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

Journal:	*IEEE Transactions on Communications*
Manuscript ID:	TCOM-TPS-12-0788
Manuscript Type:	Transactions Paper Submissions
Date Submitted by the Author:	17-Oct-2012
Complete List of Authors:	PENG, TONG; University of York, Electronics Lamare, Rodrigo; University of York, Electronics Schmeink, Anke; RWTH Aachen University, UMIC Research Centre
Keyword:	Adaptive estimation, Cooperative systems, Diversity methods, Feedback, Codes

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

Tong Peng, Rodrigo C. de Lamare, *Senior Member, IEEE,* and Anke Schmeink, *Member, IEEE*

Abstract

An adaptive distributed space-time coding (DSTC) scheme is proposed for two-hop cooperative MIMO networks. Linear minimum mean square error (MMSE) receive filters and adjustable code matrices subject to a power constraint are considered with an amplify-and-forward (AF) cooperation strategy. In the proposed adaptive DSTC scheme, 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 considered. Linear MMSE expressions are devised to compute the parameters of the adjustable code matrix and the linear receive filters. Stochastic gradient (SG) and least-squares (LS) algorithms are also developed with reduced computational complexity. An upper bound on the pairwise error probability analysis is derived and indicates the advantage of employing the adjustable code matrices at the relay node. An alternative optimization algorithm for the adaptive DSTC scheme is also derived in order to eliminate the need for the feedback. The algorithm provides a fully distributed scheme for the adaptive DSTC at the relay node based on the minimization of

Tong Peng, Rodrigo C. de Lamare is with Communications Research Group, Department of Electronics, University of York, York YO10 5DD, UK, e-mails: tp525@ohm.york.ac.uk; rcdl500@ohm.york.ac.uk

Anke Schmeink, UMIC Research Centre, RWTH Aachen University, D-52056 Aachen, Germany, e-mail: schmeink@umic.rwth-aachen.de

Part of this work has been presented at ISWCS 2012.

the error probability. Simulation results show that the proposed algorithms obtain significant performance gains as compared to existing DSTC schemes.

Index Terms

Adaptive algorithms, space time codes with feedback, cooperative systems, distributed space time codes.

I. Introduction

Cooperative multiple-input and multiple-output (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 [1]. Among the links between the relay nodes and the destination node, cooperation strategies such as Amplify-and-Forward (AF), Decode-and-Forward (DF), Compress-and-Forward (CF) [2] and various distributed space-time coding (DSTC) schemes in [3], [4] and [5] can be employed.

The utilization of a distributed space-time code (DSTC) 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 [6]. An adaptive distributed-Alamouti (D-Alamouti) space-time block code (STBC) design is proposed in [7] for non-regenerative dual-hop wireless systems which achieves the minimum outage probability. DSTC schemes for the AF protocol are discussed in [8]-[9]. In [8], the 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 [9]. In [10], 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 [11], 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 [12] one solution for this trade-off problem is derived. The utilization of limited feedback for STC encoding has been widely discussed in the literature. In [13], the phase information is sent back for STC encoding in order to maintain the full diversity, and the phase feedback is employed in [14] to improve the performance of the Alamouti STBC. A limited feedback link is used in [15] and [16] to provide the channel information for the pre-coding of an orthogonal STBC scheme.

In this paper, we propose an adaptive distributed space-time coding scheme and algorithms for cooperative MIMO relaying systems. This work is first introduced and discussed in [22]. We first develop a centralized algorithm to compute the parameters of an adjustable code matrix with limited feedback. Then adaptive optimization algorithms are derived based on the MSE and 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 utilization 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 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 to choose the optimal adjustable matrix before the transmission, and also achieves the minimum pairwise error probability 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 [10], and it is optimized according to different criteria at the destination node by the proposed algorithm. Second, in order to implement the adaptive algorithm, the adjustable code matrix is optimized with the linear receive filter vector iteratively, and then transmitted back to the relay node via a feedback channel. The impact of the feedback errors are considered and shown in the simulations.

The paper is organized as follows. Section II introduces a two-hop cooperative MIMO system with multiple relays applying the AF strategy and the adaptive DSTC scheme. In Section III the proposed optimization algorithms for the adjustable code matrix are derived, and the pairwise error probability is analyzed in Section IV. The fully distributed optimization algorithm is derived in Section V, and the results of the simulations are given in Section VI. Section VII gives the conclusions of the work.

Notation: the italic, bold lower-case and bold upper-case letters denote scalars, vectors and matrices, respectively. The operators $E[\cdot]$ and $(\cdot)^H$ stand for expected value and the Hermitian operator. The $N \times N$ identity matrix is written as $\boldsymbol{I}_N$. $\parallel \boldsymbol{X} \parallel_F = \sqrt{\mathrm{Tr}(\boldsymbol{X}^H \cdot \boldsymbol{X})} = \sqrt{\mathrm{Tr}(\boldsymbol{X} \cdot \boldsymbol{X}^H)}$ is the Frobenius norm. $\Re[\cdot]$ and $\Im[\cdot]$ stand for the real part and the imaginary part, respectively. $Tr(\cdot)$ stands for the trace of a matrix, and $(\cdot)^\dagger$ for pseudo-inverse.

II. 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. 1. 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. 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

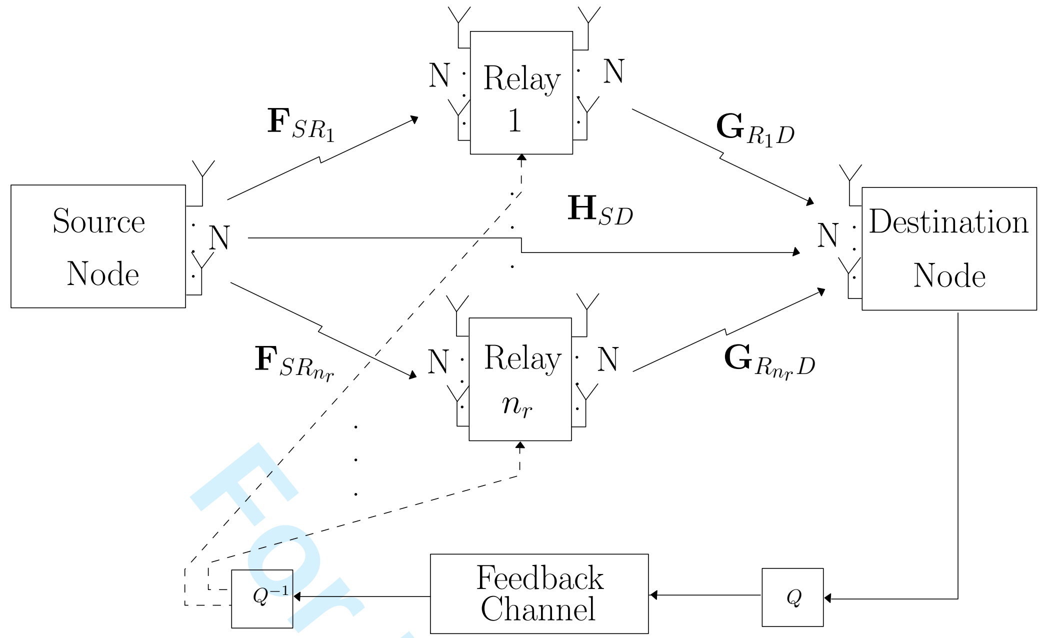

Fig. 1. Cooperative MIMO system model with n_r relay nodes

evaluate the adaptive optimization algorithms, a binary systematic channel (BSC) is considered as the feedback channel.

Consider a cooperative MIMO system with n_r relay nodes that employ the 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 $\boldsymbol{s}[i]$ 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 $E\big[\boldsymbol{s}[i]\boldsymbol{s}^H[i]\big] = \sigma_s^2\boldsymbol{I}$, 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

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

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 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 $N \times N$ channel gain matrices between the source node and the kth relay node, and between the source node and the destination node, respectively.

The received symbols are amplified and re-encoded at each relay node prior to transmission to the destination node in the second hop. The received symbol vector $\boldsymbol{r}_{SR_k}$ is pre-processed before mapped into an STC matrix according to the AF cooperative strategy. We assume that the synchronization at each node is perfect. After processing the received vector $\boldsymbol{r}_{SR_k}[i]$ at the kth relay node, the signal vector $\tilde{\boldsymbol{s}}_{SR_k}[i]$ can be obtained and then forwarded to the destination node. The symbols in $\tilde{\boldsymbol{s}}_{SR_k}[i]$ will be re-encoded by an $N \times T$ DSTC scheme $\boldsymbol{M}(\tilde{\boldsymbol{s}})$, multiplied by an $N \times N$ randomized matrix $\boldsymbol{\Phi}_k[i]$ in [10], and then forwarded to the destination node. The relationship between the kth relay and the destination node can be described as

$$\boldsymbol{R}_{R_kD}[i] = \boldsymbol{G}_{R_kD}[i]\boldsymbol{\Phi}_k[i]\boldsymbol{M}_{R_kD}[i] + \boldsymbol{N}_{R_kD}[i], \tag{2}$$

where the $N \times T$ matrix $\boldsymbol{M}_{R_kD}[i]$ is the DSTC matrix employed at the kth relay nodes whose elements are the information symbols in $\tilde{\boldsymbol{s}}_{SR_k}[i]$. The $N \times T$ received symbol matrix $\boldsymbol{R}_{R_kD}[i]$ in (2) can be written as an $NT \times 1$ vector $\boldsymbol{r}_{R_kD}[i]$ given by

$$\boldsymbol{r}_{R_kD}[i] = \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] + \boldsymbol{n}_{R_kD}[i], \tag{3}$$

where the block diagonal $NT \times NT$ matrix $\boldsymbol{\Phi}_{eq_k}[i]$ denotes the equivalent randomized matrix and the $NT \times N$ matrix $\boldsymbol{G}_{eq_k}[i]$ stands for the equivalent channel matrix which is the DSTC scheme $\boldsymbol{M}(\tilde{s}[i])$ 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 $\boldsymbol{N}_{R_kD}[i]$.

By making use of the randomized matrix $\boldsymbol{\Phi}_{eq_k}[i]$ which achieves the full diversity order and provides a lower error probability has been discussed in [10]. Three types of the randomized matrices are generated and compared in [10]. The uniform phase randomized matrix contains elements generated using $e^{j\theta}$ where θ is uniformly distributed in $[0, 2\pi)$, the Gaussian randomized matrix contains the elements which are zero-mean independent and obey the Gaussian distribution, and the uniform sphere randomized matrix contains the elements which are uniformly distributed on the surface of a complex hyper-sphere of the radius ρ. In our system, the uniform phase randomized matrix is employed because it provides the minimum BER among three randomized matrices shown in [10], and the proposed adaptive algorithms designed in the next section optimize the code matrices employed at the relay nodes in order to achieve a lower BER performance. At each relay node, the traces of the randomized matrices are normalized so that no increase in the energy is introduced at the relay nodes.

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

$$\begin{aligned} \boldsymbol{r}[i] &= \begin{bmatrix} \boldsymbol{H}_{SD}[i]\boldsymbol{s}[i] \\ \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] \end{bmatrix} + \begin{bmatrix} \boldsymbol{n}_{SD}[i] \\ \boldsymbol{n}_{RD}[i] \end{bmatrix} \\ &= \boldsymbol{D}_D[i]\tilde{\boldsymbol{s}}_D[i] + \boldsymbol{n}_D[i], \end{aligned} \tag{4}$$

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 received noise vector at the destination node, which can be modeled as an additive white Gaussian noise (AWGN) with zero mean and covariance matrix $\sigma^2(1+ \parallel \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_{R_kD}[i] \parallel_F^2)\boldsymbol{I}$, where $\boldsymbol{A}_{R_kD}[i]$ stands for the amplification matrix assigned at the kth relay node.

III. 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 [17] for determining the parameters of the adjustable code matrix with reduced complexity are also devised. The DSTC scheme used at the relay node 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 the optimization.

A. Linear MMSE Receiver Design with Adaptive DSTC Optimization

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

$$[\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_k}[i]] = \arg \min_{\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_k}[i]} E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right], \ s.t. \ \mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}[\mathrm{i}]) \leq \mathrm{P_R}, \quad (5)$$

where $\boldsymbol{r}[i]$ denotes the received symbol vector at the destination node which contains the adjustable space-time code matrix with the power constraint P_R. By employing a Lagrange multiplier λ we can obtain the Lagrange expression shown as

$$\mathscr{L} = E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right] + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}[\mathrm{i}]) - \mathrm{P_R}). \quad (6)$$

Since $\boldsymbol{w}_j[i]$ can be optimized by expanding the right-hand side of (6), by taking the gradient with respect to $\boldsymbol{w}_j^*[i]$ and equating the terms to zero, we can obtain the jth MMSE receive filter

$$\boldsymbol{w}_j[i] = \boldsymbol{R}^{-1}\boldsymbol{p}, \quad (7)$$

where the first term $\boldsymbol{R} = E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]$ denotes the auto-correlation matrix and the second one $\boldsymbol{p} = E\left[\boldsymbol{r}[i]s_j^*[i]\right]$ 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

$$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{P}}, \quad (8)$$

where λ stands for the Lagrange multiplier, and $\tilde{\boldsymbol{R}} = E\left[s_j[i]\tilde{s}_{SR_{k_j}}[i]\boldsymbol{w}_j[i]\boldsymbol{w}_j^H[i] + \lambda\boldsymbol{I}\right]$ 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 substituting $\boldsymbol{\Phi}_{eq_{kj}}[i]$ into $\lambda \mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) = \mathrm{P_R}$ and solving the power constraint function. In the proposed adaptive algorithm we employ quantization instead of using the Lagrange multiplier, which requires less computational complexity. The detailed explanation is shown in the next section.

Appendix A includes a detailed derivation of $\boldsymbol{w}_j[i]$ and $\boldsymbol{\Phi}_{eq_j}[i]$. The power constraint can be enforced by employing the Lagrange multiplier and by substituting the power constraint into the MSE cost function. In (8) a closed-form expression of the code matrix $\boldsymbol{\Phi}_{eq_{kj}}[i]$ 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 $\mathrm{O}((N(T+1))^3)$, 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 (7) and (8).

B. 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 the SG algorithm with a linear receiver design is proposed as follows.

The Lagrangian resulting from the optimization problem is derived in (6). The MMSE receive filter can be calculated by (7) which requires a matrix inversion. The Lagrange multiplier λ should be determined before the optimization so the calculation of the value of λ is another problem. In this paper, the power constraint is enforced by a normalization procedure after determining the code matrices instead of employing a Lagrange multiplier, which is a more efficient method to maintain the transmission power at the relay nodes. A simple adaptive algorithm for determining the linear receive filter vectors and the code matrices can be achieved by taking the instantaneous gradient term of (5) with respect to $\boldsymbol{w}_j^*[i]$ and with respect to $\boldsymbol{\Phi}_{eq_{kj}}^*[i]$, respectively, which are

$$\begin{aligned} \nabla \mathscr{L}_{\boldsymbol{w}_j^*[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right]_{\boldsymbol{w}_j^*[i]} \\ &= -e_j^*[i]\boldsymbol{r}[i], \\ \nabla \mathscr{L}_{\boldsymbol{\Phi}_{eq_{kj}}^*[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right]_{\boldsymbol{\Phi}_{eq_{kj}}^*[i]} \\ &= -e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}_{k_j}^H[i], \end{aligned} \tag{9}$$

where $e_j[i]$ stands for the jth detected error, 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 $\boldsymbol{G}_{R_kD}$. After we obtain (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

$$\begin{aligned} \boldsymbol{w}_j[i+1] &= \boldsymbol{w}_j[i] + \beta(e_j^*[i]\boldsymbol{r}[i]), \\ \boldsymbol{\Phi}_{eq_{kj}}[i+1] &= \boldsymbol{\Phi}_{eq_{kj}}[i] + \mu(e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}_{k_j}^H[i]), \end{aligned} \tag{10}$$

where β and μ denote the step sizes for the recursions for the estimation of the linear MMSE receive filter and the code matrices, respectively. A detailed derivation is included in Appendix B.

The energy of the code matrices in (10) will be increased with the processing of the adaptive algorithm, which will contribute to the reduction of the error probability. In order to eliminate the transmission power introduced by the optimization algorithm, a normalization of the code matrix after the optimization is required and implemented as follows

$$\boldsymbol{\Phi}_{eq_{kj}}[i+1] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}_{eq_{kj}}[i+1]}{\sqrt{\sum_{j=1}^{N}\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_{kj}}}[\mathrm{i}+1]\boldsymbol{\Phi}_{\mathrm{eq_{kj}}}^{\mathrm{H}}[\mathrm{i}+1])}}. \tag{11}$$

A summary of the C-ARMO SG algorithm is given in Table I.

According to (10), the receive filter $\boldsymbol{w}_j[i]$ and the code matrix $\boldsymbol{\Phi}_{eq_{kj}}[i]$ depend on each other, so the algorithm in [19] 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 $\boldsymbol{\Phi}_{eq_{kj}}[i]$ is $\mathrm{O}(N(T+1))$ and $\mathrm{O}(N^2T^2)$, respectively, which is

TABLE I

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 $\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}} \boldsymbol{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}) \leq \mathrm{P_R}$.
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:	$\boldsymbol{\Phi}_{eq_{k_j}}[i+1] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}_{eq_{k_j}}[i+1]}{\sqrt{\sum_{j=1}^{N}\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}[\mathrm{i+1}]\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}^{\mathrm{H}}[\mathrm{i+1}])}}$.

much less than $O(N^3(T+1)^3)$ and $O(N^4T^4)$ by using (7) and (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.

C. 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 maximum likelihood (ML) criterion, which is equivalent to a Least-squares (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

$$[\hat{s}_{d_j}[i], \hat{\boldsymbol{\Phi}}_{eq_{k_j}}[i]] = \arg \min_{s_{d_j}[i], \boldsymbol{\Phi}_{eq_{k_j}}[i]} \|\boldsymbol{r}[i] - (\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_{d_j}[i])\|^2, \quad (12)$$
$$s.t.\ \mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}[\mathrm{i}]) \leq \mathrm{P_R},\ \text{for } \mathrm{d} = 1, 2, ..., \mathrm{D},$$

where $\hat{s}_{d_j}[i]$ stands for the desired symbol and $s_{d_j}[i]$ denotes the (j,d)th element in the symbol matrix $\boldsymbol{S}$. By substituting each column of $\boldsymbol{S}$ into (12), we can obtain the most likely transmitted symbol vector $\hat{\boldsymbol{s}}$. It is worth to mention that 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, and the detection and the optimization can be implemented separately as they do not dependent on each other. The optimization algorithm can be considered as a mixed discrete-continues optimization. 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 to the performance.

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

$$\mathscr{L} = \|\boldsymbol{r}[i] - (\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]\hat{s}_{d_j}[i])\|^2 + \lambda(Tr[\boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{\Phi}_{eq_k}^H[i]] - P_R), \tag{13}$$

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

$$\begin{aligned}\nabla\mathscr{L}_{\boldsymbol{\Phi}^*_{eq_{k_j}}[i]} &= (\boldsymbol{r}[i] - \hat{\boldsymbol{r}}[i])\nabla_{\boldsymbol{\Phi}^*_{eq_{k_j}}[i]}(\boldsymbol{r}[i] - \hat{\boldsymbol{r}}[i])^H \\ &= (\boldsymbol{r}[i] - \hat{\boldsymbol{r}}[i])(-\hat{s}^*_{d_j}[i]\boldsymbol{d}_{k_j}[i]) \\ &= (\boldsymbol{r}_{e_j}[i] - \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]\hat{s}_{d_j}[i])(-\hat{s}^*_{d_j}[i]\boldsymbol{d}_{k_j}[i]),\end{aligned} \tag{14}$$

where $\hat{\boldsymbol{r}}[i] = \sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]\hat{s}_{d_j}[i]$ denotes the received symbol vector without the effect of noise, and $\boldsymbol{r}_{e_j}[i] = \boldsymbol{r}[i] - \sum_{k=1}^{n_r}\sum_{l=1,l\neq j}^{N}\boldsymbol{\Phi}_{eq_{k_l}}[i]\boldsymbol{d}_{k_l}[i]\hat{s}_{d_l}[i]$ stands for the received vector without the desired code matrix. The 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.

The optimal code matrix $\hat{\boldsymbol{\Phi}}_{eq_{k_j}}[i]$ requires $\nabla\mathscr{L}_{\boldsymbol{\Phi}^*_{eq_{k_j}}[i]} = 0$, and by substituting $\hat{\boldsymbol{r}}[i]$ into (13) we can obtain the optimal adjustable code matrix as given by

$$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \hat{s}^*_{d_j}[i]\boldsymbol{r}_{e_j}[i]\boldsymbol{d}^H_{k_j}[i](\mid \hat{s}_{d_j}[i] \mid^2 \boldsymbol{d}_{k_j}[i]\boldsymbol{d}^H_{k_j}[i])^{\dagger}. \tag{15}$$

The optimal code matrices will be normalized in order to eliminate the energy introduced during the optimization and then transmitted back to the relay nodes.

D. RLS Code Matrix Estimation Algorithm

The RLS estimation algorithm for the code matrix $\mathbf{\Phi}_{eq_{k_j}}[i]$ is derived in this section. The ML detector is employed so that the detection and the optimization procedures are separate as explained in the last section, so we focus on how to optimize the code matrix rather than the detection. The superior convergence behavior to 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

$$[\hat{\mathbf{\Phi}}_{eq_{k_j}}[n]] = \arg\min_{\mathbf{\Phi}_{eq_{k_j}}[n]} \sum_{i=1}^{n} \lambda^{n-i} \|\boldsymbol{r}[n] - (\sum_{k=1}^{n_r}\sum_{j=1}^{N} \mathbf{\Phi}_{eq_{k_j}}[n]\boldsymbol{d}_{k_j}[n]\hat{s}_{d_j}[n])\|^2, \tag{16}$$

$$s.t. \ \mathrm{Tr}(\mathbf{\Phi}_{\mathrm{eq_k}}[\mathrm{n}]\mathbf{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{n}]) \leq \mathrm{P_R},$$

where λ stands for the forgetting factor. By expanding the right-hand side of (16) and taking gradient with respect to $\mathbf{\Phi}^*_{eq_{k_j}}[i]$ and equaling the terms to zero, we obtain

$$\mathbf{\Phi}_{eq_{k_j}}[n] = (\sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n])(\sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{r}_{k_j}[n]\boldsymbol{r}^H_{k_j}[n])^{-1}, \tag{17}$$

where the $NT \times 1$ vector $\boldsymbol{r}_e[n] = \mathbf{\Phi}_{eq_{k_j}}[n]\boldsymbol{d}_{k_j}[n]\hat{s}_{d_j}[n]$ and $\boldsymbol{r}_{k_j}[n] = \boldsymbol{d}_{k_j}[n]\hat{s}_{d_j}[n]$. The power constraint is still not considered during the optimization. We define

$$\mathbf{\Psi}[n] = \sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{r}_{k_j}[n]\boldsymbol{r}^H_{k_j}[n] = \lambda\mathbf{\Psi}[i-1] + \boldsymbol{r}_{k_j}[n]\boldsymbol{r}^H_{k_j}[n], \tag{18}$$

$$\boldsymbol{Z}[n] = \sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n] = \lambda\boldsymbol{Z}[n] + \boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n], \tag{19}$$

so that we can rewrite (17) as

$$\mathbf{\Phi}_{eq_{k_j}}[n] = \boldsymbol{Z}[n]\mathbf{\Psi}^{-1}[n]. \tag{20}$$

By employing the matrix inversion lemma in [21], we can obtain

$$\mathbf{\Psi}^{-1}[n] = \lambda^{-1}\mathbf{\Psi}^{-1}[n-1] - \lambda^{-1}\boldsymbol{k}[n]\boldsymbol{r}^H_{k_j}[n]\mathbf{\Psi}^{-1}[n-1], \tag{21}$$

TABLE II

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 $\mathrm{trace}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) \leq \mathrm{P_R}$.
4:	For each instant of time, i=1, 2, ..., compute
5:	$\boldsymbol{k}[i] = \frac{\lambda^{-1}\boldsymbol{\Psi}^{-1}[n-1]\boldsymbol{r}_{k_j}[n]}{1+\lambda^{-1}\boldsymbol{r}^H_{k_j}[n]\boldsymbol{\Psi}^{-1}[n-1]\boldsymbol{r}_{k_j}[n]}$,
6:	$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \boldsymbol{\Phi}_{eq_{k_j}}[i-1] + \lambda^{-1}(\boldsymbol{r}_e[n] - \boldsymbol{Z}[n-1]\boldsymbol{k}[n])\boldsymbol{r}^H_{k_j}[n]\boldsymbol{P}[n-1]$,
7:	$\boldsymbol{P}[i] = \lambda^{-1}\boldsymbol{P}[i-1] - \lambda^{-1}\boldsymbol{k}[n]\boldsymbol{r}^H_{k_l}[i]\boldsymbol{P}[i-1]$,
8:	$\boldsymbol{Z}[i] = \lambda\boldsymbol{Z}[i-1] + \boldsymbol{r}_e[i]\boldsymbol{r}^H_{k_j}[i]$.
12:	$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}_{eq_{k_j}}[i]}{\sqrt{\sum_{j=1}^{N}\mathrm{trace}(\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_{k_j}}}[\mathrm{i}])}}$.

where $\boldsymbol{k}[n] = (\lambda^{-1}\boldsymbol{\Psi}^{-1}[n-1]\boldsymbol{r}_{k_j}[n])/(1+\lambda^{-1}\boldsymbol{r}^H_{k_j}[n]\boldsymbol{\Psi}^{-1}[n-1]\boldsymbol{r}_{k_j}[n])$. We define $\boldsymbol{P}[n] = \boldsymbol{\Psi}^{-1}[n]$ and by substituting (19) and (21) into (20), the expression of the code matrix is given by

$$\begin{aligned}\boldsymbol{\Phi}_{eq_{k_j}}[n] &= \lambda\boldsymbol{Z}[n-1]\boldsymbol{P}[n] + \boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n]\boldsymbol{P}[n] \\ &= \boldsymbol{Z}[n-1]\boldsymbol{P}[n-1] + \boldsymbol{Z}[n-1]\boldsymbol{k}[n]\boldsymbol{r}^H_{k_j}[n]\boldsymbol{P}[n-1] + \boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n]\boldsymbol{P}[n] \\ &= \boldsymbol{\Phi}_{eq_{k_j}}[n-1] + \lambda^{-1}(\boldsymbol{r}_e[n] - \boldsymbol{Z}[n-1]\boldsymbol{k}[n])\boldsymbol{r}^H_{k_j}[n]\boldsymbol{P}[n-1].\end{aligned} \tag{22}$$

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

IV. 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 the first section, 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 [18] 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 T codewords, and the codeword $\boldsymbol{C}^1$ is transmitted and decoded as another codeword $\boldsymbol{C}^i$ at the destination node, where $i = 1, 2, ..., T$. According to [18], 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

$$\mathrm{P_e} \leq \sum_{i=2}^{T} \mathrm{P}(\boldsymbol{C}^1 \to \boldsymbol{C}^i). \tag{23}$$

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 pairwise error probability as

$$\begin{aligned} \mathrm{P}(\boldsymbol{C}^1 \to \boldsymbol{C}^2 \mid \boldsymbol{\Phi}) &= \mathrm{P}(\| \boldsymbol{R}^1 - \boldsymbol{G}\boldsymbol{\Phi}\boldsymbol{C}^1 \|_F^2 - \| \boldsymbol{R}^1 - \boldsymbol{G}\boldsymbol{\Phi}\boldsymbol{C}^2 \|_F^2 > 0 \mid \boldsymbol{\Phi}_{eq}) \\ &= \mathrm{P}(\| \boldsymbol{r}^1 - \boldsymbol{\Phi}_{eq}\boldsymbol{G}_{eq}\boldsymbol{F}\boldsymbol{s}^1 \|_F^2 - \| \boldsymbol{r}^1 - \boldsymbol{\Phi}_{eq}\boldsymbol{G}_{eq}\boldsymbol{F}\boldsymbol{s}^2 \|_F^2 > 0 \mid \boldsymbol{\Phi}_{eq}), \end{aligned} \tag{24}$$

where $\boldsymbol{F}$ and $\boldsymbol{G}_{eq}$ stand for the channel coefficient matrix between the source node and the relay node, and between the relay node and the destination node, respectively. The $N \times N$ adjustable code matrix is denoted by $\boldsymbol{\Phi}$ with the equivalent matrix of $\boldsymbol{\Phi}_{eq}$. By defining $\boldsymbol{D} = \boldsymbol{G}_{eq}\boldsymbol{F}$, which stands for the total channel coefficients matrix for all links and expanding the Frobenius norm, we can rewrite the pairwise error probability expression in (24) as

$$\mathrm{P}(\boldsymbol{C}^1 \to \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq}) = \mathrm{P}(\| \boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2) \|_F^2 < \boldsymbol{Y}), \tag{25}$$

where $\boldsymbol{Y} = \mathrm{Tr}(\boldsymbol{n}^{1^H}\boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2) + (\boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2))^H\boldsymbol{n}^1)$, and $\boldsymbol{n}^1$ denotes the noise vector at the destination node with zero mean and covariance matrix $\sigma^2(\| \boldsymbol{\Phi}_{eq}\boldsymbol{G}_{eq} \|_F^2)\boldsymbol{I}$. By making use of the Q function, we can derive the pairwise error probability as

$$\mathrm{P}(\boldsymbol{C}^1 \to \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq}) = \mathrm{Q}\left(\sqrt{\frac{\gamma}{2}} \| \boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2) \|_F\right), \tag{26}$$

where $\mathrm{Q} = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u$, and γ is the received SNR at the destination node assuming the transmit power is equal to 1.

In order to obtain the upper bound of $\mathrm{P}(\boldsymbol{C}^1 \to \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq})$ we expand the formula $\| \boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2) \|_F^2$. Let $\boldsymbol{U}^H\boldsymbol{\Lambda}_s\boldsymbol{U}$ be the eigenvalue decomposition of $(\boldsymbol{s}^1 - \boldsymbol{s}^2)^H(\boldsymbol{s}^1 - \boldsymbol{s}^2)$, where $\boldsymbol{U}$ is a unitary matrix with the eigenvectors and $\boldsymbol{\Lambda}_s$ is a diagonal matrix which contains all the eigenvalues of the difference between two different codewords $\boldsymbol{s}^1$ and $\boldsymbol{s}^2$. Let $\boldsymbol{V}^H\boldsymbol{\Lambda}_{\boldsymbol{\Phi}}\boldsymbol{V}$ stand for the eigenvalue decomposition of $(\mathbf{\Phi}_{eq}\boldsymbol{DU})^H\mathbf{\Phi}_{eq}\boldsymbol{DU}$, where $\boldsymbol{V}$ is a unitary matrix that contains the eigenvectors and $\mathbf{\Lambda}_{\mathbf{\Phi}}$ is a diagonal matrix with the eigenvalues arranged in decreasing order. Therefore, the pairwise probability of error can be written as

$$\mathrm{P}(\boldsymbol{C}^1 \to \boldsymbol{C}^2 \mid \mathbf{\Phi}_{eq}) = \mathrm{Q}\left(\sqrt{\frac{\gamma}{2}\sum_{m=1}^{NT}\sum_{n=1}^{N}\lambda_{\mathbf{\Phi}_m}\lambda_{s_n}|\xi_{n,m}|^2}\right), \tag{27}$$

where $\xi_{n,m}$ is the (n, m)th element in $\boldsymbol{V}$, and $\lambda_{\mathbf{\Phi}_m}$ and λ_{s_n} are the mth and the nth eigenvalues in $\mathbf{\Lambda}_{\mathbf{\Phi}_m}$ and $\mathbf{\Lambda}_s$, respectively. According to [18], an appropriate upper bound assumption of the Q function is given by

$$\mathrm{Q}(x) \leq \frac{1}{2}e^{\frac{-x^2}{2}}. \tag{28}$$

Thus, we can derive the upper bound of the pairwise error probability for an adaptive STC scheme as

$$\mathrm{P}(\boldsymbol{C}^1 \to \boldsymbol{C}^2 \mid \mathbf{\Phi}_{eq}) \leq \frac{1}{2}\exp\left(-\frac{\gamma}{4}\sum_{m=1}^{NT}\sum_{n=1}^{N}\lambda_{\mathbf{\Phi}_m}\lambda_{s_n}|\xi_{n,m}|^2\right), \tag{29}$$

while the upper bound of the error probability expression for a traditional STC in [18] is given by

$$\mathrm{P}(\boldsymbol{C}^1 \to \boldsymbol{C}^2 \mid \boldsymbol{D}_{eq}) \leq \frac{1}{2}\exp\left(-\frac{\gamma}{4}\sum_{m=1}^{NT}\sum_{n=1}^{N}\lambda_{s_n}|\xi_{n,m}|^2\right). \tag{30}$$

By comparison of (29) and (30), it is obvious that the eigenvalue of the adjustable code matrix has to be considered in the expression of PEP, which suggests that employing an adjustable code matrix for an STC scheme at the relay node can provide an improvement in BER performance.

V. The 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 [20] 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 a space-time code has been given by Taricco and Biglieri in [20], 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 have to first find the end-to-end transmission and receive relationship. In Appendix B we obtain the received symbol vector at the destination node, which is written as

$$\boldsymbol{r}_{RD} = \sum_{k=1}^{n_r} \sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i] \boldsymbol{d}_{k_j}[i] s_j[i] + \boldsymbol{n}_{RD}[i],$$

where $\boldsymbol{d}_{k_j}$ denotes the jth column in the equivalent channel matrix that combines the links of the kth relay node connecting the source node and the destination node. If we assume for simplicity that the synchronization is perfect, 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

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq}) = \mathrm{Q}\left(\frac{\parallel \boldsymbol{\Phi}_{eq} \boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2) \parallel_F}{\sqrt{2N_o}} \right), \quad (31)$$

where $N_0 = \mathrm{Tr}(\boldsymbol{I} + \boldsymbol{\Phi}_{eq}\boldsymbol{D})$ denotes the received noise variance at the destination node. The expression in (31) is equivalent to (26) if we assume that the power of the transmitted symbol is equal to 1. We define $\boldsymbol{\Delta} = \boldsymbol{s}^1 - \boldsymbol{s}^2$ as the distance between the code words, and $\tau = \sqrt{\frac{1}{2N_o}} \boldsymbol{\Phi}_{eq} \boldsymbol{D} \boldsymbol{\Delta} \boldsymbol{\Delta}^H \boldsymbol{D}^H \boldsymbol{\Phi}_{eq}^H$ and we assume that the eigenvalue decomposition of $\Delta\Delta^H$ 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. Since the statistical information of the channel matrices $\boldsymbol{F}_{SR}$ and $\boldsymbol{G}_{RD}$ are known at the destination node and have zero mean and variance is equal to 1, their product can be considered as a Gaussian random variable with zero mean and variance equal to $\frac{\sqrt{2}}{2}$. Therfore, we have $E[\boldsymbol{D}\boldsymbol{V}\boldsymbol{V}^H\boldsymbol{D}^H] = \frac{1}{2}\boldsymbol{I}_{NT}$.

The expression of the error probability is given by

$$\begin{aligned}\Theta(c) &= E\left[\exp(-c\xi)\right] = E\left[\exp(-c\sqrt{\frac{1}{2N_o}}[\boldsymbol{\Phi}_{eq}\Delta\Delta^H\boldsymbol{\Phi}_{eq}^H])\right] \\ &= E\left[\exp\left(-c\sqrt{\frac{1}{2N_o}}[\boldsymbol{\Phi}_{eq}\boldsymbol{D}\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H\boldsymbol{D}^H\boldsymbol{\Phi}_{eq}^H]\right)\right] \\ &= \det\left(\boldsymbol{I} + \frac{c}{2\sqrt{2N_0}}\boldsymbol{\Phi}_{eq}\boldsymbol{\Lambda}\boldsymbol{\Phi}_{eq}^H\right)^{-1},\end{aligned} \tag{32}$$

where $c = a + jb$ is the variable defined in the MGF with $a = \frac{1}{4}$ and b is a constant. By inserting (32) into the pairwise error probability expression in [20], we can obtain the exact PEP of the adaptive DSTC scheme written as

$$\mathrm{P}_e = \frac{1}{2J}\sum_{i=1}^{J}\{\Re[\Phi(c)] + \frac{b}{a}\Im[\Phi(c)]\} + \mathrm{E}_J, \tag{33}$$

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

Since the PEP is proportional to (32), it is clear that minimizing the PEP is equal to maximizing the determinant of $\boldsymbol{I} + \frac{c}{2\sqrt{2N_o}}\boldsymbol{\Phi}_{eq}\boldsymbol{\Lambda}\boldsymbol{\Phi}_{eq}^H$. As a result, the optimization problem can be written as

$$\Theta_{opt}(c) = \arg\max_{l}\Theta_l(c), \qquad l = 1, 2, ... \tag{34}$$

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 (32) 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 III.

VI. 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 [18] using QPSK modulation in a quasi-static block fading channel with AWGN. It is also possible to employ the DF protocol or use different number of antennas and relay

TABLE III

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 $\boldsymbol{\Phi}$ 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 $\boldsymbol{\Phi}$ randomly with the power constraint $\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}\boldsymbol{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}) \leq \mathrm{P_R}$
5:	For all $\boldsymbol{\Phi}$, compute
	$\Theta(c) = \det\left(\boldsymbol{I} + \frac{c}{2\sqrt{2N_0}}\boldsymbol{\Phi}_{eq}\boldsymbol{\Lambda}\boldsymbol{\Phi}_{eq}^H\right)^{-1}$
6:	Choose the code matrix according to
	$\Theta_{opt}(c) = \arg\max_l \Theta_l(c)$
7:	Store the optimal code matrix $\boldsymbol{\Phi}_{opt}$ at the relay node

nodes with simple modification. The system is equipped with $n_r = 1$ relay node and $N = 2$ antennas at each node. In the simulations, we set both the symbol power and the noise variance σ^2 as equal to 1, and the power of the adjustable code matrix in the ARMO algorithms are normalized.

The upper bounds of the distributed-Alamouti (D-Alamouti), the randomized Alamouti (R-Alamouti) in [10] and the adaptive Alamouti STC in C-ARMO RLS algorithm are shown in Fig. 2. 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 randomized matrices after each transmission which contributes to a lower error probability upper bound, and the ML detection algorithm provides the optimal performance at the cost of a higher computation complexity.

The proposed C-ARMO SG algorithm with a linear MMSE receiver is compared with the SM scheme and the traditional RSTC algorithm using the D-Alamouti STBC scheme in [5] with

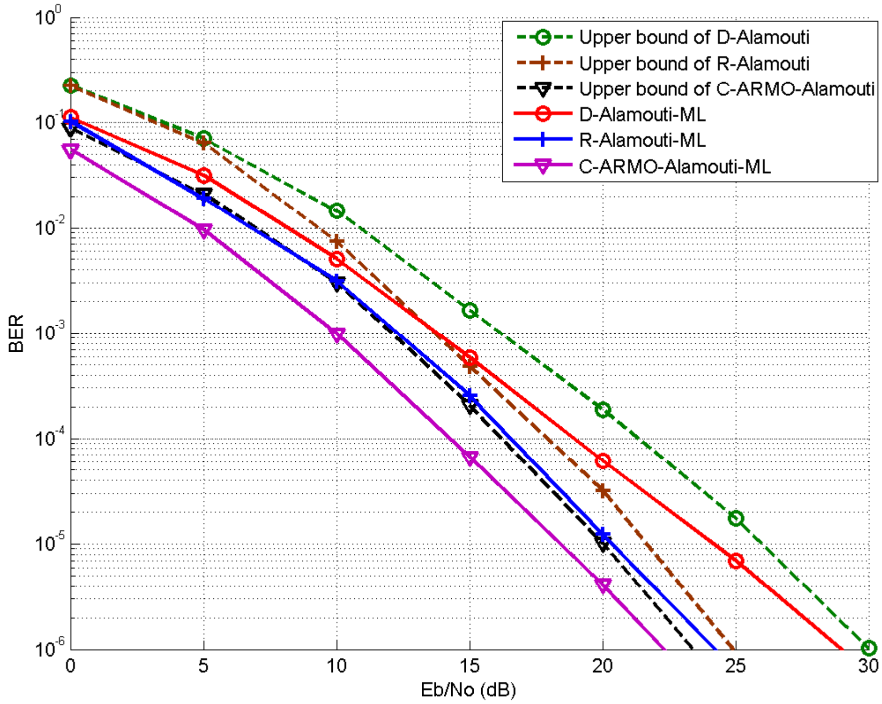

Fig. 2. BER performance vs. E_b/N_0 for the upper bound of the Alamouti schemes without the Direct Link

$n_r = 1$ relay node in Fig. 3. The number of antennas $N = 2$ at each node and the effect of the direct link is considered. The results illustrate that without the direct link, by making use of the STC or the RSTC technique, a significant performance improvement can be achieved compared to the spatial multiplexing system. The RSTC algorithm outperforms the STC-AF system, while the C-ARMO SG algorithm can improve the performance by about 3dB as compared to the RSTC algorithm. 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 and 3dB of gain as compared to employing the traditional STC-AF algorithm.

In Fig. 4, 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. By comparing

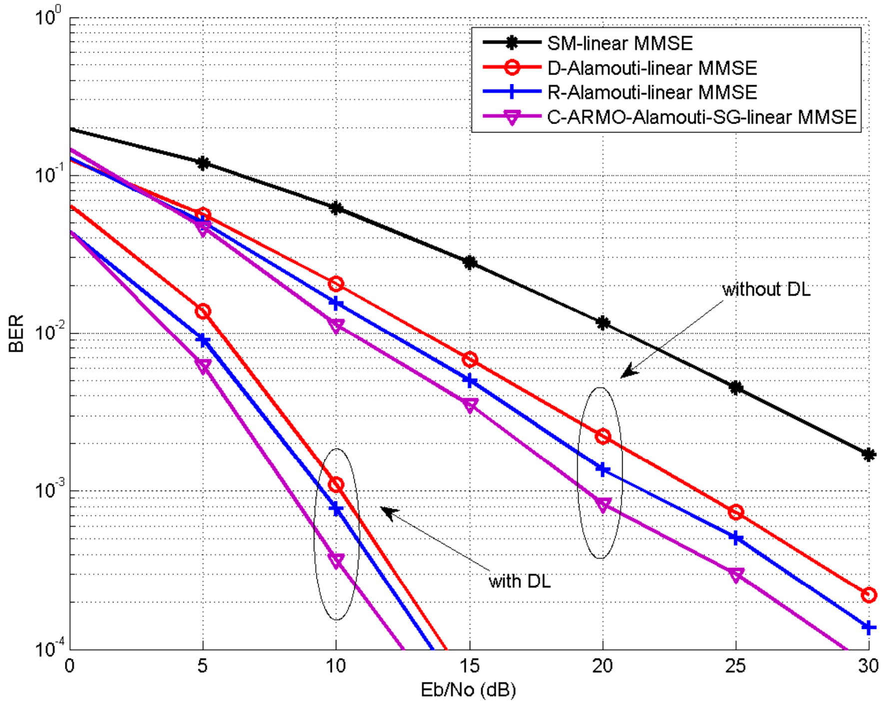

Fig. 3. BER performance vs. E_b/N_0 for C-ARMO SG Algorithm with and without the Direct Link

the curves in Fig. 3 and in Fig. 4, it is noticed that by making use of the ML detector, the performance of the different Alamouti coding schemes achieve the full diversity order and lower error probabilities. In Fig. 4, 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 two 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.

The simulation results shown in Fig. 5 illustrate the convergence property of the C-ARMO SG algorithm. The SM, D-Alamouti and R-Alamouti algorithms obtain nearly flat performance in BER as the utilization of fixed STC scheme and the randomized matrix. The SM scheme

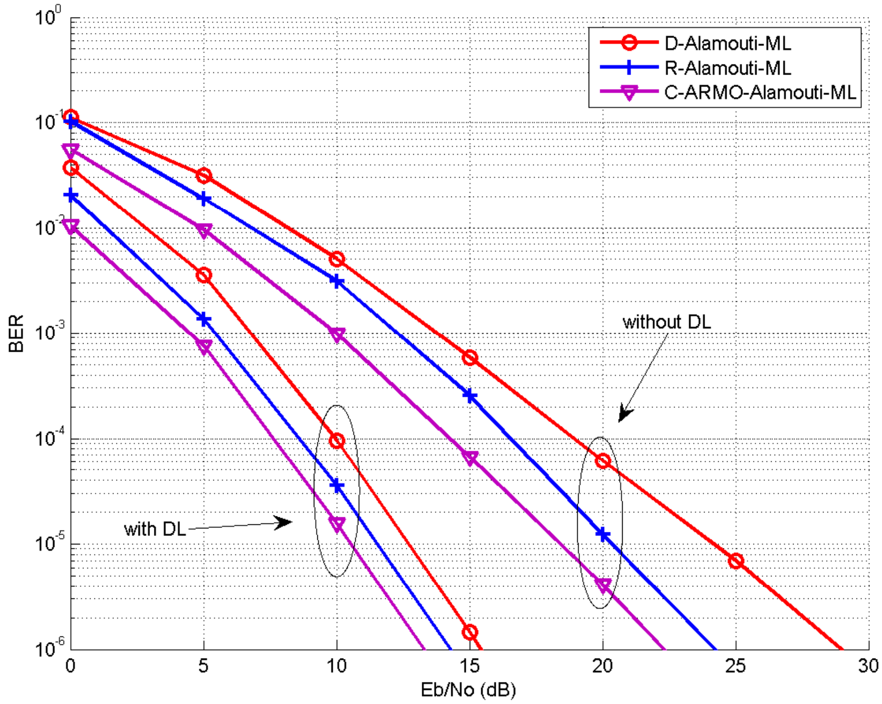

Fig. 4. BER performance vs. E_b/N_0 for C-ARMO RLS Algorithm with and without the Direct Link

has the worst performance due to the lack of coding gains, while the D-Alamouti scheme can provide a significant performance improvement in terms of the BER, and by employing the randomized matrix at the relay node the BER performance can be further improved when the transmission circumstances are the same as that of the D-Alamouti. The C-ARMO SG algorithm shows its advantage by obtaining a fast convergence and a lower BER achievement. At the beginning of the optimization process with a small number of sample vectors, the C-ARMO SG algorithm achieves the BER level of the R-Alamouti scheme because the adjustable code matrix is generated randomly as the same as the R-Alamouti scheme does, but with the increase in the received symbols, the C-ARMO SG algorithm optimizes the adjustable code matrix after each received symbol so that it achieves a better BER performance.

The simulation results shown in Fig. 6 illustrate the influence of the feedback channel on

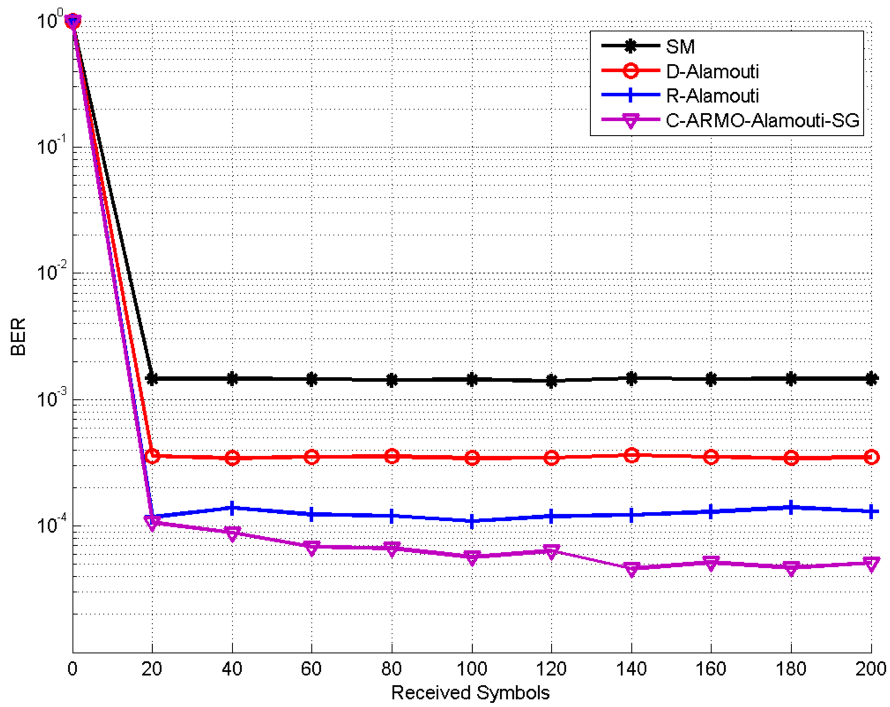

Fig. 5. BER performance vs. Number of Samples for C-ARMO SG Algorithm without the Direct Link

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. 3 and Fig. 4, so the optimized code matrix is perfectly known at the relay node after the C-ARMO SG algorithm; while in Fig. 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 $\boldsymbol{\Phi}_{eq_{kj}}[i]$, and the feedback channel is modeled as a binary symmetric channel with different error probabilities. As we can see from Fig. 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

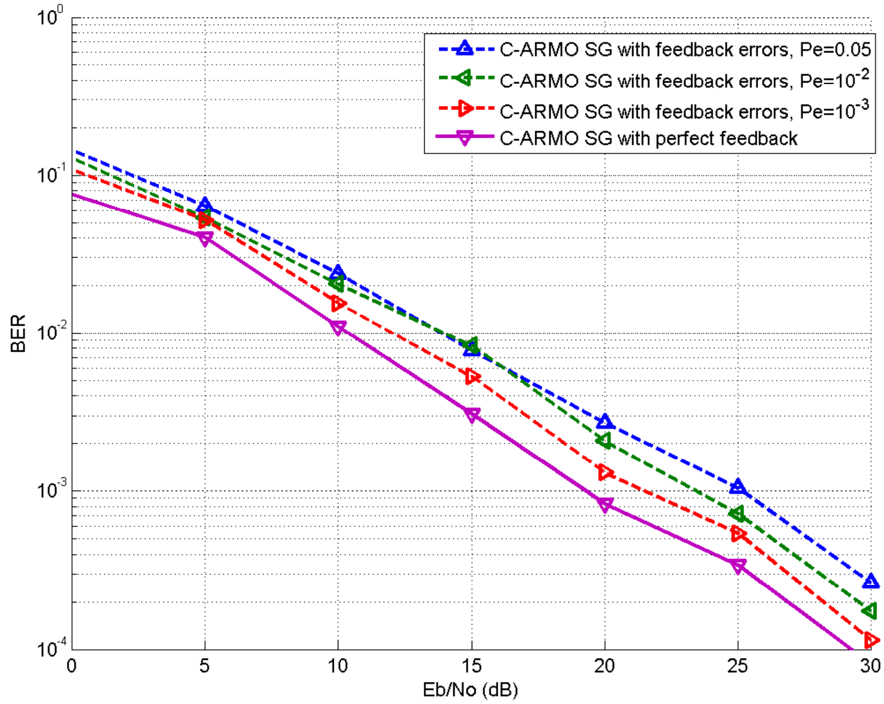

Fig. 6. BER performance vs. number of samples for C-ARMO algorithm with perfect and imperfect feedback links, quantization bits = 4

C-ARMO SG algorithm is about 1dB worse with feedback error probability of 10^{-3}.

In Fig. 7, we plot the average error probability with respect to the SNR for the FD-ARMO algorithm and the C-ARMO SG algorithm. Different adjustable code matrices are used at the relay node. As explained in the previous sections, the main difference between the FD-ARMO and the C-ARMO algorithms is the deployment of the feedback channel. In the theoretical derivation, the FD-ARMO can achieve the average minimum PEP without time for iteration and this is shown in the simulation results. In Fig. 7, the C-ARMO curve and the FD-ARMO curve are in the same shape because they optimize the adjustable code matrices with the same criterion, but 1dB of gain has obtained by the C-ARMO SG algorithm because the exact adjustable code matrix is transmitted back to the relay node in delay-free and error-free feedback channel. While the FD-ARMO, according to the algorithm introduced in the previous section, chooses the optimal

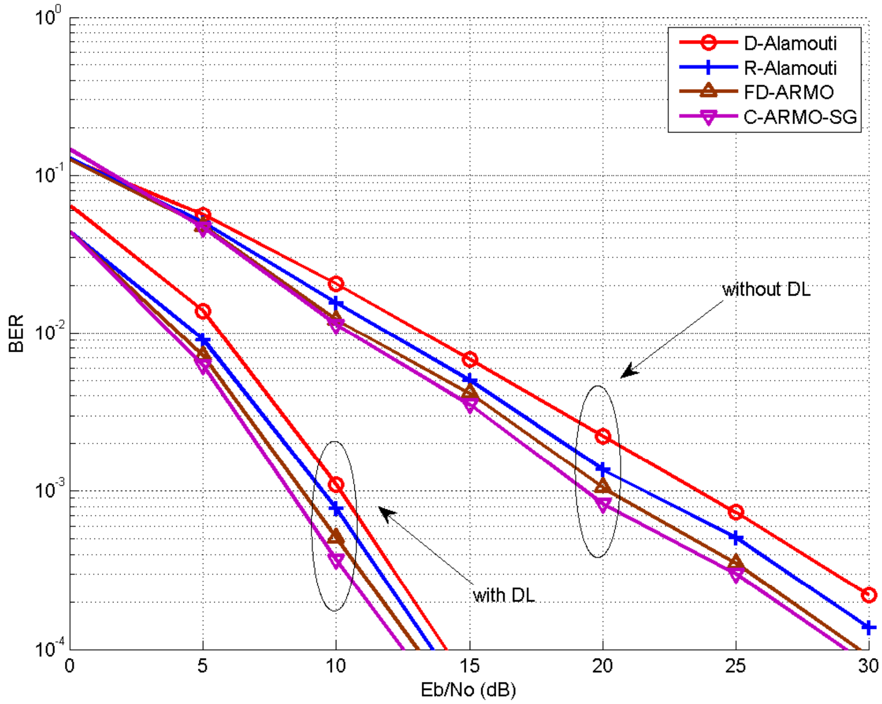

Fig. 7. Full-Distributed ARMO and C-ARMO SG

adjustable code matrix by using the statistical information of the channel before transmission so that the performance will be influenced but the loss of gain is less than 1dB.

VII. Conclusion

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.

APPENDIX A

We show how to obtain the expression of the linear MMSE receive filter $\boldsymbol{w}_j$ and the adjustable code matrix $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ in equation (7) and (8) in Section III in the following.

The MSE optimization optimization problem is given by

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

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

$$\begin{aligned}\mathscr{L} &= E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right] + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}[\mathrm{i}]) - \mathrm{P_R}) \\ &= E\left[(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i])(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i])^H\right] + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}[\mathrm{i}]) - \mathrm{P_R}) \\ &= E\left[s_j[i]s_j^*[i]\right] - \boldsymbol{w}_j^H[i]E\left[\boldsymbol{r}[i]s_j^*[i]\right] - E\left[s_j[i]\boldsymbol{r}^H[i]\right]\boldsymbol{w}_j[i] + \boldsymbol{w}_j^H[i]E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]\boldsymbol{w}_j[i] \\ &\quad + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}[\mathrm{i}]) - \mathrm{P_R}),\end{aligned} \tag{35}$$

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

$$\mathscr{L}'_{\boldsymbol{w}_j^*[i]} = -E\left[\boldsymbol{r}[i]s_j^*[i]\right] + E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]\boldsymbol{w}_j[i] = 0. \tag{36}$$

By moving the first term in (36) 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

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

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

$$\begin{aligned}\boldsymbol{r}[i] &= \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] + \boldsymbol{n}_{RD}[i] \\ &= \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] + \boldsymbol{n}_{RD}[i],\end{aligned} \tag{37}$$

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 (37) into (35), the expression of $\mathscr{L}$ can be written as

$$\begin{aligned}\mathscr{L} =& E\left[s_j[i]s_j^*[i]\right] \\ & - \boldsymbol{w}_j^H[i]E[(\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] + \boldsymbol{n}_{RD}[i])s_j^*[i]] \\ & - E[s_j[i]((\boldsymbol{w}_j^H[i]\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] + \boldsymbol{n}_{RD}[i]))^H] \\ & + E[(\boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] + \boldsymbol{n}_{RD}[i]))^H \\ & \quad \boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] + \boldsymbol{n}_{RD}[i])] \\ & + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}[\mathrm{i}]) - \mathrm{P_R}),\end{aligned}$$

and 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

$$\begin{aligned}\mathscr{L} =& - E\left[s_j[i](\boldsymbol{w}_j^H[i]\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}\tilde{s}_{SR_{k_j}}[i])^H\right] + \lambda(\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{\Phi}_{eq_{k_j}}^H[i] - \mathrm{P_R}\boldsymbol{I}) \\ & + E[(\boldsymbol{w}_j^H[i]\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i])^H\boldsymbol{w}_j^H[i]\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i]],\end{aligned} \tag{38}$$

and by taking the gradient of $\mathscr{L}$ in (38) with respect to $\boldsymbol{\Phi}_{eq_{k_j}}^H[i]$ and equating the terms to zero, we can obtain the adjustable code matrix as

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

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

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

$$\begin{aligned}\tilde{\boldsymbol{s}}_{SR_k}[i] &= \boldsymbol{A}_{R_kD}[i](\boldsymbol{F}_{SR_k}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SR_k}[i]) \\ &= \boldsymbol{A}_{R_kD}[i]\boldsymbol{F}_{SR_k}[i]\boldsymbol{s}[i] + \boldsymbol{A}_{R_kD}[i]\boldsymbol{n}_{SR_k}[i] \\ &= \boldsymbol{F}_{R_k}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{R_k}[i],\end{aligned} \tag{39}$$

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 (39) into (4), the relationship between all the relay nodes and the destination node can be written as

$$\begin{aligned}\boldsymbol{r}_{RD} &= \sum_{k=1}^{n_r}\boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i](\boldsymbol{F}_{R_k}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{R_k}[i]) + \boldsymbol{n}_{RD}[i] \\ &= \sum_{k=1}^{n_r}\boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{D}_k[i]\boldsymbol{s}[i] + \boldsymbol{n}_D[i] \\ &= \sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_D[i],\end{aligned} \tag{40}$$

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 (40) into (5), we can rewrite the MSE optimization problem as

$$\begin{aligned}[\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_{k_j}}[i]] &= \arg\min_{\boldsymbol{w}_j[i],\boldsymbol{\Phi}_{eq_{k_j}}[i]} E\left[\|s_j[i] - \boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_D[i])\|^2\right], \\ & s.t.\ \ \mathrm{Tr}(\sum_{j=1}^{\mathrm{N}}\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_{k_j}}}[\mathrm{i}]) \le \mathrm{P_R}.\end{aligned} \tag{41}$$

By taking the instantaneous gradient of $\mathscr{L}$ in (35) with respect to $\boldsymbol{w}_j^*[i]$ and $\boldsymbol{\Phi}_{eq_{k_j}}^*[i]$ we can obtain

$$\begin{aligned}
\nabla \mathscr{L}_{\boldsymbol{w}_j^*[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right]_{\boldsymbol{w}_j^*[i]} \\
&= (s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i])^H \nabla_{\boldsymbol{w}_j^*[i]}(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]) \\
&= -e_j^*[i]\boldsymbol{r}[i], \\
\nabla \mathscr{L}_{\boldsymbol{\Phi}_{eq_{k_j}}^*[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_{RD}[i])\|^2\right]_{\boldsymbol{\Phi}_{eq_{k_j}}^*[i]} \\
&= \nabla_{\boldsymbol{\Phi}_{eq_{k_j}}^*[i]}(s_j[i] - \boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_{RD}[i])\|^2)^H(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]) \\
&= -e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}_{k_j}^H[i],
\end{aligned} \tag{42}$$

where $e_j[i]$ 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

$$\boldsymbol{w}_j[i+1] = \boldsymbol{w}_j[i] + \beta(e_j^*[i]\boldsymbol{r}[i]),$$

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

References

[1] P. Clarke, R. C. de Lamare, ”Joint Transmit Diversity Optimization and Relay Selection for Multi-relay Cooperative MIMO Systems Using Discrete Stochastic Algorithms”, IEEE Communications Letters, vol. 15, p.p. 1035-1037, Oct. 2011.

[2] J. N. Laneman and G. W. Wornell, ”Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behaviour”, *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] Yulong Zou, Yu-Dong Yao, Baoyu Zheng, ”Opportunistic Distributed Space-Time Coding for Decode-and-Forward Cooperation Systems”, IEEE Trans. on Signal Processing, vol. 60, pp. 1766 - 1781, April 2012.

[7] 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.

[8] B. Maham, A. Hjorungnes, G. Abreu, ”Distributed GABBA Space-Time Codes in Amplify-and-Forward Relay Networks”, IEEE Trans. on Wireless Commun., vol.8, Issuse: 4, pp. 2036 - 2045, April 2009.

[9] Sheng Yang, J.-C. Belfiore, ”Optimal Space-Time Codes for the MIMO Amplify-and-Forward Cooperative Channel”, IEEE Transactions on Information Theory, vol. 53, Issue: 2, pp. 647-663, Feb. 2007.

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

[11] 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.

[12] 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.

[13] J. Akhtar, D. Gesbert, ”Extending Orthogonal Block Codes with Partial Feedback”, IEEE Transactions on Wireless Communications, vol. 3, p.p. 1959-1962, Nov. 2004.

[14] Choi Ilhwan, Kim Jong-Kyu, Lee Heunchul, Lee Inkyu, ”Alamouti-Codes Based Four-Antenna Transmission Schemes with Phase Feedback”, IEEE Communications Letters, vol. 13, p.p. 749-751, Oct. 2009.

[15] G. Jongren, M. Skoglund, ”Quantized Feedback Information in Orthogonal Space-Time Block Coding”, IEEE Transactions on Information Theory, vol. 50, p.p. 2473-2486, Oct. 2004.

[16] 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.

[17] S. Haykin, ”Adaptive Filter Theory”, $4th$ ed. Englewood Cliffs, NJ: Prentice- Hall, 2002.

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

[19] R.C. de Lamare, 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.

[20] G. Taricco, E. Biglieri, ”Exact Pairwise Error Probability of Space-Time Codes”, IEEE Transactions on Information Theory, vol. 48, p.p. 510-513, Feb 2002.

[21] D.J. Tylavsky and G.R.L. Sohie. ”Generalization of the matrix inversion lemma”, Proceedings of the IEEE, 74(7):1050¨C1052, July 1986.

[22] T. Peng, R.C. de Lamare and A. Schmeink, ”Adaptive Distributed Space-Time Coding for Cooperative MIMO Relaying Systems”, accepted by ISWCS 2012.

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

Tong Peng, Rodrigo C. de Lamare, *Senior Member, IEEE,* and Anke Schmeink, *Member, IEEE*

***Abstract*—An adaptive distributed space-time coding (D-STC) scheme is proposed for two-hop cooperative MIMO networks. Linear minimum mean square error (MMSE) receive filters and adjustable code matrices subject to a power constraint are considered with an amplify-and-forward (AF) cooperation strategy. In the proposed adaptive DSTC scheme, 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 considered. Linear MMSE expressions are devised to compute the parameters of the adjustable code matrix and the linear receive filters. Stochastic gradient (SG) and least-squares (LS) algorithms are also developed with reduced computational complexity. An upper bound on the pairwise error probability analysis is derived and indicates the advantage of employing the adjustable code matrices at the relay node. An alternative optimization algorithm for the adaptive DSTC scheme is also derived in order to eliminate the need for the feedback. The algorithm provides a fully distributed scheme for the adaptive DSTC at the relay node based on the minimization of the error probability. Simulation results show that the proposed algorithms obtain significant performance gains as compared to existing DSTC schemes.**

***Index Terms*—Adaptive algorithms, space time codes with feedback, cooperative systems, distributed space time codes.**

I. Introduction

Cooperative multiple-input and multiple-output (MI-MO) 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 [1]. Among the links between the relay nodes and the destination node, cooperation strategies such as Amplify-and-Forward (AF), Decode-and-Forward (DF), Compress-and-Forward (CF) [2] and various distributed space-time coding (DSTC) schemes in [3], [4] and [5] can be employed.

The utilization of a distributed space-time code (D-STC) 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 [6]. An adaptive distributed-Alamouti (D-Alamouti) space-time block code (STBC) design is proposed in [7] for non-regenerative dual-hop wireless systems which achieves the minimum outage probability. DSTC schemes for the AF protocol are discussed in [8]-[9]. In [8], the 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 [9]. In [10], 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 [11], 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 [12] one solution for this trade-off problem is derived. The utilization of limited feedback for STC encoding has been widely discussed in the literature. In [13], the phase information is sent back for STC encoding in order to maintain the full

Tong Peng, Rodrigo C. de Lamare is with Communications Research Group, Department of Electronics, University of York, York YO10 5DD, UK, e-mails: tp525@ohm.york.ac.uk; rcdl500@ohm.york.ac.uk
Anke Schmeink, UMIC Research Centre, RWTH Aachen University, D-52056 Aachen, Germany, e-mail: schmeink@umic.rwth-aachen.de
Part of this work has been presented at ISWCS 2012.

diversity, and the phase feedback is employed in [14] to improve the performance of the Alamouti STBC. A limited feedback link is used in [15] and [16] to provide the channel information for the pre-coding of an orthogonal STBC scheme.

In this paper, we propose an adaptive distributed space-time coding scheme and algorithms for cooperative MIMO relaying systems. This work is first introduced and discussed in [22]. We first develop a centralized algorithm to compute the parameters of an adjustable code matrix with limited feedback. Then adaptive optimization algorithms are derived based on the MSE and 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 utilization 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 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 to choose the optimal adjustable matrix before the transmission, and also achieves the minimum pairwise error probability 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 [10], and it is optimized according to different criteria at the destination node by the proposed algorithm. Second, in order to implement the adaptive algorithm, the adjustable code matrix is optimized with the linear receive filter vector iteratively, and then transmitted back to the relay node via a feedback channel. The impact of the feedback errors are considered and shown in the simulations.

The paper is organized as follows. Section II introduces a two-hop cooperative MIMO system with multiple relays applying the AF strategy and the adaptive

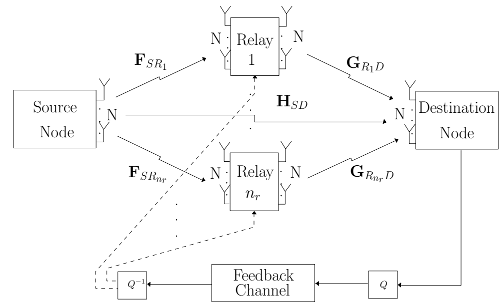

Fig. 1. Cooperative MIMO system model with n_r relay nodes

DSTC scheme. In Section III the proposed optimization algorithms for the adjustable code matrix are derived, and the pairwise error probability is analyzed in Section IV. The fully distributed optimization algorithm is derived in Section V, and the results of the simulations are given in Section VI. Section VII gives the conclusions of the work.

Notation: the italic, bold lower-case and bold upper-case letters denote scalars, vectors and matrices, respectively. The operators $E[\cdot]$ and $(\cdot)^H$ stand for expected value and the Hermitian operator. The $N \times N$ identity matrix is written as $\boldsymbol{I}_N$. $\| \boldsymbol{X} \|_F = \sqrt{\mathrm{Tr}(\boldsymbol{X}^H \cdot \boldsymbol{X})} = \sqrt{\mathrm{Tr}(\boldsymbol{X} \cdot \boldsymbol{X}^H)}$ is the Frobenius norm. $\Re[\cdot]$ and $\Im[\cdot]$ stand for the real part and the imaginary part, respectively. $Tr(\cdot)$ stands for the trace of a matrix, and $(\cdot)^\dagger$ for pseudo-inverse.

II. 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. 1. 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. 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 binary systematic channel (BSC) is considered as the feedback channel.

Consider a cooperative MIMO system with n_r relay nodes that employ the 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 $\boldsymbol{s}[i]$ 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 $E\big[\boldsymbol{s}[i]\boldsymbol{s}^H[i]\big] = \sigma_s^2\boldsymbol{I}$, 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

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

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 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 $N \times N$ channel gain matrices between the source node and the kth relay node, and between the source node and the destination node, respectively.

The received symbols are amplified and re-encoded at each relay node prior to transmission to the destination node in the second hop. The received symbol vector $\boldsymbol{r}_{SR_k}$ is pre-processed before mapped into an STC matrix according to the AF cooperative strategy. We assume that the synchronization at each node is perfect. After processing the received vector $\boldsymbol{r}_{SR_k}[i]$ at the kth relay node, the signal vector $\tilde{\boldsymbol{s}}_{SR_k}[i]$ can be obtained and then forwarded to the destination node. The symbols in $\tilde{\boldsymbol{s}}_{SR_k}[i]$ will be re-encoded by an $N \times T$ DSTC scheme $\boldsymbol{M}(\tilde{\boldsymbol{s}})$, multiplied by an $N \times N$ randomized matrix $\boldsymbol{\Phi}_k[i]$ in [10], and then forwarded to the destination node. The relationship between the kth relay and the destination node can be described as

$$\boldsymbol{R}_{R_kD}[i] = \boldsymbol{G}_{R_kD}[i]\boldsymbol{\Phi}_k[i]\boldsymbol{M}_{R_kD}[i] + \boldsymbol{N}_{R_kD}[i], \tag{2}$$

where the $N \times T$ matrix $\boldsymbol{M}_{R_kD}[i]$ is the DSTC matrix employed at the kth relay nodes whose elements are the information symbols in $\tilde{\boldsymbol{s}}_{SR_k}[i]$. The $N \times T$ received symbol matrix $\boldsymbol{R}_{R_kD}[i]$ in (2) can be written as an $NT \times 1$ vector $\boldsymbol{r}_{R_kD}[i]$ given by

$$\boldsymbol{r}_{R_kD}[i] = \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] + \boldsymbol{n}_{R_kD}[i], \tag{3}$$

where the block diagonal $NT \times NT$ matrix $\boldsymbol{\Phi}_{eq_k}[i]$ denotes the equivalent randomized 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 $\boldsymbol{N}_{R_kD}[i]$.

By making use of the randomized matrix $\boldsymbol{\Phi}_{eq_k}[i]$ which achieves the full diversity order and provides a lower error probability has been discussed in [10]. Three types of the randomized matrices are generated and compared in [10]. The uniform phase randomized matrix contains elements generated using $e^{j\theta}$ where θ is uniformly distributed in $[0, 2\pi)$, the Gaussian randomized matrix contains the elements which are zero-mean independent and obey the Gaussian distribution, and the uniform sphere randomized matrix contains the elements which are uniformly distributed on the surface of a complex hyper-sphere of the radius ρ. In our system, the uniform phase randomized matrix is employed because it provides the minimum BER among three randomized matrices shown in [10], and the proposed adaptive algorithms designed in the next section optimize the code matrices employed at the relay nodes in order to achieve a lower BER performance. At each relay node, the traces of the randomized matrices are normalized so that no increase in the energy is introduced at the relay nodes.

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

$$\begin{aligned} \boldsymbol{r}[i] &= \begin{bmatrix} \boldsymbol{H}_{SD}[i]\boldsymbol{s}[i] \\ \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] \end{bmatrix} \\ &+ \begin{bmatrix} \boldsymbol{n}_{SD}[i] \\ \boldsymbol{n}_{RD}[i] \end{bmatrix} \\ &= \boldsymbol{D}_D[i]\tilde{\boldsymbol{s}}_D[i] + \boldsymbol{n}_D[i], \end{aligned} \tag{4}$$

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 received noise vector at the destination node, which can be modeled as an additive white Gaussian noise (AWGN) with zero mean and covariance matrix $\sigma^2(1+ \parallel \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_{R_kD}[i] \parallel_F^2)\boldsymbol{I}$, where $\boldsymbol{A}_{R_kD}[i]$ stands for the amplification matrix assigned at the kth relay node.

III. 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 [17] for determining the parameters of the adjustable code matrix with reduced complexity are also devised. The DSTC scheme used at the relay node 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 the optimization.

A. Linear MMSE Receiver Design with Adaptive DSTC Optimization

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

$$\begin{aligned} &[\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_k}[i]] = \\ &\arg \min_{\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_k}[i]} E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right], \\ &s.t. \ \mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) \leq \mathrm{P_R}, \end{aligned} \tag{5}$$

where $\boldsymbol{r}[i]$ denotes the received symbol vector at the destination node which contains the adjustable space-time code matrix with the power constraint P_R. By employing a Lagrange multiplier λ we can obtain the Lagrange expression shown as

$$\begin{aligned} \mathscr{L} =& E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right] \\ &+ \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) - \mathrm{P_R}). \end{aligned} \tag{6}$$

Since $\boldsymbol{w}_j[i]$ can be optimized by expanding the right-hand side of (6), by taking the gradient with respect to $\boldsymbol{w}_j^*[i]$ and equating the terms to zero, we can obtain the jth MMSE receive filter

$$\boldsymbol{w}_j[i] = \boldsymbol{R}^{-1}\boldsymbol{p}, \tag{7}$$

where the first term $\boldsymbol{R} = E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]$ denotes the auto-correlation matrix and the second one $\boldsymbol{p} = E\left[\boldsymbol{r}[i]s_j^*[i]\right]$ 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

$$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{P}}, \tag{8}$$

where λ stands for the Lagrange multiplier, and $\tilde{\boldsymbol{R}} = E\left[s_j[i]\tilde{s}_{SR_{k_j}}[i]\boldsymbol{w}_j[i]\boldsymbol{w}_j^H[i] + \lambda\boldsymbol{I}\right]$ and $\tilde{\boldsymbol{P}} = E\left[s_j[i]\tilde{s}_{SR_{k_j}}[i]\boldsymbol{w}_j[i]\boldsymbol{g}^H_{eq_{k_j}}[i]\right]$ are $NT \times NT$ matrices. The value of the Lagrange multiplier λ can be determined by substituting $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ into $\lambda\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) = \mathrm{P_R}$ and solving the power constraint function. In the proposed adaptive algorithm we employ quantization instead of using the Lagrange multiplier, which requires less computational complexity. The detailed explanation is shown in the next section.

Appendix A includes a detailed derivation of $\boldsymbol{w}_j[i]$ and $\boldsymbol{\Phi}_{eq_j}[i]$. The power constraint can be enforced by employing the Lagrange multiplier and by substituting the power constraint into the MSE cost function. In (8) a closed-form expression of the code matrix $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ 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 $\mathrm{O}((N(T+1))^3)$, 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 (7) and (8).

B. 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 the SG algorithm with a linear receiver design is proposed as follows.

The Lagrangian resulting from the optimization problem is derived in (6). The MMSE receive filter can be calculated by (7) which requires a matrix inversion. The Lagrange multiplier λ should be determined before the optimization so the calculation of the value of λ is another problem. In this paper, the power constraint is enforced by a normalization procedure after determining the code matrices instead of employing a Lagrange multiplier, which is a more efficient method to maintain the transmission power at the relay nodes. A simple adaptive algorithm for determining the linear receive filter vectors and the code matrices can be achieved by taking the instantaneous gradient term of (5) with respect to $\boldsymbol{w}_j^*[i]$ and with respect to $\boldsymbol{\Phi}^*_{eq_{k_j}}[i]$, respectively, which are

$$\begin{aligned} \nabla\mathscr{L}_{\boldsymbol{w}_j^*[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right]_{\boldsymbol{w}_j^*[i]} \\ &= -e_j^*[i]\boldsymbol{r}[i], \\ \nabla\mathscr{L}_{\boldsymbol{\Phi}^*_{eq_{k_j}}[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right]_{\boldsymbol{\Phi}^*_{eq_{k_j}}[i]} \\ &= -e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}^H_{k_j}[i], \end{aligned} \tag{9}$$

TABLE I
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 $\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}) \leq \mathrm{P_R}$.
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:	$\nabla \mathscr{L}_{\boldsymbol{\Phi}^*_{eq_{k_j}}[i]} = -e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}^H_{k_j}[i]$,
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:	$\boldsymbol{\Phi}_{eq_{k_j}}[i+1] = \boldsymbol{\Phi}_{eq_{k_j}}[i] + \mu(e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}^H_{k_j}[i])$,
10:	$\boldsymbol{\Phi}_{eq_{k_j}}[i+1] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}_{eq_{k_j}}[i+1]}{\sqrt{\sum_{j=1}^{N}\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}[\mathrm{i+1}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_{k_j}}}[\mathrm{i+1}])}}$.

where $e_j[i]$ stands for the jth detected error, 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 $\boldsymbol{G}_{R_kD}$. After we obtain (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

$$\begin{aligned}\boldsymbol{w}_j[i+1] &= \boldsymbol{w}_j[i] + \beta(e_j^*[i]\boldsymbol{r}[i]),\\ \boldsymbol{\Phi}_{eq_{k_j}}[i+1] &= \boldsymbol{\Phi}_{eq_{k_j}}[i] + \mu(e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}^H_{k_j}[i]),\end{aligned} \quad (10)$$

where β and μ denote the step sizes for the recursions for the estimation of the linear MMSE receive filter and the code matrices, respectively. A detailed derivation is included in Appendix B.

The energy of the code matrices in (10) will be increased with the processing of the adaptive algorithm, which will contribute to the reduction of the error probability. In order to eliminate the transmission power introduced by the optimization algorithm, a normalization of the code matrix after the optimization is required and implemented as follows

$$\boldsymbol{\Phi}_{eq_{k_j}}[i+1] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}_{eq_{k_j}}[i+1]}{\sqrt{\sum_{j=1}^{N}\mathrm{Tr}\boldsymbol{\Omega}}}, \quad (11)$$

where $\boldsymbol{\Omega} = \boldsymbol{\Phi}_{eq_{k_j}}[i+1]\boldsymbol{\Phi}^H_{eq_{k_j}}[i+1]$.

A summary of the C-ARMO SG algorithm is given in Table I.

According to (10), the receive filter $\boldsymbol{w}_j[i]$ and the code matrix $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ depend on each other, so the algorithm in [19] 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 $\mathrm{O}(N(T+1))$ and $\mathrm{O}(N^2T^2)$, respectively, which is much less than $O(N^3(T+1)^3)$ and $O(N^4T^4)$ by using (7) and (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.

C. 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 maximum likelihood (ML) criterion, which is equivalent to a Least-squares (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

$$\begin{aligned}[\hat{s}_{d_j}[i], \hat{\boldsymbol{\Phi}}_{eq_{k_j}}[i]] &= \arg \min_{s_{d_j}[i], \boldsymbol{\Phi}_{eq_{k_j}}[i]}\\ &\|\boldsymbol{r}[i] - (\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_{d_j}[i])\|^2,\end{aligned} \quad (12)$$

$$s.t.\ \mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) \leq \mathrm{P_R}, \mathrm{for\ d} = 1, ..., \mathrm{D},$$

where $\hat{s}_{d_j}[i]$ stands for the desired symbol and $s_{d_j}[i]$ denotes the (j,d)th element in the symbol matrix $\boldsymbol{S}$. By substituting each column of $\boldsymbol{S}$ into (12), we can obtain the most likely transmitted symbol vector $\hat{\boldsymbol{s}}$. It is worth to mention that 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, and the detection and the optimization can be implemented separately as they do not dependent on each other. The optimization algorithm can be considered as a mixed discrete-continues optimization. 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 to the performance.

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

$$\begin{aligned}\mathscr{L} =& \|\boldsymbol{r}[i] - (\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]\hat{s}_{d_j}[i])\|^2\\ &+ \lambda(Tr[\boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{\Phi}^H_{eq_k}[i]] - P_R),\end{aligned} \quad (13)$$

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

$$\begin{aligned}\nabla \mathscr{L}_{\boldsymbol{\Phi}^*_{eq_{k_j}}[i]} &= (\boldsymbol{r}[i] - \hat{\boldsymbol{r}}[i]) \nabla_{\boldsymbol{\Phi}^*_{eq_{k_j}}[i]} (\boldsymbol{r}[i] - \hat{\boldsymbol{r}}[i])^H \\ &= (\boldsymbol{r}[i] - \hat{\boldsymbol{r}}[i])(-\hat{s}^*_{d_j}[i]\boldsymbol{d}_{k_j}[i]) \\ &= (\boldsymbol{r}_{e_j}[i] - \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]\hat{s}_{d_j}[i]) \\ &\quad (-\hat{s}^*_{d_j}[i]\boldsymbol{d}_{k_j}[i]),\end{aligned} \tag{14}$$

where $\hat{\boldsymbol{r}}[i] = \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]\hat{s}_{d_j}[i]$ denotes the received symbol vector without the effect of noise, and $\boldsymbol{r}_{e_j}[i] = \boldsymbol{r}[i] - \sum_{k=1}^{n_r}\sum_{l=1,l\neq j}^{N} \boldsymbol{\Phi}_{eq_{k_l}}[i]\boldsymbol{d}_{k_l}[i]\hat{s}_{d_l}[i]$ stands for the received vector without the desired code matrix. The 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.

The optimal code matrix $\hat{\boldsymbol{\Phi}}_{eq_{k_j}}[i]$ requires $\nabla \mathscr{L}_{\boldsymbol{\Phi}^*_{eq_{k_j}}[i]} = 0$, and by substituting $\hat{\boldsymbol{r}}[i]$ into (13) we can obtain the optimal adjustable code matrix as given by

$$\begin{aligned}\boldsymbol{\Phi}_{eq_{k_j}}[i] =& \hat{s}^*_{d_j}[i]\boldsymbol{r}_{e_j}[i]\boldsymbol{d}^H_{k_j}[i] \\ & (\mid \hat{s}_{d_j}[i] \mid^2 \boldsymbol{d}_{k_j}[i]\boldsymbol{d}^H_{k_j}[i])^{\dagger}.\end{aligned} \tag{15}$$

The optimal code matrices will be normalized in order to eliminate the energy introduced during the optimization and then transmitted back to the relay nodes.

D. 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 last section, so we focus on how to optimize the code matrix rather than the detection. The superior convergence behavior to 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

$$\begin{aligned}&[\hat{\boldsymbol{\Phi}}_{eq_{k_j}}[n]] = \arg \min_{\boldsymbol{\Phi}_{eq_{k_j}}[n]} \\ &\sum_{i=1}^{n} \lambda^{n-i} \|\boldsymbol{r}[n] - (\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[n]\boldsymbol{d}_{k_j}[n]\hat{s}_{d_j}[n])\|^2, \\ &\qquad s.t.\ \mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{n}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{n}]) \leq \mathrm{P_R},\end{aligned} \tag{16}$$

where λ stands for the forgetting factor. By expanding the right-hand side of (16) and taking gradient with respect to $\boldsymbol{\Phi}^*_{eq_{k_j}}[i]$ and equaling the terms to zero, we obtain

$$\begin{aligned}\boldsymbol{\Phi}_{eq_{k_j}}[n] =& (\sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n]) \\ & (\sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{r}_{k_j}[n]\boldsymbol{r}^H_{k_j}[n])^{-1},\end{aligned} \tag{17}$$

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

$$\boldsymbol{\Psi}[n] = \sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{r}_{k_j}[n]\boldsymbol{r}^H_{k_j}[n] = \lambda\boldsymbol{\Psi}[i-1] + \boldsymbol{r}_{k_j}[n]\boldsymbol{r}^H_{k_j}[n], \tag{18}$$

$$\boldsymbol{Z}[n] = \sum_{i=1}^{n} \lambda^{n-i}\boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n] = \lambda\boldsymbol{Z}[n] + \boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n], \tag{19}$$

so that we can rewrite (17) as

$$\boldsymbol{\Phi}_{eq_{k_j}}[n] = \boldsymbol{Z}[n]\boldsymbol{\Psi}^{-1}[n]. \tag{20}$$

By employing the matrix inversion lemma in [21], we can obtain

$$\boldsymbol{\Psi}^{-1}[n] = \lambda^{-1}\boldsymbol{\Psi}^{-1}[n-1] - \lambda^{-1}\boldsymbol{k}[n]\boldsymbol{r}^H_{k_j}[n]\boldsymbol{\Psi}^{-1}[n-1], \tag{21}$$

where $\boldsymbol{k}[n] = (\lambda^{-1}\boldsymbol{\Psi}^{-1}[n-1]\boldsymbol{r}_{k_j}[n])/(1 + \lambda^{-1}\boldsymbol{r}^H_{k_j}[n]\boldsymbol{\Psi}^{-1}[n-1]\boldsymbol{r}_{k_j}[n])$. We define $\boldsymbol{P}[n] = \boldsymbol{\Psi}^{-1}[n]$ and by substituting (19) and (21) into (20), the expression of the code matrix is given by

$$\begin{aligned}\boldsymbol{\Phi}_{eq_{k_j}}[n] &= \lambda\boldsymbol{Z}[n-1]\boldsymbol{P}[n] + \boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n]\boldsymbol{P}[n] \\ &= \boldsymbol{Z}[n-1]\boldsymbol{P}[n-1] + \boldsymbol{r}_e[n]\boldsymbol{r}^H_{k_j}[n]\boldsymbol{P}[n] \\ &\quad + \boldsymbol{Z}[n-1]\boldsymbol{k}[n]\boldsymbol{r}^H_{k_j}[n]\boldsymbol{P}[n-1] \\ &= \boldsymbol{\Phi}_{eq_{k_j}}[n-1] \\ &\quad + \lambda^{-1}(\boldsymbol{r}_e[n] - \boldsymbol{Z}[n-1]\boldsymbol{k}[n])\boldsymbol{r}^H_{k_j}[n]\boldsymbol{P}[n-1].\end{aligned} \tag{22}$$

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

IV. 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 the first section, 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 [18] is introduced for

TABLE II
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 $\mathrm{trace}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) \leq \mathrm{P_R}$.
4:	For each instant of time, i=1, 2, ..., compute
5:	$\boldsymbol{k}[i] = \frac{\lambda^{-1}\boldsymbol{\Psi}^{-1}[n-1]\boldsymbol{r}_{k_j}[n]}{1+\lambda^{-1}\boldsymbol{r}^H_{k_j}[n]\boldsymbol{\Psi}^{-1}[n-1]\boldsymbol{r}_{k_j}[n]}$,
6:	$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \boldsymbol{\Phi}_{eq_{k_j}}[i-1] + \lambda^{-1}(\boldsymbol{r}_e[n] - \boldsymbol{Z}[n-1]\boldsymbol{k}[n])\boldsymbol{r}^H_{k_j}[n]\boldsymbol{P}[n-1]$,
7:	$\boldsymbol{P}[i] = \lambda^{-1}\boldsymbol{P}[i-1] - \lambda^{-1}\boldsymbol{k}[n]\boldsymbol{r}^H_{k_j}[i]\boldsymbol{P}[i-1]$,
8:	$\boldsymbol{Z}[i] = \lambda\boldsymbol{Z}[i-1] + \boldsymbol{r}_e[i]\boldsymbol{r}^H_{k_j}[i]$.
12:	$\boldsymbol{\Phi}_{eq_{k_j}}[i] = \frac{\sqrt{\mathrm{P_R}}\boldsymbol{\Phi}_{eq_{k_j}}[i]}{\sqrt{\sum_{j=1}^{N}\mathrm{trace}(\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_{k_j}}}[\mathrm{i}])}}$.

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 T codewords, and the codeword $\boldsymbol{C}^1$ is transmitted and decoded as another codeword $\boldsymbol{C}^i$ at the destination node, where $i = 1, 2, ..., T$. According to [18], 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

$$\mathrm{P_e} \leq \sum_{i=2}^{T} \mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^i). \quad (23)$$

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 pairwise error probability as

$$\begin{aligned}
&\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{\Phi}) \\
&= \mathrm{P}(\|\ \boldsymbol{R}^1 - \boldsymbol{G\Phi C}^1\ \|_F^2 - \|\ \boldsymbol{R}^1 - \boldsymbol{G\Phi C}^2\ \|_F^2 > 0 \mid \boldsymbol{\Phi}_{eq}) \\
&= \mathrm{P}(\|\ \boldsymbol{r}^1 - \boldsymbol{\Phi}_{eq}\boldsymbol{G}_{eq}\boldsymbol{F}\boldsymbol{s}^1\ \|_F^2 - \|\ \boldsymbol{r}^1 - \boldsymbol{\Phi}_{eq}\boldsymbol{G}_{eq}\boldsymbol{F}\boldsymbol{s}^2\ \|_F^2 > 0 \mid \boldsymbol{\Phi}_{eq}),
\end{aligned} \quad (24)$$

where $\boldsymbol{F}$ and $\boldsymbol{G}_{eq}$ stand for the channel coefficient matrix between the source node and the relay node, and between the relay node and the destination node, respectively. The $N \times N$ adjustable code matrix is denoted by $\boldsymbol{\Phi}$ with the equivalent matrix of $\boldsymbol{\Phi}_{eq}$. By defining $\boldsymbol{D} = \boldsymbol{G}_{eq}\boldsymbol{F}$, which stands for the total channel coefficients matrix for all links and expanding the Frobenius norm, we can rewrite the pairwise error probability expression in (24) as

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq}) = \mathrm{P}(\|\ \boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2)\ \|_F^2 < \boldsymbol{Y}), \quad (25)$$

where $\boldsymbol{Y} = \mathrm{Tr}({\boldsymbol{n}^1}^H\boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2) + (\boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2))^H\boldsymbol{n}^1)$, and $\boldsymbol{n}^1$ denotes the noise vector at the destination node with zero mean and covariance matrix $\sigma^2(\|\ \boldsymbol{\Phi}_{eq}\boldsymbol{G}_{eq}\ \|_F^2)\boldsymbol{I}$. By making use of the Q function, we can derive the pairwise error probability as

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq}) = \mathrm{Q}\left(\sqrt{\frac{\gamma}{2}}\ \|\ \boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2)\ \|_F\right), \quad (26)$$

where $\mathrm{Q} = \frac{1}{\sqrt{2\pi}}\int_x^{\infty}\exp\left(-\frac{u^2}{2}\right)\mathrm{d}u$, and γ is the received SNR at the destination node assuming the transmit power is equal to 1.

In order to obtain the upper bound of $\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq})$ we expand the formula $\|\ \boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2)\ \|_F^2$. Let $\boldsymbol{U}^H\boldsymbol{\Lambda}_s\boldsymbol{U}$ be the eigenvalue decomposition of $(\boldsymbol{s}^1 - \boldsymbol{s}^2)^H(\boldsymbol{s}^1 - \boldsymbol{s}^2)$, where $\boldsymbol{U}$ is a unitary matrix with the eigenvectors and $\boldsymbol{\Lambda}_s$ is a diagonal matrix which contains all the eigenvalues of the difference between two different codewords $\boldsymbol{s}^1$ and $\boldsymbol{s}^2$. Let $\boldsymbol{V}^H\boldsymbol{\Lambda}_{\boldsymbol{\Phi}}\boldsymbol{V}$ stand for the eigenvalue decomposition of $(\boldsymbol{\Phi}_{eq}\boldsymbol{D}\boldsymbol{U})^H\boldsymbol{\Phi}_{eq}\boldsymbol{D}\boldsymbol{U}$, where $\boldsymbol{V}$ is a unitary matrix that contains the eigenvectors and $\boldsymbol{\Lambda}_{\boldsymbol{\Phi}}$ is a diagonal matrix with the eigenvalues arranged in decreasing order. Therefore, the pairwise probability of error can be written as

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq}) = \mathrm{Q}\left(\sqrt{\frac{\gamma}{2}\sum_{m=1}^{NT}\sum_{n=1}^{N}\lambda_{\boldsymbol{\Phi}_m}\lambda_{s_n}|\xi_{n,m}|^2}\right), \quad (27)$$

where $\xi_{n,m}$ is the (n, m)th element in $\boldsymbol{V}$, and $\lambda_{\boldsymbol{\Phi}_m}$ and λ_{s_n} are the mth and the nth eigenvalues in $\boldsymbol{\Lambda}_{\boldsymbol{\Phi}_m}$ and $\boldsymbol{\Lambda}_s$, respectively. According to [18], an appropriate upper bound assumption of the Q function is given by

$$\mathrm{Q}(x) \leq \frac{1}{2}e^{\frac{-x^2}{2}}. \quad (28)$$

Thus, we can derive the upper bound of the pairwise error probability for an adaptive STC scheme as

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq}) \leq \frac{1}{2}\exp\left(-\frac{\gamma}{4}\sum_{m=1}^{NT}\sum_{n=1}^{N}\lambda_{\boldsymbol{\Phi}_m}\lambda_{s_n}|\xi_{n,m}|^2\right), \quad (29)$$

while the upper bound of the error probability expression for a traditional STC in [18] is given by

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{D}_{eq}) \leq \frac{1}{2}\exp\left(-\frac{\gamma}{4}\sum_{m=1}^{NT}\sum_{n=1}^{N}\lambda_{s_n}|\xi_{n,m}|^2\right). \quad (30)$$

By comparison of (29) and (30), it is obvious that the eigenvalue of the adjustable code matrix has to be considered in the expression of PEP, which suggests that employing an adjustable code matrix for an STC scheme at the relay node can provide an improvement in BER performance.

V. The 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 [20] 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 a space-time code has been given by Taricco and Biglieri in [20], 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 have to first find the end-to-end transmission and receive relationship. In Appendix B we obtain the received symbol vector at the destination node, which is written as

$$\boldsymbol{r}_{RD} = \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_{RD}[i],$$

where $\boldsymbol{d}_{k_j}$ denotes the jth column in the equivalent channel matrix that combines the links of the kth relay node connecting the source node and the destination node. If we assume for simplicity that the synchronization is perfect, 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

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{\Phi}_{eq}) = \mathrm{Q}\left(\frac{\parallel \boldsymbol{\Phi}_{eq}\boldsymbol{D}(\boldsymbol{s}^1 - \boldsymbol{s}^2) \parallel_F}{\sqrt{2N_o}}\right), \tag{31}$$

where $N_0 = \mathrm{Tr}(\boldsymbol{I} + \boldsymbol{\Phi}_{eq}\boldsymbol{D})$ denotes the received noise variance at the destination node. The expression in (31) is equivalent to (26) if we assume that the power of the transmitted symbol is equal to 1. We define $\boldsymbol{\Delta} = \boldsymbol{s}^1 - \boldsymbol{s}^2$ as the distance between the code words, and $\tau = \sqrt{\frac{1}{2N_o}\boldsymbol{\Phi}_{eq}\boldsymbol{D}\boldsymbol{\Delta}\boldsymbol{\Delta}^H\boldsymbol{D}^H\boldsymbol{\Phi}_{eq}^H}$ and we assume that the eigenvalue decomposition of $\Delta\Delta^H$ 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. Since the statistical information of the channel matrices $\boldsymbol{F}_{SR}$ and $\boldsymbol{G}_{RD}$ are known at the destination node and have zero mean and variance is equal to 1, their product can be considered as a Gaussian random variable with zero mean and variance equal to $\frac{\sqrt{2}}{2}$. Therfore, we have $E[\boldsymbol{D}\boldsymbol{V}\boldsymbol{V}^H\boldsymbol{D}^H] = \frac{1}{2}\boldsymbol{I}_{NT}$. The expression of the error probability is given by

$$\begin{aligned}\Theta(c) &= E\left[\exp(-c\xi)\right] = E\left[\exp(-c\sqrt{\frac{1}{2N_o}}[\boldsymbol{\Phi}_{eq}\Delta\Delta^H\boldsymbol{\Phi}_{eq}^H])\right] \\ &= E\left[\exp\left(-c\sqrt{\frac{1}{2N_o}}[\boldsymbol{\Phi}_{eq}\boldsymbol{D}\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^H\boldsymbol{D}^H\boldsymbol{\Phi}_{eq}^H]\right)\right] \\ &= \det\left(\boldsymbol{I} + \frac{c}{2\sqrt{2N_0}}\boldsymbol{\Phi}_{eq}\boldsymbol{\Lambda}\boldsymbol{\Phi}_{eq}^H\right)^{-1},\end{aligned} \tag{32}$$

where $c = a + jb$ is the variable defined in the MGF with $a = \frac{1}{4}$ and b is a constant. By inserting (32) into the pairwise error probability expression in [20], we can obtain the exact PEP of the adaptive DSTC scheme written as

$$\mathrm{P}_e = \frac{1}{2J}\sum_{i=1}^{J}\{\Re[\Phi(c)] + \frac{b}{a}\Im[\Phi(c)]\} + \mathrm{E}_J, \tag{33}$$

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

Since the PEP is proportional to (32), it is clear that minimizing the PEP is equal to maximizing the determinant of $\boldsymbol{I} + \frac{c}{2\sqrt{2N_o}}\boldsymbol{\Phi}_{eq}\boldsymbol{\Lambda}\boldsymbol{\Phi}_{eq}^H$. As a result, the optimization problem can be written as

$$\Theta_{opt}(c) = \arg\max_{l} \Theta_l(c), \qquad l = 1, 2, ... \tag{34}$$

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 (32) 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 III.

VI. 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 [18] using

TABLE III
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 $\mathrm{Tr}(\mathbf{\Phi}_{\mathrm{eq_k}}\mathbf{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}) \leq \mathrm{P_R}$
5:	For all $\mathbf{\Phi}$, compute $\Theta(c) = \det\left(\boldsymbol{I} + \frac{c}{2\sqrt{2N_0}}\mathbf{\Phi}_{eq}\mathbf{\Lambda}\mathbf{\Phi}_{eq}^H\right)^{-1}$
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

QPSK modulation in a quasi-static block fading channel with AWGN. It is also possible to employ the DF protocol or use different number of antennas and relay nodes with simple modification. The system is equipped with $n_r = 1$ relay node and $N = 2$ antennas at each node. In the simulations, we set both the symbol power and the noise variance σ^2 as equal to 1, and the power of the adjustable code matrix in the ARMO algorithms are normalized.

The upper bounds of the distributed-Alamouti (D-Alamouti), the randomized Alamouti (R-Alamouti) in [10] and the adaptive Alamouti STC in C-ARMO RLS algorithm are shown in Fig. 2. 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 randomized matrices after each transmission which contributes to a lower error probability upper bound, and the ML detection algorithm provides the optimal performance at the cost of a higher computation complexity.

The proposed C-ARMO SG algorithm with a linear MMSE receiver is compared with the SM scheme and the traditional RSTC algorithm using the D-Alamouti STBC scheme in [5] with $n_r = 1$ relay node in Fig. 3. The number of antennas $N = 2$ at each node and the effect of the direct link is considered. The results illustrate that without the direct link, by making use of the STC or the RSTC technique, a significant performance improvement can be achieved compared to the spatial

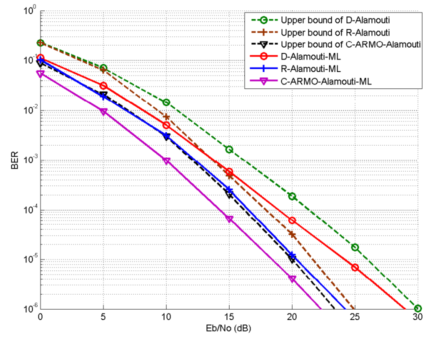

Fig. 2. BER performance vs. E_b/N_0 for the upper bound of the Alamouti schemes without the Direct Link

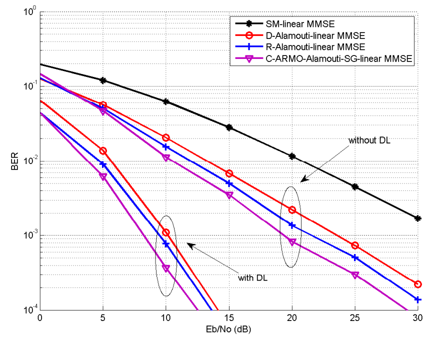

Fig. 3. BER performance vs. E_b/N_0 for C-ARMO SG Algorithm with and without the Direct Link

multiplexing system. The RSTC algorithm outperforms the STC-AF system, while the C-ARMO SG algorithm can improve the performance by about 3dB as compared to the RSTC algorithm. 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 and 3dB of gain as compared to employing the traditional STC-AF algorithm.

In Fig. 4, 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. By comparing the curves in Fig. 3 and in Fig. 4, it is noticed that by making use of the ML detector, the performance of the different Alamouti coding schemes achieve the full diversity order and lower error probabilities. In Fig. 4, the R-Alamouti scheme improves the performance by about 4dB without the direct link

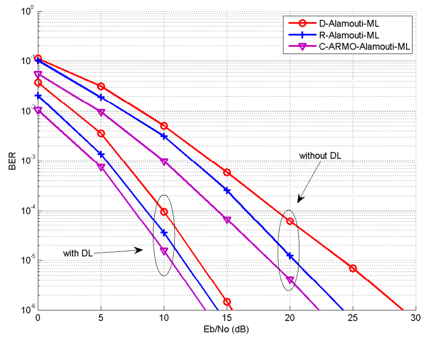

Fig. 4. BER performance vs. E_b/N_0 for C-ARMO RLS Algorithm with and without the Direct Link

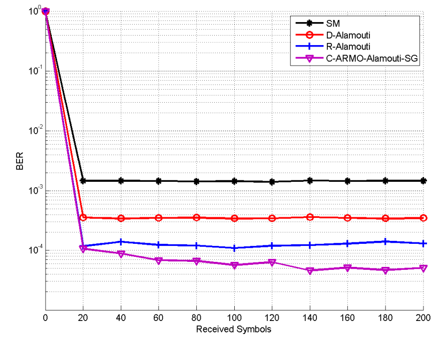

Fig. 5. BER performance vs. Number of Samples for C-ARMO SG Algorithm without the Direct Link

compared to the D-Alamouti scheme, and the C-ARMO RLS algorithm provides a significant improvement in terms of gains compared to the other two 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.

The simulation results shown in Fig. 5 illustrate the convergence property of the C-ARMO SG algorithm. The SM, D-Alamouti and R-Alamouti algorithms obtain nearly flat performance in BER as the utilization of fixed STC scheme and the randomized matrix. The SM scheme has the worst performance due to the lack of coding gains, while the D-Alamouti scheme can provide a significant performance improvement in terms of the BER, and by employing the randomized matrix at the relay node the BER performance can be further improved when the transmission circumstances are the same as that of the D-Alamouti. The C-ARMO SG algorithm shows its advantage by obtaining a fast convergence and a lower BER achievement. At the beginning of the optimization process with a small number of sample vectors, the C-ARMO SG algorithm achieves the BER level of the R-Alamouti scheme because the adjustable code matrix is generated randomly as the same as the R-Alamouti scheme does, but with the increase in the received symbols, the C-ARMO SG algorithm optimizes the adjustable code matrix after each received symbol so that it achieves a better BER performance.

The simulation results shown in Fig. 6 illustrate the influence of the feedback channel 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. 3 and Fig. 4, so the optimized code matrix is perfectly known at the relay node after the C-ARMO SG algorithm; while in Fig. 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. 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}.

In Fig. 7, we plot the average error probability with respect to the SNR for the FD-ARMO algorithm and the C-ARMO SG algorithm. Different adjustable code matrices are used at the relay node. As explained in the previous sections, the main difference between the FD-ARMO and the C-ARMO algorithms is the deployment of the feedback channel. In the theoretical derivation, the FD-ARMO can achieve the average minimum PEP without time for iteration and this is shown in the simulation results. In Fig. 7, the C-ARMO curve and the FD-ARMO curve are in the same shape because they optimize the adjustable code matrices with the same criterion, but 1dB of gain has obtained by the C-ARMO SG algorithm because the exact adjustable code matrix is transmitted back to the relay node in delay-free and error-free feedback channel. While the FD-ARMO,

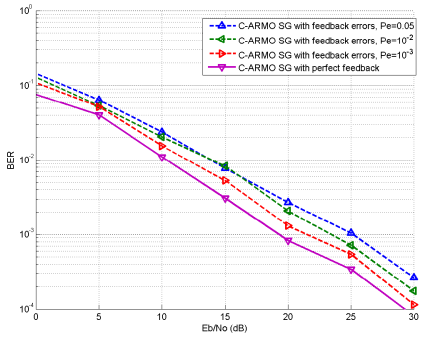

Fig. 6. BER performance vs. number of samples for C-ARMO algorithm with perfect and imperfect feedback links, quantization bits = 4

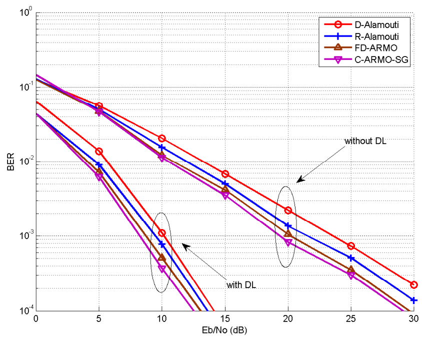

Fig. 7. Full-Distributed ARMO and C-ARMO SG

according to the algorithm introduced in the previous section, chooses the optimal adjustable code matrix by using the statistical information of the channel before transmission so that the performance will be influenced but the loss of gain is less than 1dB.

VII. Conclusion

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.

Appendix A

We show how to obtain the expression of the linear MMSE receive filter $\boldsymbol{w}_j$ and the adjustable code matrix $\boldsymbol{\Phi}_{eq_{k_j}}[i]$ in equation (7) and (8) in Section III in the following.

The MSE optimization optimization problem is given by

$$[\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_k}[i]] = \arg \min_{\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_k}[i]} E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right],$$

$$s.t. \ \mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) \leq \mathrm{P_R}.$$

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

$$\begin{aligned}\mathscr{L} &= E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right] + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) - \mathrm{P_R}) \\ &= E\left[(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i])(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i])^H\right] \\ &\quad + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) - \mathrm{P_R}) \\ &= E\left[s_j[i]s_j^*[i]\right] - \boldsymbol{w}_j^H[i]E\left[\boldsymbol{r}[i]s_j^*[i]\right] \\ &\quad - E\left[s_j[i]\boldsymbol{r}^H[i]\right]\boldsymbol{w}_j[i] + \boldsymbol{w}_j^H[i]E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]\boldsymbol{w}_j[i] \\ &\quad + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}^{\mathrm{H}}_{\mathrm{eq_k}}[\mathrm{i}]) - \mathrm{P_R}),\end{aligned} \tag{35}$$

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

$$\mathscr{L}'_{\boldsymbol{w}_j^*[i]} = -E\left[\boldsymbol{r}[i]s_j^*[i]\right] + E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]\boldsymbol{w}_j[i] = 0. \tag{36}$$

By moving the first term in (36) 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

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

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

$$\begin{aligned}\boldsymbol{r}[i] &= \sum_{k=1}^{n_r} \boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] + \boldsymbol{n}_{RD}[i] \\ &= \sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] + \boldsymbol{n}_{RD}[i],\end{aligned} \tag{37}$$

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 (37) into (35), the expression of $\mathscr{L}$ can be written as

$$\begin{aligned}\mathscr{L} =& E\left[s_j[i]s_j^*[i]\right] - \boldsymbol{w}_j^H[i] \\ & E[(\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] + \boldsymbol{n}_{RD}[i])s_j^*[i]] \\ & - E[s_j[i]((\boldsymbol{w}_j^H[i]\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] \\ & + \boldsymbol{n}_{RD}[i]))^H] \\ & + E[(\boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] \\ & + \boldsymbol{n}_{RD}[i]))^H\boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i] \\ & + \boldsymbol{n}_{RD}[i])] + \lambda(\mathrm{Tr}(\boldsymbol{\Phi}_{\mathrm{eq_k}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_k}}^{\mathrm{H}}[\mathrm{i}]) - \mathrm{P_R}),\end{aligned}$$

and 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

$$\begin{aligned}\mathscr{L} =& - E\left[s_j[i](\boldsymbol{w}_j^H[i]\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}\tilde{s}_{SR_{k_j}}[i])^H\right] \\ & + \lambda(\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{\Phi}_{eq_{k_j}}^H[i] - \mathrm{P_R}\boldsymbol{I}) \\ & + E[(\boldsymbol{w}_j^H[i]\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i])^H \\ & \quad \boldsymbol{w}_j^H[i]\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{g}_{eq_{k_j}}[i]\tilde{s}_{SR_{k_j}}[i]],\end{aligned} \quad (38)$$

and by taking the gradient of $\mathscr{L}$ in (38) with respect to $\boldsymbol{\Phi}_{eq_{k_j}}^H[i]$ and equating the terms to zero, we can obtain the adjustable code matrix as

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

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

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

$$\begin{aligned}\tilde{\boldsymbol{s}}_{SR_k}[i] &= \boldsymbol{A}_{R_kD}[i](\boldsymbol{F}_{SR_k}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SR_k}[i]) \\ &= \boldsymbol{A}_{R_kD}[i]\boldsymbol{F}_{SR_k}[i]\boldsymbol{s}[i] + \boldsymbol{A}_{R_kD}[i]\boldsymbol{n}_{SR_k}[i] \\ &= \boldsymbol{F}_{R_k}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{R_k}[i],\end{aligned} \quad (39)$$

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 (39) into (4), the relationship between all the relay nodes and the destination node can be written as

$$\begin{aligned}\boldsymbol{r}_{RD} &= \sum_{k=1}^{n_r}\boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i](\boldsymbol{F}_{R_k}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{R_k}[i]) + \boldsymbol{n}_{RD}[i] \\ &= \sum_{k=1}^{n_r}\boldsymbol{\Phi}_{eq_k}[i]\boldsymbol{D}_k[i]\boldsymbol{s}[i] + \boldsymbol{n}_D[i] \\ &= \sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_D[i],\end{aligned} \quad (40)$$

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 (40) into (5), we can rewrite the MSE optimization problem as

$$\begin{aligned}&[\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_{k_j}}[i]] = \arg\min_{\boldsymbol{w}_j[i], \boldsymbol{\Phi}_{eq_{k_j}}[i]} \\ &E\left[\|s_j[i] - \boldsymbol{w}_j^H[i](\sum_{k=1}^{n_r}\sum_{j=1}^{N}\boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_D[i])\|^2\right], \\ &\quad s.t. \;\; \mathrm{Tr}(\sum_{\mathrm{j}=1}^{\mathrm{N}}\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}[\mathrm{i}]\boldsymbol{\Phi}_{\mathrm{eq_{k_j}}}^{\mathrm{H}}[\mathrm{i}]) \leq \mathrm{P_R}.\end{aligned} \quad (41)$$

By taking the instantaneous gradient of $\mathscr{L}$ in (35) with respect to $\boldsymbol{w}_j^*[i]$ and $\boldsymbol{\Phi}_{eq_{k_j}}^*[i]$ we can obtain

$$\begin{aligned}
\nabla \mathscr{L}_{\boldsymbol{w}_j^*[i]} &= \nabla E\left[\|s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]\|^2\right]_{\boldsymbol{w}_j^*[i]} \\
&= (s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i])^H \nabla_{\boldsymbol{w}_j^*[i]}(s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]) \\
&= -e_j^*[i]\boldsymbol{r}[i], \\
\nabla \mathscr{L}_{\boldsymbol{\Phi}_{eq_{k_j}}^*[i]} &= \nabla E[\|s_j[i] - \boldsymbol{w}_j^H[i] \\
&\quad (\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_{RD}[i])\|^2]_{\boldsymbol{\Phi}_{eq_{k_j}}^*[i]} \\
&= \nabla_{\boldsymbol{\Phi}_{eq_{k_j}}^*[i]}(s_j[i] - \boldsymbol{w}_j^H[i] \\
&\quad (\sum_{k=1}^{n_r}\sum_{j=1}^{N} \boldsymbol{\Phi}_{eq_{k_j}}[i]\boldsymbol{d}_{k_j}[i]s_j[i] + \boldsymbol{n}_{RD}[i])\|^2)^H \\
&\quad (s_j[i] - \boldsymbol{w}_j^H[i]\boldsymbol{r}[i]) \\
&= -e_j[i]s_j^*[i]\boldsymbol{w}_j[i]\boldsymbol{d}_{k_j}^H[i],
\end{aligned} \tag{42}$$

where $e_j[i]$ 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

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

References

[1] P. Clarke, R. C. de Lamare, "Joint Transmit Diversity Optimization and Relay Selection for Multi-relay Cooperative MIMO Systems Using Discrete Stochastic Algorithms", IEEE Communications Letters, vol. 15, p.p. 1035-1037, Oct. 2011.

[2] J. N. Laneman and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behaviour", *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] Yulong Zou, Yu-Dong Yao, Baoyu Zheng, "Opportunistic Distributed Space-Time Coding for Decode-and-Forward Cooperation Systems", IEEE Trans. on Signal Processing, vol. 60, pp. 1766 - 1781, April 2012.

[7] 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.

[8] B. Maham, A. Hjorungnes, G. Abreu, "Distributed GABBA Space-Time Codes in Amplify-and-Forward Relay Networks", IEEE Trans. on Wireless Commun., vol.8, Issue: 4, pp. 2036 - 2045, April 2009.

[9] Sheng Yang, J.-C. Belfiore, "Optimal Space-Time Codes for the MIMO Amplify-and-Forward Cooperative Channel", IEEE Transactions on Information Theory, vol. 53, Issue: 2, pp. 647-663, Feb. 2007.

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

[11] 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.

[12] A.D. Dabbagh, D.J. Love, "Feedback Rate-Capacity Loss Trade-off for Limited Feedback MIMO Systems", IEEE Transactions on Information Theory, vol. 52, p.p. 2190-2202, May 2006.

[13] J. Akhtar, D. Gesbert, "Extending Orthogonal Block Codes with Partial Feedback", IEEE Transactions on Wireless Communications, vol. 3, p.p. 1959-1962, Nov. 2004.

[14] Choi Ilhwan, Kim Jong-Kyu, Lee Heunchul, Lee Inkyu, "Alamouti-Codes Based Four-Antenna Transmission Schemes with Phase Feedback", IEEE Communications Letters, vol. 13, p.p. 749-751, Oct. 2009.

[15] G. Jongren, M. Skoglund, "Quantized Feedback Information in Orthogonal Space-Time Block Coding", IEEE Transactions on Information Theory, vol. 50, p.p. 2473-2486, Oct. 2004.

[16] 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.

[17] S. Haykin, "Adaptive Filter Theory", $4th$ ed. Englewood Cliffs, NJ: Prentice- Hall, 2002.

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

[19] R.C. de Lamare, 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.

[20] G. Taricco, E. Biglieri, "Exact Pairwise Error Probability of Space-Time Codes", IEEE Transactions on Information Theory, vol. 48, p.p. 510-513, Feb 2002.

[21] D.J. Tylavsky and G.R.L. Sohie. "Generalization of the matrix inversion lemma", Proceedings of the IEEE, 74(7):1050¨C1052, July 1986.

[22] T. Peng, R.C. de Lamare and A. Schmeink, "Adaptive Distributed Space-Time Coding for Cooperative MIMO Relaying Systems", accepted by ISWCS 2012.

Adaptive Randomized Distributed Space-Time Coding for Cooperative MIMO Relaying Systems

Tong Peng §, Rodrigo C. de Lamare § and Anke Schmeink ♣
§ Communications Research Group, Department of Electronics, University of York, York YO10 5DD, UK
♣ UMIC Research Centre, RWTH Aachen University, D-52056 Aachen, Germany
Emails: tp525@ohm.york.ac.uk; rcdl500@ohm.york.ac.uk, schmeink@umic.rwth-aachen.de

***Abstract*—An adaptive randomized distributed space-time coding (DSTC) scheme and algorithms are proposed for two-hop cooperative MIMO networks. Linear minimum mean square error (MMSE) receivers and an amplify-and-forward (AF) cooperation strategy are considered. In the proposed DSTC scheme, a randomized matrix obtained by a feedback channel is employed to transform the space-time coded matrix at the relay node. Linear MMSE expressions are devised to compute the parameters of the adaptive randomized matrix and the linear receive filter. A stochastic gradient algorithm is also developed to compute the parameters of the adaptive randomized matrix with reduced computational complexity. We also derive the upper bound of the error probability of a cooperative MIMO system employing the randomized space-time coding scheme first. The simulation results show that the proposed algorithms obtain significant performance gains as compared to existing DSTC schemes.**

I. Introduction

Multiple-input and multiple-output (MIMO) communication systems employ multiple collocated antennas at both the source node and the destination node in order to obtain the diversity gain and combat multi-path fading in wireless links. The different methods of space-time coding (STC) schemes, which can provide a higher diversity gain and coding gain compared to an uncoded system, are also utilized in MIMO wireless systems for different numbers of antennas at the transmitter and different conditions of the channel. Cooperative MIMO systems, which employ multiple relay nodes with antennas between the source node and the destination node as a distributed antenna array, apply distributed diversity gain and provide copies of the transmitted signals to improve the reliability of wireless communication systems [1]. Among the links between the relay nodes and the destination node, cooperation strategies, such as Amplify-and-Forward (AF), Decode-and-Forward (DF), and Compress-and-Forward (CF) [2] and various distributed STC (DSTC) schemes in [3], [4] and [17] can be employed.

The utilization of a DSTC at the relay node in a cooperative network, providing more copies of the desired symbols at the destination node, can offer the system diversity gains and coding gains to combat the interference. The recent focus on the DSTC technique lies in the delay-toleration code design and the full-diversity schemes design with the minimum outage probability. In [5], the distributed delay-tolerant version of the Golden code [6] is proposed, which can provide full-diversity gain with a full coding rate. An opportunistic DSTC scheme with the minimum outage probability is designed for a DF cooperative network and compared with the fixed DSTC schemes in [7]. An adaptive distributed-Alamouti (D-Alamouti) STBC design is proposed in [8] for the non-regenerative dual-hop wireless system which achieves the minimum outage probability. DSTC schemes for the AF protocol are discussed in [9]-[10]. In [9], the GABBA STC scheme is extended to a distributed MIMO network with full-diversity and full-rate, while an optimal algorithm for design of the DSTC scheme to achieve the optimal diversity and multiplexing tradeoff is derived in [10].

The performance of cooperative networks using different strategies has been widely discussed in the literature. In [11], an exact pairwise error probability of the D-Alamouti STBC scheme is derived according to the position of the relay node. In [12], a bit error rate (BER) analysis of the D-Alamouti STBC scheme is proposed. The difference between these two works lies in the different cooperative schemes considered. A maximum likelihood (ML) detection algorithm for a MIMO relay system with DF protocol is derived in [13] with its performance analysis as well. The symbol error rate and diversity order upper bound for the scalar fixed-gain AF cooperative protocol are given in [14]. The use of single-antenna relay nodes and the DF cooperative protocol is the main difference in scenario between [14] and this work. An STC encoding process is implemented at the source node in [15], which decreases the output of the system and increases the computational complexity of the decoding at the destination node. In [16], the BER upper bound is given without a STC scheme at the relay node.

In this paper, we propose an adaptive randomized distributed space-time coding scheme and algorithms for a two-hop cooperative MIMO relaying system with the AF protocol and linear MMSE receivers. We focus on how the randomized matrix affects the DSTC during the encoding and how to optimize the parameters in the matrix. It is shown that the use of a randomized matrix benefits the performance of the system by lowering the upper bound compared to using traditional STC schemes. Linear MMSE expressions are devised to compute the parameters of the adaptive randomized matrix and the linear receive filter. Then an adaptive optimization algorithm is derived based on the MSE criterion, with the stochastic gradient (SG) algorithm in order to reduce the computational complexity of the optimization process. The updated randomized matrix is transmitted to the relay node through a feedback channel that is assumed in this work error free and delay free. The upper bound pairwise error probability of the randomized-STC schemes (RSTC) in a cooperative MIMO system which employs multi-antenna relay nodes with the AF protocol is also analyzed.

The paper is organized as follows. Section II introduces a two-hop cooperative MIMO system with multiple relays

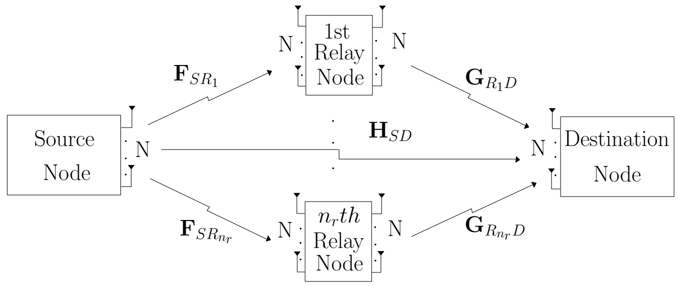

Fig. 1. Cooperative MIMO System Model with n_r Relay nodes

applying the AF strategy and the randomized DSTC scheme. In Section III the proposed MMSE expressions and the SG algorithm for the randomized matrix are derived, and the analysis of the upper bound of pairwise error probability using the randomized D-STC is shown in Section IV. Section V focuses on the results of the simulations and Section VI leads to the conclusion.

II. Cooperative MIMO System Model

The communication system under consideration is a two-hop cooperative MIMO system which employing multiple relay nodes and communicating over channels from the source node to the relay nodes and the destination node, and from the relay nodes to the destination nodes as shown in Fig. 1. A modulation scheme is used in our system to generate the transmitted symbol vector $\boldsymbol{s}[i]$ at the source node and all nodes have N antennas. There are n_r relay nodes that employ the AF cooperative strategy as well as a DSTC scheme. The system broadcasts symbols from the source to n_r relay nodes as well as to the destination node in the first phase. The symbols are amplified and re-encoded at each relay node prior to transmission to the destination node in the second phase. We consider only one user at the source node in our system that has N Spatial Multiplexing (SM)-organized data symbols contained in each packet. The received symbols at the $k-th$ relay node and the destination node are denoted as $\boldsymbol{r}_{SR_k}$ and $\boldsymbol{r}_{SD}$, respectively, where $k = 1, 2, ..., n_r$. The received symbols $\boldsymbol{r}_{SR_k}$ are amplified before mapped into an STC matrix. We assume that the synchronization at each node is perfect. The received symbols at the destination node and each relay node are described by

$$\boldsymbol{r}_{SR_k}[i] = \boldsymbol{F}_k[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SR_k}[i], \tag{1}$$

$$\boldsymbol{r}_{SD}[i] = \boldsymbol{H}[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SD}[i], \tag{2}$$

$$i = 1, 2, \ \dots, N, \quad k = 1, 2, \ \dots n_r,$$

where the $N \times 1$ vector $\boldsymbol{n}_{SR_k}[i]$ and $\boldsymbol{n}_{SD}[i]$ denote the zero mean complex circular symmetric additive white Gaussian noise (AWGN) vector generated at each relay and the destination node with variance σ^2. The transmitted symbol vector $\boldsymbol{s}[i]$ contains N parameters, $\boldsymbol{s}[i] = [s_1[i], s_2[i], ..., s_N[i]]$, which has a covariance matrix $E\big[\boldsymbol{s}[i]\boldsymbol{s}^H[i]\big] = \sigma_s^2\boldsymbol{I}$, where $E[\cdot]$ stands for expected value, $(\cdot)^H$ denotes the Hermitian operator, σ_s^2 is the signal power which we assume to be equal to 1 and $\boldsymbol{I}$ is the identity matrix. $\boldsymbol{F}_k[i]$ and $\boldsymbol{H}[i]$ are the $N \times N$ channel gain matrices between the source node and the kth relay node, and between the source node and the destination node, respectively.

After processing and amplifying the received vector $\boldsymbol{r}_{SR_k}[i]$ at the kth relay node, the signal vector $\tilde{\boldsymbol{s}}_{SR_k}[i] = \boldsymbol{A}_{R_kD}[i](\boldsymbol{F}_k[i]\boldsymbol{s}[i] + \boldsymbol{n}_{SR_k}[i])$ can be obtained and then forwarded to the destination node. The amplified symbols in $\tilde{\boldsymbol{s}}_{SR_k}[i]$ will be re-encoded by a $N \times T$ DSTC scheme $\boldsymbol{M}(\tilde{\boldsymbol{s}}[i])$ and then multiplied by an $N \times N$ randomized matrix $\boldsymbol{\mathfrak{R}}[i]$ in [19], then forwarded to the destination node. The relationship between the $k-th$ relay and the destination node can be described as

$$\boldsymbol{R}_{R_kD}[i] = \boldsymbol{G}_k[i]\boldsymbol{\mathfrak{R}}[i]\boldsymbol{M}_{R_kD}[i] + \boldsymbol{N}_{R_kD}[i], \tag{3}$$
$$k = 1, 2, ..., n_r,$$

where the $N \times T$ matrix $\boldsymbol{M}_{R_kD}[i]$ is the DSTC matrix employed at the relay nodes whose elements are the amplified symbols in $\tilde{\boldsymbol{s}}_{SR_k}[i]$. The $N \times T$ received symbol matrix $\boldsymbol{R}_{R_kD}[i]$ in (3) can be written as an $NT \times 1$ vector $\boldsymbol{r}_{R_kD}[i]$ given by

$$\boldsymbol{r}_{R_kD}[i] = \boldsymbol{\mathfrak{R}}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] + \boldsymbol{n}_{R_kD}[i], \tag{4}$$

where the block diagonal $NT \times NT$ matrix $\boldsymbol{\mathfrak{R}}_{eq_k}[i]$ denotes the equivalent randomized 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 $\boldsymbol{N}_{R_kD}[i]$. After rewriting $\boldsymbol{R}_{R_kD}[i]$ we can consider the received symbol vector at the destination node as a $N(n_r+1)$ 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

$$\begin{aligned}\boldsymbol{r}[i] &= \begin{bmatrix} \boldsymbol{H}[i]\boldsymbol{s}[i] \\ \sum_{k=1}^{n_r}\boldsymbol{\mathfrak{R}}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\tilde{\boldsymbol{s}}_{SR_k}[i] \end{bmatrix} + \begin{bmatrix} \boldsymbol{n}_{SD}[i] \\ \boldsymbol{n}_{RD}[i] \end{bmatrix} \\ &= \boldsymbol{D}_D[i]\tilde{\boldsymbol{s}}_D[i] + \boldsymbol{n}_D[i],\end{aligned} \tag{5}$$

where the $(T+1)N \times (n_r+1)N$ 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 $\boldsymbol{H}[i]$ 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. The $(n_r+1)N \times 1$ noise vector $\boldsymbol{n}_D[i]$ contains the received noise vector at the destination node and the amplified noise vectors from each relay node, which can be modeled as additive white Gaussian noise (AWGN) with zero mean and covariance matrix $\sigma^2(1+ \parallel \boldsymbol{\mathfrak{R}}_{eq_k}[i]\boldsymbol{G}_{eq_k}[i]\boldsymbol{A}_{R_kD}[i] \parallel_F^2)\boldsymbol{I}$, where $\parallel \boldsymbol{X} \parallel_F = \sqrt{\mathrm{Tr}(\boldsymbol{X}^H \cdot \boldsymbol{X})} = \sqrt{\mathrm{Tr}(\boldsymbol{X} \cdot \boldsymbol{X}^H)}$ is the Frobenius norm.

III. Design of Linear MMSE Receivers and Randomized Matrices

In this section, we design an adaptive linear MMSE receive filter and an MMSE randomized matrix for use with the proposed DSTC scheme. An adaptive SG algorithm [18] for determining the parameters of the randomized matrix with reduced complexity is also devised. The DSTC scheme used at the relay node employs an MMSE randomized matrix, which is computed at the destination node and obtained by a feedback channel and processes the data symbols prior to transmission to the destination node.

A. Optimization Method Based on the MSE Criterion

Let us consider the MMSE design of the receive filter and the randomized matrix according to the optimization problem

$$[\boldsymbol{W}[i], \mathfrak{R}_{eq}[i]] = \arg \min_{\boldsymbol{W}[i], \mathfrak{R}_{eq}[i]} E\left[\|\boldsymbol{s}[i] - \boldsymbol{W}^H[i]\boldsymbol{r}[i]\|^2\right],$$

where $\boldsymbol{r}[i]$ is the received symbol vector at the destination node which contains the randomized matrix to be optimized. If we only consider the received symbols from the relay node, the received symbol vector at the destination node can be derived as

$$\begin{aligned}\boldsymbol{r}[i] &= \boldsymbol{D}_D[i]\tilde{\boldsymbol{s}}_D[i] + \boldsymbol{n}_D[i] \\ &= \mathfrak{R}_{eq}[i]\boldsymbol{G}_{eq}[i]\boldsymbol{A}[i]\boldsymbol{F}[i]\boldsymbol{s}[i] + \mathfrak{R}_{eq}[i]\boldsymbol{G}_{eq}[i]\boldsymbol{A}[i]\boldsymbol{n}_{SR}[i] \\ &\quad + \boldsymbol{n}_{RD}[i] \\ &= \mathfrak{R}_{eq}[i]\boldsymbol{C}[i]\boldsymbol{s}[i] + \boldsymbol{n}_D[i],\end{aligned} \tag{6}$$

where $\boldsymbol{C}[i]$ is an $NT \times N$ matrix that contains all the complex channel gains and the amplified matrix assigned to the received vectors at the relay node, and the noise vector $\boldsymbol{n}_D$ is a Gaussian noise with zero mean and variance $\sigma^2(1+ \parallel \mathfrak{R}_{eq}[i]\boldsymbol{G}_{eq}[i]\boldsymbol{A}[i] \parallel_F^2)$. We can then recast the optimization as

$$\begin{aligned}&[\boldsymbol{W}[i], \mathfrak{R}_{eq}[i]] = \\ &\arg \min_{\boldsymbol{W}[i], \mathfrak{R}_{eq}[i]} E\left[\|\boldsymbol{s}[i] - \boldsymbol{W}^H[i](\mathfrak{R}_{eq}[i]\boldsymbol{C}[i]\boldsymbol{s}[i] + \boldsymbol{n}_D[i])\|^2\right].\end{aligned} \tag{7}$$

By expanding the righthand side of (7) and taking the gradient with respect to $\boldsymbol{W}^*[i]$ and equating the terms to zero, we can obtain the linear MMSE receive filter

$$\boldsymbol{W}[i] = \left(E\left[\boldsymbol{r}[i]\boldsymbol{r}^H[i]\right]\right)^{-1} E\left[\boldsymbol{r}[i]\boldsymbol{s}^H[i]\right], \tag{8}$$

where the first term denotes the inverse of the auto-correlation matrix and the second one is the cross-correlation matrix. Define $\tilde{\boldsymbol{r}} = \boldsymbol{C}[i]\boldsymbol{s}[i] + \boldsymbol{C}[i]\boldsymbol{n}_{SR}$, then the randomized matrix can be calculated by taking the gradient with respect to $\mathfrak{R}^*[i]$ and equating the terms to zero, resulting in

$$\mathfrak{R}[i] = \left(\boldsymbol{W}^H[i](E\left[\tilde{\boldsymbol{r}}[i]\tilde{\boldsymbol{r}}^H[i]\right])\boldsymbol{W}[i]\right)^{-1} E\left[\boldsymbol{s}[i]\tilde{\boldsymbol{r}}^H[i]\right]\boldsymbol{W}[i], \tag{9}$$

where $E\left[\tilde{\boldsymbol{r}}[i]\tilde{\boldsymbol{r}}^H[i]\right]$ is the auto-correlation of the space-time coded received symbol vector at the relay node, and $E\left[\boldsymbol{s}[i]\tilde{\boldsymbol{r}}^H[i]\right]$ is the cross-correlation. The expression above requires a matrix inversion with a high computational complexity.

B. Adaptive Randomized Matrix Optimization Algorithm

In order to reduce the computational complexity and achieve the optimal performance, an adaptive randomized matrix optimization (ARMO) algorithm based on an SG algorithm is devised. The MMSE problem is derived in (7), and the MMSE filter matrix can be calculated by (8) first during the optimization process. The simple ARMO algorithm can be obtained by taking the instantaneous gradient term of (7) with respect to the randomized matrix $\mathfrak{R}_{eq}{}^*[i]$, which is given by

$$\begin{aligned}&\nabla \mathrm{L}_{\mathfrak{R}_{eq}{}^*[i]} \\ &= \nabla E\left[\|\boldsymbol{s}[i] - \boldsymbol{W}^H[i](\mathfrak{R}_{eq}[i]\boldsymbol{C}[i]\boldsymbol{s}[i] + \boldsymbol{n}_D[i])\|^2\right]_{\mathfrak{R}_{eq}{}^*[i]} \\ &= -(\boldsymbol{s}[i] - \boldsymbol{W}^H[i]\boldsymbol{r}[i])\boldsymbol{s}^H[i]\boldsymbol{C}^H[i]\boldsymbol{W}[i] \\ &= -\boldsymbol{e}[i]\boldsymbol{s}^H[i]\boldsymbol{C}^H[i]\boldsymbol{W}[i],\end{aligned} \tag{10}$$

where $\boldsymbol{e}[i]$ stands for the detected error vector. After computing (10), the ARMO algorithm can be obtained by introducing a step size into an SG algorithm to update the result until the convergence is reached as given by

$$\mathfrak{R}[i+1] = \mathfrak{R}[i] + \mu(\boldsymbol{e}[i]\boldsymbol{s}^H[i]\boldsymbol{C}^H[i]\boldsymbol{W}[i]), \tag{11}$$

where μ stands for the step size of the ARMO algorithm. The complexity of calculating the randomized matrix is $\mathrm{O}(2N)$, which is much less than that of the calculation method derived in (9). As mentioned in Section I, the randomized matrix will be sent back to the relay nodes via a feedback channel which is assumed to be error-free in this work. However, in practical circumstances, the errors caused by the broadcasting and the diversification of the feedback channel with time changes will affect the accuracy of the received randomized matrix at the relay nodes.

IV. Probability of Error Analysis

In this section, the upper bound of the pairwise error probability of the system employing the randomized DSTC will be derived. As we mentioned in the first section, the randomized matrix will be considered in the derivation as it affects the performance by reducing the upper bound of the pairwise error probability. For the sake of simplicity, we consider a 2 by 2 MIMO system with 1 relay node, and the direct link is ignored in order to concentrate on the effect of the randomized matrix. The expression of the upper bound is also stable for the increase of the system size and the number of relay nodes.

Consider an $N \times N$ STC scheme we use at the relay node with L codewords. The codeword $\boldsymbol{C}^1$ is transmitted and decoded to another codeword $\boldsymbol{C}^i$ at the destination node, where $i = 1, 2, ..., L$. According to [20], the probability of error can be upper bounded by the sum of all the probabilities of incorrect decoding, which is given by

$$\mathrm{P_e} \leq \sum_{i=2}^{L} \mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^i). \tag{12}$$

Assuming the codeword $\boldsymbol{C}^2$ is decoded at the destination node and we know the channel information perfectly at the destination node, we can derive the pairwise error probability as

$$\begin{aligned}&\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \mathfrak{R}) \\ &= \mathrm{P}(\parallel \boldsymbol{R}^1 - \boldsymbol{G}\mathfrak{R}\boldsymbol{C}^1 \parallel_F^2 - \parallel \boldsymbol{R}^1 - \boldsymbol{G}\mathfrak{R}\boldsymbol{C}^2 \parallel_F^2 > 0 \mid \mathfrak{R}_{eq}) \\ &= \mathrm{P}(\parallel \boldsymbol{r}^1 - \mathfrak{R}_{eq}\boldsymbol{G}_{eq}\boldsymbol{F}\boldsymbol{s}^1 \parallel_F^2 \\ &\qquad - \parallel \boldsymbol{r}^1 - \mathfrak{R}_{eq}\boldsymbol{G}_{eq}\boldsymbol{F}\boldsymbol{s}^2 \parallel_F^2 > 0 \mid \mathfrak{R}_{eq}),\end{aligned} \tag{13}$$

where $\boldsymbol{F}$ and $\boldsymbol{G}_{eq}$ stand for the channel coefficient matrix between the source node and the relay node, and between the relay node and the destination node, respectively. The randomized matrix is denoted by $\mathfrak{R}_{eq}$. Define $\boldsymbol{H} = \boldsymbol{G}_{eq}\boldsymbol{F}$, which stands for the total channel coefficients matrix. After the calculation, we can transfer the pairwise error probability expression in (13) to

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \mathfrak{R}_{eq}) = \mathrm{P}(\| \mathfrak{R}_{eq}\boldsymbol{H}(\boldsymbol{s}^1 - \boldsymbol{s}^2) \|_F^2 < \boldsymbol{Y}), \quad (14)$$

where $\boldsymbol{Y} = \mathrm{Tr}(\boldsymbol{n^1}^H\mathfrak{R}_{eq}\boldsymbol{H}(\boldsymbol{s}^1-\boldsymbol{s}^2)+(\mathfrak{R}_{eq}\boldsymbol{H}(\boldsymbol{s}^1-\boldsymbol{s}^2))^H\boldsymbol{n}^1)$, and $\boldsymbol{n}^1$ denotes the noise vector at the destination node with zero mean and covariance matrix $\sigma^2(1+ \| \mathfrak{R}_{eq}\boldsymbol{G}_{eq} \|_F^2)\boldsymbol{I}$. By making use of the Q function, we can derive the error probability function as

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \mathfrak{R}_{eq}) = \mathrm{Q}\left(\sqrt{\frac{\gamma}{2}} \| \mathfrak{R}_{eq}\boldsymbol{H}(\boldsymbol{s}^1 - \boldsymbol{s}^2) \|_F\right), \quad (15)$$

where

$$\mathrm{Q} = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} \exp\left(-\frac{u^2}{2}\right)\mathrm{d}u, \quad (16)$$

and γ is the received SNR at the destination node assuming the transmit power is equal to 1.

In order to obtain the upper bound of $\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \mathfrak{R}_{eq})$ we expand the formula $\| \mathfrak{R}_{eq}\boldsymbol{H}(\boldsymbol{s}^1-\boldsymbol{s}^2) \|_F^2$. Let $\boldsymbol{U}^H\boldsymbol{\Lambda}_s\boldsymbol{U}$ be the eigenvalue decomposition of $(\boldsymbol{s}^1 - \boldsymbol{s}^2)^H(\boldsymbol{s}^1 - \boldsymbol{s}^2)$, where $\boldsymbol{U}$ is a Hermitian matrix and $\boldsymbol{\Lambda}_s$ contains all the eigenvalues of the difference between two different codewords $\boldsymbol{s}^1$ and $\boldsymbol{s}^2$. Let $\boldsymbol{V}^H\boldsymbol{\Lambda}_{\mathfrak{R}}\boldsymbol{V}$ stand for the eigenvalue decomposition of $(\mathfrak{R}_{eq}\boldsymbol{H}\boldsymbol{U})^H\mathfrak{R}_{eq}\boldsymbol{H}\boldsymbol{U}$, where $\boldsymbol{V}$ is a random Hermitian matrix and $\boldsymbol{\Lambda}_{\mathfrak{R}}$ is the ordered diagonal eigenvalue matrix. Therefore, the probability of error can be written as

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \mathfrak{R}_{eq}) = \mathrm{Q}\left(\sqrt{\frac{\gamma}{2}\sum_{m=1}^{NT}\sum_{n=1}^{N}\lambda_{\mathfrak{R}_m}\lambda_{s_n}|\xi_{n,m}|^2}\right), \quad (17)$$

where $\xi_{n,m}$ is the $(n,m)-th$ element in $\boldsymbol{V}$, and $\lambda_{\mathfrak{R}_m}$ and λ_{s_n} are eigenvalues in $\boldsymbol{\Lambda}_{\mathfrak{R}}$ and $\boldsymbol{\Lambda}_s$, respectively. According to [20], a good upper bound assumption of the Q function is given by

$$\mathrm{Q}(x) \leq \frac{1}{2}e^{\frac{-x^2}{2}}. \quad (18)$$

Thus, we can derive the upper bound of pairwise error probability for a randomized STC scheme as

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \mathfrak{R}_{eq}) \leq \frac{1}{2}\exp\left(-\frac{\gamma}{4}\sum_{m=1}^{NT}\sum_{n=1}^{N}\lambda_{\mathfrak{R}_m}\lambda_{s_n}|\xi_{n,m}|^2\right), \quad (19)$$

while the upper bound of the error probability expression for a traditional STC is given by

$$\mathrm{P}(\boldsymbol{C}^1 \rightarrow \boldsymbol{C}^2 \mid \boldsymbol{H}_{eq}) \leq \frac{1}{2}\exp\left(-\frac{\gamma}{4}\sum_{m=1}^{NT}\sum_{n=1}^{N}\lambda_{s_n}|\xi_{n,m}|^2\right). \quad (20)$$

With comparison of (19) and (20), it is obvious to note that the eigenvalue of the randomized matrix is the difference, which suggests that employing a randomized matrix for a STC scheme at the relay node can provide an improvement in BER performance.

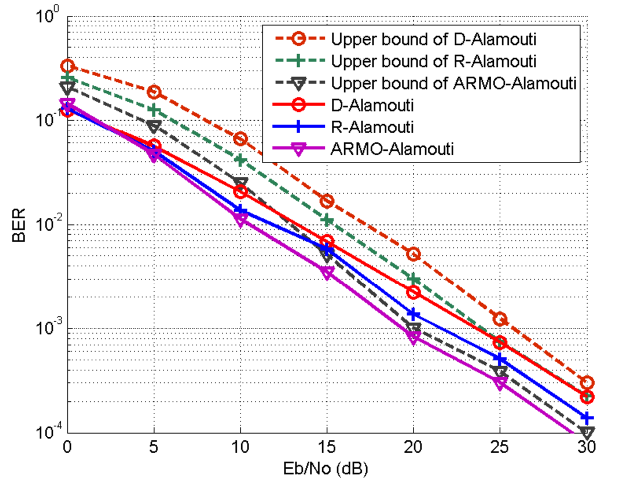

Fig. 2. BER performance v.s. E_b/N_0 for the upper bound of the R-Alamouti scheme without the Direct Link

V. 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 [20] using QPSK modulation in a quasi-static block fading channel with AWGN. The simulation system with 1 relay node and 2 antennas at each node. In the simulations, we define both the symbol power and the noise variance σ^2 for as equal to 1, and the power of the adaptive randomized matrix in the ARMO algorithm are normalized to the same transmission power of that in the R-Alamouti.

The upper bounds of the D-Alamouti and the randomized D-Alamouti derived in the previous section are shown in Fig. 2. The theoretical pairwise error probabilities provide the largest decoding errors of the two 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 comparison of the simulation results in BER performance of the R-Alamouti and the D-Alamouti indicates the advantage of using the randomized matrix.

The proposed ARMO algorithm is compared with the SM scheme and the traditional RSTC algorithm using the distributed-Alamouti (D-Alamouti) STBC scheme in [17] with $n_r = 1$ relay nodes in Fig. 3. The number of antennas $N = 2$ at each node and the effect of the direct link are considered. The results illustrate that without the direct link, by making use of the STC or the RSTC technique, a significant performance improvement can be achieved compared to the spatial multiplexing system. The RSTC algorithm outperforms the STC-AF system, while the ARMO algorithm can improve the performance by about 3dB as compared to the RSTC algorithm. With the consideration of the direct link, the results indicate that the cooperative diversity order can be increased, and using the ARMO algorithm achieves an improved performance with 2dB of gain as compared to employing the RSTC algorithm and 3dB of gain as compared to employing the traditional STC-AF algorithm.

The simulation results shown in Fig. 4 illustrate the convergence property of the ARMO algorithm. The SM, D-

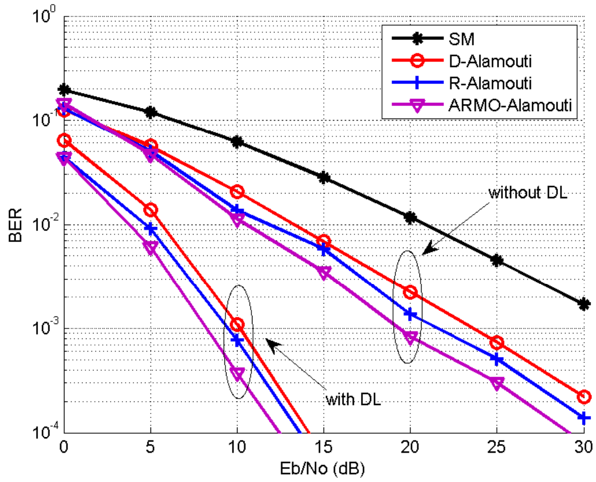

Fig. 3. BER performance v.s. E_b/N_0 for ARMO Algorithm with and without the Direct Link

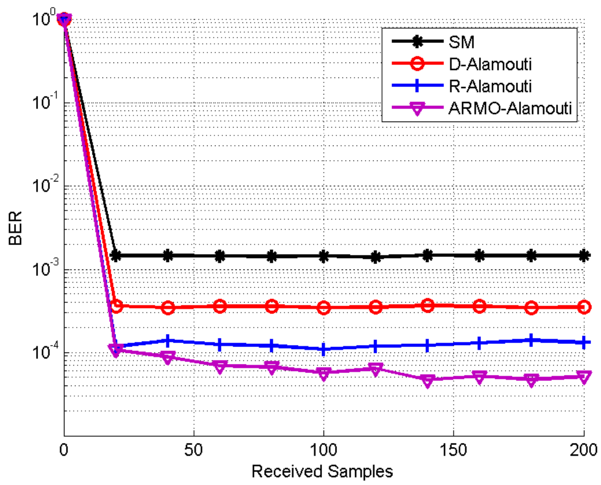

Fig. 4. BER performance v.s. Number of Samples for ARMO Algorithm without the Direct Link

Alamouti and the randomized D-Alamouti algorithms obtain nearly flat performance in BER as the utilization of fixed STC scheme and the randomized matrix. The SM scheme has the worst performance due to the lack of coding gains, while the D-Alamouti scheme can provide a significant performance improvement in terms of the BER improvement, and by employing the randomized matrix at the relay node the BER performance can decrease further when the transmission circumstances are the same as that of the D-Alamouti. The ARMO algorithm shows its advantage in a fast convergence and a lower BER achievement. At the beginning of the optimization process with a small number of samples, the ARMO algorithm achieves the BER level of the D-Alamouti one, but with the increase of the received symbols, the ARMO algorithm achieves a better BER performance.

VI. Conclusion

We have proposed an adaptive randomized matrix optimization (ARMO) algorithm for the randomized DSTC using a linear MMSE receive filter at the destination node. The pairwise error probability of introducing the randomized DSTC in a cooperative MIMO network with the AF protocol has been derived. The simulation results illustrate the advantage of the proposed ARMO algorithm by comparing it with the cooperative network employing the traditional DSTC scheme and the fixed randomized STC scheme. The proposed algorithm can be used with different distributed STC schemes using the AF strategy and can also be extended to the DF cooperation protocol.

References

[1] 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.

[2] J. N. Laneman and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behaviour", *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] M. Sarkiss, G.R.-B. Othman, M.O. Damen, J.-C. Belfiore, "Construction of New Delay-Tolerant Space-Time Codes", IEEE Transactions on Information Theory, vol. 57, Issue: 6, pp. 3567 - 3581, June 2011.

[6] J.-C. Belfiore, G. Rekaya, E. Viterbo, "The Golden Code: A 2¡Á2 Full-Rate Space-Time Code with Nonvanishing Determinants", IEEE Transactions on Information Theory, vol. 51 , Issue: 4, April 2005.

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

[8] 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.

[9] B. Maham, A. Hjorungnes, G. Abreu, "Distributed GABBA Space-Time Codes in Amplify-and-Forward Relay Networks", IEEE Trans. on Wireless Commun., vol.8, Issue: 4, pp. 2036 - 2045, April 2009.

[10] Sheng Yang, J.-C. Belfiore, "Optimal Space-Time Codes for the MIMO Amplify-and-Forward Cooperative Channel", IEEE Transactions on Information Theory, vol. 53, Issue: 2, pp. 647-663, Feb. 2007.

[11] T.Q. Duong, Ngoc-Tien Nguyen, Viet-Kinh Nguyen, "Exact Pairwise Error Probability of Distributed Space-Time Coding in Wireless Relays Networks", International Symposium on Communications and Information Technologies 2007, pp. 279 - 283, 17-19 Oct. 2007.

[12] Minchul Ju, Hyoung-Kyu Song, Il-Min Kim, "Exact BER Analysis of Distributed Alamouti's Code for Cooperative Diversity Networks", IEEE Trans. on Commun., vol. 57, Issue: 8, pp. 2380 - 2390, Aug. 2009.

[13] G.V.V. Sharma, V. Ganwani, U.B. Desai, S.N. Merchant, "Performance Analysis of Maximum Likelihood Detection for Decode and Forward MIMO Relay Channels in Rayleigh Fading", IEEE Trans. on Wireless Commun., vol. 9, Issue: 9, pp. 2880 - 2889, September 2010.

[14] Youngpil Song, Hyundong Shin, Een-Kee Hong, "MIMO Cooperative Diversity with Scalar-Gain Amplify-and-Forward Relaying", IEEE Trans. on Communications, vol. 57, Issue: 7, pp. 1932 - 1938, July 2009.

[15] Liang Yang, Q.T. Zhang, "Performance Analysis of MIMO Relay Wireless Networks With Orthogonal STBC", IEEE Transactions on Vehicular Technology, vol. 59, Issue: 7, pp. 3668 - 3674, Sept. 2010.

[16] T. Unger,A. Klein, "On the Performance of Distributed Space-Time Block Codes in Cooperative Relay Networks", IEEE Communications Letters, vol. 11, Issue: 5, pp. 411 - 413, May 2007.

[17] 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

[18] S. Haykin, "Adaptive Filter Theory", $4th$ ed. Englewood Cliffs, NJ: Prentice- Hall, 2002.

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

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