Joint Optimization for One and Two-Way MIMO AF Multiple-Relay Systems

Kyoung-Jae Lee, *Student Member, IEEE*, Hakjea Sung, *Member, IEEE*, Eunsung Park, *Student Member, IEEE*, and Inkyu Lee, *Senior Member, IEEE*

***Abstract*—This paper considers both one-way and two-way relaying systems with multiple relays between two terminal nodes where all nodes have multiple-input multiple-output (MIMO) antennas. We propose a unified algorithm which computes the optimal linear transceivers jointly at the source node and the relay nodes for amplify-and-forward (AF) protocols. First, optimization designs based on the sum-rate and the mean-square error (MSE) criteria are formulated for the two-way AF relaying channel. Due to non-convexity of the given problems, the proposed schemes iteratively identify local-optimal source and relay filters by deriving the gradients of the cost functions for a gradient descent algorithm. Then, the proposed algorithm can optimize a one-way multiple relay system as a special case of the two-way channel. Finally, we prove the global optimality of the maximum sum-rate scheme under an asymptotically large antenna assumption. From simulation results, it is confirmed that the proposed methods yield the near optimum result for the MIMO multiple relay channel even with a moderate number of antennas. Consequently, we show that the proposed algorithm outperforms conventional schemes in terms of the sum-rate and the error performance for both one-way and two-way protocols.**

***Index Terms*—Amplify-and-forward (AF) relaying, multiple-input multiple-output (MIMO), multiple relays, two-way protocol, maximum sum-rate, minimum mean-square error (MMSE).**

I. INTRODUCTION

WIRELESS relaying transmission is a promising technique which can be applied to extend the coverage or increase the system capacity. For this reason, relay based wireless networks have been studied with a lot of interest. The information theoretic analysis of relay systems has been reported in [1], [2] and [3]. In order to enhance the throughput of the relay network, multiple antennas can be considered for obtaining a similar performance benefit observed in point-to-point multiple-input multiple-output (MIMO) systems [32]. The capacity of the MIMO relay systems has also been studied in [4] and [5].

Manuscript received July 9, 2009; revised June 11, 2010; accepted September 2, 2010. The associate editor coordinating the review of this paper and approving it for publication was C. Papadias.

This research was supported in part by Seoul R&BD Program (ST090852) and in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2010-0017909). The materials in this paper were presented in part at the IEEE VTC, Barcelona, Spain, April 2009 and at the IEEE VTC, Anchorage, USA, September 2009.

K.-J. Lee, E. Park and I. Lee are with the School of Electrical Engineering, Korea University, Seoul, Korea (e-mail: {kyoungjae, kupes2000, inkyu}@korea.ac.kr).

H. Sung is with Samsung Electronics, Suwon, Korea (e-mail: jay.sung@samsung.com).

Digital Object Identifier 10.1109/TWC.2010.102210.091021

Furthermore, the capacity scaling analysis in [6] shows that the capacity of multiple relay systems grows with the number of relay nodes. Some relaying techniques are proposed to achieve a cooperative gain in terms of the capacity and the error performance [7][8]. For MIMO multiple relay systems, the mean-square error (MSE) minimization and the rate maximization problems have been investigated in [9]. Most relay systems are assumed to operate in the half-duplex mode where relay nodes do not transmit and receive signals simultaneously to avoid loop interference in the relay nodes. Such half-duplex relay systems suffer from a substantial loss in terms of spectral efficiency due to the pre-log factor $1/2$, which dominates the capacity at high signal-to-noise ratio (SNR).

Two-way relaying protocol has been proposed recently to overcome such a spectral efficiency loss in the half-duplex one-way system [10]–[13]. In the first channel phase, two terminal nodes transmit their data signals to the relay nodes at the same time. Subsequently in the second channel phase, the relay nodes transmit the processed signal to both terminals. Since each terminal node knows its own transmitted data, the self-interference in the transmitted signals can be eliminated from the received baseband signal [10][12], which is also called as analogue network coding (ANC) [13]. In this way, the two-way protocol achieves a spectral efficiency gain over the one-way protocol at the expense of an increased complexity.

In practical relay systems, an amplify-and-forward (AF) method shows advantages of simple implementation compared to decode-and-forward (DF) systems [14]. For single relay one-way AF protocols, it was shown in [15] and [16] that a singular value decomposition (SVD) technique with power-loading achieves the capacity as in point-to-point MIMO systems. In this case, the ergodic capacity has been analyzed in a large antenna regime in [17]. Also, a relay filter which minimizes the MSE has been solved in [18] through a similar approach. For the case where channel state information (CSI) is available at all nodes, an iterative power-loading scheme has been applied to find the optimal source and relay filters [19].

For systems with either two-way protocols or multiple relays, the SVD-based technique is not applicable since all link channels cannot be simultaneously decomposed. Two-way single relay systems have been investigated in terms of the sum-rate [12][13] and the sum-MSE [20]. In [13], authors have identified the optimal rate region in a relay network with single-input single-output (SISO) terminals. For the case of two-way protocols with multiple relays, the sum-rate capacity of AF channels is achieved by using an iterative power minimization algorithm assuming the complete knowledge of all CSIs, and the matched filter design with a distributed local CSI requirement is proposed [21]. However, the joint optimization algorithm with the full CSI requirement in [21] is still confined in SISO terminals as in [13], and thus any spatial multiplexing gain cannot be obtained.

In this paper, a new design method is proposed to identify the source and relay filters which jointly optimize the sum-rate or the sum-MSE in multiple relay MIMO spatial multiplexing systems for both one-way and two-way protocols. We first formulate the sum-rate maximization and the sum-MSE minimization problem for a two-way relay channel. Since the formulated problems are not generally convex or concave, it cannot be solved analytically. Hence, the proposed schemes iteratively compute the source and relay filter matrices by applying a gradient descent algorithm [22]. In our scheme, we derive a gradient expression of the weighted sum-MSE, and then determine the weighted sum-rate gradients from a relation between the differentials of the sum-MSE and the sum-rate.

Since the one-way relay channel model can be represented as a special case of the formulated two-way channel model, the proposed algorithm can be applied to the one-way multiple relay system. Moreover, exploiting the bit-error-rate (BER) minimization technique introduced in [23], the proposed minimum sum-MSE solutions can be transformed to minimize the BER in MIMO multiple relay systems. The proposed algorithms arrive at local-optimum solutions and are shown to asymptotically approach the global optimum value as the number of source antennas increases. Simulation results demonstrate that the proposed methods obtain the performance very close to the global optimum solution for multiple relay systems even with the moderate number of antennas. Additionally, we show that the proposed schemes with a single iteration outperform conventional schemes in terms of both capacity and error probability for various system configurations.

This paper is organized as follows: Section II describes the system model for two-way MIMO multiple relay channels and formulates the sum-MSE and the sum-rate problems. In Section III, we propose iterative schemes to optimize the sum-MSE or sum-rate performance. The optimality of the proposed scheme is investigated in Section IV. Section V presents the numerical results. Finally, the paper is terminated with conclusions in Section VI.

Throughout this paper, the superscripts $(\cdot)^T$, $(\cdot)^\dagger$ and $(\cdot)^*$ stand for transpose, conjugate transpose, and element-wise conjugate, respectively. $\mathcal{E}(\cdot)$ and $\Re(\cdot)$ denote the expectation and the real component, respectively. $\mathbf{I}_N$ indicates an $N \times N$ identity matrix. $\mathrm{Tr}(\mathbf{A})$, $|\mathbf{A}|$ and $\mathrm{vec}(\mathbf{A})$ represent the trace, the determinant and the stacked columns of a matrix $\mathbf{A}$, respectively, and the Frobenius norm of $\mathbf{A}$ is defined as $\|\mathbf{A}\|_F^2 = \mathrm{Tr}(\mathbf{A}\mathbf{A}^\dagger)$.

II. System Model and Problem Formulation

Consider a relay network where there are two terminal nodes $\mathcal{T}_A$, $\mathcal{T}_B$ and K relay nodes $\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_K$. The

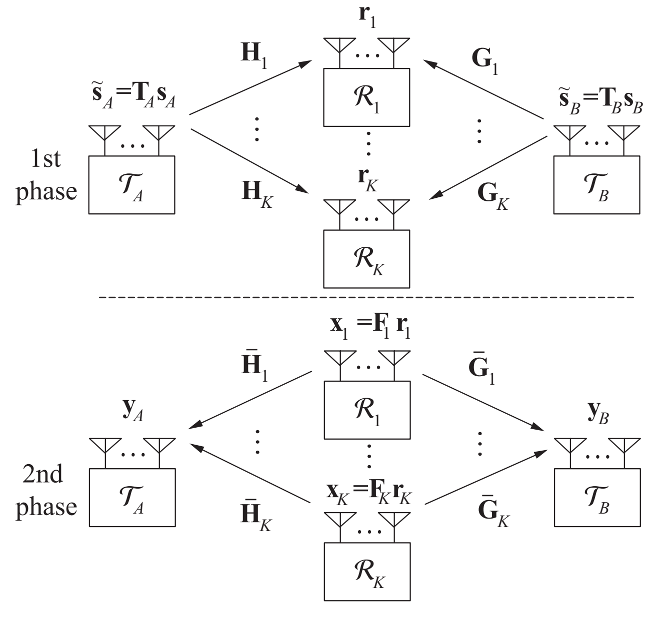


Fig. 1. Channel model of multiple-antenna multiple relay systems in the two-way protocol.

terminal nodes $\mathcal{T}_A$ and $\mathcal{T}_B$ are equipped with M_A and M_B antennas, and the relay nodes $\mathcal{R}_1, \cdots, \mathcal{R}_K$ have $N_1, \cdots, N_K$ antennas, respectively. First we describe the two-way relay channel model as in Fig. 1. Then the one-way relay protocol is considered as a special case of two-way relay systems. In this paper, it is assumed that a direct link between two terminal nodes is ignored due to a large path loss. Also, we assume that both terminals $\mathcal{T}_A$ and $\mathcal{T}_B$ acquire the complete channel knowledge of all MIMO links, and the optimized relay filters are forwarded from either of two terminals to the corresponding relay nodes rather than computed simultaneously at all relay nodes with the full CSI [13][21][1].

In the two-way relay systems, data transmission takes place in two separate channel phases as shown in Fig. 1. In the first channel phase, two terminal nodes $\mathcal{T}_A$ and $\mathcal{T}_B$ precode the signal vectors $\mathbf{s}_A \in \mathbb{C}^{M_A}$ and $\mathbf{s}_B \in \mathbb{C}^{M_B}$ by the source filters $\mathbf{T}_A \in \mathbb{C}^{M_A \times M_A}$ and $\mathbf{T}_B \in \mathbb{C}^{M_B \times M_B}$, where $\mathbf{s}_A$ and $\mathbf{s}_B$ are assumed to have $\mathcal{E}(\mathbf{s}_A \mathbf{s}_A^\dagger) = \mathbf{I}_{M_A}$ and $\mathcal{E}(\mathbf{s}_B \mathbf{s}_B^\dagger) = \mathbf{I}_{M_B}$, respectively. Then two terminal nodes simultaneously transmit the precoded signals $\tilde{\mathbf{s}}_A = \rho_A \mathbf{T}_A \mathbf{s}_A$ and $\tilde{\mathbf{s}}_B = \rho_B \mathbf{T}_B \mathbf{s}_B$ to K relay nodes $\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_K$ where the power normalizing coefficients ρ_A and ρ_B which adjust the transmission power at $\mathcal{T}_A$ and $\mathcal{T}_B$ to P_A and P_B are given as

$$\rho_A = \sqrt{\frac{P_A}{\mathrm{Tr}(\mathbf{T}_A \mathbf{T}_A^\dagger)}} \quad \text{and} \quad \rho_B = \sqrt{\frac{P_B}{\mathrm{Tr}(\mathbf{T}_B \mathbf{T}_B^\dagger)}}. \tag{1}$$

At the k-th relay node, the received signal vector $\mathbf{r}_k$ of

[1]To estimate CSI in two-way multiple relay systems, some training designs in [24], [25] and [26] can be applied. Also, CSI at the transmit phase can be obtained by utilizing channel reciprocity through sounding signals in the time-division duplex (TDD) mode or feedback techniques in the frequency-division duplex (FDD). Although two-way relay systems require more overheads for CSI than one-way protocols, the analysis and design for acquiring the CSI is outside the scope of this paper and will be an interesting topic for future work.

length N_k for $k = 1, 2, \cdots, K$ is obtained as

$$\begin{aligned}\mathbf{r}_k &= \mathbf{H}_k \tilde{\mathbf{s}}_A + \mathbf{G}_k \tilde{\mathbf{s}}_B + \mathbf{n}_k \\ &= \rho_A \mathbf{H}_k \mathbf{T}_A \mathbf{s}_A + \rho_B \mathbf{G}_k \mathbf{T}_B \mathbf{s}_B + \mathbf{n}_k\end{aligned}$$

where $\mathbf{H}_k \in \mathbb{C}^{N_k \times M_A}$ and $\mathbf{G}_k \in \mathbb{C}^{N_k \times M_B}$ are the channel matrices for links $\mathcal{T}_A \rightarrow \mathcal{R}_k$ and $\mathcal{T}_B \rightarrow \mathcal{R}_k$, respectively, and $\mathbf{n}_k$ denotes the additive complex Gaussian noise vector with zero mean and $\mathcal{E}(\mathbf{n}_k \mathbf{n}_k^\dagger) = \sigma_\mathbf{n}^2 \mathbf{I}_{N_k}$.

In the second channel phase, the received signal $\mathbf{r}_k$ is multiplied by the $N_k \times N_k$ relay filter matrix $\mathbf{F}_k$ at the k-th relay node. Then, the transmit signal vector $\mathbf{x}_k \in \mathbb{C}^{N_k}$ at the relay node is computed by

$$\begin{aligned}\mathbf{x}_k &= \gamma_k \mathbf{F}_k \mathbf{r}_k \\ &= \gamma_k \rho_A \mathbf{F}_k \mathbf{H}_k \mathbf{T}_A \mathbf{s}_A + \gamma_k \rho_B \mathbf{F}_k \mathbf{G}_k \mathbf{T}_B \mathbf{s}_B + \gamma_k \mathbf{F}_k \mathbf{n}_k.\end{aligned}$$

Here, γ_k denotes the power normalizing coefficient expressed as

$$\gamma_k = \sqrt{\frac{P_k}{\mathrm{Tr}\left\{\mathbf{F}_k \left(\rho_A^2 \mathbf{H}_k \mathbf{T}_A \mathbf{T}_A^\dagger \mathbf{H}_k^\dagger + \rho_B^2 \mathbf{G}_k \mathbf{T}_B \mathbf{T}_B^\dagger \mathbf{G}_k^\dagger + \sigma_\mathbf{n}^2 \mathbf{I}_{N_k}\right) \mathbf{F}_k^\dagger\right\}}} \quad (2)$$

where P_k indicates the transmission power of the k-th relay node.

Now, the processed signals $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K$ are transmitted from the relay nodes to two terminal nodes $\mathcal{T}_A$ and $\mathcal{T}_B$. We denote the MIMO channel matrices for links $\mathcal{R}_k \rightarrow \mathcal{T}_A$ and $\mathcal{R}_k \rightarrow \mathcal{T}_B$ at the second channel phase as $\bar{\mathbf{H}}_k \in \mathbb{C}^{M_A \times N_k}$ and $\bar{\mathbf{G}}_k \in \mathbb{C}^{M_B \times N_k}$, respectively. Then, the received signals at two terminal nodes $\mathbf{y}_A \in \mathbb{C}^{M_A}$ and $\mathbf{y}_B \in \mathbb{C}^{M_B}$ can be written as

$$\begin{aligned}\mathbf{y}_A = {} & \rho_A \sum_{k=1}^{K} \gamma_k \bar{\mathbf{H}}_k \mathbf{F}_k \mathbf{H}_k \mathbf{T}_A \mathbf{s}_A + \rho_B \sum_{k=1}^{K} \gamma_k \bar{\mathbf{H}}_k \mathbf{F}_k \mathbf{G}_k \mathbf{T}_B \mathbf{s}_B \\ & + \sum_{k=1}^{K} \gamma_k \bar{\mathbf{H}}_k \mathbf{F}_k \mathbf{n}_k + \mathbf{z}_A \end{aligned} \quad (3)$$

and

$$\begin{aligned}\mathbf{y}_B = {} & \rho_A \sum_{k=1}^{K} \gamma_k \bar{\mathbf{G}}_k \mathbf{F}_k \mathbf{H}_k \mathbf{T}_A \mathbf{s}_A + \rho_B \sum_{k=1}^{K} \gamma_k \bar{\mathbf{G}}_k \mathbf{F}_k \mathbf{G}_k \mathbf{T}_B \mathbf{s}_B \\ & + \sum_{k=1}^{K} \gamma_k \bar{\mathbf{G}}_k \mathbf{F}_k \mathbf{n}_k + \mathbf{z}_B \end{aligned} \quad (4)$$

where $\mathbf{z}_A \in \mathbb{C}^{M_A}$ and $\mathbf{z}_B \in \mathbb{C}^{M_B}$ are complex white Gaussian noise vectors with zero mean and variance $\sigma_\mathbf{z}^2$. In (3) and (4), $\rho_A \sum_{k=1}^{K} \gamma_k \bar{\mathbf{H}}_k \mathbf{F}_k \mathbf{H}_k \mathbf{T}_A \mathbf{s}_A$ and $\rho_B \sum_{k=1}^{K} \gamma_k \bar{\mathbf{G}}_k \mathbf{F}_k \mathbf{G}_k \mathbf{T}_B \mathbf{s}_B$ represent the back-propagating *self-interferences*. Since terminal nodes know their own symbols transmitted at the previous channel phase as well as the corresponding effective channels, these self-interference terms can be canceled when received at both terminals.

After removing the self-interferences, the received signals $\hat{\mathbf{y}}_A$ and $\hat{\mathbf{y}}_B$ are then given as

$$\hat{\mathbf{y}}_A = \rho_B \sum_{k=1}^{K} \gamma_k \bar{\mathbf{H}}_k \mathbf{F}_k \mathbf{G}_k \mathbf{T}_B \mathbf{s}_B + \sum_{k=1}^{K} \gamma_k \bar{\mathbf{H}}_k \mathbf{F}_k \mathbf{n}_k + \mathbf{z}_A \quad (5)$$

and

$$\hat{\mathbf{y}}_B = \rho_A \sum_{k=1}^{K} \gamma_k \bar{\mathbf{G}}_k \mathbf{F}_k \mathbf{H}_k \mathbf{T}_A \mathbf{s}_A + \sum_{k=1}^{K} \gamma_k \bar{\mathbf{G}}_k \mathbf{F}_k \mathbf{n}_k + \mathbf{z}_B. \quad (6)$$

It should be noted that in the one-way protocol, the channel model for the communication of $\mathcal{T}_A \rightarrow \mathcal{T}_B$ is equivalent to (6) with $P_B = 0$ (i.e., $\rho_B = 0$). Hence, a one-way system can be considered as a special case of two-way systems with the self-interference canceling process.

Now, we will describe the problem formulations. Our first main goal is to jointly minimize the sum-MSE of the bidirectional links (5) and (6) at the source, relay and destination nodes. Here, we consider the linear decoders $\mathbf{D}_A$ and $\mathbf{D}_B$ are applied at the terminal nodes $\mathcal{T}_A$ and $\mathcal{T}_B$ in the second channel phase, respectively. Given specific source and relay filters $\mathbf{T}_A, \mathbf{T}_B, \mathbf{F}_1, \cdots, \mathbf{F}_K$, the optimal linear decoder $\mathbf{D}_l$ at $\mathcal{T}_l$ for $l = A$ and B can be independently expressed by the minimum MSE (MMSE) receiver as [18]

$$\hat{\mathbf{D}}_l = \rho_{\underline{l}} \mathbf{\Sigma}_{\underline{l}}^\dagger \left(\rho_{\underline{l}}^2 \mathbf{\Sigma}_{\underline{l}} \mathbf{\Sigma}_{\underline{l}}^\dagger + \mathbf{\Omega}_{\underline{l}}^{-1}\right)^{-1} \quad (7)$$

where the subscript $\underline{l}$ indicates $\underline{l} = B$ if $l = A$ and $\underline{l} = A$ if $l = B$, and $\mathbf{\Sigma}_l$ and $\mathbf{\Omega}_l$ for $l = A$ and B are defined as

$$\mathbf{\Sigma}_A \triangleq \sum_{k=1}^{K} \gamma_k \bar{\mathbf{G}}_k \mathbf{F}_k \mathbf{H}_k \mathbf{T}_A, \quad \mathbf{\Sigma}_B \triangleq \sum_{k=1}^{K} \gamma_k \bar{\mathbf{H}}_k \mathbf{F}_k \mathbf{G}_k \mathbf{T}_B,$$

and

$$\mathbf{\Omega}_A \triangleq \left(\sigma_\mathbf{n}^2 \sum_{k=1}^{K} \gamma_k^2 \bar{\mathbf{G}}_k \mathbf{F}_k \mathbf{F}_k^\dagger \bar{\mathbf{G}}_k^\dagger + \sigma_\mathbf{z}^2 \mathbf{I}_{M_B}\right)^{-1},$$

$$\mathbf{\Omega}_B \triangleq \left(\sigma_\mathbf{n}^2 \sum_{k=1}^{K} \gamma_k^2 \bar{\mathbf{H}}_k \mathbf{F}_k \mathbf{F}_k^\dagger \bar{\mathbf{H}}_k^\dagger + \sigma_\mathbf{z}^2 \mathbf{I}_{M_A}\right)^{-1}.$$

Employing this MMSE linear receiver in (7) at the destination, the MSE matrix $\mathbf{\Pi}_l$ for the signal $\mathbf{s}_l$ for $l = A$ and B is written as

$$\begin{aligned}\mathbf{\Pi}_l &= \mathcal{E}\left\{\left(\hat{\mathbf{D}}_{\underline{l}} \hat{\mathbf{y}}_{\underline{l}} - \mathbf{s}_l\right)\left(\hat{\mathbf{D}}_{\underline{l}} \hat{\mathbf{y}}_{\underline{l}} - \mathbf{s}_l\right)^\dagger\right\} \\ &= \left(\mathbf{I}_{M_l} + \rho_l^2 \mathbf{\Sigma}_l^\dagger \mathbf{\Omega}_l \mathbf{\Sigma}_l\right)^{-1}. \end{aligned} \quad (8)$$

It should be noted that ρ_A and ρ_B are given as a function of the source filters $\mathbf{T}_A$ and $\mathbf{T}_B$ in (1), and $\gamma_1, \gamma_2, \cdots, \gamma_K$ in (2) are determined by the source and the relay filter matrices $\mathbf{T}_A, \mathbf{T}_B, \mathbf{F}_1, \cdots, \mathbf{F}_K$.

Then, the joint minimization problem of the weighted sum-MSE can be formulated as

$$\{\mathbf{T}_A^{\text{mse}}, \mathbf{T}_B^{\text{mse}}, \mathbf{F}_1^{\text{mse}}, \cdots, \mathbf{F}_K^{\text{mse}}\} = \underset{\mathbf{T}_A, \mathbf{T}_B, \mathbf{F}_1, \cdots, \mathbf{F}_K}{\arg \min} \Xi_{\text{sum}} \quad (9)$$

where Ξ_{sum} is defined by

$$\Xi_{\text{sum}} \triangleq \mathrm{Tr}\left(\mathbf{W}_A \mathbf{\Pi}_A\right) + \mathrm{Tr}\left(\mathbf{W}_B \mathbf{\Pi}_B\right), \quad (10)$$

and $\mathbf{W}_A \in \mathbb{C}^{M_A \times M_A}$ and $\mathbf{W}_B \in \mathbb{C}^{M_B \times M_B}$ denote constant weight matrices associated with the links $\mathcal{T}_A \rightarrow \mathcal{T}_B$ and $\mathcal{T}_B \rightarrow \mathcal{T}_A$, respectively. Here, the weight matrices are introduced to express the connection between the gradients of the sum-MSE and the sum-rate which will be derived in the next section.

$$\begin{aligned} d\Xi_{\text{sum}} = -\Big[& \rho_A^2\gamma_k \text{vec}\left\{\bar{\mathbf{G}}_k^\dagger \mathbf{\Omega}_A \mathbf{\Sigma}_A \mathbf{\Pi}_A \mathbf{W}_A \mathbf{\Pi}_A (\mathbf{T}_A^\dagger \mathbf{H}_k^\dagger - \sigma_{\mathbf{n}}^2 \gamma_k \mathbf{\Sigma}_A^\dagger \mathbf{\Omega}_A \bar{\mathbf{G}}_k \mathbf{F}_k)\right\}^T \\ & + \rho_B^2\gamma_k \text{vec}\left\{\bar{\mathbf{H}}_k^\dagger \mathbf{\Omega}_B \mathbf{\Sigma}_B \mathbf{\Pi}_B \mathbf{W}_B \mathbf{\Pi}_B (\mathbf{T}_B^\dagger \mathbf{G}_k^\dagger - \sigma_{\mathbf{n}}^2 \gamma_k \mathbf{\Sigma}_B^\dagger \mathbf{\Omega}_B \bar{\mathbf{H}}_k \mathbf{F}_k)\right\}^T \\ & + \frac{\gamma_k^3}{P_k} \Re\Big(\rho_A^2 \text{Tr}\left\{\mathbf{\Pi}_A \mathbf{W}_A \mathbf{\Pi}_A \mathbf{\Sigma}_A^\dagger \mathbf{\Omega}_A \bar{\mathbf{G}}_k \mathbf{F}_k (\sigma_{\mathbf{n}}^2 \gamma_k \mathbf{F}_k^\dagger \bar{\mathbf{G}}_k^\dagger \mathbf{\Omega}_A \mathbf{\Sigma}_A - \mathbf{H}_k \mathbf{T}_A)\right\} \\ & \quad + \rho_B^2 \text{Tr}\left\{\mathbf{\Pi}_B \mathbf{W}_B \mathbf{\Pi}_B \mathbf{\Sigma}_B^\dagger \mathbf{\Omega}_B \bar{\mathbf{H}}_k \mathbf{F}_k (\sigma_{\mathbf{n}}^2 \gamma_k \mathbf{F}_k^\dagger \bar{\mathbf{H}}_k^\dagger \mathbf{\Omega}_B \mathbf{\Sigma}_B - \mathbf{G}_k \mathbf{T}_B)\right\} \Big) \\ & \quad \times \text{vec}\left\{\mathbf{F}_k (\rho_A^2 \mathbf{H}_k \mathbf{T}_A \mathbf{T}_A^\dagger \mathbf{H}_k^\dagger + \rho_B^2 \mathbf{G}_k \mathbf{T}_B \mathbf{T}_B^\dagger \mathbf{G}_k^\dagger + \sigma_{\mathbf{n}}^2 \mathbf{I}_{N_k})\right\}^T \Big] d\text{vec}(\mathbf{F}_k^*) \end{aligned} \tag{11}$$

Note that when using identity weight matrices, the sum-MSE minimization problem is obtained.

The second criterion which we try to optimize is the weighted sum-rate of the bidirectional links in (5) and (6). Assuming complex Gaussian signaling, the weighted sum-rate maximizing problem can be expressed by [15]

$$\{\mathbf{T}_A^{\text{rate}}, \mathbf{T}_B^{\text{rate}}, \mathbf{F}_1^{\text{rate}}, \cdots, \mathbf{F}_K^{\text{rate}}\} = \underset{\mathbf{T}_A, \mathbf{T}_B, \mathbf{F}_1, \cdots, \mathbf{F}_K}{\arg\max} R_{\text{sum}} \tag{12}$$

where we define R_{sum} as

$$R_{\text{sum}} \triangleq \frac{w_A}{2}\log_2\left|\mathbf{\Pi}_A^{-1}\right| + \frac{w_B}{2}\log_2\left|\mathbf{\Pi}_B^{-1}\right|, \tag{13}$$

and w_A and w_B are nonnegative weight constants. Here the pre-log factor $1/2$ is introduced for the half-duplex mode.

In [19], specific cases of our formulated problems (9) and (12) are investigated for optimizing one-way single relay systems ($K = 1$ and $P_B = 0$), where the SVD is used to decompose the effective channel into parallel subchannels and the power distribution on the subchannels is allocated alternately in the source and the relay node [19]. However, for a multiple relay channel, this SVD-based technique cannot be applied since cooperation among separated relay nodes is impossible. Furthermore, in the two-way relay protocol, the SVD cannot decompose both bidirectional channels in (5) and (6) simultaneously into diagonal channels. In the following section, we propose a method which can generate solutions for this problem for general one- and two-way MIMO systems with multiple relay nodes.

III. Proposed Optimization Schemes for One- and Two-Way Multiple Relays

In this section, we propose a linear processing design for solving the problems formulated in Section II for one-way and two-way multiple relay MIMO systems. First, the design of the source and relay filters for minimizing the sum-MSE or maximizing the sum-rate is investigated for two-way MIMO multiple relay links in (9) and (12). Then one-way protocols are described as a special case of two-way systems.

Both optimization problems in (9) and (12) are difficult to solve analytically since the cost functions Ξ_{sum} and R_{sum} are not generally a convex or concave function with respect to $\mathbf{T}_A, \mathbf{T}_B, \mathbf{F}_1, \cdots, \mathbf{F}_K$. Hence, we apply a gradient descent algorithm to solve the unconstrained optimization problems (9) and (12). After deriving the gradient expressions of the cost functions in (10) and (13), the proposed optimization algorithm is presented which iteratively searches for the optimized filters. Also, the optimization of the BER is constructed based on the proposed algorithm.

A. Gradient Expressions of Sum-MSE and Sum-Rate

We first derive gradient expressions of the sum-MSE Ξ_{sum} in (10) with given weights of $\mathbf{W}_A$ and $\mathbf{W}_B$. Then, by determining the weight matrices $\mathbf{W}_A$ and $\mathbf{W}_B$ from the relation between the MSE and the mutual information, the gradients of the weighted sum-rate R_{sum} are obtained.

Since the weighted sum-MSE Ξ_{sum} in (10) is a real-valued function, the gradient is written as $\nabla_{\mathbf{X}}\Xi_{\text{sum}} = 2\frac{\partial \Xi_{\text{sum}}}{\partial \mathbf{X}^*}$ [27]. To obtain the derivative of the weighted sum-MSE function $\frac{\partial \Xi_{\text{sum}}}{\partial \mathbf{F}_k^*}$, we first compute the differential of the function. Using some rules for the differential $d\text{Tr}(\mathbf{Y}) = \text{Tr}(d\mathbf{Y})$ and $d\mathbf{Y}^{-1} = -\mathbf{Y}^{-1}d\mathbf{Y}\mathbf{Y}^{-1}$ [28], the partial differential of the weighted sum-MSE with respect to $\mathbf{F}_k^*$ is given by

$$\begin{aligned} d\Xi_{\text{sum}} = & -\text{Tr}\left\{\mathbf{W}_A \mathbf{\Pi}_A d\left(\rho_A^2 \mathbf{\Sigma}_A^\dagger \mathbf{\Omega}_A \mathbf{\Sigma}_A\right)\mathbf{\Pi}_A\right\} \\ & -\text{Tr}\left\{\mathbf{W}_B \mathbf{\Pi}_B d\left(\rho_B^2 \mathbf{\Sigma}_B^\dagger \mathbf{\Omega}_B \mathbf{\Sigma}_B\right)\mathbf{\Pi}_B\right\}. \end{aligned} \tag{14}$$

Note that $d\mathbf{F}_k = 0$ because $\frac{\partial \mathbf{A}\mathbf{F}_k\mathbf{B}}{\partial \mathbf{F}_k^*} = 0$ for arbitrary matrices $\mathbf{A}$ and $\mathbf{B}$. Thus it follows $d\mathbf{\Sigma}_A^\dagger = d\gamma_k \mathbf{T}_A^\dagger \mathbf{H}_k^\dagger \mathbf{F}_k^\dagger \bar{\mathbf{G}}_k^\dagger + \gamma_k \mathbf{T}_A^\dagger \mathbf{H}_k^\dagger d\mathbf{F}_k^\dagger \bar{\mathbf{G}}_k^\dagger$, $d(\mathbf{\Omega}_A^{-1}) = \sigma_{\mathbf{n}}^2 d\gamma_k^2 \bar{\mathbf{G}}_k \mathbf{F}_k \mathbf{F}_k^\dagger \bar{\mathbf{G}}_k^\dagger + \sigma_{\mathbf{n}}^2 \gamma_k^2 \bar{\mathbf{G}}_k \mathbf{F}_k d\mathbf{F}_k^\dagger \bar{\mathbf{G}}_k^\dagger$ and $d\mathbf{\Sigma}_A = d\gamma_k \bar{\mathbf{G}}_k \mathbf{F}_k \mathbf{H}_k \mathbf{T}_A$ where $d\gamma_k = -\frac{\gamma_k^3}{2P_k}\text{Tr}\{\mathbf{F}_k(\rho_A^2 \mathbf{H}_k \mathbf{T}_A \mathbf{T}_A^\dagger \mathbf{H}_k^\dagger + \rho_B^2 \mathbf{G}_k \mathbf{T}_B \mathbf{T}_B^\dagger \mathbf{G}_k^\dagger + \sigma_{\mathbf{n}}^2 \mathbf{I}_{N_k}) d\mathbf{F}_k^\dagger\}$ and $d\gamma_k^2 = 2\gamma_k d\gamma_k$. Here γ_k is a function of $\mathbf{F}_k^*$ through (2).

After some matrix manipulations using the above results, the partial differential in (14) is derived as (11) at the top of this page. This result can be verified by the properties $\text{Tr}(\mathbf{X}^T\mathbf{Y}) = \text{vec}(\mathbf{X})^T\text{vec}(\mathbf{Y})$, $\text{Tr}(\mathbf{X}^T\mathbf{Y}) = \text{Tr}(\mathbf{X}\mathbf{Y}^T)$ and $\text{Tr}(2\mathbf{Z} + \mathbf{Y} + \mathbf{Y}^\dagger) = 2\Re\{\text{Tr}(\mathbf{Z} + \mathbf{Y})\}$ for a Hermitian matrix $\mathbf{Z}$. Here, the coefficients of $d\text{vec}(\mathbf{F}_k^*)$ in (11) directly lead to the derivative $\partial\Xi_{\text{sum}}/\partial\mathbf{F}_k^*$.

Thus, with respect to the k-th relay filter $\mathbf{F}_k$, the gradient of the weighted sum-MSE $\nabla_{\mathbf{F}_k}\Xi_{\text{sum}}$ is given as

$$\begin{aligned} \nabla_{\mathbf{F}_k}\Xi_{\text{sum}} = & -2\mathbf{C}_k + 2\frac{\gamma_k^2}{P_k}\Re\left\{\text{Tr}\left(\mathbf{C}_k \mathbf{F}_k^\dagger\right)\right\}\mathbf{F}_k \\ & \times\left(\rho_A^2 \mathbf{H}_k \mathbf{T}_A \mathbf{T}_A^\dagger \mathbf{H}_k^\dagger + \rho_B^2 \mathbf{G}_k \mathbf{T}_B \mathbf{T}_B^\dagger \mathbf{G}_k^\dagger + \sigma_{\mathbf{n}}^2 \mathbf{I}_{N_k}\right) \end{aligned} \tag{15}$$

where $\mathbf{C}_k$ denotes

$$\mathbf{C}_k = \rho_A^2\gamma_k\bar{\mathbf{G}}_k^\dagger\boldsymbol{\Omega}_A\boldsymbol{\Sigma}_A\boldsymbol{\Pi}_A\mathbf{W}_A\boldsymbol{\Pi}_A(\mathbf{T}_A^\dagger\mathbf{H}_k^\dagger - \sigma_{\mathbf{n}}^2\gamma_k\boldsymbol{\Sigma}_A^\dagger\boldsymbol{\Omega}_A\bar{\mathbf{G}}_k\mathbf{F}_k) + \rho_B^2\gamma_k\bar{\mathbf{H}}_k^\dagger\boldsymbol{\Omega}_B\boldsymbol{\Sigma}_B\boldsymbol{\Pi}_B\mathbf{W}_B\boldsymbol{\Pi}_B(\mathbf{T}_B^\dagger\mathbf{G}_k^\dagger - \sigma_{\mathbf{n}}^2\gamma_k\boldsymbol{\Sigma}_B^\dagger\boldsymbol{\Omega}_B\bar{\mathbf{H}}_k\mathbf{F}_k).$$

Now, considering the partial differential with respect to $\mathbf{T}_A^*$ in (14), we obtain $d\rho_A^2 = -\frac{\rho_A^4}{P_A}\text{Tr}(\mathbf{T}_A d\mathbf{T}_A^\dagger)$, $d\boldsymbol{\Sigma}_A^\dagger = \sum_{k=1}^K d\gamma_k\mathbf{T}_A^\dagger\mathbf{H}_k^\dagger\mathbf{F}_k^\dagger\bar{\mathbf{G}}_k^\dagger + \gamma_k d\mathbf{T}_A^\dagger\mathbf{H}_k^\dagger\mathbf{F}_k^\dagger\bar{\mathbf{G}}_k^\dagger$, $d(\boldsymbol{\Omega}_A^{-1}) = \sigma_{\mathbf{n}}^2\sum_{k=1}^K d\gamma_k^2\bar{\mathbf{G}}_k\mathbf{F}_k\mathbf{F}_k^\dagger\bar{\mathbf{G}}_k^\dagger$ and $d\boldsymbol{\Sigma}_A = \sum_{k=1}^K d\gamma_k\bar{\mathbf{G}}_k\mathbf{F}_k\mathbf{H}_k\mathbf{T}_A$ where $d\gamma_k = \frac{\rho_A^2\gamma_k^3}{2P_AP_k}\text{Tr}[\{\rho_A^2\text{Tr}(\mathbf{F}_k\mathbf{H}_k\mathbf{T}_A\mathbf{T}_A^\dagger\mathbf{H}_A^\dagger\mathbf{F}_k^\dagger)\mathbf{I}_{M_A} - P_A\mathbf{H}_k^\dagger\mathbf{F}_k^\dagger\mathbf{F}_k\mathbf{H}_k\}\mathbf{T}_A d\mathbf{T}_A^\dagger]$ and $d\gamma_k^2 = 2\gamma_k d\gamma_k$. Then, the partial differential of the weighted sum-MSE with respect to $\mathbf{T}_A^*$ is derived as (16) at the top of the next page. From the coefficients of $d\text{vec}(\mathbf{T}_A^*)$ in (16), the gradient with respect to $\mathbf{T}_A$ is computed. Also, we can find the gradient for $\mathbf{T}_B$ using the symmetry between $\mathbf{T}_A$ and $\mathbf{T}_B$.

Hence, the gradient of the weighted sum-MSE Ξ_{sum} for the transmit filter $\mathbf{T}_l$ for $l = A$ and B is expressed by

$$\nabla_{\mathbf{T}_l}\Xi_{\text{sum}} = -2\rho_l^2\tilde{\boldsymbol{\Sigma}}_l^\dagger\boldsymbol{\Omega}_l\boldsymbol{\Sigma}_l\boldsymbol{\Pi}_l\mathbf{W}_l\boldsymbol{\Pi}_l + \frac{2\rho_l^4}{P_l}\text{Tr}(\boldsymbol{\Sigma}_l^\dagger\boldsymbol{\Omega}_l\boldsymbol{\Sigma}_l\boldsymbol{\Pi}_l\mathbf{W}_l\boldsymbol{\Pi}_l)\mathbf{T}_l + \frac{2\rho_l^2}{P_l}\sum_{k=1}^K\frac{\gamma_k^2}{P_k}\Re\{\text{Tr}(\mathbf{C}_k\mathbf{F}_k^\dagger)\}\{P_l\boldsymbol{\Gamma}_{l,k} - \rho_l^2\text{Tr}(\mathbf{T}_l^\dagger\boldsymbol{\Gamma}_{l,k}\mathbf{T}_l)\mathbf{I}_{M_l}\}\mathbf{T}_l \tag{17}$$

where $\tilde{\boldsymbol{\Sigma}}_l$ and $\boldsymbol{\Gamma}_{l,k}$ for $l = A$ and B are denoted by

$$\tilde{\boldsymbol{\Sigma}}_A = \sum_{k=1}^K\gamma_k\bar{\mathbf{G}}_k\mathbf{F}_k\mathbf{H}_k \quad\text{and}\quad \tilde{\boldsymbol{\Sigma}}_B = \sum_{k=1}^K\gamma_k\bar{\mathbf{H}}_k\mathbf{F}_k\mathbf{G}_k,$$

and

$$\boldsymbol{\Gamma}_{A,k} = \mathbf{H}_k^\dagger\mathbf{F}_k^\dagger\mathbf{F}_k\mathbf{H}_k \quad\text{and}\quad \boldsymbol{\Gamma}_{B,k} = \mathbf{G}_k^\dagger\mathbf{F}_k^\dagger\mathbf{F}_k\mathbf{G}_k.$$

Now, we derive the gradients of the weighted sum-rate. From the differential rule $d\ln|\mathbf{Y}| = \text{Tr}(\mathbf{Y}^{-1}d\mathbf{Y})$ [28], the partial differential of the weighted sum-rate in (13) for $\mathbf{F}_k^*$ or $\mathbf{T}_l^*$ is written by

$$dR_{\text{sum}} = \frac{w_A}{2\ln 2}\text{Tr}\{\boldsymbol{\Pi}_A d(\rho_A^2\boldsymbol{\Sigma}_A^\dagger\boldsymbol{\Omega}_A\boldsymbol{\Sigma}_A)\} + \frac{w_B}{2\ln 2}\text{Tr}\{\boldsymbol{\Pi}_B d(\rho_B^2\boldsymbol{\Sigma}_B^\dagger\boldsymbol{\Omega}_B\boldsymbol{\Sigma}_B)\}.$$

Comparing this result with (14), we can find that the differential of the weighted sum-rate dR_{sum} is equivalent to the differential of the weighted sum-MSE $d\Xi_{\text{sum}}$ if the weight matrices are chosen as $\mathbf{W}_A = -\frac{w_A}{2\ln 2}\boldsymbol{\Pi}_A^{-1}$ and $\mathbf{W}_B = -\frac{w_B}{2\ln 2}\boldsymbol{\Pi}_B^{-1}$ for a given set of the source and relay filters $\mathbf{T}_A, \mathbf{T}_B, \mathbf{F}_1, \cdots, \mathbf{F}_K$.

Thus, substituting $\mathbf{W}_A = -\frac{w_A}{2\ln 2}\boldsymbol{\Pi}_A^{-1}$ and $\mathbf{W}_B = -\frac{w_B}{2\ln 2}\boldsymbol{\Pi}_B^{-1}$ in (15) and (17), we get the gradients of the sum-rate R_{sum} with respect to $\mathbf{F}_k$ and $\mathbf{T}_l$ as

$$\nabla_{\mathbf{F}_k}R_{\text{sum}} = \tilde{\mathbf{C}}_k - \frac{\gamma_k^2}{P_k}\Re\left\{\text{Tr}\left(\tilde{\mathbf{C}}_k\mathbf{F}_k^\dagger\right)\right\}\mathbf{F}_k \times\left(\rho_A^2\mathbf{H}_k\mathbf{T}_A\mathbf{T}_A^\dagger\mathbf{H}_k^\dagger + \rho_B^2\mathbf{G}_k\mathbf{T}_B\mathbf{T}_B^\dagger\mathbf{G}_k^\dagger + \sigma_{\mathbf{n}}^2\mathbf{I}_{N_k}\right) \tag{18}$$

and

$$\nabla_{\mathbf{T}_l}R_{\text{sum}} = \frac{w_l\rho_l^2}{\ln 2}\tilde{\boldsymbol{\Sigma}}_l^\dagger\boldsymbol{\Omega}_l\boldsymbol{\Sigma}_l\boldsymbol{\Pi}_l - \frac{w_l\rho_l^4}{P_l\ln 2}\text{Tr}(\boldsymbol{\Sigma}_l^\dagger\boldsymbol{\Omega}_l\boldsymbol{\Sigma}_l\boldsymbol{\Pi}_l)\mathbf{T}_l - \frac{\rho_l^2}{P_l}\sum_{k=1}^K\frac{\gamma_k^2}{P_k}\Re\{\text{Tr}(\tilde{\mathbf{C}}_k\mathbf{F}_k^\dagger)\}\{P_l\boldsymbol{\Gamma}_{l,k} - \rho_l^2\text{Tr}(\mathbf{T}_l^\dagger\boldsymbol{\Gamma}_{l,k}\mathbf{T}_l)\mathbf{I}_{M_l}\}\mathbf{T}_l \tag{19}$$

where $\tilde{\mathbf{C}}_k$ is defined as

$$\tilde{\mathbf{C}}_k = \frac{\gamma_k}{\ln 2}\Big[w_A\rho_A^2\bar{\mathbf{G}}_k^\dagger\boldsymbol{\Omega}_A\boldsymbol{\Sigma}_A\boldsymbol{\Pi}_A(\mathbf{T}_A^\dagger\mathbf{H}_k^\dagger - \sigma_{\mathbf{n}}^2\gamma_k\boldsymbol{\Sigma}_A^\dagger\boldsymbol{\Omega}_A\bar{\mathbf{G}}_k\mathbf{F}_k) + w_B\rho_B^2\bar{\mathbf{H}}_k^\dagger\boldsymbol{\Omega}_B\boldsymbol{\Sigma}_B\boldsymbol{\Pi}_B(\mathbf{T}_B^\dagger\mathbf{G}_k^\dagger - \sigma_{\mathbf{n}}^2\gamma_k\boldsymbol{\Sigma}_B^\dagger\boldsymbol{\Omega}_B\bar{\mathbf{H}}_k\mathbf{F}_k)\Big].$$

Consequently, the gradient of the weighted sum-MSE with the proper weight matrices is equivalent to the gradient of the sum-rate. Note that this relation between the sum-MSE and the mutual information in MIMO relay networks agrees with the discussion made in [29].

B. Proposed Iterative Algorithm

The proposed optimization method exploits a fact that a real-valued cost function $C(\mathbf{X})$ with a decision variable matrix $\mathbf{X}$ decreases fastest when an arbitrary point $\mathbf{X}_o$ moves in a direction of $-\nabla_{\mathbf{X}}C(\mathbf{X}_o)$ where $\nabla_{\mathbf{X}}C(\mathbf{X}_o)$ indicates the gradient of $C(\mathbf{X})$ at the point $\mathbf{X}_o$ [22]. With the derived gradient expressions of the cost functions in (10) and (13), the proposed algorithm iteratively searches for the optimized filters. By considering the cost function C as either $C = \Xi_{\text{sum}}$ or $C = -R_{\text{sum}}$, we can solve the problem (9) or (12) using the same iterative algorithm which is summarized below.[2]

1) Initialize $\mathbf{T}_A, \mathbf{T}_B, \mathbf{F}_1, \cdots, \mathbf{F}_K$
2) for $k = 1, \cdots, K$
   Calculate the gradient $\nabla_k \triangleq \nabla_{\mathbf{F}_k}C$ from (15) or (18)
   Update $\mathbf{F}_k \leftarrow \mathbf{F}_k - \delta_k\nabla_k$
   end
3) for $l = A, B$
   Calculate the gradient $\nabla_l \triangleq \nabla_{\mathbf{T}_l}C$ from (17) or (19)
   Update $\mathbf{T}_l \leftarrow \mathbf{T}_l - \delta_l\nabla_l$
   end
4) If $\sum_{k=1}^K\|\nabla_k\|_F^2 + \sum_{l=A}^B\|\nabla_l\|_F^2 < \epsilon$, stop the iteration
   Otherwise go back to step 2)

In this algorithm, ϵ is the tolerance factor for terminating the iteration.

Several line search methods are introduced in [22] to efficiently determine the step sizes δ_k and δ_l. We employ a method called Armijo's rule which provides provable convergence [22] by setting $\delta_k = \nu^m$ and $\delta_k = \nu^l$ where m and l are the smallest integers such that $C(\mathbf{F}_k) - C(\mathbf{F}_k - \nu^m\nabla_k) \geq \varepsilon\,\nu^m\|\nabla_k\|_F^2$ and $C(\mathbf{T}_l) - C(\mathbf{T}_l -$

[2] In the proposed algorithm, the step 3) for optimizing the source filters can be replaced by the point-to-point MIMO SVD-based beamforming technique in [23] in order to reduce the processing complexity. However, such a source beamforming method still requires an iterative process for jointly optimizing the source and the relay filters since the SVD cannot diagonalize the MIMO channels at the relay nodes.

$$\begin{aligned} d\Xi_{\text{sum}} = -\Big[\ & \rho_A^2 \text{vec}\left(\sum_{k=1}^{K}\gamma_k \bar{\mathbf{G}}_k\mathbf{F}_k\mathbf{H}_k\mathbf{\Omega}_A\mathbf{\Sigma}_A\mathbf{\Pi}_A\mathbf{W}_A\mathbf{\Pi}_A\right)^T - \frac{\rho_A^4}{P_A}\text{Tr}\left(\mathbf{\Sigma}_A^\dagger\mathbf{\Omega}_A\mathbf{\Sigma}_A\mathbf{\Pi}_A\mathbf{W}_A\mathbf{\Pi}_A\right)\text{vec}(\mathbf{T}_A)^T \\ & + \frac{\rho_A^2}{P_A}\sum_{k=1}^{K}\frac{\gamma_k^3}{P_k}\Re\Big(\rho_A^2\text{Tr}\left\{\mathbf{\Pi}_A\mathbf{W}_A\mathbf{\Pi}_A\mathbf{\Sigma}_A^\dagger\mathbf{\Omega}_A\bar{\mathbf{G}}_k\mathbf{F}_k(\sigma_{\mathbf{n}}^2\gamma_k\mathbf{F}_k^\dagger\bar{\mathbf{G}}_k^\dagger\mathbf{\Omega}_A\mathbf{\Sigma}_A-\mathbf{H}_k\mathbf{T}_A)\right\} \\ & \qquad + \rho_B^2\text{Tr}\left\{\mathbf{\Pi}_B\mathbf{W}_B\mathbf{\Pi}_B\mathbf{\Sigma}_B^\dagger\mathbf{\Omega}_B\bar{\mathbf{H}}_k\mathbf{F}_k(\sigma_{\mathbf{n}}^2\gamma_k\mathbf{F}_k^\dagger\bar{\mathbf{H}}_k^\dagger\mathbf{\Omega}_B\mathbf{\Sigma}_B-\mathbf{G}_k\mathbf{T}_B)\right\}\Big) \\ & \qquad \times \text{vec}\left\{P_A\mathbf{H}_k^\dagger\mathbf{F}_k^\dagger\mathbf{F}_k\mathbf{H}_k\mathbf{T}_A-\rho_A^2\text{Tr}(\mathbf{T}_A^\dagger\mathbf{H}_k^\dagger\mathbf{F}_k^\dagger\mathbf{F}_k\mathbf{H}_k\mathbf{T}_A)\mathbf{T}_A\right\}^T\Big]d\text{vec}(\mathbf{T}_A^*) \end{aligned} \quad (16)$$

$\nu^n\nabla_l) \geq \varepsilon\,\nu^n\|\nabla_l\|_F^2$, respectively. Here, ν and ε are fixed constants between zero and one. With the step size from the Armijo's rule, the proposed algorithm obtains a non-increasing cost function value with respect to the number of iterations [22]. Therefore, it can be shown that the proposed iterative algorithms achieve at least a local optimum solution.

Consequently, using the derived sum-MSE gradients in (15) and (17) with $\mathbf{W}_A = \mathbf{I}_{M_A}$ and $\mathbf{W}_B = \mathbf{I}_{M_B}$, the proposed iterative algorithm provides a *joint MMSE* (JMMSE) solution. Also, from the derived sum-rate gradient expressions (18) and (19) with $w_A = w_B = 1$, the proposed iterative algorithm maximizes the sum-rate, which is called the *joint maximum-rate* (JMR) scheme.

C. Applications

The proposed JMMSE scheme using (15) and (17) minimizes the sum of the diagonal elements of the MSE matrix when employing the MMSE linear equalizer at the destination nodes. The sum-MSE minimizing criteria is important for optimizing the error performance. It was shown in [23] that applying a discrete Fourier transform (DFT) matrix or a Hardmard matrix at the transmitter makes the elements of the diagonal MSE matrix have the same value while the sum-MSE is maintained. Applying the proposed JMMSE scheme, the two-way relay channels are transformed to two separate point-to-point MIMO links satisfying the MMSE condition. Hence, from the filters $\mathbf{T}_A^{\text{mse}}, \mathbf{T}_B^{\text{mse}}, \mathbf{F}_1^{\text{mse}}, \cdots, \mathbf{F}_K^{\text{mse}}$ in the proposed JMMSE, the BER minimizing source transmit filter design is given as

$$\mathbf{T}_A^{\text{ber}} = \mathbf{T}_A^{\text{mse}}\mathbf{U}\mathbf{\Theta} \quad (20)$$

where $\mathbf{U}$ is an $M_A \times M_A$ unitary matrix and $\mathbf{\Theta}$ indicates the $M_A \times M_A$ normalized DFT matrix. Here, to make the MSE matrix $\mathbf{\Pi}_A$ in (8) diagonal, $\mathbf{U}$ is obtained from the eigenvalue decomposition of $\mathcal{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger$ where $\mathcal{H}$ is defined as

$$\begin{aligned} \mathcal{H} \triangleq & \left(\sum_{k=1}^{K}\gamma_k\bar{\mathbf{G}}_k\mathbf{F}_k^{\text{mse}}\mathbf{H}_k\mathbf{T}_A^{\text{mse}}\right)^\dagger \\ & \times\left(\sigma_{\mathbf{n}}^2\sum_{k=1}^{K}\gamma_k^2\bar{\mathbf{G}}_k\mathbf{F}_k^{\text{mse}}\mathbf{F}_k^{\text{mse}\dagger}\bar{\mathbf{G}}_k^\dagger+\sigma_{\mathbf{z}}^2\mathbf{I}_{M_B}\right)^{-1} \\ & \times\left(\sum_{k=1}^{K}\gamma_k\bar{\mathbf{G}}_k\mathbf{F}_k^{\text{mse}}\mathbf{H}_k\mathbf{T}_A^{\text{mse}}\right). \end{aligned}$$

Using the same method, we can also compute $\mathbf{T}_B^{\text{ber}}$. Note that since $\mathbf{U}\mathbf{\Theta}$ is unitary, it does not affect the sum-MSE Ξ_{sum} in (10), and makes the diagonal elements of the MSE matrix identical as in [23]. Consequently, the corresponding source and relay filters $\mathbf{T}_A^{\text{ber}}, \mathbf{T}_B^{\text{ber}}, \mathbf{F}_1^{\text{mse}}, \cdots, \mathbf{F}_K^{\text{mse}}$ minimize the average BER of each link for the MMSE linear detection, which is referred to as the *joint minimum-BER* (JMBER) scheme. Here, the weight matrices can be determined as $\mathbf{W}_A = w_A\mathbf{I}_{M_A}$ and $\mathbf{W}_B = w_B\mathbf{I}_{M_B}$ by adjusting non-negative weights w_A and w_B to support the different quality of service (QoS) requirements in both bidirectional links.

Also, the JMR scheme using (18) and (19) can obtain the rate region with the fixed transmit power by adjusting the non-negative weights w_A and w_B, and then the achievable rate region is illustrated as the convex hull of rate region points with any P_A and P_B satisfying the given power constraint as in [13]. In a practical system with a limited buffer size for the packet queues or a QoS demand of each link, the weight factors w_A and w_B may be computed by considering proportional fairness.

As for one-way protocol systems the transmission power of $\mathcal{T}_B$ is set to zero (i.e., $P_B = 0$), in order to find the optimum source relay filters of the one-way protocol, we can simply set $\rho_B = 0$ when calculating the gradients of the cost functions in the proposed iterative algorithm, and skip the step updating $\mathbf{T}_B$. Then, the proposed iterative algorithm can generate the optimized filters for the one-way protocol. Moreover, fixing $\mathbf{T}_A = \mathbf{I}_{M_A}$ and $\mathbf{T}_B = \mathbf{I}_{M_B}$ at step 3) of our proposed algorithm, we can optimize only the relay node filters for the case where the transmit CSI at both terminals is not available.

It is important to note that the proposed optimization algorithm is applicable for general one- and two-way multiple relay systems while the SVD-based technique in [19] can be used only in one-way single relay systems. In the case of the one-way single relay system, it will be shown in the simulation section that both the optimization methods have the same performance with the equal CSI requirement.

IV. Asymptotic Optimality

In this section, we investigate the optimality of the proposed algorithm for an asymptotic case. In general, to approach the global optimum solution in a non-convex problem, we should extensively repeat local optimization with randomly chosen different initial points and select the best one among the computed local solutions. Since the formulated cost functions in (9) and (12) are not generally convex or concave, the proposed algorithm may require an optimization process with multiple initial points. However, in the following theorem, we show that the sum-rate problem becomes convex in the asymptotic antenna regime.

Theorem 1: In multiple relay MIMO AF systems without CSI at the source terminals for both one-way and two-way protocols, a local-optimum solution achieves the global maximum in terms of the weighted sum-rate if the number of the source antennas goes to infinity.

Proof: First we consider two-way protocols. If the CSI is not available at the source terminal nodes, the source precoders are given as $\mathbf{T}_A = \mathbf{I}_{M_A}$ and $\mathbf{T}_B = \mathbf{I}_{M_B}$. The elements of $\mathbf{H}$ and $\mathbf{G}$ are independent and identically distributed (i.i.d.) Gaussian random variables. Thus various conditions for strong convergence are readily satisfied and we note that [30]

$$\lim_{M_A\to\infty}\frac{1}{M_A}\mathbf{H}_k\mathbf{H}_i^\dagger = \lim_{M_B\to\infty}\frac{1}{M_B}\mathbf{G}_k\mathbf{G}_i^\dagger = \begin{cases} \mathbf{I}_{N_k} & \text{for } k=i \\ \mathbf{0}_{N_kN_i} & \text{for } k\neq i \end{cases}$$

where $\mathbf{0}_{N_kN_i}$ denotes an $N_k \times N_i$ zero matrix. Applying the above asymptotic behavior to the cost function (13) with the relay power constraint, the weighted sum-rate problem in (12) as $M_A \to \infty$ and $M_B \to \infty$ becomes equivalent to

$$\max_{\mathbf{Q}_1,\cdots,\mathbf{Q}_K} R_{\text{sum}} = \max_{\mathbf{Q}_1,\cdots,\mathbf{Q}_K} \frac{w_A}{2}\log_2|\tilde{\mathbf{\Pi}}_A| + \frac{w_B}{2}\log_2|\tilde{\mathbf{\Pi}}_B| \quad (21)$$

$$\text{subject to } \operatorname{Tr}\{\mathbf{Q}_k\} = \frac{P_k}{P_A + P_B + \sigma_n^2} \quad \text{for } k=1,\cdots,K \quad (22)$$

where $\mathbf{Q}_k$ is defined as $\mathbf{Q}_k = \mathbf{F}_k\mathbf{F}_k^\dagger$, and then $\tilde{\mathbf{\Pi}}_A$ and $\tilde{\mathbf{\Pi}}_B$ are given by

$$\begin{aligned}\tilde{\mathbf{\Pi}}_A &= \mathbf{I}_{M_B} + P_A\sum_{k=1}^K \bar{\mathbf{G}}_k\mathbf{Q}_k\bar{\mathbf{G}}_k^\dagger\left(\sigma_n^2\sum_{k=1}^K\bar{\mathbf{G}}_k\mathbf{Q}_k\bar{\mathbf{G}}_k^\dagger + \sigma_z^2\mathbf{I}_{M_B}\right)^{-1},\\ \tilde{\mathbf{\Pi}}_B &= \mathbf{I}_{M_A} + P_B\sum_{k=1}^K \bar{\mathbf{H}}_k\mathbf{Q}_k\bar{\mathbf{H}}_k^\dagger\left(\sigma_n^2\sum_{k=1}^K\bar{\mathbf{H}}_k\mathbf{Q}_k\bar{\mathbf{H}}_k^\dagger + \sigma_z^2\mathbf{I}_{M_A}\right)^{-1}.\end{aligned} \quad (23)$$

Here $|\mathbf{I}+\mathbf{AB}| = |\mathbf{I}+\mathbf{BA}|$ is used in (23).

Let us consider $\{\mathbf{Q}_1,\cdots,\mathbf{Q}_K\}$ as the convex combination of two different feasible sets $\{\mathbf{X}_1,\cdots,\mathbf{X}_K\}$ and $\{\mathbf{Z}_1,\cdots,\mathbf{Z}_K\}$, represented by $\mathbf{Q}_k = \mathbf{X}_k + t\mathbf{Y}_k$ for $0 \le t \le 1$ where $\mathbf{Y}_k = \mathbf{Z}_k - \mathbf{X}_k$. Here $\mathbf{Y}_k$ $(k=1,\cdots,K)$ are Hermitian matrices, and $\mathbf{Q}_k$, $\mathbf{X}_k$ and $\mathbf{Z}_k$ are positive semi-definite (PSD) matrices. Then the cost function R_{sum} in (21) is a concave function of $\{\mathbf{Q}_1,\cdots,\mathbf{Q}_K\}$ if and only if $d^2R_{\text{sum}}/dt^2 \le 0$ for any feasible $\{\mathbf{X}_1,\cdots,\mathbf{X}_K\}$ and $\{\mathbf{Y}_1,\cdots,\mathbf{Y}_K\}$ [22]. Now, $\tilde{\mathbf{\Pi}}_l$ for $l = A$ and B in (23) can be rewritten as

$$\tilde{\mathbf{\Pi}}_l = \mathbf{I}_{M_{\underline{l}}} + P_l\tilde{\mathbf{Q}}_l\left(\sigma_n^2\tilde{\mathbf{Q}}_l + \sigma_z^2\mathbf{I}_{M_{\underline{l}}}\right)^{-1} \quad (24)$$

where $\tilde{\mathbf{Q}}_l = \tilde{\mathbf{X}}_l + t\tilde{\mathbf{Y}}_l$ for $l = A$ and B, $\tilde{\mathbf{X}}_A = \sum_{k=1}^K\bar{\mathbf{G}}_k\mathbf{X}_k\bar{\mathbf{G}}_k^\dagger$, $\tilde{\mathbf{Y}}_A = \sum_{k=1}^K\bar{\mathbf{G}}_k\mathbf{Y}_k\bar{\mathbf{G}}_k^\dagger$, $\tilde{\mathbf{X}}_B = \sum_{k=1}^K\bar{\mathbf{H}}_k\mathbf{X}_k\bar{\mathbf{H}}_k^\dagger$ and $\tilde{\mathbf{Y}}_B = \sum_{k=1}^K\bar{\mathbf{H}}_k\mathbf{Y}_k\bar{\mathbf{H}}_k^\dagger$.

Using the differential rules $\frac{d}{dt}\ln|\mathbf{Y}| = \operatorname{Tr}(\mathbf{Y}^{-1}\frac{d\mathbf{Y}}{dt})$, $\frac{d}{dt}\operatorname{Tr}(\mathbf{Y}) = \operatorname{Tr}(\frac{d\mathbf{Y}}{dt})$ and $\frac{d}{dt}\mathbf{Y}^{-1} = -\mathbf{Y}^{-1}\frac{d\mathbf{Y}}{dt}\mathbf{Y}^{-1}$ [28], the second order derivative of the weighted sum-rate in (21) d^2R_{sum}/dt^2 becomes

$$\frac{d^2R_{\text{sum}}}{dt^2} = \sum_{l=A,B}\frac{w_l}{2\ln 2}\operatorname{Tr}\left(-\tilde{\mathbf{\Pi}}_l^{-1}\frac{d\tilde{\mathbf{\Pi}}_l}{dt}\tilde{\mathbf{\Pi}}_l^{-1}\frac{d\tilde{\mathbf{\Pi}}_l}{dt} + \tilde{\mathbf{\Pi}}_l^{-1}\frac{d^2\tilde{\mathbf{\Pi}}_l}{dt^2}\right) \quad (25)$$

where $d\tilde{\mathbf{\Pi}}_l/dt$ and $d^2\tilde{\mathbf{\Pi}}_l/dt^2$ are calculated as

$$\frac{d\tilde{\mathbf{\Pi}}_l}{dt} = \sigma_z^2P_l\left(\sigma_n^2\tilde{\mathbf{Q}}_l + \sigma_z^2\mathbf{I}_{M_{\underline{l}}}\right)^{-1}\tilde{\mathbf{Y}}_l\left(\sigma_n^2\tilde{\mathbf{Q}}_l + \sigma_z^2\mathbf{I}_{M_{\underline{l}}}\right)^{-1}, \quad (26)$$

$$\begin{aligned}\frac{d^2\tilde{\mathbf{\Pi}}_l}{dt^2} &= -2\sigma_n^2\sigma_z^2P_l\left(\sigma_n^2\tilde{\mathbf{Q}}_l + \sigma_z^2\mathbf{I}_{M_{\underline{l}}}\right)^{-1}\\ &\quad\times\tilde{\mathbf{Y}}_l\left(\sigma_n^2\tilde{\mathbf{Q}}_l + \sigma_z^2\mathbf{I}_{M_{\underline{l}}}\right)^{-1}\tilde{\mathbf{Y}}_l\left(\sigma_n^2\tilde{\mathbf{Q}}_l + \sigma_z^2\mathbf{I}_{M_{\underline{l}}}\right)^{-1}.\end{aligned} \quad (27)$$

Here a matrix inversion lemma $\mathbf{I} - \mathbf{A}(\mathbf{I}+\mathbf{A})^{-1} = (\mathbf{I}+\mathbf{A})^{-1}$ is applied.

Since $\tilde{\mathbf{Q}}_l$ is PSD and $\tilde{\mathbf{Y}}_l$ is Hermitian, $d\tilde{\mathbf{\Pi}}_l/dt$ and $d^2\tilde{\mathbf{\Pi}}_l/dt^2$ in (26) and (27) are shown to be Hermitian and negative semi-definite, respectively. Also, $\tilde{\mathbf{\Pi}}_l^{-1}$ is a PSD matrix from (24). We can show that from these facts, the matrix inside the trace operation in (25) is negative semi-definite, and thus we get $d^2R_{\text{sum}}/dt^2 \le 0$. Consequently, since the cost function R_{sum} in (21) is a concave function of $\{\mathbf{Q}_1,\cdots,\mathbf{Q}_K\}$ and the equality constraints in (22) are affine functions, the formulated problem is a convex optimization problem which has a unique maximum. Therefore a local-optimal solution with respect to $\{\mathbf{F}_1,\cdots,\mathbf{F}_K\}$ achieves the globally maximum of the weighted sum-rate. For a one-way protocol of $\mathcal{T}_A \to \mathcal{T}_B$, the proof is identical for $P_B = 0$, $w_A = 1$ and $w_B = 0$. ■

This theorem implies that the proposed JMR scheme which obtains the local-optimum relay filters with a single initial point provides the global maximum sum-rate when the number of the source antennas increases. It is important to observe that the local-optimum filters are not unique, but have the same maximum. For example, if we optimize a set of relay filters $\{\mathbf{F}_1^{\text{rate}},\cdots,\mathbf{F}_K^{\text{rate}}\}$, it is obvious that $\{e^{j\theta}\mathbf{F}_1^{\text{rate}},\cdots,e^{j\theta}\mathbf{F}_K^{\text{rate}}\}$ for any real scalar θ also has the same optimum. Although we consider an asymptotic case here, we find from extensive simulations that our schemes generate solutions very close to global optimal solutions even with a moderate number of antennas, which will be demonstrated in the following section.

V. Simulation Results

In this section, we present numerical results for the proposed schemes in terms of the sum-rate and the BER. In our simulation, it is assumed $P_A = P_B = P/2$, $P_k = P/K$ for all k and $\sigma_{\mathbf{n}}^2 = \sigma_{\mathbf{z}}^2 = 1$, and the SNR is defined as $P/\sigma_{\mathbf{n}}^2$. Also, we assume the TDD mode so that elements of the channel matrices $\mathbf{H}_k$ and $\mathbf{G}_k$ at the first phase have an i.i.d. complex Gaussian distribution with zero mean and unit variance, and the channel matrices $\bar{\mathbf{H}}_k$ and $\bar{\mathbf{G}}_k$ at the second phase are $\bar{\mathbf{H}} = \mathbf{H}^T$ and $\bar{\mathbf{G}} = \mathbf{G}^T$ due to channel reciprocity [10]. We use a notation of $M_A \times \{N_1,\cdots,N_K\} \times M_B$ for representing antenna configurations in this section. For the proposed iterative algorithm, the tolerance factor ϵ is set to 10^{-4}, and the constants ν and ε are fixed as 0.5 and 0.2, respectively. Throughout simulations, the proposed algorithm employs identity matrices as an initial point.

In Fig. 2, we exhibit the results of the proposed JMR algorithm in terms of iterations for both one-way and two-way systems in $4 \times \{4,4\} \times 4$ relaying networks. First in Fig. 2 (a), the y-axis indicates $\sum_{k=1}^K\|\nabla_k\|_F^2 + \sum_{l=A}^B\|\nabla_l\|_F^2$. Thus, this figure confirms that the proposed algorithm satisfies

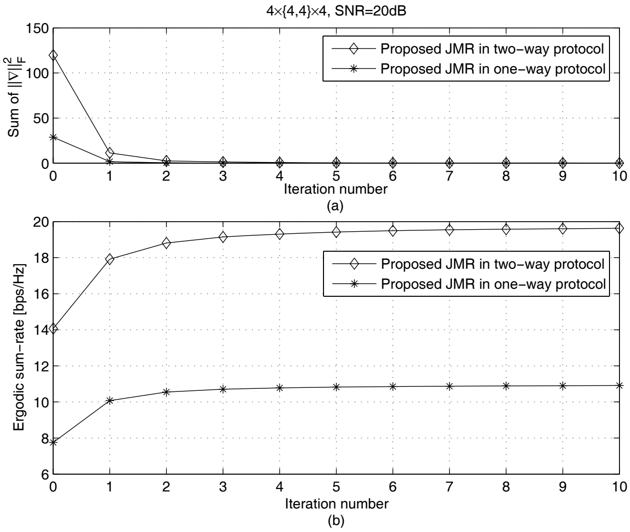


Fig. 2. Convergence property of the proposed algorithm for maximizing the sum-rate in multiple relay systems.

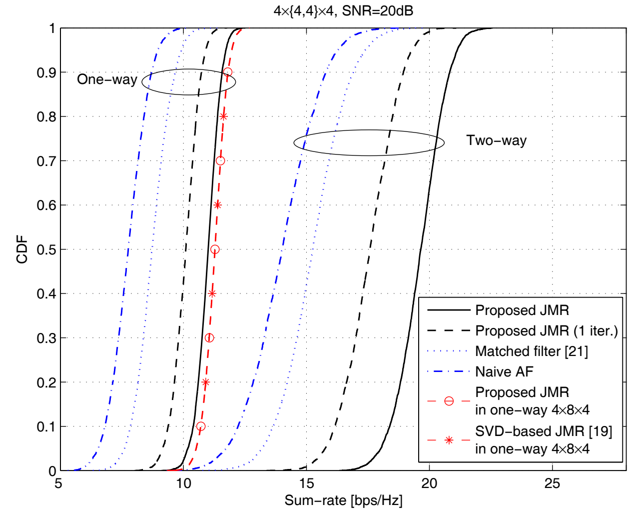


Fig. 4. CDFs of sum-rates of multiple relay systems at SNR= 20dB.

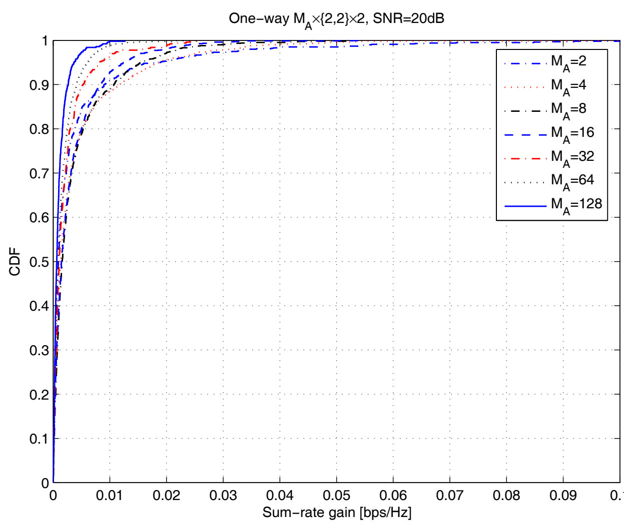


Fig. 3. CDFs of the sum-rate gains with various number of source antennas M_A at SNR $=$ 20dB.

zero gradient condition with a few iterations. Also, in Fig. 2 (b), the proposed algorithm exhibits a very good convergence property.[3]

Fig. 3 is plotted to illustrate the asymptotic optimality in Theorem 1, where one-way $M_A \times \{2, 2\} \times 2$ systems without the CSI at the source node are considered. In this figure, we plot the cumulative distribution function (CDF) of the sum-rate gain of the global optimum solution over the proposed method. Here the global optimal solution is obtained by using 1000 different initial points, while the proposed scheme employs a single identity initial point. We can see that as the number of source antennas M_A increases, the sum-rate gain becomes close to zero. This result shows that with large source antennas, our proposed solutions do not require multiple initial points, as expected in Theorem 1. Furthermore, it is interesting to see that even for a moderate antenna size such as $2 \times \{2, 2\} \times 2$, most results with extensive initial points exhibit negligible gains of less than 0.01bps/Hz. These results confirm that the proposed algorithm obtains solutions very close to the globally optimal value in a practical antenna regime.

In Fig. 4, the CDFs of the instantaneous sum-rates of $4 \times \{4, 4\} \times 4$ systems are depicted. In order to demonstrate the performance of our scheme, we compare the conventional naive AF scheme with $\mathbf{F}_k = \mathbf{I}_{N_k}$ and the matched filter method using $\mathbf{F}_k = \bar{\mathbf{G}}_k^\dagger \mathbf{H}_k^\dagger$ and $\bar{\mathbf{G}}_k^\dagger \mathbf{H}_k^\dagger + \bar{\mathbf{H}}_k^\dagger \mathbf{G}_k^\dagger$ for the one-way and two-way protocols, respectively [21]. The plots show that the proposed JMR scheme performs best in both one-way and two-way multiple relay systems. For a sanity check, we also plot the CDF of the rate of the SVD-based optimization with the rate maximizing criterion in [19] for one-way channels with $4 \times 8 \times 4$ systems. In this case, we can see that the proposed JMR algorithm yields the sum-rate identical to the SVD-based JMR scheme. Furthermore, it is obvious that since full cooperation among antennas in the relay is possible, the rate performance of the $4 \times 8 \times 4$ system can be considered as an upper bound for the $4 \times \{4, 4\} \times 4$ systems. We can see from this figure that the performance of the proposed JMR scheme is only 0.2bps/Hz away from the upper bound. Also, the proposed scheme with one iteration outperforms the conventional techniques. It should be noted that the scheme with one iteration can be considered as a non-iterative scheme, and can be preferred in a practical system.

Fig. 5 exhibits an ergodic sum-rate curve as a function of SNR for the $4 \times \{4, 4\} \times 4$ system. For one-way multiple relay channels, the proposed scheme is compared with the conventional MMSE filter [31], the QR-based design [7], the matched filter scheme [21] and the naive AF. The proposed JMR scheme obtains SNR gains of more than 3dB at 10bps/Hz over the conventional methods for both one-way and two-way systems.

Next, we present the sum-rate graphs with respect to the number of relay nodes K at the SNR of 20dB in Fig. 6. The sum-rate gain of the proposed JMR algorithm over the naive AF scheme becomes much larger as the number of

[3] The proposed JMMSE scheme shows a similar convergence property with the JMR scheme.

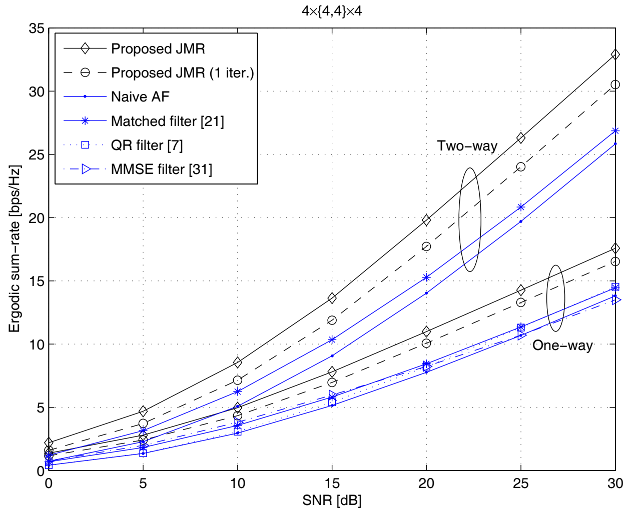


Fig. 5. Sum-rate comparison as a function of SNR in multiple relay systems.

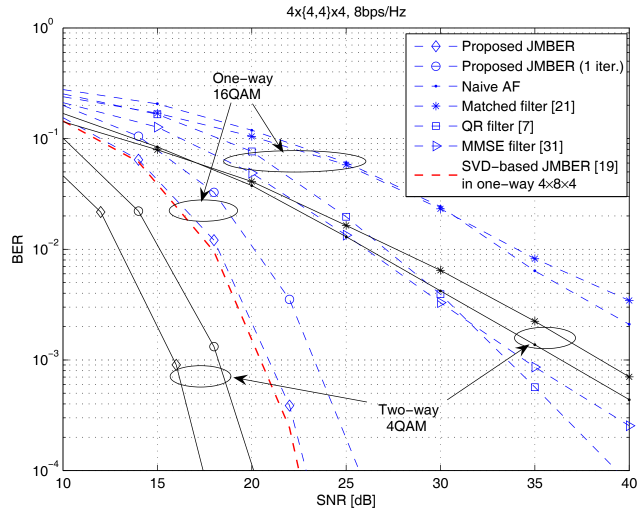


Fig. 7. BER performance for multiple relay systems at 8bps/Hz.

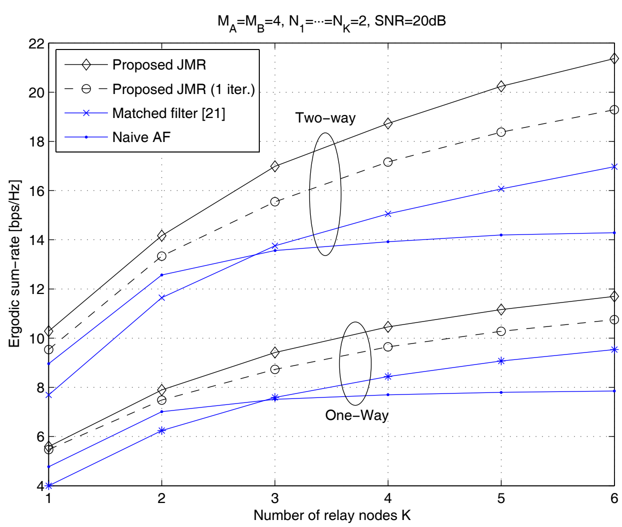


Fig. 6. Sum-rate comparison with various relay nodes K in multiple relay systems at SNR= 20dB.

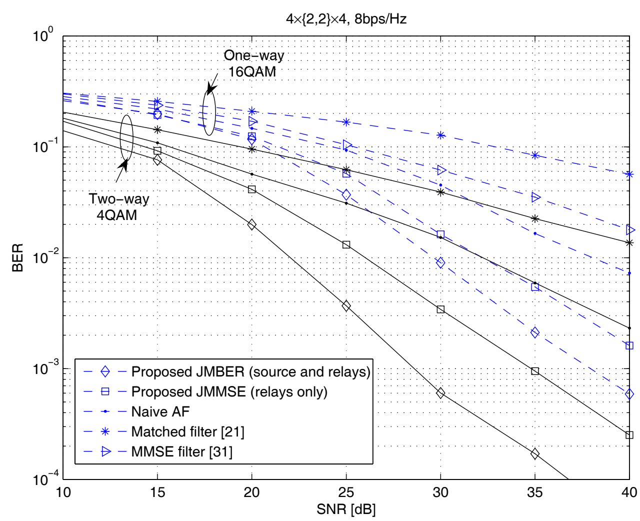


Fig. 8. BER performance for multiple relay systems at 8bps/Hz.

relay nodes increases. Also, even with a single iteration, the proposed scheme still outperforms the naive AF scheme and the matched filter method.

Fig. 7 compares the BER curve of the proposed JMBER scheme with the conventional relay schemes in $4 \times \{4,4\} \times 4$ systems for one-way and two-way protocols, where the MMSE linear receiver in (7) is employed. Here, to support the same spectral efficiency of 8bps/Hz, we use 16QAM and 4QAM constellations for the one-way and the two-way protocols, respectively. The proposed JMBER algorithms with only one iteration outperforms the conventional multiple relay techniques by more than 14dB and 25dB at a BER of 10^{-4} for one-way and two-way protocols, respectively. For comparison, we also plot the BER curve of the SVD-based optimization with the BER minimizing criteria [19] in one-way $4 \times 8 \times 4$ systems. It is interesting to see that the proposed JMBER scheme in one-way $4\times\{4,4\}\times4$ systems performs within only a few tenth of a dB compared to the single relay optimization method, although cooperation between two relay nodes is impossible.

Finally, in Fig. 8, the BERs are plotted for $4 \times \{2,2\} \times 4$ systems. We can see that in this configuration, the conventional matched filter and MMSE filter design are significantly degraded even compared with the naive AF scheme since the number of the data streams ($M_A = M_B = 4$) is larger than the antenna dimension of each relay node ($N_k = 2$). Note that the relay QR filter design cannot be applied in this antenna configuration. The proposed JMMSE (relays only) scheme which optimizes the relay filters only without the source optimization shows a SNR gain of about 10dB at a BER of 10^{-4} over the conventional naive AF for both one-way and two-way protocols. Also, it is shown that in the two-way protocol, the JMBER scheme (source and relays) obtains a gain of more than 7dB over the JMMSE (relays only) by jointly optimizing the source and relay filters and applying the BER minimization technique. These results confirm the efficacy of the proposed JMBER scheme in multiple relay systems.

VI. Conclusion

In this paper, we have proposed a method to optimize linear processing designs for several criterions in one- and two-way multiple relay systems. Using the proposed iterative algorithm, we obtain the JMR and the JMMSE schemes which maximizes the sum-rate or minimizes the sum-MSE at the source and relay nodes. Furthermore, when employing a linear MMSE receiver at the destinations, the proposed JMMSE scheme is transformed to the JMBER which provides the minimum BER performance. It is shown that the proposed algorithms converge to a local optimum solution, which approaches a global optimal solution as the number of source antennas increases. Simulation results demonstrate that the proposed algorithms provide solutions very close to the optimal value even in a moderate antenna configuration for both one-way and two-way multiple relay systems. Moreover, our schemes with one iteration are shown to be quite effective in terms of both the sum-rate and the BER.

References

[1] T. M. Cover and A. A. El Gamal, “Capacity theorems for the relay channels," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572-584, Sep. 1979.

[2] M. Gastpar and M. Vetterli, “On the capacity of wireless networks: the relay case," in *Proc. IEEE INFOCOM*, vol. 3, pp. 1577-1586, June 2002.

[3] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037-3063, Sep. 2005.

[4] B. Wang, J. Zhang, and A. Host-Madsen, “On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29-43, Jan. 2005.

[5] A. Host-Madsen and J. Zhang, “Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020-2040, June 2005.

[6] H. Bolcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj, “Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433-1444, June 2006.

[7] H. Shi, T. Abe, T. Asai, and H. Yoshino, “Relaying schemes using matrix triangularization for MIMO wireless networks," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1683-1688, Sep. 2007.

[8] Z. Yi and I.-M. Kim, “Joint optimization of relay-precoders and decoders with partial channel side information in cooperative networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 447-458, Feb. 2007.

[9] A. S. Behbahani, R. Merched, and A. M. Eltawil, “Optimizations of a MIMO relay network," *IEEE Trans. Signal Process.*, pp. 5062-5073, Oct. 2008.

[10] B. Rankov and A. Wittneben, “Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379-389, Feb. 2007.

[11] T. J. Oechtering, C. Schnurr, I. Bjelakovic, and H. Boche, “Broadcast capacity region of two-phase bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 454-458, Jan. 2008.

[12] K.-J. Lee, K. W. Lee, H. Sung, and I. Lee, “Sum-rate maximization for two-way MIMO amplify-and-forward relaying systems," in *Proc. IEEE VTC*, Apr. 2009.

[13] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, “Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 699-712, June 2009.

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

[15] X. Tang and Y. Hua, “Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1398-1407, Apr. 2007.

[16] O. Munoz-Medina, J. Vidal, and A. Agustin, “Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2593-2604, June 2007.

[17] K.-J. Lee, J.-S. Kim, G. Caire, and I. Lee, “Asymptotic ergodic capacity analysis for MIMO amplify-and-forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2712 2717, Sep. 2010.

[18] W. Guan and H. Luo, “Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Commun. Lett.*, vol. 12, no. 7, pp. 517-519, July 2008.

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

[20] N. Lee, H. Park, and J. Chun, “Linear precoder and decoder design for two-way AF MIMO relaying system," in *Proc. IEEE VTC*, May 2008, pp. 1221-1225.

[21] R. Vaze and R. W. Heath, “On the capacity and diversity-multiplexing tradeoff of the two-way relay channel." [Online]. Available: http://arxiv.org/PS cache/arxiv/pdf/0810/0810.3900v2.pdf, Oct. 2008.

[22] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd edition. John Wiley & Sons, 2006.

[23] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, “Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381-2401, Sep. 2003.

[24] H. Yomo and E. de Carvalho, “A CSI estimation method for wireless relay network," *IEEE Commun. Lett.*, vol. 11, no. 6, pp. 480-482, June 2007.

[25] F. Gao, T. Cui, and A. Nallanathan, “On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907-1916, May 2008.

[26] F. Gao, R. Zhang, and Y.-C. Liang, “Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024-3033, Oct. 2009.

[27] S. Haykin, *Adaptive Filter Theory*, 4th edition. Prentice Hall, 2001.

[28] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, revised edition. John Wiley & Sons, 2002.

[29] D. Guo, S. Shamai (Shitz), and S. Verdú, “Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261-1282, Apr. 2005.

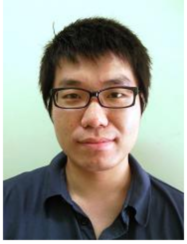
[30] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, “Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893-1909, Sep. 2004.

[31] O. Oyman and A. J. Paulraj, “Design and analysis of linear distributed MIMO relaying algorithms," *IEE Proc. Commun.*, vol. 153, no. 4, pp. 565-572, Aug. 2006.

[32] H. Lee, B. Lee, and I. Lee, “Iterative detection and decoding with an improved V-BLAST for MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp.504-513, Mar. 2006.

Kyoung-Jae Lee (S’06) received the B.S. and M.S. degrees in electrical engineering from Korea University, Seoul, Korea, in 2005 and 2007, where he is currently working toward the Ph.D. degree in the School of Electrical Engineering. He worked as an intern at Beceem Communications, Santa Clara, CA, USA during the winter of 2006, and visited University of Southern California, Los Angeles, CA, USA to conduct a collaborative research during the summer of 2009. His research interests are in communication theory and signal processing for wireless communications, including multiuser MIMO, relay, and cognitive radio networks. He received the Gold Paper Award at the IEEE Seoul Section Student Paper Contest in 2007, and the Best Paper Award at the IEEE VTC Fall in 2009.

Hakjea Sung (S’06-M’10) received the B.S. and M.S. degrees in electrical engineering from Hongik University, Seoul, Korea, in 1998 and 2000, respectively, and the Ph.D. degree in the School of Electrical Engineering at Korea University, Seoul, Korea, in 2010. He has been with the Samsung Electronics, Suwon, Korea, as a research engineer in the mobile communication R&D group, since 2000. His research interests include signal processing techniques for MIMO-OFDM systems and multi-user MIMO wireless networks. He received the Best Paper Award at the IEEE VTC Fall in 2009.



Eunsung Park (S'09) received the B.S. and M.S. degrees in electrical engineering from Korea University, Seoul, Korea, in 2007 and 2009, where he is currently working toward the Ph.D. degree in the School of Electrical Engineering. During the summer of 2009, he visited University of Southern California, Los Angeles, CA, USA to conduct a collaborative research. His research interests are in communication theory and signal processing for wireless communications such as MIMO-OFDM systems and wireless relay networks.

Inkyu Lee (S'92-M'95-SM'01) received the B.S. degree (Hon.) in control and instrumentation engineering from Seoul National University, Seoul, Korea, in 1990, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1992 and 1995, respectively. From 1991 to 1995, he was a Research Assistant at the Information Systems Laboratory, Stanford University. From 1995 to 2001, he was a Member of Technical Staff at Bell Laboratories, Lucent Technologies, where he studied the high-speed wireless system design. He later worked for Agere Systems (formerly Microelectronics Group of Lucent Technologies), Murray Hill, NJ, as a Distinguished Member of Technical Staff from 2001 to 2002. In September 2002, he joined the faculty of Korea University, Seoul, Korea, where he is currently a Professor in the School of Electrical Engineering. During 2009, he visited University of Southern California, LA, USA, as a visiting Professor. He has published over 50 journal papers in IEEE, and has 30 U.S. patents granted or pending. His research interests include digital communications, signal processing, and coding techniques applied to wireless systems with an emphasis on MIMO-OFDM. Dr. Lee currently serves as an Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. Also, he has been a Chief Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Special Issue on 4G Wireless Systems). He received the IT Young Engineer Award as the IEEE/IEEK joint award in 2006. Also he received the Best Paper Award at APCC in 2006 and IEEE VTC in 2009.