2020 IEEE/CIC International Conference on Communications in China (ICCC)

Multiuser AF MIMO Multi-Relay System Design with Direct Links and MMSE-DFE Receiver

Yang Lv*, Zhiqiang He*, and Yue Rong§
*Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China
§School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Bentley, WA 6102, Australia
Email: lvyangcn@foxmail.com, hezq@bupt.edu.cn, y.rong@curtin.edu.au

***Abstract*—Adopting the amplify-and-forward (AF) relay protocol, this paper focuses on the design of a multiuser multiple-input multiple-output (MIMO) multi-relay system with direct source-destination links taken into consideration. In order to improve the quality of signal detection at the receiver, we make use of the minimal mean-squared error (MMSE)-decision feedback equalization (DFE) technique. With the constraints of the signal transmission power at both source and relay nodes, the minimization of the sum mean-squared error (MSE) for the signal waveform estimation of all users' data streams is employed as our design criterion. Through the block coordinate descent (BCD) method of Gauss-Seidel type, we develop an iterative algorithm with guaranteed convergence to conduct the joint optimization of all the source precoding, relay amplifying, feed-forward and decision feedback matrices, where each step we take is to solve a convex problem. It is shown from simulation results that, in comparison to the linear MMSE receiver-based algorithm, the proposed nonlinear one achieves better performance in terms of both MSE and bit-error-rate (BER). Moreover, the system reliability can be significantly improved when the signal-to-noise ratios (SNRs) of direct links become relatively good.**

***Index Terms*—multiuser, MIMO relay, AF, multi-relay, direct links, MMSE, DFE**

I. Introduction

In the past decade, there has been a rapid development in the research field of multiple-input multiple-output (MIMO) relay communications [1], where the linear non-regenerative amplify-and-forward (AF) relay protocol attracts considerable attention since it not only can expand the system coverage but also has simple implementation structure and high processing speed [2]. For a three-terminal half-duplex (HD) AF MIMO relay system with no direct link between the source and destination nodes, [3] presented the analytical characterization of its ergodic capacity. [4] developed a unified framework, including Schur-concave and Schur-convex objective functions, to optimize multicarrier signal transmission. [5] investigated the transceiver design of a dual-hop AF MIMO-orthogonal frequency division multiplexing (OFDM) relay system under channel uncertainties. Considering both one-way and two-way relay systems with multiple relay nodes, [6] proposed joint optimization schemes based on sum-rate maximization and mean-squared error (MSE) minimization criteria. In the presence of direct links, [7] provided low complexity transceiver design algorithms for three-terminal relay systems. [8] developed iterative methods to solve the robust source and relay design problem under multiuser scenarios. To further improve the performance of a multi-hop relay system with any number of hops, instead of the linear minimal mean-squared error (MMSE) receiver, [9] adopted the nonlinear MMSE-decision feedback equalization (DFE) receiver and proved its effectiveness for Schur-convex objective functions. Targeting at a multiuser multi-hop relay system with MMSE-DFE receiver, [10] decomposed the system design problem and developed two distributed transceiver optimization algorithms.

In this paper, for the first time as far as we know, the nonlinear MMSE-DFE receiving technique is introduced to the design of a multiuser HD AF MIMO multi-relay system with not only multiple parallel relay nodes but also the direct links between source and destination nodes taken into account. To be specific, following the block coordinate descent (BCD) method of Gauss-Seidel type [11], we jointly optimize the precoding matrices of source nodes, the amplifying matrices of relay nodes as well as the feed-forward and feedback matrices in MMSE-DFE receiver by using the MSE minimization criterion under transmission power constraints. Numerical simulations verify the superiority of the proposed iterative algorithm by comparison with the linear MMSE receiver-based one, which indicates the potential of our research to the future development of 5G communications, especially the information aggregation services for Internet of Things [12].

This paper is organized as follows. Section II sets forth the system model and target problem. Section III carries out the design of an iterative transceiver optimization algorithm. Section IV gives the numerical simulation results. Section V draws a conclusion.

Throughout the paper, $(\cdot)^T$, $(\cdot)^H$ denote the transpose and complex conjugate transpose of a vector or matrix. $|x|$, x^{-1}, x^* represent the modulus, reciprocal and complex conjugate of scalar x. For matrix $\boldsymbol{X}$, $\mathrm{rank}(\boldsymbol{X})$, $\mathrm{tr}(\boldsymbol{X})$, $\boldsymbol{X}^{-1}$, $\boldsymbol{X}^{\dagger}$, $\boldsymbol{X}^{1/2}$ stand for its rank, trace, inverse, pseudo-inverse and squareroot [13], besides, $[\boldsymbol{X}]_n$, $[\boldsymbol{X}]_{m,n}$, $[\boldsymbol{X}]_{1:n}$, $[\boldsymbol{X}]_{1:m,1:n}$ represent its nth column vector, mth row and nth column element, leftmost n columns and the submatrix composed of its first m rows and first n columns. $\mathcal{U}[\boldsymbol{X}]$ denotes the strictly upper triangular part of matrix $\boldsymbol{X}$. $\mathrm{bd}(\cdot)$ stands for a block diagonal matrix. $\lambda_i(\boldsymbol{X})$ represents the ith biggest eigenvalue of a Hermitian matrix

This work was supported by the National Key Research and Development Program of China under Grant 2018YFB1801101.

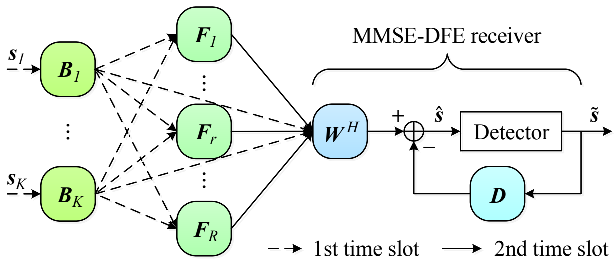

Fig. 1. System model for multiuser AF MIMO multi-relay communications with direct links and MMSE-DFE receiver.

$\boldsymbol{X}$. $\mathrm{E}[\cdot]$ is the statistical expectation operator. $\boldsymbol{I}_n$, $\boldsymbol{0}_{m\times n}$ are nth-order identity matrix and $m\times n$ zero matrix. $\mathcal{CN}(\boldsymbol{0},\boldsymbol{X})$ ($\mathcal{CN}(0,x)$) denotes the distribution of a circularly symmetric complex Gaussian (CSCG) random vector (variable) with zero mean and covariance matrix $\boldsymbol{X}$ (variance x).

II. System Model and Target Problem

As shown in Fig. 1, we consider a multiuser AF MIMO multi-relay system with direct links and MMSE-DFE receiver, where the kth source node, $k=1,\cdots,K$, transmits N_k data streams through $N_{s,k}$ antennas ($N_k\leqslant N_{s,k}$), the rth relay node, $r=1,\cdots,R$, has M_r antennas and the destination node has N_d antennas. With all nodes working in HD mode, the communication process here can be completed within two time slots. For the first time slot, the modulated signal vector $\boldsymbol{s}_k\in\mathbb{C}^{N_k}$ at the kth source node is linearly precoded into $\boldsymbol{x}_k=\boldsymbol{B}_k\boldsymbol{s}_k$, where $\boldsymbol{B}_k\in\mathbb{C}^{N_{s,k}\times N_k}$ is the kth source precoding matrix and $\boldsymbol{x}_k$ is transmitted towards relay and destination nodes. Here we assume that all users are symbol-synchronous and $\boldsymbol{s}_k$ satisfies $\mathrm{E}\big[\boldsymbol{s}_k\boldsymbol{s}_k^H\big]=\boldsymbol{I}_{N_k}$. Thus the received signal vector at the rth relay node is given by

$$\boldsymbol{y}_r=\sum_{k=1}^{K}\boldsymbol{H}_{s,rk}\boldsymbol{x}_k+\boldsymbol{n}_r=\boldsymbol{H}_{s,r}\boldsymbol{B}\boldsymbol{s}+\boldsymbol{n}_r \tag{1}$$

where $\boldsymbol{H}_{s,rk}\in\mathbb{C}^{M_r\times N_{s,k}}$ is the channel matrix between the kth source node and the rth relay node, $\boldsymbol{n}_r\in\mathbb{C}^{M_r}$ is the noise vector at the rth relay node, $\boldsymbol{s}\triangleq\big[\boldsymbol{s}_1^T,\cdots,\boldsymbol{s}_K^T\big]^T\in\mathbb{C}^N$, $\boldsymbol{B}\triangleq\mathrm{bd}(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_K)\in\mathbb{C}^{N_s\times N}$, $\boldsymbol{H}_{s,r}\triangleq\big[\boldsymbol{H}_{s,r1},\cdots,\boldsymbol{H}_{s,rK}\big]\in\mathbb{C}^{M_r\times N_s}$ with $N\triangleq\sum_{k=1}^{K}N_k$, $N_s\triangleq\sum_{k=1}^{K}N_{s,k}$. Besides, through non-negligible direct source-destination links, the received signal vector at the destination node for the first time slot is given by

$$\boldsymbol{y}_{sd}=\sum_{k=1}^{K}\boldsymbol{H}_{sd,k}\boldsymbol{x}_k+\boldsymbol{n}_{sd}=\boldsymbol{H}_{sd}\boldsymbol{B}\boldsymbol{s}+\boldsymbol{n}_{sd} \tag{2}$$

where $\boldsymbol{H}_{sd,k}\in\mathbb{C}^{N_d\times N_{s,k}}$ is the channel matrix between the kth source node and the destination node, $\boldsymbol{n}_{sd}\in\mathbb{C}^{N_d}$ is the noise vector at the destination node for the first time slot and $\boldsymbol{H}_{sd}\triangleq\big[\boldsymbol{H}_{sd,1},\cdots,\boldsymbol{H}_{sd,K}\big]\in\mathbb{C}^{N_d\times N_s}$. For the second time slot, the rth relay node amplifies its received signal vector into $\boldsymbol{x}_r=\boldsymbol{F}_r\boldsymbol{y}_r$, where $\boldsymbol{F}_r\in\mathbb{C}^{M_r\times M_r}$ is the rth relay amplifying matrix and $\boldsymbol{x}_r$ is transmitted towards the destination node. Hence the received signal vector at the destination node for the second time slot is given by

$$\boldsymbol{y}_d=\sum_{r=1}^{R}\boldsymbol{H}_{d,r}\boldsymbol{x}_r+\boldsymbol{n}_d=\boldsymbol{H}_d\boldsymbol{F}\boldsymbol{H}_s\boldsymbol{B}\boldsymbol{s}+\boldsymbol{H}_d\boldsymbol{F}\boldsymbol{n}+\boldsymbol{n}_d \tag{3}$$

where $\boldsymbol{H}_{d,r}\in\mathbb{C}^{N_d\times M_r}$ is the channel matrix between the rth relay node and the destination node, $\boldsymbol{n}_d\in\mathbb{C}^{N_d}$ is the noise vector at the destination node for the second time slot, meanwhile, we define $\boldsymbol{H}_s\triangleq\big[\boldsymbol{H}_{s,1}^T,\cdots,\boldsymbol{H}_{s,R}^T\big]^T\in\mathbb{C}^{M\times N_s}$, $\boldsymbol{F}\triangleq\mathrm{bd}(\boldsymbol{F}_1,\cdots,\boldsymbol{F}_R)\in\mathbb{C}^{M\times M}$, $\boldsymbol{H}_d\triangleq\big[\boldsymbol{H}_{d,1},\cdots,\boldsymbol{H}_{d,R}\big]\in\mathbb{C}^{N_d\times M}$, $\boldsymbol{n}\triangleq\big[\boldsymbol{n}_1^T,\cdots,\boldsymbol{n}_R^T\big]^T\in\mathbb{C}^M$ with $M\triangleq\sum_{r=1}^{R}M_r$.

Over the above two time slots, the composite received signal vector at the destination node can be written as

$$\boldsymbol{y}\triangleq\begin{bmatrix}\boldsymbol{y}_d\\ \boldsymbol{y}_{sd}\end{bmatrix}=\begin{bmatrix}\boldsymbol{H}_d\boldsymbol{F}\boldsymbol{H}_s\\ \boldsymbol{H}_{sd}\end{bmatrix}\boldsymbol{B}\boldsymbol{s}+\begin{bmatrix}\boldsymbol{H}_d\boldsymbol{F}\boldsymbol{n}+\boldsymbol{n}_d\\ \boldsymbol{n}_{sd}\end{bmatrix}=\boldsymbol{A}\boldsymbol{s}+\boldsymbol{v} \tag{4}$$

with $\boldsymbol{A}\triangleq\begin{bmatrix}\boldsymbol{H}_d\boldsymbol{F}\boldsymbol{H}_s\\ \boldsymbol{H}_{sd}\end{bmatrix}\boldsymbol{B}\in\mathbb{C}^{(2N_d)\times N}$, $\boldsymbol{v}\triangleq\begin{bmatrix}\boldsymbol{H}_d\boldsymbol{F}\boldsymbol{n}+\boldsymbol{n}_d\\ \boldsymbol{n}_{sd}\end{bmatrix}\in\mathbb{C}^{2N_d}$. In our system, $\boldsymbol{H}_{s,rk}$, $\boldsymbol{H}_{d,r}$, $\boldsymbol{H}_{sd,k}$ are considered as quasi-static block fading channel matrices which remain constant for each transmission of N data streams from all source nodes to the destination node. For achieving an acceptable system performance, we typically require $\min\{\mathrm{rank}(\boldsymbol{H}_s),\mathrm{rank}(\boldsymbol{H}_d),\mathrm{rank}(\boldsymbol{H}_{sd})\}\geqslant N$, hence $\min\{N_s,M,N_d\}\geqslant N$. Besides, $\boldsymbol{n}_r$, $\boldsymbol{n}_d$, $\boldsymbol{n}_{sd}$ are independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) vectors with distributions $\mathcal{CN}(\boldsymbol{0},\boldsymbol{I}_{M_r})$ for $\boldsymbol{n}_r$ and $\mathcal{CN}(\boldsymbol{0},\boldsymbol{I}_{N_d})$ for $\boldsymbol{n}_d$, $\boldsymbol{n}_{sd}$. Therefore we have

$$\boldsymbol{C}\triangleq\mathrm{E}\big[\boldsymbol{v}\boldsymbol{v}^H\big]=\begin{bmatrix}\boldsymbol{H}_d\boldsymbol{F}\boldsymbol{F}^H\boldsymbol{H}_d^H+\boldsymbol{I}_{N_d}, & \boldsymbol{0}_{N_d\times N_d}\\ \boldsymbol{0}_{N_d\times N_d}, & \boldsymbol{I}_{N_d}\end{bmatrix}. \tag{5}$$

In the MMSE-DFE receiver at the destination node, first of all, the Nth symbol in $\boldsymbol{s}$ is estimated as $\hat{s}_N=\boldsymbol{w}_N^H\boldsymbol{y}$ and detected as $\tilde{s}_N$. Then, for $i=N-1,\cdots,1$, the ith symbol in $\boldsymbol{s}$ is estimated as $\hat{s}_i=\boldsymbol{w}_i^H\boldsymbol{y}-\sum_{j=i+1}^{N}d_{i,j}\tilde{s}_j$. Here, $\boldsymbol{w}_i\in\mathbb{C}^{2N_d}$ is the ith feed-forward vector for $i=1,\cdots,N$ and $d_{i,j}$ is the (i,j)th decision feedback coefficient for $j=i+1,\cdots,N$ and $i=1,\cdots,N-1$. Thus, with the estimated signal vector $\hat{\boldsymbol{s}}\triangleq[\hat{s}_1,\cdots,\hat{s}_N]^T\in\mathbb{C}^N$, the detected signal vector $\tilde{\boldsymbol{s}}\triangleq[\tilde{s}_1,\cdots,\tilde{s}_N]^T\in\mathbb{C}^N$, the feed-forward matrix $\boldsymbol{W}\triangleq[\boldsymbol{w}_1,\cdots,\boldsymbol{w}_N]\in\mathbb{C}^{(2N_d)\times N}$ and the strictly upper triangular decision feedback matrix $\boldsymbol{D}\in\mathbb{C}^{N\times N}$ whose ith row and jth column element is $d_{i,j}$, we have $\hat{\boldsymbol{s}}=\boldsymbol{W}^H\boldsymbol{y}-\boldsymbol{D}\tilde{\boldsymbol{s}}$, from which, by further assuming no error propagation as in [9]–[10] with $\tilde{\boldsymbol{s}}=\boldsymbol{s}$, we can obtain

$$\hat{\boldsymbol{s}}=\boldsymbol{W}^H\boldsymbol{y}-\boldsymbol{D}\boldsymbol{s}=\big(\boldsymbol{W}^H\boldsymbol{A}-\boldsymbol{D}\big)\,\boldsymbol{s}+\boldsymbol{W}^H\boldsymbol{v}. \tag{6}$$

Now, the MSE of the signal waveform estimation of each data stream can be derived as

$$E_i\triangleq\mathrm{E}\Big[\big|\hat{s}_i-s_i\big|^2\Big]=\left(\sum_{j=1}^{i-1}\Big|\boldsymbol{w}_i^H\,[\boldsymbol{A}]_j\Big|^2\right)+\Big|\boldsymbol{w}_i^H\,[\boldsymbol{A}]_i-1\Big|^2+\left(\sum_{j=i+1}^{N}\Big|\boldsymbol{w}_i^H\,[\boldsymbol{A}]_j-d_{i,j}\Big|^2\right)+\boldsymbol{w}_i^H\boldsymbol{C}\boldsymbol{w}_i \tag{7}$$

for $i = 1, \cdots, N$, where if $i = N$ the third term in (7) will become zero. So the sum MSE for all data streams is given by

$$E_s \triangleq \sum_{i=1}^{N} E_i = \mathrm{tr}\Big\{\mathrm{E}\Big[(\hat{\boldsymbol{s}} - \boldsymbol{s})(\hat{\boldsymbol{s}} - \boldsymbol{s})^H\Big]\Big\}$$
$$= \mathrm{tr}\Big[\big(\boldsymbol{W}^H\boldsymbol{A} - \boldsymbol{U}\big)\big(\boldsymbol{W}^H\boldsymbol{A} - \boldsymbol{U}\big)^H + \boldsymbol{W}^H\boldsymbol{C}\boldsymbol{W}\Big] \quad (8)$$

with $\boldsymbol{U} \triangleq \boldsymbol{I}_N + \boldsymbol{D}$ being called the decision feedback matrix as well. Moreover, we can derive the signal transmission power of the kth source node as $Q_k \triangleq \mathrm{tr}\{\mathrm{E}[\boldsymbol{x}_k\boldsymbol{x}_k^H]\} = \mathrm{tr}(\boldsymbol{B}_k\boldsymbol{B}_k^H)$ and that of the rth relay node as $P_r \triangleq \mathrm{tr}\{\mathrm{E}[\boldsymbol{x}_r\boldsymbol{x}_r^H]\} = \mathrm{tr}[\boldsymbol{F}_r(\boldsymbol{H}_{s,r}\boldsymbol{B}\boldsymbol{B}^H\boldsymbol{H}_{s,r}^H + \boldsymbol{I}_{M_r})\boldsymbol{F}_r^H]$, which are set to be not greater than q_k and p_r, respectively. Therefore, with $\{\boldsymbol{B}_k\} \triangleq \{\boldsymbol{B}_1, \cdots, \boldsymbol{B}_K\}$, $\{\boldsymbol{F}_r\} \triangleq \{\boldsymbol{F}_1, \cdots, \boldsymbol{F}_R\}$, the target problem for optimizing the relay system in Fig. 1 is given by

$$\min_{\{\boldsymbol{B}_k\},\{\boldsymbol{F}_r\},\boldsymbol{W},\boldsymbol{U}} \mathrm{tr}\Big[\big(\boldsymbol{W}^H\boldsymbol{A} - \boldsymbol{U}\big) \times \big(\boldsymbol{W}^H\boldsymbol{A} - \boldsymbol{U}\big)^H + \boldsymbol{W}^H\boldsymbol{C}\boldsymbol{W}\Big] \quad (9)$$

$$\text{s.t. } \mathrm{tr}\big(\boldsymbol{B}_k\boldsymbol{B}_k^H\big) \leqslant q_k, \quad k = 1, \cdots, K, \quad (10)$$

$$\mathrm{tr}\big[\boldsymbol{F}_r\left(\boldsymbol{H}_{s,r}\boldsymbol{B}\boldsymbol{B}^H\boldsymbol{H}_{s,r}^H + \boldsymbol{I}_{M_r}\right)\boldsymbol{F}_r^H\big] \leqslant p_r, \quad r = 1, \cdots, R, \quad (11)$$

$$[\boldsymbol{U}]_{i,j} = \begin{cases} 0, & i > j, \\ 1, & i = j. \end{cases} \quad (12)$$

Note that the problem formulated above is the first one as far as we know in the joint optimization of the matrix parameters in a multiuser AF MIMO multi-relay system with direct links and MMSE-DFE receiver. For solving this challenging nonconvex problem, in the next section, we employ the Gauss-Seidel BCD method and propose an iterative algorithm.

III. Algorithm Design

Since the key principle of the Gauss-Seidel BCD method is to iteratively optimize blocks of variables, the algorithm developed here is made up of three main steps in one iteration, aiming to optimize $\{\boldsymbol{B}_k\}$, $\{\boldsymbol{F}_r\}$ as well as $\boldsymbol{W}$ and $\boldsymbol{U}$, respectively, and before executing these steps, we initialize $\boldsymbol{B}_k$ as $\boldsymbol{B}_k = \big[\sqrt{q_k/N_k}\,\boldsymbol{I}_{N_k}, \boldsymbol{0}^T_{(N_{s,k}-N_k)\times N_k}\big]^T$ and $\boldsymbol{F}_r$ as $\boldsymbol{F}_r = \sqrt{p_r/M_r}\left(\boldsymbol{H}_{s,r}\boldsymbol{B}\boldsymbol{B}^H\boldsymbol{H}_{s,r}^H + \boldsymbol{I}_{M_r}\right)^{-1/2}$ for $k = 1, \cdots, K$ and $r = 1, \cdots, R$.

The first step is to optimize $\boldsymbol{W}$ and $\boldsymbol{U}$ with fixed $\{\boldsymbol{B}_k\}$ and $\{\boldsymbol{F}_r\}$. It can be seen that the optimal $d_{i,j}$ which minimizes E_i in (7) should meet $d_{i,j} = \boldsymbol{w}_i^H[\boldsymbol{A}]_j$, consequently we have

$$\boldsymbol{D} = \mathcal{U}\big[\boldsymbol{W}^H\boldsymbol{A}\big] \quad (13)$$

and E_i for $i = 1, \cdots, N$ becomes

$$E_i = \boldsymbol{w}_i^H\Bigg(\sum_{j=1}^{i}[\boldsymbol{A}]_j[\boldsymbol{A}]_j^H + \boldsymbol{C}\Bigg)\boldsymbol{w}_i - \boldsymbol{w}_i^H[\boldsymbol{A}]_i - [\boldsymbol{A}]_i^H\boldsymbol{w}_i + 1. \quad (14)$$

The above is a convex quadratic function with respect to $\boldsymbol{w}_i$, whose minimum is obtained by making its gradient, $\nabla_{\boldsymbol{w}_i}E_i = \Big(\sum_{j=1}^{i}[\boldsymbol{A}]_j[\boldsymbol{A}]_j^H + \boldsymbol{C}\Big)\boldsymbol{w}_i - [\boldsymbol{A}]_i$ [14], equal to zero, thus

$$\boldsymbol{w}_i = \Big[\Big([\boldsymbol{A}]_{1:i}[\boldsymbol{A}]_{1:i}^H + \boldsymbol{C}\Big)^{-1}[\boldsymbol{A}]_{1:i}\Big]_i$$
$$= \Big[\Big[\boldsymbol{C}^{-1} - \boldsymbol{C}^{-1}[\boldsymbol{A}]_{1:i}\Big(\boldsymbol{I}_i + [\boldsymbol{A}]_{1:i}^H\boldsymbol{C}^{-1}[\boldsymbol{A}]_{1:i}\Big)^{-1} \times [\boldsymbol{A}]_{1:i}^H\boldsymbol{C}^{-1}\Big][\boldsymbol{A}]_{1:i}\Big]_i$$
$$= \Big[\boldsymbol{C}^{-1}[\boldsymbol{A}]_{1:i}\Big(\boldsymbol{I}_i + [\boldsymbol{A}]_{1:i}^H\boldsymbol{C}^{-1}[\boldsymbol{A}]_{1:i}\Big)^{-1}\Big]_i \quad (15)$$

where for obtaining the last two equations, we make use of the matrix inversion lemma, namely, $(\boldsymbol{X} + \boldsymbol{Z}_1\boldsymbol{Y}\boldsymbol{Z}_2)^{-1} = \boldsymbol{X}^{-1} - \boldsymbol{X}^{-1}\boldsymbol{Z}_1\left(\boldsymbol{Y}^{-1} + \boldsymbol{Z}_2\boldsymbol{X}^{-1}\boldsymbol{Z}_1\right)^{-1}\boldsymbol{Z}_2\boldsymbol{X}^{-1}$ for invertible matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$.

Now, the following QR factorization [13, Theorem 2.1.14] can be introduced:

$$\begin{bmatrix} \boldsymbol{C}^{-1/2}\boldsymbol{A} \\ \boldsymbol{I}_N \end{bmatrix} = \boldsymbol{Q}\boldsymbol{R} = \begin{bmatrix} \dot{\boldsymbol{Q}} \\ \ddot{\boldsymbol{Q}} \end{bmatrix}\boldsymbol{R} \quad (16)$$

where factor $\boldsymbol{Q} \in \mathbb{C}^{(2N_d+N)\times N}$ has orthonormal columns and $\boldsymbol{R} \in \mathbb{C}^{N\times N}$ is an upper triangular matrix with positive main diagonal elements, besides, all the rows of $\boldsymbol{Q}$ are divided into two parts, i.e., $\dot{\boldsymbol{Q}} \in \mathbb{C}^{(2N_d)\times N}$ and $\ddot{\boldsymbol{Q}} \in \mathbb{C}^{N\times N}$, as a result,

$$\boldsymbol{C}^{-1/2}\boldsymbol{A} = \dot{\boldsymbol{Q}}\boldsymbol{R}, \quad \boldsymbol{I}_N = \ddot{\boldsymbol{Q}}\boldsymbol{R}. \quad (17)$$

According to (17), we have

$$\boldsymbol{C}^{-1/2}[\boldsymbol{A}]_{1:i} = \big[\boldsymbol{C}^{-1/2}\boldsymbol{A}\big]_{1:i} = \big[\dot{\boldsymbol{Q}}\big]_{1:i}[\boldsymbol{R}]_{1:i,1:i}, \quad (18)$$

$$\boldsymbol{I}_i = [\boldsymbol{I}_N]_{1:i}^H[\boldsymbol{I}_N]_{1:i} = [\boldsymbol{R}]_{1:i,1:i}^H\big[\ddot{\boldsymbol{Q}}\big]_{1:i}^H\big[\ddot{\boldsymbol{Q}}\big]_{1:i}[\boldsymbol{R}]_{1:i,1:i} \quad (19)$$

from which,

$$\boldsymbol{w}_i = \Big[\boldsymbol{C}^{-1/2}\big[\dot{\boldsymbol{Q}}\big]_{1:i}[\boldsymbol{R}]_{1:i,1:i}\Big[[\boldsymbol{R}]_{1:i,1:i}^H\Big(\big[\ddot{\boldsymbol{Q}}\big]_{1:i}^H \times \big[\ddot{\boldsymbol{Q}}\big]_{1:i} + \big[\dot{\boldsymbol{Q}}\big]_{1:i}^H\big[\dot{\boldsymbol{Q}}\big]_{1:i}\Big)[\boldsymbol{R}]_{1:i,1:i}\Big]^{-1}\Big]_i. \quad (20)$$

Here, since $\big[\ddot{\boldsymbol{Q}}\big]_{1:i}^H\big[\ddot{\boldsymbol{Q}}\big]_{1:i} + \big[\dot{\boldsymbol{Q}}\big]_{1:i}^H\big[\dot{\boldsymbol{Q}}\big]_{1:i} = \big[\ddot{\boldsymbol{Q}}^H\ddot{\boldsymbol{Q}}\big]_{1:i,1:i} + \big[\dot{\boldsymbol{Q}}^H\dot{\boldsymbol{Q}}\big]_{1:i,1:i} = \big[\boldsymbol{Q}^H\boldsymbol{Q}\big]_{1:i,1:i} = \big[\boldsymbol{I}_N\big]_{1:i,1:i} = \boldsymbol{I}_i$,

$$\boldsymbol{w}_i = \Big[\boldsymbol{C}^{-1/2}\big[\dot{\boldsymbol{Q}}\big]_{1:i}[\boldsymbol{R}]_{1:i,1:i}^{-H}\Big]_i = \boldsymbol{C}^{-1/2}\big[\dot{\boldsymbol{Q}}\big]_i[\boldsymbol{R}]_{i,i}^{-*} \quad (21)$$

for $i = 1, \cdots, N$. Thus the optimal feed-forward matrix is

$$\boldsymbol{W} = \boldsymbol{C}^{-1/2}\dot{\boldsymbol{Q}}\boldsymbol{D}_{\boldsymbol{R}}^{-H} \quad (22)$$

where $\boldsymbol{D}_{\boldsymbol{R}} \in \mathbb{C}^{N\times N}$ represents a diagonal matrix with the same diagonal elements as those in $\boldsymbol{R}$. Now, by using (13), (16), (17) and (22), we can further get the optimal decision feedback matrix as

$$\boldsymbol{D} = \mathcal{U}\Big[\boldsymbol{D}_{\boldsymbol{R}}^{-1}\dot{\boldsymbol{Q}}^H\dot{\boldsymbol{Q}}\boldsymbol{R}\Big] = \mathcal{U}\Big[\boldsymbol{D}_{\boldsymbol{R}}^{-1}\Big(\boldsymbol{I}_N - \ddot{\boldsymbol{Q}}^H\ddot{\boldsymbol{Q}}\Big)\boldsymbol{R}\Big]$$
$$= \mathcal{U}\big[\boldsymbol{D}_{\boldsymbol{R}}^{-1}\boldsymbol{R} - \boldsymbol{D}_{\boldsymbol{R}}^{-1}\boldsymbol{R}^{-H}\big] = \boldsymbol{D}_{\boldsymbol{R}}^{-1}\boldsymbol{R} - \boldsymbol{I}_N, \quad (23)$$

therefore the optimal $\boldsymbol{U}$ is given by

$$\boldsymbol{U} = \boldsymbol{D}_{\boldsymbol{R}}^{-1}\boldsymbol{R}. \tag{24}$$

It is worth noting that the nonlinear MMSE-DFE receiver optimized above can be readily changed into the degenerate linear MMSE receiver. To this end, we set $\boldsymbol{U}$ to be $\boldsymbol{I}_N$ and estimate $\boldsymbol{s}$ as $\bar{\boldsymbol{W}}^H\boldsymbol{y}$, where $\bar{\boldsymbol{W}} \in \mathbb{C}^{(2N_d)\times N}$ denotes the linear receiving matrix. Hence, the sum MSE becomes $\bar{E}_s = \mathrm{tr}\Big[\big(\bar{\boldsymbol{W}}^H\boldsymbol{A} - \boldsymbol{I}_N\big)\big(\bar{\boldsymbol{W}}^H\boldsymbol{A} - \boldsymbol{I}_N\big)^H + \bar{\boldsymbol{W}}^H\boldsymbol{C}\bar{\boldsymbol{W}}\Big]$, of which the gradient with respect to $\bar{\boldsymbol{W}}$ is $\nabla_{\bar{\boldsymbol{W}}}\bar{E}_s = \boldsymbol{A}\big(\bar{\boldsymbol{W}}^H\boldsymbol{A} - \boldsymbol{I}_N\big)^H + \boldsymbol{C}\bar{\boldsymbol{W}}$. By making it equal to zero, we can get the optimal $\bar{\boldsymbol{W}}$ as

$$\bar{\boldsymbol{W}} = \big(\boldsymbol{A}\boldsymbol{A}^H + \boldsymbol{C}\big)^{-1}\boldsymbol{A}. \tag{25}$$

The second step is to optimize $\{\boldsymbol{B}_k\}$ with fixed $\{\boldsymbol{F}_r\}$, $\boldsymbol{W}$ and $\boldsymbol{U}$. Here we define

$$\boldsymbol{G} \triangleq \boldsymbol{W}^H \begin{bmatrix} \boldsymbol{H}_d\boldsymbol{F}\boldsymbol{H}_s \\ \boldsymbol{H}_{sd} \end{bmatrix} = [\boldsymbol{G}_1, \cdots, \boldsymbol{G}_K] \tag{26}$$

where for $k = 1, \cdots, K$, $\boldsymbol{G}_k \in \mathbb{C}^{N\times N_{s,k}}$ consists of the $\big(\sum_{j=0}^{k-1} N_{s,j} + 1\big)$th column to the $\big(\sum_{j=0}^{k} N_{s,j}\big)$th column of $\boldsymbol{G}$ with $N_{s,0} = 0$. Additionally, we set $\boldsymbol{U} = [\boldsymbol{U}_1, \cdots, \boldsymbol{U}_K]$, where for $k = 1, \cdots, K$, $\boldsymbol{U}_k \in \mathbb{C}^{N\times N_k}$ consists of the $\big(\sum_{j=0}^{k-1} N_j + 1\big)$th column to the $\big(\sum_{j=0}^{k} N_j\big)$th column of $\boldsymbol{U}$ with $N_0 = 0$. Therefore the optimization problem with respect to $\{\boldsymbol{B}_k\}$ can be written as

$$\min_{\{\boldsymbol{B}_k\}} \sum_{k=1}^{K} \mathrm{tr}\Big[(\boldsymbol{G}_k\boldsymbol{B}_k - \boldsymbol{U}_k)(\boldsymbol{G}_k\boldsymbol{B}_k - \boldsymbol{U}_k)^H\Big] \tag{27}$$

$$\text{s.t. } \mathrm{tr}\big(\boldsymbol{B}_k\boldsymbol{B}_k^H\big) \leqslant q_k, \quad k = 1, \cdots, K, \tag{28}$$

$$\sum_{k=1}^{K} \mathrm{tr}\big(\boldsymbol{F}_r\boldsymbol{H}_{s,rk}\boldsymbol{B}_k\boldsymbol{B}_k^H\boldsymbol{H}_{s,rk}^H\boldsymbol{F}_r^H\big) \leqslant \tilde{p}_r, \quad r = 1, \cdots, R \tag{29}$$

where $\tilde{p}_r \triangleq p_r - \mathrm{tr}\big(\boldsymbol{F}_r\boldsymbol{F}_r^H\big)$. The above is a convex quadratically constrained quadratic programming (QCQP) problem and several approaches, like the famous interior-point method [15], can be utilized to solve it. Noteworthily, the Matlab-based convex programming software CVX [16] is an efficient tool to help obtain the optimal solution of the problem (27)–(29).

The third step is to optimize each $\boldsymbol{F}_r$ for $r = 1, \cdots, R$ with fixed $\{\boldsymbol{F}_j \,|\, j \neq r\}$, $\boldsymbol{W}$, $\boldsymbol{U}$ and $\{\boldsymbol{B}_k\}$. Here we set $\boldsymbol{W} = \big[\boldsymbol{W}_1^T, \boldsymbol{W}_2^T\big]^T$, where $\boldsymbol{W}_1, \boldsymbol{W}_2 \in \mathbb{C}^{N_d\times N}$ are made up of the first N_d rows and the last N_d rows of $\boldsymbol{W}$, respectively. Then, with $\boldsymbol{G}_{s,r} \triangleq \boldsymbol{H}_{s,r}\boldsymbol{B}$, $\boldsymbol{G}_{d,r} \triangleq \boldsymbol{W}_1^H\boldsymbol{H}_{d,r}$ and $\boldsymbol{J}_r \triangleq \boldsymbol{U} - \sum_{j=1,j\neq r}^{R}\boldsymbol{W}_1^H\boldsymbol{H}_{d,j}\boldsymbol{F}_j\boldsymbol{H}_{s,j}\boldsymbol{B} - \boldsymbol{W}_2^H\boldsymbol{H}_{sd}\boldsymbol{B}$, the optimization problem with respect to $\boldsymbol{F}_r$ is given by

$$\min_{\boldsymbol{F}_r} \; f_r(\boldsymbol{F}_r) = \mathrm{tr}\Big[(\boldsymbol{G}_{d,r}\boldsymbol{F}_r\boldsymbol{G}_{s,r} - \boldsymbol{J}_r) \times (\boldsymbol{G}_{d,r}\boldsymbol{F}_r\boldsymbol{G}_{s,r} - \boldsymbol{J}_r)^H + \boldsymbol{G}_{d,r}\boldsymbol{F}_r\boldsymbol{F}_r^H\boldsymbol{G}_{d,r}^H\Big] \tag{30}$$

$$\text{s.t. } g_r(\boldsymbol{F}_r) = \mathrm{tr}\big[\boldsymbol{F}_r\big(\boldsymbol{G}_{s,r}\boldsymbol{G}_{s,r}^H + \boldsymbol{I}_{M_r}\big)\boldsymbol{F}_r^H\big] - p_r \leqslant 0. \tag{31}$$

The above is also a convex QCQP problem, which can be solved via the Karush-Kuhn-Tucker (KKT) conditions [15] with the Lagrange multiplier μ_r, i.e.,

$$g_r(\boldsymbol{F}_r) \leqslant 0, \quad \mu_r \geqslant 0, \quad \mu_r\, g_r(\boldsymbol{F}_r) = 0, \quad \nabla_{\boldsymbol{F}_r}L_r(\boldsymbol{F}_r, \mu_r) = \boldsymbol{0}_{M_r\times M_r} \tag{32}$$

where the gradient of the Lagrangian $L_r(\boldsymbol{F}_r, \mu_r) \triangleq f_r(\boldsymbol{F}_r) + \mu_r\, g_r(\boldsymbol{F}_r)$ with respect to $\boldsymbol{F}_r$ is derived as

$$\nabla_{\boldsymbol{F}_r}L_r(\boldsymbol{F}_r, \mu_r) = \big(\boldsymbol{G}_{d,r}^H\boldsymbol{G}_{d,r} + \mu_r\boldsymbol{I}_{M_r}\big)\boldsymbol{F}_r\boldsymbol{S}_r - \boldsymbol{T}_r \tag{33}$$

with $\boldsymbol{S}_r \triangleq \boldsymbol{G}_{s,r}\boldsymbol{G}_{s,r}^H + \boldsymbol{I}_{M_r}$ and $\boldsymbol{T}_r \triangleq \boldsymbol{G}_{d,r}^H\boldsymbol{J}_r\boldsymbol{G}_{s,r}^H$.

At this point, there are two possible cases to obtain the optimal $\boldsymbol{F}_r$ from (32). One case is $\mu_r = 0$, where we have

$$\boldsymbol{F}_r = \big(\boldsymbol{G}_{d,r}^H\boldsymbol{G}_{d,r}\big)^{\dagger}\boldsymbol{T}_r\boldsymbol{S}_r^{-1}. \tag{34}$$

If (34) satisfies $g_r(\boldsymbol{F}_r) \leqslant 0$, then it is the unique optimal solution. Otherwise, we should consider the other case, i.e., $\mu_r > 0$, from which,

$$\boldsymbol{F}_r = \big(\boldsymbol{G}_{d,r}^H\boldsymbol{G}_{d,r} + \mu_r\boldsymbol{I}_{M_r}\big)^{-1}\boldsymbol{T}_r\boldsymbol{S}_r^{-1} \tag{35}$$

and $g_r(\boldsymbol{F}_r) = 0$ ought to be satisfied. In order to determine the optimal μ_r in (35), we substitute (35) into $g_r(\boldsymbol{F}_r) = 0$, resulting in

$$g_r(\mu_r) = \mathrm{tr}\Big[\big(\boldsymbol{G}_{d,r}^H\boldsymbol{G}_{d,r} + \mu_r\boldsymbol{I}_{M_r}\big)^{-2}\boldsymbol{T}_r\boldsymbol{S}_r^{-1}\boldsymbol{T}_r^H\Big] - p_r = 0. \tag{36}$$

Note that $g_r(\mu_r)$ is a monotonically decreasing function and an upper bound of μ_r can be found for $g_r(\mu_r) = 0$. Specifically, according to (36) and Theorem 4.3.53 in [13], we have

$$\begin{aligned} p_r &\leqslant \lambda_1\Big[\big(\boldsymbol{G}_{d,r}^H\boldsymbol{G}_{d,r} + \mu_r\boldsymbol{I}_{M_r}\big)^{-2}\Big]\,\mathrm{tr}\big(\boldsymbol{T}_r\boldsymbol{S}_r^{-1}\boldsymbol{T}_r^H\big) \\ &= \big[\lambda_{M_r}\big(\boldsymbol{G}_{d,r}^H\boldsymbol{G}_{d,r}\big) + \mu_r\big]^{-2}\,\mathrm{tr}\big(\boldsymbol{T}_r\boldsymbol{S}_r^{-1}\boldsymbol{T}_r^H\big) \\ &\leqslant \mu_r^{-2}\,\mathrm{tr}\big(\boldsymbol{T}_r\boldsymbol{S}_r^{-1}\boldsymbol{T}_r^H\big) \end{aligned} \tag{37}$$

which further leads to

$$\mu_r \leqslant B_u \triangleq \sqrt{\mathrm{tr}\big(\boldsymbol{T}_r\boldsymbol{S}_r^{-1}\boldsymbol{T}_r^H\big)/\,p_r}. \tag{38}$$

Now, with $0 < \mu_r \leqslant B_u$, the solution of $g_r(\mu_r) = 0$ can be readily gotten through the bisection method [15], thus the optimal $\boldsymbol{F}_r$ in (35) is determined.

So far, all the major procedures for one iteration of our proposed algorithm have already been covered, which include computing $\boldsymbol{W}$, $\boldsymbol{U}$ as (22), (24), solving the problem (27)–(29) to obtain $\{\boldsymbol{B}_k\}$ and solving the problem (30)–(31) to obtain $\boldsymbol{F}_r$ for $r = 1, \cdots, R$. Hereafter, we call this iterative BCD algorithm "the DFE-Rx algorithm" for the nonlinear receiver it adopts. Note that during the iterations of the DFE-Rx algorithm, each step we take is to solve a convex problem with respect to a certain block of variables, which yields a unique optimal solution to make the value of the objective function E_s decrease. Since E_s also has a lower bound of at least zero, the convergence of this algorithm is assured and every limit point converged by the iterative process is a Nash point [11]. To implement the DFE-Rx algorithm in

TABLE I
SYSTEM SETTINGS

Ex.	K	$\tilde{N}$	$\tilde{N}_s$	R	$\tilde{M}$	N_d	ΔSNR (dB)
1	2	4	4	1	8	8	30
2	2	4	4	1	8	8	20
3	2	4	4	1	8	8	10
4	2	4	4	2	8	8	30
5	2	4	4	2	4	8	30
6	2	3	4	2	4	8	30

practice, one of the relay nodes can be assigned to collect the channel state information, perform the developed iterative algorithm and deliver the optimized system parameters to their corresponding source, relay and destination nodes. In this paper, we make comparisons between the DFE-Rx algorithm and the linear MMSE receiver-based algorithm, which follows the same procedures to optimize $\{\boldsymbol{B}_k\}$ and $\{\boldsymbol{F}_r\}$, yet adopts the linear receiving matrix given by (25) and sets $\boldsymbol{U}$ as $\boldsymbol{I}_N$. This algorithm is hereafter called "the L-Rx algorithm", noteworthily, whose procedures are similar to the Tri-Step algorithm proposed in [8]. Through simulation tests, it is deemed appropriate to carry out both the DFE-Rx algorithm and the L-Rx algorithm for 10 iterations, since after that, the performance gains can almost be neglected, which also reflects the fast convergence speed of the considered algorithms.

IV. NUMERICAL SIMULATIONS

The superior MSE and bit-error-rate (BER) performance of the DFE-Rx algorithm over the L-Rx algorithm is verified by Monte Carlo simulations, which are carried out with Intel Core i7-10510U processor and Matlab R2017b software under 64-bit Windows 10 operating system. Here, the i.i.d. Rayleigh fading channel environment is assumed, where all the elements in each of $\boldsymbol{H}_{s,rk}$, $\boldsymbol{H}_{d,r}$, $\boldsymbol{H}_{sd,k}$ are i.i.d. random variables, subject to $\mathcal{CN}\big(0, \sigma^2_{s,rk}/N_{s,k}\big)$, $\mathcal{CN}\big(0, \sigma^2_{d,r}/M_r\big)$, $\mathcal{CN}\big(0, \sigma^2_{sd,k}/N_{s,k}\big)$, respectively, with the signal propagation path loss represented by the variances $\sigma^2_{s,rk}/N_{s,k}$, $\sigma^2_{d,r}/M_r$, $\sigma^2_{sd,k}/N_{s,k}$, which are normalized by the numbers of transmitting antennas and also implicitly contain the noise powers. Hence the signal-to-noise ratios (SNRs) of the (k, r)th source-relay link, the rth relay-destination link and the kth source-destination link can be defined as $\text{SNR}_{s,rk} \triangleq q_k \sigma^2_{s,rk}/N_{s,k}$, $\text{SNR}_{d,r} \triangleq p_r \sigma^2_{d,r}/M_r$ and $\text{SNR}_{sd,k} \triangleq q_k \sigma^2_{sd,k}/N_{s,k}$. For the sake of simplicity, we set $N_k = \tilde{N}$, $N_{s,k} = \tilde{N}_s$, $M_r = \tilde{M}$, $\text{SNR}_{s,rk} = \text{SNR}_{d,r} = \text{SNR}$ and $\text{SNR}_{sd,k} = \text{SNR} - \Delta\text{SNR}$ for $k = 1, \cdots, K$ and $r = 1, \cdots, R$, where ΔSNR denotes the difference of SNR between the indirect source-relay, relay-destination links and the direct source-destination links. All the simulation results are obtained from the average of 1500 independent channel realizations with the range of SNR being 0–30 dB as well as 6 examples (Exs.) of system settings as given in Table I.

In the following, Figs. 2–3 and Figs. 4–5 respectively show the MSE performance and the BER performance of the DFE-Rx and L-Rx algorithms. Specifically, Figs. 2–3 display the average MSE per data stream, i.e., E_s/N, and in order to obtain the BER curves in Figs. 4–5, half a million bits per data stream under QPSK modulation are transmitted through the relay system, where we consider the influence of the error propagation within the MMSE-DFE receiver, whose decision feedback symbols are not the exactly correct ones as assumed in mathematical derivations, but the practical ones regenerated from previously detected bits with detection errors.

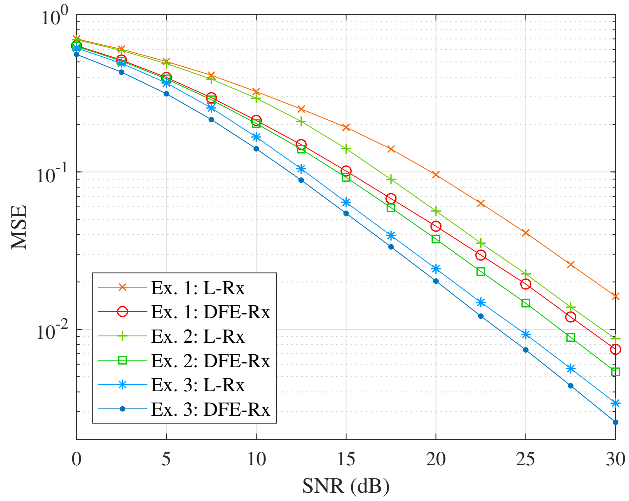

Fig. 2. MSE versus SNR performance comparisons for Exs. 1–3.

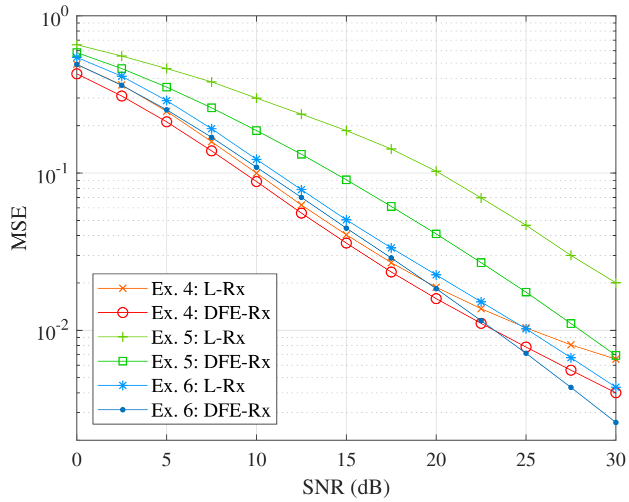

Fig. 3. MSE versus SNR performance comparisons for Exs. 4–6.

According to the simulation results in Figs. 2–5, for both the DFE-Rx and L-Rx algorithms, Ex. 2 has better MSE and BER performance than that of Ex. 1, which is due to the increased SNRs of direct links, and for the same reason, Ex. 3 outperforms Ex. 2. Compared with Ex. 1, Ex. 4 increases the number of relay nodes, thus brings in more power resources and has evident performance improvement. Since the relay nodes in Ex. 5 are equipped with less antennas than those in Ex. 4, the spatial diversity order is reduced, thereby lowering

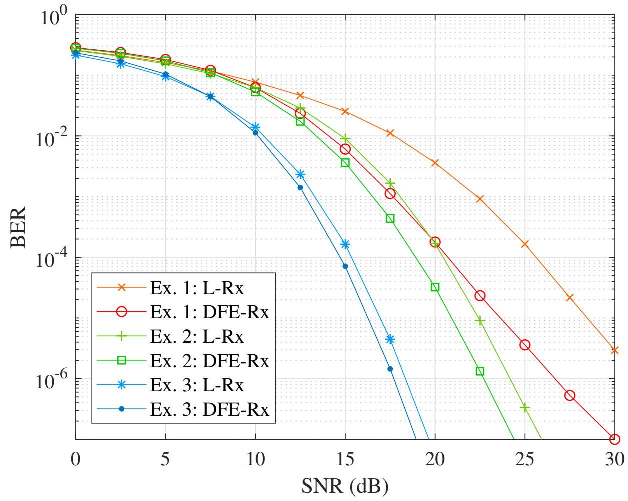

Fig. 4. BER versus SNR performance comparisons for Exs. 1–3.

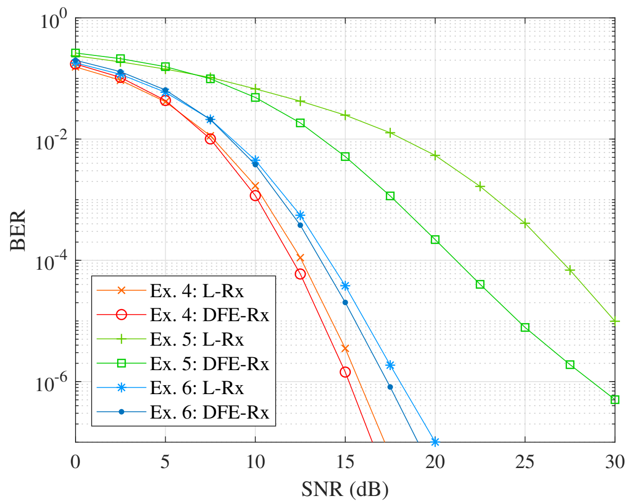

Fig. 5. BER versus SNR performance comparisons for Exs. 4–6.

the performance. Ex. 6, by comparison to Ex. 5, decreases the number of data streams at each source node, consequently, raises the spatial diversity order and improves the performance.

It can be seen from all the examples that, although, caused by the error propagation within the MMSE-DFE receiver, a little BER performance degradation appears in the low range of SNR (i.e., as SNR $\leqslant 7.5$ dB), the DFE-Rx algorithm can always perform better than the L-Rx algorithm in terms of both MSE and BER when SNR becomes moderately high. Such performance superiority is particularly obvious for Ex. 1 and Ex. 5, where there is more than an order of magnitude BER performance improvement as SNR > 17.5 dB.

V. Conclusion

Considering the direct links between source and destination nodes, this paper used the MMSE-DFE receiving technique to enhance the reliability of the signal transmission in a multiuser AF MIMO multi-relay system, whose parameter matrices were jointly optimized by our proposed DFE-Rx algorithm. Through numerical simulations, we verified the superior MSE and BER performance of the DFE-Rx algorithm by comparison with the L-Rx algorithm, which adopts the linear MMSE receiver, and also demonstrated that when direct links have relatively high SNRs, they can notably improve the system reliability.

Acknowledgment

The authors would like to thank Prof. Jiaru Lin and Prof. Kai Niu for their insightful suggestions as well as the anonymous reviewers for their valuable comments.

References

[1] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, “A survey on MIMO transmission with finite input signals: Technical challenges, advances, and future trends,” *Proc. IEEE*, vol. 106, no. 10, pp. 1779–1833, Oct. 2018.

[2] M. A. Girnyk, M. Vehkaperä, and L. K. Rasmussen, “Asymptotic performance analysis of a K-hop amplify-and-forward relay MIMO channel,” *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3532–3546, Jun. 2016.

[3] S. Jin, M. R. McKay, C. Zhong, and K.-K. Wong, “Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems,” *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2204–2224, May 2010.

[4] Y. Rong, X. Tang, and Y. Hua, “A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems,” *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4837–4851, Dec. 2009.

[5] C. Xing, S. Ma, Y.-C. Wu, and T.-S. Ng, “Transceiver design for dual-hop nonregenerative MIMO-OFDM relay systems under channel uncertainties,” *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6325–6339, Dec. 2010.

[6] K.-J. Lee, H. Sung, E. Park, and I. Lee, “Joint optimization for one and two-way MIMO AF multiple-relay systems,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3671–3681, Dec. 2010.

[7] H.-B. Kong, H. M. Shin, T. Oh, and I. Lee, “Joint MMSE transceiver designs for MIMO AF relaying systems with direct link,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3547–3560, Jun. 2017.

[8] Z. He, X. Zhang, Y. Bi, W. Jiang, and Y. Rong, “Optimal source and relay design for multiuser MIMO AF relay communication systems with direct links and imperfect channel information,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2025–2038, Mar. 2016.

[9] Y. Rong, “Optimal linear non-regenerative multi-hop MIMO relays with MMSE-DFE receiver at the destination,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 7, pp. 2268–2279, Jul. 2010.

[10] Y. Lv, Z. He, and Y. Rong, “Multiuser multi-hop AF MIMO relay system design based on MMSE-DFE receiver,” *IEEE Access*, vol. 7, pp. 42 518–42 535, Apr. 2019.

[11] Y. Xu and W. Yin, “A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion,” *SIAM J. Imag. Sci.*, vol. 6, no. 3, pp. 1758–1789, Sep. 2013.

[12] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, “Internet of Things: A survey on enabling technologies, protocols, and applications,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[13] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York, USA: Cambridge Univ. Press, 2013.

[14] A. Hjørungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. Cambridge, UK: Cambridge Univ. Press, 2011.

[15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge Univ. Press, 2004.

[16] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.2,” CVX Research, Inc., Jan. 2020. [Online]. Available: http://cvxr.com/cvx