Minimum Error Rate Linear Dispersion Codes for Cooperative Relays

Kuo-ching Liang[†], *Inaki Berenguer*[§ †], *and Xiaodong Wang*[†]

[†]Columbia University, New York, NY
[§]University of Cambridge, Cambridge, UK

***Abstract*— Cooperative diversity systems have been recently proposed as a solution to provide spatial diversity for terminals where multiple antennas are not feasible to be implemented. As in MIMO systems, space-time codes can be used to efficiently exploit the increase in capacity provided in cooperative diversity systems. In this paper we propose a two-layer linear dispersion (LD) code for cooperative diversity systems and derive a simulation-based optimization algorithm to optimize the LD code and power allocation in terms of block error rate. The proposed code design paradigm can obtain optimal codes under arbitrary fading statistics. The effect that distances between source, relays, and destination terminals have on the energy allocation between the broadcast and cooperative intervals is also studied.**

I. Introduction

The demand for low-cost and small sized terminals have prohibited practical implementations of multiple antennas on the consumer side. Cooperative diversity has been proposed for systems with multiple user terminals as an attempt to realize the spatial diversity gain similar to that of a MIMO system [1]. The mobile terminals share their antennas with other users to create a virtual antenna array and provide spatial diversity. To fully utilize the increased capacity promised by cooperative diversity systems, judicious code design is necessary. Due to the similarity between cooperative diversity and MIMO systems, space-time codes have been proposed as a possible solution [2]. These codes utilize both the spatial and time domains to introduce correlation between signals transmitted from the different antennas at different time slots. An important class of space-time code is the space-time block codes (STBC). In a cooperative diversity system, space-time coding can also be used to take advantage of its MIMO-like properties to obtain spatial diversity, coding gains, and higher spectral efficiency.

To retransmit the received signal to the destination, relays can choose from two relay schemes: amplify-and-forward (AF) and decode-and-forward (DF) [3]. Using an AF relay scheme, the relays generate the space-time codewords using the received signal, and transmit the space-time codewords to the destination at a predetermined energy level. When using a DF relay scheme, the received signal is first decoded, and then the decoded symbols are used to generate the space-time codewords to be transmitted. Depending on the type of relay scheme employed, there are two types of errors that can be introduced into the system. In AF systems, the signal received during the broadcast interval is amplified by the relays before being retransmitted. Since no other processing is performed, noise from the received signal is also amplified and transmitted to the destination. In DF systems, relays decode the received signals prior to retransmission. Since there is a non-zero probability that the received signal will be decoded incorrectly, incorrect space-time codewords can be formed at relays where decoding errors have occurred. In [4] it is shown that the capacity of AF systems is higher than that of DF systems in most SNR range of interest; and therefore, in our work we focus only on cooperative systems using AF relay scheme.

In this work, we propose a new coding scheme for AF cooperative diversity systems based on linear dispersion (LD) codes [5]. Using LD codes, the energy of the transmitted symbols is spread out along the temporal and spatial dimensions. In our proposed scheme, we take a two-layered approach to design a LD code for cooperative diversity systems. The first layer of the LD code is generated at the source by Q r-QAM symbols with the first set of dispersion vectors. The second layer of LD code is then generated by the cooperative nodes with the second set of dispersion matrices using the received signals. The power allocated to each node and communication link is also an important design parameter in the performance of the cooperative diversity system [6]. For cooperative diversity systems using LD code, an explicit analytical expression for the block error rate (BLER) does not exist, and therefore deterministic optimization techniques cannot be used. Here we employ a simulation-based stochastic approximation technique together with gradient estimation [7] to jointly optimize the LD code and power allocation with respect to the BLER. This method is useful when the objective function and the gradient cannot be evaluated analytically but can be estimated.

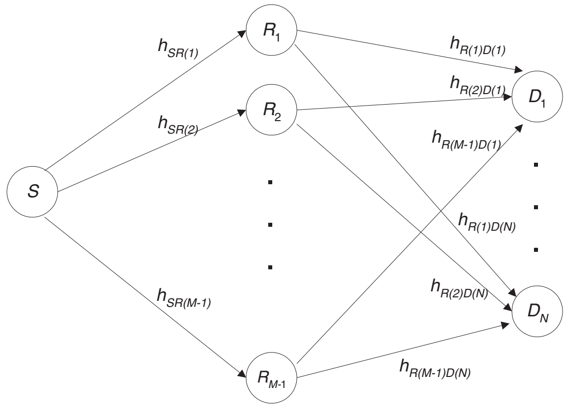

Fig. 1. Cooperative diversity system.

The remainder of the paper is organized as follows. In Section 2, we describe the system model for cooperative diversity systems using a two-layered LD code. In Section 3, we employ simulation-based stochastic approximation and gradient estimation techniques to optimize the LD code and power allocation with respect to the BLER. In Section 4 we present simulation results to compare the performance of the optimized LD code and power allocation with other space-time code designs. Finally Section 4 contains the conclusions.

II. System Model

Unlike MIMO systems, where the transmit antennas all are part of a single antenna-array, transmit antennas in a cooperative diversity system can belong to several independent terminals where only the source has perfect knowledge of the symbols to be transmitted. To be able to use the cooperative nodes as relays to transmit information jointly to the destination, the information first need to be distributed to the cooperative nodes.

Consider the cooperative network illustrated in Fig. 1 consisting of 1 source terminal, $(M-1)$ relay terminals, and 1 destination terminal, denoted as S, R_m, $m = 1, ..., M-1$, and D respectively. The source has one transmit antenna, each relay has one antenna, and the destination has N receive antennas, with the n-th receive antenna denoted as D_n, $n = 1, ..., N$. The system thus consists of M transmit antennas and N receive antennas, which we will denote as an (M, N) system. The source-to-relay, source-to-destination, and relay-to-destination channels are assumed to be mutually independent block-fading channels with arbitrary fading statistics, and are denoted as h_{SR_m}, h_{SD_n}, and $h_{R_mD_n}$, respectively.

In our proposed scheme, each transmission frame consists of two intervals:

1) Broadcast interval: source broadcasts information using first-layer LD code to cooperating relays and destination.
2) Cooperative interval: source and relays transmit with second-layer LD code to the destination.

Unlike for MIMO systems, a code design for cooperative diversity system needs to take into consideration the broadcast channels between the source and relays. Moreover, the energy allocated in the broadcast interval to transmit the first-layer LD code to the relays has significant impact on system performance. The total energy consumption E_0 is fixed for each transmission frame. The optimal distribution of energy between the two intervals depends on the space-time code, the statistics of the channels between source, relays, and destination, and the physical distances between the terminals. If insufficient energy is allocated to the broadcast interval, the first-layer LD code received by the relays during the broadcast interval can become so corrupted that the performance of the overall system is degraded despite having more energy allocated to the source-to-destination and relay-to-destination links in the cooperative interval. On the other hand, if too much energy is assigned to the broadcast interval for reliable transmission of the first-layer LD code to the relays, the lack of energy during the cooperative interval can still degrade the overall system performance. It is then clear that energy allocation between the two intervals should also be an optimization parameter.

A. Transmission Scheme

Let us define $s_1, s_2, ..., s_Q$ as the Q different r-QAM symbols that the source terminal wishes to transmit to the destination terminal, where $s_q = \alpha_q + j\beta_q$, $q = 1, ..., Q$, and $\mathbb{E}\{|s_q^2|\} = 1$. For a cooperative diversity system with one source antenna, $M-1$ relays, and N destination antennas, we can construct the following cooperative transmission scheme:

1) Source forms τ linearly combined symbols $\boldsymbol{k} = [k_1, ..., k_\tau]^T$ using Q r-QAM symbols, $s_1, ..., s_Q$, and the first-layer dispersion $\tau \times 1$ vectors $\boldsymbol{c}_q$, $\boldsymbol{d}_q$, $q = 1, \ldots, Q$.
2) Source transmits $k_1, ..., k_\tau$ to the relays and destination during τ consecutive symbol intervals.
3) Source and relays form the second-layer LD codeword using their received signals corresponding to $k_1, ..., k_\tau$ and with dispersion matrices $\boldsymbol{A}_t$, $\boldsymbol{B}_t$, $t = 1, ..., \tau$, where $\boldsymbol{A}_t$, $\boldsymbol{B}_t$ have dimensions $(T-\tau) \times M$, and each relay uses one column of the dispersion matrices.

Notice that τ, the number of linearly combined symbols transmitted during the broadcast interval, is also the length of the broadcast interval. Since the total length of the frame is fixed at T, the choice of τ will determine the size of both the broadcast and cooperative intervals. By choosing $\tau > T/2$, we are devoting more resources to ensure that information received by the cooperative terminals is less error prone. Intuitively, this is done when the the source sees poorer channels to the relays compared to the relay-to-destination channels. Conversely, we can make $\tau < T/2$ when the relays see poorer channels compared to the source-to-relay channels. Thus it is clear that τ is also a design variable that needs to be optimized.

B. Energy Constraints

The energy allocated to the broadcast interval, E_1, and cooperative interval, E_2, is constrained by the fixed total energy constraint for one transmission frame E_0 where

$$E_0 = E_1 + E_2, \quad E_1 > 0, \quad E_2 > 0, \tag{1}$$

and we can write (1) as a function of an angular coordinate

$$E_1 = E_0 \cos^2 \alpha, \quad E_2 = E_0 \sin^2 \alpha, \quad \alpha \in (0, \pi). \tag{2}$$

For an (M, N) cooperative diversity system, let us denote the distance between source and relay R_m as d_{SR_m}, $m = 1, ..., M-1$, distance between source and destination as d_{SD}, and distance between relay R_m and destination as d_{R_mD}. Denote $\rho_{SD,1}$, $\rho_{SD,2}$, ρ_{SR_m}, and ρ_{R_mD} as SNR between the source and destination during the broadcast and cooperative phase, SNR between source and relay R_m, and SNR between relay R_m and destination, respectively. By assuming that the received noise has unit variance, and incorporating path loss into our model it follows that:

$$\begin{aligned} \rho_{SD,1} &= \frac{E_1}{\tau}\left(\frac{1}{d_{SD}}\right)^\nu, & \rho_{SR_m} &= \frac{E_1}{\tau}\left(\frac{1}{d_{SR_m}}\right)^\nu, \\ \rho_{SD,2} &= \frac{1}{M}\frac{E_2}{T-\tau}\left(\frac{1}{d_{SD}}\right)^\nu, & \rho_{R_mD} &= \frac{1}{M}\frac{E_2}{T-\tau}\left(\frac{1}{d_{R_mD}}\right)^\nu, \end{aligned} \tag{3}$$

where ν is the path loss exponent. The factor of $1/M$ divides the energy allocated in the cooperative interval evenly among the M transmitting terminals. In the following analysis, unless stated otherwise, we assume that $\nu = 4$ for urban environment [8].

C. Broadcast interval

For a given τ, let us denote the $\tau \times 1$ complex-valued linear dispersion vectors for the source-to-relay transmission as $\boldsymbol{c}_q = [c_{1q}, \ldots, c_{\tau q}]^T$ and $\boldsymbol{d}_q = [d_{1q}, \ldots, d_{\tau q}]^T$, $q = 1, \ldots, Q$. Recall that $s_q = \alpha_q + j\beta_q$, the $\tau \times 1$ linearly combined symbol vector to be transmitted is then

$$\boldsymbol{k} = \sum_{q=1}^{Q} \left(\alpha_q \boldsymbol{c}_q + j\beta_q \boldsymbol{d}_q\right), \quad q = 1, ..., Q. \tag{4}$$

The energy constraint is $\mathbb{E}\{\boldsymbol{k}^H\boldsymbol{k}\} \le \tau$. Using the fact that α_q, β_q are i.i.d. with zero mean and variance $\frac{1}{2}$, we get the following constraint

$$\sum_{q=1}^{Q} \left(\boldsymbol{c}_q^H \boldsymbol{c}_q + \boldsymbol{d}_q^H \boldsymbol{d}_q\right) \le 2\tau. \tag{5}$$

Denote h_{SR_m}, $m = 1, \ldots, M-1$ as the fading channel coefficients for the source-to-relay channels. Let $\boldsymbol{r}_{R_m}$ be the $\tau \times 1$ received signal vector at relay R_m after passing through a matched filter and normalizing by $|h_{SR_m}|$. The received signal at relay R_m is then given by

$$\boldsymbol{r}_{R_m} = |h_{SR_m}|\sqrt{\rho_{SR_m}}\boldsymbol{k} + \boldsymbol{n}_{R_m}, \tag{6}$$

with $\boldsymbol{n}_{R_m} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{I})$, $m = 1, \ldots, M-1$. In the broadcast interval, the destination antenna D_n also receives the transmission from the source, and the $\tau \times 1$ received signal vector at antenna D_n is given as

$$\boldsymbol{r}_{D_n,B} = h_{SD_n}\sqrt{\rho_{SD,1}}\boldsymbol{k} + \boldsymbol{n}_{D_n}, \tag{7}$$

with $\boldsymbol{n}_{D_n} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{I})$, $n = 1, \ldots, N$. Denoting $k_t = \tilde{\alpha}_t + j\tilde{\beta}_t$, $t = 1, \ldots, \tau$, then from (4), we have the following relationship

$$\eta = \boldsymbol{G}\boldsymbol{x} \tag{8}$$

where $\eta = [\tilde{\alpha}_1, ..., \tilde{\alpha}_\tau, \tilde{\beta}_1, ..., \tilde{\beta}_\tau]^T$ is a $(2\tau \times 1)$ vector, $\boldsymbol{x} = [\alpha_1, \beta_1, ..., \alpha_Q, \beta_Q]^T$ is a $(2Q \times 1)$ vector and

$$\boldsymbol{G} = \begin{bmatrix} \Re\{c_{11}\} & -\Im\{d_{11}\} & \Re\{c_{12}\} & \cdots & \Re\{c_{1Q}\} & -\Im\{d_{1Q}\} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \Re\{c_{\tau 1}\} & -\Im\{d_{\tau 1}\} & \Re\{c_{\tau 2}\} & \cdots & \Re\{c_{\tau Q}\} & -\Im\{d_{\tau Q}\} \\ \Im\{c_{11}\} & \Re\{d_{11}\} & \Im\{c_{12}\} & \cdots & \Im\{c_{1Q}\} & \Re\{d_{1Q}\} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \Im\{c_{\tau 1}\} & \Re\{d_{\tau 1}\} & \Im\{c_{\tau 2}\} & \cdots & \Im\{c_{\tau Q}\} & \Re\{d_{\tau Q}\} \end{bmatrix} \tag{9}$$

We further denote $\boldsymbol{y}_1$ as the real-valued received signal vector during the broadcast interval at the relays, i.e.,

$$\boldsymbol{y}_1 \triangleq \left[\Re\{\boldsymbol{r}_{R_1}^T\}, \Im\{\boldsymbol{r}_{R_1}^T\}, \cdots, \Re\{\boldsymbol{r}_{R_{M-1}}^T\}, \Im\{\boldsymbol{r}_{R_{M-1}}^T\}\right]^T_{2(M-1)\tau\times 1}. \tag{10}$$

Using (6) and (8), the received signal $\boldsymbol{y}_1$ at the relays during the broadcast interval can now be written as

$$\boldsymbol{y}_1 = \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x} + \boldsymbol{n}, \quad \boldsymbol{n} \sim \mathcal{N}\left(\boldsymbol{0}, \frac{1}{2}\boldsymbol{I}\right), \tag{11}$$

with the real-valued equivalent channel matrix $\boldsymbol{H}_1$ given as

$$\boldsymbol{H}_1 \triangleq \begin{bmatrix} |h_{SR_1}|\sqrt{\rho_{SR_1}}\boldsymbol{I}_{2\tau\times 2\tau} \\ \vdots \\ |h_{SR_{M-1}}|\sqrt{\rho_{SR_{M-1}}}\boldsymbol{I}_{2\tau\times 2\tau} \end{bmatrix}_{2(M-1)\tau\times 2\tau}. \tag{12}$$

D. Cooperative Interval

In the cooperative interval, source and relays construct a new LD codeword using the received first-layer LD codeword generated during the broadcast interval. We first need to normalize the energy of the received signal during the broadcast interval. From (6) we have

$$\mathbb{E}\{\boldsymbol{r}_m^H\boldsymbol{r}_m\} = |h_{SR_m}|^2\rho_{SR_m}\tau + \tau. \tag{13}$$

Thus before performing the linear combination on the received signals at the relays using the second-layer dispersion matrices, we multiply by the normalization constant

$$\gamma_{R_m} \triangleq \sqrt{\frac{\tau}{|h_{SR_m}|^2\rho_{SR_m}\tau + \tau}}, \quad m = 1, \ldots, M-1, \tag{14}$$

such that

$$\gamma_{R_m}^2\mathbb{E}\{\boldsymbol{r}_{R_m}^H\boldsymbol{r}_{R_m}\} = \tau. \tag{15}$$

For the cooperative interval, we use a set of dispersion matrices $\{\boldsymbol{A}_t, \boldsymbol{B}_t\}_{t=1}^{\tau}$, with dimension $(T-\tau)\times M$. From (15), the τ received symbols to be linearly combined have total energy of τ, thus we normalize the second-layer dispersion matrices as

$$\sum_{t=1}^{\tau}\operatorname{tr}\left(\boldsymbol{A}_t^H\boldsymbol{A}_t + \boldsymbol{B}_t^H\boldsymbol{B}_t\right) \le 2M(T-\tau). \tag{16}$$

For transmission in the cooperative interval, the source employs the first column of the the LD matrices and transmit the following $(T-\tau)\times 1$ signal vector

$$\boldsymbol{x}_S = \sum_{t=1}^{\tau}\Re\{k_t\}\boldsymbol{a}_{1,t} + j\Im\{k_t\}\boldsymbol{b}_{1,t}, \tag{17}$$

The relay R_m will use the $(m+1)$-th column of the LD matrices and transmit the following $(T-\tau)\times 1$ signal vector

$$\boldsymbol{x}_{R_m} = \gamma_{R_m}\sum_{t=1}^{\tau}\left(\Re\{r_{R_m,t}\}\boldsymbol{a}_{m+1,t} + j\Im\{r_{R_m,t}\}\boldsymbol{b}_{m+1,t}\right), \tag{18}$$

for $m = 1, \ldots, M-1$, and where $\boldsymbol{a}_{m,t}$ and $\boldsymbol{b}_{m,t}$ are the m-th column of the dispersion matrices $\boldsymbol{A}_t$ and $\boldsymbol{B}_t$, respectively. Denote $h_{R_mD_n}$ and h_{SD_n} as the relay-to-destination and source-to-destination channel coefficients. Then the $(T-\tau)\times 1$ received signal vector at the destination antenna D_n during the cooperative interval is

$$\boldsymbol{r}_{D_n,C} = h_{SD_n}\sqrt{\rho_{SD,2}}\boldsymbol{x}_S + \sum_{m=1}^{M-1}h_{R_mD_n}\sqrt{\rho_{R_mD_n}}\boldsymbol{x}_{R_m} + \boldsymbol{v}_{D_n}, \tag{19}$$

with $\boldsymbol{v}_{D_n} \sim \mathcal{N}_{\mathcal{C}}(\boldsymbol{0}, \boldsymbol{I})$. Define $\boldsymbol{y}_2$ as the real-valued received signal vector at the destination in both broadcast and cooperative intervals

$$\begin{aligned}\boldsymbol{y}_2 \triangleq & \left[\Re\{\boldsymbol{r}_{D_1,B}^T\}, \Im\{\boldsymbol{r}_{D_1,B}^T\}, \cdots, \Re\{\boldsymbol{r}_{D_N,B}^T\}, \Im\{\boldsymbol{r}_{D_N,B}^T\}\right. \\ & \left.\Re\{\boldsymbol{r}_{D_1,C}^T\}, \Im\{\boldsymbol{r}_{D_1,C}^T\}, \cdots, \Re\{\boldsymbol{r}_{D_N,C}^T\}, \Im\{\boldsymbol{r}_{D_N,C}^T\}\right]^T_{2NT\times 1}\end{aligned} \tag{20}$$

Let us further define

$$\mathcal{A}_{m,t} \triangleq \begin{bmatrix} \Re\{\boldsymbol{a}_{m,t}\} & -\Im\{\boldsymbol{a}_{m,t}\} \\ \Im\{\boldsymbol{a}_{m,t}\} & \Re\{\boldsymbol{a}_{m,t}\} \end{bmatrix}_{2(T-\tau)\times 2},$$

$$\mathcal{B}_{m,t} \triangleq \begin{bmatrix} -\Im\{\boldsymbol{b}_{m,t}\} & -\Re\{\boldsymbol{b}_{m,t}\} \\ \Re\{\boldsymbol{b}_{m,t}\} & -\Im\{\boldsymbol{b}_{m,t}\} \end{bmatrix}_{2(T-\tau)\times 2}, \tag{21}$$

and

$$\underline{\boldsymbol{h}}_{SD_n} \triangleq \sqrt{\rho_{SD,2}}\begin{bmatrix} \Re\{h_{SD_n}\} \\ \Im\{h_{SD_n}\} \end{bmatrix},$$

$$\underline{\boldsymbol{h}}_{R_mD_n} \triangleq \gamma_{R_m}\sqrt{\rho_{RD}}\begin{bmatrix} \Re\{h_{R_mD_n}\} \\ \Im\{h_{R_mD_n}\} \end{bmatrix}. \tag{22}$$

From (7), (10), and (19) we can write $\boldsymbol{y}_2$ as

$$\boldsymbol{y}_2 = \boldsymbol{H}_2\boldsymbol{G}\boldsymbol{x} + \boldsymbol{H}_3\boldsymbol{y}_1 + \boldsymbol{u}, \quad \boldsymbol{u} \sim \mathcal{N}\left(\boldsymbol{0}, \frac{1}{2}\boldsymbol{I}\right), \tag{23}$$

where $\boldsymbol{H}_2$ is the real-valued equivalent channel matrix for the received signal component at the destination from the source during the broadcast and cooperative intervals

$$\boldsymbol{H}_2 \triangleq \begin{bmatrix} \boldsymbol{P} \\ \mathcal{A}_{1,1}\underline{\boldsymbol{h}}_{SD_1}, \cdots, \mathcal{A}_{1,\tau}\underline{\boldsymbol{h}}_{SD_1}, \mathcal{B}_{1,1}\underline{\boldsymbol{h}}_{SD_1}, \cdots, \mathcal{B}_{1,\tau}\underline{\boldsymbol{h}}_{SD_1} \\ \vdots \quad \vdots \\ \mathcal{A}_{1,1}\underline{\boldsymbol{h}}_{SD_N}, \cdots, \mathcal{A}_{1,\tau}\underline{\boldsymbol{h}}_{SD_N}, \mathcal{B}_{1,1}\underline{\boldsymbol{h}}_{SD_N}, \cdots, \mathcal{B}_{1,\tau}\underline{\boldsymbol{h}}_{SD_N} \end{bmatrix} \tag{24}$$

where $\boldsymbol{H}_2$ has dimensions $2NT\times 2\tau$, and $\boldsymbol{H}_3$ is the real-valued equivalent channel matrix for the received signal component at the destination from the relays during the cooperative interval

$$\boldsymbol{H}_3 \triangleq \begin{bmatrix} \boldsymbol{0}_{2N\tau\times 2(M-1)\tau} \\ \mathcal{A}_{2,1}\underline{\boldsymbol{h}}_{R_1D_1}, \cdots, \mathcal{A}_{2,\tau}\underline{\boldsymbol{h}}_{R_1D_1}, \mathcal{B}_{2,1}\underline{\boldsymbol{h}}_{R_1D_1}, \cdots, \mathcal{B}_{M,\tau}\underline{\boldsymbol{h}}_{R_{M-1}D_1} \\ \mathcal{A}_{2,1}\underline{\boldsymbol{h}}_{R_1D_2}, \cdots, \mathcal{A}_{2,\tau}\underline{\boldsymbol{h}}_{R_1D_2}, \mathcal{B}_{2,1}\underline{\boldsymbol{h}}_{R_1D_2}, \cdots, \mathcal{B}_{M,\tau}\underline{\boldsymbol{h}}_{R_{M-1}D_2} \\ \vdots, \vdots, \vdots, \vdots, \ddots, \vdots \\ \mathcal{A}_{2,1}\underline{\boldsymbol{h}}_{R_1D_N}, \cdots, \mathcal{A}_{2,\tau}\underline{\boldsymbol{h}}_{R_1D_N}, \mathcal{B}_{2,1}\underline{\boldsymbol{h}}_{R_1D_N}, \cdots, \mathcal{B}_{M,\tau}\underline{\boldsymbol{h}}_{R_{M-1}D_N} \end{bmatrix}, \tag{25}$$

where $\boldsymbol{H}_3$ has dimensions $(2NT\times 2(M-1)\tau)$; and with

$$\boldsymbol{P} \triangleq \sqrt{\rho_{SD,1}}\begin{bmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_2 & \cdots & \boldsymbol{P}_N \end{bmatrix}^T_{2N\tau\times 2\tau}, \tag{26}$$

where

$$\boldsymbol{P}_n \triangleq \begin{bmatrix} \Re\{h_{SD_n}\}\boldsymbol{I}_{\tau\times\tau} & \Im\{h_{SD_n}\}\boldsymbol{I}_{\tau\times\tau} \\ -\Im\{h_{SD_n}\}\boldsymbol{I}_{\tau\times\tau} & \Re\{h_{SD_n}\}\boldsymbol{I}_{\tau\times\tau} \end{bmatrix}_{2\tau\times 2\tau}. \tag{27}$$

In order to perform data detection at the destination, we need to write $\boldsymbol{y}_2$ in terms of $\boldsymbol{x}$, the real-valued transmitted symbol vector. Substituting (11) into (23), we have

$$\begin{aligned}\boldsymbol{y}_2 &= \boldsymbol{H}_2\boldsymbol{G}\boldsymbol{x} + \boldsymbol{H}_3\left(\boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x} + \boldsymbol{n}\right) + \boldsymbol{u} \\ &= \left(\boldsymbol{H}_2 + \boldsymbol{H}_3\boldsymbol{H}_1\right)\boldsymbol{G}\boldsymbol{x} + \left(\boldsymbol{H}_3\boldsymbol{n} + \boldsymbol{u}\right).\end{aligned} \tag{28}$$

In (28) the effective total noise is colored, owing to the noise amplification and recombination at the relays, with covariance $\boldsymbol{\Sigma} = \frac{1}{2}\left(\boldsymbol{H}_3\boldsymbol{H}_3^T + \boldsymbol{I}\right)$. In order to perform detection, we need to whiten the noise first. That is, define

$$\boldsymbol{z} \triangleq \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{y}_2 = \boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{H}_2 + \boldsymbol{H}_3\boldsymbol{H}_1\right)\boldsymbol{G}\boldsymbol{x} + \bar{\boldsymbol{u}}, \quad \bar{\boldsymbol{u}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}). \tag{29}$$

We can then employ the sphere decoder [9] to perform ML detection on $\boldsymbol{z}$ to obtain $\hat{\boldsymbol{x}}$, the ML estimate of $\boldsymbol{x}$.

III. Optimization of Two-layer Cooperative LD Code

A. Stochastic Approximation and Gradient Estimation

In this section we develop an algorithm to find the two-layered dispersion matrices and the energy allocation to minimize the BLER for a cooperative diversity system in an arbitrary fading scenario. Since the exact analytical expression of the average BLER for an arbitrary set of dispersion matrices and arbitrary fading statistics does not exist, we resort to stochastic gradient algorithms to optimize the average BLER performance with respect to the dispersion matrices and the energy allocation. Here we employ an optimization scheme based on the Robbins-Monro (R-M) [10] algorithm and use the score-function algorithm for gradient estimation. The R-M algorithm takes the recursive form

$$\psi_{k+1} = \psi_k - a_k \hat{g}(\psi_k), \tag{30}$$

where ψ_k is the estimated parameter value at iteration k, $\hat{g}(\psi_k)$ is the gradient estimate of the objective function at ψ_k and $\{a_k\}$ is a decreasing step size sequence of positive numbers such that

$$\sum_{k=1}^{\infty} a_k = \infty, \quad \sum_{k=1}^{\infty} a_k < \infty. \tag{31}$$

By choosing $a_k = a/k$ where a is a positive scalar, the above stochastic gradient algorithm will converge with probability 1 to a local optimum.

Consider an (M, N) cooperative diversity system with Q symbols to be transmitted, broadcast interval length τ, and total transmission interval of T. Let us define the real-valued channel vectors corresponding respectively to channels from source to relay, and channels from relay to destination and source to destination

$$\begin{aligned} \boldsymbol{h}_1 &\triangleq \left[\Re\{h_{SR_1}\}, \Im\{h_{SR_1}\}, \cdots, \Re\{h_{SR_{M-1}}\}, \Im\{h_{SR_{M-1}}\}\right]^T, \\ \boldsymbol{h}_2 &\triangleq \Big[\Re\{h_{R_1D_1}\}, \Im\{h_{R_1D_1}\}, \cdots, \Im\{h_{R_{M-1}D_N}\}, \Im\{h_{R_{M-1}D_N}\} \\ &\quad \Re\{h_{SD_1}\}, \Im\{h_{SD_1}\}, \cdots, \Re\{h_{SD_N}\}, \Im\{h_{SD_N}\}\Big]^T. \end{aligned} \tag{32}$$

For the cooperative LD code design problem, the optimization parameter set $\boldsymbol{\theta}$ consists of the broadcast dispersion vectors, the cooperative dispersion matrices, and the angular coordinate defining the energy allocation (2), i.e.,

$$\boldsymbol{\theta} \triangleq \Big\{ \{\boldsymbol{c}_q, \boldsymbol{d}_q, q = 1, \ldots, Q\}, \quad \{\boldsymbol{A}_t, \boldsymbol{B}_t, t = 1, \ldots, \tau\}, \quad \alpha\Big\}, \tag{33}$$

with constraints (5) and (16). Define the empirical BLER as $\gamma(\boldsymbol{z}, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})$ for the given sets of noise-whitened receive signal vector $\boldsymbol{z}$, information symbol vector $\boldsymbol{x}$, channel realizations $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$, and the given parameter $\boldsymbol{\theta}$. The empirical BLER is then given by an indicator function

$$\gamma(\boldsymbol{z}, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta}) \triangleq \mathbb{I}\left(\hat{\boldsymbol{x}} \neq \boldsymbol{x} \mid \boldsymbol{z}, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta}\right), \tag{34}$$

where $\hat{\boldsymbol{x}}$ denotes the decoded symbol vector. Recalling from (29) that $\boldsymbol{z} = \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{y}_2$, we can thus write the empirical BLER as a function of $\boldsymbol{y}_2$. For given $\boldsymbol{\theta}$, the average BLER is then

$$\Upsilon(\boldsymbol{\theta}) \triangleq \mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{h}_1}\mathbb{E}_{\boldsymbol{h}_2}\mathbb{E}_{\boldsymbol{y}_1,\boldsymbol{y}_2|\boldsymbol{x},\boldsymbol{h}_1,\boldsymbol{h}_2}\{\gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\}. \tag{35}$$

We want to solve the following optimization problem

$$\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \Upsilon(\boldsymbol{\theta}), \tag{36}$$

with the constraint set given by

$$\begin{aligned} \boldsymbol{\Theta} &\triangleq \Big\{ \sum_{q=1}^{Q} \mathrm{tr}\left(\boldsymbol{c}_q^H \boldsymbol{c}_q + \boldsymbol{d}_q^H \boldsymbol{d}_q\right) \leq 2\tau, \\ &\quad \sum_{t=1}^{\tau} \mathrm{tr}\left(\boldsymbol{A}_t^H \boldsymbol{A}_t + \boldsymbol{B}_t^H \boldsymbol{B}_t\right) \leq 2M(T-\tau)\Big\}. \end{aligned} \tag{37}$$

In (35), we have

$$\begin{aligned} &\mathbb{E}_{\boldsymbol{y}_1,\boldsymbol{y}_2|\boldsymbol{x},\boldsymbol{h}_1,\boldsymbol{h}_2}\{\gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\} \\ =& \int\int \gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\, p(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\, d\boldsymbol{y}_1 d\boldsymbol{y}_2 \\ =& \int\int \gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\, p(\boldsymbol{y}_1 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{\theta})\, p(\boldsymbol{y}_2 \mid \boldsymbol{y}_1, \boldsymbol{x}, \boldsymbol{h}_2, \boldsymbol{\theta})\, d\boldsymbol{y}_1 d\boldsymbol{y}_2 \end{aligned} \tag{38}$$

From (11) and (23), it follows that $p(\boldsymbol{y}_1 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{\theta})$ and $p(\boldsymbol{y}_2 \mid \boldsymbol{y}_1\boldsymbol{x}, \boldsymbol{h}_2, \boldsymbol{\theta})$ are both white Gaussian pdf. Let us denote

$$\begin{aligned} \mathcal{P}_1 &\triangleq p(\boldsymbol{y}_1 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{\theta}) \\ &= \frac{1}{\pi^{(M-1)\tau}} \exp\left[-(\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x})^T (\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x})\right], \end{aligned} \tag{39}$$

$$\begin{aligned} \mathcal{P}_2 &\triangleq p(\boldsymbol{y}_2 \mid \boldsymbol{y}_1, \boldsymbol{x}, \boldsymbol{h}_2, \boldsymbol{\theta}) \\ &= \frac{1}{\pi^{NT}} \exp\left[-(\boldsymbol{y}_2 - \boldsymbol{H}_2\boldsymbol{G}\boldsymbol{x} - \boldsymbol{H}_3\boldsymbol{y}_1)^T (\boldsymbol{y}_2 - \boldsymbol{H}_2\boldsymbol{G}\boldsymbol{x} - \boldsymbol{H}_3\boldsymbol{y}_1)\right] \end{aligned} \tag{40}$$

Using (35), the gradient of the average BLER with respect to $\boldsymbol{\theta}$ is then given by

$$\begin{aligned} \nabla_{\boldsymbol{\theta}}\Upsilon(\boldsymbol{\theta}) =& \mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{h}_1}\mathbb{E}_{\boldsymbol{h}_2}\int\int \left[\nabla_{\boldsymbol{\theta}}\gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\right]\mathcal{P}_1\mathcal{P}_2 \\ &+\gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\left[(\nabla_{\boldsymbol{\theta}}\mathcal{P}_1)\mathcal{P}_2 + \mathcal{P}_1(\nabla_{\boldsymbol{\theta}}\mathcal{P}_2)\right] d\boldsymbol{y}_1 d\boldsymbol{y}_2 \end{aligned} \tag{41}$$

For ML detection, it can be proven that

$$\mathbb{E}_{\boldsymbol{x}}\int\int \nabla_{\boldsymbol{\theta}}\gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\,\mathcal{P}_1\mathcal{P}_2 d\boldsymbol{y}_1 d\boldsymbol{y}_2 = 0. \tag{42}$$

(owing to space limitation we omit a rigorous proof). Then (41) can be written as

$$\begin{aligned} &\nabla_{\boldsymbol{\theta}}\Upsilon(\boldsymbol{\theta}) \\ =& \mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{h}_1}\mathbb{E}_{\boldsymbol{h}_2}\int\int \gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})[(\nabla_{\boldsymbol{\theta}}\mathcal{P}_1)\mathcal{P}_2 + \mathcal{P}_1(\nabla_{\boldsymbol{\theta}}\mathcal{P}_2)]d\boldsymbol{y}_1 d\boldsymbol{y}_2 \\ =& \mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{h}_1}\mathbb{E}_{\boldsymbol{h}_2}\int\int \gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}p(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})d\boldsymbol{y}_1 d\boldsymbol{y}_2 \\ =& \mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{h}_1}\mathbb{E}_{\boldsymbol{h}_2}\int\int \gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta}) \times \\ &\frac{\nabla_{\boldsymbol{\theta}}p(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})}{p(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})}p(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})d\boldsymbol{y}_1 d\boldsymbol{y}_2 \\ =& \mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{h}_1}\mathbb{E}_{\boldsymbol{h}_2}\mathbb{E}_{\boldsymbol{y}_1,\boldsymbol{y}_2|\boldsymbol{x},\boldsymbol{h}_1,\boldsymbol{h}_2} \\ &\Big\{\gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}\log p(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\Big\} \\ =& \mathbb{E}_{\boldsymbol{x}}\mathbb{E}_{\boldsymbol{h}_1}\mathbb{E}_{\boldsymbol{h}_2}\mathbb{E}_{\boldsymbol{y}_1,\boldsymbol{y}_2|\boldsymbol{x},\boldsymbol{h}_1,\boldsymbol{h}_2} \\ &\Big\{\gamma(\boldsymbol{y}_2, \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta})\nabla_{\boldsymbol{\theta}}[\log\mathcal{P}_1 + \log\mathcal{P}_2]\Big\}. \end{aligned} \tag{43}$$

B. Simulation-based LD Code Optimization Algorithm

We now present the iterative simulation based algorithm to optimize the dispersion matrices and the energy allocation. The optimal value for the design variable τ is chosen by evaluating the following algorithm at different values of τ and selecting the one which gives the lowest BLER.

For a given τ in the k-th iteration, let $\boldsymbol{\theta}_k$ be the set

$$\begin{aligned} \boldsymbol{\theta}_k = & \Big\{\{\boldsymbol{c}_q^{(k)}, \boldsymbol{d}_q^{(k)}, q = 1, \ldots, Q\}, \\ & \{\boldsymbol{A}_t^{(k)}, \boldsymbol{B}_t^{(k)}, t = 1, \ldots, \tau\}, \quad \alpha^{(k)}\Big\}. \end{aligned} \tag{44}$$

Perform the following steps to update the parameter $\boldsymbol{\theta}_{k+1}$ for the next iteration:

1) Generate symbol and signal samples:
   1.1) Draw L symbol vectors $\boldsymbol{x}(1)$, $\boldsymbol{x}(2)$, $\ldots$, $\boldsymbol{x}(L)$, uniformly from the constellation set.
   1.2) Simulate L observations $\boldsymbol{y}_1(1)$, $\boldsymbol{y}_1(2)$, $\ldots$, $\boldsymbol{y}_1(L)$, where each $\boldsymbol{y}_1(\ell)$ is generated by [cf. (8), (11), (12)]

$$\boldsymbol{y}_1(\ell) = \boldsymbol{H}_1(\ell)\boldsymbol{G}(\ell)\boldsymbol{x}(\ell) + \boldsymbol{n}(\ell), \quad \ell = 1, 2, \ldots, L. \tag{45}$$

1.3) Simulate L observations $\boldsymbol{y}_2(1)$, $\boldsymbol{y}_2(2)$, ..., $\boldsymbol{y}_2(L)$, where each $\boldsymbol{y}_2(\ell)$ is generated by [cf. (21)-(27)]

$$\boldsymbol{y}_2(\ell) = \boldsymbol{H}_2(\ell)\boldsymbol{G}(\ell)\boldsymbol{x}(\ell) + \boldsymbol{H}_3(\ell)\boldsymbol{y}_1(\ell) + \boldsymbol{u}(\ell), \quad (46)$$

for $\ell = 1, 2, ..., L$.

1.4) Decode $\boldsymbol{x}(\ell)$ based on (29) and compute the empirical BLER $\gamma\Big(\boldsymbol{z}(\ell), \boldsymbol{x}(\ell), \boldsymbol{h}_1(\ell), \boldsymbol{h}_2(\ell), \boldsymbol{\theta}_k\Big)$.

2) Score function method for gradient estimation: Generate the estimate of (43)

$$\begin{aligned}\hat{\boldsymbol{g}}(\boldsymbol{\theta}_k) \quad = \quad & \frac{1}{L}\sum_{\ell=1}^{L} \gamma\Big(\boldsymbol{y}_2(\ell), \boldsymbol{x}(\ell), \boldsymbol{h}_1(\ell), \boldsymbol{h}_2(\ell), \boldsymbol{\theta}_k\Big) \\ & \times\Big\{\nabla_{\boldsymbol{\theta}}\Big[\log p\left(\boldsymbol{y}_1(\ell) \mid \boldsymbol{x}(\ell), \boldsymbol{h}_1(\ell), \boldsymbol{\theta}\right) \\ & + \log p\left(\boldsymbol{y}_2(\ell) \mid \boldsymbol{y}_1(\ell), \boldsymbol{x}(\ell), \boldsymbol{h}_2(\ell), \boldsymbol{\theta}\right)\Big]|_{\boldsymbol{\theta}=\boldsymbol{\theta}_k}\Big\} \quad (47)\end{aligned}$$

The expressions to compute the gradients required in (47) are given in Section III-C.

3) Update parameters: The parameters are updated as

$$\boldsymbol{\theta}_{k+1} = \Pi_{\boldsymbol{\Theta}}\left[\boldsymbol{\theta}_k - a_k\hat{\boldsymbol{g}}(\boldsymbol{\theta}_k)\right], \quad (48)$$

where $\Pi_{\boldsymbol{\Theta}}(\cdot)$ is a projection operator onto the set $\boldsymbol{\Theta}$. That is, $\Pi_{\boldsymbol{\Theta}}(\cdot)$ normalizes the first-layer dispersion vectors and the second-layer dispersion matrices such that the equalities in (5) and (16) are satisfied, respectively.

C. Gradient Calculations

Gradient for energy allocation: Recall from (43) that

$$\nabla_{\boldsymbol{\theta}} \log p\left(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta}\right) = \nabla_{\boldsymbol{\theta}}\left[\log \mathcal{P}_1 + \log \mathcal{P}_2\right]. \quad (49)$$

From (39) we have $\mathcal{P}_1$ distributed as multivariate Gaussian. Define

$$f_1 \triangleq -\left(\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x}\right)^T\left(\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x}\right). \quad (50)$$

Then from (2), (3), and (12), we have

$$\begin{aligned}\frac{\partial f_1\Big(\alpha\Big)}{\partial\alpha} \quad = \quad & -\frac{\partial\Big(\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x}\Big)^T}{\partial\alpha}\Big(\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x}\Big) \\ & -\Big(\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x}\Big)^T\frac{\partial\Big(\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x}\Big)}{\partial\alpha} \\ = \quad & -\frac{\sin\alpha}{\cos\alpha}\Big[\Big(\boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x}\Big)^T\Big(\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x}\Big) \\ & +\Big(\boldsymbol{y}_1 - \boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x}\Big)^T\Big(\boldsymbol{H}_1\boldsymbol{G}\boldsymbol{x}\Big)\Big]. \quad (51)\end{aligned}$$

In the cooperative interval we have from (40) that $\mathcal{P}_2$ is also multivariate Gaussian distributed. Define

$$f_2 \triangleq -\left(\boldsymbol{y}_2 - \boldsymbol{H}_2\boldsymbol{G}\boldsymbol{x} - \boldsymbol{H}_3\boldsymbol{y}_1\right)^T\left(\boldsymbol{y}_2 - \boldsymbol{H}_2\boldsymbol{G}\boldsymbol{x} - \boldsymbol{H}_3\boldsymbol{y}_1\right), \quad (52)$$

and similarly we can obtain an expression for $\frac{\partial f_1(\alpha)}{\partial\alpha}$ (we do not include here the derivation owing to space limitations). Therefore, the gradient of $\log p\left(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta}\right)$ with respect to α is given by

$$\nabla_{\boldsymbol{\theta}} \log p\left(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{\theta}\right)|_{\alpha} = \frac{\partial f_1\left(\alpha\right)}{\partial\alpha} + \frac{\partial f_2\left(\alpha\right)}{\partial\alpha}. \quad (53)$$

Gradient for first-layer dispersion matrices: Next we show how to obtain the gradient of f_1 and f_2 with respect to $\boldsymbol{c}_{R,q}$. The t-th entry of the gradient of $f_1\left(\boldsymbol{c}_{R,q}\right)$ is

$$\left[\frac{\partial f_1\left(\boldsymbol{c}_{R,q}\right)}{\partial\boldsymbol{c}_{R,q}}\right]_t = \lim_{\delta\to 0}\frac{f_1\left(\boldsymbol{c}_{R,q} + \delta\boldsymbol{\psi}_t\right) - f_1\left(\boldsymbol{c}_{R,q}\right)}{\delta}, \quad (54)$$

where $\boldsymbol{\psi}_t$ is a $\tau \times 1$ column vector with one at the t-th position and zero elsewhere else. Similarly we can obtain the gradients of f_1 with respect to $\boldsymbol{c}_{I,q}$, $\boldsymbol{d}_{R,q}$, and $\boldsymbol{d}_{I,q}$. Similar rationale is followed to compute the gradient of $\log\mathcal{P}_2$ with respect to $\boldsymbol{c}_{R,q}, \boldsymbol{c}_{I,q}$, $\boldsymbol{d}_{R,q}$, and $\boldsymbol{d}_{I,q}$. Again, we do not include the final expressions owing to space limitations.

Gradient for second-layer dispersion matrices: We can see that $p\left(\boldsymbol{y}_1 \mid \boldsymbol{x}, \boldsymbol{h}_1, \boldsymbol{\theta}\right)$ is independent of the second-layer dispersion matrices, therefore the gradient with respect to the second-layer dispersion matrices is zero, and we only have to evaluate the gradient $\nabla_{\boldsymbol{\theta}} \log p\left(\boldsymbol{y}_2 \mid \boldsymbol{y}_1, \boldsymbol{x}, \boldsymbol{h}_2, \boldsymbol{\theta}\right)$. That is, we need to evaluate the gradient of $\log p\left(\boldsymbol{y}_2 \mid \boldsymbol{y}_1, \boldsymbol{x}, \boldsymbol{h}_2, \boldsymbol{\theta}\right)$ with respect to $\boldsymbol{A}_{R,t}, \boldsymbol{A}_{I,t}, \boldsymbol{B}_{R,t}$ and $\boldsymbol{B}_{I,t}$. Note from (24) and (25) that $\boldsymbol{H}_2$ depends only upon the first columns of the second-layer dispersion matrices, while $\boldsymbol{H}_3$ is independent of the first columns of the second-layer dispersion matrices. We omit final expressions owing to space limitation.

IV. Simulation Results

In this section we present simulations to demonstrate the performance of cooperative diversity systems using the BLER optimized LD code obtained by the algorithm given in Section III-B. The two-layer LD code designed for cooperative diversity systems is optimized with respect to BLER for different system configurations. Their BLER performance is compared against other coding schemes for the range of source-to-destination SNR of interest. For all examples, the two-layer LD codes are designed at $\rho_{SD,1} = 15$dB. It will be seen that codes designed at a particular SNR also works well for a wide range of SNR.

Example 1 - Cooperative LD code for a (2,2) system: We consider a cooperative LD code in a system with one source terminal, one relay terminal, and a destination terminal with two receive antennas. Figure 2 compares the cooperative LD code with that of the Alamouti code for cooperative relays [11] in different fading channels. In this example, the terminals are all equal-distance and have i.i.d. fading to all other terminals. We can see that the optimized cooperative LD code outperforms the Alamouti code in a wide range of SNR values. The BLER performances of cooperative LD code and Alamouti code are compared in Rayleigh, Rician $K = 2$, Nakagami $m = 0.5$ and $m = 2$ fading channels. The Rician $K = 2$ and Nakagami $m = 0.5$ channels represents better-than-Rayleigh channels while Nakagami $m = 0.5$ channel represents worse-than-Rayleigh fading conditions. For the cooperative LD code we have chosen $T = 4$, $Q = 2$, and 16-QAM constellation for rate $R = 4$, and the dispersion matrices are optimized for the given channel statistics. We have found that the optimal length of the broadcast interval is at $\tau = 3$. This means that for this particular physical setup, more resources are needed in the broadcast interval for optimal BLER performance. For fair comparison we consider the same rate, using 16-QAM constellation for the Alamouti code, and energy is equally divided between the broadcast interval and cooperative interval by choosing $\alpha = \frac{\pi}{2}$.

In [5], it was determined that LD codes with good performance typically have

$$Q = \min(M, N) \times T. \quad (55)$$

In this example, we can see that this constraint is met for both layers of LD code. Since for the broadcast interval the transmitted symbol size is 3 and $\min(M, N) \times T = 3$. In the cooperative interval, we fall short of this limit since we are transmitting 3 linearly combined symbols while $\min(M, N) \times (T - \tau) = 2$. However, since the first-layer LD code is simply the linear combination of the two 16-QAM symbols, we can see that from the overall system point of view, we are transmitting two 16-QAM symbols over 1 time periods in a $(2, 2)$ system, thus satisfying the constraint. For the cooperative LD code, the constraint in (55) becomes

$$Q = \min(M, N) \times (T - \tau). \quad (56)$$

Example 2 - Cooperative LD code for a (4,2) system: We present a cooperative LD code for a system of one source, three relays, and one two-receive-antenna destination terminal. The terminals are

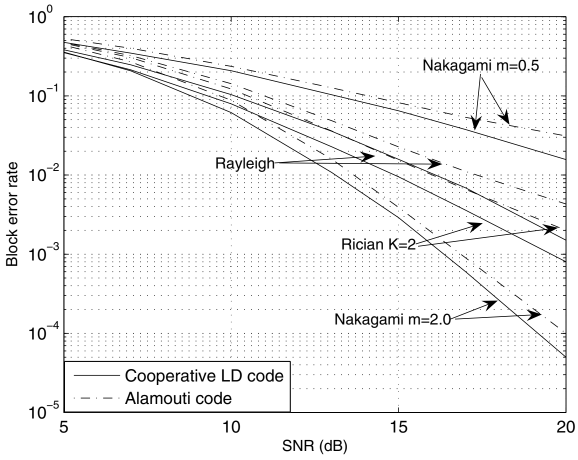

Fig. 2. Cooperative LD code vs Alamouti code for system with 1 source, 1 relay, and 2 destination receive antenna.

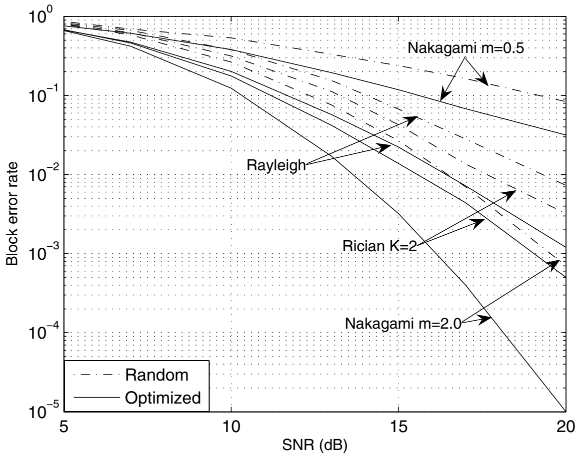

Fig. 3. Optimized cooperative LD code vs randomly chosen cooperative LD code (1 source, 3 relays, and 2 destination receive antennas).

again all equal-distance and have i.i.d. fading to all other terminals. In Figure 3 we compare the optimized cooperative LD code with randomly generated LD code in different fading enviroments. For both the optimized cooperative LD code and randomly generated code we have code length $T = 12$, broadcast interval length $\tau = 8$, and number of substreams transmitted $Q = 8$. When QPSK constellation is used we have rate $R = 16/12$. The energy is evenly divided between the broadcast and cooperative intervals for the randomly chosen code, while for the optimized cooperative LD code it is determined by the proposed algorithm. We can see that in general, a randomly chosen code does not provide good performance, and the performance of the cooperative LD code is shown to be significantly improved for all fading environments.

Example 3 - Cooperative LD code for systems of a different distance configuration: In this example, we show that a code designed for one physical distribution of terminals is not necessarily optimal for systems with different physical distributions. Figure 4 compares two cooperative LD codes optimized at different physical system setups. The cooperative LD code 1 is optimized for $d_{SR_m} = 2$, $d_{R_m D_n} = 1$, whereas cooperative LD code 2 is optimized for $d_{SR_m} = 1$, $d_{R_m D_n} = 2$. Both codes are then compared in a system where $d_{SR_m} = 2$, $d_{R_m D_n} = 1$. We can see from the figure that while both are optimized codes, the optimization algorithm is sensitive to the physical setup of the systems. In this case, at BLER of 10^{-2}, there is a 3 dB loss of performance by the mismatched code compared to the performance of the code that is matched to its environment.

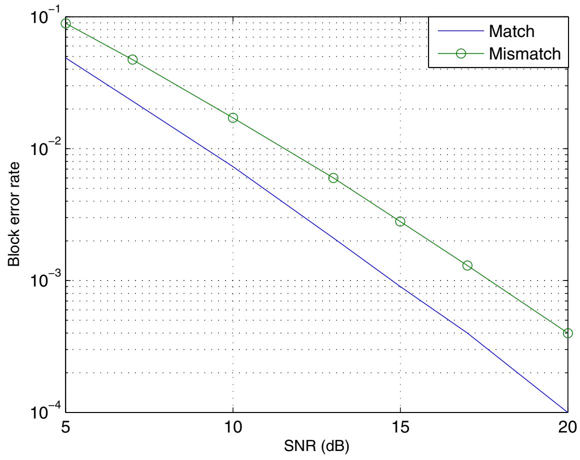

Fig. 4. Cooperative LD code optimized for different terminal distances.

V. Conclusions

In this paper we have proposed a two-layered linear dispersion code-based space-time coding scheme for cooperative diversity systems. We also proposed a simulation-based optimization algorithm to construct the optimal code based on BLER performance. The proposed code design algorithm can obtain optimal codes under arbitrary fading statistics and arbitrary system configuration in terms of both number of terminals and distances between the terminals by finding the optimal set of code matrices, energy allocation scheme, and broadcast interval length. By applying the first-layer LD code and making the broadcast interval length a design variable, we can design optimal codes for a wide range of distance/fading configurations between the terminals. Simulation results have shown that the improvement in performance of the optimized cooperative LD code.

References

[1] A. Nosratinia and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, pp. 74–80, Oct. 2004.

[2] B. Vucetic and J. Yuan, *Space-time coding*. John Wiley and Sons Inc., 2003.

[3] I. Maric and R. Yates, "Forwarding strategies for gaussian parallel-relay networks," in *Conference on Information Sciences and Systems CISS' 04*, Mar. 2004.

[4] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[5] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1804–1824, July 2002.

[6] M. Dohler, A. Gkeilas, and H. Aghvami, "Resource allocation for FDMA-based regenerative multi-hop links," *IEEE Trans. Commun.*, vol. 3, no. 6, pp. 1989–1993, Nov. 2004.

[7] J. Spall, *Introduction to stochastic search and optimization*. Wiley, 2003.

[8] T. S. Rappaport, *Wireless Communications: Priciples and Practice*. New Jersey: Prentice Hall, 1996.

[9] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in lattice, including a computational complexity," *Mathematics for Computation*, vol. 44, pp. 463–471, Apr. 1995.

[10] H. J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*. New York: Springer, 2003.

[11] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1099–1109, Aug. 2004.