination

In this section, we assume that the destination is a base station that has the capability to support $A_D \geqslant 1$ antennas, while the source, and relays are small nodes with limited power and few resources, thus, each of them is equipped only with one antenna. Our aim is to perform data detection at the destination.

4.4.1 Proposed Optimal Receiver Structure

The proposed optimal receiver mainly consists of N_d parallel demappers which provide the bit metrics of the transmitted frame to P parallel detectors, as shown in Figure 4.5. The receiver iterates between the demappers and the detectors for performance improvement.

Demappers

From (3.1) and (3.3), one can notice that each transmitted symbol $d^{(p)}(n)$ is included in $y_D^{(f,p)}(n)$ and $y_D^{(f,p+1)}(n)$, $f = 1, 2, \cdots, A_D$. Accordingly, in order to exploit the repetition of the transmitted symbols, we provide the nth demapper with the nth symbols of each received packet of each receive antenna, $\mathbf{y}_D^{(f)}(n) = \left[y_D^{(f,1)}(n), y_D^{(f,2)}(n), \cdots, y_D^{(f,P)}(n)\right]$, $f = 1, 2, \cdots, A_D$, and the a priori probabilities of the nth symbol of each packet in the frame. Each demapper employs serial BCJR algorithms; each is associated with one unique receive antenna, as shown in Figure 4.6. Each algorithm delivers a posteriori probabilities (APPs) about the

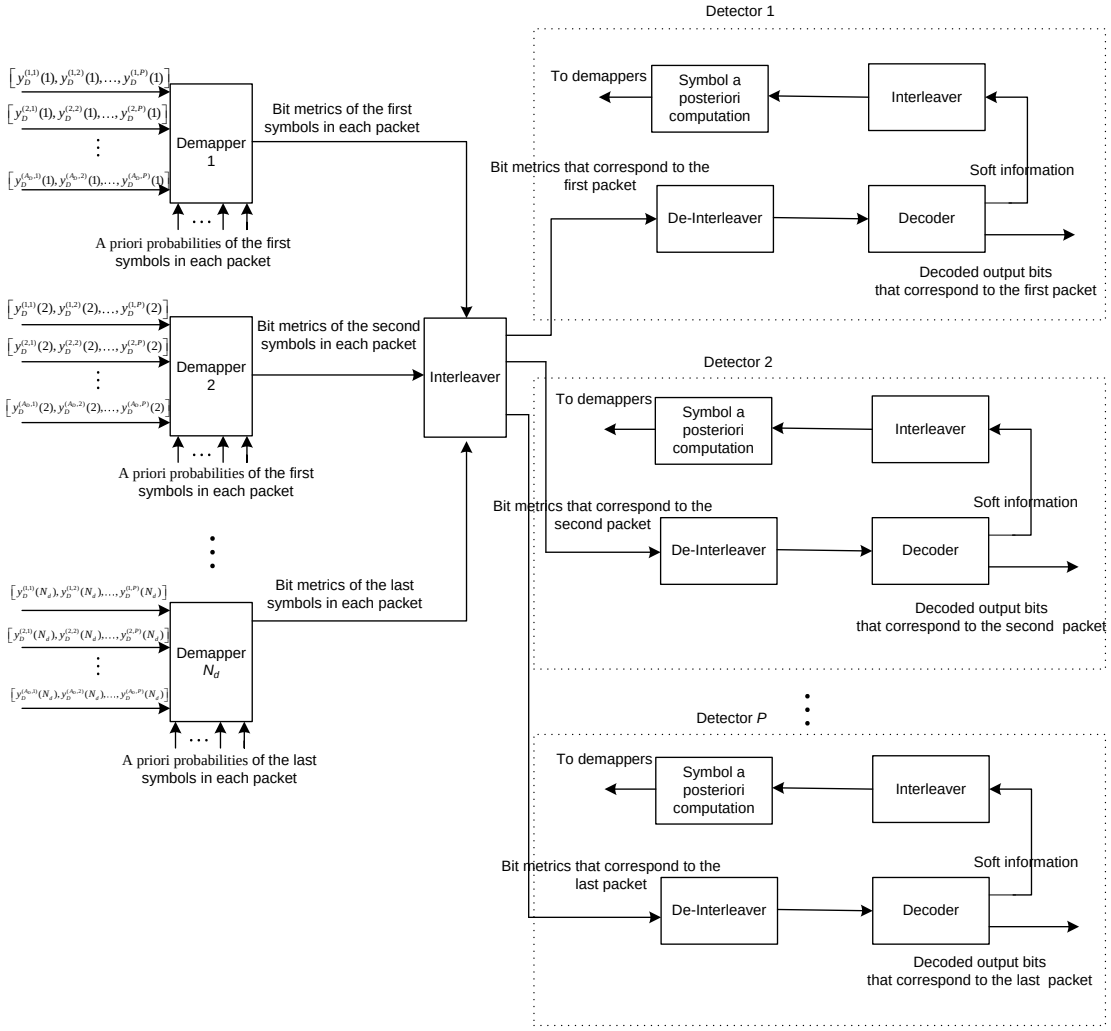

Figure 4.5: Optimal receiver for multiple receive antennas DF ARC cooperative system.

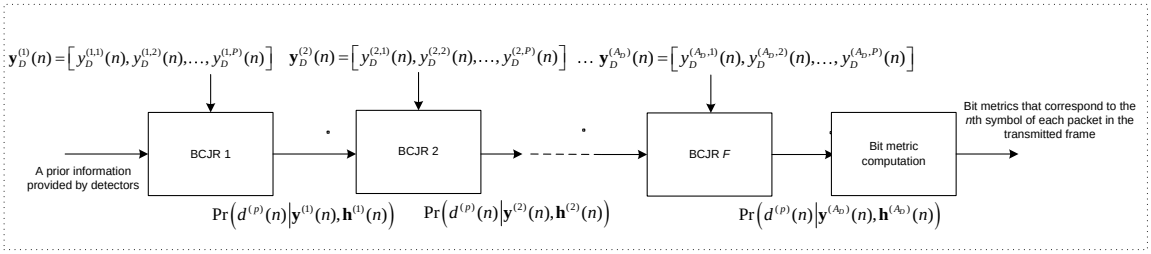

Figure 4.6: The structure of the nth demapper.

transmitted data which is used as a priori information for the next algorithm. For the first iteration, we assume no a prior information available for the first BCJR algorithm. While in the following iterations, the APPs provided by the detectors are employed as a priori information for the first BCJR algorithm. Upon completion, the APPs provided by the last BCJR algorithm are used to compute the bit metrics as follows.

The fth BCJR algorithm computes the APPs of the nth symbol of each received packet on the fth antenna, $\Pr\left(d^{(p)}(n)\left|\mathbf{y}^{(f)}(n),\mathbf{h}^{(f)}(n)\right.\right)$, where $\mathbf{h}^{(f)}(n) = \left[h_{SD}^{(f,1)}(n), h_{R_xD}^{(f,1)}(n), h_{SD}^{(f,2)}(n), h_{R_yD}^{(f,2)}(n), \cdots, h_{SD}^{(f,P)}(n), h_{R_xD}^{(f,P)}(n)\right]$ represents the channel coefficients and $p = 1, \cdots, P$. Using (3.1) and following the principles of the BCJR algorithm which is explained in Section 2.2.4, one can write the branch metric between the states s' and s for the nth symbol in the pth packet given the received symbol $y_D^{(f,p)}(n)$, is defined as

$$\begin{aligned}\gamma(s',s) = &\tfrac{1}{\sqrt{\pi\sigma^2}}\Pr\left(d^{(p)}(n)\right)\times\\ &\exp\left(-\left|y_D^{(f,p)}(n) - d^{(p)}(n)h_{SD}^{(f,p)}(n) - d_{R_x}^{(p-1)}(n)\psi_{R_x}(p-1)h_{R_xD}^{(f,p)}(n)\right|/\sigma^2\right)\end{aligned}, \tag{4.18}$$

where $\left[d^{(p)}(n)\, d_{R_x}^{(p-1)}(n)\psi_{R_x}(p-1)\right]$ represents the output associated with this branch and $\Pr\left(d^{(p)}(n)\right)$ is the a priori probability. The APPs, $\Pr\left(d^{(p)}(n)\left|\mathbf{y}^{(A_D)}(n),\mathbf{h}^{(A_D)}(n)\right.\right)$, provided by the last BCJR algorithm, are used to compute the bit metrics $\lambda\left(v_n^{(p)}(k)=b\right)$ as

$$\lambda\left(v_n^{(p)}(k)=b\right) = \sum_{d^{(p)}(n)\in\Psi(k,b)} \mathrm{P}\left(d^{(p)}(n) \mid \mathbf{y}_D^{(A_D)}(n), \mathbf{h}^{(A_D)}(n)\right), \tag{4.19}$$

where the subset $\Psi(k,b) = \left\{\mu\left(\left[v_n^{(p)}(1), v_n^{(p)}(2), \cdots, v_n^{(p)}(m)\right] \mid v_n^{(p)}(k)=b\right)\right\}$.

Detectors

The bit metrics is passed to the decoder to provide the a posteriori probabilities of coded bits. These probabilities are then interleaved and forwarded to the symbol a posteriori computation unit. Assuming that the probabilities $\mathrm{P}\left(v_n^{(p)}(1)\right)$, $\mathrm{P}\left(v_n^{(p)}(2)\right), \cdots, \mathrm{P}\left(v_n^{(p)}(m)\right)$ are independent by using a good interleaver, the symbol a posteriori probability $\mathrm{P}\left(d^{(p)}(n)\right)$ can be computed as

$$\mathrm{P}\left(d^{(p)}(n)\right) = \prod_{u=1}^{m} \mathrm{P}\left(v_n^{(p)}(u)\right). \tag{4.20}$$

These a posteriori symbol probabilities are provided to the demappers as a priori information, as seen in (4.18). At the last iteration, the final decoded outputs are the hard decisions based on the a posteriori probabilities.

Implementation Aspects

- Since the BCJR algorithm is a trace back algorithm, the end state must be known. An additional N_d zero symbols are added to the end of the transmitted frame to drive the encoder to state zero. In the literature, these additional zeros are called the tail [49, 51–53, 81, 85]. This will result in a sightly bandwidth loss compared with the direct transmission. This loss is given by $1/P$, which approaches zero for a large number of packets per frame.

- The proposed optimal receiver requires a large memory and delay. The destination has to receive and store the entire frame before starting data detection. These limitations are an increasing function on the modulation order and number of packets per frame. Accordingly, these constraints may prevent it from practical implementation. This motivates us to propose a sub-optimal detector.

4.5 Sub-Optimal Receiver

In this section, we employ the successive interference cancellation principle to avoid the delay problem associated with the optimal receiver, while maintaining the diversity gain. The basic idea behind the sub-optimal receiver is to use A_D sub-optimal algorithms to compute the bit metrics of the the transmitted frame, as shown in Figure 4.7; each algorithm is associated with one unique receive antenna. The sub-optimal receiver exploits two consecutive received packets on each receive antenna, $\left\{y_D^{(f,p)}(n)\right\}_{n=1}^{N_d}$ and $\left\{y_D^{(f,p+1)}(n)\right\}_{n=1}^{N_d}$, $f = 1, \cdots, A_D$, given in (3.1) and (3.3), respectively to detect the packet $\left\{d^{(p)}(n)\right\}_{n=1}^{N_d}$. The fth algorithm operates as follows. We remove the contribution of $\left\{\hat{d}^{(p-1)}(n)\right\}_{n=1}^{N_d}$ from $\left\{y_D^{(f,p)}(n)\right\}_{n=1}^{N_d}$ and form

$$z_D^{(f,p)}(n) = y_D^{(f,p)}(n) - d^{(p-1)}(n)\psi_{R_x}(p-1)h_{R_xD}^{(f,p)}(n). \tag{4.21}$$

Using (3.1), we can write

$$z_D^{(f,p)}(n) = d^{(p)}(n)h_{SD}^{(f,p)}(n) + w_D^{(f,p)}(n). \tag{4.22}$$

Here we follow the common assumption associated with most of the successive detection algorithms found in the literature (e.g., [91,92]) that the previous symbol $\hat{d}^{(p-1)}(n)$ is perfectly detected, when we perform detection for $d^{(p)}(n)$. This assumption is necessary for the theoretical development of the sub-optimal receiver; however, in the simulations we relax this assumption. With (3.3) rewritten as

$$y_D^{(f,p+1)}(n) = d^{(p+1)}(n)h_{SD}^{(f,p+1)}(n) + d^{(p)}(n)\psi_{R_y}(p)h_{R_yD}^{(f,p+1)}(n) + w_D^{(f,p+1)}(n). \tag{4.23}$$

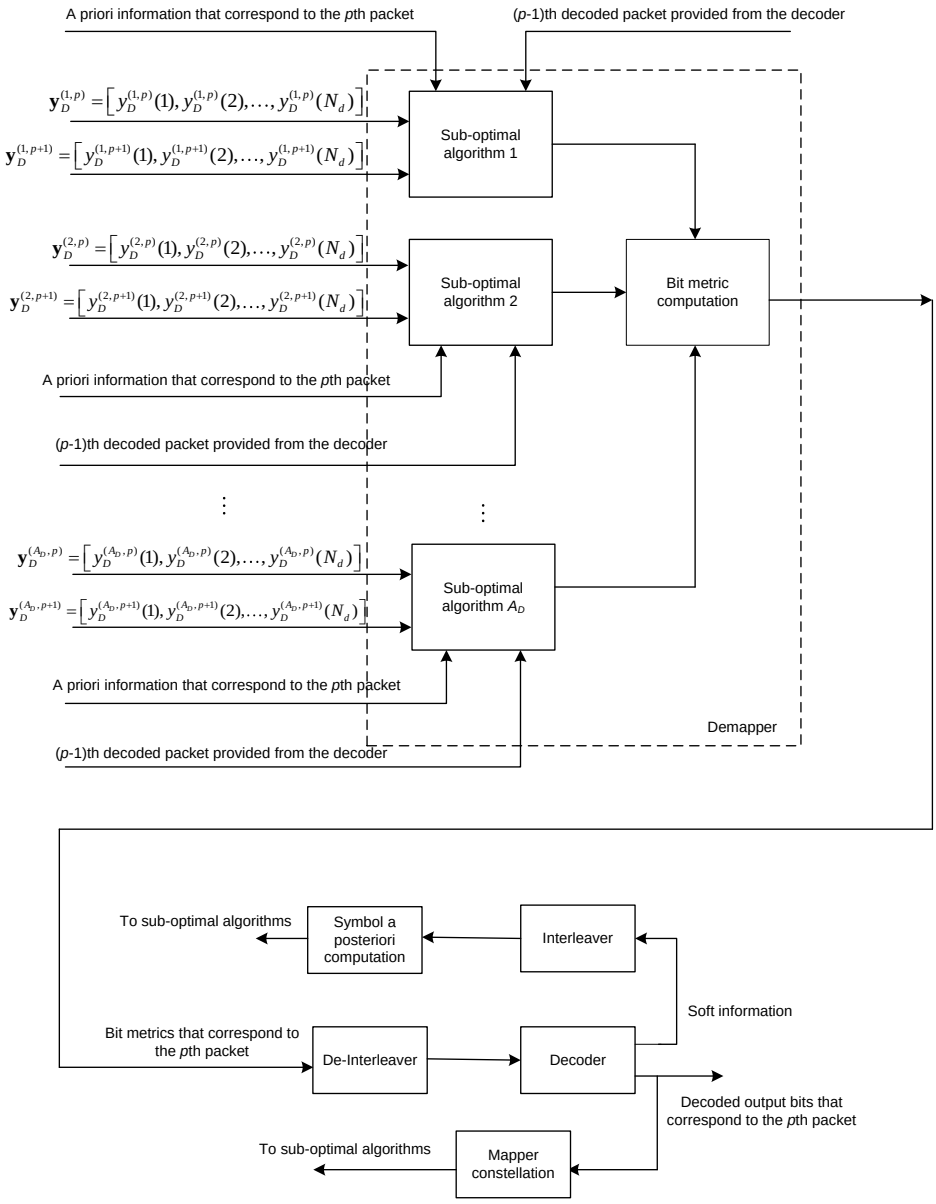

Figure 4.7: Sub-optimal receiver for DF ARC systems.

Therefore, we can write

$$\begin{aligned}\Pr\left(d^{(p)}(n)\left|z_D^{(f,p)}(n), y_D^{(f,p+1)}(n), \mathbf{h}'^{(f,p)}(n)\right.\right) = \\ \sum_{d^{(p+1)}(n)\in\Omega} \Pr\left(z_D^{(f,p)}(n), y_D^{(f,p+1)}(n)\left|d^{(p)}(n), d^{(p+1)}(n), \mathbf{h}'^{(f,p)}(n)\right.\right) \\ \times \Pr\left(d^{(p)}(n)\right)\Pr\left(d^{(p+1)}(n)\right),\end{aligned} \tag{4.24}$$

where

$$\mathbf{h}'^{(f,p)}(n) = \begin{bmatrix} h_{SD}^{(f,p)}(n) & 0 \\ h_{R_yD}^{(f,p+1)}(n)\psi_{R_y}(p) & h_{SD}^{(f,p+1)}(n) \end{bmatrix}, \tag{4.25}$$

and

$$\begin{aligned}&\Pr\left(z_D^{(f,p)}(n), y_D^{(f,p+1)}(n)\left|d^{(p)}(n), d^{(p+1)}(n), \mathbf{h}'^{(f,p)}(n)\right.\right) = \\ &\tfrac{1}{\pi\sigma^2}\exp\left(-\left\|\left[z_D^{(f,p)}(n), y_D^{(f,p+1)}(n)\right]^T - \mathbf{h}'^{(f,p)}(n)\left[d^{(p)}(n), d^{(p+1)}(n)\right]^T\right\|^2/\sigma^2\right),\end{aligned} \tag{4.26}$$

where the superscript T refers to the vector transpose. After computing the a posteriori probabilities of the data symbols as in (4.24), (4.19) can be applied to produce the bits metrics. As for the optimal detector, these bit metrics are then deinterleaved and passed to the SISO detector. The soft information provided by the decoder is fed back to (4.24) to refine the computation.

After detecting $\left\{d^{(p)}(n)\right\}_{n=1}^{N_d}$, its contribution can be removed from $\left\{y_D^{(f,p+1)}(n)\right\}_{n=1}^{N_d}$ forming $z_D^{(f,p)}(n+1) = y_D^{(f,p+1)}(n) - d^{(p)}(n)\psi_{R_x}(p)h_{R_xD}^{(f,p+1)}(n)$. Similarly, from $\left\{z_D^{(f,p+1)}(n)\right\}_{n=1}^{N_d}$ and $\left\{y_D^{(f,p+2)}(n)\right\}_{n=1}^{N_d}$ and by averaging over $\left\{d^{(p+2)}(n)\right\}_{n=1}^{N_d}$, we can detect $\left\{d^{(p+1)}(n)\right\}_{n=1}^{N_d}$, and so on.

Table 4.2: A comparison of the optimal and sub-optimal detectors.

	Complexity	Delay	MS	BW loss
Optimal	$70M^2NPA_D + \Upsilon$	NP	NPA_D	$1/P$
Sub-optimal	$\left(33M^2NP + 8N(P-1)\right)A_D + \Upsilon$	N	NA_D	zero

A comparison between the proposed optimal and sub-optimal detectors in terms of complexity, delay, MS, and BW loss is provided in Table 4.2. The computational complexity of the proposed detectors is evaluated in terms of the number of required flops per iteration. In the table, $\Upsilon = P\left(MN_d m + (m-1)N_d + \varpi\right)$ refers to the term that is common in both detectors, where ϖ is the number of flops that requires for one decoder; ϖ depends on encoder parameters such as constraint length, code rate and generator polynomial. One can notice that the complexity of both detectors is proportional to the number of symbols per packet N_d, number of packets per frame P and square of the modulation order, M^2. We also notice that the sub-optimal detector outperforms the ML detectors in terms of the required delay, MS, and BW loss at the expense of a slightly computational complexity.

4.6 Simulation Results

In this section, we validate the proposed detectors through Monte Carlo computer simulations. We consider an alternate-relaying DF cooperative communication system, using a convolutional code with constraint length 5, rate 1/2 and polynomial generators $(23)_8$ and $(35)_8$. The BCJR algorithm [76] is used for decoding. A frame based transmission is assumed; each has 20 packets. A packet length of N_b = 150 information bits is chosen, leading to N_c = 300 coded bits. The coded bits are set partition-mapped on a 8-PSK constellation, resulting in N_d= 100 symbols. The channel coefficients are modeled as zero-mean complex Gaussian random variables with variances, $\sigma^2_{SD} = 1$, $\sigma^2_{SR_1} = \sigma^2_{SR_2} = 6.25$, and $\sigma^2_{R_1D} = \sigma^2_{R_2D} = 2.8$.

Figure 4.8 shows the BER performance of the proposed detector at the relays as a function of E_b/N_0. For the sake of comparison, we show the performance when the interference signal can be treated as additional noise, and this is referred

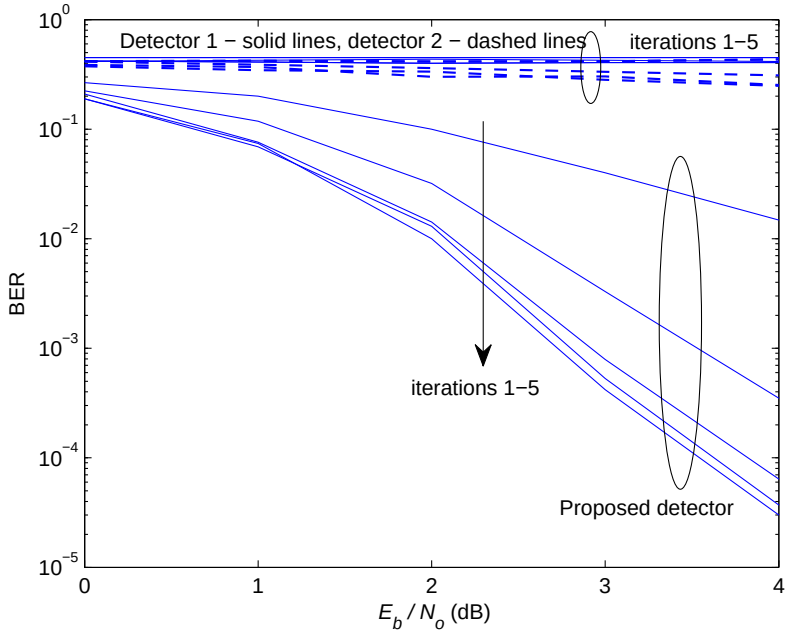

Figure 4.8: BER performance of the proposed detector, detector 1, and detector 2 as a function of E_b/N_0, at the relays.

to as detector 1. In addition, we show the BER performance of the successive interference cancellation detector proposed in [38, 59], which is based on detecting the strongest signal first, then subtract its contributions from the received interfered signal before detecting the other signal. In the sequel, this is referred to as detector 2. As one can observe, both detector 1 and detector 2 lead to an unacceptable performance. We also notice that the proposed detector achieves a strong improvement of the performance after only two iterations. In addition, there is no significant improvement in the performance after three iterations for any E_b/N_0 values.

Figures 4.9 and 4.10 respectively depict the BER performance of the proposed optimal and sub-optimal detectors at the destination with single antenna as a function of E_b/N_0, where E_b and N_0 are the energy per bit and noise power spectral density, respectively. Since two relays are used in the alternate-relaying coopera-

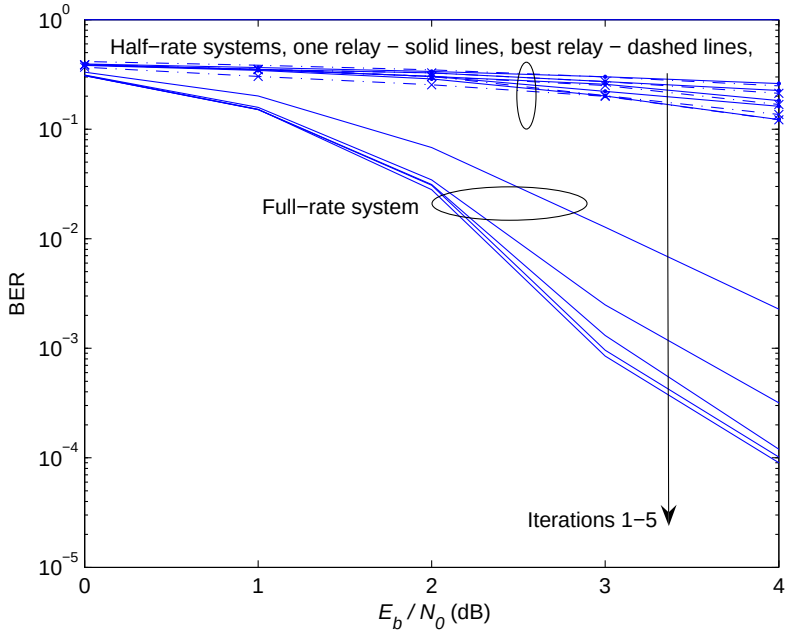

Figure 4.9: BER performance of the DF ARC cooperative system with the optimal detector and half-rate cooperative systems as a function of E_b/N_0, at the destination.

tive systems, the BER performance of the half-rate one relay and best relay from a set of two relays are also shown. For a fair comparison between the half-rate and full-rate (alternate-relaying) systems, we keep the same data rate and transmitted power for both systems. Hence, we use 64-PSK modulation for the half-rate systems and 8-PSK for the full-rate systems. As one can observe, for the full-rate cooperative systems, iterative processing achieves a significant performance improvement for the proposed optimal and sub-optimal detectors. Furthermore, it is seen that the performance of the proposed full-rate cooperative systems outperforms that of the half-rate cooperative systems.

Figure 4.12 shows the BER performance of the proposed optimal and suboptimal detectors with different relays' positions, while the distance between the two relays is kept constant. As one can observe, the BER at the destination degrades

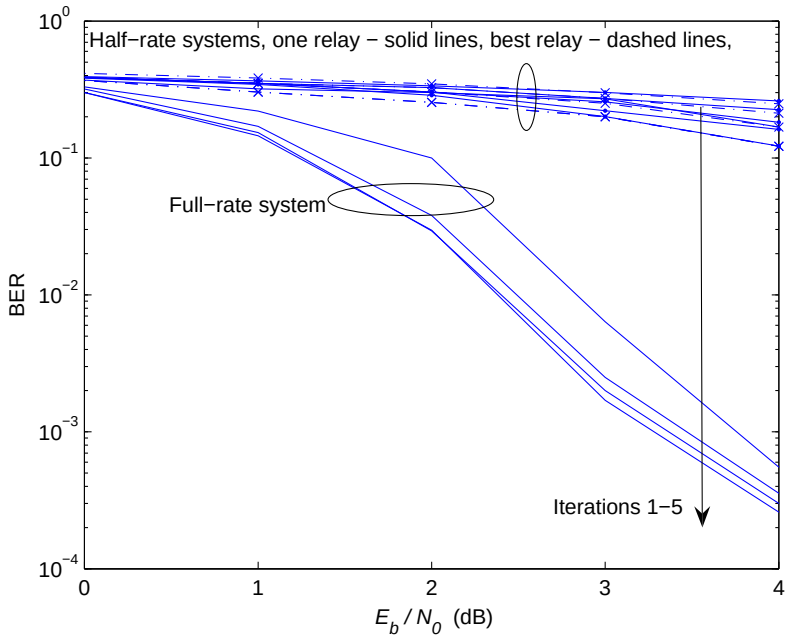

Figure 4.10: BER performance of the DF ARC system with the sub-optimal detector and half-rate cooperative systems as a function of E_b/N_0, at the destination. Note that the first iteration of the alternate-relaying system locates in the region of the half-rate system.

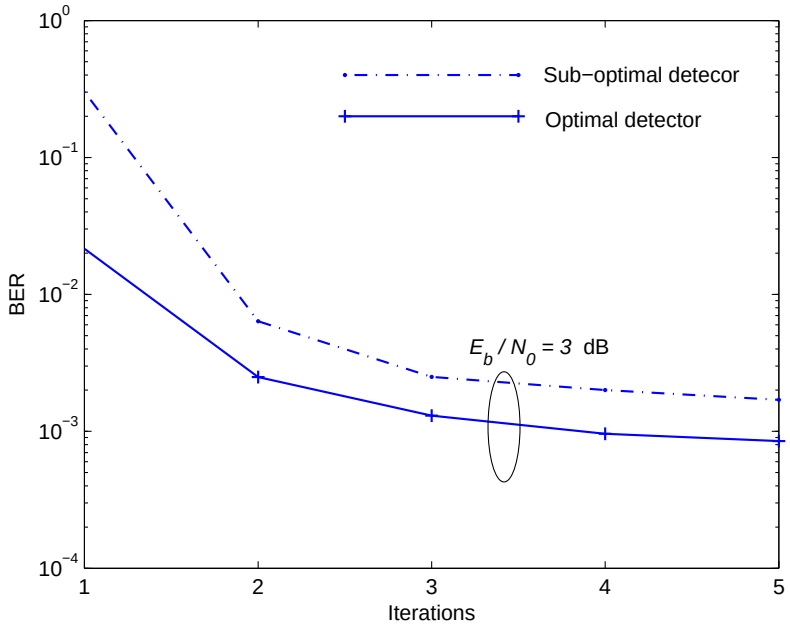

Figure 4.11: BER performance of the DF ARC system with the optimal and sub-optimal detectors as a function of the number of iterations, at the destination.

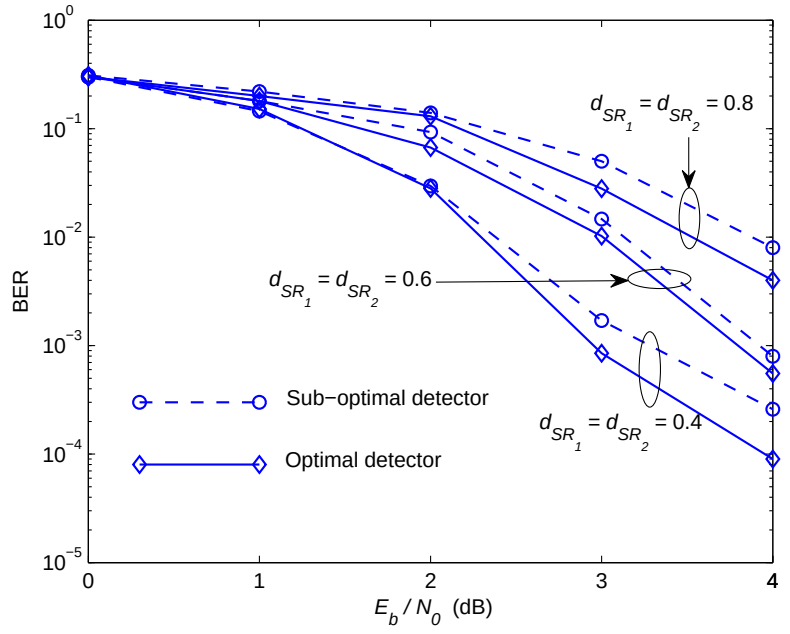

Figure 4.12: BER performance at the destination with different relays' positions. Here we set $d_{R_1R_2}$ to 0.2 and d_{R_1D} and d_{R_2D} are changed accordingly.

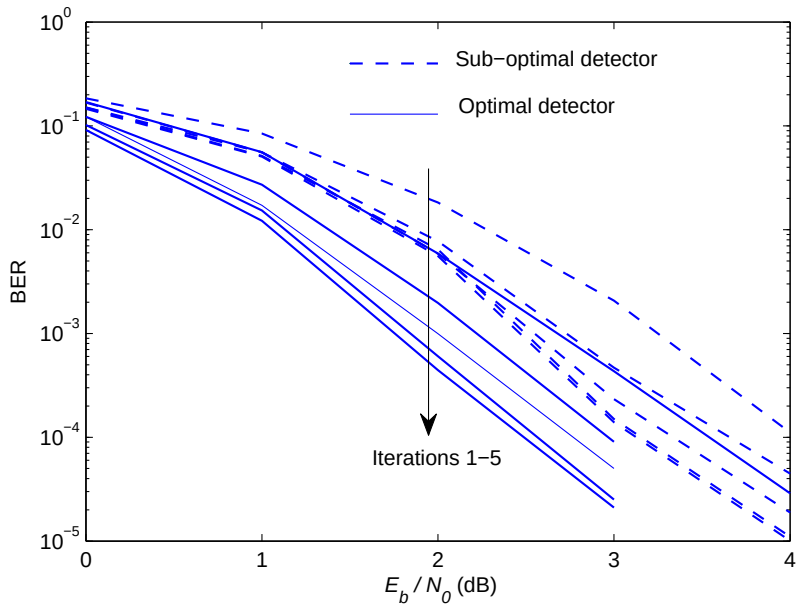

Figure 4.13: BER performance of the optimal and sub-optimal detectors for a destination supporting two antennas.

with d_{SR_1} and d_{SR_2}. It is because the BER at the relays increases with d_{SR_1} and d_{SR_2}; this, in turn, increases the BER at the destination.

Figures 4.11, 4.13, and 4.14 compare the BER performance of the proposed optimal detector with that of the proposed sub-optimal detector for a destination supporting one, two, and three antennas, respectively. As one can observe, at the first iteration the performance of the sub-optimal detector is much worse than that of the optimal one. However, by increasing the number of iterations, the performance difference between the sub-optimal and optimal detectors reduces. This occurs because at the first iteration the a posteriori probabilities provided by the sub-optimal algorithm are not reliable to initialize the SISO decoder. With the aid of iterative processing, the reliability of the sub-optimal detector increases, and its performance approaches the performance of the optimal detector.

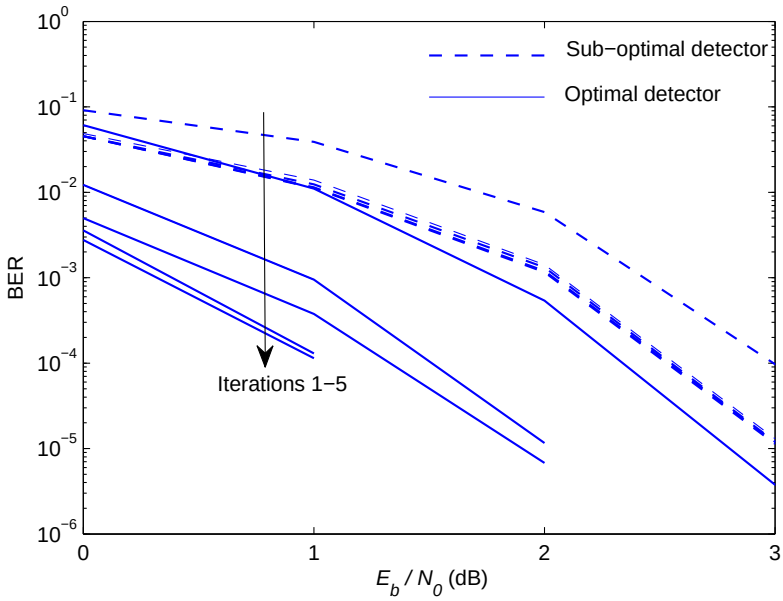

Figure 4.14: BER performance of the optimal and sub-optimal detectors for a destination supporting three antennas.

4.7 Summary

In this chapter, we propose the use of bit-interleaved coded modulation (BICM) in DF-ARC systems. A modified BICM decoder has been developed at the relays to mitigate the influence of the interference signal. Furthermore, at the destination, we exploit the interference signal, which results from the simultaneous transmission of data streams through both direct and one of the relay channels, to develop an optimal detector. It is shown that the proposed detector can be implemented by parallel concatenating maximum a posteriori (MAP) algorithms and demappers to the decoders. The detector exchanges soft information between decoders and MAP algorithms in an iterative way for performance improvement. The major drawback of the proposed optimal detector is the delay, i.e., the destination has to receive and store all received packets before performing data detection. Due to the inevitable delay restriction, a sub-optimal detector is developed. In contrast with the optimal

one, the sub-optimal detector exploits two consecutive received packets to decode one packet.

In addition, the proposed algorithms have been extended for a new development of a multiple antenna destination. It is shown that the optimal receiver can be implemented by parallel demappers, each is based on a family of the BCJR algorithms, connected with parallel detectors in an iterative way. Unlike the optimal detector which requires to receive and store the entire packets before performing data detection, the sub-optimal one exploits two consecutive received packets to decode one packet.

The performance of the proposed receivers is assessed via Monte Carlo simulations, and the results illustrate their effectiveness. It turns out that the sub-optimal receiver outperforms the optimal one in terms of the required memory size, delay, bandwidth loss, and computational complexity.

Chapter 5

Conclusions and Future Work

Cooperative communications has recently received wide recognition as a simple way to improve system error rate performance and capacity. The key idea is that multiple terminals in a wireless network cooperate by relaying each other's information, forming a virtual antenna array and thus realizing spatial diversity in a distributed fashion. These cooperative techniques take advantage of the broadcast nature of the wireless channel by using the fact that a source signal intended for a particular destination can be overheard at neighboring nodes. It has been shown that cooperative communications bring several network enhancements such as improved coverage, increased capacity, and improved reliability in terms of diversity gain. Cooperative communications has found applications in various networks such as cellular, ad hoc, and sensor networks [15–17].

Cooperative communications have brought forth many interesting issues, challenges and open questions. One of the most important challenges is the reduction in spectral efficiency, which results from half duplex constraint at the relays and orthogonal relay transmission. Among various cooperative systems proposed to recover spectral efficiency, it was shown that alternate-relaying coop-

erative (ARC) system can recover a significant portion of the spectral efficiency loss [34–36, 38–43]). A key feature of this system is that the source continues to transmit data, while one of two relays receives and transmits the data from the source in turn. Due to simultaneous transmission of data streams through both direct and one of relay channels, harmful interference occur. This represents a major drawback for this system.

To the best of our knowledge, all the reported works for ARC systems have in common that they do not provide the optimal data detection solution and they restrict themselves to uncoded transmission and do not exploit any properties of the underlying error correcting codes. In this thesis, we have proposed novel data detection algorithms for uncoded and coded decode-and-forward (DF) ARC systems, with a destination node supporting single and multiple antenna.

In Chapter 3, we have proposed a maximum-likelihood detector at the relays. This is based on averaging out the interference signal. Furthermore, unlike previous work in which interference cancellation is required at the destination, we exploit the interference signal as a beneficial resource to develop the optimal detector at the destination. It is shown that the optimal detector can be implemented by parallel Viterbi algorithms. The major drawback of the proposed optimal detector is the delay because the destination has to receive and store the entire frame before performing data detection. Due to the inevitable delay restriction, a sub-optimal detector is developed. In contrast with the optimal detector, the sub-optimal detector exploits two consecutive received packets to decode one packet.

In addition, the proposed algorithms have been extended for a new development of a multiple antenna destination. We demonstrated that the optimal detector can be performed by parallel detectors, each is based on a family of BCJR algorithms. A sub-optimal detector has been also proposed to avoid the delay lim-

itation that is associated with the optimal one.

Generally, simulation results have been indicated that the performance of the optimal and sub-optimal detectors is very close to that of the ideal case with perfect separation of direct and relaying links. In addition, the sub-optimal detector outperforms the optimal one in terms of the required delay, memory size, bandwidth loss, and computational complexity.

In Chapter 4, we propose the use of bit-interleaved coded modulation (BICM) in DF-ARC systems. A modified BICM decoder has been developed at the relays to mitigate the influence of the interference signal. Furthermore, at the destination, we exploit the interference signal, which results from the simultaneous transmission of data streams through both direct and one of the relay channels, to develop an optimal detector. It is shown that the proposed detector can be implemented by parallel concatenating maximum a posteriori (MAP) algorithms and demappers to the decoders. The detector exchanges soft information between decoders and MAP algorithms in an iterative way for performance improvement. The major drawback of the proposed optimal detector is the delay, i.e., the destination has to receive and store all received packets before performing data detection. Due to the inevitable delay restriction, a sub-optimal detector is developed. In contrast with the optimal one, the sub-optimal detector exploits two consecutive received packets to decode one packet.

In addition, the proposed algorithms have been extended for a new development of a multiple antenna destination. It is shown that the optimal receiver can be implemented by parallel demappers, each is based on a family of the BCJR algorithms, connected with parallel detectors in an iterative way. Unlike the optimal detector which requires to receive and store the entire packets before performing data detection, the sub-optimal one exploits two consecutive received packets to

decode one packet.

The performance of the proposed receivers is assessed via Monte Carlo simulations, and the results illustrate their effectiveness. It turns out that the sub-optimal receiver outperforms the optimal one in terms of the required memory size, delay, bandwidth loss, and computational complexity.

Possible extensions of the work

- Developing optimal data detection algorithms for uncoded and coded ARC cooperative systems over frequency-selective channels.
- Developing new interference cancellation algorithms at the relays.
- Proposing parameters estimation algorithms for uncoded and coded ARC cooperative transmission.

Bibliography

[1] J. Parsons, M. Henze, P. Ratliff, and M. Withers, "Diversity Techniques for Mobile Radio Reception," *IEEE Transactions on Vehicular Technology*, vol. 25, no. 3, pp. 75–85, Aug. 1976.

[2] S. Lin, T. Lee, and M. Gardina, "Diversity protections for digital radio-summary of ten-year experiments and studies," *IEEE Communications Magazine*, vol. 26, no. 2, pp. 51–63, 1988.

[3] D. Gesbert, M. Shafi, S. Da-shan Shiu, P. Smith, and A. Naguib, "From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems," *IEEE Transactions on Selected Areas in Communications*, vol. 21, no. 3, pp. 281 – 302, Apr. 2003.

[4] A. Paulraj, D. Gore, R. Nabar, and H. Bolcskei, "An Overview of MIMO Communications - A Key to Gigabit Wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–218, No. 2004.

[5] X. Wang, X. Hong, X. Ge, X. Cheng, G. Zhang, and J. Thompson, "Cooperative MIMO Channel Models: A Survey," *IEEE Communications Magazine*, vol. 48, no. 2, pp. 80–87, 2010.

[6] John Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2000.

[7] Theodore S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Prentice-Hill, 2000.

[8] J. Mietzner, R. Schober, L. Lampe, W. Gerstacker, and P. Hoeher, "Multiple Antenna Techniques For Wireless Communications - A Comprehensive Literature Survey," *IEEE Communications Surveys and Tutorials*, vol. 11, no. Second Quarter, pp. 87–105, Mar. 2009.

[9] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space Time Block Codes From Orthogonal Designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[10] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative Communication in Wireless Networks," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 74–80, 2004.

[11] N. Laneman, D. Tse, and G. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[12] B. Talha and M. Patzold, "Channel Models for Mobile-to-Mobile Cooperative Communication Systems: A State of the Art Review," *IEEE Vehicular Technology Magazine*, vol. 6, no. 2, pp. 33–43, Mar. 2011.

[13] Z. Sheng, K. Leung, and Z. Ding, "Cooperative Wireless Networks: From Radio to Network Protocol Designs," *IEEE Communications Magazine*, vol. 49, no. 5, pp. 64–69, 2011.

[14] L. Yonghui, "Distributed Coding for Cooperative Wireless Networks: An Overview and Recent Advances," *IEEE Communications Magazine*, vol. 47, no. 8, pp. 71–77, 2009.

[15] M. Dohler and Y. Li, *Cooperative Communications: Hardware, Channel and PHY*. John Wiley - Sons Inc, 2010.

[16] V. Stankovic, A. Host-Madsen, and Z. Xiong, "Cooperative Diversity for Wireless Ad Hoc Networks," *IEEE Signal Processing Magazine*, vol. 23, no. 5, pp. 37–49, Sept. 2006.

[17] M. Lee, Z. Jianliang, K. Young-Bae, and D. Shrestha, "Cooperative Diversity for Wireless Ad Hoc Networks," *IEEE Wireless Communications*, vol. 13, no. 2, pp. 56–63, Apr. 2006.

[18] *IEEE Standard 802.11s [Online]. Available: http://www.ieee802.org/11/.*

[19] *IEEE Standard 802.15.3 [Online]. Available: http://www.ieee802.org/15/pub/TG3.html*.

[20] *IEEE Standard 802.15.4 [Online]. Available: http://www.ieee802.org/15/pub/TG4.html*.

[21] *IEEE Standard 802.15.5 [Online]. Available: http://www.ieee802.org/15/pub/TG5.html*.

[22] *IEEE Standard 802.16 [Online]. Available: http://standards.ieee.org/about/get/802/802.16.html*.

[23] *IEEE Standard 802.22 [Online]. Available: http://ieee802.org/22/.*

[24] K. Hwang, Y. Ko, and M. Alouini, "Performance Analysis of Incremental Opportunistic Relaying Over Identically and Non-identically Distributed Cooperative Paths," *IEEE Transactions on Wireless Communication*, vol. 8, no. 4, pp. 1953–1961, Apr. 2009.

[25] T. Nechiporenko, P. Kalansuriya, and C. Tellambura, "Performance of Optimum Switching Adaptive M-QAM for Amplify-and-forward Relays," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 5, pp. 2258–2268, Jun. 2009.

[26] A. Bastami and A. Olfat, "Selection Relaying Schemes for Cooperative Wireless Networks With Adaptive Modulation," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 5, pp. 1539–1558, May 2011.

[27] Y. Ma, R. Tafazolli, Y. Zhang, and C. Qian, "Adaptive Modulation for Opportunistic Decode-and-Forward Relaying," *IEEE Transactions on Wireless Communication*, vol. 10, no. 7, pp. 2017–2022, Jul. 2011.

[28] E. Larsson and B. Vojcic, "Cooperative Transmit Diversity Based on Superposition Modulation," *IEEE Communications Letters*, vol. 9, no. 9, pp. 778 –780, Nov. 2005.

[29] L. Xiao, T. Fuja, J. Kliewer, and D. Costello, "Error Performance Analysis of Signal Superposition Coded Cooperative Diversity," *IEEE Transactions on Communications*, vol. 57, no. 10, pp. 3123–3131, Oct. 2009.

[30] I. Krikidis, "Analysis and Optimization Issues for Superposition Modulation in Cooperative Networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 9, pp. 4837–4847, Nov. 2009.

[31] T. Yune, D. Kim, and G.-H. Im, "Iterative Detection for Spectral Efficient User Cooperative Transmissions over Multipath Fading Channels," *IEEE Transactions on Communication*, vol. 58, no. 4, pp. 1121–1128, Apr. 2010.

[32] Z Ding, I Krikidis, B Rong, J Thompson, C Wang, and S Yang, "On Combating the Half-duplex Constraint in Modern Cooperative Networks: Protocols and

Techniques," *IEEE Transactions on Wireless Communcations*, vol. 19, no. 6, pp. 20–27, Dec. 2012.

[33] L. Rodriguez, N. Tran, and T. Le-Ngoc, "Multiple Frame Precoding for NAF Relaying over Rayleigh Fading Channels," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 1, pp. 398–404, Jan. 2012.

[34] P. Rost and G. Fettweis, "A Cooperative Relaying Scheme without the Need for Modulation with Increased Spectral Efficiency," in *Proc. of the IEEE Vehicular Technology,*, pp. 1- 5, 2006.

[35] C. Luo, Y. Gong, and F. Zheng, "Full Interference Cancellation for Two-Path Cooperative Communications," in *Proc. of the IEEE Wireless Communications and Networking*, pp. 1- 5, 2009.

[36] P. Wang, S. Hassan, and Y. Li, "A Full Rate Symmetrical Cooperative Relay Approach for Wireless Systems," in *Proc. of the IEEE Circuits and Systems for Communications*, pp. 104 - 108, 2008.

[37] J. Kwon, Y. Ko, and H. Yang, "Maximum Spectral Efficiency of Amplify-and-Forward Cooperative Transmission With Multiple Relays," *IEEE Transactions on Wireless Communication*, vol. 10, no. 1, pp. 49–54, Jan. 2011.

[38] B. Rankov and A. Wittneben, "Spectral Efficient Protocols for Half Duplex Fading Relay Channels," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[39] C. Luo, Y. Gong, and F. Zheng, "Full Interference Cancellation for Two-Path Relay Cooperative Networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 343–347, Jan. 2011.

[40] H. Wicaksana, S. Ting , C. Ho, W. Chin, and Y. Guan, "AF Two-Path Half Duplex Relaying with Inter-Relay Self Interference Cancellation: Diversity Analysis and its Improvement," *IEEE Transactions on Wireless Communication*, vol. 8, no. 9, pp. 4720–4729, Sep. 2009.

[41] A. Ribeiro, X. Cai, and G. Giannakis, "Opportunistic Multipath for Bandwidth Efficient Cooperative Multiple Access," *IEEE Transactions on Information Theory*, vol. 5, no. 9, pp. 2321–2327, Sept. 2006.

[42] C. Wang, Y. Fan, J. Thompson, M. Skoglund, and H. Poor, "Approaching the Optimal Diversity-Multiplexing Tradeoff in a Four-node Cooperative Network ," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 3690–3700, Dec. 2010.

[43] L. Sun, T. Zhang, H. Niu, "Inter-Relay Interference in Two-Path Digial Relaying Systems: Detrimental or Beneficial?" *IEEE Transactions on Wireless Communication*, vol. 10, no. 8, pp. 2468–2473, Aug. 2011.

[44] M. Chen and A. Yener, "Multiuser Two-Way Relaying: Detection and Interference Management Strategies," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 4269–4305, Aug. 2009.

[45] P. Xu, X. Dai, Z. Ding, I. Krikidis, and K. Leung, "Approaching MISO Upper Bound: Design of New Wireless Cooperative Transmission Protocols," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2725–2737, Aug. 2011.

[46] C. Wang, Y. Fan, and H. Poor, "A Comprehensive Study of Repetition-coded Protocols in Multi-user Multi-relay Networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2725–2737, Aug. 2009.

[47] R. Tannious and R. Nosratinia, "Spectrally-Efficient Relay Selection with Limited Feedback," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1419–1428, Oct. 2008.

[48] R. Louie, Y. Li, and B. Vucetic, "Practical Physical Layer Network Coding for Two-Way Relay Channels: Performance Analysis and Comparison," *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, pp. 764–777, Feb. 2010.

[49] E. Zehavi, "8-PSK Trellis Codes for a Rayleigh Fading Channels," *IEEE Transactions on Communications*, vol. 40, no. 5, pp. 873–883, May 1992.

[50] S. Pfletschinger and F. Sanzi, "Error Floor Removal for Bit-Interleaved Coded Modulation with Iterative Detection," *IEEE Transactions on Wireless Communications*, vol. 11, no. 5, pp. 3174–3181, Nov. 2006.

[51] X. Li and J. Ritcey, "Bit-Interleaved Coded Modulation with Iterative Decoding," *IEEE Communications Letters*, vol. 1, no. 6, pp. 169–171, Nov. 1997.

[52] ——, "Trellis-Coded Modulation with Bit Interleaving and Iterative Decoding," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 715–724, Apr. 1999.

[53] ——, "Bit Interleaved Coded Modulation with Iterative Decoding and 8PSK Signaling," *IEEE Transactions on Communications*, vol. 50, no. 8, pp. 1250–1257, Aug. 2002.

[54] D. Divsalar and M. Simon, "The Design of Trellis Coded MPSK for Fading Channel: Performance Criteria," *IEEE Transactions on Communications*, vol. 36, no. 9, pp. 1004–1012, Sep. 1988.

[55] ——, "The Design of Trellis Coded MPSK for Fading Channel: Set Partitioning for Optimum Code Design," *IEEE Transactions on Communications*, vol. 36, no. 9, pp. 1013–1021, Sep. 1988.

[56] L. Wei, "Trellis Coded Modulation with Multi-Dimensional Constellations," *IEEE Transactions on Information Theory*, vol. 33, no. 6, pp. 483–501, Jul. 1987.

[57] S. Alamouti and S. Kallel, "Adaptive Trellis Coded Multiple Phased Shift Keying Rayleigh Fading Channels," *IEEE Transactions on Communications*, vol. 42, no. 6, pp. 2305–2341, Jun. 1994.

[58] L. Hanzo, T. Liew, and B. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels*. Wiley, 2002.

[59] Y. Fan, C. Wang, J. Thompson, and H. Poor, "Recovering Multiplexing Loss through Successive Relaying Using Repetition Coding," *IEEE Transactions on Wireless Communications*, vol. 6, no. 12, pp. 4484–4493, Dec. 2007.

[60] W. Chang, S. Chung, and Y. H. Lee, "Capacity Bounds for Alternating Two-path Relay Channels," in *Proc. Allerton Conf. Communication, Control Computing, Monticello, IL*, pp. 1- 5, 2007.

[61] F. Xue and S. Sandhu, "Cooperation in a Half-duplex Gaussian Diamond Relay Channel," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3806–3814, Oct. 2007.

[62] R. Zhang, "Characterizing Achievable Rates for Two-path Digital Relaying," in *Proc. IEEE Int. Conf. Communication*, pp. 1- 5, 2008.

[63] Y. Han, S. Ting, C. Ho, W. Chin, and H. Poor, "Performance Bounds for Two-way Amplify-and-Forward Relaying," *IEEE Transactions on Wireless Communications*, vol. 8, no. 1, pp. 432–439, Jan. 2009.

[64] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, 1993.

[65] H. Trees, *Detection, Estimation, and Modulation Theory*. Wiley, 2004.

[66] J. Heller and I. Jacobs, "Viterbi Decoding for Satellite and Space Communication," *IEEE Journal on Communication Technology*, vol. 19, no. 5, pp. 835–848, Oct. 1971.

[67] S. Lin and D. Costello, *Error Control Coding: Fundamentals and Applications*, 2nd ed. Pearson-Prentice Hall, 2004.

[68] G. Ungerboeck, "Trellis Coded Modulation with Redundant Signal Sets Part I: Introduction," *IEEE Communications Magazine*, vol. 25, no. 2, pp. 5–11, Feb. 1987.

[69] ——, "Trellis Coded Modulation with Redundant Signal Sets Part II: State of the Art," *IEEE Communications Magazine*, vol. 25, no. 2, pp. 12–21, Feb. 1987.

[70] P. Elias, "Coding for Noisy Channels," *IRE Conv. Rec., Part 4*, pp. 37–46, Nov. 1955.

[71] A. J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," *IEEE Transactions on Information Theory*, vol. 13, no. 2, pp. 260–269, Apr. 1967.

[72] A. Viterbi and J. Odenwalder, "Further Results on Optimal Decoding of Convolutional Codes," *IEEE Transactions on Information Theory*, vol. 15, pp. 732–734, Nov. 1969.

[73] G. Forney, "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.

[74] J. Wolf, "A Survey of Coding Theory: 1967-1972," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 381–389, Nov. 1973.

[75] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. McGraw-Hill, 1991.

[76] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[77] C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding: Turbo-codes," *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.

[78] B. Sklar, "A Primer on Turbo Code Concepts ," *IEEE Communications Magazine*, vol. 35, no. 12, pp. 94–102, Nov. 1997.

[79] N. Bonello, C. Sheng, and L. Hanzo, "Low-Density Parity-Check Codes and Their Rateless Relatives," *IEEE Communications Surveys and Tutorials*, vol. 13, no. 1, pp. 3–26, First Quarter 2011.

[80] P. Robertson, P. Hoeher, and E. Villebrun, "Optimal and Suboptimal Maximum A posteriori Algorithms Suitable for Turbo Decoding," *European Transactions Telecommunications*, vol. 8, no. 2, pp. 119–125, Mar-Apr. 1997.

[81] G. Caire, G. Taricco and E. Biglieri, "Bit Interleaved Coded Modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[82] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "A Soft-Input Soft-Output APP Module for Iterative Decoding of Concatenated Codes," *IEEE Communications Letters*, vol. 1, no. 1, pp. 22–24, Jan. 1997.

[83] J. Hagenauer, "The Turbo Principle: Tutorial Introduction and State of the Art ," in *Proc. of International Symposium on Turbo-Codes*, Brest, France, Sep. 1997.

[84] P. Regalia, "Iterative Decoding of Concatenated Codes: A Tutorial," *EURASIP Journal on Applied Signal Processing*, no. 6, pp. 762–774, 2005.

[85] N. Nghi and H. Nguyen, "Signal Mappings of 8-ary Constellations for Bit Interleaved Coded Modulation with Iterative Decoding," *IEEE Transactions on Broadcasting*, vol. 52, no. 1, pp. 92–99, Mar. 2006.

[86] A. Chindapol and J. Ritcey, "Design, Analysis and Performance Evaluation for BICM-ID with Square QAM Constellations in Rayleigh Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. vol. 19, no. 5, pp. 944–757, May 2001.

[87] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels*. Wiley, 2002.

[88] S. Wei, D. L. Goeckel, and M. Valenti, "Asynchronous Cooperative Diversity," *IEEE Transactions on Communication*, vol. 5, no. 6, pp. 1547–1557, Jun. 2006.

[89] Y. Shang and X.-G. Xia, "Shift-full-rank Matrices and Applications in Space-time Trellis Codes for Relay Networks with Asynchronous Cooperative Di-

versity," *IEEE Transactions on Infomation Theory*, vol. 52, no. 7, pp. 3153–3167, Jul. 2006.

[90] A. Viterbi, "Convolutional Codes and Their Performance in Communication Systems," *IEEE Transactions on Communication*, vol. 19, no. 5, pp. 751–772, Oct. 1971.

[91] P. Patel and J. Holtzman, "Analysis of a Simple Successive Interference Cancellation Scheme in a DS/CDMA System," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 5, pp. 796–807, Jun. 1994.

[92] T. C. Yoon, R. Kohno, and H. Imai, "A Spread-spectrum Multi-access System with Cochannel Interference Cancellation for Multipath Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 7, pp. 1067–1075, Sep. 1993.

[93] S. Verdu, *Multiuser Detection*. Cambridge University Press, 1998.

[94] J. Heller and I. Jacobs, "Viterbi Decoding for Satellite and Space Communication," *IEEE Transactions on Communications*, vol. 19, pp. 835–848, Oct. 1971.

[95] D. Shiu, G. Foschini, M. Gans, and J. Kahn, "Fading Correlation and Its Effect on the Capacity of Multielement Antenna Systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[96] E. Ko and D. Hong, "A Robust STBC-Based Transmit Diversity Scheme for OFDM Systems Over Spatially Transmit Correlated Fading Channels," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 984–991, Mar. 2007.

[97] S. Watkins, *Fundamentals of Matrix Computations*. Wiley, 2002.

[98] L. Sanguinetti, M. Morelli, and G. Imbarlina, "An EM-based Frequency Offset Estimator for OFDM Systems with Unknown Interference," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4470–4475, Sept. 2009.