Transceiver Optimization for Multi-hop MIMO Relay Multicasting from Multiple Sources

Muhammad R. A. Khandaker, *Member, IEEE*, and Yue Rong, *Senior Member, IEEE*

Abstract—In this paper, we consider a multicasting multiple-input multiple-output (MIMO) relay system where multiple transmitters multicast their own messages to a group of receivers over multiple hops, and all nodes are equipped with multiple antennas. Joint transmit and relay precoding design problem has been investigated for multicasting multiple data streams based on min-max mean-squared error (MSE) criterion. We aim at minimizing the maximal MSE of the signal waveform estimation among all receivers subjecting to power constraints at the transmitters and all the relay nodes. This problem is highly nonconvex with matrix variables and the exactly optimal solution is very hard to obtain. We develop an iterative algorithm to jointly optimize the transmitter, relay, and receiver matrices through solving convex subproblems. By exploiting the optimal structure of the relay precoding matrices, we then propose a low complexity solution for the problem under some mild approximation. In particular, we show that under (moderately) high signal-to-noise ratio (SNR) assumption, the min-max optimization problem can be solved using the semidefinite programming (SDP) technique. Numerical simulations demonstrate the effectiveness of the proposed algorithms.

I. Introduction

In many practical communication systems, multiple users (transmitters) need to send their messages to a group of receivers simultaneously. The simplest way to send data to multiple receivers simultaneously is to transmit individual copies of the data to each receiver. However, this is highly inefficient, since multiple copies of the same data are sent from the source through one or more networks. Multicasting enables a single transmission to be received by multiple users, significantly reducing the required bandwidth. For example, in war fields, military troops need to share their current status, pass secret messages to the allied groups. In an online interactive gaming scenario, all the participants are interested to know the current status of their rivals. If the participants multicast their information, the interested users can receive all the messages simultaneously. Multicasting from multiple sources can also be used to support video conferencing and webcasts among multiple users.

The broadcasting nature of the wireless channel makes it naturally suitable for multicasting applications, since a single transmission may be simultaneously received by multiple

M. R. A. Khandaker is with the Department of Electronic and Electrical Engineering, University College London, Gower Street, London, WC1E 7JE, UK (e-mail: m.khandaker@ucl.ac.uk).

Y. Rong is with the Department of Electrical and Computer Engineering, Curtin University, Bentley, WA 6102, Australia (e-mail: y.rong@curtin.edu.au).

users. Recently, wireless multicasting technology has attracted great research interest, due to the increasing demand for mobile applications such as streaming media, software updates, and location-based services involving group communications. Particular multicasting applications include live IP-TV, Internet radio, video conferencing and webcasts. However, wireless channel is subject to fading. By exploiting the spatial diversity, multi-antenna techniques can be applied to combat channel fading [1], [2]. Next generation wireless communication standards such as WiMAX 802.16m and 3GPP LTE-Advanced have already included technologies which enable better multicasting solutions based on multi-antenna and beamforming techniques [3].

Due to its nonconvex nature, the problem of designing optimal beamforming vectors for multicasting is hard in general. The authors of [4] have designed transmit beamformers for physical layer multicasting using rank relaxations, where two design criteria were adopted, namely minimizing the transmit power subject to minimum received signal-to-noise ratio (SNR) at each of the intended receivers and a related max-min SNR problem subject to a transmit power constraint. It has been proven in [4] that both problems are NP-hard. Using lower complexity transmission schemes, the information theoretic capacity of the multi-antenna multicasting channel was studied in [5] with a particular focus on the scaling of the capacity and achievable rates as the number of antennas and/or users approaches infinity. The effect of channel spatial correlation on the multicasting capacity has been investigated in [6]. The asymptotic capacity limits of multi-antenna multicasting channel have been studied in [7] based on antenna subset selection. The authors of [8] investigated transmit precoding design for multi-antenna multicasting systems where the channel state information (CSI) is obtained via limited feedback. The authors of [9] considered transmit covariance design for a secrecy rate maximization problem, where a multi-antenna transmitter delivers a confidential message to multiple single-antenna receivers in the presence of multiple multi-antenna eavesdroppers. In [9], the tight upper bound and the lower bound of the multicast secrecy rate are defined via convex approximation.

The works in [4]-[9] solved the max-min SNR/rate beamforming problems with the aid of semidefinite relaxation (SDR) and rank-one approximation. Note that the rank-relaxation technique is suboptimal in general. In [10]-[11], a stochastic beamforming strategy is proposed for multi-antenna multicasting where the randomization is guided by SDR, but without the need of rank-one approximation. While the use of channel coding and the assumption of sufficiently long

code lengths play a vital role in achieving the above result, a combination of transmit beamforming and the Alamouti space-time code has also been considered in [11] which yields a rank-two generalization of the SDR-based beamforming. The fundamental limit of the max-min beamforming is that as the number of users increases to infinity, the achievable rate decreases to zero [5]. To solve this problem, a joint beamforming and admission control approach has been developed in [12] and [13], where a subset of users is selected so that certain quality-of-service (QoS) requirements can be satisfied. An iterative transmit beamforming algorithm was proposed in [13] for multiple cochannel multicasting groups to minimize the total power consumed by the antenna array subjecting to signal-to-interference-plus-noise ratio (SINR) constraints at the receivers. However, the method used in [13] has a high computational complexity. The authors in [14] attempted to reduce the complexity of [13] by combining the concept of the iterative second-order cone programming (SOCP) with that of interior-point methods.

While the works in [4]-[14] investigated multicasting systems with single-antenna receivers, recently multi-antenna receivers have been considered [15]-[17] for multicasting systems since receiver beamforming can significantly improve the system performance. In particular, coordinated beamforming techniques have been investigated in [15], where a generalized form of block diagonalization has been proposed to make orthogonal transmissions to distinct multicasting groups with multi-antenna receivers. The scaling of the achievable rate with increasing number of users was investigated in [16] for multiple-input multiple-output (MIMO) multicasting where the transmission is coded at the application layer over a number of channel realizations. In [17], non-iterative nearly optimal transmit beamformers are designed for wireless link layer multicasting with real-valued channels, and for complex-valued channels an upper-bound on the multicasting rate is derived.

The above works [3]-[17] considered single-hop multicasting systems. However, as the transmitter-receiver distance increases, it becomes necessary to adopt relay nodes to efficiently combat the pathloss of wireless channel. Relay nodes are also essential to overcome the shadowing effect of wireless links in large urban areas with giant buildings and other obstacles, behind-the-hill areas, and so on. Hence, efforts are being made to design optimal beamformers for multicasting over more than one hops using relay nodes. A multi-group multicasting relay network has been considered in [18] and a distributed beamforming algorithm was proposed to minimize the total relay power where each node is equipped with a single antenna. The authors in [19] studied the lower-bound for the outage probability of cooperative multi-antenna multicasting schemes based on the amplify-and-forward (AF) strategy where the users are equipped with a single antenna. In [20], multicast scheduling with multiple sessions and multiple channels was investigated where the base station may multicast data in two sessions using MIMO simultaneously through the same channel and the users are allowed to cooperatively help each other on orthogonal channels. Thus, the scheme in [20] leads to a higher multicasting rate than single-session transmis-

sions. Joint transmit and relay precoding design problems were investigated in [21]-[22] for a two-hop multicasting MIMO relay system where all nodes are equipped with multiple antennas. An iterative algorithm has been developed in [22] to jointly optimize the source, relay, and receiver matrices. In order to reduce the computational complexity of the iterative algorithm, a simplified algorithm has also been proposed in [21]-[22] for the two-hop multicasting system. Multicasting from multiple sources in a dual-hop MIMO relay system has been considered in [23].

In this paper, we consider multi-hop multicasting MIMO relay systems where multiple transmitters multicast their messages to a group of receivers with the aid of multiple relay nodes located in series. The transmitters, relay nodes, and receivers are all equipped with multiple antennas. To the best of our knowledge, such multicasting (from multiple sources) MIMO relay system has not been investigated in existing works. Note that our paper generalizes the multicasting scheme in [21]-[22] in two ways. Firstly, we consider multicasting from multiple sources instead of the single-transmitter multicasting in [21]-[22]. Secondly, we generalize the two-hop MIMO relay multicasting scheme to multi-hop systems with any number of hops. Such extension is important in the case of long source-destination distance where a two-hop relay is not sufficient and multi-hop relays are necessary to establish a reliable source-destination link. It is obvious that due to the introduction of multiple users and multiple relay nodes, the mean-squared error (MSE) matrix decomposition and hence the source and relay matrices optimization procedure become much more challenging than that for the single-transmitter two-hop system. For the sake of the implementation simplicity, we choose the AF relaying strategy at all relay nodes. We consider the joint transmit and relay precoding design problem based on the min-max MSE criterion. We aim at minimizing the maximal MSE of the signal waveform estimation among all receivers subjecting to power constraints at the transmitters and the relay nodes. The problem is highly nonconvex with matrix variables and the exactly optimal solution is very difficult to obtain. We develop an iterative algorithm to jointly optimize the transmitter, relay, and receiver matrices through solving convex subproblems. By exploiting the optimal structure of the relay precoding matrices, we propose a low-complexity solution to the problem under some mild approximation. We apply the same concept of high SNR assumption as in [22] in order to decompose the complicated original optimization problem into smaller easily solvable subproblems. In particular, we show that under (moderately) high SNR assumption, the problem can be solved using standard semidefinite programming (SDP) techniques. Numerical simulations demonstrate the effectiveness of the proposed algorithms. Note that the proposed algorithms support multicasting multiple data streams in contrast to the existing single data stream multicasting schemes [3]-[20]. In this paper, for notational convenience, we consider a narrow-band single-carrier system. However, our results can be straightforwardly generalized to each subcarrier of a broadband multi-carrier multicasting MIMO relay system.

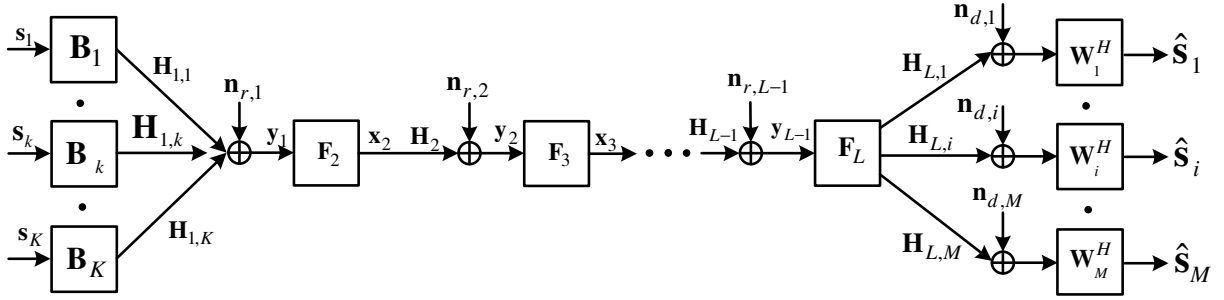The rest of this paper is organized as follows. In Section II,

Fig. 1. Block diagram of a multi-hop multiuser multicasting MIMO relay system.

the system model of a multi-hop multicasting MIMO relay network is introduced. The joint transmit and relay precoding matrices design algorithms are developed in Section III. Section IV shows the simulation results which justify the significance of the proposed algorithms under various scenarios. Conclusions are drawn in Section V.

II. SYSTEM MODEL

We consider an L-hop ($L \geq 2$) MIMO multicasting system where K transmitters simultaneously multicast their information to M receivers with the aid of $L - 1$ relay nodes as illustrated in Fig. 1. The kth transmitter and the lth relay node are equipped with $N_{\mathrm{s},k}$ and N_l antennas, respectively. For the sake of notational simplicity, we assume that each receiver has N_{d} antennas. The algorithms developed in this paper can be straightforwardly extended to multicasting systems where receivers have different number of antennas. The transmitters multicast their information-carrying symbols to all receivers with the aid of $L-1$ relay nodes. The direct links between the transmitters and the receivers and those between any two non-consecutive relays are not considered since we assume that these direct links undergo much larger path attenuations compared with the links via the intermediate (relaying) node(s).

We assume that the relay nodes work in half-duplex mode. Thus the communication between the transmitters and receivers is accomplished in L time slots. In the first time slot, the kth transmitter linearly precodes an $N_{\mathrm{b},k} \times 1 (1 \leq N_{\mathrm{b},k} \leq N_{\mathrm{s},k})$ modulated signal vector $\mathbf{s}_k$ by an $N_{\mathrm{s},k} \times N_{\mathrm{b},k}$ precoding matrix $\mathbf{B}_k$ and transmits the precoded vector $\mathbf{B}_k\mathbf{s}_k$ to the first relay node. We assume that $\mathrm{E}[\mathbf{s}_k\mathbf{s}_k^H] = \mathbf{I}_{N_{\mathrm{b},k}}$, where $\mathrm{E}[\cdot]$ denotes statistical expectation, $(\cdot)^H$ stands for the matrix Hermitian transpose, and $\mathbf{I}_n$ is an $n \times n$ identity matrix. We denote $N_{\mathrm{b}} = \sum_{k=1}^{K} N_{\mathrm{b},k}$ as the total number of independent data streams from all transmitters. In order to transmit N_{b} independent data streams simultaneously, there must be $N_{\mathrm{b}} \leq \min(\{N_l\}, N_{\mathrm{d}})$ where $\{N_l\} = \{N_l, l = 1, \cdots, L-1\}$. The received signal vector at the first relay node is given by

$$\mathbf{y}_1 = \sum_{k=1}^{K} \mathbf{H}_{1,k}\mathbf{B}_k\mathbf{s}_k + \mathbf{n}_{\mathrm{r},1} \triangleq \mathbf{H}_1\mathbf{x}_1 + \mathbf{n}_{\mathrm{r},1} \qquad (1)$$

where $\mathbf{H}_{1,k}$ is the $N_1 \times N_{\mathrm{s},k}$ MIMO channel matrix between the kth transmitter and the first relay node, $\mathbf{y}_1$ and $\mathbf{n}_{\mathrm{r},1}$ are the $N_1 \times 1$ received signal and additive Gaussian noise vectors introduced at the first relay node, respectively, $\mathbf{H}_1 =$

$[\mathbf{H}_{1,1}, \cdots, \mathbf{H}_{1,K}]$, and $\mathbf{x}_1 = \mathbf{F}_1\mathbf{s}$, $\mathbf{F}_1 \triangleq \mathrm{bd}(\mathbf{B}_1, \cdots, \mathbf{B}_K)$, $\mathbf{s} \triangleq [\mathbf{s}_1^T, \cdots, \mathbf{s}_K^T]^T$. Here $\mathrm{bd}(\cdot)$ stands for a block diagonal matrix and $(\cdot)^T$ denotes the matrix (vector) transpose.

The input-output relationship at the lth relay node is

$$\mathbf{x}_{l+1} = \mathbf{F}_{l+1}\mathbf{y}_l, \qquad l = 1, \cdots, L-1 \qquad (2)$$

where $\mathbf{F}_{l+1}, l = 1, \cdots, L-1$, is the $N_l \times N_l$ amplifying matrix at the lth relay node and $\mathbf{y}_l, l = 1, \cdots, L-1$, is the $N_l \times 1$ signal vector received at the lth relay node written as

$$\mathbf{y}_l = \mathbf{H}_l\mathbf{x}_l + \mathbf{n}_{\mathrm{r},l}, \qquad l = 1, \cdots, L-1 \qquad (3)$$

where $\mathbf{H}_l, l = 1, \cdots, L-1$ is the $N_l \times N_{l-1}$ MIMO channel matrix of the lth hop ($N_0 = N_{\mathrm{b}}$), $\mathbf{y}_l$ and $\mathbf{n}_{\mathrm{r},l}$ are the $N_l \times 1$ received signal and additive Gaussian noise vectors introduced at the lth relay node, respectively. Using (1) and (2), for the case of $L \geq 3$, (3) can be rewritten as

$$\begin{aligned} \mathbf{y}_l &= \prod_{i=l}^{1}(\mathbf{H}_i\mathbf{F}_i)\,\mathbf{s} + \sum_{j=2}^{l}\left(\prod_{i=l}^{j}(\mathbf{H}_i\mathbf{F}_i)\,\mathbf{n}_{\mathrm{r},j-1}\right) + \mathbf{n}_{\mathrm{r},l} \\ &\triangleq \mathbf{A}_l\mathbf{s} + \bar{\mathbf{n}}_{\mathrm{r},l}, \qquad l = 2, \cdots, L-1. \end{aligned} \qquad (4)$$

Here for matrices $\mathbf{X}_i$, $\prod_{i=l}^{k}(\mathbf{X}_i) \triangleq \mathbf{X}_l \cdots \mathbf{X}_k$, $\mathbf{A}_l \triangleq \prod_{i=l}^{1}(\mathbf{H}_i\mathbf{F}_i)$ is the $N_l \times N_{\mathrm{b}}$ equivalent MIMO channel matrix from the source nodes to the lth hop, and $\bar{\mathbf{n}}_{\mathrm{r},l} \triangleq \sum_{j=2}^{l}\left(\prod_{i=l}^{j}(\mathbf{H}_i\mathbf{F}_i)\,\mathbf{n}_{\mathrm{r},j-1}\right) + \mathbf{n}_{\mathrm{r},l}$ is the $N_l \times 1$ equivalent noise vector. Thus the $N_l \times N_l$ covariance matrix of $\bar{\mathbf{n}}_{\mathrm{r},l}$, $\mathbf{C}_l = \mathrm{E}[\bar{\mathbf{n}}_{\mathrm{r},l}\bar{\mathbf{n}}_{\mathrm{r},l}^H], l = 1, \cdots, L-1$, is given by $\mathbf{C}_1 = \mathbf{I}_{N_1}$, and for systems with $L \geq 3$

$$\mathbf{C}_l = \sum_{j=2}^{l}\left(\prod_{i=l}^{j}(\mathbf{H}_i\mathbf{F}_i)\prod_{i=j}^{l}(\mathbf{F}_i^H\mathbf{H}_i^H)\right) + \mathbf{I}_{N_l},$$
$$l = 2, \cdots, L-1.$$

Finally, the last relay node multicasts the precoded signal vector $\mathbf{x}_L$ to all the M destination nodes. From (1)-(4), the

received signal vector at the ith receiver can be written as[1]

$$
\begin{aligned}
\mathbf{y}_{\mathrm{d},i} &= \mathbf{H}_{L,i}\mathbf{x}_L + \mathbf{n}_{\mathrm{d},i} \\
&= \mathbf{H}_{L,i}\mathbf{F}_L \prod_{l=L-1}^{1} (\mathbf{H}_l\mathbf{F}_l)\,\mathbf{s} + \mathbf{H}_{L,i}\mathbf{F}_L \\
&\quad \times \left(\sum_{j=2}^{L-1} \left(\prod_{l=L-1}^{j} (\mathbf{H}_l\mathbf{F}_l)\,\mathbf{n}_{\mathrm{r},j-1} \right) + \mathbf{n}_{\mathrm{r},L-1} \right) + \mathbf{n}_{\mathrm{d},i} \\
&\triangleq \bar{\mathbf{A}}_i\mathbf{s} + \bar{\mathbf{n}}_i, \qquad i = 1,\cdots,M
\end{aligned}
\tag{5}
$$

where $\mathbf{H}_{L,i}$ is the $N_{\mathrm{d}} \times N_{L-1}$ MIMO channel matrix between the last relay node and the ith receiver and $\mathbf{n}_{\mathrm{d},i}$ is the $N_{\mathrm{d}} \times 1$ additive Gaussian noise vector at the ith receiver. In (5), $\bar{\mathbf{A}}_i \triangleq \mathbf{H}_{L,i}\mathbf{F}_L \prod_{l=L-1}^{1}(\mathbf{H}_l\mathbf{F}_l)$ is the $N_{\mathrm{d}} \times N_{\mathrm{b}}$ equivalent MIMO channel between the transmitters and the ith receiver, and

$$
\bar{\mathbf{n}}_i \triangleq \mathbf{H}_{L,i}\mathbf{F}_L\left(\sum_{j=2}^{L-1} \left(\prod_{l=L-1}^{j} (\mathbf{H}_l\mathbf{F}_l)\,\mathbf{n}_{\mathrm{r},j-1} \right) + \mathbf{n}_{\mathrm{r},L-1} \right) + \mathbf{n}_{\mathrm{d},i}
$$

is the $N_{\mathrm{d}} \times 1$ equivalent noise vector at the ith receiver. We assume that all noises are independent and identically distributed (i.i.d.) complex circularly symmetric Gaussian noise with zero mean and unit variance. Thus the $N_{\mathrm{d}} \times N_{\mathrm{d}}$ covariance matrix of $\bar{\mathbf{n}}_i$, $\bar{\mathbf{C}}_i = \mathrm{E}\left[\bar{\mathbf{n}}_i\bar{\mathbf{n}}_i^H\right]$ is given by

$$
\begin{aligned}
\bar{\mathbf{C}}_i &= \mathbf{H}_{L,i}\mathbf{F}_L \left(\sum_{j=2}^{L-1} \left(\prod_{l=L-1}^{j} (\mathbf{H}_l\mathbf{F}_l) \prod_{l=j}^{L-1} (\mathbf{F}_l^H\mathbf{H}_l^H) \right) \right. \\
&\quad \left. + \mathbf{I}_{N_{L-1}} \right) \mathbf{F}_L^H\mathbf{H}_{L,i}^H + \mathbf{I}_{N_{\mathrm{d}}}, \qquad i = 1,\cdots,M.
\end{aligned}
$$

We assume that the relay nodes and the receivers calculate their own weighting matrices. In addition, the first relay node calculates the optimal transmit matrices $\mathbf{B}_i, i = 1,\cdots,K$. Thus the transmitters do not need any channel knowledge, they obtain their precoding matrices by a feedback from the first relay node instead. The lth node, $l = 2,\cdots,L-1$, knows the CSI of its backward channel $\mathbf{H}_{l-1}$ and its forward channel $\mathbf{H}_l$. In practice, the backward CSI can be obtained through standard training methods. The forward CSI required at the lth node is exactly the backward CSI at the $(l+1)$-th node, and thus can be obtained by a feedback from the $(l+1)$-th node. The last relay node can obtain the CSI of $\mathbf{H}_{L,i}, i = 1,\cdots,M$, by a feedback from the ith receiver and the ith receiver knows $\bar{\mathbf{A}}_i$.

We also assume that all channels are quasi-static, i.e., the channel matrices $\mathbf{H}_{1,k}, k = 1,\cdots,K$, $\mathbf{H}_l, l = 2,\cdots,L-1$, and $\mathbf{H}_{L,i}, i = 1,\cdots,M$, are constant throughout a block of transmission. The quasi-static channel model is valid in practice since the fading of wireless relay channels is often relatively slow whenever the mobility of the relays is relatively low, and for static relays, the CSI can be almost constant. Thus, in this way, the necessary CSI can be obtained at each node with a reasonably high precision during the channel training period. Investigating the sensitivity of the algorithms proposed in this paper with respect to outdated CSI in high mobility

scenarios can be an interesting future work. Note that each receiver only needs the CSI of its own channel with the relay and that of the backward $(L-1)$ hop channels. This is a very important assumption for multicasting communication since in a multicasting scenario the receivers are distributed and cannot cooperate.

We aim at improving the system performance through optimizing the transmit and relay precoding matrices. Usually, the system performance is quantified by its QoS and the resources it consumes. The most commonly used QoS metrics include the MSE of the signal waveform estimation, bit-error-rate (BER), system capacity and the output SNR. Interestingly, the aforementioned QoS measures can be expressed in terms of MSE [24]. On the other hand, resources that a multicasting system consumes include the spectrum and transmission power. In the next section, we consider optimizing the source and relay precoding matrices in order to improve the MSE performance of the system. The problem deals with minimizing the worst-user MSE of the signal waveform estimation subjecting to transmission power constraints at the transmitters and the relay nodes.

III. MIN-MAX MSE-BASED TRANSMITTER AND RELAY DESIGN

Due to its simplicity, a linear receiver is used at each destination node to retrieve the transmitted signals. Denoting $\mathbf{W}_i$ as an $N_{\mathrm{d}} \times N_{\mathrm{b}}$ weight matrix at the ith receiver, the estimated signal vector $\hat{\mathbf{s}}_i$ is given by

$$
\hat{\mathbf{s}}_i = \mathbf{W}_i^H\mathbf{y}_{\mathrm{d},i}, \qquad i = 1,\cdots,M.
\tag{6}
$$

From (6), the MSE of the signal waveform estimation at the ith receiver is given by

$$
\begin{aligned}
E_i &= \mathrm{tr}\big(\mathrm{E}\left[(\hat{\mathbf{s}}_i - \mathbf{s})(\hat{\mathbf{s}}_i - \mathbf{s})^H\right]\big) \\
&= \mathrm{tr}\big((\mathbf{W}_i^H\bar{\mathbf{A}}_i - \mathbf{I}_{N_{\mathrm{b}}})(\mathbf{W}_i^H\bar{\mathbf{A}}_i - \mathbf{I}_{N_{\mathrm{b}}})^H + \mathbf{W}_i^H\bar{\mathbf{C}}_i\mathbf{W}_i\big), \\
&\quad i = 1,\cdots,M
\end{aligned}
\tag{7}
$$

where $\mathrm{tr}(\cdot)$ denotes matrix trace. Let us introduce $N_l \times N_l$ matrices

$$
\begin{aligned}
\boldsymbol{\Psi}_l &= \mathbf{A}_l\mathbf{A}_l^H + \mathbf{C}_l = \sum_{j=1}^{l} \left(\prod_{i=l}^{j} (\mathbf{H}_i\mathbf{F}_i) \prod_{i=j}^{l} (\mathbf{F}_i^H\mathbf{H}_i^H) \right) \\
&\quad + \mathbf{I}_{N_l}, \qquad l = 1,\cdots,L-1.
\end{aligned}
\tag{8}
$$

Obviously, the power consumed by the kth transmitter is $\mathrm{tr}(\mathbf{B}_k\mathbf{B}_k^H), k = 1,\cdots,K$. And from (2), the transmission power consumed by the lth relay node is given by

$$
\mathrm{tr}\left(\mathrm{E}\left[\mathbf{x}_{l+1}\mathbf{x}_{l+1}^H\right]\right) = \mathrm{tr}\left(\mathbf{F}_{l+1}\boldsymbol{\Psi}_l\mathbf{F}_{l+1}^H\right), l = 1,\cdots,L-1.
\tag{9}
$$

In the following, we consider minimizing the maximal MSE among all receivers subjecting to power constraints at the transmitter and the relay node for optimizing the transmit and relay matrices.

Given the power constraints at the transmitters and the relay nodes, we aim at minimizing the maximal MSE of the signal waveform estimations among all receivers. This problem formulation is important when the power consumption is a

---

[1]To unify equations for two-hop systems ($L = 2$) and multi-hop systems ($L \geq 3$), we define summations whose upper-limit is smaller than the lower-limit to be zero, e.g., $\sum_{j=2}^{L-1} x_j = 0$, if $L = 2$.

strict system constraint that cannot be relaxed. In this case, the transmit, relay, and receive matrices optimization problem can be formulated as

$$\min_{\{\mathbf{B}_k\},\{\mathbf{F}_l\},\{\mathbf{W}_i\}} \max_i \ E_i \tag{10a}$$

$$\text{s.t. } \text{tr}\left(\mathbf{F}_l \boldsymbol{\Psi}_{l-1} \mathbf{F}_l^H\right) \le P_{\text{r},l}, l=2,\cdots,L \tag{10b}$$

$$\text{tr}(\mathbf{B}_k \mathbf{B}_k^H) \le P_{\text{s},k}, \ k=1,\cdots,K \tag{10c}$$

where $\{\mathbf{W}_i\} \triangleq \{\mathbf{W}_i, i = 1,\cdots,M\}$, $\{\mathbf{F}_l\} \triangleq \{\mathbf{F}_l, l = 2,\cdots,L\}$, $\{\mathbf{B}_k\} \triangleq \{\mathbf{B}_k, k = 1,\cdots,K\}$, (10b) and (10c) are the transmission power constraints at the $(l-1)$-th relay node and the kth transmitter, respectively, and $P_{\text{r},l} > 0$, $P_{\text{s},k} > 0$ are the corresponding power budgets. In the following, we first develop an iterative algorithm to solve the problem (10) and then we propose a low complexity solution to the problem (10) based on some mild approximation.

A. Iterative Approach

This algorithm starts with random $\{\mathbf{F}_l\}$ and $\{\mathbf{B}_k\}$ satisfying (10b) and (10c). In each iteration, the transmitter, relay, and receiver matrices are updated alternatingly through solving convex subproblems.

Firstly, with given $\{\mathbf{F}_l\}$ and $\{\mathbf{B}_k\}$, the optimal $\mathbf{W}_i, i = 1,\cdots,M$, are obtained by solving the unconstrained convex problem of $\min_{\mathbf{W}_i} E_i$, since E_i does not depend on $\mathbf{W}_j$, $j = 1,\cdots,M$, $j \ne i$, and $\mathbf{W}_i$ does not appear in constraints (10b) and (10c). The solution is the well-known linear minimal mean-squared error (MMSE) filter and given by [25]

$$\mathbf{W}_i = \left(\bar{\mathbf{A}}_i \bar{\mathbf{A}}_i^H + \bar{\mathbf{C}}_i\right)^{-1} \bar{\mathbf{A}}_i, \qquad i=1,\cdots,M \tag{11}$$

where $(\cdot)^{-1}$ denotes matrix inversion.

Secondly, with given $\{\mathbf{W}_i\}$ and $\{\mathbf{B}_k\}$, we optimize $\mathbf{F}_l$, $l = 2,\cdots,L$, iteratively. Let us introduce

$$\boldsymbol{\Phi}_{ij} \triangleq \mathbf{W}_i^H \mathbf{H}_{L,i} \boldsymbol{\Delta}_{Lj}, \quad i=1,\cdots,M, j=2,\cdots,L \tag{12}$$

$$\boldsymbol{\Delta}_{lj} \triangleq \begin{cases} \prod_{i=l}^{j+1}(\mathbf{F}_i \mathbf{H}_{i-1}), & j=2,\cdots,l-1, \\ \mathbf{I}_{N_{l-1}}, & j=l, \end{cases} \tag{13}$$
$$l=2,\cdots,L$$

$$\boldsymbol{\Gamma}_{lj} \triangleq \begin{cases} \prod_{i=l-1}^{j}(\mathbf{H}_i \mathbf{F}_i), & j=1,\cdots,l-1, \\ \mathbf{I}_{N_{l-1}}, & j=l, \end{cases} \tag{14}$$
$$l=2,\cdots,L$$

$$\boldsymbol{\Omega}_{il} \triangleq \begin{cases} \mathbf{0}, & l=L, \\ \sum_{j=l+1}^{L} \boldsymbol{\Phi}_{ij} \mathbf{F}_j \mathbf{F}_j^H \boldsymbol{\Phi}_{ij}^H, & l=2,\cdots,L-1, \end{cases} \tag{15}$$
$$i=1,\cdots,M.$$

With fixed $\mathbf{F}_j$, $j = 2,\cdots,L$, $j \ne l$, the MSE in (7) as a function of $\mathbf{F}_l$ can be rewritten as

$$E_i = \text{tr}\Big((\boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Gamma}_{l1} - \mathbf{I}_{N_\text{b}})(\boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Gamma}_{l1} - \mathbf{I}_{N_\text{b}})^H$$
$$+ \sum_{j=2}^{l} \boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Gamma}_{lj} \boldsymbol{\Gamma}_{lj}^H \mathbf{F}_l^H \boldsymbol{\Phi}_{il}^H + \boldsymbol{\Omega}_{il} + \mathbf{W}_i^H \mathbf{W}_i\Big)$$
$$= \text{tr}\big(\boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Psi}_{l-1} \mathbf{F}_l^H \boldsymbol{\Phi}_{il}^H - \boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Gamma}_{l1} - (\boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Gamma}_{l1})^H\big)$$
$$+ \eta_{il} \tag{16}$$

where $\eta_{il} \triangleq N_\text{b} + \text{tr}(\boldsymbol{\Omega}_{il} + \mathbf{W}_i^H \mathbf{W}_i)$ and we used the fact that $\boldsymbol{\Psi}_{l-1} = \sum_{j=1}^{l} \boldsymbol{\Gamma}_{lj} \boldsymbol{\Gamma}_{lj}^H$.

The power constraint in (10b) relevant to $\mathbf{F}_l$ can be written as

$$\text{tr}\left(\sum_{j=1}^{l} \boldsymbol{\Delta}_{nl} \mathbf{F}_l \boldsymbol{\Gamma}_{lj} \boldsymbol{\Gamma}_{lj}^H \mathbf{F}_l^H \boldsymbol{\Delta}_{nl}^H\right) \le P_{\text{r},n} - \beta_{n,l}, n=l,\cdots,L \tag{17}$$

where

$$\beta_{n,l} \triangleq \begin{cases} 0, & n=l \\ \text{tr}(\mathbf{F}_n \mathbf{F}_n^H), & n=l+1 \\ \text{tr}\Big(\mathbf{F}_n\Big(\sum_{j=l+1}^{n-1}\Big(\prod_{i=n-1}^{j}(\mathbf{H}_i \mathbf{F}_i) \\ \quad \times \prod_{i=j}^{n-1}(\mathbf{F}_i^H \mathbf{H}_i^H)\Big) + \mathbf{I}_{N_{n-1}}\Big)\mathbf{F}_n^H\Big), \\ \qquad n=l+2,\cdots,L. \end{cases}$$

Using (16) and (17), the problem of optimizing $\mathbf{F}_l$ can be written as

$$\min_{\mathbf{F}_l} \max_i \ \text{tr}\big(\boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Psi}_{l-1} \mathbf{F}_l^H \boldsymbol{\Phi}_{il}^H - \boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Gamma}_{l1}$$
$$-(\boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Gamma}_{l1})^H\big) + \eta_{il} \tag{18a}$$
$$\text{s.t. } \text{tr}\big(\boldsymbol{\Delta}_{nl} \mathbf{F}_l \boldsymbol{\Psi}_{l-1} \mathbf{F}_l^H \boldsymbol{\Delta}_{nl}^H\big) \le \bar{P}_{\text{r},n,l}, n=l,\cdots,L \tag{18b}$$

where $\bar{P}_{\text{r},n,l} \triangleq P_{\text{r},n} - \beta_{n,l}$.

By introducing a real-valued slack variable t_l, the problem (18) can be rewritten as an SDP problem

$$\min_{t_l, \mathbf{F}_l, \{\boldsymbol{\Theta}_{il}\}, \{\boldsymbol{\Pi}_{nl}\}} \ t_l \tag{19a}$$
$$\text{s.t. } \text{tr}(\boldsymbol{\Theta}_{il}) + \eta_{il} \le t_l, \qquad i=1,\cdots,M \tag{19b}$$
$$\begin{pmatrix} \boldsymbol{\Theta}_{il} + \boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Gamma}_{l1} + (\boldsymbol{\Phi}_{il} \mathbf{F}_l \boldsymbol{\Gamma}_{l1})^H & \boldsymbol{\Phi}_{il} \mathbf{F}_l \\ \mathbf{F}_l^H \boldsymbol{\Phi}_{il}^H & \boldsymbol{\Psi}_{l-1}^{-1} \end{pmatrix}$$
$$\succeq 0, \quad i=1,\cdots,M \tag{19c}$$
$$\text{tr}(\boldsymbol{\Pi}_{nl}) \le \bar{P}_{\text{r},n,l}, \quad n=l,\cdots,L \tag{19d}$$
$$\begin{pmatrix} \boldsymbol{\Pi}_{nl} & \boldsymbol{\Delta}_{nl} \mathbf{F}_l \\ \mathbf{F}_l^H \boldsymbol{\Delta}_{nl}^H & \boldsymbol{\Psi}_{l-1}^{-1} \end{pmatrix} \succeq 0, n=l,\cdots,L \tag{19e}$$

where $\mathbf{A} \succeq 0$ indicates that matrix $\mathbf{A}$ is positive semidefinite (PSD). The problem (19) is a convex SDP problem and can be efficiently solved by the disciplined convex programming toolbox CVX [29], where interior-point method-based solvers such as SeDuMi or SDPT3 are called internally.

Thirdly, we obtain the optimal $\{\mathbf{B}_k\}$ with given $\{\mathbf{F}_l\}$ and $\{\mathbf{W}_i\}$. Let us introduce $\mathbf{G}_{ik} \triangleq \mathbf{W}_i^H \mathbf{H}_{L,i} \mathbf{F}_L$ $\times \prod_{l=L-1}^{2}(\mathbf{H}_l \mathbf{F}_l)\mathbf{H}_{1,k}$, $i = 1,\cdots,M$, $k = 1,\cdots,K$. We can rewrite E_i in (7) as

$$E_i = \sum_{k=1}^{K} \Big[\mathbf{b}_k^H(\mathbf{I}_{N_{\text{b},k}} \otimes (\mathbf{G}_{ik}^H \mathbf{G}_{ik}))\mathbf{b}_k - \big(\text{vec}(\mathbf{G}_{ik,k}^T)\big)^T \mathbf{b}_k$$
$$- \mathbf{b}_k^H \text{vec}(\mathbf{G}_{ik,k}^H)\Big] + \theta_i \tag{20}$$

where $\mathbf{b}_k = \text{vec}(\mathbf{B}_k)$ stands for a vector obtained by stacking all column vectors of $\mathbf{B}_k$ on top of each other, $\theta_i \triangleq \text{tr}(\mathbf{W}_i^H \bar{\mathbf{C}}_i \mathbf{W}_i) + N_\text{b}$, $\mathbf{G}_{ik,k}$ is a matrix containing the $\big(\sum_{j=1}^{k-1} N_{\text{b},j} + 1\big)$-th to $\big(\sum_{j=1}^{k} N_{\text{b},j}\big)$-th rows of $\mathbf{G}_{ik}$, and $\otimes$ denotes the matrix Kronecker product. Here we used the identity of $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$. The MSE in (20) can be rewritten as

$$E_i = \mathbf{b}^H \mathbf{G}_i \mathbf{b} - \mathbf{c}_i^H \mathbf{b} - \mathbf{b}^H \mathbf{c}_i + \theta_i \tag{21}$$

where

$$\mathbf{G}_i \triangleq \mathrm{bd}\big(\mathbf{I}_{N_{\mathrm{b},1}} \otimes (\mathbf{G}_{i1}^H \mathbf{G}_{i1}), \cdots, \mathbf{I}_{N_{\mathrm{b},K}} \otimes (\mathbf{G}_{iK}^H \mathbf{G}_{iK})\big)$$

$$\mathbf{c}_i \triangleq \big[(\mathrm{vec}(\mathbf{G}_{i1,1}^H))^T, \cdots, (\mathrm{vec}(\mathbf{G}_{iK,K}^H))^T\big]^T$$

$$\mathbf{b} \triangleq \big[\mathbf{b}_1^T, \cdots, \mathbf{b}_K^T\big]^T.$$

By introducing $\mathbf{M}_{lk} \triangleq \mathbf{F}_l \prod_{i=l-1}^{2}(\mathbf{H}_i \mathbf{F}_i)\mathbf{H}_{1,k}$, $l = 2, \cdots, L$, the power constraints in (10b) can be rewritten as

$$\mathbf{b}^H \mathbf{M}_l \mathbf{b} \leq Q_{\mathrm{r},l}, \qquad l = 2, \cdots, L \qquad (22)$$

where $\mathbf{M}_l \triangleq \mathrm{bd}\big(\mathbf{I}_{N_{\mathrm{b},1}} \otimes (\mathbf{M}_{l1}^H \mathbf{M}_{l1}), \cdots, \mathbf{I}_{N_{\mathrm{b},K}} \otimes (\mathbf{M}_{lK}^H \mathbf{M}_{lK})\big)$, $Q_{\mathrm{r},l} = P_{\mathrm{r},l} - \mathrm{tr}(\mathbf{F}_l \mathbf{C}_{l-1} \mathbf{F}_l^H)$. Using (21) and (22), the problem (10) can be written as

$$\min_{\mathbf{b}} \quad \max_{i} \quad \mathbf{b}^H \mathbf{G}_i \mathbf{b} - \mathbf{c}_i^H \mathbf{b} - \mathbf{b}^H \mathbf{c}_i + \theta_i \qquad (23a)$$

$$\mathrm{s.t.} \quad \mathbf{b}^H \mathbf{M}_l \mathbf{b} \leq Q_{\mathrm{r},l}, \qquad l = 2, \cdots, L \qquad (23b)$$

$$\mathbf{b}^H \mathbf{D}_k \mathbf{b} \leq P_{\mathrm{s},k}, \qquad k = 1, \cdots, K \qquad (23c)$$

where $\mathbf{D}_k \triangleq \mathrm{bd}(\mathbf{D}_{k1}, \mathbf{D}_{k2}, \cdots, \mathbf{D}_{kK})$ with $\mathbf{D}_{kk} = \mathbf{I}_{N_{\mathrm{s},k} N_{\mathrm{b},k}}$ and $\mathbf{D}_{kj} = \mathbf{0}$, $j = 1, \cdots, K$, $j \neq k$.

By introducing a real-valued slack variable t_{s}, the problem (23) can be cast as an SDP problem as

$$\min_{t_{\mathrm{s}}, \mathbf{b}} \quad t_{\mathrm{s}} \qquad (24a)$$

$$\mathrm{s.t.} \quad \begin{pmatrix} t_{\mathrm{s}} - \theta_i + \mathbf{c}_i^H \mathbf{b} + \mathbf{b}^H \mathbf{c}_i & \mathbf{b}^H \\ \mathbf{b} & \mathbf{G}_i^{-1} \end{pmatrix} \succcurlyeq 0,$$

$$i = 1, \cdots, M \qquad (24b)$$

$$\begin{pmatrix} Q_{\mathrm{r},l} & \mathbf{b}^H \\ \mathbf{b} & \mathbf{M}_l^{-1} \end{pmatrix} \succcurlyeq 0, \qquad l = 2, \cdots, L \qquad (24c)$$

$$\begin{pmatrix} P_{\mathrm{s},k} & \mathbf{b}^H \mathbf{D}_k^{\frac{1}{2}} \\ \mathbf{D}_k^{\frac{1}{2}} \mathbf{b} & \mathbf{I}_p \end{pmatrix} \succcurlyeq 0, \qquad k = 1, \cdots, K \qquad (24d)$$

where $p \triangleq \sum_{k=1}^{K} N_{\mathrm{s},k} N_{\mathrm{b},k}$ The problem (24) can be efficiently solved by the disciplined convex programming toolbox CVX [29]. Now the original transmitter, relay, and receiver matrices optimization problem (10) can be solved by an iterative technique as shown in Table I, where $\|\cdot\|_1$ denotes the matrix maximum absolute column sum norm, ε is a small positive number close to zero and the superscript (n) denotes the number of iterations.

TABLE I
PROCEDURE OF SOLVING THE PROBLEM (10) BY THE PROPOSED
ITERATIVE ALGORITHM

1) Randomly initialize the algorithm with $\{\mathbf{F}_l^{(0)}\}$ and $\{\mathbf{B}_k^{(0)}\}$ satisfying the constraints (10b) and (10c); Set $n = 0$.
2) Update $\{\mathbf{W}_i^{(n)}\}$ using $\{\mathbf{B}_k^{(n)}\}$ and $\{\mathbf{F}_l^{(n)}\}$ as in (11).
3) Solve the subproblem (19) for $l = 2, \cdots, L$, to update $\mathbf{F}_l^{(n+1)}$ using given $\mathbf{F}_j^{(n)}$, $j = 2, \cdots, L$, $j \neq l$, $\{\mathbf{W}_i^{(n)}\}$ and $\{\mathbf{B}_k^{(n)}\}$.
4) Solve the subproblem (24) using known $\{\mathbf{F}_l^{(n+1)}\}$ and $\{\mathbf{W}_i^{(n)}\}$ to obtain $\{\mathbf{B}_k^{(n+1)}\}$.
5) If $\max_k \|\mathbf{B}_k^{(n+1)} - \mathbf{B}_k^{(n)}\|_1 \leq \varepsilon$, then end. Otherwise, let $n := n + 1$ and go to step 2.

In the iterative algorithm, the source, relay, and receiver matrices are updated alternatingly through solving the corresponding subproblems. Since all the subproblems are convex

optimization problems, the conditional update of each variable can not increase the corresponding objective function and hence the objective function (10a). Therefore, each conditional update of $\{\mathbf{W}_i^{(n)}\}$, $\{\mathbf{F}_l^{(n)}\}$, and $\{\mathbf{B}_k^{(n)}\}$ may either decrease or maintain but cannot increase the objective function (10a). Note that the constraints in the problem (10) are always satisfied with every conditional update. Thus a monotonic convergence of $\{\mathbf{W}_i^{(n)}\}$, $\{\mathbf{F}_l^{(n)}\}$, and $\{\mathbf{B}_k^{(n)}\}$ towards (at least) a locally optimal solution follows directly from this observation.

The major computational complexity of the iterative approach involves solving the SDP problems (19) and (24) in each iteration. Thus the overall computational complexity of the iterative algorithm is significantly high. In the following section, we develop a simplified algorithm for the min-max MSE problem such that the nearly optimal transmit and relay matrices can be designed with a significantly reduced computational complexity.

B. Simplified Approach

By substituting (11) back into (7), we have

$$E_i = \mathrm{tr}\Big(\big[\mathbf{I}_{N_{\mathrm{b}}} + \bar{\mathbf{A}}_i^H \bar{\mathbf{C}}_i^{-1} \bar{\mathbf{A}}_i\big]^{-1}\Big), \quad i = 1, \cdots, M. \quad (25)$$

Thus we can equivalently rewrite the problem (10) as

$$\min_{\{\mathbf{B}_k\}, \{\mathbf{F}_l\}} \quad \max_{i} \quad \mathrm{tr}\Big(\big[\mathbf{I}_{N_{\mathrm{b}}} + \bar{\mathbf{A}}_i^H \bar{\mathbf{C}}_i^{-1} \bar{\mathbf{A}}_i\big]^{-1}\Big) \qquad (26a)$$

$$\mathrm{s.t.} \quad \mathrm{tr}\big(\mathbf{F}_l \mathbf{\Psi}_{l-1} \mathbf{F}_l^H\big) \leq P_{\mathrm{r},l}, \ l = 2, \cdots, L \quad (26b)$$

$$\mathrm{tr}(\mathbf{B}_k \mathbf{B}_k^H) \leq P_{\mathrm{s},k}, \quad k = 1, \cdots, K. \quad (26c)$$

The min-max problem (26) is highly nonconvex with matrix variables, and an exactly optimal solution is very hard to obtain with a reasonable computational complexity (non-exhaustive searching). In the following, we propose a low complexity solution to the problem (26).

It can be shown similar to [26] that the optimal $\mathbf{F}_l, l = 2, \cdots, L$, for *each link* with the input-output relationship given by (5) has the generic structure of

$$\mathbf{F}_l = \mathbf{T}_l \mathbf{D}_l^H, \quad l = 2, \cdots, L \qquad (27)$$

where $\mathbf{D}_l = (\mathbf{A}_{l-1} \mathbf{A}_{l-1}^H + \mathbf{C}_{l-1})^{-1} \mathbf{A}_{l-1}, l = 2, \cdots, L$, is the $N_{l-1} \times N_{\mathrm{b}}$ weight matrix of the linear MMSE filter for the received signal vector at the $(l-1)$-th relay node given by $\mathbf{y}_{l-1} = \mathbf{A}_{l-1} \mathbf{s} + \bar{\mathbf{n}}_{\mathrm{r},l-1}$ (see (4)), and the linear filter $\mathbf{T}_l$ can be treated as the $N_{l-1} \times N_{\mathrm{b}}$ transmit precoding matrix for the effective lth hop MIMO channel, which will be designed later.

Using the optimal $\mathbf{F}_l, l = 2, \cdots, L$, in (27), the MSE of the signal waveform estimation at the ith receiver in (25) can be equivalently decomposed to

$$\begin{aligned} E_i &= \mathrm{tr}\Big(\big[\mathbf{I}_{N_{\mathrm{b}}} + \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{H}_1 \mathbf{F}_1\big]^{-1}\Big) \\ &\quad + \sum_{l=2}^{L-1} \mathrm{tr}\Big(\big[\mathbf{R}_l^{-1} + \mathbf{T}_l^H \mathbf{H}_l^H \mathbf{H}_l \mathbf{T}_l\big]^{-1}\Big) \\ &\quad + \mathrm{tr}\Big(\big[\mathbf{R}_L^{-1} + \mathbf{T}_L^H \mathbf{H}_{L,i}^H \mathbf{H}_{L,i} \mathbf{T}_L\big]^{-1}\Big), \\ &\qquad\qquad\qquad i = 1, \cdots, M \end{aligned} \qquad (28)$$

where the $N_{l-1} \times N_{l-1}$ matrix

$$\mathbf{R}_l = \mathbf{A}_{l-1}^H \boldsymbol{\Psi}_{l-1}^{-1} \mathbf{A}_{l-1}, \quad l = 2, \cdots, L. \tag{29}$$

Note that the first term in (28) $\mathrm{tr}\left(\left[\mathbf{I}_{N_\mathrm{b}} + \mathbf{F}_1^H \mathbf{H}_1^H \mathbf{H}_1 \mathbf{F}_1\right]^{-1}\right)$ is actually the MSE of estimating the signal vector $\mathbf{s}$ from the received signal vector (1) at the first relay node using the linear MMSE receiver $\mathbf{D}_2$, while the remaining terms in (28) can be viewed as the increment of the MSE introduced by the corresponding hop. Interestingly, matrix $\mathbf{R}_l$, $l = 2, \cdots, L$, in (29) is in fact the covariance matrix of $\mathbf{z}_l \triangleq \mathbf{D}_l^H \mathbf{y}_{l-1}$ as $\mathbf{R}_l = \mathrm{E}[\mathbf{z}_l \mathbf{z}_l^H] = \mathbf{D}_l^H \mathrm{E}[\mathbf{y}_{l-1} \mathbf{y}_{l-1}^H] \mathbf{D}_l$. It can be seen from (28) that the effect of noise in the first-hop is reflected by $\mathbf{I}_{N_\mathrm{b}}$ in the first term and that of the lth hop is reflected by $\mathbf{R}_l^{-1}$. Using the optimal structure of $\mathbf{F}_l$ in (27), the transmission power consumed by each relay node can be rewritten as

$$\mathrm{tr}\left(\mathbf{F}_l \boldsymbol{\Psi}_{l-1} \mathbf{F}_l^H\right) = \mathrm{tr}(\mathbf{T}_l \mathbf{R}_l \mathbf{T}_l^H), \quad l = 2, \cdots, L. \tag{30}$$

And the first term in (28) can be rewritten as

$$\mathrm{tr}\left(\left[\mathbf{I}_{N_1} + \sum_{k=1}^K \mathbf{H}_{1,k} \mathbf{P}_k \mathbf{H}_{1,k}^H\right]^{-1}\right) + N_\mathrm{b} - N_1 \tag{31}$$

where $\mathbf{P}_k = \mathbf{B}_k \mathbf{B}_k^H$, $k = 1, \cdots, K$, is the $N_{\mathrm{s},k} \times N_{\mathrm{s},k}$ covariance matrix of the signal transmitted by the kth user. Therefore, the problem (26) can be equivalently rewritten as

$$\min_{\{\mathbf{P}_k\},\{\mathbf{T}_l\}} \max_i \quad \mathrm{tr}\left(\left[\mathbf{I}_{N_1} + \sum_{k=1}^K \mathbf{H}_{1,k} \mathbf{P}_k \mathbf{H}_{1,k}^H\right]^{-1}\right)$$
$$+ \sum_{l=2}^{L-1} \mathrm{tr}\left(\left[\mathbf{R}_l^{-1} + \mathbf{T}_l^H \mathbf{H}_l^H \mathbf{H}_l \mathbf{T}_l\right]^{-1}\right)$$
$$+ \mathrm{tr}\left(\left[\mathbf{R}_L^{-1} + \mathbf{T}_L^H \mathbf{H}_{L,i}^H \mathbf{H}_{L,i} \mathbf{T}_L\right]^{-1}\right) \tag{32a}$$
$$\text{s.t.} \quad \mathrm{tr}(\mathbf{T}_l \mathbf{R}_l \mathbf{T}_l^H) \leq P_{\mathrm{r},l}, \quad l = 2, \cdots, L \tag{32b}$$
$$\mathrm{tr}(\mathbf{P}_k) \leq P_{\mathrm{s},k}, \quad \mathbf{P}_k \succcurlyeq 0, \quad k = 1, \cdots, K \tag{32c}$$

By applying the matrix inversion lemma $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} \left(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1}\right)^{-1} \mathbf{DA}^{-1}$, matrix $\mathbf{R}_l$ in (29) can be rewritten as

$$\mathbf{R}_l = \mathbf{A}_{l-1}^H \left(\mathbf{C}_{l-1}^{-1} - \mathbf{C}_{l-1}^{-1} \mathbf{A}_{l-1} \left(\mathbf{A}_{l-1}^H \mathbf{C}_{l-1}^{-1} \mathbf{A}_{l-1} + \mathbf{I}_{N_\mathrm{b}}\right)^{-1}\right.$$
$$\left. \times \mathbf{A}_{l-1}^H \mathbf{C}_{l-1}^{-1}\right) \mathbf{A}_{l-1}$$
$$= \mathbf{A}_{l-1}^H \mathbf{C}_{l-1}^{-1} \mathbf{A}_{l-1} \left(\mathbf{A}_{l-1}^H \mathbf{C}_{l-1}^{-1} \mathbf{A}_{l-1} + \mathbf{I}_{N_\mathrm{b}}\right)^{-1},$$
$$l = 2, \cdots, L. \tag{33}$$

An interesting observation from (33) is that with increasing SNR, the term $\mathbf{A}_{l-1}^H \mathbf{C}_{l-1}^{-1} \mathbf{A}_{l-1}$ approaches infinity. And at a (moderately) high SNR level, there is $\mathbf{A}_{l-1}^H \mathbf{C}_{l-1}^{-1} \mathbf{A}_{l-1} \gg \mathbf{I}_{N_\mathrm{b}}$. Here for matrices $\mathbf{X}$ and $\mathbf{Y}$, $\mathbf{X} \gg \mathbf{Y}$ indicates that the eigenvalues of $(\mathbf{X} - \mathbf{Y})$ are much greater than zero. Thus, we can approximate $\mathbf{R}_l$, $l = 2, \cdots, L$, as $\mathbf{I}_{N_\mathrm{b}}$ for the high SNR case [27]. As a consequence, in (32a), $\mathrm{tr}\left(\left[\mathbf{R}_l^{-1} + \mathbf{T}_l^H \mathbf{H}_l^H \mathbf{H}_l \mathbf{T}_l\right]^{-1}\right)$ can be closely upper-bounded by $\mathrm{tr}\left(\left[\mathbf{I}_{N_\mathrm{b}} + \mathbf{T}_l^H \mathbf{H}_l^H \mathbf{H}_l \mathbf{T}_l\right]^{-1}\right)$ and $\mathrm{tr}\left(\left[\mathbf{R}_L^{-1} + \mathbf{T}_L^H \mathbf{H}_{L,i}^H \mathbf{H}_{L,i} \mathbf{T}_L\right]^{-1}\right)$ can be closely upper-bounded by $\mathrm{tr}\left(\left[\mathbf{I}_{N_\mathrm{b}} + \mathbf{T}_L^H \mathbf{H}_{L,i}^H \mathbf{H}_{L,i} \mathbf{T}_L\right]^{-1}\right)$, $l = 2, \cdots, L-$

1, $i = 1, \cdots, M$, and the tightness of these bounds increase with the increase in SNR. Therefore, the problem (32) can be approximated as

$$\min_{\{\mathbf{P}_k\},\{\mathbf{T}_l\}} \max_i \quad \mathrm{tr}\left(\left[\mathbf{I}_{N_1} + \sum_{k=1}^K \mathbf{H}_{1,k} \mathbf{P}_k \mathbf{H}_{1,k}^H\right]^{-1}\right)$$
$$+ \sum_{l=2}^{L-1} \mathrm{tr}\left(\left[\mathbf{I}_{N_\mathrm{b}} + \mathbf{T}_l^H \mathbf{H}_l^H \mathbf{H}_l \mathbf{T}_l\right]^{-1}\right)$$
$$+ \mathrm{tr}\left(\left[\mathbf{I}_{N_\mathrm{b}} + \mathbf{T}_L^H \mathbf{H}_{L,i}^H \mathbf{H}_{L,i} \mathbf{T}_L\right]^{-1}\right) \tag{34a}$$
$$\text{s.t.} \quad \mathrm{tr}(\mathbf{T}_l \mathbf{T}_l^H) \leq P_{\mathrm{r},l}, \quad l = 2, \cdots, L \tag{34b}$$
$$\mathrm{tr}(\mathbf{P}_k) \leq P_{\mathrm{s},k}, \quad \mathbf{P}_k \succcurlyeq 0, \ k = 1, \cdots, K. \tag{34c}$$

We would like to mention that since $\mathrm{tr}(\mathbf{T}_l \mathbf{T}_l^H) > \mathrm{tr}(\mathbf{T}_l \mathbf{R}_l \mathbf{T}_l^H)$, $l = 2, \cdots, L$, if $\mathrm{tr}(\mathbf{T}_l \mathbf{T}_l^H) = p$, then $\mathrm{tr}(\mathbf{T}_l \mathbf{R}_l \mathbf{T}_l^H) < p$. This indicates that due to the approximation in (34b), the transmission power available at each relay node is not fully utilized in the case of the low SNR. We can simply scale the relay matrices obtained from solving the problem (34) to compensate such loss and make the best use of the available power budget (10b) at the relay nodes. Let $\alpha_l = \sqrt{\frac{P_{\mathrm{r},l}}{\mathrm{tr}\left(\mathbf{F}_l \boldsymbol{\Psi}_{l-1} \mathbf{F}_l^H\right)}}$, $l = 2, \cdots, L$. In order to satisfy the power constraints with equality, we obtain $\mathbf{F}_l^{\mathrm{opt}} = \alpha_l \mathbf{F}_l$, $l = 2, \cdots, L$.

Interestingly, it can be seen from the problem (34) that $\mathbf{T}_l$, $l = 2, \cdots, L$, do not affect the first term of the objective function (34a) and $\mathbf{P}_k$, $k = 1, \cdots, K$, are irrelevant to the remaining terms of (34a). This fact implies that the objective function (34a) and the constraints (34b) and (34c) are decoupled with respect to the optimization variables $\{\mathbf{P}_k\}$ and $\{\mathbf{T}_l\}$. In this case, matrices $\{\mathbf{P}_k\}$ can be determined independent of $\{\mathbf{T}_l\}$, and vice-versa, which greatly simplifies the design of the transmit and relay matrices. Therefore, with the (relatively) high SNR assumption, the problem (34) can be decomposed into the following source covariance matrices optimization problem

$$\min_{\{\mathbf{P}_k\}} \quad \mathrm{tr}\left(\left[\mathbf{I}_{N_1} + \sum_{k=1}^K \mathbf{H}_{1,k} \mathbf{P}_k \mathbf{H}_{1,k}^H\right]^{-1}\right) \tag{35a}$$
$$\text{s.t.} \quad \mathrm{tr}(\mathbf{P}_k) \leq P_{\mathrm{s},k}, \quad \mathbf{P}_k \succcurlyeq 0, \quad k = 1, \cdots, K \tag{35b}$$

and for $L \geq 3$, the relay amplifying matrix optimization problem for each $l = 2, \cdots, L-1$

$$\min_{\mathbf{T}_l} \quad \mathrm{tr}\left(\left[\mathbf{I}_{N_\mathrm{b}} + \mathbf{T}_l^H \mathbf{H}_l^H \mathbf{H}_l \mathbf{T}_l\right]^{-1}\right) \tag{36a}$$
$$\text{s.t.} \quad \mathrm{tr}(\mathbf{T}_l \mathbf{T}_l^H) \leq P_{\mathrm{r},l} \tag{36b}$$

and for the last relay node as

$$\min_{\mathbf{T}_L} \max_i \quad \mathrm{tr}\left(\left[\mathbf{I}_{N_\mathrm{b}} + \mathbf{T}_L^H \mathbf{H}_{L,i}^H \mathbf{H}_{L,i} \mathbf{T}_L\right]^{-1}\right) \tag{37a}$$
$$\text{s.t.} \quad \mathrm{tr}(\mathbf{T}_L \mathbf{T}_L^H) \leq P_{\mathrm{r},L}. \tag{37b}$$

The source covariance matrices optimization problem (35) can be solved as follows. By introducing a PSD matrix $\mathbf{X}$ with $\mathbf{X} \succcurlyeq \left[\mathbf{I}_{N_1} + \sum_{k=1}^K \mathbf{H}_{1,k} \mathbf{P}_k \mathbf{H}_{1,k}^H\right]^{-1}$, where $\mathbf{A} \succcurlyeq \mathbf{B}$ means

that $\mathbf{A} - \mathbf{B} \succcurlyeq 0$, and using the Schur complement [28], the problem (35) can be converted to the problem of

$$\min_{\{\mathbf{P}_k\},\mathbf{X}} \quad \text{tr}(\mathbf{X}) \tag{38a}$$

$$\text{s.t.} \quad \begin{pmatrix} \mathbf{X} & \mathbf{I}_{N_1} \\ \mathbf{I}_{N_1} & \mathbf{I}_{N_1} + \sum_{k=1}^{K} \mathbf{H}_{1,k} \mathbf{P}_k \mathbf{H}_{1,k}^H \end{pmatrix} \succcurlyeq 0 \tag{38b}$$

$$\text{tr}(\mathbf{P}_k) \leq P_{\text{s},k}, \quad \mathbf{P}_k \succcurlyeq 0, \ k = 1, \cdots, K. \tag{38c}$$

The problem (38) is a convex SDP problem which can be efficiently solved by the disciplined convex programming toolbox CVX [29], where interior-point method-based solvers such as SeDuMi or SDPT3 are called internally, at a complexity order of $\mathcal{O}\left((\sum_{k=1}^{K} N_{\text{s},k}^2 + K)^{3.5}\right)$ [30]. Since we assume that the first relay node calculates the optimal transmit matrices $\mathbf{B}_i, i = 1, \cdots, K$, the complexity at the first relay node will increase at a complexity order of $\mathcal{O}\left((\sum_{k=1}^{K} N_{\text{s},k}^2 + K)^{3.5}\right)$ as the number of transmitters K and/or the number of transmit antennas $N_{\text{s},k}, \forall k$, increase.

For multi-hop systems with $L \geq 3$, let us now introduce the eigenvalue decomposition (EVD) of $\mathbf{H}_l^H \mathbf{H}_l = \mathbf{V}_l \mathbf{\Lambda}_l \mathbf{V}_l^H, l = 2, \cdots, L - 1$, where the dimensions of $\mathbf{V}_l$ and $\mathbf{\Lambda}_l$ are $N_{l-1} \times N_{l-1}$ and the diagonal elements of $\mathbf{\Lambda}_l$ are sorted in decreasing order. According to Lemma 2 in [26], the solution to the problem (36) is given by

$$\mathbf{T}_l = \mathbf{V}_{l,1} \mathbf{\Sigma}_l \mathbf{\Pi}, \qquad l = 2, \cdots, L - 1 \tag{39}$$

where $\mathbf{V}_{l,1}$ contains the leftmost N_{b} columns of $\mathbf{V}_l$, $\mathbf{\Pi}$ can be any $N_{\text{b}} \times N_{\text{b}}$ unitary matrix, and $\mathbf{\Sigma}_l$ is an $N_{\text{b}} \times N_{\text{b}}$ diagonal matrix. Substituting (39) back into (36), we find that the ith diagonal element of $\mathbf{\Sigma}_l$ is given by $\sigma_{l,i} = \left[\frac{1}{\lambda_{l,i}}\left(\sqrt{\frac{\lambda_{l,i}}{\mu_l}} - 1\right)^+\right]^{\frac{1}{2}}, i = 1, \cdots, N_{\text{b}}$, where $\lambda_{l,i}$ denotes the ith diagonal element of $\mathbf{\Lambda}_l$, $(x)^+ \triangleq \max(x, 0)$, and $\mu_l > 0$ is the Lagrangian multiplier and the solution to the nonlinear equation of $\sum_{i=1}^{N_{\text{b}}} \frac{1}{\lambda_{l,i}}\left(\sqrt{\frac{\lambda_{l,i}}{\mu_l}} - 1\right)^+ = P_{\text{r},l}$.

By introducing $\mathbf{T}_L \mathbf{T}_L^H \triangleq \mathbf{Q}$, the problem (37) can be equivalently rewritten as

$$\min_{\mathbf{Q}} \ \max_i \text{tr}\left(\left[\mathbf{I}_{N_{\text{d}}} + \mathbf{H}_{L,i} \mathbf{Q} \mathbf{H}_{L,i}^H\right]^{-1}\right) + N_{\text{b}} - N_{\text{d}} \tag{40a}$$

$$\text{s.t.} \quad \text{tr}(\mathbf{Q}) \leq P_{\text{r},L}, \qquad \mathbf{Q} \succcurlyeq 0. \tag{40b}$$

By introducing $\mathbf{Y}_i \succcurlyeq \left[\mathbf{I}_{N_{\text{d}}} + \mathbf{H}_{L,i} \mathbf{Q} \mathbf{H}_{L,i}^H\right]^{-1}, i = 1, \cdots, M$, and a real-valued slack variable t, the problem (40) can be equivalently transformed to

$$\min_{t,\mathbf{Q},\{\mathbf{Y}_i\}} \ t \tag{41a}$$

$$\text{s.t.} \quad \text{tr}(\mathbf{Y}_i) \leq t, \qquad i = 1, \cdots, M \tag{41b}$$

$$\text{tr}(\mathbf{Q}) \leq P_{\text{r},L} \tag{41c}$$

$$\begin{pmatrix} \mathbf{Y}_i & \mathbf{I}_{N_{\text{d}}} \\ \mathbf{I}_{N_{\text{d}}} & \mathbf{I}_{N_{\text{d}}} + \mathbf{H}_{L,i} \mathbf{Q} \mathbf{H}_{L,i}^H \end{pmatrix} \succcurlyeq 0,$$
$$i = 1, \cdots, M \tag{41d}$$

$$t \geq 0, \qquad \mathbf{Q} \succcurlyeq 0 \tag{41e}$$

where $\{\mathbf{Y}_i\} \triangleq \{\mathbf{Y}_i, i = 1, \cdots, M\}$ and we use the Schur complement to obtain (41d). Note that in the above formulation, t provides an MSE upper-bound for the last-hop channels.

The problem (41) is an SDP problem which can be efficiently solved by the disciplined convex programming toolbox CVX [29] at a maximal complexity order of $\mathcal{O}\left((N_{L-1}^2 + M + 1)^{3.5}\right)$ [30]. Since most of the computation task in solving the problem (36) involves performing SVD and calculating the power loading parameters, the computation overhead is negligible compared with that of solving the problems (38) and (41). Note that the problem (34) can also be directly formulated as an SDP problem which can be solved using interior point-based solvers at a complexity order that is at most $\mathcal{O}\left((\sum_{k=1}^{K} N_{\text{s},k}^2 + \sum_{l=1}^{L-1} N_l^2 + K + L + M)^{3.5}\right)$. Therefore, solving the decoupled transmit and relay precoding problems (38), (36), and (41) has a much smaller computational complexity compared with directly solving the problem (34). The proposed simplified algorithm is summarized in Table II.

TABLE II
PROCEDURE OF SOLVING THE PROBLEM (10) BY THE PROPOSED
SIMPLIFIED ALGORITHM

1) Optimize transmit precoding matrices:
   a) Solve the subproblem (38) to obtain source covariance matrices $\mathbf{P}_k, k = 1, \cdots, K$.
   b) The source precoding matrices $\mathbf{B}_k, k = 1, \cdots, K$, are calculated as $\mathbf{B}_k = \mathbf{U}_{k,1} \mathbf{\Upsilon}_{k,1}^{\frac{1}{2}}$, where $\mathbf{P}_k = \mathbf{U}_k \mathbf{\Upsilon}_k \mathbf{U}_k^H$ is the eigenvalue decomposition (EVD) of $\mathbf{P}_k$, $\mathbf{\Upsilon}_{k,1}$ contains $N_{\text{b},k}$ largest eigenvalues of $\mathbf{P}_k$, and $\mathbf{U}_{k,1}$ contains $N_{\text{b},k}$ columns of $\mathbf{U}_k$ associated with $\mathbf{\Upsilon}_{k,1}$.

2) Optimize the first $L - 2$ relay amplifying matrices:
   a) Obtain $\mathbf{T}_l, l = 2, \cdots, L - 1$, from (39).
   b) Calculate $\mathbf{D}_l = (\mathbf{A}_{l-1} \mathbf{A}_{l-1}^H + \mathbf{C}_{l-1})^{-1} \mathbf{A}_{l-1}, l = 2, \cdots, L - 1$.
   c) Obtain $\mathbf{F}_l, l = 2, \cdots, L - 1$, using (27).
   d) Scale $\mathbf{F}_l, l = 2, \cdots, L - 1$, to utilize the maximal available transmit power at the $(l - 1)$th relay node.

3) Optimize the last relay matrix:
   a) Solve the subproblem (41) to obtain $\mathbf{T}_L$ from $\mathbf{Q}$.
   b) Calculate $\mathbf{D}_L = (\mathbf{A}_{L-1} \mathbf{A}_{L-1}^H + \mathbf{C}_{L-1})^{-1} \mathbf{A}_{L-1}$.
   c) Obtain $\mathbf{F}_L$ using (27).
   d) Scale $\mathbf{F}_L$ to utilize the maximal available transmit power at the $(L - 1)$th relay node.

4) Calculate the receiver weight matrices $\mathbf{W}_i, i = 1, \cdots, M$, using (11).

We would like to mention that for a point-to-point two-hop MIMO relay system, it has been shown in [27] through numerical examples that the high SNR approximation provides negligible performance loss in all SNR range in comparison to the optimal designs. For the multicasting MIMO relay system addressed in this paper, the exactly optimal solution for the transmit and relay precoding matrices are intractable. However, by using the high SNR approximation, the nearly optimal transmit and relay matrices can be designed with a significantly reduced computational complexity.

IV. NUMERICAL EXAMPLES

In this section, we study the performance of the proposed multi-hop multicasting MIMO relay optimization algorithms through numerical simulations. For simplicity, we assume that the transmitters and the receivers are equipped with N_{s} and N_{d} antennas each, respectively, $P_{\text{s},k} = P_{\text{s}}, k = 1, \cdots, K$, and $P_{\text{r},l} = P_{\text{r}}, l = 1, \cdots, L - 1$. We simulate a flat Rayleigh fading environment where the channel matrices have zero-mean entries with variances $1/N_{\text{s}}$ for $\mathbf{H}_{1,k}, k = 1, \cdots, K$,

$1/N_{l-1}$ for $\mathbf{H}_l$, $l = 2, \cdots, L-1$, and $1/N_{L-1}$ for $\mathbf{H}_{L,i}$, $i = 1, \cdots, M$, respectively. All simulation results are averaged over 500 independent channel realizations.

Since there is no existing solution for the problem addressed in this paper, we compare the performance of the proposed min-max MSE algorithms in Section III with the naive amplify-and-forward (NAF) algorithm in terms of both MSE and BER. The NAF algorithm simply amplifies and forwards the modulated signals at each transmitter and the received signals at the relay nodes assigning equal power to each data stream such that $\mathbf{B}_k = \sqrt{P_s/N_s}\,\mathbf{I}_{N_s}$, $k = 1, \cdots, K$, $\mathbf{F}_l = \sqrt{P_r/\mathrm{tr}(\mathbf{\Psi}_{l-1})}\,\mathbf{I}_{N_{l-1}}$, $l = 2, \cdots, L$.

In the first example, we compare the performance of the proposed algorithms with the NAF approach in terms of the MSE normalized by the number of data streams (NMSE) for $K = 2$, $L = 3$, $M = 2$, $N_s = 2, N_1 = 8, N_2 = 4$, and $N_d = 8$. Fig. 2 shows the MSE performance of the algorithms versus P_s with $P_r = 20$dB. For the simplified algorithm, we plot the NMSE of the user with the worst channel (Worst) and the average of all the users (Avg.). For the proposed iterative algorithm, the NMSE of the user with the worst channel is shown. Our results clearly demonstrate the better performance of the proposed joint transmitter and relay optimization algorithms. It can be seen that the proposed iterative algorithm consistently yields the lowest worst-user MSE over the entire P_s region. The NAF algorithm has much higher MSE compared with the proposed schemes even with very high transmission power. Note that the average- and the worst-user MSEs from the simplified algorithm are almost identical in this example.

It can also be observed from Fig. 2 that the iterative and simplified algorithms have a similar worst-user MSE performance, even at low SNRs. This indicates that the (moderately) high SNR assumption in the simplified algorithm introduces negligible performance loss in the whole SNR range in comparison to the iterative design. Note that the computational complexity of the simplified algorithm is less than the complexity of one iteration of the iterative algorithm, making it very attractive for practical MIMO relay multicasting systems. We will focus on the simplified algorithm in the following simulation examples.

In the second example, we compare the MSE performance of the proposed simplified algorithm for different number of receivers. We set $K = 2$, $L = 3, N_s = 2, N_1 = 6, N_2 = 4$, and $N_d = 6$. Fig. 3 shows the worst-user MSE and the average MSE of all the users for the simplified algorithm versus P_r for $P_s = 20$dB. It can be clearly seen from Fig. 3 that as the number of receivers increases, the average- and the worst-user MSE keep increasing. This is reasonable since it is more likely to find a worse relay-receiver channel among the increased number of users and we choose the worst-user MSE as the objective function. On the other hand, the average MSEs for different number of receivers are almost similar.

In the next example, we compare the performance of the simplified min-max MSE algorithm with the NAF algorithm in terms of BER. QPSK signal constellations are used to modulate the transmitted signals. We set $K = 3$, $L = 3$, $M = 2, N_s = 2, N_1 = 10, N_2 = 6$, $N_d = 10$, and multicast $1000N_s$ randomly generated bits from each transmitter in each
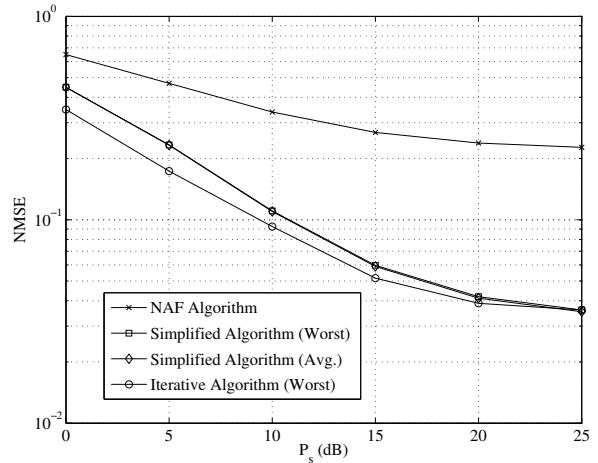

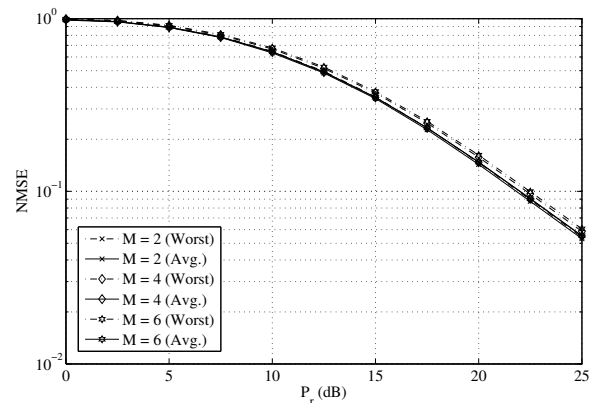Fig. 2. Example 1: Normalized MSE versus P_s. $K = 2$, $L = 3$, $M = 2$, $N_s = 2, N_1 = 8, N_2 = 4$, $N_d = 8$, $P_r = 20$dB.

Fig. 3. Example 2: Normalized MSE versus P_r. Varying number of receivers, $K = 2$, $L = 3, N_s = 2, N_1 = 6, N_2 = 4, N_d = 6$, $P_s = 20$dB.
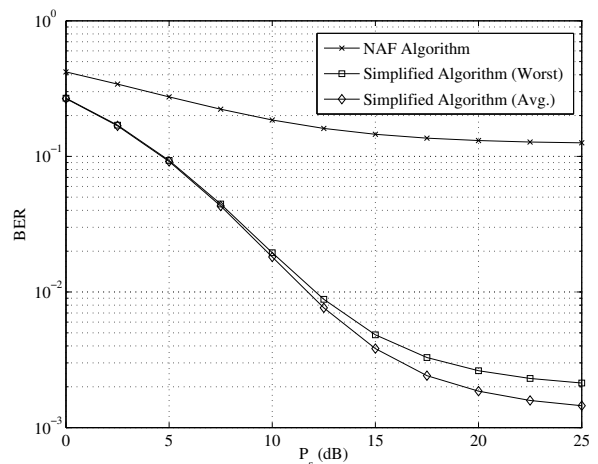

Fig. 4. Example 3: BER versus P_s. $K = 3$, $L = 3$, $M = 2$, $N_s = 2, N_1 = 10, N_2 = 6$, $N_d = 10$, $P_r = 20$dB.

channel realization. Fig. 4 shows the BER performance of both algorithms versus P_s with $P_r = 20$dB. It can be seen from

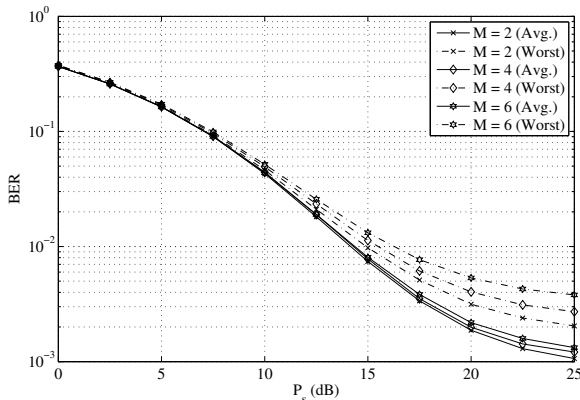Fig. 5. Example 4: BER versus P_s. Varying number of receivers, $K = 2$, $L = 3, N_\mathrm{s} = 2, N_1 = 6, N_2 = 4, N_\mathrm{d} = 6, P_\mathrm{r} = 20$dB.

Fig. 4 that the proposed simplified joint transmitter and relay optimization algorithm obtains a lower BER compared with the NAF approach. Even the worst-user BER of the simplified algorithm is always much lower than that of the NAF scheme.

In the last example, we compare the BER performance of the proposed simplified algorithm for different number of receivers. This time we set $K = 2$, $L = 3, N_\mathrm{s} = 2, N_1 = 6, N_2 = 4$, and $N_\mathrm{d} = 6$. Fig. 5 shows the BER of the simplified algorithm versus P_s with $P_\mathrm{r} = 20$dB for different number of receivers. It can be clearly seen from Fig. 5 that as we increase the number of receivers, the worst-user BER keeps increasing which is analogous to the results obtained in [22]. Interestingly, the average BERs of the users are almost similar up to $P_\mathrm{s} = 15$ dB for different number of receivers.

V. CONCLUSIONS

We considered a multi-hop multicasting MIMO relay system with multi-antenna nodes and proposed transmit and relay precoding matrices based on the min-max MSE criterion. The worst-case MSE is minimized subjecting to power constraints at the transmitters and the relay nodes when multiple transmitters intend to multicast their messages to a group of receivers. Under some mild approximation, we show that the problem can be solved with a significantly lower computational complexity. Numerical simulations have been carried out to evaluate the performance of the proposed algorithm. Simulation results demonstrate that the proposed transmitter and relay design algorithm outperforms the existing techniques.

REFERENCES

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
[2] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
[3] M. A. Khojastepour, A. Salehi-Golsefidi, and S. Rangarajan, "Towards an optimal beamforming algorithm for physical layer multicasting," in *Proc. IEEE Inf. Theory Workshop*, Oct. 16-20, 2011, pp. 395–399.
[4] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, pp. 2239–2251, Jun. 2006.
[5] N. Jindal and Z.-Q. Luo, "Capacity limits of multiple antenna multicast," in *Proc. IEEE ISIT*, Seattle, USA, Jul. 9-14, 2006, pp. 1841–1845.
[6] S. Y. Park, D. J. Love, and D. H. Kim, "Capacity limits of multi-antenna multicasting under correlated fading channels," *IEEE Trans. Commun.*, vol. 58, pp. 2002–2013, Jul. 2010.
[7] S. Y. Park and D. J. Love, "Capacity limits of multiple antenna multicasting using antenna subset selection," *IEEE Trans. Signal Process.*, vol. 56, pp. 2524–2534, Jun. 2008.
[8] E. Chiu and V. Lau, "Precoding design for multi-antenna multicast broadcast services with limited feedback," *IEEE Systems Journal*, vol. 4, pp. 550–560, Dec. 2010.
[9] Q. Li and W.-K. Ma, "Multicast secrecy rate maximization for MISO channels with multiple multi-antenna eavesdroppers," in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 5-9, 2011.
[10] S. X. Wu and W.-K. Ma, "Multicast transmit beamforming using a randomize-in-time strategy," in *Proc. IEEE ICASSP*, Prague, Czech Republic, May 22-27, 2011, pp. 3376–3379.
[11] S. X. Wu, W.-K. Ma, and A. M.-C. So, "Physical-layer multicasting by stochastic transmit beamforming and Alamouti space-time coding," *IEEE Trans. Signal Process.*, vol. 61, pp. 4230–4245, Sep. 2013.
[12] E. Matskani, N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, "Efficient batch and adaptive approximation algorithms for joint multicast beamforming and admission control," *IEEE Trans. Signal Process.*, vol. 57, pp. 4882–4894, Dec. 2009.
[13] N. Bornhorst and M. Pesavento, "An iterative convex approximation approach for transmit beamforming in multi-group multicasting," in *Proc. IEEE 12th Int. Workshop Signal Process. Adv. Wireless Commun.*, San Francisco, USA, Jun. 26-29, 2011, pp. 426–430.
[14] N. Bornhorst, P. Davarmanesh, M. Pesavento, "An extended interior-point method for transmit beamforming in multi-group multicasting," in *Proc. 20th European Signal Process. Conference*, Aug. 27-31, 2012, pp. 6–10.
[15] D. Senaratne and C. Tellambura, "Beamforming for physical layer multicasting," in *Proc. IEEE WCNC*, Cancun, Mexico, Mar. 28-31, 2011, pp. 1776–1781.
[16] M. Kaliszan, E. Pollakis, and S. Stańczak, "Efficient beamforming algorithms for MIMO multicast with application-layer coding," in *Proc. IEEE ISIT*, St. Petersburg, Russia, Jul. 31-Aug. 5, 2011, pp. 928–932.
[17] M. A. Khojastepour, A. Khajehnejad, K. Sundaresan, and S. Rangarajan, "Adaptive beamforming algorithms for wireless link layer multicasting," in *Proc. IEEE PIMRC*, Toronto, Canada, Sep. 11-14, 2011, pp. 1994–1998.
[18] N. Bornhorst, M. Pesavento, and A. B. Gershman, "Distributed beamforming for multi-group multicasting relay networks," *IEEE Trans. Signal Process.*, vol. 60, pp. 221–232, Jan. 2012.
[19] H. Zhang, X. You, G. Wu, and H. Wang, "Cooperative multi-antenna multicasting for wireless networks," in *Proc. IEEE GLOBECOM*, Miami, FL, USA, Dec. 6-10, 2010.
[20] H. Xu, J. Jin, and B. Li, "YMMV: Multiple session multicast with MIMO," in *Proc. IEEE GLOBECOM*, Texas, USA, Dec. 5-9, 2011.
[21] M. R. A. Khandaker and Y. Rong, "Multicasting MIMO relay optimization based on min-max MSE criterion," in *Proc. IEEE Int. Conf. Commun. Systems (ICCS'2012)*, Singapore, Nov. 21-23, 2012, pp. 16-20.
[22] M. R. A. Khandaker and Y. Rong, "Precoding design for MIMO relay multicasting," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 3544–3555, Jul. 2013.
[23] M. R. A. Khandaker and Y. Rong, "Simplified MIMO relay design for multicasting from multiple sources," in *Proc. IEEE ICASSP*, Florence, Italy, May 4-9, 2014.
[24] D. P. Palomar and Y. Jiang, *MIMO Transceiver Design via Majorization Theory.* now Publishers, 2007.
[25] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Englewood Cliffs, NJ: Prentice Hall, 1993.
[26] Y. Rong, "Simplified algorithms for optimizing multiuser multi-hop MIMO relay systems," *IEEE Trans. Commun.*, vol. 59, pp. 2896–2904, Oct. 2011.
[27] C. Song, K.-J. Lee, and I. Lee, "MMSE based transceiver designs in closed-loop non-regenerative MIMO relaying systems," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2310–2319, Jul. 2010.
[28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U. K.: Cambridge University Press, 2004.
[29] M. Grant and S. Boyd, "Cvx: Matlab software for disciplined convex programming (web page and software)." *http://cvxr.com/cvx*, April, 2010.
[30] Y. Nesterov and A. Nemirovski, *Interior Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, 1994.

PLACE PHOTO HERE

Muhammad R. A. Khandaker (S'10-M'13) received the bachelor of science (Honours) degree in computer science and engineering from Jahangirnagar University, Dhaka, Bangladesh, in 2006, the master of science degree in telecommunications engineering from East West University, Dhaka, Bangladesh, in 2007, and the Ph.D. degree in electrical and computer engineering from Curtin University, Australia, in 2013.

Dr. Khandaker started his career as a Jr. Hardware Design Engineer for Visual Magic Corporation Limited (VMCL-BD) in November 2005. He joined as a Lecturer in the Department of Computer Science and Engineering, IBAIS University, Dhaka, Bangladesh, in January 2006. In November 2007, he joined as a Lecturer in the Department of Information and Communication Technology of Mawlana Bhasani Science and Technology University, Bangladesh. Then he joined as a Lecturer in the Institute of Information Technology, Jahangirnagar University, Dhaka, Bangladesh, in January 2008. Since July 2013, he has been working as a postdoctoral Research Associate in the Department of Electronic and Electrical Engineering, University College London (UCL), United Kingdom. He was awarded the Curtin International Postgraduate Research Scholarship (CIPRS) for his Ph.D. study in 2009. He also received the Best Paper Award at the 16th Asia-Pacific Conference on Communications, Auckland, New Zealand, 2010.

PLACE PHOTO HERE

Yue Rong (S'03-M'06-SM'11) received the Ph.D. degree (summa cum laude) in electrical engineering from the Darmstadt University of Technology, Darmstadt, Germany, in 2005.

He was a Post-Doctoral Researcher with the Department of Electrical Engineering, University of California, Riverside, from February 2006 to November 2007. Since December 2007, he has been with the Department of Electrical and Computer Engineering, Curtin University, Bentley, Australia, where he is currently an Associate Professor. His research interests include signal processing for communications, wireless communications, underwater acoustic communications, applications of linear algebra and optimization methods, and statistical and array signal processing.

Dr. Rong was a recipient of the Best Paper Award at the 2011 International Conference on Wireless Communications and Signal Processing, the Best Paper Award at the 2010 Asia-Pacific Conference on Communications, and the Young Researcher of the Year Award of the Faculty of Science and Engineering at Curtin University in 2010. He is an Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS, a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS special issue on theories and methods for advanced wireless relays, and was a TPC Member for the IEEE ICC, WCSP, IWMC, and ChinaCom.