Distributed Space-Time Coding Techniques with Limited Feedback in Cooperative MIMO Networks

Tong Peng

Doctor of Philosophy

University of York

Electronics

February 2014

Abstract

DSTC designs with high diversity and coding gains and efficient detection and code matrices optimization algorithms in cooperative MIMO networks are proposed in this thesis. Firstly, adaptive power allocation (PA) algorithms with different criteria for a cooperative MIMO system equipped with DSTC schemes are proposed and evaluated. Linear receive filter and maximum likelihood (ML) detection are considered with amplify-and-forward (AF) and decode-and-forward (DF) cooperation strategies. In the proposed algorithms, the elements in the PA matrices are optimized at the destination node and then transmitted back to the relay nodes via a feedback channel. Linear minimum mean square error (MMSE) receive filter expressions and the PA matrices depend on each other and are updated iteratively. Stochastic gradient (SG) algorithms are developed with reduced detection complexity. Secondly, an DSTC scheme is proposed for two-hop cooperative MIMO networks. An adjustable code matrix obtained by a feedback channel is employed to transform the space-time coded matrix at the relay node. The effects of the limited feedback and the feedback errors are assessed. An upper bound on the pairwise error probability analysis is derived and indicates the advantage of employing the adjustable code matrices at the relay nodes. An alternative optimization algorithm for the adaptive DSTC scheme is also derived in order to eliminate the need for feedback. Thirdly, an adaptive delay-tolerant DSTC (DT-DSTC) scheme is proposed for two-hop cooperative MIMO networks. An ML receiver and adjustable code matrices are considered for different DSTC configuration schemes subject to a power constraint with a DF cooperation strategy. An upper bound on the pairwise error probability and rank criteria analysis are derived and indicates the advantage of the proposed coding algorithm. Adaptive DT-DSTC algorithms are extended to the cooperative MIMO systems using an AF strategy and opportunistic relaying algorithms in order to achieve a delay-tolerant coding scheme combined with the optimal power allocation strategies.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to show my sincere gratitude to my supervisor, Prof. Rodrigo C. de Lamare, for his support, supervision and guidance not only on my research but also on career and future development.

I am very grateful to my thesis advisor, Prof. Alister G. Burr, whose insightful discussions and suggestions have benefited me.

I would also like to thank Dr. Peng Li, Dr. Yi Wang, Dr. JingJing Liu and Dr. Li Li and other colleagues in the Communications Research Group.

This thesis is dedicated to my parents and my love Ling Ding.

Declaration

This work has not previously been presented for an award at this, or any other, University. Some of the research presented in this thesis has resulted in some publications. These publications are listed as follows.

All work presented in this thesis as original is so, to the best knowledge of the author. References and acknowledgements to other researchers have been given as appropriate.

Journal Papers

1. T. Peng, R. C. de Lamare, A. Schmeink, "Adaptive Power Allocation Strategies for DSTC in Cooperative MIMO Networks", *IET Commun.*, 2013 (accepted).
2. T. Peng, R. C. de Lamare, A. Schmeink, "Distributed Space-Time Coding Based on Adjustable Code Matrices for Cooperative MIMO Relaying Systems", *IEEE Trans. Commun.*, 2012 (accepted).
3. T. Peng, R. C. de Lamare, "Adaptive Delay-Tolerant Distributed Space-Time Coding Based on Adjustable Code Matrices for Cooperative MIMO Relaying Systems", *IEEE Trans. Veh. Technol.*, 2014 (under review).

Conference Papers

1. T. Peng, R. C. de Lamare, A. Schmeink, "Joint Power Allocation and Receiver Design for Distributed Space-Time Coded Cooperative MIMO Systems", *2011 8th International Symposium on Wireless Communication Systems (ISWCS)*, Nov. 2011.
2. T. Peng, R. C. de Lamare, A. Schmeink, "Joint Minimum BER Power Allocation and Receiver Design for Distributed Space-Time Coded Cooperative MIMO Relay-

ing Systems", *2012 International ITG Workshop on Smart Antennas (WSA)*, March 2012.

3. T. Peng, R. C. de Lamare, A. Schmeink, "Adaptive Distributed Space-Time Coding for Cooperative MIMO Relaying Systems", *2012 International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2012.

4. T. Peng, R. C. de Lamare, A. Schmeink, "Adaptive Distributed Space-Time Coding for Cooperative MIMO Relaying Systems with Limited Feedback", *2013 IEEE Vehicular Technology Conference (VTC Spring)* , Jun. 2013.

5. T. Peng, R. C. de Lamare, "Adaptive Delay-Tolerant Distributed Space-Time Coding in Opportunistic Relaying Cooperative MIMO Systems", *2014 International Telecommunications Symposium (ITS)* , 2014 (under preparation).

Chapter 1

Introduction

Contents

1.1 Overview

Multiple-input multiple-output (MIMO) wireless communication systems employ multiple collocated antennas at both source and destination nodes in order to obtain diversity gains and combat fading. Various space-time coding (STC) schemes, which provide high diversity and coding gains compared to uncoded schemes, are also utilized in MIMO wireless systems [1]. However, applying MIMO in mobile communication systems suffers from high cost of computational complexity and the size of mobile terminals.

Cooperative MIMO systems can obtain diversity gains by providing copies of the transmitted signals with the help of relays to improve the reliability of wireless communication systems [2] - [9]. The basic idea behind these cooperative relaying systems is to employ multiple relay nodes between the source node and the destination node to form

a distributed antenna array, which can provide significant advantages in terms of diversity gains. Several cooperation strategies that exploit the links between the relay nodes and the destination node such as amplify-and-forward (AF), decode-and-forward (DF), compress-and-forward (CF) [2] and distributed space-time coding (DSTC) schemes [3, 4] that employ space-time codes [4, 5] have been extensively studied in the literature. A key problem that arises in cooperative MIMO systems and which degrades the performance of such systems is the existence of delays between the signals that are space-time coded at the relays and decoded at the destination.

In this thesis, optimal power allocation (PA) algorithms, novel distributed STC (D-STC) designs with adaptive code matrices optimization algorithms are proposed which achieve higher coding gains and lower complexity detection algorithms compared to existing techniques. Firstly, PA optimization algorithms [10] - [13] are considered in a unified framework, and then joint adaptive power allocation (JAPA) algorithms according to different optimization criteria with a linear receiver or a maximum likelihood (ML) detector for cooperative MIMO systems employing multiple relay nodes with multiple antennas are designed. Analyses on the computational complexity and the influence of feedback are presented which confirms the advantages of the proposed algorithms.

Secondly, adaptive centralized DSTC schemes and algorithms with limited feedback for cooperative MIMO relaying systems are investigated. Low-complexity encoding and optimization algorithms are developed and compared to existing coding schemes. In particular, different DSTC schemes in [14] - [16] are tested and compared with the designed coding schemes. In order to eliminate the need for feedback channels and achieve higher coding gains, a fully distributed optimization algorithm is presented and analyzed based on the centralized schemes.

Thirdly, the challenges of delay-tolerant code design with optimal PA and relay selection algorithms are addressed by novel delay-tolerant adaptive coding matrices optimization algorithms. Two basic configurations of distributed space-time coding schemes in cooperative MIMO systems are analyzed to indicate the advantages of the space-time coding designs and algorithms. By combining the proposed delay-tolerant coding scheme and opportunistic relaying algorithms in [14], a delay-tolerant coding scheme with optimal power allocation and relay selection algorithms is devised and investigated in several

scenarios of practical interest.

1.2 Contributions

The main contributions of this thesis can be structured and summarized as follows:

- Joint adaptive power allocation (JAPA) algorithms according to different optimization criteria with a linear receiver or an ML detector for cooperative MIMO systems employing multiple relay nodes with multiple antennas to design cooperating strategies. The diagonal power allocation matrices are employed in which the parameters stand for the power allocated to each transmit antenna. Adaptive stochastic gradient (SG) algorithms are devised in order to determine the linear receiver and power allocation parameters iteratively with low computational complexity. A normalization procedure is employed in the optimization algorithms in order to enforce the power constraint in both transmission phases and reduce the computational complexity of computing the required Lagrange multiplier. The effect of the feedback errors is considered in the analyses and in the simulation study, where it leads to a degradation in the mean square error (MSE) performance. The requirement of the limited feedback is significantly reduced in terms of the number of required bits per coefficient as compared to the algorithms in the literature. The JAPA strategies derived in our algorithms are two-phase optimization approaches, which optimize the power assigned at the source node and at the relay nodes in the first phase and the second phase iteratively, and the JAPA algorithms also can be used as power allocation strategies for the second phase only.

- Adaptive distributed space-time coding scheme and algorithms for cooperative MIMO relaying systems. A centralized algorithm with limited feedback is developed to compute the parameters of an adjustable code matrix, which requires sending the adjustable code matrices back to the relay nodes after the optimization via a feedback channel that is modeled as a binary symmetric channel (BSC). Then, adaptive optimization algorithms are derived based on the MSE and the ML criteria subject to constraints on the transmitted power at the relays, in order to release the destination node from the high computational complexity of the optimization process.

We focus on how the adjustable code matrix influences the DSTC during the encoding procedure and how to optimize the linear receive filter with the code matrix iteratively or, alternatively, by employing an ML detector and adjusting the code matrices. An upper bound of the pairwise error probability (PEP) of the designed adaptive DSTC is derived in order to show its advantages as compared to the traditional DSTC schemes. The influence of the imperfect feedback is discussed after the analyses of the code. It is shown that the use of an adjustable code matrix benefits the performance of the system compared to employing traditional STC schemes. Then, we derive a fully distributed matrix optimization algorithm which does not require feedback from the destination. The PEP of the adaptive DSTC is employed in order to devise a fully distributed algorithm and to eliminate the need for feedback channels. The fully distributed matrix optimization algorithm allows the system to determine the optimal adjustable matrix before the transmission, and also achieves the minimum PEP when the statistical information of the channel does not change.

- We introduce adjustable code matrices into the space-time encoding and optimization process. An optimal adjustable code matrix is multiplied by an existing space-time coding scheme at the relay node and the encoded data are forwarded to the destination node. The code matrix is first generated randomly as discussed in [6], and it is optimized according to different criteria at the destination node by the optimization algorithms. The decoding algorithms for different DSTC schemes are not affected by employing adjustable code matrices. In order to implement the adaptive algorithms, the adjustable code matrix is optimized with the linear receive filter iteratively, and then transmitted back to the relays via a feedback channel. The impact of the feedback errors is considered and shown in the simulations. The fully distributed optimization algorithm eliminates the need for feedback by choosing the optimal code matrix before transmission, and the receiver is released from the design task.

- A delay-tolerant adjustable code matrices optimization (DT-ACMO) algorithm based on the ML criterion subject to constraints on the transmitted power at the relays is proposed for different cooperative systems. Adaptive optimization algorithms using SG and RLS estimation methods are developed for the DT-ACMO algorithm in order to release the destination node from the high computational complexity of the optimization procedure. Studies of how the adjustable code matrix

affects the DSTBCs during the encoding process and how to optimize the adjustable code matrices by employing an ML detector are carried out. Then analyses of the differences in terms of the rank criterion and pairwise error probabilities of the D-STBCs in these two system configurations with the same number of antennas and the same delay profiles are presented. We study how the different systems affect the delay tolerance of the DSTBCs and conclude that the DSTBCs in the cooperative systems with multi-antenna relay nodes can address the delays from the relay nodes compared to the cooperative system employing single-antenna relays.

- The DT-ACMO algorithms are extended to cooperative systems using an AF protocol with the opportunistic relaying selection algorithms in [14]. Delay-tolerant adjustable code matrices opportunistic relaying optimization (DT-ACMORO) algorithms are developed in order to address the delay issue among relay nodes. Two basic configurations of distributed space-time coding schemes are studied in terms of rank criterion and error probability of STC scheme. The DT-ACMORO algorithms can be implemented with different types of STC schemes in DF and AF protocols and overcome the issues caused by the delay among relay nodes.

1.3 Thesis Outline

The structure of the thesis is listed as follows:

- Chapter 2 presents a literature review of the techniques that are relevant to this thesis in wireless and mobile communications and introduce the system models considered in the thesis. The review of MIMO techniques in terms of capacity and diversity order, and STC and DSTC design criteria, and detection and estimation techniques are given in this chapter.

- Chapter 3 presents novel joint adaptive power allocation algorithms according to different optimization criteria with a linear receiver or an ML detector for cooperative MIMO systems employing multiple relay nodes with multiple antennas to design cooperating strategies. The computational complexity of different power allocation techniques is detailed and compared with joint adaptive optimization algorithms.

- Chapter 4 presents centralized and fully-distributed adaptive space-time coding scheme and algorithms for cooperative MIMO relaying systems with and without the requirement of feedback channels. SG and RLS methods are employed in the DSTC scheme and algorithms to reduce the computational complexity. The error probability of the adaptive DTC design is derived along with a study of the effect of feedback channels.

- Chapter 5 presents delay-tolerant adjustable code matrices optimization (DT-ACMO) algorithms based on ML criterion subject to constraints on the transmitted power at the relays for different cooperative systems. Two types of distributed MIMO networks are considered. The extension of the DT-ACMO algorithms combined with the opportunistic relaying algorithms in [33] is studied. Analyses of the proposed and existing schemes in terms of rank criterion and pairwise error probability of the algorithms are given as well.

- Chapter 6 presents conclusions and discusses possible future work based on the content of the thesis.

1.4 Notation

$E[\cdot]$	expectation operator
$(\cdot)^H$	Hermitian operator
$()^*$	complex conjugate
$\boldsymbol{I}_N$	$N \times N$ identity matrix
$\| \cdot \|_F$	Frobenius norm, $\| \boldsymbol{X} \|_F = \sqrt{\mathrm{Tr}(\boldsymbol{X}^H \cdot \boldsymbol{X})} = \sqrt{\mathrm{Tr}(\boldsymbol{X} \cdot \boldsymbol{X}^H)}$
$\Re[\cdot]$	real part
$\Im[\cdot]$	imaginary part
$Tr(\cdot)$	trace of a matrix
$(\cdot)^\dagger$	pseudo inverse
$\bigotimes$	Kronecker product

Chapter 2

Literature Review

Contents

In this chapter, an introduction to fundamental techniques related to the research carried out during the preparation of this thesis such as MIMO, STC and DSTC techniques, cooperative relaying systems, and parameter estimation and detection algorithms is presented.

2.1 MIMO Systems

MIMO systems employ multiple antennas at both the transmitting node and the receiving device in order to achieve a considerable improvement in signal quality and increase the capacity in terms of data transmission without the cost of extra spectrum in wireless communications [17] - [20]. Unlike the traditional single-input single-output (SISO) systems with communication between single antennas, MIMO systems effectively solve the

problems in wireless communications caused by the limited availability of radio frequency spectrum and time-variant propagation environments by employing multiple antennas. Additionally, taking full advantage of multipath effects to benefit communication is another outstanding feature of MIMO systems [17]. Due to these properties and their great potential in wireless communications, MIMO systems have been used in a wide range of applications and standards (such as 3GPP, WIFI, etc.), gradually replacing SISO systems.

2.1.1 Capacity Aspects

The most important contribution of MIMO systems lies in the increase of the capacity of a wireless communication system compared to that of a SISO system [19]. We consider a MIMO system which contains a transmitter with N antennas and a receiver with M antennas. A wireless block random fading channel represented by an $N \times M$ matrix $\boldsymbol{H}[i]$ is considered between the transmitter and the receiver, and we assume $M \leq N$. Define a symbol stream represented by an $N \times 1$ vector $\boldsymbol{s}[i]$ which contains N modulated elements by a K-PSK or a K-QAM scheme. The transmission is described by the model given by [19]

$$\boldsymbol{r}[i] = \sqrt{\frac{P_T}{N}}\boldsymbol{H}[i]\boldsymbol{s}[i] + \boldsymbol{n}[i], \tag{2.1}$$

where $\boldsymbol{r}$ denotes the $M \times 1$ received signal vector at the receiver, $\boldsymbol{n}[i]$ denotes the $M \times 1$ additive white Gaussian noise (AWGN) vector with entries having zero mean and variance σ_n^2 generated at the receiver and P_T is the transmit power assigned to all the antennas at the transmitter over a symbol period. The element we are most interested in is the channel matrix $\boldsymbol{H}[i]$, more specifically the eigenvectors and the singular values of the covariance matrix $\boldsymbol{H}[i]\boldsymbol{H}^H[i]$. If the channel matrix is unknown at the transmitter, the capacity of the above MIMO system is derived as [19]

$$\mathrm{C} = E\left[\log_2 \det\left(\boldsymbol{I}_N + \frac{1}{N}\frac{P_T}{\sigma_n^2}\boldsymbol{H}[i]\boldsymbol{H}^H[i]\right)\right]. \tag{2.2}$$

By performing an eigenvalue decomposition of $\boldsymbol{H}[i]\boldsymbol{H}^H[i]$, we can rewrite the expression of the MIMO capacity in (2.2) as

$$\begin{aligned} \mathrm{C} &= \log_2 \det\left(\boldsymbol{I}_N + \frac{1}{N}\frac{P_T}{\sigma_n^2}\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^H\right) \\ &= \log_2 \det\left(\boldsymbol{I}_N + \frac{1}{N}\frac{P_T}{\sigma_n^2}\boldsymbol{\Lambda}\right) \\ &= \sum_{i=1}^{r} \log_2(1 + \frac{\rho}{N}\lambda_i), \end{aligned} \tag{2.3}$$

where $r = \mathrm{rank}(\boldsymbol{H}[i]\boldsymbol{H}^H[i])$, $\rho = P_T/\sigma_n^2$, and λ_i denotes the $i-th$ positive eigenvalue of the matrix $\boldsymbol{H}[i]\boldsymbol{H}^H[i]$. It is worth to mention that according to (2.3) the capacity of an $N \times M$ MIMO system is equal to the sum of the capacity of M parallel SISO links with signal-to-noise ratio (SNR) equal to $\frac{\rho}{N}$.

The capacity derived in (2.3) is known as the capacity of an open-loop MIMO system which requires the channel state information (CSI) to be known at the receiver but not at the transmitter. If the channel is known at both the transmitter and the receiver, it is called as a closed-loop scenario, and the ergodic capacity is given by

$$\mathrm{C} = E\left[\log_2 \det\left(\boldsymbol{I}_N + \frac{1}{N}\frac{P_T}{\sigma_n^2}\boldsymbol{H}[i]\boldsymbol{R}\boldsymbol{H}^H[i]\right)\right], \tag{2.4}$$

where $\boldsymbol{R} = E\{\boldsymbol{s}[i]\boldsymbol{s}^H[i]\}$ denotes the autocorrelation of the symbol vector $\boldsymbol{s}[i]$. It is worth to mention that the wireless channel between the source and the destination is flat fading. In the open-loop case, equal power is assigned among the transmit antennas so that the autocorrelation is an identity matrix $\boldsymbol{I}_N$. On the contrary in the closed-loop scenario, the CSI is known at the transmitter which allows the transmitter to assign different transmit power to different antennas. By utilization of the singular value decomposition (SVD), we can rewrite $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H$ by using $M \times M$ and $N \times N$ complex unitary matrices $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively, and an $M \times N$ rectangular matrix whose elements in the main diagonal are singular values of the channel matrix $\boldsymbol{H}[i]$. The system capacity can be written as

$$\mathrm{C} = \sum_{i=1}^{r} \log_2(1 + \frac{\rho\gamma_i}{N}\lambda_i), \tag{2.5}$$

where γ_i denotes the optimal power. Thus, the capacity of a closed-loop MIMO system

related to the power assigned to each transmit antenna is given by

$$\mathrm{C}_{max} = \underset{\sum_{i=1}^{r}\gamma_i=N}{\operatorname{argmax}} \sum_{i=1}^{r} \left(\log_2(1 + \frac{\rho\gamma_i}{N}\lambda_i) \right). \tag{2.6}$$

The Water Filling algorithm [20] is employed in order to obtain the optimal value of γ_i which is derived as

$$\gamma_i = \left(\mu - \frac{N}{\rho\lambda_i} \right)^+, \ \sum_{i=1}^{r} \gamma_i = P_T, \quad \text{for } i = 1, 2, ..., r, \tag{2.7}$$

where μ is a constant and

$$(x)^+ = \begin{cases} x, & \text{if x} \geq 0, \\ 0, & \text{if x} < 0. \end{cases} \tag{2.8}$$

Therefore, the capacity of a closed-loop MIMO system is given by

$$\mathrm{C} = \sum_{i=1}^{r} \log_2 \left(1 + \frac{\rho\gamma_i}{N}\lambda_i \right). \tag{2.9}$$

In Fig. 2.1, the ergodic capacity of MIMO systems (2.5) with different configurations of antennas is shown. The channel is assumed to be Rayleigh fading. As shown in the figure, a SISO system has the worst capacity performance, while the utilization of multiple antennas only at the receiver or at the receiver can improve the capacity of a communication system compared to that of a SISO system. By using multiple antennas both at the transmitter and the receiver, the ergodic capacity increases dramatically. The difference between the MISO and SIMO is that due to the absence of CSI at the transmitter, the MISO system does not offer array gain.

2.1.2 Diversity and Spatial Multiplexing

Studies of diversity and multiplexing in a MIMO system are presented in [19] - [20]. Multiple antennas are used in wireless systems in order to increase the diversity to combat the channel fading. The transmission of signals carrying the same information through different antennas and paths allows the receiver to obtain multiple independently faded copies of the symbols, which indicates it enhances the reception reliability at the receiver. If

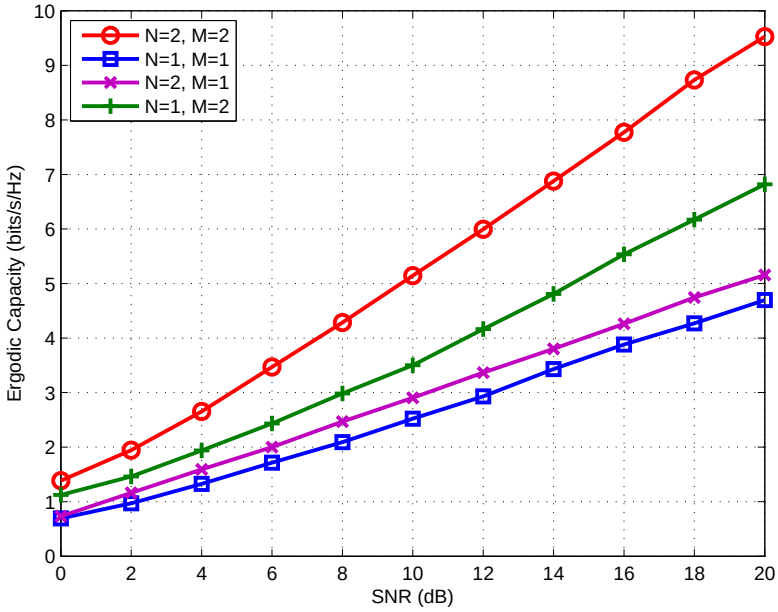

Figure 2.1: Ergodic Capacity of MIMO System with Different Number of Antennas

the fading between the pair of transmit and receive antennas is independent, the maximal diversity gain which is equal to MN will be achieved in the MIMO system discussed in the previous subsection [20]. Different types of diversity can be achieved in different multi-antenna systems. In a single-input multiple-output (SIMO) system, receive diversity can be obtained if the channels between the transmit antenna and the receive antennas are independent. In this case, the maximum receive diversity order is equal to the number of antennas at the receiver. If multiple antennas are used at the transmitter and a single antenna is employed at the receiver, a multiple-input single-input (MISO) system is constructed with the maximum transmit diversity being equal to the number of transmit antennas. Space-time coding techniques, such as [21] and [22], focused on achieving the maximum of the diversity gain have been developed and employed in MIMO systems.

However, the transmission of the symbols with the same number of information content will decrease the data rate of a MIMO system. By transmitting independent symbols with different information content can increase the available degrees of freedom for MIMO communication which leads to increased or higher transmission rates. It is called spatial multiplexing and is important in the high-SNR scenario in a MIMO system as the degrees of freedom in the high-SNR scenario is equal to $\min(M, N)$ [23,24]. In the high-SNR regime, the capacity of an $N \times M$ MIMO network with independent and identically

distributed (i.i.d) Rayleigh fading channel is studied in [23] and is given by

$$\mathrm{C}(SNR) = \min\{N, M\} \log_2(SNR) + \mathrm{O}(1). \tag{2.10}$$

The degrees of freedom is equal to the minimum of N and M which is related to the throughput of a MIMO system so that the spatial multiplexing transmission focuses on the maximization of the system capacity. By increasing the number of transmit and receive antennas, the spatial multiplexing transmission allows the linear increase of transmission rate [25]. For example, the Bell Laboratory Layered Space-Time (BLAST) system in [36] is a typical spatial multiplexing transmission model whose maximum multiplexing gain is equal to $\min\{M, N\}$.

2.2 Cooperative MIMO Systems

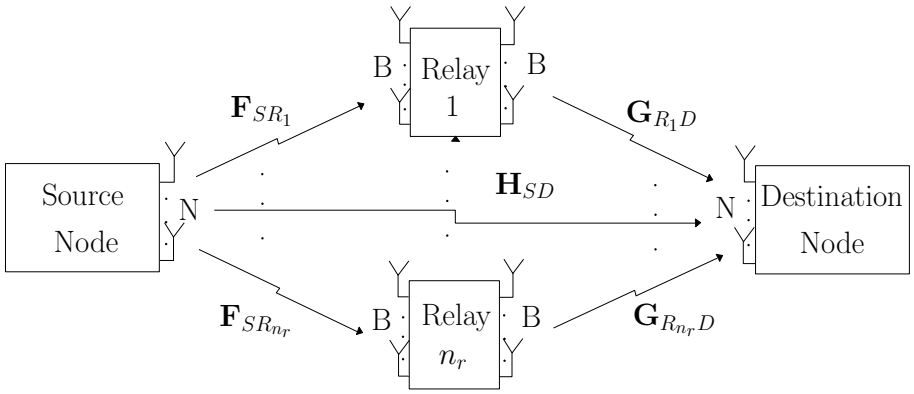

Figure 2.2: Cooperative MIMO system model with n_r multiple-antenna relay nodes

Cooperative communications employ multiple transmission phases and relays in half-duplex mode in order to improve the coverage and the performance of wireless links [27] - [31]. Different relaying protocols, such as AF and DF, are employed in cooperative systems in order to achieve different needs. A cooperative MIMO system model is shown in Fig. 2.2 which consists of one source node, n_r relay nodes (Relay 1, Relay 2, ..., Relay n_r), and one destination node. All the nodes employ N antennas and can either transmit or receive at one time.

The initial idea of using antenna arrays to assist the communication between terminals is due to the long distance associated with the data transmission and the effects of the

noise [27]. With the study of cooperative communications, traditional MIMO and SISO networks which employ mobile or fixed relay nodes have been studied due to the high diversity and low cost. Communication protocols employed in cooperative networks are defined according to different implementation requirements, such as AF protocol focuses on low energy cost which simply amplifies what the relays receive and then forwards the data to the destination, and DF focuses on a different strategy which allows the relays to detect and re-encode the information symbols with powerful coding schemes.

More complex cooperative systems have been reported in the last few years which combine and change the original protocols. In [32] and [33], STC schemes are used at both the source node and the relay nodes in order to achieve novel relay selection algorithms without loss of the diversity gain. In [34] STC schemes are employed in a cooperative system with the AF protocol in order to achieve the optimal diversity and multiplexing tradeoff. The flexibility of using different cooperative protocols and powerful STC schemes indicates the reasons for the popularity of cooperative systems.

We take a two-hop transmission with the AF protocol as an example to show how a cooperative system works. The first hop is devoted to the source transmission, which broadcasts the information symbols to the relay nodes and to the destination node. The second hop forwards the amplified received information symbols at the relay nodes to the destination node. Mathematically, the received symbols at the relays and the destination node in the first hop are given by

$$\boldsymbol{r}_{SD}[i] = \boldsymbol{H}_{SD}[i]\boldsymbol{A}_{S}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SD}[i], \tag{2.11}$$

$$\begin{aligned} \boldsymbol{r}_{SR_k}[i] = \boldsymbol{F}_{SR_k}[i]\boldsymbol{A}_{S}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SR_k}[i], \\ \text{for } k = 1, 2, ..., n_r, \end{aligned} \tag{2.12}$$

where $\boldsymbol{A}_{S}[i] = \mathrm{diag}[\mathrm{a_{S_1}[i]}, \mathrm{a_{S_2}[i]}, ..., \mathrm{a_{S_N}[i]}]$ denotes the diagonal $N \times N$ power allocation matrix assigned to the source node. According to the AF protocol, the received symbols at the relays are amplified prior to transmission to the destination node in the second hop,

which is given by

$$\begin{aligned}\boldsymbol{r}_{RD}[i] &= \sum_{k=1}^{n_r} \boldsymbol{G}_{R_kD}[i]\boldsymbol{A}_k[i]\boldsymbol{r}_{SR_k}[i] + \boldsymbol{n}[i] \\ &= \sum_{k=1}^{n_r} \boldsymbol{G}_{R_kD}[i]\boldsymbol{A}_k[i]\boldsymbol{F}_{SR_k}[i]\boldsymbol{A}_S[i]\boldsymbol{s}[i] + \sum_{k=1}^{n_r} \boldsymbol{G}_{R_kD}[i]\boldsymbol{A}_k[i]\boldsymbol{n}_{SR_k}[i] + \boldsymbol{n}[i],\end{aligned} \tag{2.13}$$

where $\boldsymbol{A}_k[i]$ denotes the diagonal $N \times N$ power allocation matrix assigned for the kth relay node. It is clear to see that the noise at the relay nodes are amplified and forwarded to the destination which contributes to the increased of detection errors.

2.3 Space-Time Coding

Space-time coding (STC) schemes are usually used in MIMO wireless communication systems to defeat the multipath fading and decrease the probability of error. Different STC schemes are designed for MIMO wireless systems due to different needs with different design criteria [41]. During the last 15 years, many STC schemes have been studied and implemented in the real world. In this thesis, some of them are considered and optimized by the proposed algorithms. The orthogonal space-time block coding (OSTBC) designs with low-cost decoding algorithms are the main subject of this thesis, whereas linear dispersion codes (LDCs) with high throughput that require high-computational-complexity detection methods are also considered as references and some comparisons.

The study of STC techniques starts with the introduction of multiple antennas. In 1996, the first space-time architecture was investigated by Foschini and named vertical Bell Laboratories Layered Space-Time (V-BLAST) scheme [36]. The symbols are assigned to multiple antennas first and then transmitted to the receiver. An orthogonal design for 2×2 MIMO systems was developed in 1998. The coding scheme is known as the Alamouti STBC scheme [21] which achieves the full rate and full diversity order. The design criterion is quite simple which ensures the inner product of the 2×2 code matrix equals zero and obtain a linear ML decoding scheme with low computational complexity. The orthogonal STBC designs for different number of antennas are then presented in [40]. The authors define the criterion and extend the work of Alamouti from the 2×2

MIMO to any MIMO systems. The general OSTBCs achieve full diversity order and low complexity encoding and decoding process as the same as the Alamouti code; however, the design criterion keeps the redundancy of the code high which means the full encoding rate and linear ML decoding cannot be achieved simultaneously for a MIMO system with the number of antennas greater than 2. The design of quasi-OSTBC (QOSTBC) codes in [41] solved the problem of improving the encoding rate for systems with more than 2 antennas. However, the full diversity achievement is another issue to be addressed. In order to achieve a full rate and full diversity STBC with efficient decoding algorithm, Jafarkhani designed a quasi-OSTBC (QOSTBC) in [41]. In the QOSTBC schemes, the rotation of the constellation is introduced in order to achieve quasi-orthogonality between the columns of the code matrix and which leads to joint decoding instead of separately decoding for OSTBCs. The full rate and full diversity oder are achieved with the cost of increasing the decoding complexity in a QOSTBC. As mentioned in the previous section, the diversity and spatial multiplexing trade-off is critical to code design. The OSTBCs and QOSTBCs focus on the full diversity order achievement, while the linear dispersion codes (LDCs) designed in [42], [43] and [44] focus on the throughput of a MIMO system. The linear combination of the real part and the imaginary part of the modulated symbols provides a high throughput in the cost of high computational complexity decoding process. Number theory and random matrix theory are introduced in the design of LDCs in order to maintain the high coding rate and achieve full diversity order at the same time, such as the Golden code in [44].

It is important to understand the design criteria for the space-time codes before introduction of different space-time codes. In distributed and cooperative networks, the antennas at the relays can be seen as one antenna array if they are i.i.d. which allows the traditional STC schemes to be adapted to DSTC schemes without changing the design criterion or affecting the properties of a coding scheme.

2.3.1 Design Criteria for Space-Time Codes

The first criterion of STC design is related to the rank and determinant criteria, which describes the coding gain and the diversity gain, the most important properties of a space-time code. We first consider an $N \times M$ MIMO system with AWGN interference employ-

ing $N \times T$ STC scheme $\boldsymbol{C}$, where T stands for the number of time slots in the code matrix design $\boldsymbol{C}$, and the expression of the system model is given by

$$\boldsymbol{R} = \boldsymbol{H}\boldsymbol{C} + \boldsymbol{N}, \tag{2.14}$$

where $\boldsymbol{R}$ denotes the $M \times T$ received matrix at the destination, the $\boldsymbol{H}$ stands for the $M \times N$ channel matrix, and $\boldsymbol{C}$ is the $N \times T$ STC matrix. We assume $N \leq T$. The AWGN matrix contains circularly symmetric complex Gaussian random variables with zero mean and variance N_0. If the codebook of the STC scheme contains K codewords, the upper bound of the error probability when the code matrix $\boldsymbol{C}_1$ is sent but others are received is given by

$$P_e(\boldsymbol{C}_1 \text{ is sent}) \leq \sum_{\mathrm{k}=2}^{\mathrm{K}} \mathrm{P}(\boldsymbol{C}_1 \rightarrow \boldsymbol{C}_{\mathrm{k}}). \tag{2.15}$$

For simplicity we assume the transmitted matrix is $\boldsymbol{C}_1$ and the received matrix $\boldsymbol{R}_1 = \boldsymbol{H}\boldsymbol{C}_1 + \boldsymbol{N}$, the pairwise error probability is calculated by

$$\begin{aligned} P(\boldsymbol{C}_1 \rightarrow \boldsymbol{C}_2 \mid \boldsymbol{H}) &= P(\parallel \boldsymbol{R}_1 - \boldsymbol{H}\boldsymbol{C}_1 \parallel_F^2 - \parallel \boldsymbol{R}_1 - \boldsymbol{H}\boldsymbol{C}_2 \parallel_F^2 > 0 \mid \boldsymbol{H}) \\ &= P(\boldsymbol{X} > \parallel \boldsymbol{H}(\boldsymbol{C}_2 - \boldsymbol{C}_1) \parallel_F^2 \mid \boldsymbol{H}), \end{aligned} \tag{2.16}$$

where $\boldsymbol{X} = Tr(\boldsymbol{N}^H\boldsymbol{H}(\boldsymbol{C}_2 - \boldsymbol{C}_1) + (\boldsymbol{C}_2 - \boldsymbol{C}_1)^H\boldsymbol{H}^H\boldsymbol{N})$ is a zero mean Gaussian random variable with variance $2N_0 \parallel \boldsymbol{H}(\boldsymbol{C}_2 - \boldsymbol{C}_1) \parallel_F^2$. By using the Q function, we can obtain

$$P(\boldsymbol{C}_1 \rightarrow \boldsymbol{C}_2 \mid \boldsymbol{H}) = Q\left(\sqrt{\frac{1}{2N_0}} \parallel (\boldsymbol{C}_2 - \boldsymbol{C}_1)\boldsymbol{H} \parallel_F\right), \tag{2.17}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty exp(\frac{-y^2}{2})dy. \tag{2.18}$$

Define a matrix $\boldsymbol{A}(\boldsymbol{C}_1, \boldsymbol{C}_2) = \boldsymbol{D}(\boldsymbol{C}_1, \boldsymbol{C}_2)^H\boldsymbol{D}(\boldsymbol{C}_1, \boldsymbol{C}_2) = (\boldsymbol{C}_2 - \boldsymbol{C}_1)^H(\boldsymbol{C}_2 - \boldsymbol{C}_1)$, and the nonnegative eigenvalues of $\boldsymbol{A}(\boldsymbol{C}_1, \boldsymbol{C}_2)$ are λ_n, $n = 1, 2, ..., N$. By making use of the SVD theorem [45], we can rewrite (2.16) as

$$\begin{aligned} P(\boldsymbol{C}_1 \rightarrow \boldsymbol{C}_2 \mid \boldsymbol{H}) &= Q\left(\sqrt{\frac{1}{2N_0}\sum_{m=1}^{M}\sum_{n=1}^{N}\lambda_n \mid \beta_{n,m} \mid^2}\right) \\ &\leq \frac{1}{2}exp\left(-\frac{1}{4N_0}\sum_{m=1}^{M}\sum_{n=1}^{N}\lambda_n \mid \beta_{n,m} \mid^2\right), \end{aligned} \tag{2.19}$$

where $\beta_{n,m}$ is the $n-th$ row, $m-th$ column element in matrix $\boldsymbol{VH}$, and $\boldsymbol{V}$ is the unitary matrix generated during the singular value decomposition of $\boldsymbol{A}(\boldsymbol{C}_1, \boldsymbol{C}_2)$. Because $\boldsymbol{VH}$ is Gaussian, the parameters $\mid \beta_{n,m} \mid$ should be Rayleigh and their probability density function is $f(\mid \beta_{n,m} \mid) = 2 \mid \beta_{n,m} \mid exp(- \mid \beta_{n,m} \mid^2)$ [17]. As a result, the upper bound of the pairwise error probability can be written as

$$P(\boldsymbol{C}_1 \rightarrow \boldsymbol{C}_2 \mid \boldsymbol{H}) \leq \frac{1}{\prod_{n=1}^{N}[1 + (\gamma\lambda_n/4)]^M}, \tag{2.20}$$

where $\gamma = 1/N_0$ is the SNR of the system. In high SNR situation, we can ignore the 1 in (2.20) and obtain the expression of the pairwise error probability of the codeword $\boldsymbol{C}$ as

$$P(\boldsymbol{C}_1 \rightarrow \boldsymbol{C}_2 \mid \boldsymbol{H}) \leq \frac{4^{rM}}{(\prod_{n=1}^{r}\lambda_n)^M \gamma^{rM}}, \tag{2.21}$$

where r denotes the rank of the distance matrix $\boldsymbol{A}(\boldsymbol{C}_1, \boldsymbol{C}_2)$. The diversity gain of the space-time code $\boldsymbol{C}$ is equal to rM, the product of the rank of the distance matrix and the number of receive antennas, and a good coding design can be obtained by maximizing the minimum determinant of $\boldsymbol{A}(\boldsymbol{C}_1, \boldsymbol{C}_2)$ in order to achieve larger coding gain.

Another important criterion for STC scheme design is the maximum mutual information criterion. The $N \times T$ transmitted coding matrix $\boldsymbol{C}$ can be derived as $\boldsymbol{C} = \sum_{i=1}^{n} \boldsymbol{A}_i s_i + \boldsymbol{B}_i s_i^*$, and the $N_t \times T$ matrix $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ denote the encoding matrix for the $i-th$ information symbol s_i and its conjugate form s_i^*, respectively. As a result, the expression of a MIMO system can be written as

$$\boldsymbol{r} = \boldsymbol{G}\boldsymbol{s} + \boldsymbol{n}, \tag{2.22}$$

where the $MT \times 1$ vector $\boldsymbol{r}$ is the transformation vector from $\boldsymbol{R}$ in (2.14), and the $MT \times N$ matrix $\boldsymbol{G}$ denotes the equivalent channel matrix which combined the channel matrix $\boldsymbol{H}$ in (2.14) with the encoding matrices $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$. For example, the Alamouti 2×2 space-time block code (STBC) in [22], which have a coding matrix given by

$$\boldsymbol{C} = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix}, \tag{2.23}$$

and the equivalent received symbol vector at the receiver of an Alamouti STBC can be

derived as

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{G}\boldsymbol{s} + \boldsymbol{n} \\ &= \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \\ h_{1,2}^* & -h_{1,1}^* \\ h_{2,2}^* & -h_{1,2}^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_{1,1} \\ n_{2,1} \\ n_{1,2}^* \\ n_{2,2}^* \end{bmatrix}. \end{aligned} \tag{2.24}$$

The maximum mutual information criterion can be described by

$$C(\rho, N, M) = \max E\left[\log\det(\boldsymbol{I}_M + \frac{\rho}{N}\boldsymbol{G}\boldsymbol{G}^H)\right]. \tag{2.25}$$

The equivalent channel matrix $\boldsymbol{G}$ can be found by maximizing the channel capacity, and after collecting the encoding parameters in $\boldsymbol{G}$ we can obtain the STC encoding scheme.

2.3.2 Alamouti Space-Time Block Code

The Alamouti 2×2 STBC [22] is the simplest and the most important STC design as various space-time codes are designed based on this scheme. The information bits are first modulated by mapping to a constellation and we have the modulated symbol vector $\boldsymbol{s} = [s_1, s_2, ..., s_n]$. Then a block of modulated symbols s_1 and s_2 are encoded with the encoding matrix and we obtain the STBC scheme shown as

$$\boldsymbol{C} = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix}. \tag{2.26}$$

The symbols in the first column of the coding matrix will be sent out from the transmitter in the first time slot, and the symbols in the second column will be sent out in the next time slot. If the coding matrix is employed in a 2×2 MIMO system, the received matrix

at the destination can be written as

$$\begin{aligned}\boldsymbol{R} &= \boldsymbol{HC} + \boldsymbol{N} \\ &= \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{bmatrix} \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} + \begin{bmatrix} n_{1,1} & n_{1,2} \\ n_{2,1} & n_{2,2} \end{bmatrix} \\ &= \begin{bmatrix} h_{1,1}s_1 + h_{1,2}s_2 + n_{1,1} & h_{1,2}s_1^* - h_{1,1}s_2^* + n_{1,2} \\ h_{2,1}s_1 + h_{2,2}s_2 + n_{2,1} & h_{2,2}s_1^* - h_{2,1}s_2^* + n_{2,2} \end{bmatrix} \\ &= \begin{bmatrix} r_{1,1} & r_{1,2} \\ r_{2,1} & r_{2,2} \end{bmatrix}.\end{aligned} \tag{2.27}$$

The orthogonal property of the Alamouti STBC scheme is illustrated by

$$\boldsymbol{c}_1^H \boldsymbol{c}_2 = \begin{bmatrix} s_1^* & s_2^* \end{bmatrix} \begin{bmatrix} -s_2^* \\ s_1^* \end{bmatrix} = -s_1^* s_2^* + s_2^* s_1^* = 0. \tag{2.28}$$

The product of two Alamouti code matrices is given by

$$\begin{aligned}\boldsymbol{C}^H\boldsymbol{C} &= \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix}^H \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} \\ &= \begin{bmatrix} s_1^* & s_2^* \\ -s_2 & s_1 \end{bmatrix} \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} \\ &= \begin{bmatrix} \mid s_1 \mid^2 + \mid s_2 \mid^2 & -s_1^* s_2^* + s_1^* s_2^* \\ -s_1 s_2 + s_1 s_2 & \mid s_1 \mid^2 + \mid s_2 \mid^2 \end{bmatrix} \\ &= \begin{bmatrix} \mid s_1 \mid^2 + \mid s_2 \mid^2 & 0 \\ 0 & \mid s_1 \mid^2 + \mid s_2 \mid^2 \end{bmatrix}.\end{aligned} \tag{2.29}$$

The inner product of the columns of the Alamouti scheme is equal to zero as shown in (2.28) so that at the receiver, a maximum-likelihood (ML) decoding technique with low computational complexity can be utilized. By assuming the channel information is

perfectly known at the receiver, the decoding algorithm is given by

$$\begin{aligned}\tilde{s}_1 &= h_{1,1}^* r_{1,1} + h_{1,2} r_{1,2}^* + h_{2,1}^* r_{2,1} + h_{2,2} r_{2,2}^* \\ &= (\mid h_{1,1} \mid^2 + \mid h_{1,2} \mid^2 + \mid h_{2,1} \mid^2 + \mid h_{2,2} \mid^2) s_1 + n_1, \\ \tilde{s}_2 &= h_{1,2}^* r_{1,1} - h_{1,1} r_{1,2}^* + h_{2,2}^* r_{2,1} - h_{2,1} r_{2,2}^* \\ &= (\mid h_{1,1} \mid^2 + \mid h_{1,2} \mid^2 + \mid h_{2,1} \mid^2 + \mid h_{2,2} \mid^2) s_2 + n_2,\end{aligned} \tag{2.30}$$

where $n_1 = h_{1,1}^* n_{1,1} + h_{1,2} n_{1,2}^* + h_{2,1}^* n_{2,1} + h_{2,2} n_{2,2}^*$ and $n_2 = h_{1,2}^* n_{1,1} + h_{1,1} n_{1,2}^* + h_{2,2}^* n_{2,1} + h_{2,1} n_{2,2}^*$ can be considered as noise elements. After testing all the combinations of s_1 and s_2 for ML decision, the most likely symbols are selected.

The Alamouti STBC can reach the full diversity because the codeword distance matrix is given by

$$\boldsymbol{A}(\boldsymbol{C}_1, \boldsymbol{C}_2) = \begin{bmatrix} \mid s_1^{(1)} - s_1^{(2)} \mid^2 + \mid s_2^{(1)} - s_2^{(2)} \mid^2 & 0 \\ 0 & \mid s_1^{(1)} - s_1^{(2)} \mid^2 + \mid s_2^{(1)} - s_2^{(2)} \mid^2 \end{bmatrix}, \tag{2.31}$$

and the determinant of $\boldsymbol{A}(\boldsymbol{C}_1, \boldsymbol{C}_2)$ is equal to $(\mid s_1^{(1)} - s_1^{(2)} \mid^2 + \mid s_2^{(1)} - s_2^{(2)} \mid^2)^2$ which is nonzero as $(s_1^{(1)}, s_2^{(1)}) \neq (s_1^{(2)}, s_2^{(2)})$. By using the SVD theorem [45] we can obtain that the eigenvalue of $\boldsymbol{A}(\boldsymbol{C}_1, \boldsymbol{C}_2)$ equals 1, which indicates that the coding gain of the Alamouti STBC is 1 [22].

2.3.3 Orthogonal Space-Time Block Codes

From the 2×2 Alamouti STBC scheme, the construction of complex OSTBC is developed in [41], which can be considered as the combination of a real OSTBC generator matrix and its conjugate form. The key property of OSTBC is the inner product between columns of the coding matrix equals to zero, which means the $N \times T$ coding matrix of OSTBC $\boldsymbol{G}$ should have the following property

$$\boldsymbol{G}^H \boldsymbol{G} = \kappa(\mid s_1 \mid^2 + \mid s_2 \mid^2 + ... + \mid s_K \mid^2) \boldsymbol{I}_T, \tag{2.32}$$

where $s_k, k = 1, 2, ..., K$ denote the indeterminate variables in $\boldsymbol{G}$ and κ is a constant. In order to reach this property, the real orthogonal generator matrix should be found.

The study of real orthogonal designs in [47] and [48] provides a method of generating an orthogonal matrix. First, find a set of L Hurwitz-Radon family matrices which are $N \times N$ real matrices $\{\boldsymbol{B}_1, \boldsymbol{B}_2, ..., \boldsymbol{B}_L\}$ satisfying

$$\begin{aligned} \boldsymbol{B}_l^T \boldsymbol{B}_l &= \boldsymbol{B}_N, \quad l = 1, 2, ..., L \\ \boldsymbol{B}_l^T &= -\boldsymbol{B}_l, \quad l = 1, 2, ..., L \\ \boldsymbol{B}_l \boldsymbol{B}_{l'} &= -\boldsymbol{B}_{l'} \boldsymbol{B}_l, \quad 1 \leq l < l' \leq L. \end{aligned} \tag{2.33}$$

Then the $N \times N$ real orthogonal design can be generated by the following algorithm

$$\boldsymbol{G}_N = s_1 \boldsymbol{I}_N + \sum_{n=2}^{N} s_n \boldsymbol{B}_n, \tag{2.34}$$

where $s_n, n = 1, 2, ...N$ are information symbols. The complex OSTBC scheme is obtained by combining the real design with its conjugate transformation which is given by

$$\boldsymbol{G}_{complex} = \begin{bmatrix} \boldsymbol{G}_N \\ \boldsymbol{G}_N^* \end{bmatrix}. \tag{2.35}$$

OSTBC can achieve the full-diversity, i.e. NM, for any antenna number of M and N, and a simple linear ML decoding algorithm similar to that of the Alamouti STBC can be used at the receiver. However, the redundancy of the OSTBC matrix is high when N is greater than 2, which reduces the throughput of a MIMO system significantly. Some OSTBC designs presented in [37] - [40] focused on increasing the coding rate of the QSTBC and improving the throughput with the maintenance of the orthogonal scheme. The highest rate of the high-rate OSTBCs is $\frac{3}{4}$. Thus, the quasi-orthogonal STBC (QOSTBC) was invented in order to address the redundancy of the code and improve the throughput of the system.

In Figure 2.3, the simulation results of different coding schemes are shown. As we can observe, the single-antenna system obtains the worst BER performance compared to others. By using 2 antennas at the transmitter and the receiver, the BER performance improved dramatically, especially when using the Alamouti STBC scheme. The V-BLAST scheme contains the multiplexing gain rather than the diversity gain; while, the advantage of the STC is the code schemes have the diversity gain and allows the BER curve achieves a steeper slope compare to the V-BLAST scheme. The higher diversity order a STC can

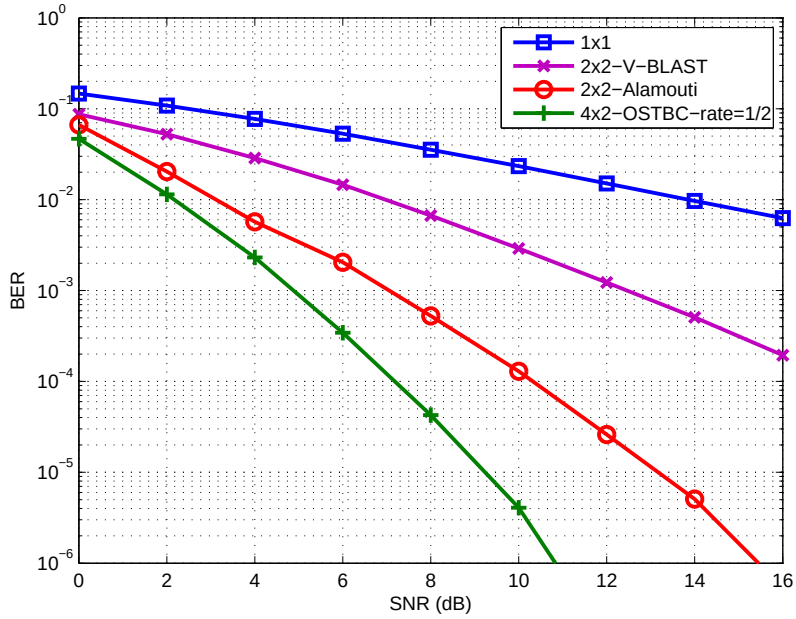

Figure 2.3: SNR versus BER for STBCs

achieve, a steeper slope of the BER curve will achieve. By increasing the number of transmitting antennas, more powerful STBC schemes can be used and achieve higher diversity gain according to the result in Figure 2.3.

2.3.4 Quasi-Orthogonal Space-Time Block Codes

QOSTBC have an orthogonal property between some of its columns but not all of its columns. One example of a QOSTBC matrix in [41] is shown by

$$\boldsymbol{G}_{QOSTBC} = \begin{bmatrix} \boldsymbol{G}(s_1, s_2) & \boldsymbol{G}(s_3, s_4) \\ -\boldsymbol{G}^H(s_3, s_4) & \boldsymbol{G}^*(s_1, s_2) \end{bmatrix}, \tag{2.36}$$

where $\boldsymbol{G}(s_1, s_2)$ is a 2×2 Alamouti STBC scheme. The coding rate is 1 as 4 symbols are transmitted via 4 time slots; however, due to the non-orthogonality between some columns, a separate ML decoding scheme cannot be employed at the receiver to detect the codeword. The solution of this problem lies in the pairwise decoding algorithm. The orthogonality of the subspaces in $\boldsymbol{G}_{QOSTBC}$ ensures the codewords can be decoded in pairs independently [41]. The ML decoding problem for the system described in the

previous section can be derived as the following

$$\min_{s_1,s_2,s_3,s_4} \left(\boldsymbol{G}^H \boldsymbol{H}^H \boldsymbol{H} \boldsymbol{G} - \boldsymbol{R}^H \boldsymbol{H} \boldsymbol{G} - \boldsymbol{G}^H \boldsymbol{H}^H \boldsymbol{R} \right), \tag{2.37}$$

where $\boldsymbol{G}$ denotes the QOSTBC matrix. The ML decoding of $\boldsymbol{G}$ in (2.37) is equivalent to minimizing the sum of 2 independent formulas containing s_1 and s_4, s_2 and s_3, respectively [41].

Another problem of the QOSTBC scheme is the loss of diversity order due to the quasi-orthogonal design. The rank of the distance matrix of encoding scheme $\boldsymbol{G}$ in (2.36) is only 2, as a result the maximum diversity it can achieve is equal to $2M$, where M is the number of receive antennas. The solution reported in the literature for achieving the full-diversity for a QOSTBC is using a rotated constellation for different codewords [49] [50]. The determinant of the distance matrix of $\boldsymbol{G}$ is given by [41]

$$\det \boldsymbol{A} = \left(\sum_{k=1}^{2} \mid (s_k - s'_k) + (s_{k+2} - s'_{k+2}) \mid^2 \right)^2 \left(\sum_{k=1}^{2} \mid (s_k - s'_k) - (s_{k+2} - s'_{k+2}) \mid^2 \right)^2, \tag{2.38}$$

where $s_k, k = 1, 2, 3, 4$ is the information symbols in $\boldsymbol{G}$. If we collect the codeword s_1 and s_2 in one constellation and collect s_3 and s_4 in another, the full-diversity can be reached if and only if $s_k - s'_k \neq s_{k+2} - s'_{k+2}$ [41]. The method of finding the optimal rotation of a constellation is shown in [51], and it has been proved that the optimal rotation for a constellation in order to achieve a full-diversity QOSTBC should follow the theorem

$$\det\{\boldsymbol{A}\}_{min}(\phi) \leq \mid 2 \sin \phi \mid^4 d_{min}^8, \tag{2.39}$$

where ϕ is the optimal rotation angle and d_{min} denotes the minimum Euclidean distance of the constellation which is equal to $\mid s - s' \mid$, the distance between the closest two points in the constellation. The simulation results of the OSTBC and QOSTBC schemes are shown in Figure 2.4. In this figure, 4 antennas are employed at the transmitter and 1 at the receiver. The benefit of using the QOSTBC at the low SNR can be observed from the figure; however, when the SNR is increased to 20dB, the performance of the QOSTBC using non-rotated constellation is getting worse, and the QOSTBC employing rotated constellation achieves full diversity and low BER performance.

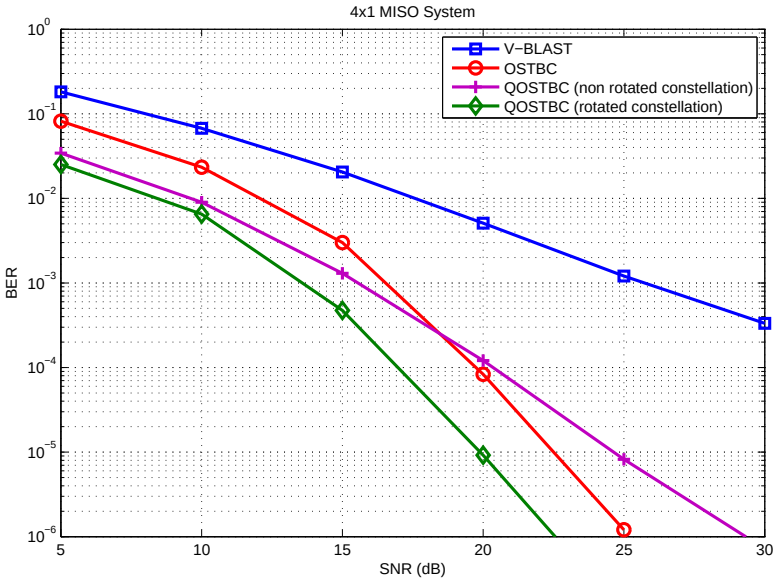

Figure 2.4: SNR versus BER for QOSTBCs in systems with 4 antennas

2.3.5 Linear Dispersion Codes

The 2×2 Alamouti orthogonal design and the quasi-orthogonal design with rotated constellation can achieve the full-diversity and full coding rate, which means the rank and determinant criteria are achieved. However, the maximum mutual information criterion is difficult for orthogonal and quasi-orthogonal designs as redundancy should be used in the encoding matrix in order to maintain the orthogonality. The LDC designs in [42] - [44] can achieve the maximum MIMO channel capacity without the orthogonal design, and provides the full diversity and high coding gain with the help of algebraic number theory.

The encoding scheme of an LDC can be derived as the sum of the original and conjugate transform of the information symbols multiplied by its coding matrix, which is given by

$$\boldsymbol{C} = \sum_{n=1}^{N} (s_n \boldsymbol{A}_n + s_n^* \boldsymbol{B}_n) = \sum_{n=1}^{N} (\Re[s_n] \boldsymbol{C}_n + j\Im[s_n] \boldsymbol{D}_n), \tag{2.40}$$

where $\boldsymbol{C}_n = \boldsymbol{A}_n + \boldsymbol{B}_n$ and $\boldsymbol{D}_n = \boldsymbol{A}_n - \boldsymbol{B}_n$ are $N \times T$ encoding matrices for the real and imaginary parts, respectively. At the destination node, the received LDC matrix is derived as

$$\boldsymbol{R} = (\boldsymbol{H}^R + j\boldsymbol{H}^I) \left(\sum_{n=1}^{N} (s_n^R (\boldsymbol{C}_n^R + j\boldsymbol{C}_n^I) + j s_n^I (\boldsymbol{D}_n^R + j\boldsymbol{D}_n^I)) \right) + \boldsymbol{N}. \tag{2.41}$$

If we define the transmitted symbol vector $\underline{\boldsymbol{s}} = (s_1^R, s_1^I, ..., s_N^R, s_N^I)^T$ and the received vector $\underline{\boldsymbol{r}} = (r_1^R, r_1^I, ..., r_{MT}^R, r_{MT}^I)^T$, and $\underline{\boldsymbol{N}} = (\boldsymbol{N}_1^R, \boldsymbol{N}_1^I, ..., \boldsymbol{N}_M^R, \boldsymbol{N}_M^I)^T$ is the noise vector, the equivalent channel matrix

$$\underline{\boldsymbol{H}} = \begin{bmatrix} \boldsymbol{E}_1\underline{\boldsymbol{h_1}} & \boldsymbol{F}_1\underline{\boldsymbol{h_1}} & \cdots & \boldsymbol{E}_N\underline{\boldsymbol{h_1}} & \boldsymbol{F}_N\underline{\boldsymbol{h_1}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{E}_1\underline{\boldsymbol{h_M}} & \boldsymbol{F}_1\underline{\boldsymbol{h_M}} & \dots & \boldsymbol{E}_N\underline{\boldsymbol{h_M}} & \boldsymbol{F}_N\underline{\boldsymbol{h_M}} \end{bmatrix}, \tag{2.42}$$

where

$$\begin{aligned} \boldsymbol{E}_n &= \begin{bmatrix} \boldsymbol{C}_n^R & -\boldsymbol{C}_n^I \\ \boldsymbol{C}_n^I & \boldsymbol{C}_n^R \end{bmatrix}, \boldsymbol{F}_n = \begin{bmatrix} -\boldsymbol{D}_n^I & -\boldsymbol{D}_n^R \\ \boldsymbol{D}_n^R & -\boldsymbol{D}_n^I \end{bmatrix}, for\ n = 1, 2, ..., N, \\ \underline{\boldsymbol{h_i}} &= \begin{bmatrix} -\boldsymbol{H}_i^R \\ \boldsymbol{H}_i^I \end{bmatrix}, \text{for } m = 1, 2, ..., M. \end{aligned} \tag{2.43}$$

The expression in (2.41) can be written as

$$\underline{\boldsymbol{r}} = \underline{\boldsymbol{H}}\underline{\boldsymbol{s}} + \underline{\boldsymbol{N}}, \tag{2.44}$$

and the sphere decoding algorithm [52, 53] can be used for detection with lower computational complexity.

According to different design algorithms, various encoding matrices $\boldsymbol{A}_n$ and $\boldsymbol{B}_n$ can be designed. According to [42], if the MIMO system has N transmit antennas and M receive antennas, the number of information symbols in LDC matrix is $L = \min(N, M)T$, where T is the number of time slots for transmitting an LDC matrix. This is because the greater L is, the closer it can approach the capacity; however, the smaller L the system uses, the more coding gain we can obtain [42]. Then the coding design should maximize the capacity expression which is given by

$$C_{LDC}(\rho, T, N, M) = \max_{\boldsymbol{A}_n, \boldsymbol{B}_n, n=1,2,...,L} \frac{1}{2T} E \log\det\left(\boldsymbol{I}_{2MT} + \frac{\rho}{N}\boldsymbol{H}\boldsymbol{H}^H\right), \tag{2.45}$$

where ρ denotes the SNR. Some power constraints for the transmitted symbols should be

satisfied for the encoding matrix, which are

$$\begin{aligned}
&\sum_{n=1}^{L}(Tr\{\boldsymbol{A}_n^H\boldsymbol{A}_n\}+Tr\{\boldsymbol{B}_n^H\boldsymbol{B}_n\})=2TN\\
&Tr\{\boldsymbol{A}_n^H\boldsymbol{A}_n\}=Tr\{\boldsymbol{B}_n^H\boldsymbol{B}_n\}=\frac{TN}{L},n=1,2,...,L\\
&\boldsymbol{A}_n^H\boldsymbol{A}_n=\boldsymbol{B}_n^H\boldsymbol{B}_n=\frac{T}{L}\boldsymbol{I}_N,n=1,2,...,L.
\end{aligned} \tag{2.46}$$

An example of LDC encoding matrix is given in [42], which can achieve the maximum mutual information expression with transmission power constraints. The encoding matrices are designed according to the following algorithm

$$\boldsymbol{A}_{N(k-1)+l}=\boldsymbol{B}_{N(k-1)+l}=\frac{1}{\sqrt{N}}\boldsymbol{D}^{k-1}\boldsymbol{\Pi}^{l-1},k=1,2,...,N,l=1,2,...,N, \tag{2.47}$$

where

$$\boldsymbol{D}=\begin{bmatrix}1&0&\cdots&0\\0&e^{j\frac{2\pi}{N}}&0&\cdots\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&e^{j\frac{2\pi(N-1)}{N}}\end{bmatrix},\Pi=\begin{bmatrix}0&\cdots&0&1\\1&0&\cdots&0\\\vdots&\ddots&\vdots&\vdots\\0&\cdots&1&0\end{bmatrix}. \tag{2.48}$$

This LDC encoding matrix is the simplest and most widely used as it can maximize the channel capacity with a simple design criterion given by (2.47); however, according to the generation of the encoding matrix, the rank and determinant criteria in [41] are not satisfied.

In order to design a full-rate and full-diversity STC, number theory and other optimal constellation rotation techniques are employed to design the encoding matrix such as [42] - [44]. The main idea of the code design is based on the construction of an ideal constellation, or the rotation of a constellation, so that the non-vanishing determinant of the distance matrix can be obtained. For example, the STC scheme in [43] is shown as

$$\boldsymbol{B}_{2,\phi}=\frac{1}{\sqrt{2}}\begin{pmatrix}s_1+\phi s_2&\theta(s_3+\phi s_4)\\\theta(s_3-\phi s_4)&s_1-\phi s_2\end{pmatrix}, \tag{2.49}$$

where $\theta^2=\phi$ and $\phi=e^{i\lambda}$, and λ is a real parameter chosen by

$$\underset{\lambda\in Z}{\operatorname{argmax}}\frac{1}{2}\left(\inf_{\boldsymbol{s}\neq[0,0,0,0]^T\in Z[i]^4}\mid s_1^2-s_3^2\phi-s_2^2\phi^2+s_4^2\phi^3\mid\right), \tag{2.50}$$

where $Z[i]^4$ is the integer field or constellation, such as QAM or PAM constellation, and Z denotes the integer field. Different LDCs using different constellations are compared in Figure 2.5.

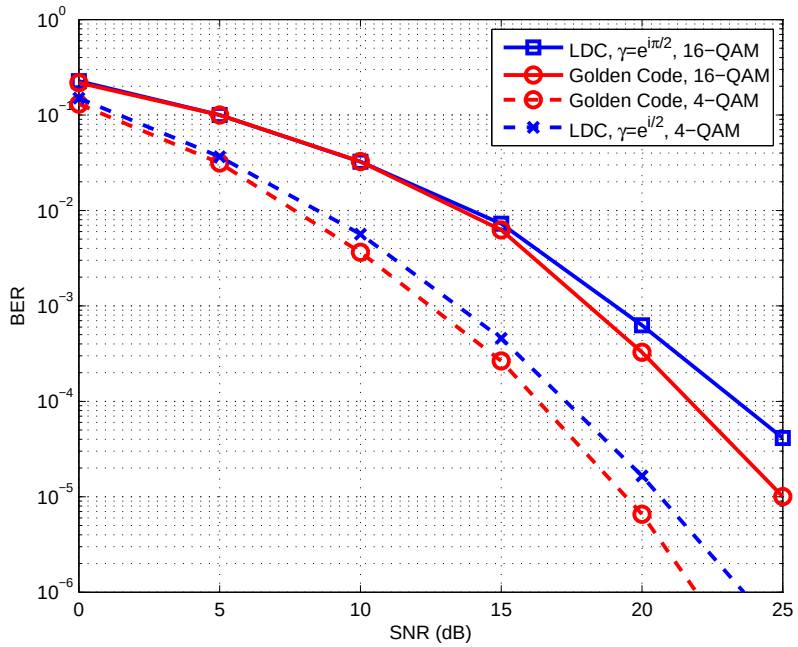

Figure 2.5: SNR versus BER for LDCs

2.4 Parameter Estimation

Parameter estimation is a fundamental element of signal processing which is essential to the implementation of a wireless communication system. Different estimation methods have different performances and computational complexities. As a result, the choice of a proper estimation method is crucial to a wireless communication system.

2.4.1 Maximum Likelihood Estimation

Maximum likelihood (ML) estimation is asymptotically efficient technique which chooses the parameters that maximize the probability density function (PDF) distribution in the observed data so that it often requires high computational complexity in a MIMO system

[54, 55]. The ML estimation procedure is equivalent to finding a parameter vector that maximizes the likelihood function $p(\boldsymbol{x};\boldsymbol{\theta})$, and the $\log$ of the likelihood function is usually used in many estimation scenarios for simplicity. The ML estimation problem is given by

$$\boldsymbol{\theta} = \operatorname{argmax} \ln p(\boldsymbol{x};\boldsymbol{\theta}). \tag{2.51}$$

The ML estimation is a simple estimation method that is implemented by substituting the observed data into (2.51) and taking the data which returns the maximal value. However, when the number of candidates is very large, the high computational complexity restricts the utilization of an ML estimator in practical situations. On the other hand, a closed form solution for the ML estimator will not be always obtained so that different iterative search methods are often required. When multiple parameters need to be estimated, a further increase in the computational complexity limits the implementation of an ML estimator.

2.4.2 Least Squares Estimation

The most significant benefit of using a least-squares (LS) estimator lies on releasing the system from the requirement of additional assumption and statistical information about the received noise [54]. The LS estimation method can converge to the minimum variance unbiased estimator under correct conditions [55, 56].

Consider an LS estimator obtains the received symbols $\boldsymbol{x}$ derived as

$$\boldsymbol{x} = \boldsymbol{H}\boldsymbol{\theta} + \boldsymbol{n}, \tag{2.52}$$

where $\boldsymbol{\theta}$ denotes the desired symbol vector, and $\boldsymbol{n}$ denotes the additive noise. The cost function of the LS estimator is given by

$$J(\boldsymbol{\theta}) = \|\boldsymbol{x} - \boldsymbol{H}\hat{\boldsymbol{\theta}}\|^2. \tag{2.53}$$

As shown in (2.53), the LS estimator obtains the $\hat{\boldsymbol{\theta}}$ which minimizes the squared difference between the observed signal vector. By expanding the righthand side of (2.53), we

will obtain the solution of the LS estimator derived as

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{H}\boldsymbol{H}^H)^{-1}\boldsymbol{H}^H\boldsymbol{x}. \tag{2.54}$$

The estimated solution $\hat{\boldsymbol{\theta}}$ in (2.54) is obtained without the requirement of additive information or assumption. In addition, if the errors represented by the noise vector $\boldsymbol{n}$ are wide-sense stationary, independent and normally distributed, the LS estimator is equivalent to the ML estimator [54]. Therefore the LS estimator can be widely used in wireless communications due to its simplicity and equivalence to establishing the ML estimator in various scenarios.

2.4.3 Adaptive Linear Estimation

In order to avoid the matrix inversion in MMSE and LS estimation and detection procedures, adaptive linear estimation techniques are used. Moreover, adaptive linear estimation algorithms are useful in time-varying channels because they can learn the statistics of the channels and track the desired parameters in time-varying scenarios. The requirement of the knowledge of the symbol and channel are replaced by simply sending initial training symbols.

In the least mean-square (LMS) algorithm, which is also a stochastic gradient (SG) algorithm, the objective function is the mean-square error (MSE) given by

$$J[i] = E \mid s_k[i] - \boldsymbol{w}_k^H[i]\boldsymbol{r}[i] \mid^2, \tag{2.55}$$

where $s_k[i]$ denotes the kth training symbol, $\boldsymbol{w}_k[i]$ and $\boldsymbol{r}[i]$ stand for the receiver filter vector assigned for the kth symbol and the received vector, respectively. By introducing a step size parameter μ, the LMS algorithm is given by

$$\boldsymbol{w}_k[i+1] = \boldsymbol{w}_k[i] + \mu\nabla J[i], \tag{2.56}$$

where

$$\begin{aligned}\nabla_{\boldsymbol{w}_k^*[i]} J[i] &= \nabla_{\boldsymbol{w}_k^*[i]} \left(E \mid s_k[i] - \boldsymbol{w}_k^H[i]\boldsymbol{r}[i] \mid^2 \right) \\ &= \nabla_{\boldsymbol{w}_k^*[i]} \left(E(s_k[i] - \boldsymbol{w}_k^H[i]\boldsymbol{r}[i])^H (s_k[i] - \boldsymbol{w}_k^H[i]\boldsymbol{r}[i]) \right) \\ &= E(s_k[i] - \boldsymbol{w}_k^H[i]\boldsymbol{r}[i])^H \nabla_{\boldsymbol{w}_k^*[i]} \left(E(s_k[i] - \boldsymbol{w}_k^H[i]\boldsymbol{r}[i]) \right) \\ &= -e_k[i]\boldsymbol{r}[i],\end{aligned} \tag{2.57}$$

where $e_k[i]$ stands for the kth detected error symbol. After we obtain (2.57), the LMS algorithm is written as

$$\boldsymbol{w}_k[i+1] = \boldsymbol{w}_k[i] - \mu e_k[i]\boldsymbol{r}[i]. \tag{2.58}$$

The LMS algorithm uses a stochastic gradient approach to reduce the computational complexity of calculating the receive filter $\boldsymbol{w}_k[i]$ as compared to the cost required by Wiener filter which needs to compute a matrix inversion. However, the low convergence rate problem of LMS due to the nature of stochastic approximation of the gradient is the limitation [54, 59].

In order to achieve a faster convergence rate, the recursive least squares (RLS) algorithm which requires a higher computational cost than the LMS algorithm has been developed [54]. All of the received symbols are considered in the RLS algorithm and a forgetting factor is used to weight the past data. Instead of using (2.58) to calculate $\boldsymbol{w}_k[i]$ the expression of the RLS receiver filter is given by

$$\begin{aligned}\boldsymbol{w}_k[i] &= \boldsymbol{R}[i]^{-1}\boldsymbol{p}[i] \\ &= (\lambda\boldsymbol{R}[i-1] + \lambda^{i-j}\boldsymbol{r}[j]\boldsymbol{r}^H[j])^{-1}(\lambda\boldsymbol{p}[i-1] + \lambda^{i-j}s_k^*[j]\boldsymbol{r}[j]),\end{aligned} \tag{2.59}$$

where λ denotes the forgetting factor. The matrix inversion is still required as shown in (2.59) and in order to avoid its high complexity, the matrix inversion lemma in [59] is introduced to determine $\boldsymbol{w}_k[i]$. The lemma is given by

$$(\boldsymbol{A} + \boldsymbol{BDC})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1} + \boldsymbol{DA}^{-1}\boldsymbol{B})\boldsymbol{DA}^{-1}. \tag{2.60}$$

By using (2.60) the expression of the $\boldsymbol{w}_k[i]$ in (2.59) will be changed. The full RLS algorithm is given in Table 2.1.

Other advanced estimation algorithms have been reported in the literature , such as

Table 2.1: Summary of the RLS Algorithm

1:	Initialize: $\boldsymbol{P}[0] = \delta^{-1}\boldsymbol{I}$,
2:	For each instant of time, i=1, 2, ..., compute
3:	$\boldsymbol{k}[i] = \frac{\lambda^{-1}\boldsymbol{P}[i-1]\boldsymbol{r}[i]}{1+\lambda^{-1}\boldsymbol{r}^H[i]\boldsymbol{P}[i-1]\boldsymbol{r}[i]}$,
4:	$e_k[i] = s_k[i] - \boldsymbol{w}_k^H[i]\boldsymbol{r}[i]$,
4:	$\boldsymbol{P}[i] = \lambda^{-1}\boldsymbol{P}[i-1] - \lambda^{-1}\boldsymbol{k}[i]\boldsymbol{r}^H[i]\boldsymbol{P}[i-1]$,
8:	$\boldsymbol{w}_k[i] = \boldsymbol{w}_k[i-1] + \boldsymbol{k}[i]e_k^H[i]$.

the reduced-rank techniques in [57, 58] which reflects the eigenvalue matrix or the singular value matrix of the desired parameters to a matrix with much smaller size in order to reduce the number of data samples required for the estimation and further reduces the computational complexity and the interference from the noise. In this thesis, the basic linear estimation algorithms are employed in order to achieve joint optimization algorithms of the receive filter and DSTC scheme designs.

2.5 Detection Techniques

Another fundamental element of signal processing is detection. Different detection methods with different performance and computational complexity requirements have been reported in the last decades [55]. In this subsection, the fundamental criteria of detection methods and the detection algorithms considered in the thesis are introduced.

2.5.1 Maximum A Posteriori Probability Detection

Consider an $N \times N$ MIMO system with a single user, in order to implement the ML detector, the receiver requires the knowledge of the channel state information (CSI) between the transmitter and itself, the knowledge of power of the symbol vector at the transmitter and the noise vector at the receiver. The received symbol vector $\boldsymbol{r}$ in a MIMO system with sufficient statistics for detection is derived as

$$\boldsymbol{r} = \sqrt{\frac{P_T}{N}}\boldsymbol{H}\boldsymbol{s} + \boldsymbol{n}, \tag{2.61}$$

where $\boldsymbol{H}$ denotes the $M \times N$ channel information matrix, $\boldsymbol{s}$ denotes the $N \times 1$ information symbol vector and $\boldsymbol{n}$ stands for the $M \times 1$ AWGN vector with zero mean and variance σ_n^2 at the receiver. The power at the transmitter is denoted by P_T.

The maximum a posteriori probability (MAP) detector achieves an optimal detection performance due to the detection of a symbol with the highest a posteriori probability [55]. In the system derived in (2.61), given $\boldsymbol{s}$ and $\boldsymbol{H}$, the joint probability density function of the received symbol vector $\boldsymbol{r}$ can be written as

$$P(\boldsymbol{r}|\boldsymbol{s}, \boldsymbol{H}) = \frac{1}{(\pi\sigma_n^2)^N} \exp\left(-\frac{\|\boldsymbol{r} - \boldsymbol{H}\boldsymbol{s}\|^2}{\sigma_n^2}\right). \tag{2.62}$$

As a result, the detected symbol vector $\hat{\boldsymbol{s}}_{MAP}$ is derived as

$$\hat{\boldsymbol{s}}_{MAP} = \underset{\hat{\boldsymbol{s}} \in \boldsymbol{S}}{\operatorname{argmax}}\, P(\hat{\boldsymbol{s}}|\boldsymbol{r}, \boldsymbol{H}) = \underset{\hat{\boldsymbol{s}} \in \boldsymbol{S}}{\operatorname{argmax}}\, P(\hat{\boldsymbol{r}}|\hat{\boldsymbol{s}}, \boldsymbol{H}) P(\hat{\boldsymbol{s}}), \tag{2.63}$$

where $P(\hat{\boldsymbol{s}})$ denotes the a priori information, and $\boldsymbol{S}$ denotes the candidates symbol matrix which contains all the possible combinations of the information symbols. The size of $\boldsymbol{S}$ depends on the number of antennas N and the modulation scheme used at the transmitter. Specifically, for K-PSK or K-QAM modulation constellation, the number of all possible combinations is equal to 2^{N+K-1}. Although the MAP detector achieves an optimal performance, the high computational complexity increases exponentially with the increase of the number of transmit antennas and the order of modulation scheme which restricts the implementation of the MAP detector in mobile communications. It is worth to mention that if all the transmitted symbols are equally likely, the MAP detector is equivalent to the ML detector.

2.5.2 Maximum Likelihood Detection

The ML detection algorithm obtains the symbol vector which maximizes the likelihood or log-likelihood function which outputs the most-likely transmitted symbol vector, and also it is equivalent to the solution of a minimum noise energy [60]. As a result, if we know the received symbol vector $\boldsymbol{r}$ and we want to find the transmitted symbols $\boldsymbol{s}$, the MAP detection maximizes $p(\boldsymbol{s}|\boldsymbol{r})$, while the ML detection maximizes $p(\boldsymbol{r}|\boldsymbol{s})$. Consider

the MIMO system derived in the previous section, if the channel matrix $\boldsymbol{H}$ is known at the receiver, the ML detection is expressed by

$$\hat{\boldsymbol{s}}_{ML} = \underset{\boldsymbol{s}\in\boldsymbol{S}}{\operatorname{argmax}} \|\boldsymbol{r} - \boldsymbol{H}\boldsymbol{s}\|^2. \tag{2.64}$$

By searching the candidates symbol vector $\boldsymbol{s}$, the solution of the ML problem will be obtained. The ML detection method returns the most likely transmitted symbol vector from the receiver in a MIMO system which is an optimal solution with the lowest error rate in the symbol vector. However, the searching task with extremely high computational complexity is sometimes not practical for real MIMO wireless communication systems. In order to solve this problem, several techniques for finding a reduced set of possible solutions in the candidates matrix $\boldsymbol{S}$ are presented in [52, 53]. One of these techniques is the sphere decoder. This sphere decoding method defines the lattice of potential solutions and then selects the most probable symbol combination. After determining the point on the lattice, the decoder searches the candidate vectors within a predefined-radius sphere so the combinations within the sphere will be tested [52]. The sphere decoder can be seen as an ML decoder with reduced candidates, and it is widely used in the practical MIMO networks.

2.5.3 Linear Detection

Compared to the ML detection and sphere decoder, the sub-optimal linear detection methods require lower computational complexity which have advantages in low battery and less memory length requirement, and are more feasible for practical MIMO systems. The BER performance of a linear detector is always inferior to that of the ML detection; however, the low computational complexity is a significant advantages of implementation a linear receiver offers in MIMO systems.

The most widely used linear detection algorithms are zero forcing (ZF) and minimum mean square error (MMSE) detection. If we consider an $N \times N$ MIMO system described in (2.61), by using the ZF criterion we can obtain the linear ZF filter matrix which is given by

$$\boldsymbol{G}_{ZF} = (\boldsymbol{H}^H\boldsymbol{H})^{-1}\boldsymbol{H}^H. \tag{2.65}$$

It is assumed that the channel matrix $\boldsymbol{H}$ is invertible. After we obtain the ZF filter matrix we can rewrite the filtered received symbol vector as

$$\hat{\boldsymbol{s}}_{ZF} = \boldsymbol{s} + \boldsymbol{G}_{ZF}\boldsymbol{n}. \tag{2.66}$$

As shown in (2.66), the detected symbol vector contains the original information symbol vector plus the filtered noise vector, and we can obtain the error covariance matrix of ZF detection shown as

$$\boldsymbol{\Theta}_{ZF} = E[(\hat{\boldsymbol{s}}_{ZF} - \boldsymbol{s})(\hat{\boldsymbol{s}}_{ZF} - \boldsymbol{s})^H] = \boldsymbol{G}_{ZF}\boldsymbol{G}_{ZF}^H E[\boldsymbol{n}\boldsymbol{n}^H] = \sigma_n^2 \boldsymbol{G}_{ZF}\boldsymbol{G}_{ZF}^H, \tag{2.67}$$

where σ_n^2 stands for the variance of the received noise. It is easy to conclude from (2.67) that the increase of the noise power is due to the ZF filter matrix. Thus, the BER performance of a ZF receiver will be affected and degraded significantly by the increase of energy of the ZF filter matrix [25].

Another linear detection criterion, named MMSE, is designed in order to solve the problem of noise power increase. The noise vector is considered in the MMSE filter design so that the BER performance degradation is reduced. The linear MMSE cost function is given by

$$\boldsymbol{G}_{MMSE} = \underset{\boldsymbol{G}}{\operatorname{argmin}}\, E\left[\|\boldsymbol{G}\boldsymbol{r} - \boldsymbol{s}\|^2\right]. \tag{2.68}$$

The optimal solution of $\boldsymbol{G}_{MMSE}$ makes (2.68) equal to zero, so that we can obtain the linear MMSE filter matrix given by

$$\boldsymbol{G}_{MMSE} = (\boldsymbol{H}^H\boldsymbol{H} + \sigma_n^2\boldsymbol{I}_N)^{-1}\boldsymbol{H}^H. \tag{2.69}$$

By comparing (2.65) with (2.69) it is shown that the difference in ZF and MMSE lies in the noise elements $\sigma_n^2\boldsymbol{I}_N$, and the linear MMSE detector is equivalent to the ZF detector if we extend the channel matrix $\boldsymbol{H}$ to $\bar{\boldsymbol{H}}$ [61], where

$$\bar{\boldsymbol{H}} = \begin{bmatrix} \boldsymbol{H} \\ \sigma_n^2\boldsymbol{I}_N \end{bmatrix}. \tag{2.70}$$

The extended channel matrix $\bar{\boldsymbol{H}}$ contains the information of the noise so that it determines the noise application which leads to an enhanced reliability and BER performance. Note

that the MMSE filter is also equivalent to the ZF filter when the SNR tends to infinity.

The LS criterion can be used as a parameter estimation method which was introduced in the previous section, and it can be also employed as a detection algorithm. Instead of the requirement of the statistical assumption about the input data, the LS based receive filter introduces a dependency on the number of data samples. The cost function of the LS receiver filter is written as

$$J(\boldsymbol{w}_k[i]) = \sum_{j=1}^{i} \lambda^{i-j} \|s_k[j] - \boldsymbol{w}_k^H[j]\boldsymbol{r}[j]\|^2, \text{ for } k = 1, 2, ..., N, \tag{2.71}$$

where λ denotes forgetting factor which weights the received samples. From (2.71) it is shown that the LS based receiver filter uses all the received symbols from the beginning to the ith time slot to calculate the filter vector, and the more recent received symbols more useful according to the forgetting factor [59]. The solution of (2.71) is given by

$$\boldsymbol{w}_k[i] = \left(\sum_{j=1}^{i} \lambda^{i-j} \boldsymbol{r}[j]\boldsymbol{r}^H[j] \right)^{-1} \left(\sum_{j=1}^{i} \lambda^{i-j} s_k^*[j]\boldsymbol{r}[j] \right). \tag{2.72}$$

2.5.4 Successive Interference Cancelation

The successive interference cancelation (SIC) technique is a non-linear detection algorithm with sorted cancelation ordering [36], which achieves a better BER performance compared to the linear detection algorithms at the cost of extra computational complexity. Besides the basic SIC technique introduced in [36], several advanced SIC detection algorithms [66]- [68] have been developed in the recent decade. In [66], adaptive and iterative multi-branch MMSE decision feedback detection algorithms for MIMO systems are designed which achieves performance close to the ML detector. A multiple feedback SIC strategy is proposed for the uplink of multiuser MIMO systems which introduces constellation points as the candidates to achieve enhanced interference cancelation and can be combined with multi-branch technique to achieve higher detection diversity in [67].

The SIC detection can be seen as a trade-off between the linear detection and the ML detection which achieves better BER performance compared to the linear detection and lower computational complexity compared to the ML or sphere detection [62]. The key

idea of the SIC detection is to introduce layer peeling, which is equivalent to removing the first detected symbol before the detection in the next layer peeling and repeating the process until the last symbol is detected in the existing linear detectors to achieve a BER performance improvement. Instead of detecting the symbol stream simultaneously, by utilization of layer peeling technique, the SIC algorithm detects symbol by symbol, and the interference from the previously detected symbol can be completely removed before the detection of the next one if the data symbols are correctly detected.

Consider the transmitted symbol stream $\boldsymbol{s} = [s_1, s_2, ..., s_N]$ in (2.61) which contains N symbols, and the receiver which uses an MMSE filter with SIC algorithm to detect the received symbols. The MMSE filter vector for the kth symbol in $\boldsymbol{s}$ is calculated by

$$\boldsymbol{w}_k = \left(\bar{\boldsymbol{H}}_k \bar{\boldsymbol{H}}_k^H + \frac{\sigma_k^2}{\sigma_s^2} \boldsymbol{I} \right)^{-1} \boldsymbol{h}_k, \tag{2.73}$$

where $\bar{\boldsymbol{H}}_k$ denotes the channel matrix containing the kth to the Nth column in $\boldsymbol{H}$. In the SIC algorithm, the kth symbol s_k is detected by

$$\hat{s_k} = \boldsymbol{w}_k^H \bar{\boldsymbol{r}}_k, \tag{2.74}$$

where

$$\bar{\boldsymbol{r}}_k = \begin{cases} \boldsymbol{r}, & k = 1, \\ \boldsymbol{r} - \sum_{i=1}^{k-1} \boldsymbol{h}_i \hat{s}_i, & k \geq 2. \end{cases} \tag{2.75}$$

By subtracting the detected symbol from the received symbol vector, the remaining symbols are detected without the interference from the detected symbols.

In [63], the order and priority of the symbol to be detected has been studied. Basically, the detected symbol is removed through a feedback loop that can be considered as interference to other symbols and improve the BER performance in orders. If s_n contains the highest energy in the received symbol vector and it is correctly detected and subtracted from the received symbol vector which gives $\hat{s_n} = s_n$, the rest of the symbols are released from the most significant interference. However, if $\hat{s_n} \neq s_n$ which means the incorrect detected symbol is removed from the received symbol vector and this leads to an error burst in the overall BER performance. In order to address this error propagation problem [64, 65], the order of detection needs to be determined carefully which refers to the detection according to the reliability of the received symbol.

The SNR of a symbol can be used to decide the order of detection in a SIC based detector. The SNR of the nth received symbol is derived as

$$SNR_k = \frac{\sigma_s^2}{\sigma_n^2 \|\boldsymbol{w}_k\|}, \tag{2.76}$$

where $\boldsymbol{w}_k$ stands for the filter vector which is calculated using (2.73). By choosing the layer with the highest SNR, the SIC based detector can determine the order of detection. After detection of the most reliable symbol, this strongest symbol and its corresponding channel vector will be removed from the received symbol vector and the channel matrix. After the first detection, the symbol with the second highest SNR will be detected and removed. The process will be repeated until all the symbols are detected.

The received signal-to-interference-plus-noise ratio (SINR) which considered the influence of the interference and the noise can be used to determine the order of a SIC based detector. The SINR of the kth symbol is derived as

$$SINR_k = \frac{\sigma_s^2 |\boldsymbol{w}_k \boldsymbol{h}_k|^2}{\sigma_s^2 \sum_{n \neq k} |\boldsymbol{w}_n \boldsymbol{h}_n|^2 + \sigma_n^2 \|\boldsymbol{w}_k\|}, \tag{2.77}$$

where $\boldsymbol{h}_k$ denotes the kth column of the channel matrix $\boldsymbol{H}$. The symbol with the highest received SINR will be considered to be the most reliable symbol and will be removed from the received symbol vector after being detected. BER performances of different detection algorithms are shown in Fig. 2.6. It is worth to mention that the SD can obtain a performance identical to the ML detector if the radius is properly selected. In the example given in the Fig. 2.6, the radius of the SD was selected in such a way that the ML solution was not always obtained by the SD detector.

The performance of different detection algorithms are shown in Figure 2.6. Single-antenna transmitters and receivers are used. By using SIC techniques, the detection performance of MMSE improves. The detection error probabilities of SD and ML are quite close, and ML achieves the best detection performance, at the cost of the highest computational complexity.

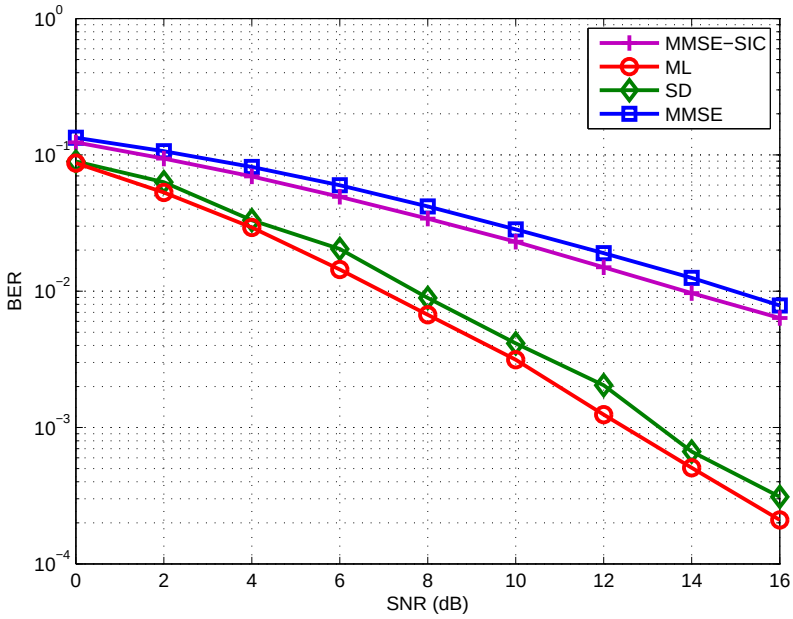

Figure 2.6: SNR versus BER for different detection algorithms

Chapter 3

Adaptive Power Allocation Strategies for DSTC in Cooperative MIMO Networks

Contents

3.1 Introduction

Due to the benefits of cooperative multiple-input and multiple-output (MIMO) systems [8], extensive studies of cooperative MIMO networks have been undertaken [69] - [72]. In [69], an adaptive joint relay selection and power allocation algorithm based on the minimum mean square error (MMSE) criterion is designed. A joint transmit diversity

optimization and relay selection algorithm for the Decode-and-Forward (DF) cooperating strategy [2] is designed in [88]. A transmit diversity selection matrix is introduced at each relay node in order to achieve a better MSE performance by deactivating some relay nodes. A central node which controls the transmission power for each link is employed in [70]. Although the centralized power allocation can improve the performance significantly, the complexity of the calculation increases with the size of the system. The works on the power allocation problem for the DF strategy measuring the outage probability in each relay node with a single antenna and determining the power for each link between the relay nodes and the destination node, have been reported in [12] - [13]. The diversity gain can be improved by using relay nodes with multiple antennas. When the number of relay nodes is the same, the cooperative gain can be improved by using the DF strategy compared with a system employing the AF strategy. However, the interference at the destination will be increased if the relay nodes forward the incorrectly detected symbols in the DF strategy. The power allocation optimization algorithms in [10] and [11] provide improved BER performance at the cost of requiring an eigenvalue decomposition to obtain the key parameters.

In this chapter, we propose joint adaptive power allocation (JAPA) algorithms according to different optimization criteria with a linear receiver or an ML detector for cooperative MIMO systems employing multiple relay nodes with multiple antennas that perform cooperating strategies. This work was first introduced and discussed in [74] and [75]. The power allocation matrices utilized in [74] are full rank and after the optimization, all the parameters are transmitted back to the relay nodes and the source node with an error-free and delay-free feedback channel. In this chapter, we employ the diagonal power allocation matrices in which the parameters stand for the power allocated to each transmit antenna. The requirement of the limited feedback is significantly reduced as compared to the algorithms in the previously reported works. It is worth to mention that the JAPA strategies derived in our algorithms are two-phase optimization techniques, which optimize the power assigned at the source node and at the relay nodes in the first phase and the second phase iteratively, and the proposed JAPA algorithms can be used as a power allocation strategy for the second phase only.

Three optimization criteria, namely, MMSE, minimum bit error rate (MBER) and maximum sum rate (MSR), are employed in the proposed JAPA optimization algorithms in

this chapter. We first develop joint optimization algorithms for the power allocation matrices and the linear receive filter according to these three criteria, which require matrix inversions and bring a high computational burden to the receiver. In the proposed JAPA algorithms with the MMSE, MBER, and MSR criteria, an SG method [59] is employed in order to reduce the computational complexity of the proposed algorithm. A comparison of the computational complexity of the algorithms is considered in this chapter. A normalization procedure is employed by the optimization algorithm in order to enforce the power constraint in both transmission phases. After the normalization, the PA parameters are transmitted back to each transmit node through a feedback channel. The effect of the feedback errors is considered in the analysis and in the simulation sections, where we indicate an increased MSE performance due to feedback inaccuracy.

3.2 Cooperative System Model

Consider a two-hop cooperative MIMO system as in Fig. 2.2 with n_r relay nodes that employ an AF cooperative strategy as well as a DSTC scheme. The source node and the destination node have N antennas to transmit and receive data. An arbitrary number of antennas can be used at the relays which is denoted by B shown in Fig. 2.2. We consider only one user at the source node in our system that operates in a spatial multiplexing configuration. Let $\boldsymbol{s}[i]$ denote the transmitted information symbol vector at the source node which contains N symbols $\boldsymbol{s}[i] = [s_1[i], s_2[i], ..., s_N[i]]$, and has a covariance matrix $E\big[\boldsymbol{s}[i]\boldsymbol{s}^H[i]\big] = \sigma_s^2\boldsymbol{I}_N$, where σ_s^2 is the signal power which we assume to be equal to 1. The source node broadcasts $\boldsymbol{s}[i]$ from the source to n_r relay nodes as well as to the destination node in the first hop, which can be described by

$$\begin{aligned} \boldsymbol{r}_{SD}[i] =& \boldsymbol{H}_{SD}[i]\boldsymbol{A}_S[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SD}[i], \\ \boldsymbol{r}_{SR_k}[i] =& \boldsymbol{F}_{SR_k}[i]\boldsymbol{A}_S[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SR_k}[i], \\ k =& 1, 2, \ ... \ n_r, \ \ i = 1, 2, \ ... \ , \end{aligned} \tag{3.1}$$

where $\boldsymbol{A}_S[i] = \mathrm{diag}[\mathrm{a_{S_1}[i]}, \mathrm{a_{S_2}[i]}, ..., \mathrm{a_{S_N}[i]}]$ denotes the diagonal $N \times N$ power allocation matrix assigned for the source node, and $\boldsymbol{r}_{SR_k}[i]$ and $\boldsymbol{r}_{SD}[i]$ denote the received symbol vectors at the kth relay node and at the destination node, respectively. The $B \times 1$ vec-

tors $\boldsymbol{n}_{SR_k}[i]$ and $\boldsymbol{n}_{SD}[i]$ denote the zero mean complex circular symmetric additive white Gaussian noise (AWGN) vector generated at the kth relay node and at the destination node with variance σ^2. The matrices $\boldsymbol{F}_{SR_k}[i]$ and $\boldsymbol{H}_{SD}[i]$ are the $B \times N$ channel coefficient matrices. It is worth to mention that an orthogonal transmission protocol is considered which requires that the source node does not transmit during the time period of the second hop.

The received symbols are amplified and re-encoded at each relay node prior to transmission to the destination node in the second hop. We assume that the synchronization at each node is perfect. The received vector $\boldsymbol{r}_{SR_k}[i]$ at the kth relay node is assigned a $B \times B$ diagonal power allocation matrix $\boldsymbol{A}_k[i] = \mathrm{diag}[\mathrm{a_{k_1}}, \mathrm{a_{k_2}}, ..., \mathrm{a_{k_B}}]$ which leads to $\tilde{\boldsymbol{s}}_{SR_k}[i] = \boldsymbol{A}_k[i]\boldsymbol{r}_{SR_k}[i]$. The $B \times 1$ signal vector $\tilde{\boldsymbol{s}}_{SR_k}[i]$ will be re-encoded by a $B \times T$ DSTC matrix $\boldsymbol{M}(\tilde{\boldsymbol{s}})$, and then forwarded to the destination node. The relationship between the kth relay and the destination node can be described as

$$\boldsymbol{R}_{R_kD}[i] = \boldsymbol{G}_{R_kD}[i]\boldsymbol{M}_{R_kD}[i] + \boldsymbol{N}_{RD}[i]. \tag{3.2}$$

The $N \times T$ received symbol matrix $\boldsymbol{R}_{R_kD}[i]$ in (3.2) can be written as an $NT \times 1$ vector $\boldsymbol{r}_{R_kD}[i]$ given by

$$\begin{aligned} \boldsymbol{r}_{R_kD}[i] &= \boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] + \boldsymbol{n}_{R_kD}[i] = \boldsymbol{G}_{eq_k}[i]\big(\boldsymbol{A}_k[i]\boldsymbol{r}_{SR_k}[i]\big) + \boldsymbol{n}_{RD}[i] \\ &= \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i]\boldsymbol{F}_{SR_k}[i]\boldsymbol{A}_S[i]\boldsymbol{s}[i] + \boldsymbol{n}_{R_k}[i] + \boldsymbol{n}_{RD}[i], \end{aligned} \tag{3.3}$$

where the $NT \times B$ matrix $\boldsymbol{G}_{eq_k}[i]$ stands for the equivalent channel matrix which is the DSTC scheme $\boldsymbol{M}(\tilde{\boldsymbol{s}}[i])$ combined with the channel matrix $\boldsymbol{G}_{R_kD}[i]$. The second term $\boldsymbol{n}_{R_k}[i] = \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i]\boldsymbol{n}_{SR_k}[i]$ in (3.3) stands for the amplified noise received from the relay node, and the $NT \times 1$ equivalent noise vector $\boldsymbol{n}_{RD}[i]$ generated at the destination node contains the noise parameters in $\boldsymbol{N}_{RD}[i]$.

After rewriting $\boldsymbol{R}_{R_kD}[i]$ we can consider the received symbol vector at the destination node as a $(T+1)N \times 1$ vector with two parts, one is from the source node and another one is the superposition of the received vectors from each relay node. Therefore, we can

write the received symbol at the destination node as

$$\begin{aligned}\boldsymbol{r}[i] &= \begin{bmatrix} \boldsymbol{r}_{SD}[i] \\ \boldsymbol{r}_{RD}[i] \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_{SD}[i]\boldsymbol{A}_S[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SD}[i] \\ \sum_{k=1}^{n_r} \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i]\boldsymbol{F}_{SR_k}[i]\boldsymbol{A}_S[i]\boldsymbol{s}[i] + \boldsymbol{n}_{RD}[i] \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{H}_{eq_{SD}}[i] \\ \sum_{k=1}^{n_r} \boldsymbol{H}_{eq_k}[i] \end{bmatrix} \boldsymbol{s}[i] + \begin{bmatrix} \boldsymbol{n}_{SD}[i] \\ \boldsymbol{n}_{RD}[i] \end{bmatrix} \\ &= \boldsymbol{H}_D[i]\boldsymbol{s}[i] + \boldsymbol{n}_D[i], \end{aligned} \tag{3.4}$$

where the $(T+1)N \times N$ matrix $\boldsymbol{H}_D[i]$ denotes the channel gain matrix with the power allocation of all the links in the system. The $N \times N$ channel matrix $\boldsymbol{H}_{eq_{SD}}[i] = \boldsymbol{H}_{SD}[i]\boldsymbol{A}_S[i]$, while the kth equivalent channel matrix $\boldsymbol{H}_{eq_k}[i] = \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i]\boldsymbol{F}_{SR_k}[i]\boldsymbol{A}_S[i]$. We assume that the coefficients in all channel matrices are statistically independent and remain constant over the transmission and the destination node has the information of all links in each phase. The $(T+1)N \times 1$ noise vector $\boldsymbol{n}_D[i]$ contains the equivalent received noise vector at the destination node, which can be modeled as AWGN with zero mean and covariance matrix $\sigma^2(1+ \parallel \sum_{k=1}^{n_r} \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i] \parallel_F^2)\boldsymbol{I}_{(T+1)N}$. It is worth to mention that the value in B is variable and, in this work, we focus on the power allocation optimization algorithms in cooperative MIMO systems. For simplicity, we consider scenarios in which there are $B = N$ antennas at the relays.

3.3 Adaptive Power Allocation Matrix Optimization Strategies

In this section, we consider the design of a two-phase adjustable power allocation matrix according to various criteria using a DSTC scheme in cooperative MIMO systems. The linear receive filter is determined jointly with the power allocation matrices. A feedback channel is considered in order to convey the information about the power allocation matrix.

3.3.1 Joint Linear MMSE Receiver Design with Power Allocation

The linear MMSE receiver design with power allocation matrices is derived as follows. By defining the $(T+1)N \times 1$ parameter vector $\boldsymbol{w}_j[i]$ to determine the jth symbol $s_j[i]$ in the signal vector $\boldsymbol{s}[i]$, we propose the MSE based optimization with a power constraint described by

$$\begin{aligned}[\boldsymbol{w}_j[i], \boldsymbol{A}_S[i], \boldsymbol{A}_k[i]]] = \underset{\boldsymbol{w}_j[i], \boldsymbol{A}_S[i], \boldsymbol{A}_k[i]}{\operatorname{argmin}} E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right], \\ s.t.\ \mathrm{Tr}(\sum_{\mathrm{k=1}}^{\mathrm{n_r}} \boldsymbol{A}_{\mathrm{k}}[\mathrm{i}]\boldsymbol{A}_{\mathrm{k}}^H[\mathrm{i}]) \leq \mathrm{P_R},\ \mathrm{Tr}(\boldsymbol{A}_{\mathrm{S}}[\mathrm{i}]\boldsymbol{A}_{\mathrm{S}}^H[\mathrm{i}]) \leq \mathrm{P_T},\end{aligned} \tag{3.5}$$

where P_R and P_T denote the transmit power assigned to the relay nodes and to the source node, respectively. In equal power allocation, P_R is equally allocated among all the relay nodes. If the optimal power allocation is considered, the total power P_R will be divided among different relays according to different power allocation algorithms. The proposed optimization algorithms will be implemented at the destination node so that the total power among the relays can be normalized and the power allocated to each relay node will be determined and sent back to each relay after the optimization. The values of the parameters in the power allocation (PA) matrices are restricted by P_T and P_R. By employing the Lagrange multipliers λ_1 and λ_2 we can obtain the Lagrangian function shown as

$$\mathscr{L} = E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right] + \lambda_1(\sum_{j=1}^{N} a_{S_j}[i] - P_T) + \lambda_2(\sum_{k=1}^{n_r}\sum_{j=1}^{N} a_{k_j}[i] - P_R), \tag{3.6}$$

where $a_{S_j}[i]$ denotes the jth parameters in the diagonal of $\boldsymbol{A}_S[i]$ while $a_{k_j}[i]$ stands for the jth parameters in the diagonal of $\boldsymbol{A}_k[i]$.

By expanding the right-hand side of (3.6), taking the gradient with respect to $\boldsymbol{w}_j^*[i]$, $a_{S_j}^*[i]$ and $a_{k_j}^*[i]$, respectively, and equating the terms to zero, we can obtain

$$\begin{aligned}\boldsymbol{w}_j[i] &= \boldsymbol{R}^{-1}\boldsymbol{p}, \\ a_{S_j}[i] &= \tilde{R}_S^{-1}\tilde{P}_S, \\ a_{k_j}[i] &= \tilde{R}^{-1}\tilde{P},\end{aligned} \tag{3.7}$$

where

$$\begin{aligned}
\boldsymbol{R} &= E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right],\ \boldsymbol{p} = E\left[\boldsymbol{r}[i]s_j^*[i]\right], \\
\tilde{R}_S &= E\left[\boldsymbol{w}_j^H[i]\boldsymbol{h}_{SDA_j}[i]s_j[i]s_j^*[i]\boldsymbol{h}_{SDA_j}^H[i]\boldsymbol{w}_j[i] + \lambda_1 a_{S_j}[i]\right], \\
\tilde{P}_S &= E\left[\boldsymbol{h}_{SDA_j}^H[i]\boldsymbol{w}_j[i]s_j^*[i]s_j[i]\right], \\
\tilde{P} &= E\left[s_j[i]s_j^*[i]a_{S_j}^*[i]f_{k_j}^*[i]\boldsymbol{g}_{eq_{k_j}}^H[i]\boldsymbol{w}_j[i]\right], \\
\tilde{R} &= E\left[\boldsymbol{w}_j^H[i]\boldsymbol{g}_{eq_{k_j}}[i]f_{k_j}[i]a_{S_j}[i]s_j[i]s_j^*[i]a_{S_j}^*[i]f_{k_j}^*[i]\boldsymbol{g}_{eq_{k_j}}^H[i]\boldsymbol{w}_j[i] + \lambda_2 a_{k_j}[i]\right].
\end{aligned} \tag{3.8}$$

The vector $\boldsymbol{h}_{SDA_j}$ denotes the channel vector assigned to the parameter a_{S_j} and is the jth column of the equivalent channel matrix $\boldsymbol{H}_{SDA}[i] = \left[\boldsymbol{H}_{SD}[i]; \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i]\boldsymbol{F}_{SR_k}[i]\right]$, and $f_{k_j}[i]$ and $\boldsymbol{g}_{eq_{k_j}}[i]$ denotes the jth parameter in $\boldsymbol{F}_k[i]$ and the jth column in $\boldsymbol{G}_{eq_k}[i]$, respectively. The value of the Lagrange multipliers λ_1 and λ_2 can be determined by substituting $\boldsymbol{A}_S[i]$ and $\boldsymbol{A}_k[i]$ into $\mathrm{Tr}(\boldsymbol{A}_{\mathrm{S}}[\mathrm{i}]\boldsymbol{A}_{\mathrm{S}}^H[\mathrm{i}]) \leq \mathrm{P_T}$ and $\lambda\mathrm{Tr}(\sum_{\mathrm{k}=1}^{\mathrm{n_r}} \boldsymbol{A}_{\mathrm{k}}[\mathrm{i}]\boldsymbol{A}_{\mathrm{k}}^H[\mathrm{i}]) = \mathrm{P_R}$, respectively, and then solving the power constraint equations. The problem is that a high computational complexity of $\mathrm{O}(((T+1)N)^3)$ is required, and it will increase cubically with the number of antennas or the use of more powerful STC encoders. The computational complexity is calculated by counting the number of multiplies.

3.3.2 Joint Linear MBER Receiver Design with Power Allocation

The MBER receiver design [102,103] with power allocation in the second phase is derived as follows. The BPSK modulation scheme is utilized for simplicity. The derivation in the following can be used for other modulation schemes, and the only difference is the requirement of calculating the imaginary part. According to the expression in (3.4), the desired information symbols at the destination node can be computed as

$$b_j[i] = \mathrm{sgn}(\boldsymbol{w}_{\mathrm{j}}^H[\mathrm{i}]\boldsymbol{r}[\mathrm{i}]) = \mathrm{sgn}(\tilde{\mathrm{s}}_{\mathrm{j}}[\mathrm{i}]), \tag{3.9}$$

where $\tilde{s}_j[i]$ denotes the detected symbol at the receiver which can be further written as

$$\begin{aligned}
\tilde{s}_j[i] &= \Re\left[\boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\right] = \Re\left[\boldsymbol{w}_j^H[i](\boldsymbol{H}_D[i]\boldsymbol{s}[i] + \boldsymbol{n}_D[i])\right] \\
&= \Re\left[\boldsymbol{w}_j^H[i]\boldsymbol{H}_D[i]\boldsymbol{s}[i] + \boldsymbol{w}_j^H[i]\boldsymbol{n}_D[i]\right] \\
&= \Re\left[\tilde{s}_j'[i] + e_j[i]\right],
\end{aligned} \tag{3.10}$$

where $\tilde{s}'_j[i]$ is the noise-free detected symbol, and $e_j[i]$ denotes the error factor for the jth detected symbol. Define an $N \times N_b$ matrix $\bar{\boldsymbol{S}}$ which is constructed by a set of vectors $\bar{\boldsymbol{s}}_l = [s_{l_1}, s_{l_2}, ..., s_{l_N}]^T,\ l = 1, 2, ..., N_b$ and $N_b = 2^N$, containing all the possible combinations of the transmitted symbol vector $\boldsymbol{s}[i]$ and we can obtain

$$\bar{s}_{l_j}[i] = \Re\left[\boldsymbol{w}_j^H[i]\boldsymbol{H}_D[i]\bar{\boldsymbol{s}}_l + \boldsymbol{w}_j^H[i]\boldsymbol{n}_D[i]\right] = \Re\left[\boldsymbol{w}_j^H[i]\bar{\boldsymbol{r}}_l[i]\right] + e_{l_j}[i]\right] = \bar{s}'_{l_j}[i] + e_{l_j}[i], \tag{3.11}$$

where $\bar{s}'_{l_j}[i]$ denotes the noise-free detected symbol in the lth column and the jth row of $\bar{\boldsymbol{S}}$. Since the probability density function (pdf) of $\bar{\boldsymbol{r}}[i]$ is given by

$$p_{\bar{\boldsymbol{r}}[i]} = \frac{1}{N_b\sqrt{2\pi\sigma_n^2\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}}\sum_{l=1}^{N_b}\exp\big(-\frac{(\bar{s}_{l_j}[i] - \bar{s}'_{l_j}[i])^2}{2\sigma_n^2\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}\big), \tag{3.12}$$

by employing the Q function, we can obtain the BER expression of the cooperative MIMO system which is

$$P_E(\boldsymbol{w}_j[i], a_{S_j}[i], a_{k_j}[i]) = \frac{1}{N_b}\sum_{l=1}^{N_b}\boldsymbol{Q}(c_{l_j}[i]), \tag{3.13}$$

where

$$c_{l_j}[i] = \frac{sgn(s_{l_j})\bar{s}'_{l_j}[i]}{\sigma_n\sqrt{\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}} = \frac{sgn(s_{l_j})\Re\left[\boldsymbol{w}_j^H[i]\bar{\boldsymbol{r}}_l[i]\right]}{\sigma_n\sqrt{\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}}. \tag{3.14}$$

The joint power allocation with linear receiver design problem is given by

$$\begin{aligned}[\boldsymbol{w}_j[i], a_{S_j}[i], a_{k_j}[i]] &= \underset{\boldsymbol{w}_j[i], a_{S_j}[i], a_{k_j}[i]}{\operatorname{argmin}} P_E(\boldsymbol{w}_j[i], a_{S_j}[i], a_{k_j}[i]),\\ s.t.\ \sum_{j=1}^{N} a_{S_j}[i] &\le \mathrm{P_T},\ \sum_{\mathrm{k=1}}^{\mathrm{n_r}}\sum_{\mathrm{j=1}}^{\mathrm{N}}\mathrm{a_{k_j}[i]} \le \mathrm{P_R}.\end{aligned} \tag{3.15}$$

According to (3.13) and (3.14), the solution of the design problem in (3.15) with respect to $\boldsymbol{w}_j[i]$, $a_{S_j}[i]$ and $a_{k_j}[i]$ is not a closed-form one. Therefore, we design an adaptive JAPA strategy according to the MBER criterion using the SG algorithm in order to update the parameters iteratively to achieve the optimal solution in the next section.

3.3.3 Joint Linear MSR Receiver Design with Power Allocation

We develop a joint power allocation strategy which focuses on maximizing the sum rate at the destination node. The expression of the sum rate after the detection is derived in [116] as

$$I = \frac{1}{2}\log_2(1 + SNR_{ins}), \tag{3.16}$$

where

$$SNR_{ins} = \frac{E[\boldsymbol{s}^H\boldsymbol{s}]Tr(\boldsymbol{W}^H[i]\boldsymbol{H}_D[i]\boldsymbol{H}_D^H[i]\boldsymbol{W}[i])}{E[\boldsymbol{n}_D^H[i]\boldsymbol{n}_D[i]]}, \tag{3.17}$$

and $\boldsymbol{W}[i] = [\boldsymbol{w}_1[i], \boldsymbol{w}_2[i], ..., \boldsymbol{w}_N[i]]$ denotes the $N(T+1)\times N$ linear receive filter matrix, and $\boldsymbol{n}[i]$ denotes the received noise vector. By substituting (3.4) into (3.17), we can obtain

$$\begin{aligned} &SNR_{ins} \\ &= \frac{\sigma_s^2 Tr(\boldsymbol{W}^H[i](\sum_{k=1}^{n_r}\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i]\boldsymbol{F}_{SR_k}[i]\boldsymbol{A}_S[i])(\sum_{k=1}^{n_r}\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i]\boldsymbol{F}_{SR_k}[i]\boldsymbol{A}_S[i])^H\boldsymbol{W}[i])}{\sigma_n^2 Tr(\boldsymbol{W}^H[i](\boldsymbol{I}_{N(T+1)} + (\sum_{k=1}^{n_r}\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_{R_kD}[i])(\sum_{k=1}^{n_r}\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_{R_kD}[i])^H)\boldsymbol{W}[i])}. \end{aligned} \tag{3.18}$$

Since the logarithm is an increasing function, maximizing the sum rate is equivalent to maximizing the instantaneous SNR. The optimization problem can be written as

$$\begin{aligned} &[\boldsymbol{W}[i], \boldsymbol{A}_S[i], \boldsymbol{A}_k[i]] = \underset{\boldsymbol{W}[i],\boldsymbol{A}_S[i],\boldsymbol{A}_k[i]}{\operatorname{argmax}} SNR_{ins}, \\ &s.t.\ \mathrm{Tr}(\boldsymbol{A}_{\mathrm{k}}[\mathrm{i}]\boldsymbol{A}_{\mathrm{k}}^H[\mathrm{i}]) \le \mathrm{P_R},\ \mathrm{Tr}(\boldsymbol{A}_{\mathrm{S}}[\mathrm{i}]\boldsymbol{A}_{\mathrm{S}}^H[\mathrm{i}]) \le \mathrm{P_T}, \end{aligned} \tag{3.19}$$

where SNR_{ins} is given by (3.18).

As expressed in (3.18), the solution of (3.19) with respect to the matrices $\boldsymbol{W}[i]$, $\boldsymbol{A}_S[i]$ and $\boldsymbol{A}_k[i]$ does not result in closed-form expressions. Therefore, in the next section we propose a JAPA SG algorithm to obtain the joint optimization algorithm for determining the linear receiver filter parameters and power allocation matrices to maximize the sum rate.

3.4 Low Complexity Joint Linear Receiver Design with Power Allocation

In this section, we jointly design an adjustable power allocation matrix and the linear receiver for the DSTC scheme in cooperative MIMO systems. Adaptive SG algorithms [59] with reduced complexity are devised.

3.4.1 Joint Adaptive SG Estimation for MMSE Receive Filter and Power Allocation

According to (3.5) and (3.6), the joint optimization problem for power allocation matrices and receiver parameter vectors depend on each other. By computing the instantaneous gradient terms of (3.6) with respect to $\boldsymbol{w}_j[i]$, $a_{S_j}[i]$ and $a_{k_j}[i]$, respectively, we can obtain

$$\begin{aligned}
\nabla \mathscr{L}_{\boldsymbol{w}_j^*[i]} &= -\boldsymbol{r}[i](s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]))^* = -\boldsymbol{r}[i]e_j^*[i], \\
\nabla \mathscr{L}_{a_{S_j}^*[i]} &= -\nabla_{a_{S_j}^*[i]}(\boldsymbol{w}_j^H[i]\boldsymbol{h}_{SDA_j}[i]a_{S_j}[i]s_j[i])^H(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]) \\
&= \boldsymbol{h}_{SDA_j}^H[i]\boldsymbol{w}_j[i]s_j^*[i]e_j[i], \\
\nabla \mathscr{L}_{a_{k_j}^*[i]} &= -\nabla_{a_{k_j}^*[i]}(\boldsymbol{w}_{R_j}^H[i]\boldsymbol{g}_{eq_{k_j}}[i]\boldsymbol{f}_{k_j}[i]a_{k_j}[i]\boldsymbol{s}[i])^H(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]) \\
&= -(\boldsymbol{g}_{eq_{k_j}}[i]\boldsymbol{f}_{k_j}[i]\boldsymbol{s}[i])^H\boldsymbol{w}_{R_j}[i]e_j[i],
\end{aligned} \tag{3.20}$$

where $\boldsymbol{h}_{SDA_j}[i]$ denotes the jth column with dimension $N(T+1) \times 1$ of the equivalent channel matrix $\boldsymbol{H}_{SDA}[i]$, and $\boldsymbol{g}_{eq_{k_j}}[i]$ and $\boldsymbol{f}_{k_j}[i]$ denote the jth column and the jth row of the channel matrices $\boldsymbol{F}_k[i]$ and $\boldsymbol{G}_{eq_{k_j}}[i]$, respectively. The $NT \times 1$ vector $\boldsymbol{w}_{R_j}[i]$ is the parameter vector for the received symbols from the relay nodes. The error signal is denoted by $e_j[i] = s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]$. We can devise an adaptive SG estimation algorithm by using the instantaneous gradient terms of the Lagrangian which were previously derived with SG descent rules [59]:

$$\begin{aligned}
\boldsymbol{w}_j[i+1] &= \boldsymbol{w}_j[i] - \mu\nabla\mathscr{L}_{\boldsymbol{w}_j^*[i]}, \\
a_{S_j}[i+1] &= a_{S_j}[i] - \nu\nabla\mathscr{L}_{a_{S_j}^*[i]}, \\
a_{k_j}[i+1] &= a_{k_j}[i] - \tau\nabla\mathscr{L}_{a_{k_j}^*[i]},
\end{aligned} \tag{3.21}$$

where μ, ν and τ are the step sizes of the recursions for the estimation procedure. The computational complexity of $\boldsymbol{w}_j[i]$, $a_{S_j}[i]$ and $a_{k_j}[i]$ in (3.21) is $(\boldsymbol{O}(NT))$, $(\boldsymbol{O}(3NT))$ and $(\boldsymbol{O}(N^2T^2))$, respectively, which is much less than that of the algorithm we described in Section III. The computational complexity given here is the number of multiplies in each iteration and the number of iteration we consider in the simulation is equal to $N+T$. The analysis of the convergence is derived in Section 4.3.5.

It is worth to mention that instead of calculating the Lagrange multiplier λ, a normalization of the power allocation matrices after the optimization which ensures that the energy is not increased is required. A simple division is utilized and the normalization procedure is implemented as

$$\begin{aligned} \boldsymbol{A}_S[i+1] =& \frac{\sqrt{\mathrm{P_T}}\boldsymbol{A}_S[i+1]}{\parallel \boldsymbol{A}_S[i+1] \parallel_F}, \\ \boldsymbol{A}_k[i+1] =& \frac{\sqrt{\mathrm{P_R}}\boldsymbol{A}_k[i+1]}{\parallel \sum_{k=1}^{n_r} \mathrm{Tr}(\boldsymbol{A}_{\mathrm{k}}[\mathrm{i}+1] \parallel_{\mathrm{F}}}. \end{aligned} \tag{3.22}$$

3.4.2 Joint Adaptive MBER SG Estimation and Power Allocation

The key strategy to derive an adaptive estimation algorithm for solving (3.15) is to find out an efficient and reliable method to calculate the pdf of the received symbol vector $\boldsymbol{r}[i]$ at the destination node. According to the algorithms in [77], kernel density estimation provides an effective method for accurately estimating the required pdf.

By transmitting a block of v training samples $\hat{\boldsymbol{s}} = \mathrm{sgn}(\hat{\boldsymbol{b}})$, the kernel density estimated pdf of $\hat{\boldsymbol{s}}[i]$ is given by

$$p_{\hat{s}} = \frac{1}{v\sqrt{2\pi}\rho_n\sqrt{\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}} \sum_{j=1}^{v} exp\big(- \frac{(\tilde{s}_j - \hat{s}_j)^2}{2\rho_n^2 \boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}\big), \tag{3.23}$$

where ρ_n is the coefficient which has a lower bound of $\rho_n = \left(\frac{4}{3v}\right)^{\frac{1}{5}}\sigma_n$ suggested in [77]. The symbol $\tilde{s}_j$ is calculated by (3.11), and $\hat{s}_j$ stands for the jth element in the $M \times 1$

training samples $\hat{\boldsymbol{s}}$. The expression of the BER can be derived as

$$\hat{P_E}(\boldsymbol{w}_j[i], a_{S_j}[i], a_{k_j}[i]) = \frac{1}{v}\sum_{j=1}^{v}\boldsymbol{Q}(c_j[i]), \tag{3.24}$$

where

$$c_j[i] = \frac{sgn(\hat{s}_j)\hat{s}_j}{\rho_n\sqrt{\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}}. \tag{3.25}$$

By substituting (3.25) into (3.24) and taking the gradient with respect to different arguments, we can obtain

$$\nabla P_{E_{\boldsymbol{w}_j}}[i] = \frac{1}{v\sqrt{2\pi}\sqrt{\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}}\sum_{j=1}^{v}\exp\big(-\frac{c_j^2[i]}{2}\big)sgn(s_j)\frac{\bar{\boldsymbol{r}}[i]-\frac{1}{2}\bar{s}_j[i]\boldsymbol{w}_j[i]}{\sigma_n\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}, \tag{3.26}$$

$$\nabla P_{E_{a_{S_j}}}[i] = \frac{1}{v\sqrt{2\pi}\sigma_n\sqrt{\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}}\sum_{j=1}^{v}exp\big(-\frac{c_j^2[i]}{2}\big)sgn(s_j)\Re[\boldsymbol{w}_j^H[i]\boldsymbol{h}_{SDA_j}[i]s_j], \tag{3.27}$$

$$\nabla P_{E_{a_{k_j}}}[i] = \frac{1}{v\sqrt{2\pi}\sigma_n\sqrt{\boldsymbol{w}_{D_j}^H[i]\boldsymbol{w}_{R_j}[i]}}\sum_{j=1}^{v}exp\big(-\frac{c_j^2[i]}{2}\big)sgn(s_j)\Re[\boldsymbol{w}_{R_j}^H[i]\boldsymbol{h}_{k_j}[i]s_j], \tag{3.28}$$

where $\boldsymbol{h}_{k_j}[i] = \boldsymbol{g}_{eq_{k_j}}[i]f_{k_j}[i]$ denotes the equivalent channel vector assigned for s_{m_j}. By making use of an SG algorithm in [59], the updated $\boldsymbol{w}_j[i]$, $a_{S_j}[i]$ and$a_{k_j}[i]$ can be calculated by (3.21). The convergence property of the joint iterative optimization problems have been tested and proved by Niesen et al. in [78]. In the proposed design problem the receive filter parameter vectors and the power allocation parameters depend on each other, and the proposed JAPA algorithms provide an iterative update process and finally both of the desired items will reach at least a local optimum of the BER cost function.

3.4.3 Joint Adaptive MSR SG Algorithm for Power Allocation and Receiver Design

The proposed power allocation algorithm that maximizes the sum rate at the destination node is derived as follows. We consider the design problem in (3.19) and the instantaneous received SNR_{ins} as given in (3.18). According to the property of the trace $\mathrm{Tr}(\cdot)$ we can

obtain

$$SNR_{ins} = \frac{\sigma_s^2}{n_{eq}[i]} Tr(\boldsymbol{R}_{SDA}[i]\boldsymbol{A}_S^H[i]) = \frac{\sigma_s^2}{n_{eq}[i]} Tr(\sum_{k=1}^{n_r} \boldsymbol{R}_{\boldsymbol{G}_{eq_k}}[i]\boldsymbol{A}_k^H[i]), \tag{3.29}$$

where

$$\begin{aligned}
&\boldsymbol{R}_{SDA}[i] = \boldsymbol{H}_{SDA}^H[i]\boldsymbol{W}[i]\boldsymbol{W}^H[i]\boldsymbol{H}_{SDA}[i]\boldsymbol{A}_S[i], \\
&\boldsymbol{R}_{\boldsymbol{G}_{eq_k}}[i] = \boldsymbol{G}_{eq_k}^H[i]\boldsymbol{W}[i]\boldsymbol{W}^H[i]\boldsymbol{H}_{SDA}[i]\boldsymbol{A}_S[i]\boldsymbol{A}_S^H[i]\boldsymbol{F}_{SR_k}^H[i], \\
&n_{eq}[i] = \sigma_n^2 Tr(\boldsymbol{W}^H[i]\boldsymbol{W}[i] + \boldsymbol{W}^H[i](\sum_{k=1}^{n_r} \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i])(\sum_{k=1}^{n_r} \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i])^H\boldsymbol{W}[i]).
\end{aligned}$$

Since the power allocation matrices $\boldsymbol{A}_S[i]$ and $\boldsymbol{A}_k[i]$ are diagonal, we just focus on the terms containing the conjugate of the jth parameter in order to simplify the derivation, and rewrite (3.29) as

$$\begin{aligned}
SNR_{ins} &= \frac{\sigma_s^2 \sum_{k=1}^{n_r} \sum_{j=1}^{N} r_{SDA_j}[i] a_{S_j}^*[i]}{\sigma_n^2 \sum_{j=1}^{N} (\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i] + \sum_{k=1}^{n_r} r_{\boldsymbol{N}_{kj}}[i] a_{k_j}^*[i])} \\
&= \frac{\sigma_s^2 \sum_{k=1}^{n_r} \sum_{j=1}^{N} r_{\boldsymbol{G}_{eq_{k_j}}}[i] a_{k_j}^*[i]}{\sigma_n^2 \sum_{j=1}^{N} (\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i] + \sum_{k=1}^{n_r} r_{\boldsymbol{N}_{kj}}[i] a_{k_j}^*[i])},
\end{aligned} \tag{3.30}$$

where $r_{\boldsymbol{G}_{eq_{k_j}}}[i]$ and $r_{\boldsymbol{N}_{kj}}[i]$ denotes the jth element in the diagonal of $\boldsymbol{R}_{\boldsymbol{G}_{eq_k}}[i]$ and $\boldsymbol{R}_{\boldsymbol{N}_k}[i]$, respectively. $\boldsymbol{R}_{\boldsymbol{N}_k}[i] = \boldsymbol{G}_{eq_k}^H[i]\boldsymbol{W}[i]\boldsymbol{W}^H[i]\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i]$ denotes the equivalent matrix assigned for the noise at the kth relay node.

By taking the stochastic gradient of (3.30) with respect to $a_{S_j}^*[i]$, $a_{k_j}^*[i]$ and $\boldsymbol{W}^H[i]$ we can obtain

$$\begin{aligned}
\nabla_{\boldsymbol{W}[i]} &= \frac{\sigma_s^2}{n_{eq}[i]} (Tr(\| \boldsymbol{H}_{SDA}[i]\boldsymbol{A}_S[i] \|_F^2 \boldsymbol{W}[i]) n_{eq}[i] \\
&\quad - \| \boldsymbol{W}^H[i]\boldsymbol{H}_{SDA}[i]\boldsymbol{A}_S[i] \|_F^2 Tr(\boldsymbol{W}[i] + \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i](\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i])\boldsymbol{W}[i])), \\
\nabla_{a_{S_j}[i]} &= \frac{\sigma_s^2}{n_{eq}[i]} r_{SDA_j}[i], \\
\nabla_{a_{R_k D}[i]} &= \frac{\sigma_s^2}{n_{eq}[i]} (r_{\boldsymbol{G}_{eq}}[i] n_{eq}[i] - \sigma_n^2 r_{\boldsymbol{N}_{kj}}[i] \sum_{j=1}^{N} r_{\boldsymbol{G}_{eq_{k_j}}}[i] a_{k_j}^*[i]).
\end{aligned} \tag{3.31}$$

By using (3.21) and (3.22) the proposed algorithm is achieved. Table I shows a summary

Table 3.1: The JAPA SG Algorithms

1: **Initialization:**

$\boldsymbol{W}[0] = \boldsymbol{I}_{(T+1)N\times 1}$,
$a_{S_j}[0] = 1,\ a_{k_j}[0] = 1$,
$\boldsymbol{H}_{SDA}[i] = \sum_{k=1}^{n_r} \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i]\boldsymbol{F}_{SR_k}[i]$,

2: **for** $j = 1$ **to** N **do**

2-1: **JAPA SG MMSE Algorithm**

$e_j[i] = s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]$,
$\nabla\mathscr{L}_{\boldsymbol{w}_j^H[i]} = -\boldsymbol{r}[i]e_j^*[i]$,
$\nabla\mathscr{L}_{a_{S_j}^*[i]} = \boldsymbol{h}_{SDA_j}^H[i]\boldsymbol{w}_j[i]s_j^*[i]e_j[i]$,
$\nabla\mathscr{L}_{a_{k_j}^*[i]} = -(\boldsymbol{g}_{eq_{k_j}}[i]\boldsymbol{f}_{k_j}[i]\boldsymbol{s}[i])^H\boldsymbol{w}_{D_j}[i]e_j[i]$,

2-2: **JAPA SG MBER Algorithm**

$c_{m_j}[i] = \frac{sgn(\hat{s}_{m_j})\bar{\tilde{s}}_{m_j}}{\rho_n\sqrt{\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}},\ \boldsymbol{h}_{k_j}[i] = \boldsymbol{g}_{eq_{k_j}}[i]f_{k_j}[i]$
$\nabla P_{E_{\boldsymbol{w}_j}}[i] = \frac{1}{M\sqrt{2\pi}\sqrt{\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}}\sum_{j=1}^{M}\exp\big(-\frac{c_j^2[i]}{2}\big)sgn(s_j)\frac{\bar{\boldsymbol{r}}_{\boldsymbol{j}}[i]-\frac{1}{2}\bar{s}_j[i]\boldsymbol{w}_j[i]}{\sigma_n\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}$,
$\nabla P_{E_{a_{S_j}}}[i] = \frac{1}{M\sqrt{2\pi}\sigma_n\sqrt{\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i]}}\sum_{j=1}^{M}exp\big(-\frac{c_j^2[i]}{2}\big)sgn(s_j)\Re[\boldsymbol{w}_j^H[i]\boldsymbol{h}_{SDA_j}[i]s_j]$,
$\nabla P_{E_{a_{k_j}}}[i] = \frac{1}{M\sqrt{2\pi}\sigma_n\sqrt{\boldsymbol{w}_{D_j}^H[i]\boldsymbol{w}_{D_j}[i]}}\sum_{j=1}^{M}exp\big(-\frac{c_j^2[i]}{2}\big)sgn(s_j)\Re[\boldsymbol{w}_{D_j}^H[i]\boldsymbol{h}_{k_j}[i]s_j]$,

2-3: **JAPA SG MSR Algorithm**

$\boldsymbol{R}_{SDA}[i] = \boldsymbol{H}_{SDA}^H[i]\boldsymbol{W}[i]\boldsymbol{W}^H[i]\boldsymbol{H}_{SDA}[i]\boldsymbol{A}_S[i]$,
$\boldsymbol{R}_{\boldsymbol{G}_{eq_k}}[i] = \boldsymbol{G}_{eq_k}^H[i]\boldsymbol{W}[i]\boldsymbol{W}^H[i]\boldsymbol{H}_{SDA}[i]\boldsymbol{A}_S[i]\boldsymbol{A}_S^H[i]\boldsymbol{F}_{SR_k}^H[i]$,
$n_{eq}[i] = \sigma_n^2 Tr(\boldsymbol{W}^H[i]\boldsymbol{W}[i] + \boldsymbol{W}^H[i](\sum_{k=1}^{n_r}\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i])(\sum_{k=1}^{n_r}\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i])^H\boldsymbol{W}[i])$,
$\nabla_{\boldsymbol{W}[i]} = \frac{\sigma_s^2}{n_{eq}[i]}(Tr(\parallel \boldsymbol{H}_{SDA}[i]\boldsymbol{A}_S[i] \parallel_F^2 \boldsymbol{W}[i])n_{eq}[i]$
$\quad - \parallel \boldsymbol{W}^H[i]\boldsymbol{H}_{SDA}[i]\boldsymbol{A}_S[i] \parallel_F^2 Tr(\boldsymbol{W}[i] + \boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i](\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_k[i])\boldsymbol{W}[i]))$,
$\nabla_{a_{S_j}[i]} = \frac{\sigma_s^2}{n_{eq}[i]}r_{SDA_j}[i]$,
$\nabla_{a_{R_kD}[i]} = \frac{\sigma_s^2}{n_{eq}[i]}(r_{\boldsymbol{G}_{eq}}[i]n_{eq}[i] - \sigma_n^2 r_{\boldsymbol{N}_{kj}}[i]\sum_{j=1}^{N}r_{\boldsymbol{G}_{eq_{k_j}}}[i]a_{k_j}^*[i])$,

end for

3: **Update:**

$\boldsymbol{w}_j[i+1] = \boldsymbol{w}_j[i] - \mu\nabla_{\boldsymbol{w}_j^*[i]}$,
$a_{S_j}[i+1] = a_{S_j}[i] - \nu\nabla_{a_{S_j}^*[i]}$,
$a_{k_j}[i+1] = a_{k_j}[i] - \tau\nabla_{a_{k_j}^*[i]}$,

4: **Normalization:**

$\boldsymbol{A}_S[i+1] = \frac{\sqrt{\mathrm{P_T}}\boldsymbol{A}_S[i+1]}{\|\boldsymbol{A}_S[i+1]\|_F}$,
$\boldsymbol{A}_k[i+1] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{A}_k[i+1]}{\|\sum_{k=1}^{n_r}\mathrm{Tr}(\boldsymbol{A}_k[i+1]\|_F}$.

of the JAPA SG algorithms with different criteria. A low complexity channel estimation method derived in [74] can be also employed to obtain the channel matrices required in the proposed algorithms.

Table 3.2: Computational Complexity of the Algorithms

Algorithm	Number of operations per symbol	
	Multiplications	Additions
PA MMSE(3.3.1)	$(T+1)^6N^6 + (T+1)N + 8(T+1)N$	$7(T+1)N + 2$
JAPA MMSE SG(3.4.1)	$(7T+5)N$	$4(T+1)N$
JAPA MBER SG(3.4.2)	$(M+1)(T+1)N + M$	$(2M+1)(T+1)N$
OPA [11]	$N^4 + 2N^2 + N^2T^2$	$2NT - 1$
JAPA MSR SG(3.4.3)	$7(T+1)N + N + 1$	$7(T+1)N + N + 2$
PO-PR-SIM [10]	$N^4 + 2N^2$	$2NT$

3.5 Analysis

The proposed JAPA SG algorithms employ three different criteria to compute the power allocation matrices iteratively at the destination node and then send them back via a feedback channel. In this section, we will illustrate the low computational complexity required by the proposed JAPA SG algorithms compared to the existing power allocation optimization algorithms using the same criteria and will examine their feedback requirement.

3.5.1 Computational Complexity Analysis

In Table II, we compute the number of additions and multiplications to compare the complexity of the proposed JAPA SG algorithms with the conventional power allocation strategies. The computational complexity of the proposed algorithms is calculated by summing the number of additions and multiplications, which is related to the number of antennas N, the number of relay nodes n_r, and the $N \times T$ STC scheme employed in the network. Note that the computational complexity in [10] and [11] is high because the key parameters in the algorithms can only be obtained by eigenvalue decomposition, which requires a high-cost computing process when the matrices are large [79].

3.5.2 Feedback Requirements

The proposed JAPA SG algorithms require communication between the relay nodes and the destination node according to different algorithms. The feedback channel we considered is modeled as an AWGN channel. A 4-bit quantization scheme, which quantizes the real part and the imaginary part by 4 bits, respectively, is utilized prior to the feedback channel. More efficient schemes employing vector quantization [104, 105] and that take into account correlations between the coefficients are also possible.

For simplicity we show how the feedback errors in power allocation matrices at the relay nodes affect the accuracy of the detection when only one relay node is employed. The $N \times N$ diagonal power allocation matrix with feedback errors at the kth relay node is derived as

$$\hat{\boldsymbol{A}}[i] = \boldsymbol{A}[i] + \boldsymbol{E}[i], \tag{3.32}$$

where $\boldsymbol{A}[i]$ denotes the accurate power allocation matrix and $\boldsymbol{E}[i]$ stands for the error matrix. We assume the parameters in $\boldsymbol{E}[i]$ are Gaussian with zero mean and variance σ_f. Then the received symbol vector is given by

$$\begin{aligned} \hat{\boldsymbol{r}}[i] &= \boldsymbol{G}_{eq}[i]\hat{\boldsymbol{A}}[i]\boldsymbol{F}[i]\boldsymbol{s}[i] + \boldsymbol{G}_{eq}[i]\hat{\boldsymbol{A}}[i]\boldsymbol{n}_{SR}[i] + \boldsymbol{n}_{RD}[i] \\ &= \boldsymbol{G}_{eq}[i]\hat{\boldsymbol{A}}[i]\boldsymbol{F}[i]\boldsymbol{s}[i] + \hat{\boldsymbol{n}}_D[i], \end{aligned} \tag{3.33}$$

where $\hat{\boldsymbol{n}}_D[i]$ denotes the received noise with zero mean and variance $\sigma_f(\boldsymbol{I} + \| \boldsymbol{G}_{eq}[i]\hat{\boldsymbol{A}}[i] \|_F)$. By defining $\hat{\boldsymbol{p}} = E\left[\hat{\boldsymbol{r}}\boldsymbol{s}^H\right]$ and $\hat{\boldsymbol{R}}\boldsymbol{x} = E\left[\hat{\boldsymbol{r}}\hat{\boldsymbol{r}}^H\right]$, we can obtain the MSE with the feedback errors as

$$\begin{aligned} m_e &= \mathrm{Tr}(\hat{\boldsymbol{p}}^H \hat{\boldsymbol{R}}\boldsymbol{x}^{-1}\hat{\boldsymbol{p}}) \\ &= \mathrm{Tr}\big((\boldsymbol{G}_{\mathrm{eq}}[\mathrm{i}](\boldsymbol{A}[\mathrm{i}] + \boldsymbol{E}[\mathrm{i}])\boldsymbol{F}[\mathrm{i}]\sigma_{\mathrm{s}})^H \\ &\qquad (\| \boldsymbol{G}_{eq}[i](\boldsymbol{A}[i] + \boldsymbol{E}[i])\boldsymbol{F}[i] \|_F^2 \, \sigma_s + (\boldsymbol{I} + \| \boldsymbol{G}_{eq}[i](\boldsymbol{A}[i] + \boldsymbol{E}[i]) \|_F^2)\sigma_f)^{-1} \\ &\qquad (\boldsymbol{G}_{eq}[i](\boldsymbol{A}[i] + \boldsymbol{E}[i])\boldsymbol{F}[i]\sigma_s)\big), \end{aligned} \tag{3.34}$$

while the MSE expression of the system with accurate power allocation parameters is

given by

$$\begin{aligned} m = \mathrm{Tr}\big((\boldsymbol{G}_{\mathrm{eq}}[\mathrm{i}]\boldsymbol{A}[\mathrm{i}]\boldsymbol{F}[\mathrm{i}]\sigma_{\mathrm{s}})^H \\ (\parallel \boldsymbol{G}_{eq}[i]\boldsymbol{A}[i]\boldsymbol{F}[i] \parallel_F^2 \sigma_s + (\boldsymbol{I} + \parallel \boldsymbol{G}_{eq}[i]\boldsymbol{A}[i] \parallel_F^2)\sigma_n)^{-1} \\ (\boldsymbol{G}_{eq}[i]\boldsymbol{A}[i]\boldsymbol{F}[i]\sigma_s)\big). \end{aligned} \tag{3.35}$$

By substituting (3.35) into (3.34), we can obtain the difference between the MSE expressions with accurate and inaccurate power allocation matrices which is given by

$$\begin{aligned} m_e &= m + \mathrm{Tr}\big((\boldsymbol{G}_{\mathrm{eq}}[\mathrm{i}]\boldsymbol{E}[\mathrm{i}]\boldsymbol{F}[\mathrm{i}]\sigma_{\mathrm{s}})^H \\ &\qquad (\parallel \boldsymbol{G}_{eq}[i]\boldsymbol{E}[i]\boldsymbol{F}[i] \parallel_F^2 \sigma_s + (\boldsymbol{I} + \parallel \boldsymbol{G}_{eq}[i]\boldsymbol{E}[i] \parallel_F^2)\sigma_n)^{-1} \\ &\qquad (\boldsymbol{G}_{eq}[i]\boldsymbol{E}[i]\boldsymbol{F}[i]\sigma_s)\big) \\ &= m + m_{eo}. \end{aligned} \tag{3.36}$$

The received power allocation matrices are positive definite according to the power constraint, which indicates m_{eo} is a positive scalar. The expression in (3.36) denotes an analytical derivation of the MSE at the destination node, which indicates the impact of the limited feedback employed in the JAPA SG algorithms.

3.6 Simulations

The simulation results are provided in this section to assess the proposed JAPA SG algorithms. The equal power allocation (EPA) algorithm in [13] is employed in order to identify the benefits achieved by the proposed power allocation algorithms. The cooperative MIMO system considered employs an AF protocol with the Alamouti STBC scheme in [74] using BPSK modulation in a quasi-static block fading channel with AWGN. The effect of the direct link is also considered. It is possible to employ the DF protocol or use a different number of antennas and relay nodes with a simple modification. The ML detection is considered at the destination node to indicate the achievement of full receive diversity even though other detection algorithms [85–87] can also be adopted. The system is equipped with $n_r = 1$ relay node and $N = 2$ antennas at each node. In the simulations, we set the symbol power σ_s^2 to 1. The SNR in the simulations is the received SNR which is calculated by (3.29).

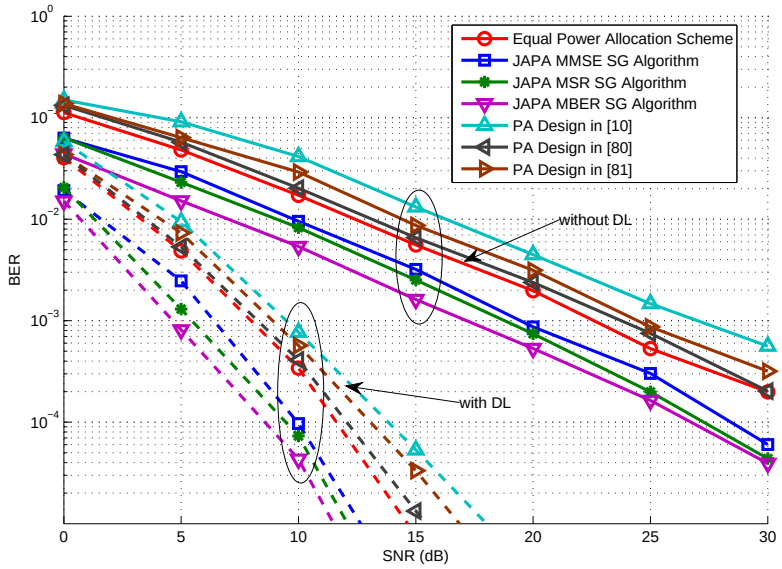

Figure 3.1: SNR versus BER for JAPA SG Algorithms

The proposed JAPA SG algorithms derived in Section 3.4 are compared with the EPA algorithm and the power allocation algorithms in [10], [80] and [81] with and without the direct link (DL) in Fig. 3.1. The results illustrate that the performance of the proposed JAPA SG algorithms is superior to the EPA algorithm by more than 3dB. The power allocation algorithms in the literature are designed for AF systems without re-encoding at the relays and in order to obtain a fair comparison, they have been adapted to the system considered in Fig. 3.1. However, as shown in the plot, the performance of the existing power allocation algorithms cannot achieve a BER performance as good as the proposed algorithms. In the low SNR scenario, the JAPA MSR SG algorithm can achieve a better BER performance compared with the JAPA MMSE SG algorithm, while with the increase of the SNR, the BER curves of the JAPA MSR and MMSE SG algorithms approach the BER performance of the JAPA MBER SG algorithm. The BER of the JAPA MBER SG algorithm achieves the best performance because of the received BER is minimized by the algorithm in Section 3.4. The performance improvement of the proposed JAPA SG algorithms is achieved with more relays employed in the system as an increased spatial diversity is provided by the relays. It is worth to mention that the power allocation algorithms in [10] and [81] are designed for other transmission scenarios so that when we adapt these algorithms into our system their performances are worse than that with the equal power allocation scheme.

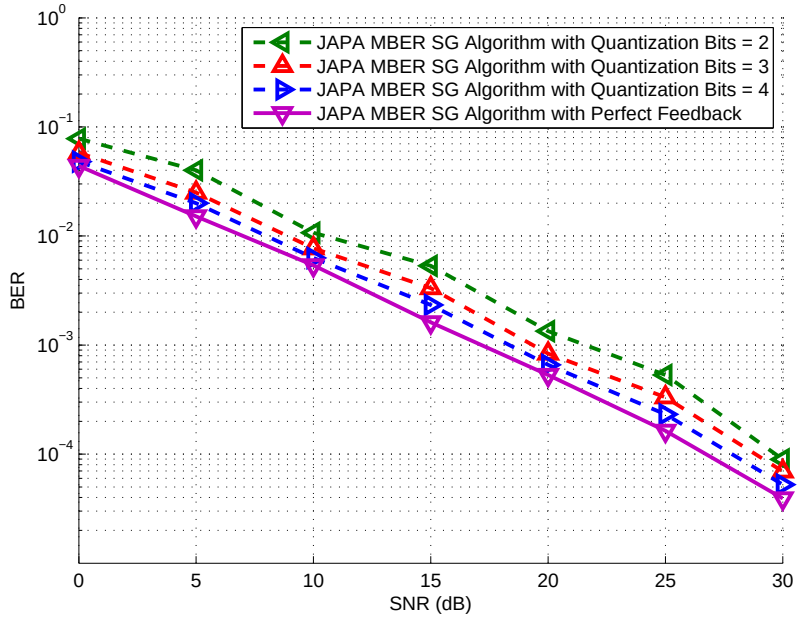

Figure 3.2: JAPA MBER SG Algorithm BER versus SNR

The simulation results shown in Fig. 3.2 illustrate the influence of the feedback channel on the JAPA MBER SG algorithm. As mentioned in Section 3.5, the optimized power allocation matrices will be sent back to each relay node and the source node through an AWGN feedback channel. In the simulation, we assume the feedback channel contains noise interference and model the noise as AWGN in order to test our algorithms. The real part and the imaginary part of the elements in the code matrices are quantized with different numbers of bits to indicates the effects of the imperfect feedback and the limitation of the feedback. The quantization and feedback errors are not considered in the simulation results in Fig. 3.1, so the optimized power allocation matrices are perfectly known at the relay node and the source node after the JAPA SG algorithm; while in Fig. 3.2, it indicates that the performance of the proposed algorithm will be affected by the accuracy of the feedback information. In the simulation, we use $2,\ 3,\ 4$ bits to quantize the real part and the imaginary part of the element in $\boldsymbol{A}_S[i]$ and $\boldsymbol{A}_k[i]$, and the feedback channel is modeled as an AWGN channel. As we can see from Fig. 3.2, by increasing the number of quantization bits for the feedback, the BER performance approaches the performance with perfect feedback, and by making use of 4 quantization bits for the real and imaginary part of each parameter in the matrices, the performance of the JAPA SG algorithm is about 1dB worse. According to the simulation, the more quantization bits we use, the better performance we will obtain.

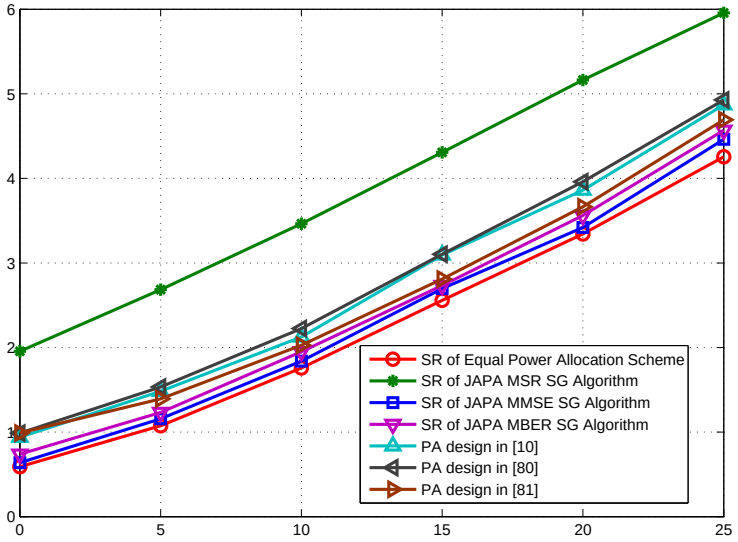

Figure 3.3: JAPA SG Algorithms Sum Rate versus SNR

The transmission rate of the cooperative MIMO network with EPA and PA schemes in [10], [80] and [81] and the proposed JAPA SG algorithms in Section 3.4.3 is given by Fig. 3.3. The number of relay nodes is equal to 1 for all the algorithms. The proposed JAPA MSR SG optimization algorithm adjusts the power allocated to each antenna in order to achieve the maximum of the sum rate in the system. From the simulation results, it is obvious that a higher throughput can be achieved by the existing PA algorithms in Fig. 3.3 compared to the proposed JAPA MMSE and MBER SG algorithms. The reason for that lies in the design criterion of the existing and the proposed algorithms. However, the improvement in the sum rate by employing the JAPA MSR SG algorithm can be observed as well. The rate improvement of the JAPA MMSE and MBER SG algorithms is not as much as the JAPA MSR SG algorithm because the optimization of the proposed JAPA MMSE and MBER optimization algorithms are not suitable for the maximization of the sum rate.

The simulation results shown in Fig. 3.4 illustrate the convergence property of the proposed JAPA SG algorithm. All the schemes have an error probability of 0.5 at the beginning, and after the first 20 symbols are received and detected, the JAPA MMSE scheme achieves a better BER performance compared with the JAPA MSR scheme and the JAPA MBER scheme a better BER than the other algorithms. With the number of received symbols increasing, the BER curve of all the schemes are almost straight, while

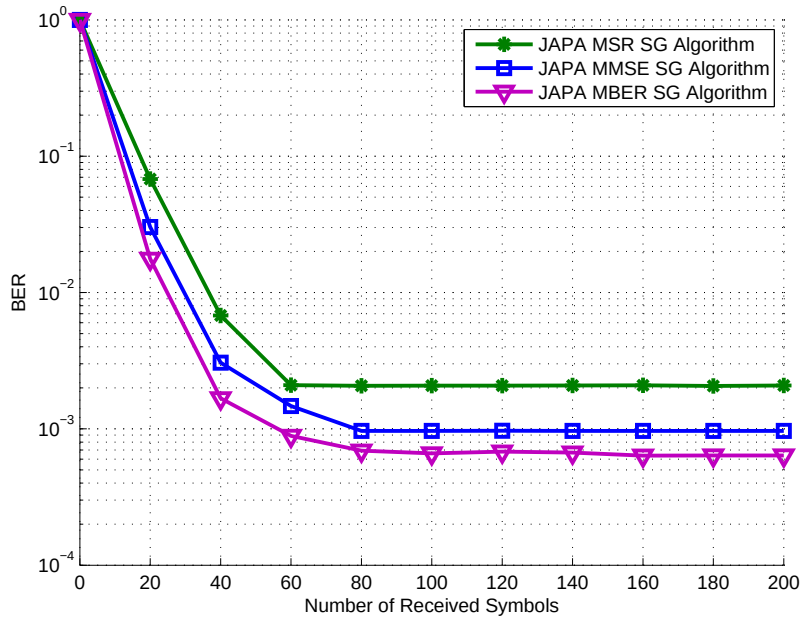

Figure 3.4: BER performance vs. Number of Symbols for JAPA SG Algorithms

the BER performance of the JAPA MBER algorithm can be further improved and obtain a fast convergence after receiving 80 symbols.

3.7 Summary

We have proposed joint adaptive power allocation and receiver design algorithms according to different criteria with the power constraint between the source node and the relay nodes, and between relay nodes and the destination node to achieve low BER performance. Joint iterative estimation algorithms with low computational complexity for computing the power allocation parameters and the linear receive filter have been derived. The simulation results illustrated the advantage of the proposed power allocation algorithms by comparing them with the equal power allocation algorithm. The proposed algorithm can be utilized with different DSTC schemes and a variety of detectors [82] [83] and estimation algorithms [84] in cooperative MIMO systems with AF strategy and can also be extended to the DF cooperation protocols.

Chapter 4

Adaptive Distributed Space-Time Coding Based on Adjustable Code Matrices for Cooperative MIMO Relaying Systems

Contents

4.1 Introduction

Cooperative MIMO systems, which employ multiple relay nodes with antennas between the source node and the destination node as a distributed antenna array, can obtain diversity gains by providing copies of the transmitted signals to improve the reliability of wireless communication systems [88]. Among the links between the relay nodes and the destination node, cooperation strategies such as AF, DF, CF [2] and various distributed space-time coding (DSTC) schemes [3], [4], and [5] can be employed.

The use of DSTC schemes at the relay node in a cooperative network, providing more copies of the desired symbols at the destination node, can offer the system diversity and coding gains to mitigate the interference. A recent focus on DSTC techniques lies in the design of delay-tolerant codes and full-diversity schemes with minimum outage probability. An opportunistic DSTC scheme with the minimum outage probability is designed for a DF cooperative network and compared with the fixed DSTC schemes in [32], while in [33] a novel opportunistic relaying algorithm is achieved by employing DSTC in an AF cooperative MIMO network. An adaptive distributed-Alamouti (D-Alamouti) space-time block code (STBC) design is proposed in [89] for non-regenerative dual-hop wireless systems which achieves the minimum outage probability. DSTC schemes for the AF protocol are discussed in [14,34,90]. In [90], the generalized ABBA (GABBA) STC scheme is extended to a distributed MIMO network with full-diversity and full-rate, while an optimal algorithm for the design of the DSTC scheme to achieve the optimal diversity and multiplexing tradeoff is derived in [34]. A quasi-orthogonal DSTBC for cooperative MIMO networks is presented and shown to achieve full rate and full diversity with any number of antennas in [14]. In [6], a new STC scheme that multiplies a randomized matrix by the STC matrix at the relay node before the transmission is derived and analyzed. The randomized space-time code (RSTC) can achieve the performance of a centralized space-time code in terms of coding gain and diversity order.

Optimal space-time codes can be obtained by transmitting the channel or other useful information for code design back to the source node, in order to achieve higher coding gains by pre-processing the symbols. In [91], the trade-off between the length of the feedback symbols, which is related to the capacity loss and the transmission rate is discussed,

whereas in [92] one solution for this trade-off problem is derived. The use of limited feedback for STC has been widely discussed in the literature. In [93], the phase information is sent back for space-time encoding in order to maintain the full diversity, and the phase feedback is employed in [94] to improve the performance of the Alamouti STBC. A limited feedback link is used in [95] and [96] to provide the channel information for the pre-coding of an orthogonal STBC scheme. Another limited feedback strategy has been considered for relay power selection in [9].

In this chapter, we propose an adaptive distributed space-time coding scheme and algorithms for cooperative MIMO relaying systems. This work was first introduced and discussed in [16]. We first develop a centralized algorithm with limited feedback to compute the parameters of an adjustable code matrix, which requires sending the adjustable code matrices back to the relay nodes after the optimization via a feedback channel that is modeled as a Binary Symmetric Channel (BSC). Then, adaptive optimization algorithms are derived based on the MSE and the ML criteria subject to constraints on the transmitted power at the relays, in order to release the destination node from the high computational complexity of the optimization process. We focus on how the adjustable code matrix affects the DSTC during the encoding and how to optimize the linear receive filter with the code matrix iteratively or, alternatively, by employing an ML detector and adjusting the code matrix. The upper bound of the error probability of the proposed adaptive DSTC is derived in order to show its advantages as compared to the traditional DSTC schemes and the influence of the imperfect feedback is discussed. It is shown that the use of an adjustable code matrix benefits the performance of the system compared to employing traditional STC schemes. Then, we derive a fully distributed matrix optimization algorithm which does not require feedback. The pairwise error probability (PEP) of the adaptive DSTC is employed in order to devise a distributed algorithm and to eliminate the need for feedback channels. The fully distributed matrix optimization algorithm allows the system to use the optimal adjustable matrix before the transmission, and also achieves the minimum PEP when the statistical information of the channel does not change. The differences of our work compared with the existing works are discussed as follows. First, an optimal adjustable code matrix will be multiplied by an existing space-time coding scheme at the relay node and the encoded data are forwarded to the destination node. The code matrix is first generated randomly as discussed in [6], and it is optimized according to different criteria at the destination node by the proposed algorithms. Second, in order

to implement the adaptive algorithms, the adjustable code matrix is optimized with the linear receive filter iteratively, and then transmitted back to the relay node via a feedback channel. The impact of the feedback errors is considered and shown in the simulations. Third, the proposed fully distributed optimization algorithm eliminates the effect of the feedback by choosing the optimal code matrix before transmission, and the receiver is released from the design task. Moreover, the power allocation matrices used in the previous section are diagonal but the code matrices in the proposed ARMO algorithms are not.

4.2 Cooperative MIMO System Model

The communication system under consideration is a two-hop cooperative MIMO system which employs multiple relay nodes as shown in Fig. 2.2. The first hop is devoted to the source transmission, which broadcasts the information symbols to the relay nodes and to the destination node. The second hop forwards the amplified and re-encoded information symbols from the relay nodes to the destination node. Each hop between the source and the relays contains 1 time slot and each hop between the relays and the destination contains T time slots, where T is the length of the codeword of the STC scheme used at relays. An orthogonal transmission protocol is considered which requires that the source node does not transmit during the time period of the second hop. In order to evaluate the adaptive optimization algorithms, a BSC is considered as the feedback channel.

Consider a cooperative MIMO system with n_r relay nodes that employ an AF cooperative strategy as well as a DSTC scheme. All nodes have N antennas to transmit and receive. We consider only one user at the source node in our system that operates in a spatial multiplexing configuration. Let $\boldsymbol{s}[i]$ denotes the transmitted information symbol vector at the source node, which contains N parameters, $\boldsymbol{s}[i] = [s_1[i], s_2[i], ..., s_N[i]]$, and has a covariance matrix $E\big[\boldsymbol{s}[i]\boldsymbol{s}^H[i]\big] = \sigma_s^2 \boldsymbol{I}_N$, where σ_s^2 is the signal power which we assume to be equal to 1. The source node broadcasts $\boldsymbol{s}[i]$ from the source to n_r relay nodes as well as to the destination node in the first hop, which can be described by

$$\boldsymbol{r}_{SD}[i] = \boldsymbol{H}_{SD}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SD}[i], \quad \boldsymbol{r}_{SR_k}[i] = \boldsymbol{F}_{SR_k}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SR_k}[i], \tag{4.1}$$

$$i = 1, 2, \ ... \ , N, \ \ k = 1, 2, \ ... \ n_r,$$

where $\boldsymbol{r}_{SR_k}[i]$ and $\boldsymbol{r}_{SD}[i]$ denote the received symbol vectors at the kth relay node and at the destination node, respectively. The $N \times 1$ vector $\boldsymbol{n}_{SR_k}[i]$ and $\boldsymbol{n}_{SD}[i]$ denote the zero mean complex circular symmetric AWGN vector generated at the kth relay node and at the destination node with variance σ^2. The matrices $\boldsymbol{F}_{SR_k}[i]$ and $\boldsymbol{H}_{SD}[i]$ are the $N \times N$ channel coefficient matrices between the source node and the kth relay node, and between the source node and the destination node, respectively.

The received symbols are re-encoded at each relay node prior to transmission to the destination node in the second hop. We assume that the synchronization at each node is perfect. After amplifying the received vector $\boldsymbol{r}_{SR_k}[i]$ at the kth relay node, the signal vector $\tilde{\boldsymbol{s}}_{SR_k}[i] = \boldsymbol{A}_{R_kD}[i]\boldsymbol{r}_{SR_k}[i]$ can be obtained, where $\boldsymbol{A}_{R_kD}[i]$ stands for the $N \times N$ diagonal amplification matrix assigned at the kth relay node. The $N \times 1$ signal vector $\tilde{\boldsymbol{s}}_{SR_k}[i]$ will be re-encoded by an $N \times T$ DSTC scheme $\boldsymbol{M}(\tilde{\boldsymbol{s}})$, multiplied by an $N \times N$ adjustable code matrix $\boldsymbol{\Phi}_k[i]$ generated randomly [6], and then forwarded to the destination node. The advantage of multiplying a randomized matrix or vector at the relays, such as reduction of the error probability and full diversity achievement, are introduced and discussed in [6]. The relationship between the kth relay and the destination node can be described as

$$\boldsymbol{R}_{R_kD}[i] = \boldsymbol{G}_{R_kD}[i]\boldsymbol{\Phi}_k[i]\boldsymbol{M}_{R_kD}[i] + \boldsymbol{N}_{R_kD}[i]. \tag{4.2}$$

The $N \times T$ received symbol matrix $\boldsymbol{R}_{R_kD}[i]$ in (3.2) can be written as an $NT \times 1$ vector $\boldsymbol{r}_{R_kD}[i]$ given by

$$\boldsymbol{r}_{R_kD}[i] = \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] + \boldsymbol{n}_{R_kD}[i], \tag{4.3}$$

where the block diagonal $NT \times NT$ matrix $\boldsymbol{\Phi}_{eq_k}[i]$ denotes the equivalent adjustable code matrix and the $NT \times N$ matrix $\boldsymbol{G}_{eq_k}[i]$ stands for the equivalent channel matrix which is the DSTC scheme $\boldsymbol{M}(\tilde{\boldsymbol{s}}[i])$ combined with the channel matrix $\boldsymbol{G}_{R_kD}[i]$. The $NT \times 1$ equivalent noise vector $\boldsymbol{n}_{R_kD}[i]$ generated at the destination node contains the noise parameters in $\boldsymbol{N}_{R_kD}[i]$.

The utilization of an adjustable code matrix or a randomized matrix $\boldsymbol{\Phi}_{eq_k}[i]$ which achieves the full diversity order and provides a lower error probability has been discussed in [6]. The uniform sphere randomized matrix which achieves the lowest BER of the analyzed schemes and contains elements that are uniformly distributed on the surface of a complex hyper-sphere of radius ρ is used in our system. The proposed adaptive

algorithms detailed in the next section optimize the code matrices employed at the relay nodes in order to achieve a lower BER. At each relay node, the adjustable code matrices are normalized so that no increase in the energy is introduced at the relay nodes and the comparison between different schemes is fair.

After rewriting $\boldsymbol{R}_{R_kD}[i]$ we can consider the received symbol vector at the destination node as a $(T+1)N \times 1$ vector with two parts, one is from the source node and another one is the superposition of the received vectors from each relay node. Therefore, the received symbol vector for the cooperative MIMO system can be written as

$$\begin{aligned} \boldsymbol{r}[i] &= \begin{bmatrix} \boldsymbol{H}_{SD}[i]\boldsymbol{s}[i] \\ \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] \end{bmatrix} + \begin{bmatrix} \boldsymbol{n}_{SD}[i] \\ \boldsymbol{n}_{RD}[i] \end{bmatrix} \\ &= \boldsymbol{D}_D[i]\tilde{\boldsymbol{s}}_D[i] + \boldsymbol{n}_D[i], \end{aligned} \tag{4.4}$$

where the $(T+1)N \times 2N$ block diagonal matrix $\boldsymbol{D}_D[i]$ denotes the channel gain matrix of all the links in the network which contains the $N \times N$ channel coefficients matrix $\boldsymbol{H}_{SD}[i]$ between the source node and the destination node, the $NT \times N$ equivalent channel matrix $\boldsymbol{G}_{eq_k}[i]$ for $k = 1, 2, ..., n_r$ between each relay node and the destination node. We assume that the coefficients in all channel matrices are independent and remain constant over the transmission. The $(T+1)N \times 1$ noise vector $\boldsymbol{n}_D[i]$ contains the equivalent received noise vector at the destination node, which can be modeled as an AWGN with zero mean and covariance matrix $\sigma^2(1+ \parallel \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_{R_kD}[i] \parallel_F^2)\boldsymbol{I}_{(T+1)N}$.

4.3 Joint Adaptive Code Matrix Optimization and Receiver Design

In this section, we jointly design an MMSE adjustable code matrix and the receiver for the proposed DSTC scheme. Adaptive SG and RLS algorithms [59] for determining the parameters of the adjustable code matrix with reduced complexity are also devised. The DSTC scheme used at the relay nodes employs an MMSE-based adjustable code matrix, which is computed at the destination node and obtained by a feedback channel in order to process the data symbols prior to transmission to the destination node. It is worth to

mention that the code matrices are only used at the relay node so the direct link from the source node to the destination node is not considered in this section.

4.3.1 Linear MMSE Receiver Design with Adaptive DSTC Optimization

The linear MMSE receiver design with optimal code matrices at the destination node is derived as follows. By defining the $TN \times 1$ parameter vector $\boldsymbol{w}_j[i]$ to determine the jth symbol $s_j[i]$, we propose the MSE based optimization with a power constraint described by

$$[\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_k}[i]] = \arg \min_{\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_k}[i]} E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right], \; s.t. \; \mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^H[\mathrm{i}]) \leq \mathrm{P_R}, \tag{4.5}$$

where $\boldsymbol{r}[i]$ denotes the received symbol vector at the destination node derived in (4.4). By employing a Lagrange multiplier λ we can obtain the Lagrangian function expression shown as

$$\mathscr{L} = E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right] + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^H[\mathrm{i}]) - \mathrm{P_R}). \tag{4.6}$$

By expanding the right-hand side of (4.6) and taking the gradient with respect to $\boldsymbol{w}_j^*[i]$ and equating the terms to zero, we can obtain the jth MMSE receive filter vector for the jth symbol

$$\boldsymbol{w}_j[i] = \boldsymbol{R}^{-1}\boldsymbol{p}, \tag{4.7}$$

where the first term $\boldsymbol{R} = E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]$ denotes the auto-correlation matrix and the second term $\boldsymbol{p} = E\left[\boldsymbol{r}[i]s_j^*[i]\right]$ stands for the cross-correlation vector. To optimize the code matrix $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ for each symbol at each relay node, we can calculate the code matrix by taking the gradient with respect to $\boldsymbol{\Phi}_{eq_{k_j}}^*[i]$ and equating the terms to zero, resulting in

$$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{P}}, \tag{4.8}$$

where $\tilde{\boldsymbol{R}} = E\left[s_j[i]\tilde{s}_{SR_{k_j}}[i]\boldsymbol{w}_j[i]\boldsymbol{w}_j^H[i] + \lambda\boldsymbol{I}\right]$ and $\tilde{\boldsymbol{P}} = E\left[s_j[i]\tilde{s}_{SR_{k_j}}[i]\boldsymbol{w}_j[i]\boldsymbol{g}_{eq_{k_j}}^H[i]\right]$ are $NT \times NT$ matrices. The value of the Lagrange multiplier λ can be determined by substi-

tuting $\mathbf{\Phi}_{eq_{k_j}}[i]$ into $\lambda \mathrm{Tr}(\mathbf{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\mathbf{\Phi}^{H}_{\mathrm{eq_k}}[\mathrm{i}]) = \mathrm{P_R}$ and solving the power constraint function. In the proposed adaptive algorithm we employ quantization, which requires less computational complexity, instead of using the Lagrange multiplier. The detailed explanation is shown in the next section. Note that non-linear detection algorithms [83] can also be employed at the receiver for an improved performance.

Appendix A includes a detailed derivation of $\boldsymbol{w}_j[i]$ and $\mathbf{\Phi}_{eq_j}[i]$. The power constraint can be enforced by employing the Lagrange multiplier and by substituting the power constraint into the MSE cost function. In (4.8) a closed-form expression of the code matrix $\mathbf{\Phi}_{eq_{k_j}}[i]$ assigned for the jth received symbol at the kth relay node is derived. The problem is that the optimization method requires the calculation of a matrix inversion with a high computational complexity of $\mathrm{O}((NT)^3)$, and with the increase in the number of antennas employed at each node or the use of more complicated STC encoders at the relay nodes, the computational complexity increases cubically according to the matrix sizes in (4.7) and (4.8).

4.3.2 Adaptive Stochastic Gradient Optimization Algorithm

In order to reduce the computational complexity and achieve an optimal performance, a centralized adaptive robust matrix optimization (C-ARMO) algorithm based on an SG algorithm with a linear receiver design is proposed as follows.

The Lagrangian resulting from the optimization problem is derived in (4.6), and a simple adaptive algorithm for determining the linear receive filters and the code matrices can be derived by taking the instantaneous gradient term of (4.6) with respect to $\boldsymbol{w}^*_j[i]$ and with respect to $\mathbf{\Phi}^*_{eq_{k_j}}[i]$, respectively, which are

$$\begin{aligned} \nabla \mathscr{L}_{\boldsymbol{w}^*_j[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}^H_j[i]\boldsymbol{r}[i]\|^2\right]_{\boldsymbol{w}^*_j[i]} = -e^*_j[i]\boldsymbol{r}[i], \\ \nabla \mathscr{L}_{\mathbf{\Phi}^*_{eq_{k_j}}[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}^H_j[i]\boldsymbol{r}[i]\|^2\right]_{\mathbf{\Phi}^*_{eq_{k_j}}[i]} = -e_j[i]s^*_j[i]\boldsymbol{w}_j[i]\boldsymbol{d}^H_{k_j}[i], \end{aligned} \tag{4.9}$$

where $e_j[i] = s_j[i] - \boldsymbol{w}^H_j[i]\boldsymbol{r}[i]$ stands for the jth error signal, and the $NT \times 1$ vector $\boldsymbol{d}_{k_j}[i]$ denotes the jth column of the channel matrix which contains the product of the channel matrices $\boldsymbol{F}_{SR_k}$ and $\boldsymbol{G}_{R_kD}$ and the power allocation matrices $\boldsymbol{A}_{R_kD}$. After we

Table 4.1: Summary of the C-ARMO SG Algorithm

1:	Initialize: $\boldsymbol{w}_j[0] = \mathbf{0}_{NT\times 1}$,
2:	$\boldsymbol{\Phi}[0]$ is generated randomly with the power constraint $\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}\boldsymbol{\Phi}_{\mathrm{eq_k}}^H) \leq \mathrm{P_R}$.
3:	For each instant of time, i=1, 2, ..., compute
4:	$\nabla\mathscr{L}_{\boldsymbol{w}_j^*[i]} = -e_j^*[i]\boldsymbol{r}[i]$,
5:	$\nabla\mathscr{L}_{\boldsymbol{\Phi}_{eq_{k_j}}^*[i]} = -e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}_{k_j}^H[i]$,
6:	where $e_j[i] = s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]$.
7:	Update $\boldsymbol{w}_j[i]$ and $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ by
8:	$\boldsymbol{w}_j[i+1] = \boldsymbol{w}_j[i] + \beta(e_j^*[i]\boldsymbol{r}[i])$,
9:	$\boldsymbol{\Phi}_{eq_{k_j}}[i+1] = \boldsymbol{\Phi}_{eq_{k_j}}[i] + \mu(e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}_{k_j}^H[i])$,
10:	$\boldsymbol{\Phi}_{eq_{k_j}}[i+1] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}_{eq_{k_j}}[i+1]}{\sqrt{\sum_{j=1}^{N}\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}[\mathrm{i+1}]\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}^H[\mathrm{i+1}])}}$.

obtain (4.9) the proposed algorithm can be obtained by introducing a step size into a gradient optimization algorithm to update the result until the convergence is reached, and the algorithm is given by

$$\boldsymbol{w}_j[i+1] = \boldsymbol{w}_j[i]+\beta(e_j^*[i]\boldsymbol{r}[i]), \quad \boldsymbol{\Phi}_{eq_{k_j}}[i+1] = \boldsymbol{\Phi}_{eq_{k_j}}[i]+\mu(e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}_{k_j}^H[i]), \quad (4.10)$$

where β and μ denote the step sizes in the recursions for the estimation. A detailed derivation is included in Appendix B.

The energy of the code matrices in (4.10) will be increased with the processing of the adaptive algorithm, which will contribute to the reduction of the error probability. A normalization of the code matrix after the optimization is required and implemented as $\boldsymbol{\Phi}_{eq_{k_j}}[i+1] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}_{eq_{k_j}}[i+1]}{\sqrt{\sum_{j=1}^{N}\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}[\mathrm{i+1}]\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}^H[\mathrm{i+1}])}}$ to ensure that the energy is not increased and for a fair comparison among the analyzed DSTC schemes. A summary of the C-ARMO SG algorithm is given in Table 4.1.

According to (4.10), the receive filter $\boldsymbol{w}_j[i]$ and the code matrix $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ depend on each other. Therefore, alternating optimization algorithms [84, 97] can be used to determine the linear MMSE receive filter and the code matrix iteratively, and the optimization procedure can be completed. The complexity of calculating the optimal $\boldsymbol{w}_j[i]$ and $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ is $\mathrm{O}(NT)$ and $\mathrm{O}(N^2T^2)$, respectively, which is much less than $O(N^4T^4)$ and $O(N^5T^5)$ by using (4.7) and (4.8). As mentioned in Section I, the optimal MMSE code matrices will be sent back to the relay nodes via a feedback channel, and the influence of the imperfect feedback is shown and discussed in simulations.

4.3.3 ML Detection and LS Code Matrix Estimation Algorithm

The criterion for optimizing the adjustable code matrices and performing symbol detection in the C-ARMO algorithm can be changed to the ML criterion, which is equivalent to a LS criterion in this case. For example, if we take the ML instead of the MSE criterion to determine the code matrices, then we have to store an $N \times D$ matrix $\boldsymbol{S}$ at the destination node which contains all the possible combinations of the transmitted symbol vectors. The ML optimization problem can be written as

$$\begin{aligned} &[\hat{s}_{d_j}[i], \hat{\mathbf{\Phi}}_{eq_{k_j}}[i]] = \arg \min_{s_{d_j}[i], \mathbf{\Phi}_{eq_{k_j}}[i]} \|\boldsymbol{r}[i] - \hat{\boldsymbol{r}}[i]\|^2, \\ &s.t.\ \mathrm{Tr}(\mathbf{\Phi}_{\mathrm{eqk}}[\mathrm{i}]\mathbf{\Phi}^H_{\mathrm{eqk}}[\mathrm{i}]) \leq \mathrm{P_R},\ \text{for d} = 1, 2, ..., \mathrm{D}, \end{aligned} \tag{4.11}$$

where $\hat{\boldsymbol{r}}[i] = \sum_{k=1}^{n_r} \sum_{j=1}^{N} \mathbf{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]\hat{s}_{d_j}[i]$ denotes the received symbol vector without noise which is determined by substituting each column of $\boldsymbol{S}$ into (4.11). The optimization algorithm contains a discrete part which refers to the ML detection and a continuous part which refers to the optimization of the code matrix. The detection and the optimization can be implemented separately. The jointly detection and estimation procedure benefits the optimization of coding design, but the computational complexity will increase. In this case, other detectors such as sphere decoders can be used in the optimization algorithm in the detection part in order to reduce the computational complexity without an impact on the performance, and the algorithm will converge after several iterations.

After determining the transmitted symbol vector, we can calculate the optimal code matrix $\mathbf{\Phi}_{eq_{k_j}}[i]$ by employing the LS estimation algorithm. The Lagrangian expression is given by

$$\mathscr{L} = \|\boldsymbol{r}[i] - (\sum_{k=1}^{n_r} \sum_{j=1}^{N} \mathbf{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]\hat{s}_{d_j}[i])\|^2 + \lambda(Tr[\mathbf{\Phi}_{eq_k}[i]\mathbf{\Phi}^H_{eq_k}[i]] - P_R), \tag{4.12}$$

and by taking the instantaneous gradient of $\mathscr{L}$ with respect to the code matrix $\mathbf{\Phi}^*_{eq_{k_j}}[i]$ we can obtain

$$\begin{aligned} \nabla\mathscr{L}_{\mathbf{\Phi}^*_{eq_{k_j}}[i]} &= (\boldsymbol{r}[i] - \hat{\boldsymbol{r}}[i])\nabla_{\mathbf{\Phi}^*_{eq_{k_j}}[i]}(\boldsymbol{r}[i] - \hat{\boldsymbol{r}}[i])^H \\ &= (\boldsymbol{r}_{e_j}[i] - \mathbf{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]\hat{s}_{d_j}[i])(-\hat{s}^*_{d_j}[i]\boldsymbol{d}\boldsymbol{H}_{k_j}[i]), \end{aligned} \tag{4.13}$$

where $\boldsymbol{r}_{e_j}[i] = \boldsymbol{r}[i] - \sum_{k=1}^{n_r}\sum_{l=1,l\neq j}^{N}\boldsymbol{\Phi}_{eq_{k_l}}[i]\boldsymbol{d}_{k_l}[i]\hat{s}_{d_l}[i]$ stands for the received vector without the desired code matrix. The optimal code matrix $\hat{\boldsymbol{\Phi}}_{eq_{k_j}}[i]$ requires $\nabla\mathscr{L}_{\boldsymbol{\Phi}^*_{eq_{k_j}}[i]} = 0$, and the optimal adjustable code matrix as given by

$$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \hat{s}^*_{d_j}[i]\boldsymbol{r}_{e_j}[i]\boldsymbol{d}^H_{k_j}[i](\mid \hat{s}_{d_j}[i] \mid^2 \boldsymbol{d}_{k_j}[i]\boldsymbol{d}^H_{k_j}[i])^{\dagger}. \tag{4.14}$$

The power constraint is not considered because the quantization method can be employed in order to reduce the high computational complexity for determining the value of the Lagrange multiplier.

4.3.4 RLS Code Matrix Estimation Algorithm

The RLS estimation algorithm for the code matrix $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ is derived in this section. The ML detector is employed so that the detection and the optimization procedures are separate as explained in the previous section, so we focus on how to optimize the code matrix rather than the detection. The superior convergence behavior of the LS algorithm when the size of the adjustable code matrix is large indicates the reason of the utilization of an RLS estimation, and it is worth to mention that the computational complexity reduces from cubic to square by employing the RLS algorithm.

According to the RLS algorithm, the optimization problem is given by

$$\begin{aligned}[\hat{\boldsymbol{\Phi}}_{eq_{k_j}}[i]] &= \arg\min_{\boldsymbol{\Phi}_{eq_{k_j}}[i]}\sum_{n=1}^{i}\lambda^{i-n}\|\boldsymbol{r}[n] - \hat{\boldsymbol{r}}[i]\|^2, \\ s.t.\ &\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^H_{\mathrm{eq_k}}[\mathrm{i}]) \leq \mathrm{P_R},\end{aligned} \tag{4.15}$$

where λ stands for the forgetting factor. By expanding the right-hand side of (3.15) and taking gradient with respect to $\boldsymbol{\Phi}^*_{eq_{k_j}}[i]$ and equaling the terms to zero, we obtain

$$\boldsymbol{\Phi}_{eq_{k_j}}[i] = (\sum_{n=1}^{i}\lambda^{i-n}\boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n])(\sum_{i=1}^{n}\lambda^{i-n}\boldsymbol{r}_{k_j}[n]\boldsymbol{r}^H_{k_j}[n])^{-1}, \tag{4.16}$$

where the $NT \times 1$ vector $\boldsymbol{r}_e[n] = \boldsymbol{\Phi}_{eq_{k_j}}[n]\boldsymbol{d}_{k_j}[n]\hat{s}_{d_j}[n]$ and $\boldsymbol{r}_{k_j}[n] = \boldsymbol{d}_{k_j}[n]\hat{s}_{d_j}[n]$. The

Table 4.2: Summary of the C-ARMO RLS Algorithm

1:	Initialize: $\boldsymbol{P}[0] = \delta^{-1}\boldsymbol{I}_{NT\times NT}$, $\boldsymbol{Z}[0] = \boldsymbol{I}_{NT\times NT}$,
2:	the value of δ is small when SNR is high and is large when SNR is low,
3:	$\boldsymbol{\Phi}[0]$ is generated randomly with the power constraint $\mathrm{trace}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^H_{\mathrm{eq_k}}[\mathrm{i}]) \leq \mathrm{P_R}$.
4:	For each instant of time, i=1, 2, ..., compute
5:	$\boldsymbol{k}[i] = \frac{\lambda^{-1}\boldsymbol{\Psi}^{-1}[i-1]\boldsymbol{r}_{k_j}[i]}{1+\lambda^{-1}\boldsymbol{r}^H_{k_j}[i]\boldsymbol{\Psi}^{-1}[i-1]\boldsymbol{r}_{k_j}[i]}$,
6:	$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \boldsymbol{\Phi}_{eq_{k_j}}[i-1] + \lambda^{-1}(\boldsymbol{r}_e[i] - \boldsymbol{Z}[i-1]\boldsymbol{k}[i])\boldsymbol{r}^H_{k_j}[i]\boldsymbol{P}[i-1]$,
7:	$\boldsymbol{P}[i] = \lambda^{-1}\boldsymbol{P}[i-1] - \lambda^{-1}\boldsymbol{k}[i]\boldsymbol{r}^H_{k_l}[i]\boldsymbol{P}[i-1]$,
8:	$\boldsymbol{Z}[i] = \lambda\boldsymbol{Z}[i-1] + \boldsymbol{r}_e[i]\boldsymbol{r}^H_{k_j}[i]$.
12:	$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}_{eq_{k_j}}[i]}{\sqrt{\sum_{j=1}^{N}\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}[\mathrm{i}]\boldsymbol{\Phi}^H_{\mathrm{eq_{k_j}}}[\mathrm{i}])}}$.

power constraint is still not considered during the optimization. We define

$$\boldsymbol{\Psi}[i] = \sum_{n=1}^{i}\lambda^{i-n}\boldsymbol{r}_{k_j}[n]\boldsymbol{r}^H_{k_j}[n] = \lambda\boldsymbol{\Psi}[i-1] + \boldsymbol{r}_{k_j}[n]\boldsymbol{r}^H_{k_j}[n], \tag{4.17}$$

$$\boldsymbol{Z}[i] = \sum_{n=1}^{i}\lambda^{i-n}\boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n] = \lambda\boldsymbol{Z}[i-1] + \boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n], \tag{4.18}$$

so that we can rewrite (4.16) as

$$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \boldsymbol{Z}[i]\boldsymbol{\Psi}^{-1}[i]. \tag{4.19}$$

By employing the matrix inversion lemma in [15], we can obtain

$$\boldsymbol{\Psi}^{-1}[i] = \lambda^{-1}\boldsymbol{\Psi}^{-1}[i-1] - \lambda^{-1}\boldsymbol{k}[i]\boldsymbol{r}^H_{k_j}[i]\boldsymbol{\Psi}^{-1}[i-1], \tag{4.20}$$

where $\boldsymbol{k}[i] = (\lambda^{-1}\boldsymbol{\Psi}^{-1}[i-1]\boldsymbol{r}_{k_j}[i])/(1+\lambda^{-1}\boldsymbol{r}^H_{k_j}[i]\boldsymbol{\Psi}^{-1}[i-1]\boldsymbol{r}_{k_j}[i])$. We define $\boldsymbol{P}[i] = \boldsymbol{\Psi}^{-1}[i]$ and by substituting (3.18) and (4.20) into (4.19), the expression of the code matrix is given by

$$\begin{aligned}\boldsymbol{\Phi}_{eq_{k_j}}[i] &= \lambda\boldsymbol{Z}[i-1]\boldsymbol{P}[i] + \boldsymbol{r}_e[i]\boldsymbol{r}^H_{k_j}[i]\boldsymbol{P}[i] \\ &= \boldsymbol{Z}[i-1]\boldsymbol{P}[i-1] + \boldsymbol{Z}[i-1]\boldsymbol{k}[i]\boldsymbol{r}^H_{k_j}[i]\boldsymbol{P}[i-1] + \boldsymbol{r}_e[i]\boldsymbol{r}^H_{k_j}[i]\boldsymbol{P}[i] \\ &= \boldsymbol{\Phi}_{eq_{k_j}}[i-1] + \lambda^{-1}(\boldsymbol{r}_e[i] - \boldsymbol{Z}[i-1]\boldsymbol{k}[i])\boldsymbol{r}^H_{k_j}[i]\boldsymbol{P}[i-1].\end{aligned} \tag{4.21}$$

Table 4.2 shows a summary of the C-ARMO RLS algorithm.

4.3.5 Convergence Analysis

The C-ARMO algorithms can be divided into two cases: the first one performs the optimization by updating the receive filter and the code matrix iteratively, i.e., the MSE based C-ARMO algorithm, the second one only optimizes the code matrix itself according to the Lagrangian function, i.e., the ML and RLS based C-ARMO algorithms. In this subsection, we will illustrate how the C-ARMO algorithms converge to the global optimum solution.

MSE based C-ARMO Algorithm

The proposed MSE based C-ARMO algorithm allows the optimization of the receive filter $\boldsymbol{w}[i]$ and the code matrix $\boldsymbol{\Phi}[i]$ iteratively. A detailed proof of the convergence of this type of algorithm is derived in [78]. We will give a brief outline on how these results can be used to prove the convergence of our algorithms.

According to [78], the optimization problem in (4.5) can be described as: Given an initial $(\boldsymbol{w}_0, \boldsymbol{\Phi}_0) \in \mathcal{W}_0 \times \mathcal{P}_0$, we have to find a sequence of points $(\boldsymbol{w}_n, \boldsymbol{\Phi}_n) \in \mathcal{W}_n \times \mathcal{P}_n$ that

$$\lim_{n\to\infty} \mathscr{L}(\boldsymbol{w}_n, \boldsymbol{\Phi}_n) = \mathscr{L}(\mathcal{W}, \mathcal{P}), \tag{4.22}$$

where the sequence of compact sets $\{(\mathcal{W}_n, \mathcal{P}_n)\}_{n\geq 0} : \mathcal{W}_n, \mathcal{P}_n$ that are revealed at time n such that as $n \to \infty$, $\mathcal{P}_n \overset{d_H}{\to} \mathcal{P}$ and $\mathcal{W}_n \overset{d_H}{\to} \mathcal{W}$, and

$$d_H(\mathcal{A}, \mathcal{B}) = \max\{\sup_{A\in\mathcal{A}} \inf_{B\in\mathcal{B}} d(A, B), \sup_{B\in\mathcal{B}} \inf_{A\in\mathcal{A}} d(A, B)\} \tag{4.23}$$

denotes the Hausdorff distance between $\mathcal{A}$ and $\mathcal{B}$. The proposed algorithm is written recursively for $n \geq 1$ as described by

$$\boldsymbol{w}_n \in \arg\min_{\boldsymbol{w}\in\mathcal{W}_n} \mathscr{L}(\boldsymbol{w}, \boldsymbol{\Phi}_{n-1}),\ \boldsymbol{\Phi}_n \in \arg\min_{\boldsymbol{\Phi}\in\mathcal{P}_n} \mathscr{L}(\boldsymbol{w}_n, \boldsymbol{\Phi}). \tag{4.24}$$

According to the three-point property and the four-point property in [78], we can obtain

$$\mathscr{L}(\boldsymbol{w}_n, \boldsymbol{\Phi}_n) + \mathscr{L}(\boldsymbol{w}, \boldsymbol{\Phi}_n) \leq \mathscr{L}(\boldsymbol{w}, \boldsymbol{\Phi}_{n-1}) + \mathscr{L}(\boldsymbol{w}, \boldsymbol{\Phi}) + \omega(\gamma_n) \tag{4.25}$$

for all $\boldsymbol{w} \in \mathcal{W}_n$ and $\boldsymbol{\Phi} \in \mathcal{P}_n$, where $\omega(\gamma_n)$ denotes the modulus of continuity of $\mathscr{L}(\boldsymbol{w}_n, \boldsymbol{\Phi}_n)$ with $\omega(\gamma_n) \to 0$ as $\gamma_n \to 0$, and $\gamma_n = \varepsilon_n + \varepsilon_{n-1}$ with $\varepsilon_n \to 0$ as $n \to \infty$.

Since $\mathcal{P}_n \overset{d_H}{\to} \mathcal{P}$ and $\mathcal{W}_n \overset{d_H}{\to} \mathcal{W}$, there must exists a sequence $(\boldsymbol{w}'_n, \boldsymbol{\Phi}'_n) \in \mathcal{W}_n \times \mathcal{P}_n$ such that $(\boldsymbol{w}'_n, \boldsymbol{\Phi}'_n) \to (\boldsymbol{w}', \boldsymbol{\Phi}') \in \arg\min \mathscr{L}(\mathcal{W}, \mathcal{P})$ and $d(\boldsymbol{w}'_n, \boldsymbol{w}') + d(\boldsymbol{\Phi}'_n, \boldsymbol{\Phi}') \leq \varepsilon_n$ for all $n \geq 0$. By replacing $(\boldsymbol{w}, \boldsymbol{\Phi})$ with $(\boldsymbol{w}', \boldsymbol{\Phi}')$ and choosing $\mathscr{L}(\boldsymbol{w}'_n, \boldsymbol{\Phi}'_n) \leq \mathscr{L}(\boldsymbol{w}', \boldsymbol{\Phi}') + \omega(\varepsilon_n)$, we can obtain

$$\mathscr{L}(\boldsymbol{w}_n, \boldsymbol{\Phi}_n) + \mathscr{L}(\boldsymbol{w}'_n, \boldsymbol{\Phi}_n) \leq \mathscr{L}(\boldsymbol{w}'_{n-1}, \boldsymbol{\Phi}_{n-1}) + \mathscr{L}(\boldsymbol{w}', \boldsymbol{\Phi}') + 2\omega(\gamma_n) + \omega(\varepsilon_n), \quad (4.26)$$

and further derive

$$\liminf_{n\to\infty} \mathscr{L}(\boldsymbol{w}_n, \boldsymbol{\Phi}_n) \leq \mathscr{L}(\boldsymbol{w}, \boldsymbol{\Phi}). \quad (4.27)$$

By defining a subsequence $\{n_k\}_{k>0}$ such that $\liminf_{n\to\infty} \mathscr{L}(\boldsymbol{w}_n, \boldsymbol{\Phi}_n) = \lim_{k\to\infty} \mathscr{L}(\boldsymbol{w}_{n_k}, \boldsymbol{\Phi}_{n_k})$, and assuming compactness of $\mathcal{W}$ and $\mathcal{P}$, we can obtain $\boldsymbol{w} \in \mathscr{W}$, $\boldsymbol{\Phi} \in \mathcal{P}$, and

$$\liminf_{n\to\infty} \mathscr{L}(\boldsymbol{w}_n, \boldsymbol{\Phi}_n) = \lim_{k\to\infty} \mathscr{L}(\boldsymbol{w}_{n_k}, \boldsymbol{\Phi}_{n_k}) \geq \mathscr{L}(\boldsymbol{w}, \boldsymbol{\Phi}). \quad (4.28)$$

Combining (4.27) and (4.28), we can obtain $\liminf_{n\to\infty} \mathscr{L}(\boldsymbol{w}_n, \boldsymbol{\Phi}_n) = \mathscr{L}(\boldsymbol{w}, \boldsymbol{\Phi})$ which indicates (4.22) converges to the optimum values.

ML and RLS based C-ARMO Algorithm

The ML and RLS based C-ARMO algorithms just optimize the code matrix, and we can analyze the Hessian matrix of (4.12) and check its positive (semi-)definiteness. By taking the second-order partial derivatives of the Lagrangian cost function in (4.12), we can obtain

$$\begin{aligned} H(\mathscr{L}) &= \frac{\partial}{\partial \boldsymbol{\Phi}_{eq_{k_j}}[i]} \left(\frac{\partial \mathscr{L}}{\partial \boldsymbol{\Phi}^*_{eq_{k_j}}[i]}\right) \\ &= \frac{\partial}{\partial \boldsymbol{\Phi}_{eq_{k_j}}[i]} \left(\boldsymbol{r}[i] s^*_j[i] \boldsymbol{d}^H_{k_j}[i] + \mid s_j \mid^2 \boldsymbol{\Phi}_{eq_{k_j}}[i] \boldsymbol{d}_{k_j}[i] \boldsymbol{d}^H_{k_j}[i]\right) \\ &= \mid s_j \mid^2 \boldsymbol{d}_{k_j}[i] \boldsymbol{d}^H_{k_j}[i], \end{aligned} \quad (4.29)$$

where the first term $\mid s_j \mid^2$ is a positive scalar and the rest of the terms denotes the multiplication of the equivalent channel vectors which is a positive-definite matrix and the problem is convex. We conclude that the Hessian matrix of the Lagrangian cost function is a positive-definite matrix so that the ML and RLS based C-ARMO algorithms converge to the global optimum under the usual assumptions used to prove the convergence of these algorithms for convex problems.

4.4 Probability of Error Analysis

In this section, the pairwise error probability (PEP) of the system employing the adaptive DSTC will be derived. As we mentioned in Section I, the adjustable code matrices will be considered in the derivation as it affects the performance by reducing the upper bound of the pairwise error probability. The PEP upper bound of the traditional STC schemes in [41] is introduced for comparison, and the main difference lies in the eigenvalues of the adjustable code matrices. Please note that the direct link is ignored in the PEP upper bound derivation in order to concentrate on the effects of the adjustable code matrix on the performance. The expression of the upper bound holds for systems with different sizes and an arbitrary number of relay nodes.

Consider an $N \times N$ STC scheme at the relay node with C_T codewords, and the codeword $\boldsymbol{C}^1$ is transmitted and decoded as another codeword $\boldsymbol{C}^i$ at the destination node, where $i = 1, 2, ..., C_T$. According to [41], the probability of error for this code can be upper bounded by the sum of all the probabilities of incorrect decoding, which is given by

$$\mathrm{P_e} \leq \sum_{i=2}^{C_T} \mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^i). \tag{4.30}$$

Assuming that the codeword $\boldsymbol{C}^2$ is decoded at the destination node and that we know the channel information perfectly, we can derive the conditional pairwise error probability of the DSTC encoded with the adjustable code matrix $\boldsymbol{\Phi}$ as [118]

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{\Phi}) = \mathrm{Q}\left(\sqrt{\frac{\gamma}{2}} \parallel \boldsymbol{\Phi}\boldsymbol{D}(\boldsymbol{C}^1 - \boldsymbol{C}^2) \parallel_F\right), \tag{4.31}$$

where $\boldsymbol{D}$ stands for the matrix with the channel coefficients for all links. Let $\boldsymbol{U}^H\boldsymbol{\Lambda_C}\boldsymbol{U}$ be the eigenvalue decomposition of $(\boldsymbol{C}^1-\boldsymbol{C}^2)^H(\boldsymbol{C}^1-\boldsymbol{C}^2)$, where $\boldsymbol{U}$ is a unitary matrix with the eigenvectors and $\boldsymbol{\Lambda_C}$ is a diagonal matrix which contains all the eigenvalues of the difference between two different codewords $\boldsymbol{C}^1$ and $\boldsymbol{C}^2$. Let $\boldsymbol{V}^H\boldsymbol{\Lambda_\Phi}\boldsymbol{V}$ stand for the eigenvalue decomposition of $(\boldsymbol{\Phi DU})^H\boldsymbol{\Phi DU}$, where $\boldsymbol{V}$ is a unitary matrix that contains the eigenvectors and $\boldsymbol{\Lambda_\Phi}$ is a diagonal matrix with the eigenvalues arranged in decreasing order. Therefore, the conditional pairwise probability of error can be written as

$$\mathrm{P}(\boldsymbol{C}^1\rightarrow\boldsymbol{C}^2\mid\boldsymbol{\Phi})=\mathrm{Q}\left(\sqrt{\frac{\gamma}{2}\sum_{m=1}^{NC_T}\sum_{n=1}^{N}\lambda_{\boldsymbol{\Phi}_n}\lambda_{\boldsymbol{C}_n}|\xi_{n,m}|^2}\right), \tag{4.32}$$

where $\xi_{n,m}$ is the (n,m)th element in $\boldsymbol{V}$, and $\lambda_{\boldsymbol{\Phi}_n}$ and $\lambda_{\boldsymbol{C}_n}$ are the nth eigenvalues in $\boldsymbol{\Lambda_\Phi}$ and $\boldsymbol{\Lambda}_s$, respectively. It is important to note that the value of $\lambda_{\boldsymbol{\Phi}}$ is positive and real because $(\boldsymbol{\Phi DU})^H\boldsymbol{\Phi DU}$ is Hermitian symmetric. According to [41], an appropriate upper bound assumption of the Q function is $\mathrm{Q}(x)\leq\frac{1}{2}e^{\frac{-x^2}{2}}$, thus we can derive the upper bound of the pairwise error probability for an adaptive STC scheme as

$$\begin{aligned}\mathrm{P}_{e_{\boldsymbol{\Phi}}}&\leq E\left[\frac{1}{2}\exp\left(-\frac{\gamma}{4}\sum_{m=1}^{NC_T}\sum_{n=1}^{N}\lambda_{\boldsymbol{\Phi}_n}\lambda_{\boldsymbol{C}_n}|\xi_{n,m}|^2\right)\right]\\&=\frac{1}{\prod_{n=1}^{N}(1+\frac{\gamma}{4}\lambda_{\boldsymbol{\Phi}_n}\lambda_{\boldsymbol{C}_n})^{NT}},\end{aligned} \tag{4.33}$$

while the upper bound of the error probability expression for a traditional STC in [41] is given by

$$\begin{aligned}\mathrm{P}_{e_{\boldsymbol{D}}}&\leq E\left[\frac{1}{2}\exp\left(-\frac{\gamma}{4}\sum_{m=1}^{NC_T}\sum_{n=1}^{N}\lambda_{\boldsymbol{C}_n}|\xi_{n,m}|^2\right)\right]\\&=\frac{1}{\prod_{n=1}^{N}(1+\frac{\gamma}{4}\lambda_{\boldsymbol{C}_n})^{NT}}.\end{aligned} \tag{4.34}$$

If we neglect the 1 in the denominator in (4.33), the exponent of the SNR γ indicates the diversity order which means the full diversity NT can be achieved in (4.33). By comparing (4.33) and (4.34), employing an adjustable code matrix for an STC scheme at the relay node introduces $\lambda_{\boldsymbol{\Phi}_n}$ in the BER upper bound. With the aid of simulations, we found that $\boldsymbol{\Lambda_\Phi}$ is diagonal with one eigenvalue less than 1 and others much greater than 1. As a result, employing the adjustable code matrices can provide a decrease in the BER upper bound since the value in the denominator increases.

4.5 Fully Distributed Adaptive Robust Matrix Optimization Algorithm

Inspired by the analysis developed in the previous section, we derive a fully distributed ARMO (FD-ARMO) algorithm which does not require the feedback channel in this section. We will extend the exact PEP expression in [98] for MIMO communication systems to the AF cooperative MIMO systems with the adaptive DSTC schemes. Then, we design the FD-ARMO algorithm to determine and store the adjustable code matrices at the relay nodes before the transmission in Phase II.

The exact PEP expression of an STC has been given by Taricco and Biglieri in [98], which contains the sum of the real part and the imaginary part of the mean value of the error probability, and the moment generating function (MGF) is employed to compute the mean value. To extend the exact PEP expression to the cooperative MIMO systems, we can rewrite the received symbol vector at the destination node as

$$\boldsymbol{R}_{RD} = \sum_{k=1}^{n_r} \boldsymbol{\Phi}_k[i]\boldsymbol{D}_k[i]\boldsymbol{C}[i] + \boldsymbol{N}_{RD}[i],$$

where $\boldsymbol{D}_k[i]$ denotes the equivalent channel matrix of the network. For simplicity, we assume the synchronization is perfect, and each relay node transmits the STC matrix simultaneously and the received symbol vector at the destination node will be the superposition of each column of each STC code. The equivalent noise vector contains the AWGN at the destination node as well as the amplified and re-encoded noise vectors at the relay nodes. As a result the PEP expression of the AF cooperative MIMO system with the adaptive DSTC can be derived as

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq}) = \mathrm{Q}\left(\frac{\| \boldsymbol{\Phi}\boldsymbol{D}(\boldsymbol{C}^1 - \boldsymbol{C}^2) \|_F}{\sqrt{2N_o}}\right), \tag{4.35}$$

where $N_0 = \mathrm{Tr}(\boldsymbol{I} + \boldsymbol{\Phi}\boldsymbol{D})$ denotes the received noise variance at the destination node. We define $\boldsymbol{\Delta} = \boldsymbol{C}^1 - \boldsymbol{C}^2$ as the distance between the code words, and $\tau = \sqrt{\frac{1}{2N_o}}\boldsymbol{\Phi}\boldsymbol{D}\boldsymbol{\Delta}\boldsymbol{\Delta}^H\boldsymbol{D}^H\boldsymbol{\Phi}^H$ and we assume that the eigenvalue decomposition of $\boldsymbol{\Delta}\boldsymbol{\Delta}^H$ can be written as $\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H$, where $\boldsymbol{V}$ stands for a unitary matrix that contains the eigenvectors of $\boldsymbol{\Delta}\boldsymbol{\Delta}^H$ and $\boldsymbol{\Lambda}$ contains all the eigenvalues of the square of the distance vector. Define

an $N \times N$ matrix $\boldsymbol{Z} = \boldsymbol{\Phi D}$, and $\boldsymbol{Z} \sim N_c(\mu_{\boldsymbol{Z}}, \boldsymbol{\Sigma_Z})$, where $\mu_{\boldsymbol{Z}} = 0$ denotes the mean and $\boldsymbol{\Sigma_Z} = E\left[\boldsymbol{\Sigma_Z \Sigma_Z^H}\right]$ stands for the covariance matrix. The expression of the error probability is given by

$$\begin{aligned}\Theta(c) &= E\left[\exp(-c\xi)\right] = E\left[\exp(-c\sqrt{\frac{1}{2N_o}}[\boldsymbol{\Phi D}\Delta\Delta^H\boldsymbol{D}^H\boldsymbol{\Phi}^H])\right] \\ &= E\left[\exp\left(-c\sqrt{\frac{1}{2N_o}}[\boldsymbol{Z\Lambda Z}^H]\right)\right] = \frac{\exp\left(-\mu_{\boldsymbol{Z}}^H\boldsymbol{B}(\boldsymbol{I}+\Sigma_{\boldsymbol{Z}}\boldsymbol{B})^{-1}\mu_{\boldsymbol{Z}}\right)}{\det\left(\boldsymbol{I}+\frac{c}{2\sqrt{2N_0}}\boldsymbol{\Phi\Lambda\Phi}^H\right)} \\ &= \det\left(\boldsymbol{I}+\frac{c}{2\sqrt{2N_0}}\boldsymbol{\Phi\Lambda\Phi}^H\right)^{-1},\end{aligned} \tag{4.36}$$

where $\boldsymbol{B} = \boldsymbol{I} \bigotimes \boldsymbol{\Delta\Delta}^H$, and $c = a + jb$ is the variable defined in the MGF with $a = \frac{1}{4}$ and b is a constant. By inserting (4.36) into the pairwise error probability expression in [98], we can obtain the exact PEP of the adaptive DSTC scheme written as

$$\mathrm{P}_e = \frac{1}{2J}\sum_{i=1}^{J}\{\Re[\Phi(c)] + \frac{b}{a}\Im[\Phi(c)]\} + \mathrm{E}_J, \tag{4.37}$$

where $\mathrm{E}_J \to 0$ as $J \to \infty$.

Since the PEP is proportional to (4.36), it is clear that minimizing the PEP is equal to maximizing the determinant of $\boldsymbol{I} + \frac{c}{2\sqrt{2N_o}}\boldsymbol{\Phi\Lambda\Phi}^H$. As a result, the optimization problem can be written as

$$\Theta_{opt}(c) = \arg\max_{l}\Theta_l(c), \quad l = 1, 2, ... \tag{4.38}$$

where $\Theta_l(c)$ stands for the lth candidate code matrix. For simplicity the candidate code matrices are generated randomly and satisfy the power constraint. In order to obtain the adjustable code matrix we can first randomly generate a set of matrices, and then substitute them into (4.36) to compute the determinant. In the simulation, we randomly generate 500 code matrices and choose the optimal one according to the FD-ARMO algorithm. The optimal code matrix with the largest value of the determinant which achieves the minimal PEP will be employed at the relay node. A summary of the FD-ARMO is given in Table 4.3.

Table 4.3: Summary of the FD-ARMO Algorithm

1:	Choose the $N \times T$ STC scheme used at the relay node
2:	Determine the dimension of the adjustable code matrix $\mathbf{\Phi}$ which is $N \times N$
3:	Compute the eigenvalue decomposition of $\Delta\Delta^H$ and store the result in Λ
4:	Generate a set of $\mathbf{\Phi}$ randomly with the power constraint $\mathrm{Tr}(\mathbf{\Phi}_{\mathrm{k}}\mathbf{\Phi}_{\mathrm{k}}^H) \leq \mathrm{P_R}$
5:	For all $\mathbf{\Phi}$, compute $\Theta(c) = \det\left(\boldsymbol{I} + \frac{c}{2\sqrt{2N_0}}\mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^H\right)^{-1}$
6:	Choose the code matrix according to $\Theta_{opt}(c) = \arg\max_l \Theta_l(c)$
7:	Store the optimal code matrix $\mathbf{\Phi}_{opt}$ at the relay node

4.6 Simulations

The simulation results are provided in this section to assess the proposed scheme and algorithms. The cooperative MIMO system considered employs an AF protocol with the Alamouti STBC scheme in [5] using QPSK modulation in a quasi-static block fading channel with AWGN. The effect of the direct link is also considered. It is worth to mention that with a simple modification, the proposed ARMO algorithms can be employed by the DF cooperative systems. The number of antennas and the number of relay nodes can also be changed. The system is equipped with $n_r = 1$ relay node and $N = 2$ antennas at each node. In the simulations, we set the symbol power σ_s^2 as equal to 1, and the power of the adjustable code matrix in the ARMO algorithms are normalized. The SNR in the simulations is the received SNR which is calculated by $SNR = \frac{\|\sum_{k=1}^{n_r} \mathbf{\Phi}_{eq_k}[i]\boldsymbol{D}\|_F^2}{1+\|\sum_{k=1}^{n_r} \mathbf{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_{R_k D}[i]\|_F^2}$. It is worth to mention that the constellation scheme and the number of antennas used in this section are different from that in the previous section.

The upper bounds of the D-Alamouti, the R-Alamouti in [6] and the adaptive Alamouti STC in C-ARMO RLS algorithm, derived in previous subsections, are shown in Fig. 4.1. The maximum diversity order is equal to 4. The theoretical pairwise error probabilities provide the largest decoding errors of the three different coding schemes and as shown in the figure, by employing a randomized matrix at the relay node it decreases the decoding error upper bound. The bounds become tighter to the respective coding schemes as the SNR increases. The comparison of the simulation results in a better BER performance of the R-Alamouti and the D-Alamouti which indicates the advantage of using the randomized matrix at relay nodes. The C-ARMO RLS algorithm optimizes the random-

ized matrices after each transmission which contributes to a lower error probability upper bound, and the ML detection algorithm, requires substitution of all possible combinations of the symbols into the cost function, provides the optimal performance at the cost of an exponential increase of computation complexity. In this chapter the computational complexity of ML detection is $\mathrm{O}(TN^3)$.

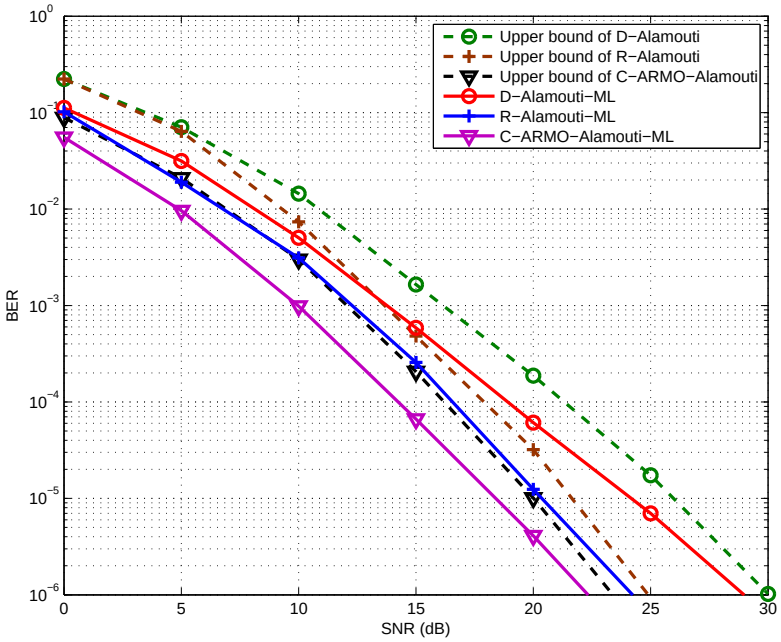

Figure 4.1: BER performance vs. SNR for the upper bound of the Alamouti schemes without the direct link

The proposed C-ARMO SG algorithm with a linear MMSE receiver is compared with the SM scheme and the DSTC algorithms in [13], [99], [6] and [33] in Fig. 4.2. It is worth mentioning that the coding schemes in the simulations are different. In the proposed algorithm and the algorithms in [13], [99] and [6], the information symbols are formed in a spatial multiplexing scheme, sent from the source node to the relays and re-encoded at the relay node. On the other hand, the full-opportunistic code [33] space-time encoded the information symbols at the source node instead of re-encoding at the relay node. The diversity order is about 2 in this figure because of the utilization of the linear detector. The step sizes for the iterative optimization are $\beta = 0.01$ and $\mu = 0.03$, which are chosen according to [100]. The results illustrate that without the direct link, by making use of the STC technique, a significant performance improvement can be achieved compared to the spatial multiplexing system. The RSTC algorithm in [6] outperforms the STC-AF

schemes in [13] and [99], while the C-ARMO SG algorithm can improve the performance by about 3dB as compared to the RSTC algorithm. The STC scheme in [33] achieves a much better performance compared to other schemes although this comparison must be considered with caution. In [33] the standard 2×2 Alamouti STBC is employed at the source node, which indicates the received matrix at the relay node is amplified without the interference to the orthogonality of the code. Moreover, encoding at the source node requires more time slots to transmit so that the transmission rate is half compared to the proposed C-ARMO algorithm. It is also worth to mention that the C-ARMO algorithm can be employed in an opportunistic scheme to achieve a better BER performance as both of the algorithms employ the STCs and can perform the optimization at the destination node. With the consideration of the direct link, the results indicate that the diversity order can be increased, and using the C-ARMO SG algorithm an improved performance is achieved with 2dB of gain as compared to employing the RSTC algorithm in [6] and 3dB of gain as compared to employing the traditional STC-AF algorithm in [13].

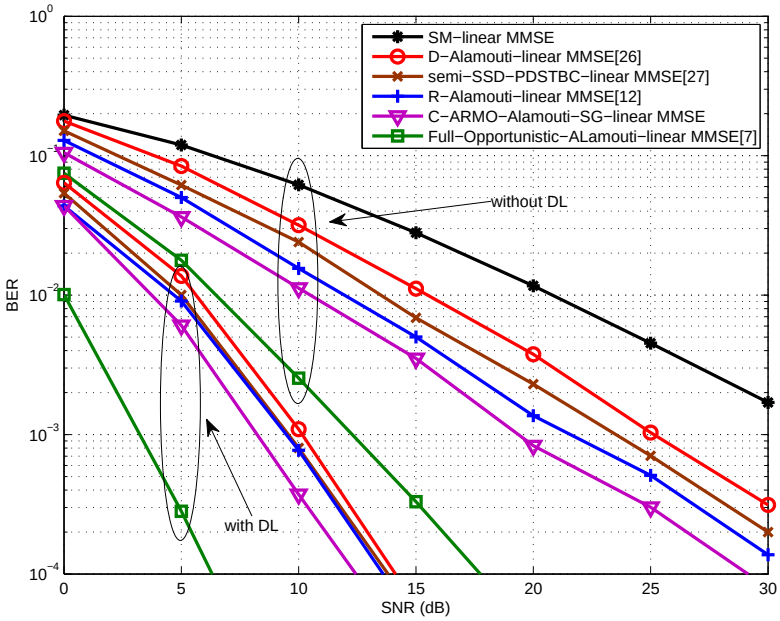

Figure 4.2: BER performance vs. SNR for C-ARMO SG algorithm with and without the direct link

In Fig. 4.3, BER curves of different Alamouti coding schemes and the proposed C-ARMO RLS algorithm with and without the direct link using an ML detector are compared. In Fig. 4.3, the R-Alamouti scheme improves the performance by about 4dB

without the direct link compared to the D-Alamouti scheme, and the C-ARMO RLS algorithm provides a significant improvement in terms of gains compared to the other DSTC schemes. When the direct link is considered, all the coding schemes can achieve the full diversity order and obtain lower BER performances compared to that without the direct link, and still the C-ARMO RLS algorithm which optimizes the adjustable code matrix achieves the lowest BER performance.

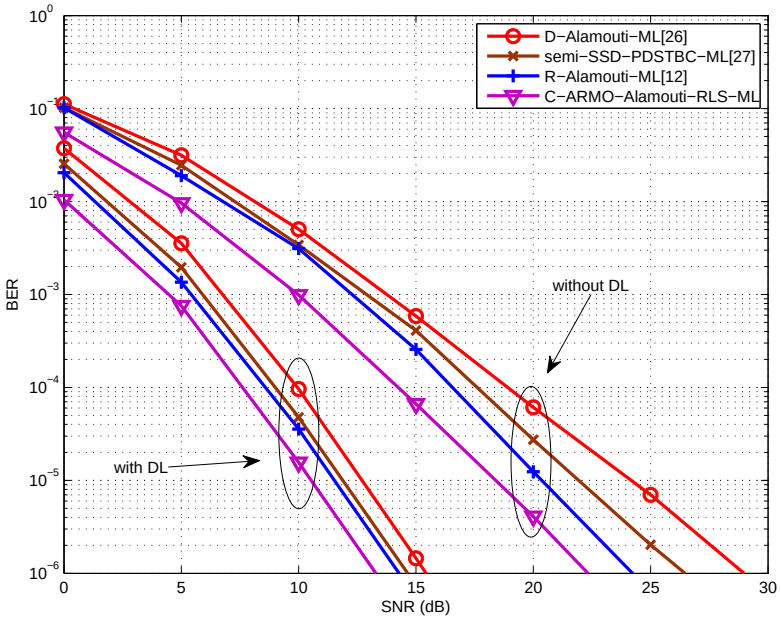

Figure 4.3: BER performance vs. SNR for C-ARMO RLS algorithm with and without the direct link

The simulation results shown in Fig. 4.4 and Fig. 4.6 illustrate the convergence property of the C-ARMO SG algorithm. All the schemes have an error probability of 1 at the beginning because at time slot 0 nothing is transmitted and received and in the first 20 time slots, it is obvious to see the advantage in BER by using the DSTC and RSTC schemes. The proposed C-ARMO-SG algorithm achieves an improvement in BER performance compared to the BER performance of original RSTC scheme. After the first 20 symbols are received and detected, the curves drop quickly to a low BER level. The SNR settles to 30dB in the simulation. The R-Alamouti scheme in [13] achieves a better BER performance compared with the spatial multiplexing scheme and the R-Alamouti scheme in [6] can reach a lower BER than the C-ARMO algorithm. With the number of received symbol increasing, the BER curve of the spatial multiplexing, the D-Alamouti

and the R-Alamouti schemes are almost straight, while the BER performance of the C-ARMO algorithm can be further improved and obtain a fast convergence after receiving 140 symbols.

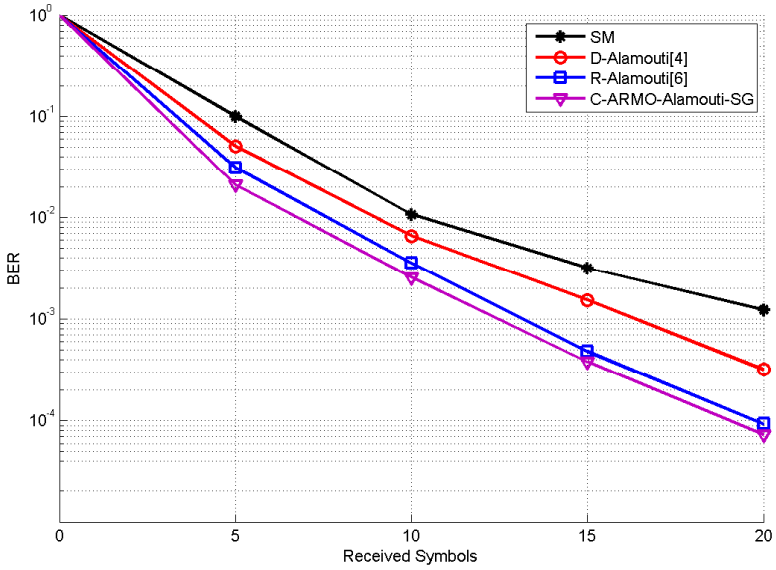

Figure 4.4: BER performance vs. number of samples for C-ARMO SG algorithm without the direct link

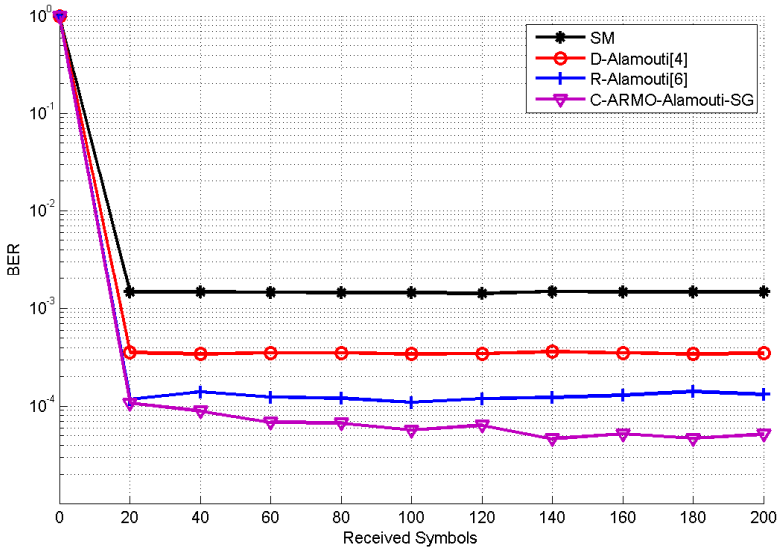

Figure 4.5: BER performance vs. number of samples for C-ARMO SG algorithm without the direct link

The simulation results shown in Fig. 4.6 illustrate the influence of the feedback chan-

nel on the C-ARMO SG algorithm. As mentioned in Section I, the optimized code matrix will be sent back to each relay node through a feedback channel. The quantization and feedback errors are not considered in the simulation results in Fig. 4.2 and Fig. 4.3, so the optimized code matrix is perfectly known at the relay node after the C-ARMO algorithm; while in Fig. 4.6, it indicates that the performance of the proposed algorithm will be affected by the accuracy of the feedback information. In the simulation, we use 4 bits to quantize the real part and the imaginary part of the element in the code matrix $\mathbf{\Phi}_{eq_{kj}}[i]$, and the feedback channel is modeled as a binary symmetric channel with different error probabilities. As we can see from Fig. 4.6, by decreasing the error probabilities for the feedback channel with fixed quantization bits, the BER performance approaches the performance with the perfect feedback, and by making use of 4 quantization bits for the real and imaginary part of each parameter in the code matrix, the performance of the C-ARMO SG algorithm is about 1dB worse with feedback error probability of 10^{-3}.

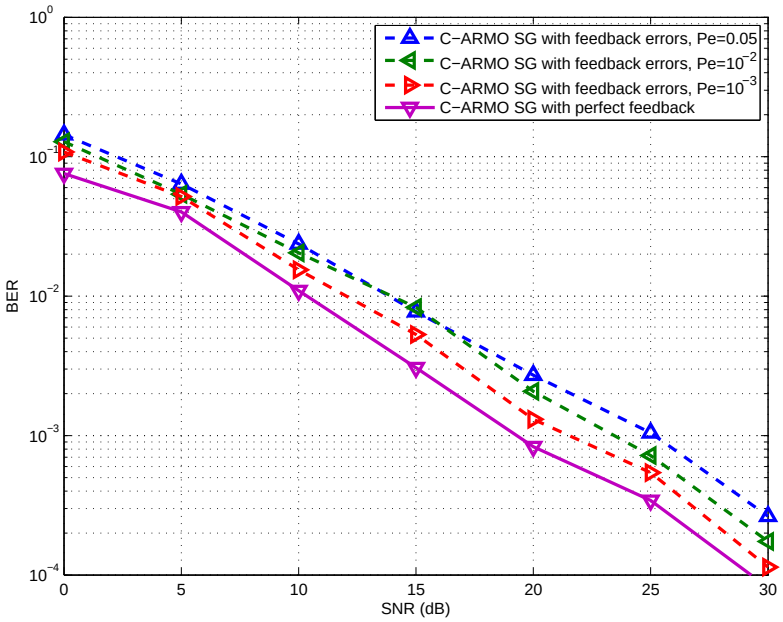

Figure 4.6: BER performance vs. SNR for C-ARMO algorithm with perfect and imperfect feedback links, quantization bits = 4

In Fig. 4.7, we plot the average error probability with respect to the SNR for the FD-ARMO algorithm and the C-ARMO SG algorithm with perfect feedback. The C-ARMO curve and the FD-ARMO curve outperforms the others because they optimize the adjustable code matrices with the same criterion, but 1dB of gain has been obtained

by the C-ARMO SG algorithm because the exact adjustable code matrix is transmitted back to the relay node in a delay-free and error-free feedback channel. While the FD-ARMO chooses the optimal adjustable code matrix by using the statistical information of the channel before transmission so that the performance will be influenced, resulting in a gain less than 1dB.

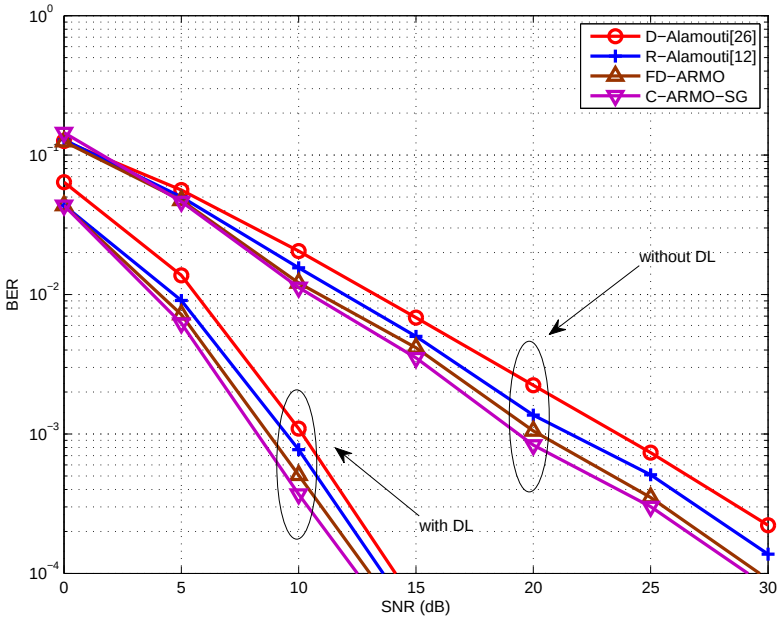

Figure 4.7: Full-distributed ARMO algorithm and C-ARMO SG algorithm

4.7 Summary

We have proposed centralized adaptive robust matrix optimization (C-ARMO) algorithms for the AF cooperative MIMO system using a linear MMSE receive filter and an ML receiver at the destination node. The pairwise error probability of introducing the adaptive DSTC in a cooperative MIMO network with the AF protocol has been derived. In order to eliminate the need for a feedback channel we have derived a fully-distributed ARMO (FD-ARMO) algorithm which can achieve a similar coding gain without the feedback as compared to the C-ARMO algorithms. The simulation results illustrate the advantage of the proposed ARMO algorithms by comparing them with the cooperative network employing the traditional DSTC scheme and the RSTC scheme. The proposed algorithms

can be used with different DSTC schemes using the AF strategy and can also be extended to the DF cooperation protocol.

Chapter 5

Adaptive Delay-Tolerant Distributed Space-Time Coding with Feedback for Cooperative MIMO Relaying Systems

Contents

5.1 Introduction

The application of DSTC schemes at relay nodes in a cooperative network can offer the system diversity and coding gains to mitigate the interference by providing more copies

of the desired symbols at the destination. However, a serious issue for distributed MIMO systems using a DSTC scheme is the asynchronous transmission from the relays or the delayed reception of the DSTC scheme at the destination. The delayed code matrices or vectors forwarded from the relays shift the DSTC structure which results in collapse of rank design criterion and leads to deterioration of decoding performance. In order to tackle the delay problem, various delay-tolerant distributed space-time coding (DT-DSTC) schemes [106] - [110] which address the delay between relay nodes have been recently considered for distributed MIMO systems. Extending the distributed threaded algebraic space-time (TAST) codes [106], Damen and Hammons designed a delay-tolerant coding scheme in [107] by the extension of the Galois field employed in the coding scheme to achieve full diversity and full rate. A further optimization which ensures that the codes in [107] obtain full diversity with the minimum length and lower decoding complexity is presented in [108]. The authors of [109] propose delay-tolerant distributed linear convolutional STC schemes which can maintain the full diversity property under any delay situation among the relays. In [110], a transmit processing technique based on linear constellation precoder (LCP) design is employed to construct an optimal DT-DSTC scheme to achieve the upper bound on the error probability.

In the literature so far, two basic configurations of distributed space-time coding schemes have been reported: one in which the coding is performed independently at the relays [6, 7] and another in which coding is performed across the relays [107, 110]. The received symbol matrix at the destination node in the first configuration is a sum of fully encoded DSTC schemes from each relay node, while in the second configuration the received symbol matrix contains only one DSTC scheme encoded across all the relays. Considering delays between the relay nodes, the sum of the DSTC schemes will be affected by imperfect asynchronous superposition of signals in the first configuration, while the structure of the received DSTC schemes in the second configuration will be switched. It is not clear the key advantages of these schemes and their suitability for situations with delays. In addition, the work on delay-tolerant space-time codes has focused on encoding using a fixed scheme, such as TAST codes in [106], which is difficult to implement with other DSTC schemes. In the literature, the DSTC designs in [14, 32–34, 89, 90, 111, 112] can achieve full diversity order and high coding gains but are vulnerable to the delays. In order to implement the existing DSTC schemes and overcome the delay issues, a general delay-tolerant encoding design and algorithm is needed. Moreover, the DT-DSTC

schemes in the literature do not employ feedback or optimal relay selection algorithms to improve the performance of the systems. The advantage of employing feedback to optimize DSTC schemes have been studied in [7] with various space-time coding schemes. However, strategies to exploit feedback and improve the design of DT-DSTC schemes remain unexplored. Opportunistic cooperation algorithms are reported in [33] and [111]. The authors in [33] provide three relay selection algorithms for distributed systems which can achieve a lower bit error rate (BER) performance compared to the traditional equal power allocation schemes, and an opportunistic resource allocation optimization design which maximizes the delay-limited capacity and minimizes the outage probability is given in [111]. However, the single-antenna relays are employed in the designs so that if a DSTC scheme is employed at the relay node it will be disturbed after selecting the optimal relay or allocating the optimal relay to the best position when delays are considered among the relays.

In this chapter, we propose an adaptive delay-tolerant DSTC scheme and algorithms for cooperative MIMO relaying systems with a feedback channel and different delays among relay nodes. The work is first introduced and discussed in [112]. Two scenarios are considered in the design, one is a cooperative system with single-antenna relay nodes, another is a cooperative system with multi-antenna relay nodes. We first propose a delay-tolerant adjustable code matrices optimization (DT-ACMO) algorithm based on the maximum-likelihood (ML) criterion subject to constraints on transmitted power at the relays for different cooperative systems. Specifically, adaptive optimization algorithms using stochastic gradient (SG) and recursive least square (RLS) estimation methods are developed for the DT-ACMO algorithm in order to release the destination from the high computational complexity of the optimization process. We study how the adjustable code matrices affect the DSTC scheme during the encoding process and how to optimize the adjustable code matrices by employing an ML detector. Then we analyze the differences in terms of the rank criterion and pairwise error probabilities (PEP) of the distributed space-time block coding (DSTBC) schemes in these two system configurations with the same number of antennas and the same delay profiles. We focus on how the different systems affect the delay tolerance of the DSTBC schemes and conclude that the DSTBC schemes in distributed systems with multi-antenna relay nodes can address the delays from relay nodes compared to the cooperative system employing single-antenna relays. After designing the proposed DT-ACMO algorithm for cooperative MIMO systems using

the DF strategy, we extend our design to cooperative systems exploiting the AF protocol and the opportunistic relaying algorithms in [33] in order to achieve a delay-tolerant adjustable code matrices opportunistic relaying optimization (DT-ACMORO) algorithm which can address the delay issue when the optimal power allocation and relay selection is employed. The proposed delay-tolerant designs can be implemented with different types of DSTC schemes in cooperative MIMO relaying systems with DF or AF protocols.

Unlike our prior work [7], we discuss and show the advantages and disadvantages of employing DSTBC in different cooperative systems with different protocols in this chapter. The analyses of the proposed designs indicate the improvement of using the proposed designs in a cooperative MIMO relaying system with asynchronous relaying. The issue of asynchronous reception at the destination node can be addressed by the proposed DT-ACMO algorithms, while the BER performance is improved by the optimization algorithms. Moreover, the combination of the proposed DT-ACMO algorithms and the opportunistic schemes in [33] provides a delay-tolerant DSTC design with optimal relay selection. All the detection errors, including errors in DF protocol and feedback errors at each relay node, have been considered in the algorithm design in order to show the potential of implementation of the proposed algorithms in practical circumstances.

5.2 Cooperative MIMO System Model

We consider a cooperative communication system, which consists of one source node, n_r relay nodes (Relay 1, Relay 2, ..., Relay n_r), a feedback channel and one destination node as shown in Fig. 5.1. Both the source and the destination employ N antennas and can either transmit or receive at one time. The number of antennas at the relays is denoted by B. We consider two types of system configurations in this chapter, one is called multiple antenna system (MAS) configuration which employs relay nodes with $B = N$ antennas, and another one is called single antenna system (SAS) configuration which employs relay nodes with $B = 1$ antenna. As mentioned in the previous section, MAS obtains a sum of the fully-encoded DSTC schemes at the destination and SAS obtains one full DSTC scheme at the destination. The diversity order of MAS and SAS with the same number of distributed antenna elements is the same shown in Section V. A delay profile which is

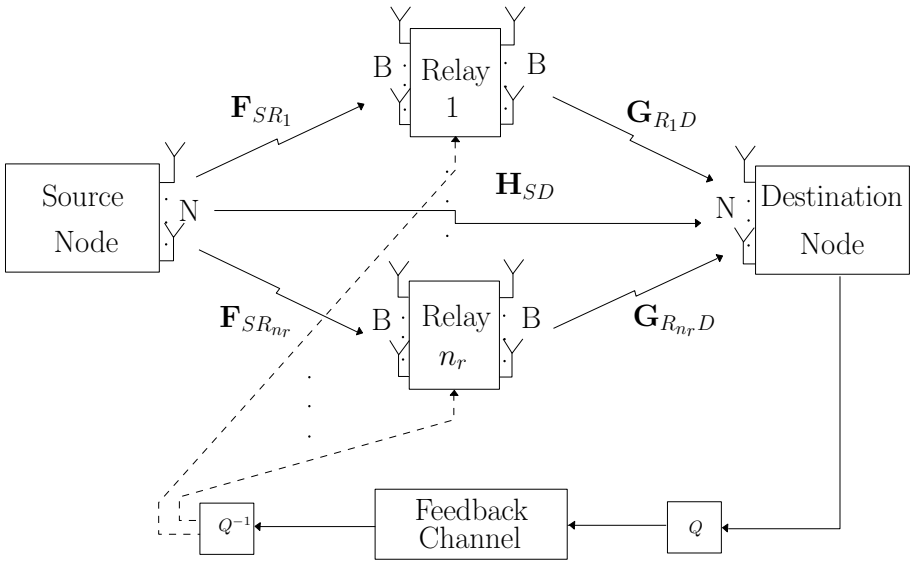

Figure 5.1: Cooperative Wireless Communication System

given by $\mathbf{\Delta} = [\delta_1, \delta_2, ..., \delta_{n_r}]$, where δ_k denotes the relative delay of the signal received from the kth relay node as a reference to the earliest received relay signal. The maximum relative delay is denoted as δ_{max}. Assumptions are considered as follows:

1. $0 \leq \delta_1 \leq \delta_2 \leq ... \leq \delta_{max}$, and δs are integers.

2. The source and the relay nodes do not know the delay profile $\mathbf{\Delta}$. However, the destination node knows $\mathbf{\Delta}$ perfectly.

3. The destination node knows the channels between the relays and the destination perfectly, and the relays know the channels between the source node and themselves perfectly.

4. The total number of antennas at the relays for the two system models is the same in order to have a fair comparison. That leads to more relay nodes in SAS compare to MAS.

It is worth to mention that if the delay coefficients are not integer, which leads to the imperfect superposition between the received space-time encoded schemes sent from different relays and the received code matrices at the destination node can no longer to be considered as individual coding matrices. We consider only one user at the source node in our system that operates in a spatial multiplexing configuration. Let $\boldsymbol{s}[i]$ denote the transmitted information symbol vector at the source node, which contains N parameters,

$\boldsymbol{s}[i] = [s_1[i], s_2[i], ..., s_N[i]]$, and has a covariance matrix $E\big[\boldsymbol{s}[i]\boldsymbol{s}^H[i]\big] = \sigma_s^2 \boldsymbol{I}_N$, where σ_s^2 denotes the signal power. The source node broadcasts $\boldsymbol{s}[i]$ to n_r relay nodes as well as to the destination in the first hop, and all the relay nodes can detect the received symbol vectors with no errors.

The received symbols are re-encoded at each relay node prior to transmission to the destination node in the second hop. We first consider the kth relay node in MAS, the $N \times 1$ signal vector $\boldsymbol{s}[i]$ will be re-encoded by an $N \times T$ DSTC scheme $\boldsymbol{M}(\boldsymbol{s})$, multiplied by an $N \times N$ adjustable code matrix $\boldsymbol{\Phi}_k[i]$ generated randomly [6], and then forwarded to the destination. The relationship between the kth relay and the destination can be described as

$$\boldsymbol{R}_{R_k D_{MAS}}[i] = \boldsymbol{G}^{\boldsymbol{\Delta}}_{R_k D_{MAS}}[i] \boldsymbol{\Phi}_{k_{MAS}}[i] \boldsymbol{M}_{R_k D_{MAS}}[i] + \boldsymbol{N}_{R_k D}[i], \tag{5.1}$$

where $\boldsymbol{G}^{\boldsymbol{\Delta}}_{R_k D_{MAS}}[i]$ denotes the $(\delta_{max} + T) \times N$ delayed version of the channel matrix between the kth relay node and the destination node and $\boldsymbol{M}_{R_k D_{MAS}}[i]$ is the $N \times T$ DSTC scheme. $\boldsymbol{N}_{R_k D}[i]$ stands for the noise matrix at the destination with variance σ_d^2. The $(\delta_{max} + T) \times N$ received symbol matrix $\boldsymbol{R}_{R_k D_{MAS}}[i]$ in (3.1) can be written as an $N(\delta_{max} + T) \times 1$ vector $\boldsymbol{r}_{R_k D_{MAS}}[i]$ given by

$$\boldsymbol{r}_{R_k D_{MAS}}[i] = \boldsymbol{\Delta}_k \boldsymbol{\Phi}_{eq_{k\,MAS}}[i] \boldsymbol{G}_{eq_{k\,MAS}}[i] \tilde{\boldsymbol{s}}[i] + \boldsymbol{n}_{R_k D}[i], \tag{5.2}$$

where the $N(\delta_{max}+T) \times NT$ delay profile matrix $\boldsymbol{\Delta}_k = [\boldsymbol{0}_{\delta_k \times N}; \boldsymbol{I}_N; \boldsymbol{0}_{(\delta_{max}-\delta_k)\times N}]$ at the kth relay nodes is considered. The block diagonal $NT \times NT$ matrix $\boldsymbol{\Phi}_{eq_{k\,MAS}}[i]$ denotes the equivalent adjustable code matrix and the $NT \times N$ matrix $\boldsymbol{G}_{eq_{k\,MAS}}[i]$ stands for the equivalent channel matrix which is the DSTC scheme $\boldsymbol{M}(\boldsymbol{s})$ combined with the channel matrix $\boldsymbol{G}_{R_k D_{MAS}}[i]$. $\tilde{\boldsymbol{s}}[i]$ stands for the data vector processed by the relay node with the AF or DF protocol. The $N(\delta_{max} + T) \times 1$ equivalent noise vector $\boldsymbol{n}_{R_k D}[i]$ generated at the destination node contains noise samples.

In SAS, the $n_r \times T$ DSTC scheme $\boldsymbol{M}(\boldsymbol{s})$ is encoded across all the relay nodes. At the kth relay node, a $1 \times N$ code vector $\boldsymbol{m}_k(\boldsymbol{s})$ is re-encoded which denotes the kth column in $\boldsymbol{M}(\boldsymbol{s})$. The re-encoded code vector will be multiplied by a $1 \times N$ adjustable code vector $\boldsymbol{\phi}_k[i]$ prior to transmission to the destination. The signal matrix sent from the kth relay

node can be described as

$$\boldsymbol{R}_{R_kD_{SAS}}[i] = \boldsymbol{g}^{\boldsymbol{\Delta}}_{R_kD_{SAS}}[i](\boldsymbol{\phi}_{k_{SAS}}[i] \cdot \boldsymbol{m}_{R_kD_{SAS}}[i]) + \boldsymbol{N}_{R_kD}[i], \tag{5.3}$$

where $\boldsymbol{g}^{\boldsymbol{\Delta}}_{R_kD_{SAS}}[i]$ denotes the $(\delta_{max}+T)\times 1$ delayed version of the channel vector between the kth relay node and the destination and $\boldsymbol{N}_{R_kD}[i]$ stands for the noise matrix with variance σ_d^2. The $(\delta_{max}+T)\times N$ received symbol matrix $\boldsymbol{R}_{R_kD_{SAS}}[i]$ in (5.3) can be written as an $N(\delta_{max}+T)\times 1$ vector $\boldsymbol{r}_{R_kD_{SAS}}[i]$ given by

$$\boldsymbol{r}_{R_kD_{SAS}}[i] = \boldsymbol{\Delta}_k\boldsymbol{G}_{eq_{kSAS}}[i]\boldsymbol{\Phi}_{eq_{kSAS}}[i]\tilde{\boldsymbol{s}}[i] + \boldsymbol{n}_{R_kD}[i], \tag{5.4}$$

where the block diagonal $N\times N$ matrix $\boldsymbol{\Phi}_{eq_{kSAS}}[i]$ denotes the diagonal equivalent adjustable code matrix and the block diagonal $NT\times N$ matrix $\boldsymbol{G}_{eq_{kSAS}}[i]$ stands for the equivalent channel matrix. The $N(\delta_{max}+T)\times 1$ equivalent noise vector $\boldsymbol{n}_{R_kD}[i]$ generated at the destination node contains the noise parameters.

After rewriting the received vectors in (5.2) and (5.4) we can consider the received symbol vector at the destination node as an $N(T+\delta_{max}+1)\times 1$ vector. Therefore, the received symbol vector for MAS and SAS can be written respectively, as

$$\begin{aligned}\boldsymbol{r}_{MAS}[i] &= \begin{bmatrix} \boldsymbol{H}_{SD}[i] \\ \sum_{k=1}^{n_r}\boldsymbol{\Delta}_k\boldsymbol{\Phi}_{eq_{kMAS}}[i]\boldsymbol{G}_{eq_{kMAS}}[i] \end{bmatrix}\tilde{\boldsymbol{s}}[i] + \begin{bmatrix} \boldsymbol{n}_{SD}[i] \\ \boldsymbol{n}_{RD}[i] \end{bmatrix} \\ &= \boldsymbol{D}_{D_{MAS}}[i]\tilde{\boldsymbol{s}}[i] + \boldsymbol{n}_D[i],\end{aligned} \tag{5.5}$$

and

$$\begin{aligned}\boldsymbol{r}_{SAS}[i] &= \begin{bmatrix} \boldsymbol{H}_{SD}[i] \\ \sum_{k=1}^{n_r}\boldsymbol{\Delta}_k\boldsymbol{G}_{eq_{kSAS}}[i]\boldsymbol{\Phi}_{eq_{kSAS}}[i] \end{bmatrix}\tilde{\boldsymbol{s}}[i] + \begin{bmatrix} \boldsymbol{n}_{SD}[i] \\ \boldsymbol{n}_{RD}[i] \end{bmatrix} \\ &= \boldsymbol{D}_{D_{SAS}}[i]\tilde{\boldsymbol{s}}[i] + \boldsymbol{n}_D[i],\end{aligned} \tag{5.6}$$

where the $N(\delta_{max}+T+1)\times N$ matrices $\boldsymbol{D}_D[i]$ in MAS and SAS denote the channel gain matrix of all the links in the network. We assume that the coefficients in all channel matrices are independent and remain constant over the transmission. The $N(\delta_{max}+T+1)\times N$ noise vector $\boldsymbol{n}_D[i]$ contains the equivalent received noise vector at the destination, which can be modeled as complex Gaussian random variables with zero mean and variance σ_d^2.

According to the proposed designs, the system we considered needs a feedback chan-

nel for transmitting the optimized code matrices and the information of the optimal relay node back to the relays. We have studied the effect of using a Binary Symmetric Channel (BSC) as the feedback channel in [7] and concluded that imperfect feedback channel with different error probabilities and different numbers of bits in quantization lead to different detection errors in the optimized code matrices at the relays and also cause degradation of the BER performance at the destination. In this chapter, we assume the feedback channel is error-free and delay-free for simplicity. A guard interval is also considered in this system.

5.3 Delay-Tolerant Adjustable Code Matrix Optimization for Delayed DSTC Schemes

In this section, we propose a delay-tolerant adjustable code matrix optimization (DT-ACMO) strategy which employs an ML receiver at the destination for designing DSTC schemes in MAS and SAS. Adaptive RLS and SG algorithms [59] for determining the parameters of the adjustable code matrix with reduced computational complexity are also devised. The DSTC scheme encoded at each relay node employs an ML-based adjustable code matrix, which is computed at the destination node and obtained by a feedback channel in order to process the data symbols prior to transmission to the destination. It is worth to mention that the code matrices will be sent back to the relays through a feedback channel after the optimization and they are only used at the relays so the direct link from the source node to the destination node is not considered for simplicity.

5.3.1 DT-ACMO Algorithm for MAS

The ML criterion is employed for optimizing the adjustable code matrices and performing symbol detection in the DT-ACMO algorithm, which is equivalent to a least-squares (LS) criterion in this case. An $N \times D$ matrix $\boldsymbol{S}$ is stored at the destination which contains all the possible combinations of the transmitted symbol vectors, where D denotes the number of

the possible combination. The constrained ML optimization problem can be written as

$$\begin{aligned}\left[\hat{\boldsymbol{s}}[i], \hat{\boldsymbol{\Phi}}_{eq_{k\,MAS}}[i]\right] &= \underset{\boldsymbol{s}[i], \boldsymbol{\Phi}_{eq_{k\,MAS}}[i]}{\operatorname{argmin}} \|\boldsymbol{r}_{MAS}[i] - \hat{\boldsymbol{r}}_{MAS}[i]\|^2, \\ s.t.\ \sum_{k=1}^{n_r} \mathrm{Tr}(\boldsymbol{\Phi}_{eq_{k\,MAS}}[i]\boldsymbol{\Phi}^H_{eq_{k\,MAS}}[i]) &\leq \mathrm{P_R},\ for\ d = 1, 2, ..., D,\end{aligned} \tag{5.7}$$

where $\hat{\boldsymbol{r}}_{MAS}[i] = \sum_{k=1}^{n_r} \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]\boldsymbol{G}_{eq_{k\,MAS}}[i]\boldsymbol{s}_d[i]$ denotes the received symbol vector without noise which is determined by substituting the dth column of $\boldsymbol{S}$ into (3.20), and $\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]$ denotes the $N(\delta_{max}+T)\times NT$ delayed adjustable code matrix at the kth relay node. As mentioned in [7], an ML detection and an optimization of the adjustable code matrix are implemented in the proposed algorithms. Different types of detectors such as MMSE and sphere decoders can be used in order to reduce the computational complexity. The ML optimization problem for determining the transmitted symbol vector in (5.7) is given by

$$\hat{\boldsymbol{s}}[i] = \underset{\boldsymbol{s}_d \in \boldsymbol{S}}{\operatorname{argmin}} \|\boldsymbol{r}_{MAS}[i] - \sum_{k=1}^{n_r} \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]\boldsymbol{G}_{eq_{k\,MAS}}[i]\boldsymbol{s}_d\|^2. \tag{5.8}$$

By substituting each column in $\boldsymbol{S}$ into (5.8) to search the most likely transmitted symbol vector $\hat{\boldsymbol{s}}[i]$ according to the ML detection, we can then calculate the optimal adjustable code matrix $\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]$. The Lagrangian expression of the optimization problem in (5.7) is given by

$$\mathscr{L} = \|\boldsymbol{r}_{MAS}[i] - \sum_{k=1}^{n_r} \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]\boldsymbol{G}_{eq_{k\,MAS}}[i]\hat{\boldsymbol{s}}[i]\|^2 + \lambda(\sum_{k=1}^{n_r} Tr(\boldsymbol{\Phi}_{eq_{k\,MAS}}[i]\boldsymbol{\Phi}^H_{eq_{k\,MAS}}[i]) - P_R), \tag{5.9}$$

and by taking the instantaneous gradient of $\mathscr{L}$ with respect to the code matrix $(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i])^*$ we can obtain

$$\begin{aligned}\nabla\mathscr{L}_{(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i])^*} =& -\boldsymbol{r}_{MAS}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{k\,MAS}}[i] \\ &+ \sum_{l=1, l\neq k}^{n_r} \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{l\,MAS}}[i]\boldsymbol{G}_{eq_{l\,MAS}}[i]\hat{\boldsymbol{s}}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{k\,MAS}}[i] \\ &+ \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]\boldsymbol{G}_{eq_{k\,MAS}}[i]\hat{\boldsymbol{s}}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{k\,MAS}}[i] \\ =&(\boldsymbol{r}_{l_{MAS}}[i] - \boldsymbol{r}_{MAS}[i])\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{k\,MAS}}[i] \\ &+ \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]\boldsymbol{G}_{eq_{k\,MAS}}[i]\hat{\boldsymbol{s}}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{k\,MAS}}[i],\end{aligned} \tag{5.10}$$

where $\boldsymbol{r}_{l_{MAS}}[i] = \sum_{l=1,l\neq k}^{n_r} \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eql_{MAS}}[i]\boldsymbol{G}_{eql_{MAS}}[i]\hat{\boldsymbol{s}}[i]$ stands for the received vector without the desired code matrix. The optimal code matrix $\hat{\boldsymbol{\Phi}}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i]$ requires $\nabla\mathscr{L}_{(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i])^*} = 0$, and the optimal adjustable code matrix is given by

$$\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i] = (\boldsymbol{r}_{MAS}[i] - \boldsymbol{r}_{l_{MAS}}[i])\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{k_{MAS}}}[i](\boldsymbol{G}_{eq_{k_{MAS}}}[i]\hat{\boldsymbol{s}}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{k_{MAS}}}[i])^{\dagger}. \tag{5.11}$$

It is worth to mention that the last term related to the power constraint in (5.9) is not considered during the optimization in (5.11) which is due to the high computational complexity of calculating the value of λ. The power constraint can be achieved by the normalization method described by

$$\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i]}{\sqrt{\sum_{k=1}^{n_r}\mathrm{Tr}(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i](\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i])^H)}}. \tag{5.12}$$

RLS Code Matrix Estimation Algorithm

The RLS estimation algorithm for the code matrix $\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i]$ is derived in this section. The superior convergence behavior of LS type algorithms when the size of the adjustable code matrix is large is the reason for its utilization. It is worth to mention that the computational complexity reduces from cubic to square by employing the RLS algorithm.

In order to derive an RLS algorithm to compute $\hat{\boldsymbol{\Phi}}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i]$, the optimization problem is given by

$$\begin{aligned}\hat{\boldsymbol{\Phi}}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i] = &\underset{\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i]}{\operatorname{argmin}} \sum_{n=1}^{i}\lambda^{i-n}\|\boldsymbol{r}_{MAS}[n] - \hat{\boldsymbol{r}}_{MAS}[i]\|^2,\\ &s.t.\ \sum_{k=1}^{n_r}\mathrm{Tr}(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i](\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i])^H[i]) \leq \mathrm{P_R},\end{aligned} \tag{5.13}$$

where λ stands for the forgetting factor. By expanding the right-hand side of (5.13), taking the gradient with respect to $(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i])^*$ and equating the terms to zero, we obtain

$$\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k_{MAS}}}[i] = \left(\sum_{n=1}^{i}\lambda^{i-n}\boldsymbol{r}_{e_{MAS}}[n]\boldsymbol{r}^H_{k_{MAS}}[n]\right)\left(\sum_{i=1}^{n}\lambda^{i-n}\boldsymbol{r}_{k_{MAS}}[n]\boldsymbol{r}^H_{k_{MAS}}[n]\right)^{-1}, \tag{5.14}$$

where the $N(\lambda_{max}+T)\times 1$ vector $\boldsymbol{r}_{e_{MAS}}[n] = \boldsymbol{\Delta}_k\boldsymbol{\Phi}_{eq_{k_{MAS}}}[i]\boldsymbol{G}_{eq_{k_{MAS}}}[i]\boldsymbol{s}[i]$ and

$\boldsymbol{r}_{k_{MAS}}[n] = \boldsymbol{G}_{eq_{k\,MAS}}[i]\boldsymbol{s}[i]$. The power constraint is still not considered during the derivation but is enforced by the normalization strategy after the optimization. We can define

$$\boldsymbol{\Psi}[i] = \sum_{n=1}^{i} \lambda^{i-n} \boldsymbol{r}_{k_{MAS}}[n]\boldsymbol{r}^{H}_{k_{MAS}}[n] = \lambda\boldsymbol{\Psi}[i-1] + \boldsymbol{r}_{k_{MAS}}[n]\boldsymbol{r}^{H}_{k_{MAS}}[n], \tag{5.15}$$

and

$$\boldsymbol{Z}[i] = \sum_{n=1}^{i} \lambda^{i-n} \boldsymbol{r}_{e_{MAS}}[n]\boldsymbol{r}^{H}_{k_{MAS}}[n] = \lambda\boldsymbol{Z}[i-1] + \boldsymbol{r}_{e_{MAS}}[n]\boldsymbol{r}^{H}_{k_{MAS}}[n], \tag{5.16}$$

so that we can rewrite (5.14) as

$$\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i] = \boldsymbol{Z}[i]\boldsymbol{\Psi}^{-1}[i]. \tag{5.17}$$

By employing the matrix inversion lemma in [15], we can obtain

$$\boldsymbol{\Psi}^{-1}[i] = \lambda^{-1}\boldsymbol{\Psi}^{-1}[i-1] - \lambda^{-1}\boldsymbol{k}[i]\boldsymbol{r}^{H}_{k_{MAS}}[i]\boldsymbol{\Psi}^{-1}[i-1], \tag{5.18}$$

where $\boldsymbol{k}[i] = (\lambda^{-1}\boldsymbol{\Psi}^{-1}[i-1]\boldsymbol{r}_{k_{MAS}}[i])/(1+\lambda^{-1}\boldsymbol{r}^{H}_{k_{MAS}}[i]\boldsymbol{\Psi}^{-1}[i-1]\boldsymbol{r}_{k_{MAS}}[i])$. We define $\boldsymbol{P}[i] = \boldsymbol{\Psi}^{-1}[i]$ and substitute (5.16) and (5.18) into (5.17) to obtain the expression of the code matrix which is given by

$$\begin{aligned}
\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i] &= \lambda\boldsymbol{Z}[i-1]\boldsymbol{P}[i] + \boldsymbol{r}_{e_{MAS}}[i]\boldsymbol{r}^{H}_{k_{MAS}}[i]\boldsymbol{P}[i] \\
&= \boldsymbol{Z}[i-1]\boldsymbol{P}[i-1] + \boldsymbol{Z}[i-1]\boldsymbol{k}[i]\boldsymbol{r}^{H}_{k_{MAS}}[i]\boldsymbol{P}[i-1] + \boldsymbol{r}_{e_{MAS}}[i]\boldsymbol{r}^{H}_{k_{MAS}}[i]\boldsymbol{P}[i] \\
&= \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i-1] + \lambda^{-1}(\boldsymbol{r}_{e_{MAS}}[i] - \boldsymbol{Z}[i-1]\boldsymbol{k}[i])\boldsymbol{r}^{H}_{k_{MAS}}[i]\boldsymbol{P}[i-1].
\end{aligned} \tag{5.19}$$

As mentioned before, a normalization process after (5.19) is implemented in order to maintain the energy of the code matrices. Table 5.1 shows a summary of the DT-ACMO RLS algorithm.

SG Code Matrix Estimation Algorithm

The SG estimation algorithm for the code matrix $\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]$ is derived in this section. As shown in (3.24), the inversion of the matrices is required and the computational complex-

Table 5.1: Summary of the DT-ACMO RLS Algorithm

1:	Initialize: $\boldsymbol{P}[0] = \delta^{-1}\boldsymbol{I}_{N(\lambda_{max}+T)\times N(\lambda_{max}+T)}$, $\boldsymbol{Z}[0] = \boldsymbol{I}_{N(\lambda_{max}+T)\times N(\lambda_{max}+T)}$,
2:	generate $\boldsymbol{\Phi}[0]$ randomly with the power constraint $\mathrm{Tr}(\boldsymbol{\Phi}_{eq_{k\,MAS}}[i]\boldsymbol{\Phi}^{H}_{eq_{k\,MAS}}[i]) \leq \mathrm{P_R}$.
3:	For each instant of time, i=1, 2, ..., compute
4:	$\boldsymbol{k}[i] = \frac{\lambda^{-1}\boldsymbol{P}[i-1]\boldsymbol{r}_{k_{MAS}}[i]}{1+\lambda^{-1}\boldsymbol{r}^{H}_{k_{MAS}}[i]\boldsymbol{\Psi}^{-1}[i-1]\boldsymbol{r}_{k_{MAS}}[i]}$,
5:	$\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i] = \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i-1] + \lambda^{-1}(\boldsymbol{r}_{e_{MAS}}[i] - \boldsymbol{Z}[i-1]\boldsymbol{k}[i])\boldsymbol{r}^{H}_{k_{MAS}}[i]\boldsymbol{P}[i-1]$,
6:	$\boldsymbol{P}[i] = \lambda^{-1}\boldsymbol{P}[i-1] - \lambda^{-1}\boldsymbol{k}[i]\boldsymbol{r}^{H}_{k_{MAS}}[i]\boldsymbol{P}[i-1]$,
7:	$\boldsymbol{Z}[i] = \lambda\boldsymbol{Z}[i-1] + \boldsymbol{r}_{e_{MAS}}[i]\boldsymbol{r}^{H}_{k_{MAS}}[i]$.
8:	$\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]}{\sqrt{\sum_{k=1}^{n_r}\mathrm{Tr}(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i](\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i])^{H})}}$.

ity will be increased by employing a large number of relay nodes or number of antennas.

By taking the instantaneous gradient term of the Lagrangian expression derived in (5.9) with respect to the adjustable code matrix results in

$$\begin{aligned}\nabla\mathscr{L}_{(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i])^{*}} &= \nabla\left(\boldsymbol{r}_{MAS}[i] - \sum_{k=1}^{n_r}\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]\boldsymbol{G}_{eq_{k\,MAS}}[i]\hat{\boldsymbol{s}}[i]\right)^{H}_{(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i])^{*}}\boldsymbol{r}_{e}[i] \\ &= -\hat{\boldsymbol{s}}^{H}[i]\boldsymbol{G}^{H}_{eq_{k\,MAS}}[i]\boldsymbol{r}_{e}[i],\end{aligned} \tag{5.20}$$

where $\boldsymbol{r}_{e}[i] = \boldsymbol{r}_{MAS}[i] - \sum_{k=1}^{n_r}\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]\boldsymbol{G}_{eq_{k\,MAS}}[i]\hat{\boldsymbol{s}}[i]$ stands for the error vector. After we obtain (5.20) the proposed SG algorithm can be obtained by introducing a step size into a gradient optimization algorithm to update the result until the convergence is reached. The DT-ACMO SG algorithm is given by

$$\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i+1] = \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i] + \beta(\hat{\boldsymbol{s}}^{H}[i]\boldsymbol{G}^{H}_{eq_{k\,MAS}}[i]\boldsymbol{r}_{e}[i]), \tag{5.21}$$

where β denotes the step size in the recursions for the estimation.

The energy of the code matrices in (5.21) will be increased with the processing of the adaptive algorithm, which will contribute to the reduction of the error probability. The normalization of the code matrix expressed in (5.12) is applied to ensure that the energy is not increased and for a fair comparison among the analyzed DSTC schemes. A summary of the DT-ACMO SG algorithm is given in Table 5.2.

Table 5.2: Summary of the DT-ACMO SG Algorithm in MAS

1:	Generate $\boldsymbol{\Phi}[0]$ randomly with the power constraint.
2:	For each instant of time, i=1, 2, ..., compute
3:	$\nabla \mathscr{L}_{(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i])^*} = -\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{k\,MAS}}[i]\boldsymbol{r}_e[i]$,
4:	where $\boldsymbol{r}_e[i] = \boldsymbol{r}_{MAS}[i] - \sum_{k=1}^{n_r} \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]\boldsymbol{G}_{eq_{k\,MAS}}[i]\hat{\boldsymbol{s}}[i]$.
5:	Update $\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]$ by
6:	$\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i+1] = \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i] + \beta\nabla \mathscr{L}_{(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i])^*}$,
7:	Normalization: $\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]}{\sqrt{\sum_{k=1}^{n_r}\mathrm{Tr}(\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i](\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i])^H)}}$.

5.3.2 DT-ACMO Algorithm for SAS

The main difference of the DT-ACMO algorithm in SAS compared with that in MAS is the computational complexity which is related to the dimension of the adjustable code matrices used at the relays. If we assume that the samples of the data received by the antennas used at the relays are independent and identically distributed (i.i.d) and consider the relay nodes as one unit transmitter, the distributed array formed by the relays can be seen as one transmitter with n_r antennas. The equivalent received vector at the destination is expressed as

$$\boldsymbol{r}_{SAS}[i] = \sum_{k=1}^{n_r N} \boldsymbol{\Delta}_k \boldsymbol{G}_{eq_{k\,SAS}}[i]\boldsymbol{\Phi}_{eq_{k\,SAS}}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{RD}[i], \tag{5.22}$$

where the block diagonal $N \times N$ matrix $\boldsymbol{\Phi}_{eq_{k\,SAS}}[i]$ denotes the equivalent adjustable code matrix for the kth antenna and the block diagonal $NT \times N$ matrix $\boldsymbol{G}_{eq_{k\,SAS}}[i]$ stands for the equivalent channel matrix combined with the DSTC scheme and the $1 \times N$ channel vector between the kth antenna and the destination.

According to the ML criterion, the optimization problem can be written as

$$\begin{aligned}\left[\hat{\boldsymbol{s}}[i], \hat{\boldsymbol{\Phi}}_{eq_{k\,SAS}}[i]\right] = \underset{\boldsymbol{s}[i], \boldsymbol{\Phi}_{eq_{k\,SAS}}[i]}{\operatorname{argmin}} \|\boldsymbol{r}_{SAS}[i] - \sum_{k=1}^{n_r N} \boldsymbol{\Delta}_k \boldsymbol{G}_{eq_{k\,SAS}}[i]\boldsymbol{\Phi}_{eq_{k\,SAS}}[i]\boldsymbol{s}[i]\|^2, \\ s.t.\ \mathrm{Tr}(\boldsymbol{\Phi}_{eq_{k\,SAS}}[i]\boldsymbol{\Phi}^H_{eq_{k\,SAS}}[i]) \leq \mathrm{P_R},\ for\ d = 1, 2, ..., D.\end{aligned} \tag{5.23}$$

After we obtain $\hat{\boldsymbol{s}}[i]$ by using an ML detector derived in the previous section, we can

rewrite the Lagrangian expression of (5.23) as

$$\mathscr{L} = \|\boldsymbol{r}_{SAS}[i] - \sum_{k=1}^{n_r N} \boldsymbol{\Delta}_k \boldsymbol{G}_{eq_{kSAS}}[i] \boldsymbol{\Phi}_{eq_{kSAS}}[i] \hat{\boldsymbol{s}}[i]\|^2 + \lambda(Tr(\boldsymbol{\Phi}_{eq_{kSAS}}[i] \boldsymbol{\Phi}^H_{eq_{kSAS}}[i]) - P_R), \tag{5.24}$$

where λ stands for the Lagrange multiplier. In order to obtain the optimal adjustable code matrices, we have to expand the right-hand side of $\mathscr{L}$, compute the gradient terms with respect to the kth matrix $\boldsymbol{\Phi}_{eq_{kSAS}}[i]$, and equate them to zero, which results in

$$\hat{\boldsymbol{\Phi}}_{eq_{kSAS}}[i] = (\boldsymbol{\Delta}_k \boldsymbol{G}_{eq_{kSAS}}[i])^{\dagger} \boldsymbol{\Theta}[i] (\hat{\boldsymbol{s}}[i] \hat{\boldsymbol{s}}^H[i] \boldsymbol{G}^H_{eq_{kSAS}}[i] \boldsymbol{\Delta}^H_k)^{\dagger}, \tag{5.25}$$

where

$$\begin{aligned} \boldsymbol{\Theta}[i] =& \boldsymbol{r}_{SAS}[i] \hat{\boldsymbol{s}}^H[i] \boldsymbol{G}^H_{eq_{kSAS}}[i] \boldsymbol{\Delta}^H_k \\ & - (\sum_{l=1, l\neq k}^{n_r N} \boldsymbol{\Delta}_l \boldsymbol{G}_{eq_{lSAS}}[i] \boldsymbol{\Phi}_{eq_{lSAS}}[i] \hat{\boldsymbol{s}}[i]) \hat{\boldsymbol{s}}^H[i] \boldsymbol{G}^H_{eq_{kSAS}}[i] \boldsymbol{\Delta}^H_k. \end{aligned} \tag{5.26}$$

The detailed derivation is shown in Appendix III. Note that the last term in (5.24) which contains the Lagrange multiplier λ is not considered in the optimization algorithm. The value of λ can be obtained by substituting (5.25) into (5.24), or this procedure can be avoided by employing the normalization to achieve the power constraint as derived in (5.12) in order to reduce the computational complexity.

RLS Code Matrix Estimation Algorithm

The RLS optimization problem of the adjustable code matrices is written as

$$\hat{\boldsymbol{\Phi}}_{eq_{kSAS}}[i] = \arg \min_{\boldsymbol{\Phi}_{eq_{kSAS}}[i]} \sum_{n=1}^{i} \lambda^{i-n} \|\boldsymbol{r}_{SAS}[n] - \hat{\boldsymbol{r}}_{SAS}[i]\|^2. \tag{5.27}$$

By expanding the right-hand side of (5.27), taking the gradient with respect to $\boldsymbol{\Phi}^{*}_{eq_{kSAS}}[i]$ and equating the terms to zero, we obtain

$$\begin{aligned}\boldsymbol{\Phi}_{eq_{kSAS}}[i] = &\left(\sum_{n=1}^{i}\lambda^{i-n}\boldsymbol{\Delta}_k\boldsymbol{G}_{eq_{kSAS}}[i]\boldsymbol{\Phi}_{eq_{kSAS}}[i]\boldsymbol{s}[i]\boldsymbol{r}^{H}_{k_{SAS}}[n]\right)\\ &\left(\sum_{i=1}^{n}\lambda^{i-n}\boldsymbol{G}_{eq_{kSAS}}[i]\boldsymbol{s}[i]\boldsymbol{r}^{H}_{k_{SAS}}[n]\right)^{-1}.\end{aligned} \tag{5.28}$$

The power constraint is enforced by the normalization strategy after the optimization. In order to implement the RLS estimation algorithm we can define

$$\boldsymbol{\Psi}_{SAS}[i] = \sum_{n=1}^{i}\lambda^{i-n}\boldsymbol{r}_{k_{SAS}}[n]\boldsymbol{r}^{H}_{k_{SAS}}[n] = \lambda\boldsymbol{\Psi}_{SAS}[i-1] + \boldsymbol{r}_{k_{SAS}}[n]\boldsymbol{r}^{H}_{k_{SAS}}[n], \tag{5.29}$$

and

$$\boldsymbol{Z}_{SAS}[i] = \sum_{n=1}^{i}\lambda^{i-n}\boldsymbol{r}_{e_{SAS}}[n]\boldsymbol{r}^{H}_{k_{SAS}}[n] = \lambda\boldsymbol{Z}_{SAS}[i-1] + \boldsymbol{r}_{e_{SAS}}[n]\boldsymbol{r}^{H}_{k_{SAS}}[n], \tag{5.30}$$

where $\boldsymbol{r}_{e_{SAS}}[n] = \boldsymbol{\Delta}_k\boldsymbol{G}_{eq_{kSAS}}[i]\boldsymbol{\Phi}_{eq_{kSAS}}[i]\boldsymbol{s}[i]$ and $\boldsymbol{r}_{k_{SAS}}[n] = \boldsymbol{G}_{eq_{kSAS}}[i]\boldsymbol{s}[i]$. Thus, we can rewrite (5.28) as

$$\boldsymbol{\Phi}_{eq_{kSAS}}[i] = \boldsymbol{Z}_{SAS}[i]\boldsymbol{\Psi}^{-1}_{SAS}[i]. \tag{5.31}$$

By employing the matrix inversion lemma in [15], we can obtain

$$\boldsymbol{\Psi}^{-1}_{SAS}[i] = \lambda^{-1}\boldsymbol{\Psi}^{-1}_{SAS}[i-1] - \lambda^{-1}\boldsymbol{k}_{SAS}[i]\boldsymbol{r}^{H}_{k_{SAS}}[i]\boldsymbol{\Psi}^{-1}_{SAS}[i-1], \tag{5.32}$$

where $\boldsymbol{k}_{SAS}[i] = (\lambda^{-1}\boldsymbol{\Psi}^{-1}_{SAS}[i-1]\boldsymbol{r}_{k_{SAS}}[i])/(1+\lambda^{-1}\boldsymbol{r}^{H}_{k_{SAS}}[i]\boldsymbol{\Psi}^{-1}_{SAS}[i-1]\boldsymbol{r}_{k_{SAS}}[i])$. We define $\boldsymbol{P}_{SAS}[i] = \boldsymbol{\Psi}^{-1}_{SAS}[i]$ and by substituting (5.30) and (5.32) into (5.31), the expression of the code matrix is given by

$$\begin{aligned}\boldsymbol{\Phi}_{eq_{kSAS}}[i] &= \lambda\boldsymbol{Z}_{SAS}[i-1]\boldsymbol{P}_{SAS}[i] + \boldsymbol{r}_{e_{SAS}}[i]\boldsymbol{r}^{H}_{k_{SAS}}[i]\boldsymbol{P}_{SAS}[i]\\ &= \boldsymbol{Z}_{SAS}[i-1]\boldsymbol{P}_{SAS}[i-1]\\ &\quad + \boldsymbol{Z}_{SAS}[i-1]\boldsymbol{k}_{SAS}[i]\boldsymbol{r}^{H}_{k_{SAS}}[i]\boldsymbol{P}_{SAS}[i-1]\\ &\quad + \boldsymbol{r}_{e_{SAS}}[i]\boldsymbol{r}^{H}_{k_{SAS}}[i]\boldsymbol{P}_{SAS}[i]\\ &= \boldsymbol{\Phi}_{eq_{kSAS}}[i-1] + \lambda^{-1}(\boldsymbol{r}_{e_{SAS}}[i] - \boldsymbol{Z}_{SAS}[i-1]\boldsymbol{k}_{SAS}[i])\boldsymbol{r}^{H}_{k_{SAS}}[i]\boldsymbol{P}_{SAS}[i-1].\end{aligned} \tag{5.33}$$

By using the algorithm in Table I we can obtain the optimal code matrices, and the only differences are the initialized matrices $\boldsymbol{P}[0] = \delta^{-1}\boldsymbol{I}_{N\times N}$, and $\boldsymbol{Z}[0] = \boldsymbol{I}_{N\times N}$ for SAS.

SG Code Matrix Estimation Algorithm

The SG estimation algorithm for the code matrix $\boldsymbol{\Phi}_{eq_{kSAS}}[i]$ is derived in this section. As shown in (5.25), the inversion of the matrices is required and the computational complexity will be increased by employing a large number of relay nodes or number of antennas.

We can rewrite the expression of the received vector for SAS in (5.22) at the destination node as

$$\boldsymbol{r}_{SAS}[i] = \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{G}^{\boldsymbol{\Delta}_k}_{eq_{k_j\,SAS}}[i]\phi_{k_{j\,SAS}}[i]\boldsymbol{s}[i] + n_D[i], \tag{5.34}$$

where $\boldsymbol{G}^{\boldsymbol{\Delta}_k}_{eq_{k_j\,SAS}}[i]$ denotes the $N(\delta_{max}+T)\times N$ equivalent channel matrix combining with the delay profile, the DSTC scheme, and the channel vector between the jth antenna of the kth relay node and the destination node. The ML based Lagrangian expression is derived as

$$\mathscr{L} = \|\boldsymbol{r}_{SAS}[i] - \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{G}^{\boldsymbol{\Delta}_k}_{eq_{k_j\,SAS}}[i]\phi_{k_{j\,SAS}}[i]\hat{\boldsymbol{s}}[i]\|^2, \tag{5.35}$$

and a simple adaptive algorithm for determining the adjustable code matrices can be derived by taking the instantaneous gradient terms of the Lagrangian expression with respect to the scalar $\phi^*_{k_{j\,SAS}}[i]$ in $\boldsymbol{\Phi}_{eq_{kSAS}}[i]$, which is

$$\begin{aligned}\nabla\mathscr{L}_{\phi^*_{k_{j\,SAS}}[i]} &= \nabla\left(\boldsymbol{r}_{SAS}[i] - \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{G}^{\boldsymbol{\Delta}_k}_{eq_{k_j\,SAS}}[i]\phi_{k_{j\,SAS}}[i]\hat{\boldsymbol{s}}[i]\right)^H_{\phi^*_{k_{j\,SAS}}[i]}\boldsymbol{r}_e[i] \\ &= -\boldsymbol{s}^H[i](\boldsymbol{G}^{\boldsymbol{\Delta}_k}_{eq_{k_j\,SAS}}[i])^H\boldsymbol{r}_e[i],\end{aligned} \tag{5.36}$$

where $\boldsymbol{r}_e[i] = \boldsymbol{r}_{SAS}[i] - \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{G}^{\boldsymbol{\Delta}_k}_{eq_{k_j\,SAS}}[i]\phi_{k_{j\,SAS}}[i]\hat{\boldsymbol{s}}[i]$ stands for the error vector. After we obtain (5.36) the proposed SG algorithm can be obtained by introducing a step size β into a gradient optimization algorithm to update the result until the convergence is reached. The DT-ACMO SG algorithm for SAS is given by

$$\phi_{k_{j\,SAS}}[i+1] = \phi_{k_{j\,SAS}}[i] + \beta(\boldsymbol{s}^H[i](\boldsymbol{G}^{\boldsymbol{\Delta}_k}_{eq_{k_j\,SAS}}[i])^H\boldsymbol{r}_e[i]), \tag{5.37}$$

By following the steps in Table 5.2 we have the DT-ACMO SG algorithm in SAS.

5.4 DT-ACMO Algorithm with Opportunistic DSTCs

In this section, we extend the DT-ACMO algorithms to cooperative systems with the AF strategy using the opportunistic relaying schemes in [33] and design delay-tolerant adjustable code matrices opportunistic relaying optimization (DT-ACMORO) algorithms. The opportunistic relaying techniques in [33] for the AF strategy chooses the best relay and antenna which can achieve a lower BER performance compared to the traditional equal power allocation schemes for the AF protocol systems. Three opportunistic relaying algorithms are designed in [33] for two-hop cooperative systems, namely opportunistic AF scheme, which uses DSTC schemes at the source node in the first phase and selects the best relay nodes in the second phase; opportunistic source, which selects the best antenna at the source node in the first phase and uses DSTCs at the relay nodes in the second phase; and fully-opportunistic, which optimizes the transmitting power both at the source node and at the relay nodes. A feedback channel is employed which allows the destination node to inform the optimal relay and to send back the optimized code matrices to the relays. The delay profiles for different relay nodes are considered in the design described in this section. For simplicity, we only show the design of the DT-ACMORO algorithms for MAS.

The system model is described as follows. The power used at the source node is defined as P_1 and the total power used among all the relay nodes is $P_2 = \sum_{k=1}^{n_r N} P_{2,k}$ where k denotes the total number of antennas among the relay nodes. In the first phase, the source node will encode the MPSK or M-QAM modulated information symbol vector $\boldsymbol{s}[i]$ to an $T \times N$ STC matrix, which gives

$$\boldsymbol{S}[i] = [\boldsymbol{A}_1 \boldsymbol{s}[i] \ \boldsymbol{A}_2 \boldsymbol{s}[i] \ ... \ \boldsymbol{A}_N \boldsymbol{s}[i]] \,, \tag{5.38}$$

where $\boldsymbol{A}_n$, $n = 1, ..., N$, are $T \times T$ unitary matrices. The source node broadcast the STC matrix to the n_r relay nodes after the encoding, and the received symbol at the kth relay

node is described by

$$\boldsymbol{X}_k[i] = \sqrt{\frac{P_1 T}{N}} \boldsymbol{S}[i] \boldsymbol{F}_k + \boldsymbol{N}_k, \tag{5.39}$$

where $\boldsymbol{F}_k$ denotes the $N \times N$ channel matrix between the source node and the kth relay node, and $\boldsymbol{N}_k$ is the matrix whose entries are complex zero-mean complex random variables with variance $\sigma_{n_1}^2$.

In the second phase, according to the AF protocol the received symbols will be re-encoded and amplified before being forwarded to the destination node. It is worth to mention that instead of transmitting the fixed linear functions of the received matrix from the source node in [33], we generate the adjustable code matrices $\boldsymbol{\Phi}_k[i]$ for the kth relay node and multiply them by the received symbol matrices before forwarding them to the destination node. The delay profile $\boldsymbol{\Delta}_k$ for each relay node is considered as well. The communication between the relays and the destination node can be described by

$$\boldsymbol{R}[i] = \sum_{k=1}^{n_r} \sum_{j=1}^{N} \rho_{k,j} \boldsymbol{g}_{k,j}[i] \boldsymbol{y}_{k,j}[i] \boldsymbol{\Delta}_{k,j}[i] + \boldsymbol{N}_d, \tag{5.40}$$

where

$$\begin{aligned} \rho_{k,j} &= \sqrt{\frac{P_{k,j}}{\sigma_{\boldsymbol{F}_k}^2 P_1 + \sigma_{n_1}^2}}, \\ \boldsymbol{y}_{k,j}[i] &= \boldsymbol{\phi}_{k,j}[i] \boldsymbol{X}_k[i], \end{aligned} \tag{5.41}$$

and $\boldsymbol{g}_{k,j}[i]$ is the $N \times 1$ channel coefficients vector between the jth antenna in the kth relay and the destination. By substituting (5.39) and (5.41) into (5.40) we can obtain the received matrix derived as

$$\begin{aligned} \boldsymbol{R}[i] &= \sqrt{\frac{P_1 T}{N}} \sum_{k=1}^{n_r} \sum_{j=1}^{N} \boldsymbol{H}_{k,j}[i] \boldsymbol{S}[i] \boldsymbol{\Delta}_{k,j}[i] + \boldsymbol{W} \\ &= \sqrt{\frac{P_1 T}{N}} \sum_{k=1}^{n_r} \sum_{j=1}^{N} \rho_{k,j} \boldsymbol{g}_{k,j}[i] \boldsymbol{\phi}_{k,j}[i] \boldsymbol{F}_k \boldsymbol{S}[i] \boldsymbol{\Delta}_{k,j}[i] \\ &\quad + \sum_{k=1}^{n_r} \sum_{j=1}^{N} \rho_{k,j} \boldsymbol{g}_{k,j}[i] \boldsymbol{\phi}_{k,j}[i] \boldsymbol{N}_k \boldsymbol{\Delta}_{k,j}[i] + \boldsymbol{N}_d, \end{aligned} \tag{5.42}$$

and we can rewrite the $(N \times (\delta_{max} + T))$ received matrix $\boldsymbol{R}[i]$ as an $(N(T + \delta_{max}) \times 1)$

vector derived as

$$\boldsymbol{r}[i] = \sqrt{\frac{P_1 T}{N}} \sum_{k=1}^{n_r} \sum_{j=1}^{N} \rho_{k,j} \boldsymbol{\Phi}_{eq_{k,j}}[i] \boldsymbol{H}_{eq_{k,j}}^{\boldsymbol{\Delta}_{k,j}}[i] \boldsymbol{s}[i] + \sum_{k=1}^{n_r} \sum_{j=1}^{N} \rho_{k,j} \boldsymbol{\Phi}_{eq_{k,j}}[i] \boldsymbol{G}_{eq_{k,j}}^{\boldsymbol{\Delta}_{k,j}}[i] \boldsymbol{n}_k + \boldsymbol{n}_d, \tag{5.43}$$

where $\boldsymbol{H}_{eq_{k,j}}^{\boldsymbol{\Delta}_{k,j}}[i]$ is the $(N(\delta_{max}+T) \times N)$ equivalent channel matrix with the delay profile and $\boldsymbol{\Phi}_{eq_{k,j}}[i]$ stands for the $(N(\delta_{max}+T) \times N(\delta_{max}+T))$ block diagonal adjustable code matrix. The noise vectors $\boldsymbol{n}_k$ and $\boldsymbol{n}_d$ are the columns of $\boldsymbol{N}_k$ and $\boldsymbol{N}_d$, respectively. Therefore, the opportunistic relaying algorithm problem can be expressed as

$$\hat{k} = \underset{k=1,2,...,n_r}{\operatorname{argmax}} SNR_{k_{ins}} = \sum_{j=1}^{N} \frac{P_1 T \rho_{k,j}^2 \|\boldsymbol{\Phi}_{eq_{k,j}}[i] \boldsymbol{H}_{eq_{k,j}}^{\boldsymbol{\Delta}_{k,j}}[i]\|_F^2 \sigma_s^2}{N \rho_{k,j}^2 \|\boldsymbol{\Phi}_{eq_{k,j}}[i] \boldsymbol{G}_{eq_{k,j}}^{\boldsymbol{\Delta}_{k,j}}[i]\|_F^2 \sigma_{n_1}^2 + \sigma_d^2}, \quad \text{for } k = 1, 2, ..., n_r. \tag{5.44}$$

According to the opportunistic relaying algorithms [33], the relays which can achieve the highest SNR_{ins} are chosen to forward the symbols with transmission power P_2. The DT-ACMORO can be implemented by firstly calculating the optimized adjustable code matrices using the DT-ACMO algorithms proposed in the Table I or Table II, followed by choosing the optimal relay node using the SNR expressed in (5.44). For example, taking the SG algorithm described in Table II, the DT-ACMORO algorithm is shown as

1. Use the ML detector to decode $\hat{\boldsymbol{s}}[i]$.

2. Calculate $\nabla \mathscr{L}_{\boldsymbol{\Phi}_{eq_{k\,MAS}}^*}[i] = -\frac{P_1 T}{N} \rho_{k,j} \boldsymbol{r}_e[i] \hat{\boldsymbol{s}}^H[i] (\boldsymbol{H}_{eq_{k,j}}^{\boldsymbol{\Delta}_{k,j}}[i])^H$, where

$\boldsymbol{r}_e[i] = \boldsymbol{r}[i] - \sqrt{\frac{P_1 T}{N}} \sum_{k=1}^{n_r} \sum_{j=1}^{N} \rho_{k,j} \boldsymbol{\Phi}_{eq_{k,j}}[i] \boldsymbol{H}_{eq_{k,j}}^{\boldsymbol{\Delta}_{k,j}}[i] \hat{\boldsymbol{s}}[i]$.

3. Choose the optimal relay node by $\hat{k} = \operatorname{argmax}_{k=1,2,...,n_r} SNR_{k_{ins}} = \sum_{j=1}^{N} \frac{P_1 T \rho_{k,j}^2 \|\boldsymbol{\Phi}_{eq_{k,j}}[i] \boldsymbol{H}_{eq_{k,j}}^{\boldsymbol{\Delta}_{k,j}}[i]\|_F^2 \sigma_s^2}{N \rho_{k,j}^2 \|\boldsymbol{\Phi}_{eq_{k,j}}[i] \boldsymbol{G}_{eq_{k,j}}^{\boldsymbol{\Delta}_{k,j}}[i]\|_F^2 \sigma_{n_1}^2 + \sigma_d^2}$.

4. Update $\boldsymbol{\Phi}_{eq_{k,j}}[i]$ by $\boldsymbol{\Phi}_{eq_{k,j}}[i+1] = \boldsymbol{\Phi}_{eq_{k,j}}[i] + \beta \nabla \mathscr{L}_{\boldsymbol{\Phi}_{eq_{k\,MAS}}^*}[i]$.

5. Normalization: $\boldsymbol{\Phi}_{eq_{k,j}}[i] = \frac{\sqrt{\mathrm{P_R}} \boldsymbol{\Phi}_{eq_{k,j}}[i]}{\sqrt{\sum_{k=1}^{n_r} \mathrm{Tr}(\boldsymbol{\Phi}_{eq_{k,j}}[i] \boldsymbol{\Phi}_{eq_{k,j}}^H[i])}}$.

A summary of the DT-ACMORO algorithms is shown in Table 5.3, and different op-

portunistic relaying algorithms in [33] are considered. Note that the DT-ACMORO fully-opportunistic and opportunistic source algorithms need the CSI between the source node and the relays $\boldsymbol{f}_n$ in order to maximize the received SNR and choose the optimal antenna at the source node [33], where $\boldsymbol{f}_n$ is the $1 \times n_r N$ channel vector between the nth antenna at the source node and all the relay nodes.

5.5 Analysis of the Proposed DSTBC Schemes and the Algorithms in MAS and SAS

In this section, we will develop an analysis that explains the differences between the delayed DSTBCs in SAS and MAS in terms of the rank criterion and the error probability. The convergence properties of the proposed algorithms are briefly derived in this section as well. The ML criterion is employed in the proposed DT-ACMO algorithms and different DSTC schemes can be implemented with the proposed design, and the fundamental diversity order is not changed but further coding gains can be achieved.

5.5.1 Rank Criterion

The DSTBC schemes in SAS implements encoding across relay nodes, while the DSTBC schemes in MAS provides full STBCs at each relay node. For simplicity, we consider 2 single-antenna relay nodes for SAS and 1 relay node with 2 antennas for MAS with the same DSTC scheme for a fair comparison. In the standard STBCs in SAS such as the Alamouti scheme, the code word consists of matrices in the form

$$\boldsymbol{C} = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix}, \tag{5.45}$$

where s_1 and s_2 are symbols from a desired PSK or QAM constellation. The first antenna sends the first row in $\boldsymbol{C}$, s_1 and $-s_2^*$, to the destination node, and the second antenna sends the symbols s_2 and s_1^* during the transmission interval. It is important that the determinant of the code matrix is different from zero unless both $s_1 = 0$ and $s_2 = 0$.

Table 5.3: The DT-ACMORO SG Algorithms

1: **Initialization:**

Generate $\mathbf{\Phi}[0]$ randomly with the power constraint $\mathrm{Tr}(\mathbf{\Phi}_{eq_{k\,MAS}}\mathbf{\Phi}^{H}_{eq_{k\,MAS}}) \leq \mathrm{P_R}$,

2: **for** $k = 1$ **to** n_r**, do**

2-1: **DT-ACMORO AF Algorithm**

$\nabla \mathscr{L}_{\mathbf{\Phi}^{*}_{eq_{k\,MAS}}[i]} = -\frac{P_1 T}{N}\rho_{k,j}\boldsymbol{r}_e[i]\hat{\boldsymbol{s}}^{H}[i](\boldsymbol{H}^{\mathbf{\Delta}_{k,j}}_{eq_{k,j}}[i])^{H}$,
choose the optimal relay by
$\hat{k} = \mathrm{argmax}_{k=1,2,\ldots,n_r}\, SNR_{k_{ins}} = \sum_{j=1}^{N} \frac{P_1 T \rho^2_{k,j}\|\mathbf{\Phi}_{eq_{k,j}}[i]\boldsymbol{H}^{\mathbf{\Delta}_{k,j}}_{eq_{k,j}}[i]\|^2_F \sigma^2_s}{N\rho^2_{k,j}\|\mathbf{\Phi}_{eq_{k,j}}[i]\boldsymbol{G}^{\mathbf{\Delta}_{k,j}}_{eq_{k,j}}[i]\|^2_F \sigma^2_{n_1} + \sigma^2_d}$,

2-2: **DT-ACMORO Full-Opportunistic Algorithm**

$\nabla \mathscr{L}_{\mathbf{\Phi}^{*}_{eq_{k\,MAS}}[i]} = -\frac{P_1 T}{N}\rho_{k,j}\boldsymbol{r}_e[i]\hat{\boldsymbol{s}}^{H}[i](\boldsymbol{H}^{\mathbf{\Delta}_{k,j}}_{eq_{k,j}}[i])^{H}$,
choose the optimal relay by
$\hat{k} = \mathrm{argmax}_{k=1,2,\ldots,n_r}\, SNR_{k_{ins}} = \sum_{j=1}^{N} \frac{P_1 T \rho^2_{k,j}\|\mathbf{\Phi}_{eq_{k,j}}[i]\boldsymbol{H}^{\mathbf{\Delta}_{k,j}}_{eq_{k,j}}[i]\|^2_F \sigma^2_s}{N\rho^2_{k,j}\|\mathbf{\Phi}_{eq_{k,j}}[i]\boldsymbol{G}^{\mathbf{\Delta}_{k,j}}_{eq_{k,j}}[i]\|^2_F \sigma^2_{n_1} + \sigma^2_d}$,
and the choose the best antenna at the source node by
$\hat{n} = \mathrm{argmax}_{n=1,\ldots,N}\,|\boldsymbol{f}_n|^2$,

2-3: **DT-ACMORO Opportunistic Source Algorithm**

$\nabla \mathscr{L}_{\mathbf{\Phi}^{*}_{eq_{k\,MAS}}[i]} = -\frac{P_1 T}{N}\rho_{k,j}\boldsymbol{r}_e[i]\hat{\boldsymbol{s}}^{H}[i](\boldsymbol{H}^{\mathbf{\Delta}_{k,j}}_{eq_{k,j}}[i])^{H}$,
choose the best antenna at the source node by
$\hat{n} = \mathrm{argmax}\, SNR_{ins} = \sum_{k=1}^{n_r}\sum_{j=1}^{N} \frac{\|\mathbf{\Phi}_{eq_{k,j}}[i]\boldsymbol{H}^{\mathbf{\Delta}_{k,j}}_{eq_{k,j}}[i]\|^2_F}{\sigma^2_{n_1} + P_1\sigma^2_{\boldsymbol{F}_k}}$,

3: **Update:**

$\mathbf{\Phi}_{eq_{k,j}}[i]$ by $\mathbf{\Phi}_{eq_{k,j}}[i+1] = \mathbf{\Phi}_{eq_{k,j}}[i] + \beta\nabla\mathscr{L}_{\mathbf{\Phi}^{*}_{eq_{k\,MAS}}[i]}$,

4: **Normalization:**

$\mathbf{\Phi}_{eq_{k\,MAS}}[i] = \frac{\sqrt{\mathrm{P_R}}\mathbf{\Phi}_{eq_{k\,MAS}}[i]}{\sqrt{\sum_{k=1}^{n_r}\mathrm{Tr}(\mathbf{\Phi}_{eq_{k\,MAS}}[i]\mathbf{\Phi}^{H}_{eq_{k\,MAS}}[i])}}$.

The difference matrix of the standard Alamouti STBC scheme between two codewords can be derived as

$$\Delta \boldsymbol{C} = \boldsymbol{C}_1 - \boldsymbol{C}_2 = \begin{bmatrix} s_1^{(1)} & -s_2^{(1)*} \\ s_2^{(1)} & s_1^{(1)*} \end{bmatrix} - \begin{bmatrix} s_1^{(2)} & -s_2^{(2)*} \\ s_2^{(2)} & s_1^{(2)*} \end{bmatrix} = \begin{bmatrix} \Delta s_1 & -\Delta s_2^* \\ \Delta s_2 & \Delta s_1^* \end{bmatrix}. \quad (5.46)$$

Unless $\boldsymbol{C}_1 = \boldsymbol{C}_2$, the rank of the difference matrix $\Delta \boldsymbol{C}$ is full. According to the rank criterion in [41] for the STCs, the Alamouti scheme achieves full diversity.

However, the standard Alamouti scheme is not delay tolerant with SAS. Consider a delay profile $\boldsymbol{\Delta} = [0\ 1]$ in which the second transmission from the second relay node is one-symbol slot later than the first one. Then the difference matrix is given by

$$\begin{aligned} \Delta \boldsymbol{C}^{\boldsymbol{\Delta}} = \boldsymbol{C}_1^{\boldsymbol{\Delta}} - \boldsymbol{C}_2^{\boldsymbol{\Delta}} &= \begin{bmatrix} s_1^{(1)} & -s_2^{(1)*} & 0 \\ 0 & s_2^{(1)} & s_1^{(1)*} \end{bmatrix} - \begin{bmatrix} s_1^{(2)} & -s_2^{(2)*} & 0 \\ 0 & s_2^{(2)} & s_1^{(2)*} \end{bmatrix} \\ &= \begin{bmatrix} \Delta s_1 & -\Delta s_2^* & 0 \\ 0 & \Delta s_2 & \Delta s_1^* \end{bmatrix}. \end{aligned} \quad (5.47)$$

If $s_1^{(1)} = s_1^{(2)}$ but $s_1^{(2)} \neq s_1^{(2)}$, the difference matrix in (5.47) is

$$\Delta \boldsymbol{C}^{\boldsymbol{\Delta}} = \begin{bmatrix} 0 & -\Delta s_2^* & 0 \\ 0 & \Delta s_2 & 0 \end{bmatrix}, \quad (5.48)$$

which suffers a rank reduction. Hence, the Alamouti scheme in SAS cannot achieve the full diversity when a time delay happens.

In contrast to the SAS case, in MAS with $n_r = 2$ relay nodes, the standard Alamouti scheme can be obtained in each individual relay node, with the delay profile $\boldsymbol{\Delta} = [0\ 1]$ the delayed matrix is described by

$$\boldsymbol{C}^{\boldsymbol{\Delta}} = \boldsymbol{C}^{\delta_1} + \boldsymbol{C}^{\delta_2} = \begin{bmatrix} s_1 & -s_2^* & 0 \\ s_2 & s_1^* & 0 \end{bmatrix} + \begin{bmatrix} 0 & s_1 & -s_2^* \\ 0 & s_2 & s_1^* \end{bmatrix} = \begin{bmatrix} s_1 & s_1 - s_2^* & -s_2^* \\ s_2 & s_1^* + s_2 & s_1^* \end{bmatrix}, \quad (5.49)$$

then the difference matrix between the 2 codewords is given by

$$\begin{aligned}\Delta \boldsymbol{C}^{\boldsymbol{\Delta}} = \boldsymbol{C}_1^{\boldsymbol{\Delta}} - \boldsymbol{C}_2^{\boldsymbol{\Delta}} &= \begin{bmatrix} s_1^{(1)} & s_1^{(1)} - s_2^{(1)^*} & -s_2^{(1)^*} \\ s_2^{(1)} & s_1^{(1)^*} + s_2^{(1)} & s_1^{(1)^*} \end{bmatrix} - \begin{bmatrix} s_1^{(2)} & s_1^{(2)} - s_2^{(2)^*} & -s_2^{(2)^*} \\ s_2^{(2)} & s_1^{(2)^*} + s_2^{(2)} & s_1^{(2)^*} \end{bmatrix} \\ &= \begin{bmatrix} \Delta s_1 & \Delta s_1 - \Delta s_2^* & -\Delta s_2^* \\ \Delta s_2 & \Delta s_1^* + \Delta s_2 & \Delta s_1^* \end{bmatrix},\end{aligned} \tag{5.50}$$

which achieves the same rank as (5.49) even when $\Delta s_1 = 0$ or $\Delta s_2 = 0$.

5.5.2 Error Probability

In the SAS environment, the received matrix at the destination node in (5.3) and (5.5) can be rewritten as

$$\boldsymbol{R}_{RD_{SAS}}[i] = \boldsymbol{G}_{RD_{SAS}}[i]\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}}[i] + \boldsymbol{N}_{RD}[i], \tag{5.51}$$

where $\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}}[i] = \boldsymbol{\Delta}[i]\boldsymbol{\Phi}_{SAS}[i]\boldsymbol{M}_{RD_{SAS}}[i]$ denotes the delayed randomized DSTCs encoding across the relays. Let $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ denotes two possible modulated symbol vectors, and given the code matrix $\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}}[i]$, the PEP of decoding $\boldsymbol{s}_1$ to $\boldsymbol{s}_2$ has the following Chernoff upper bound [41]

$$\begin{aligned} PEP(\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}}) &\leq E_{\boldsymbol{G}_{RD_{SAS}}}\left[\exp\left(-\frac{E_s}{4N_0}Tr(\boldsymbol{G}_{RD_{SAS}}\boldsymbol{R}_{SAS}\boldsymbol{G}^H_{RD_{SAS}})\right)\right] \\ &= \int_{\boldsymbol{G}_{RD_{SAS}}} \exp\left(-\frac{E_s}{4N_0}Tr(\boldsymbol{G}_{RD_{SAS}}\boldsymbol{R}_{SAS}\boldsymbol{G}^H_{RD_{SAS}})\right) \\ &\quad \exp\left(-Tr(\boldsymbol{G}_{RD_{SAS}}\boldsymbol{G}^H_{RD_{SAS}})\right) d\boldsymbol{G}_{RD_{SAS}} \\ &= \frac{1}{\det(\boldsymbol{I} + \frac{E_s}{4N_0}\boldsymbol{R}_{SAS})^N}, \end{aligned} \tag{5.52}$$

where $\boldsymbol{R}_{SAS} = (\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}}(\boldsymbol{s}_1) - \boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}}(\boldsymbol{s}_2))(\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}}(\boldsymbol{s}_1) - \boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}}(\boldsymbol{s}_2))^H$ denotes the auto-correlation matrix of the difference code matrices. The conditional average PEP of the delayed codeword matrix over all possible symbols and delay profiles can be written as

$$P_{ave}(\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}}) = E_{\boldsymbol{\Delta}}\left[E_{\boldsymbol{s}_1,\boldsymbol{s}_2\in\boldsymbol{S}}\left[PEP(\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}})\right]\right], \tag{5.53}$$

where $\boldsymbol{S}$ is the set of all possible combinations of the symbol vectors.

The Chernoff upper bound in (5.52) with a high SNR scenario can be approximately written as

$$P_{\boldsymbol{s}_1,\boldsymbol{s}_2}(\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{SAS}}) \leq \frac{1}{(\frac{E_s}{4N_0})^{rN}\prod_{n=1}^{r}\lambda_n^N}, \tag{5.54}$$

where $r = \mathrm{rank}(\boldsymbol{R}_{SAS})$ denotes the rank of the auto-correlation matrix and λ_n denotes the nth eigenvalue of $\boldsymbol{R}_{SAS}$. The diversity order of the code matrix is given by rN, and is determined by the minimum rank of the correlated code matrix. According to (5.48), the rank reduction affects the diversity order of the code matrix so that the Alamouti matrix code in SAS cannot achieve the full diversity and is not a delay tolerant code.

In MAS with $n_r = 2$ relay nodes, the received symbol matrix at the destination node can be rewritten as

$$\begin{aligned}\boldsymbol{R}_{RD_{MAS}}[i] &= \sum_{k=1}^{n_r}\boldsymbol{G}_{R_kD_{MAS}}[i]\boldsymbol{M}^{\boldsymbol{\Delta}}_{R_kD_{MAS}}[i] + \boldsymbol{N}_{RD}[i] \\ &= \boldsymbol{G}_{RD_{MAS_{eq}}}[i]\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{MAS_{eq}}}[i] + \boldsymbol{N}_{RD}[i],\end{aligned} \tag{5.55}$$

where $\boldsymbol{M}^{\boldsymbol{\Delta}}_{R_kD_{MAS}}[i]$ is the code matrix at the kth relay node with the delay profile, the equivalent channel matrix is denoted as $\boldsymbol{G}_{RD_{MAS_{eq}}}[i]$ and the delayed code matrix is $\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{MAS_{eq}}}[i]$. The PEP of the decoding at the destination node is upper bounded by the Chernoff bound derived as

$$\begin{aligned}PEP(\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{MAS_{eq}}}) &\leq E_{\boldsymbol{G}_{RD_{MAS_{eq}}}}\left[\exp\left(-\frac{E_s}{4N_0}Tr(\boldsymbol{G}_{RD_{MAS_{eq}}}\boldsymbol{R}_{MAS}\boldsymbol{G}^H_{RD_{MAS_{eq}}})\right)\right] \\ &= \frac{1}{\det(\boldsymbol{I}+\frac{E_s}{4N_0}\boldsymbol{R}_{MAS})^N},\end{aligned} \tag{5.56}$$

where $\boldsymbol{R}_{MAS} = (\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{MAS_{eq}}}(\boldsymbol{s}_1) - \boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{MAS_{eq}}}(\boldsymbol{s}_2))(\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{MAS_{eq}}}(\boldsymbol{s}_1) - \boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{MAS_{eq}}}(\boldsymbol{s}_2))^H$ denotes the auto-correlation matrix of the difference code matrices. The Chernoff upper bound of the conditional average PEP with a high SNR over all possible symbols and delay profiles can be expressed as

$$P_{ave}(\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{MAS_{eq}}}) = E_{\boldsymbol{\Delta}}\left[E_{\boldsymbol{s}_1,\boldsymbol{s}_2\in\boldsymbol{S}}\left[PEP(\boldsymbol{M}^{\boldsymbol{\Delta}}_{RD_{MAS_{eq}}})\right]\right] \leq \frac{1}{(\frac{E_s}{4N_0})^{rN}\prod_{n=1}^{r}\lambda_n^N}, \tag{5.57}$$

where $r = \mathrm{rank}(\boldsymbol{R}_{MAS})$ stands for the rank of the auto-correlation matrix and λ_n is the nth eigenvalue of $\boldsymbol{R}_{MAS}$. The diversity order depends on the minimum rank r and according to (5.50) that is $r = N$ which guarantees that the full diversity order can be achieved.

According to [41], the rank and diversity analysis can be considered as the design criterion of the delay-tolerant DSTC schemes in cooperative communication systems in a high SNR scenario. The optimization algorithm derived in the previous section ensures that the design of the space-time coding scheme is delay-tolerant for any given SNR and achieve a high coding gain, and the ML detection algorithm is employed in order to achieve the full receive diversity.

The design of the DT-ACMORO algorithms achieves a lower BER performance by employing the proposed DT-ACMORO algorithms in cooperative systems with opportunistic relaying methods in [33]. As a result, the PEP analysis of the DT-ACMORO algorithms in MAS and SAS is the same as that of the DT-ACMO algorithms. The only difference is the channel matrix $\boldsymbol{G}_{RD}[i]$ in (5.51) and (5.55) denotes the $N \times N$ channel matrix between the optimal relay node and the destination node in SAS and MAS, respectively.

5.5.3 Convergence Analysis

As shown in the previous sections, the proposed ML based DT-ACMO algorithms allow the optimization of the code matrix $\boldsymbol{\Phi}[i]$ update adaptively. In this section, we briefly study the convergence of the DT-ACMO RLS and SG algorithms.

Convergence Analysis of DT-ACMO Algorithms in MAS

The ML and RLS based DT-ACMO algorithms in MAS optimize the adjustable code matrix, and we can analyze the Hessian matrix of (5.9) and check its positive (semi-)definiteness. By taking the second-order partial derivatives of the Lagrangian cost func-

tion in (5.9), we can obtain the 2nd order conditions derived as

$$\begin{aligned} H(\mathscr{L}) &= \frac{\partial}{\partial \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]}\left(\frac{\partial \mathscr{L}}{\partial (\boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i])^*}\right) \\ &= \frac{\partial}{\partial \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]}\left(-\boldsymbol{r}_{MAS}[i]\boldsymbol{s}^H[i]\boldsymbol{G}^H_{eq_{k\,MAS}}[i] + \boldsymbol{\Phi}^{\boldsymbol{\Delta}}_{eq_{k\,MAS}}[i]\boldsymbol{G}_{eq_{k\,MAS}}[i]\boldsymbol{s}[i]\boldsymbol{s}^H[i]\boldsymbol{G}^H_{eq_{k\,MAS}}[i]\right) \\ &= \boldsymbol{G}_{eq_{k\,MAS}}[i]\boldsymbol{s}[i]\boldsymbol{s}^H[i]\boldsymbol{G}^H_{eq_{k\,MAS}}[i], \end{aligned} \tag{5.58}$$

where the first term $\boldsymbol{s}[i]\boldsymbol{s}^H[i]$ is a positive matrix and the rest of the terms denotes the multiplication of the equivalent channel vectors which is a positive-definite matrix and the problem is convex. So that we conclude that the Hessian matrix of the Lagrangian cost function is a positive-definite matrix so that the RLS based DT-ACMO algorithms converge to the global optimum under the usual assumptions used to prove the convergence of these algorithms for convex problems.

Convergence Analysis of DT-ACMO Algorithms in SAS

The ML-based SG ACMO algorithm for SAS optimize the adjustable code matrix, and by taking the second-order partial derivatives of the Lagrangian cost function in (5.24), we can obtain

$$\begin{aligned} H(\mathscr{L}) &= \frac{\partial}{\partial \boldsymbol{\Phi}_{eq_{k\,SAS}}[i]}\left(\frac{\partial \mathscr{L}}{\partial (\boldsymbol{\Phi}_{eq_{k\,SAS}}[i])^*}\right) \\ &= \frac{\partial}{\partial \boldsymbol{\Phi}_{eq_{k\,SAS}}[i]}\left((\boldsymbol{r}_{SAS}[i] - \sum_{k=1}^{n_r N}\boldsymbol{\Delta}_k\boldsymbol{G}_{eq_{k\,SAS}}[i]\boldsymbol{\Phi}_{eq_{k\,SAS}}[i]\hat{\boldsymbol{s}}[i])\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{k\,SAS}}[i]\boldsymbol{\Delta}^H_k\right) \\ &= \boldsymbol{\Delta}_k\boldsymbol{G}_{eq_{k\,SAS}}[i]\hat{\boldsymbol{s}}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{k\,SAS}}[i]\boldsymbol{\Delta}^H_k, \end{aligned} \tag{5.59}$$

As shown in the previous section, we conclude that the Hessian matrix of the Lagrangian cost function is a positive-definite matrix so that the ML based SG DT-ACMO algorithm in SAS converge to the global optimum under the usual assumptions used to prove the convergence of these algorithms for convex problems.

5.6 Simulations

The simulation results are provided in this section to assess the proposed schemes and algorithms. The cooperative MIMO system considered employs AF and DF protocols with the distributed Alamouti (D-Alamouti) STBC scheme in [5], randomized Alamouti (R-Alamouti) in [6] and linear dispersion code (LDC) in [42] using BPSK and QPSK modulation in a quasi-static block fading channel with AWGN. The effect of the direct link is also considered. It is possible to employ different DSTC schemes with a simple modification and to incorporate the proposed algorithms. The cooperative MAS configuration equipped with $n_r = 1, 2$ relay nodes and $N = 2$ antennas at each node in both cases, while the cooperative SAS scheme employed $N = 1, 2$ antennas at the source node and the destination node and $n_r = 1, 2$ relay nodes with a single antenna in both cases. In the simulations, we set the symbol power σ_s^2 as equal to 1, and the power of the adjustable code matrix in the DT-ACMO algorithm is normalized. The SNR in the simulations is the received SNR which is calculated by $SNR = \frac{\|\boldsymbol{D}_D[i]\|_F^2}{\|\boldsymbol{n}_D[i]\boldsymbol{n}_D^H[i]\|_F^2}$.

The BER performance of the D-Alamouti and the R-Alamouti in the cooperative MAS and SAS configurations without delay are shown in Fig. 5.2. The solid curves are the BER curves without the direct link (DL) and the dashed ones are that with the DL. SAS employing a single antenna at each node and the AF strategy with and without the DL have the worst BER performance compared to SAS and MAS employing multiple antennas and relay nodes. The cooperation of the DL helps SAS to achieve a higher diversity gain and lower the BER. In SAS employing $N = 2$ antennas at the source node and the destination node with $n_r = 2$ single-antenna relay nodes, and MAS employing the same number of the antennas at the source node and the destination node with $n_r = 1$ relay node using $N = 2$ antennas, the superposition of the BER curves can be observed. With the cooperation of the DL, a higher order of diversity can be achieved. The simulation results lead to the conclusion that with the same number of antennas at the source node and the destination node in SAS and MAS, if the number of the relay nodes with single antenna is the same as the number of multi-antenna relay nodes, the same diversity order and BER performance can be achieved when the synchronization is perfect at the relay node. The diversity order of the DF MAS networks without the direct link is equal to 4, while that of the AF SAS networks is equal to 2. According to the figure, if there are the same

number of antennas at source and destination in MAS and SAS, and if the total number of antennas among the relays is equal, the same diversity order can be achieved by these two different systems.

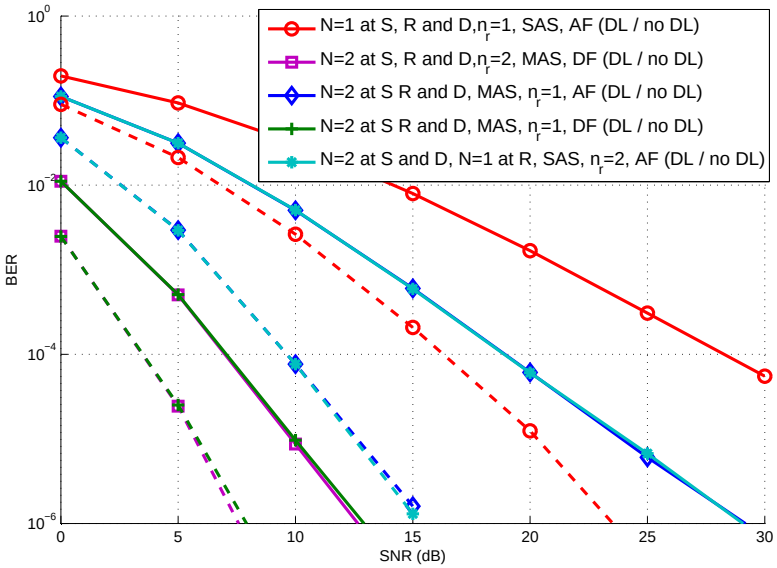

Figure 5.2: BER performance vs. SNR for SAS and MAS employing the Alamouti schemes without the delay profile

In Fig. 5.3, BER curves of different DSTC schemes without the DL using an ML detector are compared using the SAS scheme with $N = 2$ relay nodes. Different delay profiles are considered to evaluate the performance of D-Alamouti, R-Alamouti using BPSK and R-LDC using QPSK in SAS. As shown in Fig. 5.3, the R-Alamouti scheme improves the performance by about 2dB without the DL compared to the D-Alamouti scheme when perfect synchronization is considered. When the delay profile $\mathbf{\Delta} = [0, 1]$ is considered among the relay nodes, it is shown that the D-Alamouti and R-Alamouti are not delay tolerant schemes in SAS. The R-LDC schemes using QPSK modulation constellation with different delay profiles are shown in Fig. 5.3. It is shown in Fig. 5.3 that the R-LDC in SAS is not a delay tolerant scheme because of the degradation of the BER performance. When employing the proposed DT-ACMO algorithm in SAS, a performance improvement is obtained in R-Alamouti and R-LDC schemes. However, the BER curves of the Alamouti and LDC schemes using DT-ACMO algorithms are still worse than that of the Alamouti and LDC schemes with perfect synchronization among

the relays.

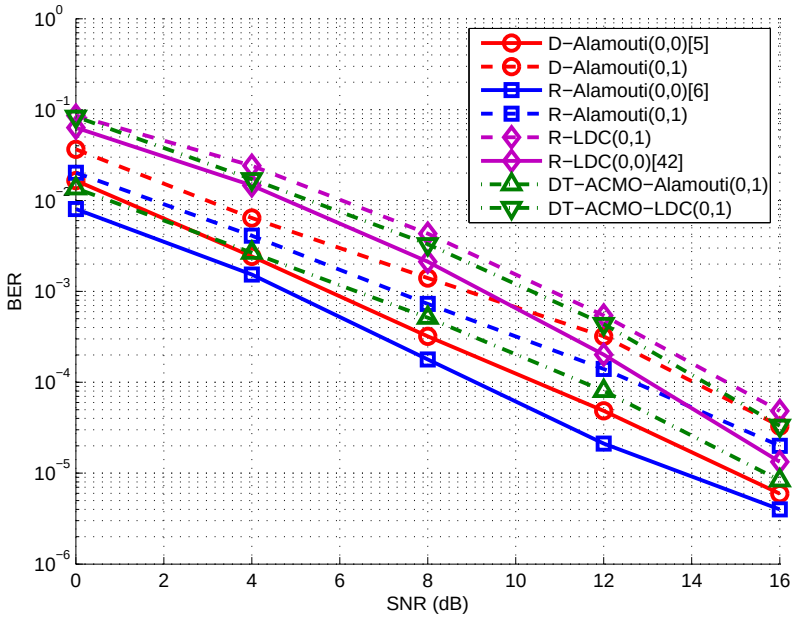

Figure 5.3: BER performance vs. SNR for SAS

The proposed DT-ACMO algorithm is compared with different DSTC schemes and different delay profiles in MAS in Fig. 5.4. In the proposed algorithms the DF strategy is employed which requires re-encoding at the relay nodes, and a full STC scheme is obtained at each of the $n_r = 2$ relay nodes. The results illustrate that without the DL, by making use of the randomized encoding technique in [6], a 2dB BER performance improvement can be achieved compared to the D-Alamouti scheme with and without the delay. The simulation results illustrate that MAS with different delay profiles when the same DSTC schemes are considered, a 1dB performance improvement is obtained by the delayed STC compared to a non-delayed one. It is because when the full DSTC schemes can be obtained at the relay nodes as employed in MAS, at the destination node a delayed symbol matrix with increased dimension is received which means the observation window is bigger than that with the non-delayed MAS. When the DT-ACMO algorithms are employed in Alamouti and LDC schemes, better BER performances can be achieved. With the consideration of the delay profiles, the results indicate that the diversity order will not be reduced, and using the DT-ACMO RLS algorithm an improved performance is achieved with 3dB of gain as compared to employing the RSTC algorithm in [6] and

5dB of gain as compared to employing the traditional DSTBC algorithm. In the case of $\delta_{max} = 2$, a increase of diversity order can be obtained according to the simulation results shown in Fig. 5.4. It is because we assume the delays are between the relays in MAS and SAS, so if the value of the maximum delay coefficient δ_{max} is bigger than the value of T in the DSTC scheme, a non-overlapped STC scheme will be received at the destination node and in this case, individual DSTC schemes without superposing with others will be received at the destination node and the diversity order increases.

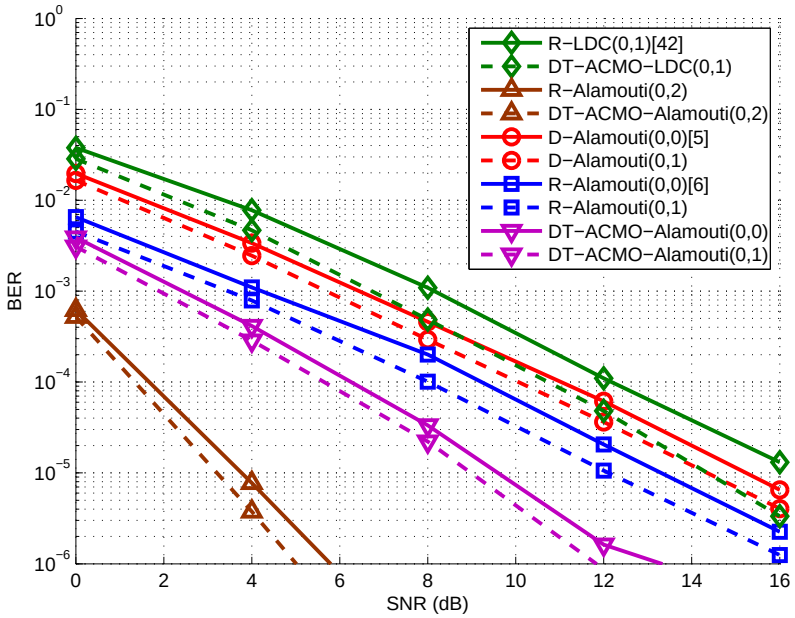

Figure 5.4: BER performance vs. SNR for DT-ACMO RLS algorithm for MAS

The proposed DT-ACMORO algorithms with Alamouti scheme and an ML receiver is compared in MAS in Fig. 5.5. $N = 2$ antennas at each node and $n_r = 2$ relay nodes are employed. Different delay profiles are considered in different relay nodes. It is shown in the figure that the BER result of the cooperative system with DSTC schemes and AF protocol achieves a diversity order of 2, and by employing opportunistic relaying (OR), fully-opportunistic (FO) and opportunistic source (OS) algorithms, a 2dB to 3dB BER improvement can be further achieved, respectively. In [33], the benefits of employing three opportunistic relaying algorithms are explained. In the proposed DT-ACMORO algorithms, the DT-ACMO algorithms are combined with different opportunistic relaying algorithms in order to achieve a BER improvement as shown in the previous section.

According to the simulation results in Fig. 5.5, a 2dB to 3dB gains can be achieved by using DT-ACMORO compared to the systems only use opportunistic relaying algorithms. It is worth to mention that different DT-ACMORO algorithms require different computational complexities as shown in Table III, and the DT-ACMORO-FO algorithm which requires the choice of the best antenna at the source node and the use of the best relay node can achieve the best BER performance as compared with the DT-ACMORO-OR and DT-ACMORO-OS algorithms.

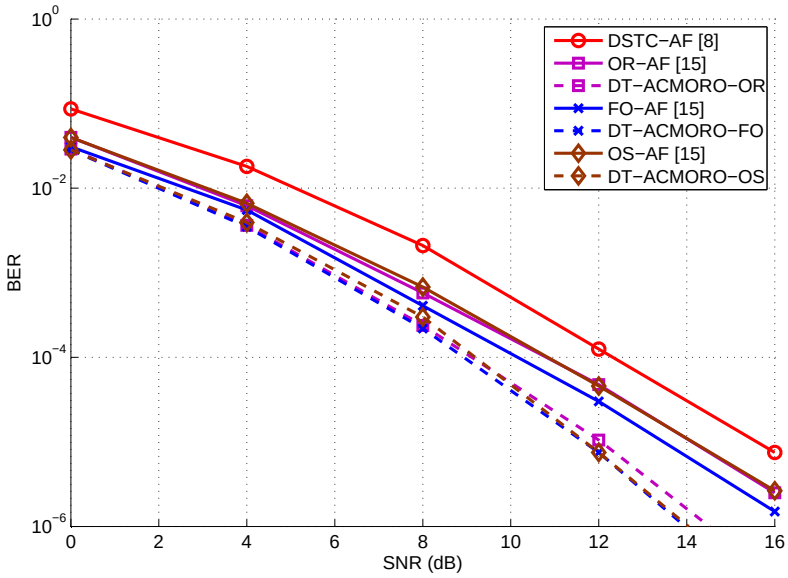

Figure 5.5: BER performance vs. SNR for DT-ACMORO SG algorithm for MAS

In Fig. 5.6, the proposed DT-ACMORO SG algorithms are employed among the relay nodes in SAS. Compared to the opportunistic algorithms in [33], the DT-ACMORO algorithms achieve a 2dB to 4dB improvement compared to the original opportunistic relaying algorithms as shown in Fig. 5.6. When comparing the curves in Fig. 5.6 to those in Fig. 5.5, it is noticed that the diversity order of the curves in Fig. 5.6 and that in Fig. 5.5 are the same because the same DSTBC schemes are employed in MAS and SAS, however the BER performances in Fig. 5.3 are worse that that in Fig. 5.5. The reason for worse BER performances in SAS is due to the lower number of antennas used in SAS according to the opportunistic relaying algorithms. In SAS $n_r = 2$ relay nodes with a single antenna are employed and after the optimization, only one relay node with $N = 1$ antenna is chosen to transmit the re-encoded information symbols. On the contrary, the best relay

node in MAS contains $N = 2$ antennas to forward the re-encoded DSTC scheme to the destination which ensures the complete DSTC scheme is received at the destination node.

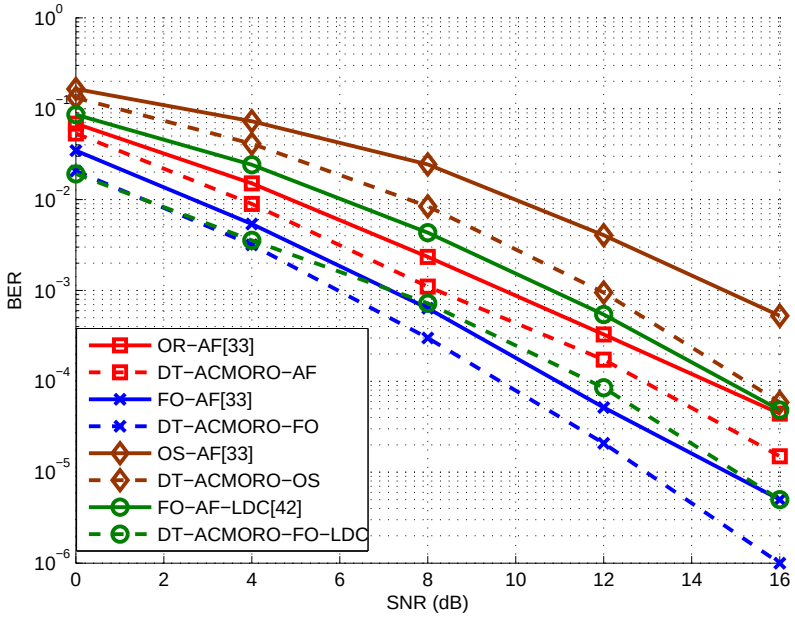

Figure 5.6: BER performance vs. SNR for DT-ACMORO SG algorithm for SAS

5.7 Summary

We have proposed a delay-tolerant adjustable code matrix optimization (DT-ACMO) scheme and algorithms for cooperative MIMO systems with a feedback channel using an ML receiver at the destination node to mitigate the effect of the delay associated with DSTCs from relay nodes. Two types of cooperative systems have been analyzed by comparing the pairwise error probability of the delayed DSTCs and the rank criterion. In order to achieve a better BER performance and eliminate the computation complexity at the destination node, we have introduced the SG and RLS algorithms in the proposed design. We have extended the proposed algorithms to cooperative MIMO systems with the AF strategy and opportunistic relaying algorithms in order to devise delay-tolerant adjustable code matrix opportunistic relaying optimization (DT-ACMORO) algorithms.

The simulation results have illustrated the performance of the two different cooperative systems with the same number of transmit and receive antennas, and the advantage of the proposed DT-ACMO and DT-ACMORO algorithms by comparing them with the cooperative network employing the traditional DSTC scheme and the RSTC scheme. The proposed algorithms can be used with different DSTC schemes and can also be extended to cooperative systems with any number of antennas.

Chapter 6

Conclusions and Future Work

6.1 Summary of the Work

In this thesis, novel distributed space-time coding schemes and optimization algorithms in cooperative relaying networks have been proposed and investigated. In particular, cooperative systems with multi-antenna relays and single-antenna relays have been considered along with optimization strategies to design space-time coding schemes and algorithms. Iterative adaptive algorithms were derived and used to reduce the computational complexity of the proposed optimization techniques. The background techniques required to the development of the main contributions of this thesis have been discussed at the beginning of Chapters 3, 4 and 5. The algorithms and designs in the Chapters 3, 4 and 5 are proposed in order to address the issues discussed in the literature and the drawbacks of prior art. Also a literature review of the designs and algorithms related to the research in this thesis has been done and the designs and algorithms have been compared with the proposed algorithms in terms of BER performance, sum rate performance and computational complexity.

In Chapter 3, joint adaptive power allocation and receiver design algorithms according to different criteria subject to a power constraint in a cooperative MIMO system are investigated. Joint iterative estimation algorithms with low computational complexity for computing the power allocation parameters and the linear receive filter has been derived,

which allows the proposed algorithms utilized in mobile communication networks and wireless sensor networks. Simulations showed the proposed algorithms can achieve improved performance compared to existing techniques.

In Chapter 4, novel distributed space-time coding schemes are designed for cooperative MIMO networks. Firstly, centralized adaptive robust matrix optimization (C-ARMO) algorithms have been designed for cooperative MIMO system with AF protocol using a linear MMSE receive filter and an ML receiver at the destination node. In order to eliminate the need for a feedback channel, a fully-distributed ARMO (FD-ARMO) algorithm has been devised and investigated which can achieve a similar coding gain without the feedback as compared to the C-ARMO algorithms. The pairwise error probability of introducing the adaptive DSTC in a cooperative MIMO network with the AF protocol has been derived. The simulation results illustrate the advantage of the proposed ARMO algorithms and the proposed algorithms can be used with different DSTC schemes as well as different cooperation protocols.

In Chapter 5, the proposed algorithms in Chapter 4 were extended to a delay-tolerant adjustable code matrix optimization (DT-ACMO) scheme and algorithms. Two types of cooperative systems have been analyzed by comparing the pairwise error probability of the delayed DSTCs and the rank criterion. The SG and RLS algorithms are utilized in the proposed designs in order to achieve a better BER performance and eliminate the computation complexity at the destination node. The opportunistic relaying algorithms are considered in the proposed designs in order to devise delay-tolerant adjustable code matrix opportunistic relaying optimization (DT-ACMORO) algorithms. The simulation results have illustrated advantages of the proposed algorithms compared to existing designs.

Modern mobile and wireless sensor network communications require energy saving and fast transmission with low errors. With advanced detection and estimation algorithms, mobile devices and sensors can restrict the energy used in the detection procedure in order to achieve a longer battery life. The adaptive optimization techniques with fast convergence property used in the proposed algorithms in this thesis can be considered as a key advantage for practical implementation in the real world. The tables of the comparison of the computational complexity of the proposed algorithms and the existing

algorithms indicate the advantage of the proposed coding designs. The novel joint design of the distributed space-time coding schemes and receive filter shown in the thesis focus on the improvement of the BER performance with low computational complexity. Different existing DSTCs can be used in the proposed coding design which widens the utilization of the proposed schemes. In addition, the delay issue in cooperative relaying networks is addressed by the proposed DT-ACMORO algorithms which provide a delay tolerant DSTC designs with promising error rate and convergence rate. With the low-computational-complexity encoding and decoding procedures shown in the thesis, the proposed algorithms can be easily adopted to existing cooperative MIMO systems with different protocols, i.e., wireless sensor networks and systems with multiple antennas.

6.2 Future Work

The algorithms described in this thesis assume the channel between the nodes are flat fading and unchanged during one transmission period. The fast fading scenario is not considered and the DSTC designs based on the fast fading channel are not considered in this thesis. In order to widen the application of the algorithms, the design and operation of the proposed algorithms in fast fading channels should be considered in the future.

The designs in Chapter 3 and 5 require feedback to transmit the power allocation matrices or optimal coding matrices to the source node and the relays. The BSC feedback and limited feedback channel are considered in the designs, and the proposed algorithms suffer a loss in performance because of the feedback errors. In the future, the robustness of the proposed algorithms should be considered and further researched.

In the proposed designs, stochastic gradient and recursive least squares adaptive optimization algorithms are used in order to reduce the computational complexity of the optimization techniques. However, the slow convergence rate of the adaptive optimization algorithms is a weakness if the channel is fast fading. Therefore, developing adaptive optimization algorithms with faster convergence rate is another area for future research.

As cooperative MIMO networks are widely used in modern mobile communications, such as wireless sensor networks, multi-user cooperative MIMO systems should be

considered. The principles and designs of a single-user scenario are different from that of a multi-user scenario so that in order to achieve a wide application of the proposed algorithms more work is needed in future research.

Appendix A

Derivation of Equation (4.7) and (4.8)

We show how to obtain the expression of the linear MMSE receive filter $\boldsymbol{w}_j[i]$ and the adjustable code matrix $\boldsymbol{\Phi}_{eq_{kj}}[i]$ in equation (4.7) and (4.8) in Section III in the following.

The MSE optimization problem is given by

$$[\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_k}[i]] = \arg \min_{\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_k}[i]} E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right], \; s.t. \; \mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^H[\mathrm{i}]) \leq \mathrm{P_R}.$$

We define a cost function associated with the optimization problem above and expand it as follows

$$\begin{aligned}\mathscr{L} &= E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right] + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^H[\mathrm{i}]) - \mathrm{P_R}) \\ &= E\left[s_j[i]s_j^*[i]\right] - \boldsymbol{w}_j^H[i]E\left[\boldsymbol{r}[i]s_j^*[i]\right] - E\left[s_j[i]\boldsymbol{r}^H[i]\right]\boldsymbol{w}_j[i] + \boldsymbol{w}_j^H[i]E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]\boldsymbol{w}_j[i] \\ &\quad + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^H[\mathrm{i}]) - \mathrm{P_R}),\end{aligned} \tag{A.1}$$

where λ stands for the Lagrange multiplier and should be determined before the calculation. It is worth to notice that the first, the third and the fifth terms are not functions of $\boldsymbol{w}_j^H[i]$, so by taking the gradient of $\mathscr{L}$ with respect to $\boldsymbol{w}_j^H[i]$ and equating the terms to 0, we can obtain

$$\mathscr{L}'_{\boldsymbol{w}_j^*[i]} = -E\left[\boldsymbol{r}[i]s_j^*[i]\right] + E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]\boldsymbol{w}_j[i] = 0. \tag{A.2}$$

By moving the first term in (A.2) to the right-hand side and by multiplying the inverse of the auto-correlation of the received symbol vector, we obtain the expression of the

linear MMSE receive filter as $\boldsymbol{w}_j[i] = \boldsymbol{R}^{-1}\boldsymbol{p}$, where the auto-correlation matrix $\boldsymbol{R} = E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]$ and the cross-correlation vector $\boldsymbol{p} = E\left[\boldsymbol{r}[i]s_j^*[i]\right]$.

In order to obtain the expression of the adjustable code matrix $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ we have to rewrite the received symbol vector $\boldsymbol{r}[i]$ as

$$\boldsymbol{r}[i] = \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] + \boldsymbol{n}_{RD}[i] = \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] + \boldsymbol{n}_{RD}[i], \tag{A.3}$$

where $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ denotes the adjustable code matrix assigned to the jth received symbol $\tilde{s}_{SR_{k_j}}[i]$ at the kth relay node, and $\boldsymbol{g}_{eq_{k_j}}[i]$ stands for the jth column of the equivalent channel matrix $\boldsymbol{G}_{eq_k}[i]$. By substituting (A.3) into (A.1), the expression of $\mathscr{L}$ can be written as

$$\begin{aligned}\mathscr{L} =& E\left[s_j[i]s_j^*[i]\right] - \boldsymbol{w}_j^H[i]E[(\tilde{\boldsymbol{r}} + \boldsymbol{n}_{RD}[i])s_j^*[i]] - E[s_j[i](\boldsymbol{w}_j^H[i](\tilde{\boldsymbol{r}} + \boldsymbol{n}_{RD}[i]))^H] \\ &+ E[(\boldsymbol{w}_j^H[i](\tilde{\boldsymbol{r}} + \boldsymbol{n}_{RD}[i]))^H\boldsymbol{w}_j^H[i](\tilde{\boldsymbol{r}} + \boldsymbol{n}_{RD}[i])] + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^H[\mathrm{i}]) - \mathrm{P_R}),\end{aligned}$$

where $\tilde{\boldsymbol{r}} = \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i]$. We do not have to consider the first and the second terms because they are not functions of $\boldsymbol{\Phi}_{eq_{k_j}}^H[i]$ so taking the gradient of $\mathscr{L}$ with respect to $\boldsymbol{\Phi}_{eq_{k_j}}^*[i]$ these terms will disappear. The last three terms contain the sum of the adjustable code matrices, and we focus on the exact jth code matrix we need and consider the rest of the sum terms as constants. We can rewrite $\mathscr{L}$ as

$$\begin{aligned}\mathscr{L} =& - E\left[s_j[i](\boldsymbol{w}_j^H[i]\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}\tilde{s}_{SR_{k_j}}[i])^H\right] + \lambda(\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{\Phi}_{eq_{k_j}}^H[i] - \mathrm{P_R}\boldsymbol{I}) \\ &+ E[(\boldsymbol{w}_j^H[i]\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i])^H\boldsymbol{w}_j^H[i]\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i]],\end{aligned} \tag{A.4}$$

and by taking the gradient of $\mathscr{L}$ in (A.4) with respect to $\boldsymbol{\Phi}_{eq_{k_j}}^*[i]$ and equating the terms to zero, we can obtain $\boldsymbol{\Phi}_{eq_{k_j}}[i] = \tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{P}}$, where $\tilde{\boldsymbol{R}} = E\left[s_j[i]\tilde{s}_{SR_{k_j}}[i]\boldsymbol{w}_j[i]\boldsymbol{w}_j^H[i] + \lambda\boldsymbol{I}\right]$ and $\tilde{\boldsymbol{P}} = E\left[s_j[i]\tilde{s}_{SR_{k_j}}[i]\boldsymbol{w}_j[i]\boldsymbol{g}_{eq_{k_j}}^H[i]\right]$.

Appendix B

Derivation of Equation (4.10)

We show the detailed derivation of the C-ARMO SG algorithm in this section. First, we have to rewrite the received symbol vector $\boldsymbol{r}_{R_kD}$ transmitted from the kth relay node. By employing the AF cooperative strategy and space-time coding schemes at the relay node, the received symbol vector at the relay nodes will be amplified and re-encoded prior to being forwarded to the destination node. Let us first define the amplified symbol vector before re-encoding as

$$\begin{aligned}\tilde{\boldsymbol{s}}_{SR_k}[i] &= \boldsymbol{A}_{R_kD}[i](\boldsymbol{F}_{SR_k}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SR_k}[i]) \\ &= \boldsymbol{A}_{R_kD}[i]\boldsymbol{F}_{SR_k}[i]\boldsymbol{s}[i] + \boldsymbol{A}_{R_kD}[i]\boldsymbol{n}_{SR_k}[i] \\ &= \boldsymbol{F}_{R_k}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{R_k}[i],\end{aligned} \tag{B.1}$$

where $\boldsymbol{A}_{R_kD}[i]$ denotes the $N \times N$ amplify matrix at the kth relay node. The symbol vector $\tilde{\boldsymbol{s}}_{SR_k}[i]$ will be mapped to an $N \times T$ space-time code matrix $\boldsymbol{M}(\tilde{\boldsymbol{s}})$, and multiplied by an adjustable code matrix which is generated randomly before being forwarded to the destination node. By substituting (B.1) into (4.4), the relationship between all the relay

nodes and the destination node can be written as

$$\begin{aligned}\boldsymbol{r}_{RD} &= \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i](\boldsymbol{F}_{R_k}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{R_k}[i]) + \boldsymbol{n}_{RD}[i] \\ &= \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{D}_k[i]\boldsymbol{s}[i] + \boldsymbol{n}_D[i] \\ &= \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_D[i],\end{aligned} \tag{B.2}$$

where the $NT \times N$ matrix $\boldsymbol{D}_k[i]$ contains all the channel information between the source node and the kth relay node, and between the kth relay node and the destination node. The noise vector at the destination node $\boldsymbol{n}_D[i]$ is Gaussian with covariance matrix $\sigma^2(1 + Tr(\sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{D}_k[i]))\boldsymbol{I}_N$. By substituting (B.2) into (4.5), we can rewrite the MSE optimization problem as

$$\begin{aligned}[\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_{k_j}}[i]] = \arg \min_{\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_{k_j}}[i]} E\left[\|s_j[i] - \boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_D[i])\|^2\right], \\ s.t.\ \ \mathrm{Tr}(\sum_{\mathrm{j}=1}^{\mathrm{N}} \boldsymbol{\Phi}_{\mathrm{eq_{kj}}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_{kj}}}^H[\mathrm{i}]) \le \mathrm{P_R}.\end{aligned} \tag{B.3}$$

By taking the instantaneous gradient of $\mathscr{L}$ in (A.1) with respect to $\boldsymbol{w}_j^H[i]$ and $\boldsymbol{\Phi}_{eq_{k_j}}^H[i]$ we can obtain

$$\begin{aligned}\nabla\mathscr{L}_{\boldsymbol{w}_j^*[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right]_{\boldsymbol{w}_j^*[i]} = (s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i])^H\nabla_{\boldsymbol{w}_j^*[i]}(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]) \\ &= -e_j^*[i]\boldsymbol{r}[i], \\ \nabla\mathscr{L}_{\boldsymbol{\Phi}_{eq_{k_j}}^*[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_{RD}[i])\|^2\right]_{\boldsymbol{\Phi}_{eq_{k_j}}^*[i]} \\ &= \nabla_{\boldsymbol{\Phi}_{eq_{k_j}}^*[i]}(s_j[i] - \boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_{RD}[i]))^H(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]) \\ &= -e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}_{k_j}^H[i],\end{aligned} \tag{B.4}$$

where $e_j[i] = s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]$ stands for the jth detected error. By employing step sizes β and μ for the receive filter and the code matrix recursions, respectively, we obtain the

C-ARMO SG algorithm derived as

$$\boldsymbol{w}_j[i+1] = \boldsymbol{w}_j[i] + \beta(e_j^*[i]\boldsymbol{r}[i]),$$
$$\boldsymbol{\Phi}_{eq_{k_j}}[i+1] = \boldsymbol{\Phi}_{eq_{k_j}}[i] + \mu(e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}_{k_j}^H[i]).$$

Appendix C

Derivation of Equation (5.23)

We show how to obtain the optimal adjustable code matrices in the SAS. The ML-based optimal problem is derived in (5.24). We can expand the righthand side of $\mathscr{L}$ in order to obtain the expression of $\boldsymbol{\Phi}_{eq_{kSAS}}[i]$, as described by

$$\begin{aligned}\mathscr{L} &= (\boldsymbol{r}_{SAS}[i] - \sum_{l=1}^{n_r N} \boldsymbol{\Delta}_l \boldsymbol{G}_{eq_{lSAS}}[i]\boldsymbol{\Phi}_{eq_{lSAS}}[i]\hat{\boldsymbol{s}}[i])(\boldsymbol{r}_{SAS}[i] - \sum_{k=1}^{n_r N} \boldsymbol{\Delta}_k \boldsymbol{G}_{eq_{kSAS}}[i]\boldsymbol{\Phi}_{eq_{kSAS}}[i]\hat{\boldsymbol{s}}[i])^H \\ &= \boldsymbol{r}_{SAS}[i]\boldsymbol{r}_{SAS}^H[i] - \boldsymbol{r}_{SAS}[i](\sum_{k=1}^{n_r N} \boldsymbol{\Delta}_k \boldsymbol{G}_{eq_{kSAS}}[i]\boldsymbol{\Phi}_{eq_{kSAS}}[i]\hat{\boldsymbol{s}}[i])^H \\ &\quad - (\sum_{l=1}^{n_r N} \boldsymbol{\Delta}_l \boldsymbol{G}_{eq_{lSAS}}[i]\boldsymbol{\Phi}_{eq_{lSAS}}[i]\hat{\boldsymbol{s}}[i])\boldsymbol{r}_{SAS}^H[i] \\ &\quad + (\sum_{l=1}^{n_r N} \boldsymbol{\Delta}_l \boldsymbol{G}_{eq_{lSAS}}[i]\boldsymbol{\Phi}_{eq_{lSAS}}[i]\hat{\boldsymbol{s}}[i])(\sum_{k=1}^{n_r N} \boldsymbol{\Delta}_k \boldsymbol{G}_{eq_{kSAS}}[i]\boldsymbol{\Phi}_{eq_{kSAS}}[i]\hat{\boldsymbol{s}}[i])^H,\end{aligned} \tag{C.1}$$

and if we focus on the kth adjustable matrix $\boldsymbol{\Phi}_{eq_{kSAS}}[i]$, take the gradient terms with respect to it and equate the result to 0, we can obtain

$$\begin{aligned}
&\frac{d\mathscr{L}}{d\boldsymbol{\Phi}_{eq_{kSAS}}[i]} \\
&= -\boldsymbol{r}_{SAS}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{kSAS}}[i]\boldsymbol{\Delta}^H_k + (\sum_{l=1}^{n_rN}\boldsymbol{\Delta}_l\boldsymbol{G}_{eq_{lSAS}}[i]\boldsymbol{\Phi}_{eq_{lSAS}}[i]\hat{\boldsymbol{s}}[i])\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{kSAS}}[i]\boldsymbol{\Delta}^H_k \\
&= -\boldsymbol{r}_{SAS}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{kSAS}}[i]\boldsymbol{\Delta}^H_k + (\sum_{l=1,l\neq k}^{n_rN}\boldsymbol{\Delta}_l\boldsymbol{G}_{eq_{lSAS}}[i]\boldsymbol{\Phi}_{eq_{lSAS}}[i]\hat{\boldsymbol{s}}[i])\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{kSAS}}[i]\boldsymbol{\Delta}^H_k \\
&\quad + \boldsymbol{\Delta}_k\boldsymbol{G}_{eq_{kSAS}}[i]\boldsymbol{\Phi}_{eq_{kSAS}}[i]\hat{\boldsymbol{s}}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{kSAS}}[i]\boldsymbol{\Delta}^H_k \\
&= 0.
\end{aligned} \tag{C.2}$$

Finally, we move the terms that do not contain $\boldsymbol{\Phi}_{eq_{kSAS}}[i]$ to the other side of the equation and take the pseudo-inverse of the equation and we can obtain the optimal adjustable code matrix as

$$\begin{aligned}
&\boldsymbol{\Phi}_{eq_{kSAS}}[i] \\
&= (\boldsymbol{\Delta}_k\boldsymbol{G}_{eq_{kSAS}}[i])^\dagger \\
&\quad (\boldsymbol{r}_{SAS}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{kSAS}}[i]\boldsymbol{\Delta}^H_k - (\sum_{l=1,l\neq k}^{n_rN}\boldsymbol{\Delta}_l\boldsymbol{G}_{eq_{lSAS}}[i]\boldsymbol{\Phi}_{eq_{lSAS}}[i]\hat{\boldsymbol{s}}[i])\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{kSAS}}[i]\boldsymbol{\Delta}^H_k) \\
&\quad (\hat{\boldsymbol{s}}[i]\hat{\boldsymbol{s}}^H[i]\boldsymbol{G}^H_{eq_{kSAS}}[i]\boldsymbol{\Delta}^H_k)^\dagger.
\end{aligned} \tag{C.3}$$

Glossary

AF	**A**mplify-and-**F**orward
ACMO	**A**djustable **C**ode **M**atrix **O**ptimization
ARMO	**A**daptive **R**obust **M**atrix **O**ptimization
AWGN	**A**dditional **W**hite **G**aussian **N**oise
BER	**B**it-**E**rror-**R**ate
BPSK	**B**inary **P**hase **S**hift **K**eying
BSC	**B**inary **S**ymmetric **C**hannel
C-ARMO	**C**entralized-**A**daptive **R**obust **M**atrix **O**ptimization
CF	**C**ompress-and-**F**orward
CSI	**C**hannel **S**tate **I**nformation
dB	Decibel
D-Alamouti	**D**istributed-Alamouti
DF	**D**ecod-and-**F**orward
DSTC	**D**istributed **S**pace-**T**ime **C**oding
DSTBC	**D**istributed **S**pace-**T**ime **B**lock **C**oding
DT-ACMORO	**D**elay-**T**olerant **A**djustable **C**ode **M**atrices **O**pportunistic **R**elaying **O**ptimization
DT-STBC	**D**elay-**T**olerant **S**pace-**T**ime **B**lock **C**ode
FD-ARMO	**F**ully-**D**istributed-**A**daptive **R**obust **M**atrix **O**ptimization
Fo	**F**ully-**O**pportunistic
JAPA	**J**oint **A**daptive **P**ower **A**llocation
Hz	**H**ert**z**
ISI	**I**nter-**S**ymbol **I**nterference
LCP	**L**inear **C**onstellation **P**recoder
LDC	**L**inear **D**ispersion **C**ode

LMS	**L**east **M**ean-**S**quare
LS	**L**east-**S**quares
MAP	**M**aximum **A** **P**osteriori probability
MAS	**M**ultiple **A**ntenna **S**ystem
MBER	**M**inimum **B**it **E**rror **R**ate
MGF	**M**oment **G**enerating **F**unction
MIMO	**M**ultiple-**I**nput **M**ultiple-**O**utput
MISO	**M**ultiple-**I**nput **S**ingle-**O**utput
ML	**M**aximum **L**ikelihood
MMSE	**M**inimum **M**ean **S**quare **E**rror
MSE	**M**ean **S**quare **E**rror
MSR	**M**aximum **S**um **R**ate
OR	**O**pportunistic **R**elaying
OS	**O**pportunistic **S**ource
OSIC	**O**rdered **SIC**
OSTBC	**O**thogonal **S**pace-**T**ime **B**lock **C**ode
PA	**P**ower **A**llocation
PAM	**P**hase **A**mplitude **M**odulation
PDF	**P**robability **D**ensity **F**unction
PSK	**P**hase-**S**hift **K**eying
QAM	**Q**uadrature **A**mplitude **M**odulation
QOSTBC	**Q**uasi **O**thogonal **S**pace-**T**ime **B**lock **C**ode
R-Alamouti	**R**andomized-Alamouti
RSTC	**R**andomized **S**pace-**T**ime **C**ode
RLS	**R**ecursive **L**east-**S**quares
SAS	**S**ingle **A**ntenna **S**ystem
SER	**S**ymbol **E**rror **R**ate
SD	**S**phere **D**ecoder
SG	**S**tochastic **G**radient
SIC	**S**uccessive **I**nterference **C**ancellation
SIMO	**S**ingle-**I**nput **M**ultiple-**O**utput
SINR	**S**ignal to **I**nterference plus **N**oise **R**atio
SISO	**S**ingle-**I**nput **S**ingle-**O**utput

SM	**S**patial **M**ultiplexing
SNR	**S**ignal to **N**oise **R**atio
STBC	**S**pace-**T**ime **B**lock **C**ode
STC	**S**pace-**T**ime **C**oding
SVD	**S**ingular **V**alue **D**ecomposition
TAST	**T**hreaded **A**lgebraic **S**pace-**T**ime
ZF	**Z**ero **F**orcing

Bibliography

[1] M. Jankiraman, ”Space-Time Codes and MIMO Systems”, Artech House, 2004.

[2] J. N. Laneman and G. W. Wornell, “Cooperative diversity in wireless networks: efficient protocols and outage behavior”, *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[3] J. N. Laneman and G. W. Wornell, “Distributed Space-Time-Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks”, IEEE Trans. Inf. Theory, vol. 49, no. 10, pp. 2415-2425, Oct. 2003.

[4] S. Yiu, R. Schober, L. Lampe, “Distributed Space-Time Block Coding”, IEEE Trans. Wir. Commun., vol. 54, no. 7, pp. 1195 - 1206, Jul. 2006.

[5] R. C. de Lamare, R. Sampaio-Neto, “Blind Adaptive MIMO Receivers for Space-Time Block-Coded DS-CDMA Systems in Multipath Channels Using the Constant Modulus Criterion”, IEEE Trans. on Commun., vol. 58, no. 1, Jan. 2010.

[6] B. Sirkeci-Mergen, A. Scaglione, “Randomized Space-Time Coding for Distributed Cooperative Communication”, IEEE Transactions on Signal Processing, vol. 55, no. 10, Oct. 2007.

[7] T. Peng, R. C. de Lamare and A. Schmeink, “Adaptive Distributed Space-Time Coding Based on Adjustable Code Matrices for Cooperative MIMO Relaying Systems”, *IEEE Transactions on Communications*, vol. 61, no. 7, July 2013.

[8] P. Clarke and R. C. de Lamare, “Joint transmit diversity optimization and relay selection for multi-relay cooperative MIMO systems using discrete stochastic algorithms”, *IEEE Commun. Lett.*, vol. 15, October 2011.

[9] P. Clarke and R. C. de Lamare, "Transmit Diversity and Relay Selection Algorithms for Multirelay Cooperative MIMO Systems", *IEEE Transactions on Vehicular Technology*, vol.61, no. 3, pp. 1084 - 1098, October 2011.

[10] J. Liu, N. B. Shroff and H. D. Sherali, "Optimal power allocation in multi-relay MIMO cooperative networks: theory and algorithms", *IEEE Journal on Selected Areas in Commun.*, vol. 30, pp. 331 - 340, February 2012.

[11] Z. Zhou and B. Vucetic, "A cooperative beamforming scheme in MIMO relay broadcast channels", *IEEE Trans. on Wireless Commun.*, vol. 10, pp. 940 - 947, March 2011.

[12] M. Chen, S. Serbetli and A. Yener, "Distributed power allocation strategies for parallel relay networks", *IEEE Trans. on Wireless Commun.*, vol. 7, no. 2, pp. 552 - 561, Feb. 2008.

[13] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks", *IEEE Trans. on Wireless Commun.*, vol. 5, no. 12, Dec. 2006.

[14] B. Maham, A. Hjϕrungnes, B. S. Rajan, "Quasi-Orthogonal Design and Performance Analysis of Amplify-And-Forward Relay Networks with Multiple-Antennas", 2010 IEEE Wireless Communications and Networking Conference (WCNC), 18 - 21 April 2010.

[15] D. J. Tylavsky and G. R. L. Sohie., "Generalization of the matrix inversion lemma", Proceedings of the IEEE, 74(7):1050 - 1052, July 1986.

[16] T. Peng, R. C. de Lamare and A. Schmeink, "Adaptive Distributed Space-Time Coding for Cooperative MIMO Relaying Systems", 2012 International Symposium on Wireless Communication Systems (ISWCS), pp. 28 - 31 Aug. 2012.

[17] I. Berenguer, X.D. Wang, "Space-Time Coding and Signal Processing for MIMO Communications". J.comput. Sci. Technol, vol. 18, No. 6, pp. 689 - 702.

[18] H. E. Gamal and A.R. Hammons Jr. "On the Design and Performance of Algebraic Space-Time Codes for BPSK and QPSK Modulation". IEEE Transactions on Communications, 50 (8): 907 - 913, June 2002.

[19] A. J. Paulraj, D.A. GORE, R.U. Nabar, H. Bolcskei, “An Overview of MIMO Communications: A Key to Gigabit Wireless”, Proceedings of the IEEE, vol. 92, pp. 198 - 218, Feb. 2004.

[20] L. Z. Zheng, D.N.C. Tse, “Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels”, IEEE Transactions on Information Theory, vol. 49, pp. 1073 - 1096, May 2003.

[21] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, “Space-time block codes from orthogonal designs”, IEEE Trans. on Information Theory, vol. 45, pp. 1456 - 67, July 1999.

[22] S. Alamouti, “A Simple Transmit Diversity Technique for Wireless Communications”, IEEE Journal on Selected Areas in Communications, vol. 16, pp. 1451 - 1458, Oct 1998.

[23] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas”, Bell Labs Technical Journal, vol. 1, no. 2, pp. 41 - 59, Autumn 1996.

[24] R. Heath, Jr. and A. Paulraj, “Switching between multiplexing and diversity based on constellation distance”, in Proc. Allerton Conf. Communication, Control and Computing, Oct. 2000.

[25] A. J. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*, Cambridge University Press, 2003.

[26] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, “V-BLAST: Architecture for realizing very high data rates over the rich-scattering wireless channel”, in Proceedings of URSI International Symposium on Signals, Systems and Electronics, pp. 295 - 300, September 1998.

[27] A. Nosratinia, T. E. Hunter, and A. Hedayat, “Cooperative communications in wireless networks”, IEEE Commun. Magazine, pp. 74 - 80, October 2004.

[28] A. Scaglione, D. L. Goeckel, and J. N. Laneman, “Cooperative communications in mobile ad hoc networks”, IEEE Signal Process.Magazine, pp. 18 - 29, September 2006.

[29] V. Stankovic, A. Host-Madsen, and Z. Xiong, "Cooperative communications in wireless networks", IEEE Signal Process.Magazine, pp. 37 - 49, September 2006.

[30] R. Krishna, Z. Xiong, and S. Lambotharan, "A cooperative MMSE relay strategy for wireless sensor networks", IEEE Signal Process. Letters, vol. 15, pp. 549 - 552, 2008.

[31] S. Berger and A. Wittneben, "Distributed multiuser MMSE relaying in wireless ad-hoc networks", in IEEE Asilomar Conf. on Signals, Systems and Computers, Monterrey, CA, November 2005.

[32] Y. Zou, Y. Yao and B. Zheng, "Opportunistic Distributed Space-Time Coding for Decode-and-Forward Cooperation Systems", IEEE Trans. on Signal Processing, vol. 60, pp. 1766 - 1781, April 2012.

[33] B. Maham, A. Hjørungnes, "Opportunistic Relaying for MIMO Amplify-and-Forward Cooperative Networks", Wireless Personal Communications, DOI 10.1007/s11277-011-0499-9, January 2012.

[34] S. Yang, J.-C. Belfiore, "Optimal Space-Time Codes for the MIMO Amplify-and-Forward Cooperative Channel", IEEE Transactions on Information Theory, vol. 53, pp. 647 - 663, Feb. 2007.

[35] H. Jafarkhani, *Space-Time Coding Theory and Practice*, Cambridge University Press, 2005.

[36] P.W.Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: Architecture for realizing very high data rates over the rich-scattering wireless channel", in Proceedings of URSI International Symposium on Signals, Systems and Electronics, pp. 295 - 300, September 1998.

[37] X.-B. Liang, "A high-rate orthogonal space-time block code", IEEE Communications Letters, vol. 7, pp. 222 - 223, May 2003.

[38] W. Su and X.-G. Xia, "Two generalized complex orthogonal space-time block codes of rates 7/11 and 3/5 for 5 and 6 transmit antennas". IEEE Trans. on Information Theory, 49(1): Jan. 2003, 313 - 316.

[39] W. Su, X.-G. Xia and K. J. R. Liu, “A systematic design of high-rate complex orthogonal space-time block codes”. IEEE Communications Letters, 8: June 2004, 380 - 382.

[40] V. Tarokh, H. Jafarkhani and A. R. Calderbank, “Space-time block codes from orthogonal designs”. IEEE Trans. on Information Theory, 45(5): July 1999, 1456 - 1467.

[41] H. Jafarkhani, “A quasi-orthogonal space-time block code”, IEEE Trans. on Communications, vol. 49, pp. 1 - 4, Jan. 2001.

[42] B. Hassibi and B. Hochwald, “High-rate codes that are linear in space and time”, *IEEE Trans. on Information Theory*, vol. 48, Issue 7, pp. 1804 - 1824, July 2002.

[43] M. O. Damen, A. Tewfik, and J.-C. Belfiore, “A construction of a space-time code based on the theory of numbers”, IEEE Trans. Inf. Theory, vol. 48, no. 3, pp. 753 - 760, Mar. 2002.

[44] J.-C. Belfiore, G. Rekaya and E. Viterbo, “The golden code: a 2x2 full-rate space-time code with nonvanishing determinants”, IEEE Transactions on Information Theory, vol. 51, pp. 1432 - 1436, April 2005.

[45] A. H. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1999.

[46] E. Agrell, T. Eriksson, A. Vardy and K. Zeger, “Closest point search in lattices”, IEEE Trans. on Information Theory, vol. 48, pp. 2201 - 2214, Aug. 2002.

[47] J. Radon, “Lineare scharen orthogonaler matrizen”, Abhandlungen aus dem Mathematischen Seminar der Hamburgishen Universitat, I: 1922, pp. 1 - 14.

[48] A. V. Geramita and J. Seberry, “Orthogonal Designs, Quadratic Forms and Hadamard Matrices”, Lecture Notes in Pure and Applied Mathematics, 1979.

[49] N. Sharma and C.B. Papadias, “Improved quasi-orthogonal codes through constellation rotation”, IEEE Trans. on Communications, vol. 51, pp. 332 - 335, Mar. 2003.

[50] O. Tirkkonen, “Optimizing space-time block codes by constellation rotations”, Finnish Wireless Communications Workshop (FWWC), Oct. 2001.

[51] W. Su and X. Xia, “Signal constellations for quasi-orthogonal space-time block codes with full diversity”, IEEE Trans. on Information Theory, vol. 50, pp. 2331 - 2347, Oct. 2004.

[52] E. Viterbo and J. Boutros, “A universal lattice code decoder for fading channels”, IEEE Trans. Info. Theory, vol. 45, no. 5, pp. 1639 - 1642, July 1999.

[53] M. Damen, H. E. Gamal, and G. Caire, “On maximum-likelihood detection and the search for the closest lattice point”, IEEE Trans. Info. Theory, vol. 49, no. 10, pp. 2389 - 2402, October 2003.

[54] S. Kay, “Fundamentals of Statistical Signal Processing: Estimation Theory”, Prentice Hall, April 5, 1993.

[55] H. V. Trees, *Detection, Estimation and Modulation Theory*, John Wiley and Sons, September 2001.

[56] H. Hsu, *Theory and Problems of Probability*, Random Variables and Random Processes. McGraw-Hill, October 1996.

[57] R.C. de Lamare and R. Sampaio-Neto, “Reduced-Rank Space¨CTime Adaptive Interference Suppression With Joint Iterative Least Squares Algorithms for Spread-Spectrum Systems”, IEEE Transactions on Vehicular Technology, vol. 59, pp. 1217 - 1228, March 2010.

[58] R.C. de Lamare, “Massive MIMO Systems: Signal Processing Challenges and Research Trends”, ARXIV Computer Science - Information Theory, Oct. 2013.

[59] S. Haykin, *Adaptive Filter Theory*, 4th ed. NJ: Prentice Hall, 2002.

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

[61] D. W$\ddot{u}$bben, R. B$\ddot{o}$hnke, V. K$\ddot{u}$hn, and K. Kammeyer, “MMSE extension of V-BLAST based on sorted QR decomposition”, in Proceedings of IEEE Vehicular Technology Conference(VTC), Orlando, USA,, October 2003, pp. 399 - 405.

[62] R. C. de Lamare, R. Sampaio-Neto, and A. Hjrungnes, “Joint iterative interference cancellation and parameter estimation for CDMA systems”, IEEE Communications Letters, vol. 11, no. 12, pp. 916 - 918, December 2007

[63] S. Kim, K. Kim, K. “Log-likelihood ratio based detection ordering in V-BLAST”, IEEE Transactions on Communications, 54(2), 302307. 2006.

[64] M. Chiani, “Introducing erasures in decision-feedback equalization to reduce error propagation”, IEEE Transactions on Communications, vol. 45, no. 7, pp. 757 - 760, July 1997.

[65] M. Reuter, J. C. Allen, J. R. Zeidler, R. C. North, “Mitigating Error Propagation Effects in a Decision Feedback Equalizer”, IEEE Transactions on Communications, vol. 49, no. 11, pp.2028 - 2041, Nov. 2001.

[66] R. C. de Lamare, “Adaptive and Iterative Multi-Branch MMSE Decision Feedback Detection Algorithms for MIMO Systems”, ARXIV Computer Science - Information Theory, Aug. 2013.

[67] P. Li and R. C. de Lamare and R. Fa, “Multiple Feedback Successive Interference Cancellation Detection for Multiuser MIMO Systems”, IEEE Transactions on Wireless Communications, vol. 10, no. 8, pp.2434 - 2439, August 2011.

[68] P. Li, R. C. de Lamare and J. Liu, “Adaptive Decision Feedback Detection with Parallel Interference Cancellation and Constellation Constraints for Multiuser MIMO systems”, IET Communications, vol. 7, no. 6, April 2013.

[69] P. Clarke and R. C. de Lamare, “Joint iterative power allocation and relay selection for cooperative MIMO systems using discrete stochastic algorithms”, 8th International Symposium on Wireless Communication Systems (ISWCS), pp. 432 - 436, Nov. 2011.

[70] O. Seong-Jun, D. Zhang and K. M. Wasserman, “Optimal resource allocation in multiservice CDMA networks”, *IEEE Trans. on Wireless Commun.*, vol. 2, no. 4, pp. 811 - 821, Jul. 2003.

[71] A. Khabbazi and S. Nader-Esfahani, “Power allocation in an amplify-and-forward cooperative network for outage probability minimization”, in *2008 International Symposium on Telecomms.*, 27 - 28 Aug. 2008.

[72] G. Farhadi and N. C. Beaulieu, “A decentralized power allocation scheme for amplify-and-forward multi-hop relaying systems”, in *2010 IEEE International Conference Communications (ICC)*, May 2010.

[73] J. Luo, R. S. Blum, L. J. Cimini, L. J. Greenstein and A. M. Haimovich, "Decode-and-forward cooperative diversity with power allocation in wireless networks", *IEEE Trans. on Wireless Commun.*, pp. 793 - 799, 2007.

[74] T. Peng, R. C. de Lamare and A. Schmeink, "Joint power allocation and receiver design for distributed space-time coded cooperative MIMO systems", in *2011 8th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 427 - 431, 6 - 9 Nov. 2011.

[75] T. Peng, R. C. de Lamare and A. Schmeink, "Joint minimum BER power allocation and receiver design for distributed space-time coded cooperative MIMO systems", *2012 International ITG Workshop on Smart Antennas (WSA)*, pp. 225 - 229, 7 - 8 March 2012.

[76] S. Chen, A. Wolfgang, Y. Shi, and L. Hanzo, "Space-time decision feedback e-qualization using a minimum bit error rate design for single-input multiple-output channels", *IET Commun.*, vol. 1, pp. 671 - 678, August 2007.

[77] A. W. Bowman and A. Azzalini, *Applied Smoothing Techniques for Data Analysis*, Oxford University Press, Oxford, 1997.

[78] U. Niesen, D. Shah and G. W. Wornell, "Adaptive alternating minimization algo-rithms", *IEEE Trans. on Inf. Theory*, vol. 55, pp. 1423 - 1429, March 2009.

[79] K. Zu and R. C. de Lamare, "Lattice reduction-aided regularized block diagonal-ization for multiuser MIMO systems", *2012 IEEE Wireless Communications and Networking Conference(WCNC)*, pp. 131 - 135, 1st - 4th April 2012.

[80] W. Guan, H. Luo and W. Chen, "Linear Relaying Scheme for MIMO Relay System With QoS Requirements", *IEEE Signal Processing Letters*, vol. 15, pp. 697 - 700, December 2008.

[81] O. Munoz-Medina, J. Vidal and A. Agustin, "Linear Transceiver Design in Non-regenerative Relays With Channel State Information", *IEEE Trans. on Signal Pro-cessing*, vol. 55, pp. 2593 - 2604, May 2007.

[82] H. Vikalo, B. Hassibi, and T. Kailath, "Iterative decoding for MIMO channels via modified sphere decoding", *IEEE Trans. Wireless Commun.*, vol. 3, pp. 2299 - 2311, Nov. 2004.

[83] R. C. de Lamare and R. Sampaio-Neto, "Minimum mean-squared error iterative successive parallel arbitrated decision feedback detectors for DS-CDMA systems", *IEEE Trans. on Commun.*, vol. 56, no. 5, pp. 778 - 789, May 2008.

[84] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank processing based on joint and iterative interpolation, decimation, and filtering", *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2503 - 2514, July 2009.

[85] J. H. Choi, H. Y. Yu, Y. H. Lee, "Adaptive MIMO decision feedback equalization for receivers with time-varying channels", *IEEE Trans. Signal Processing*, 2005, 53, no. 11, pp. 4295-4303

[86] C. Windpassinger, L. Lampe, R. F. H. Fischer, T.A Hehn, "A performance study of MIMO detectors", *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, August 2006, pp. 2004-2008.

[87] R. C. de Lamare and R. Sampaio-Neto, "Adaptive Reduced-Rank Equalization Algorithms Based on Alternating Optimization Design Techniques for MIMO Systems", *IEEE Trans. Vehicular Technology*, vol. 60, no. 6, pp.2482-2494, July 2011.

[88] P. Clarke and R. C. de Lamare, "Joint transmit diversity optimization and relay selection for multi-relay cooperative MIMO systems using discrete stochastic algorithms", *IEEE Commun. Lett.*, vol. 15, pp. 1035 - 1037, Oct. 2011.

[89] J. Abouei, H. Bagheri, A. Khandani, "An Efficient Adaptive Distributed Space-Time Coding Scheme for Cooperative Relaying", IEEE Trans. on Wireless Commun., vol. 8, Issue: 10, pp. 4957 - 4962, October 2009.

[90] B. Maham, A. Hjϕrungnes, G. Abreu, "Distributed GABBA Space-Time Codes in Amplify-and-Forward Relay Networks", IEEE Trans. on Wireless Commun., vol.8, pp. 2036 - 2045, April 2009.

[91] M. Kobayashi, G. Caire, N. Jindal, "How Much Training and Feedback are Needed in MIMO Broadcast Channels?", IEEE International Symposium on Information Theory, 2008. ISIT 2008., p.p. 2663 - 2667, 6-11 July 2008.

[92] A. D. Dabbagh, D. J. Love, "Feedback Rate-Capacity Loss Tradeoff for Limited Feedback MIMO Systems", IEEE Transactions on Information Theory, vol. 52, p.p. 2190 - 2202, May 2006.

[93] J. Akhtar, D. Gesbert, "Extending Orthogonal Block Codes with Partial Feedback", IEEE Transactions on Wireless Communications, vol. 3, p.p. 1959 - 1962, Nov. 2004.

[94] I. Choi, J-K. Kim, H. Lee, I. Lee, "Alamouti-Codes Based Four-Antenna Transmission Schemes with Phase Feedback", IEEE Communications Letters, vol. 13, p.p. 749 - 751, Oct. 2009.

[95] G. Jongren and M. Skoglund, "Quantized Feedback Information in Orthogonal Space-Time Block Coding", IEEE Transactions on Information Theory, vol. 50, p.p. 2473 - 2486, Oct. 2004.

[96] D. J. Love, R. W. Heath, Jr., "Limited Feedback Unitary Precoding for Orthogonal Space-Time Block Codes", IEEE Transactions on Signal Processing, vol. 53, p.p. 64 - 73, Jan. 2005.

[97] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank equalization algorithms based on alternating optimization design techniques for MIMO systems", *IEEE Transactions on Vehicular Technology*, vol. 60, no. 6, pp. 2482 - 2494, 2011.

[98] G. Taricco and E. Biglieri, "Exact Pairwise Error Probability of Space-Time Codes", IEEE Transactions on Information Theory, vol. 48, p.p. 510 - 513, Feb 2002.

[99] J. Harshan, B. S. Rajan, "High-Rate, Single-Symbol ML Decodable Precoded D-STBCs for Cooperative Networks", IEEE Transactions on Information Theory, vol. 55, p.p. 2004 - 2015, May 2009.

[100] A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data", IEEE Trans. Acousr., Speech, Signal Processing, vol. ASSP-33, no. 1, p.p. 222 - 230, Feb. 1985.

[101] T. Wang and R. C. de Lamare, "Joint linear receiver design and power allocation using alternating optimization algorithms for wireless sensor networks", *IEEE Trans. on Vehi. Tech.*, vol. 61, pp. 4129 - 4141, 2012.

[102] R.C. de Lamare, R. Sampaio-Neto, "Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels", *IEEE Communications Letters*, vol. 7, no. 2, Feb. 2003, pp. 73 - 75.

[103] S. Chen, A. Wolfgang, Y. Shi, and L. Hanzo, "Space-time decision feedback equalisation using a minimum bit error rate design for single-input multiple-output channels", *IET Commun.*, vol. 1, pp. 671-678, August 2007.

[104] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Press/Springer, 1992.

[105] R. C. de Lamare and A. Alcaim, "Strategies to improve the performance of very low bit rate speech coders and application to a 1.2 kb/s codec", IEEE Proceedings-Vision, image and signal processing, vol. 152, no. 1, February, 2005.

[106] H. El Gamal and M. O. Damen, "Universal space-time coding", *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1097 - 1119, May 2003.

[107] M. O. Damen and A. R. Hammons, "Delay-Tolerant Distributed-TAST Codes for Cooperative Diversity", *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3755 - 3773, Oct. 2007.

[108] M. Torbatian and M. O. Damen, "On the design of delay-tolerant distributed space-time codes with minimum length", *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 931 - 939, Feb. 2009.

[109] Z. Zhimeng, Z. Shihua and A. Nallanathan, "Delay-tolerant distributed linear convolutional space-time code with minimum memory length under frequency-selective channels", *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 3944 - 3949, August 2009.

[110] M. R. Bhatnagar, A. Hjorungnes and M. Debbah, "Delay-tolerant decode-and-forward based cooperative communication over Ricean channels", *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1277 - 1282, April 2010.

[111] D. Gunduz and E. Erkip, "Opportunistic Cooperation by Dynamic Resource Allocation", *IEEE Transactions on Wireless Communications*, vol. 6, pp. 1446 - 1454, April 2007.

[112] T. Peng and R. C. de Lamare, "Adaptive Delay-Tolerant Distributed Space-Time Coding Based on Adjustable Code Matrices for Cooperative MIMO Relaying Systems", 2014 IEEE Wireless Communications and Networking Conference (WCNC), under review.

[113] O. Damen, A. Chkeif, and J.-C. Belfiore, “Lattice code decoder for space-time codes”, IEEE Communications Letters, vol. 4, pp. 161 - 163, May 2000.

[114] M. O. Damen and N. C. Beaulieu, “On two high-rate algebraic space-time codes”, IEEE Trans. Inf. Theory, vol. 49, no. 4, pp. 1059 - 1063, Apr. 2003.

[115] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, “Perfect Space-CTime Block Codes”, IEEE Trans. Inf. Theory, vol. 52, no. 9, September 2006.

[116] T. Wang and R. C. de Lamare, “Joint linear receiver design and power allocation using alternating optimization algorithms for wireless sensor networks”, *IEEE Trans. on Vehi. Tech.*, vol. 61, pp. 4129 - 4141, 2012.

[117] R. C. de Lamare and R. Sampaio-Neto, “Minimum Mean-Squared Error Iterative Successive Parallel Arbitrated Decision Feedback Detectors for DS-CDMA Systems”, IEEE Trans. on Commun., vol. 56, p.p. 778 - 789, May 2008.

[118] J. Yuan, Z. Chen, B. S. Vucetic, and W. Firmanto, “Performance and design of space-time coding in fading channels”, IEEE Transactions on Communications, vol. 51, no. 12, p.p. 1991 - 1996, Dec. 2003.