Joint source and relay optimization for interference MIMO relay networks

Muhammad R. A. Khandaker[*] iD and Kai-Kit Wong

Abstract

This paper considers multiple-input multiple-output (MIMO) relay communication in multi-cellular (interference) systems in which MIMO source-destination pairs communicate simultaneously. It is assumed that due to severe attenuation and/or shadowing effects, communication links can be established only with the aid of a relay node. The aim is to minimize the maximal mean-square-error (MSE) among all the receiving nodes under constrained source and relay transmit powers. Both one- and two-way amplify-and-forward (AF) relaying mechanisms are considered. Since the exactly optimal solution for this practically appealing problem is intractable, we first propose optimizing the source, relay, and receiver matrices in an alternating fashion. Then we contrive a simplified *semidefinite programming (SDP) solution* based on the error covariance matrix decomposition technique, avoiding the high complexity of the iterative process. Numerical results reveal the effectiveness of the proposed schemes.

Keywords: Interference, MIMO, Two-way, Relay, Optimization

1 Introduction

Due to scarcity of frequency spectrum in practical wireless networks, multiple communicating pairs are motivated to share a common time-frequency channel to ensure efficient use of the available spectrum. Co-channel interference (CCI) is, however, one of the main deteriorating factors in such networks that adversely affect the system performance. The impact is more obvious in 5G heterogeneous networks where there is oceanic volume of interference due to hyper-dense frequency reuse among small-cell and macro cell base stations. Therefore, it is important to develop schemes to mitigate the CCI, which has been a major research direction in wireless communications over the past decades.

In the literature, various schemes have been proposed to control CCI at an acceptable level. A conventional approach in MIMO systems is to exploit spatial diversity for suppressing CCI [1]. Such spatial diversity technique has been used to solve many power control problems in interference systems for different network setups. In [2], a power control scheme has been designed

with receive diversity only, whereas joint transmit-receive beamforming has been considered in [2, 3] for interference systems. However, the incorporation of the spatial diversity at the transmitter side in [3], results in lower total transmit power compared to that in [2].

On the other hand, there is synergy between multiple antenna and relaying technologies. The latter is particularly useful to reestablish communications in case of a broken channel between source and destination. Hence, relaying has been considered in interference networks in order to afford longer source-destination distance [4–7]. Both [4, 5] considered network beamforming for minimizing total relay transmit power, whereas in [7], an iterative transceiver optimization scheme has been proposed to minimize total source and relay transmit power.

While the works in [2–5, 7] all considered minimizing the total transmit power of interference networks, another important performance metric, which concerns more about the quality of communications, is the mean-square-error (MSE) for signal estimation. In [8–10], the sum minimum MSE (MMSE) was considered to design iterative algorithms for MIMO interference relay systems taking the direct links between the source and destination nodes into consideration, and in [11], similar

*Correspondence: m.khandaker@ucl.ac.uk
Department of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, UK

problem has been considered ignoring the direct links between the communicating parties. Indeed the direct source-destination links can play a vital role in wireless communication systems when the link-strength is significant. However, multihop communication is motivated by the fact that such direct links may undergo deep shadowing effects in many practical scenarios. Hence, many existing works on multihop communications have ignored the direct links. Nonetheless, the sum MMSE criterion runs the risk that some of the receivers may suffer from unacceptably high MSEs. Also, the works in [8–10] considered one-way relaying only.

Due to the increasing demands on multimedia applications, in particular, the notion of emerging wireless communications terminologies such as Big data, ultra-high spectral efficiency is essential in future wireless networks, including 5G, to provide ADSL-like user experience aspired by 2020. The abovementioned one-way relay systems suffer from a substantial performance loss in terms of spectral efficiency due to the pre-log factor of 1/2 persuaded by the fact that two channel uses are required for each end-to-end transmission.

Two-way relay systems have hence been proposed to overcome the loss of spectral efficiency in such one-way relay methods [12–14]. Utilizing the concept of analog network coding [14], communication in a two-way relay channel can be accomplished in two phases: the multiple access (MAC) phase and the broadcast (BC) phase. During the MAC phase, all the users simultaneously send their messages to an intermediate relay node, whereas in the BC phase, the relay retransmits the received information to the users. As each user knows its own transmitted signals, each user can cancel the self-interference and decode the intended message. The capacity region of multi-pair two-way relay networks in the deterministic channel was characterized in [15]. Later in [16], the achievable total degrees of freedom in a two-way interference MIMO relay channel were also studied. Most recently in [17], the transceivers in a full-duplex MIMO interference system were optimized based on the weighted sum-rate maximization criterion.

In this paper, we consider a *K*-user MIMO interference system where each of the pairs can communicate only with the aid of a relay node thus ignoring the direct source-destination links. The direct links are understood to be in deep shadowing and hence negligible. Both one- and two-way amplify-and-forward (AF) relaying mechanisms are considered. All nodes are assumed to be equipped with multiple antennas so as to afford simultaneous transmission of multiple data streams. Our aim is to develop joint transceiver optimization algorithms for minimizing the worst-user MSE (min-max MSE)[1] subject to the source and relay power constraints. It can be verified that the problem is strictly non-convex,

and thus it is difficult to find an analytical solution. To tackle this, we first devise an algorithm to optimize the source, relay, and receiver matrices alternatingly by decomposing the original non-convex problem into convex subproblems. To avoid the complexity of the iterative process, we then extend the error covariance matrix decomposition technique applied to point-to-point MIMO relay systems in [18] to interference MIMO relay systems in this paper. More specifically, under practically reasonable high first-hop signal-to-noise ratio (SNR) assumption, we demonstrate that the problem can be decomposed into two standard semidefinite programming (SDP) problems to optimize source and relay matrices separately. Note that high SNR assumption has also been made in [19] to simplify the joint codebook design problem in single-user MIMO relay systems and in [20, 21] for multicasting MIMO relay design. Hence our work is a generalization to multi-pair communication scheme taking co-channel interference into account.

The remainder of this paper is lined-up as follows. In Section 2, the interference MIMO relay system model is introduced. The joint optimal transmitter, relay, and receiver beamforming optimization schemes are developed in Section 3 and Section 4, respectively, for one-way and two-way relaying. Section 5 provides simulation results to analyze the performance of the proposed algorithms in various system configurations before concluding remarks are made in Section 6.

2 System model

Let us consider a communication scenario, as illustrated in Fig. 1, where each of the *K* source nodes communicates with the corresponding destination node sharing the same frequency channel via a common relay node. The direct link between each transmitter-receiver pair is assumed to be broken due to strong attenuation and/or shadowing effects. The *k*th source, the relay, and the *k*th destination nodes are assumed to be equipped with $N_{s,k}$, N_r, and $N_{d,k}$ antennas, respectively.
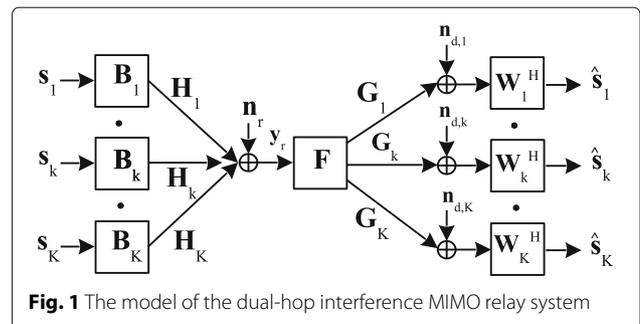
Fig. 1 The model of the dual-hop interference MIMO relay system

3 One-way relaying

In this section, we consider that communication takes place in one direction only. The relay node is assumed to work in half-duplex mode which implies that the actual communication between the source and destination nodes is accomplished in two time slots. In the first time slot, the source nodes transmit the linearly precoded signal vectors $\mathbf{B}_k\mathbf{s}_k, k = 1, \cdots, K$, to the relay node. The received signal vector at the relay node is therefore given by

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

where $\mathbf{H}_k$ denotes the $N_{\mathrm{r}} \times N_{\mathrm{s},k}$ Gaussian channel matrix between the kth source node and the intermediate relay node, $\mathbf{s}_k$ is the $N_{\mathrm{b},k} \times 1$ ($1 \leq N_{\mathrm{b},k} \leq N_{\mathrm{s},k}$) transmit symbol vector with covariance $\mathbf{I}_{N_{\mathrm{b}},k}$, $\mathbf{B}_k$ is the $N_{\mathrm{s},k} \times N_{\mathrm{b},k}$ source precoding matrix, and $\mathbf{n}_{\mathrm{r}}$ is the $N_{\mathrm{r}} \times 1$ additive white Gaussian noise (AWGN) vector introduced at the relay node. Let us denote $N_{\mathrm{b}} = \sum_{k=1}^{K} N_{\mathrm{b},k}$ as the total number of data streams transmitted by all the source nodes. In order to successfully transmit N_{b} independent data streams simultaneously through the relay, the relay node must be equipped with $N_{\mathrm{r}} \geq N_{\mathrm{b}}$ antennas.

After receiving $\mathbf{y}_{\mathrm{r}}$, the relay node simply multiplies the signal vector by an $N_{\mathrm{r}} \times N_{\mathrm{r}}$ precoding matrix $\mathbf{F}$ and transmits the amplified version of $\mathbf{y}_{\mathrm{r}}$ in the second time slot. Thus the relay's $N_{\mathrm{r}} \times 1$ transmit signal vector $\mathbf{x}_{\mathrm{r}}$ is given by

$$\mathbf{x}_{\mathrm{r}} = \mathbf{F}\mathbf{y}_{\mathrm{r}}. \qquad (2)$$

Accordingly, the signal received at the kth destination node can be expressed as

$$
\begin{aligned}
\mathbf{y}_{\mathrm{d},k} &= \mathbf{G}_k \mathbf{x}_{\mathrm{r}} + \mathbf{n}_{\mathrm{d},k} \\
&= \underbrace{\mathbf{G}_k \mathbf{F} \mathbf{H}_k \mathbf{B}_k \mathbf{s}_k}_{\text{desired signal}} + \underbrace{\mathbf{G}_k \mathbf{F} \sum_{\substack{j=1 \\ j \neq k}}^{K} \mathbf{H}_j \mathbf{B}_j \mathbf{s}_j}_{\text{interference signal}} + \underbrace{\mathbf{G}_k \mathbf{F} \mathbf{n}_{\mathrm{r}} + \mathbf{n}_{\mathrm{d},k}}_{\text{noise}}, \quad (3)
\end{aligned}
$$

$$= \bar{\mathbf{H}}_k \mathbf{s}_k + \bar{\mathbf{n}}_{\mathrm{d},k}, \text{ for } k = 1, \ldots, K, \qquad (4)$$

where $\mathbf{G}_k$ denotes the $N_{\mathrm{d},k} \times N_{\mathrm{r}}$ complex channel matrix between the relay node and the kth destination node, $\mathbf{n}_{\mathrm{d},k}$ is the $N_{\mathrm{d},k} \times 1$ AWGN vector introduced at the kth destination node, $\bar{\mathbf{H}}_k \triangleq \mathbf{G}_k \mathbf{F} \mathbf{H}_k \mathbf{B}_k$ is the equivalent source-destination channel matrix, and $\bar{\mathbf{n}}_{\mathrm{d},k} \triangleq \mathbf{G}_k \mathbf{F}(\sum_{\substack{j=1 \\ j \neq k}}^{K} \mathbf{H}_j \mathbf{B}_j \mathbf{s}_j + \mathbf{n}_{\mathrm{r}}) + \mathbf{n}_{\mathrm{d},k}$ is the equivalent noise vector. All noises are assumed to be independent and identically distributed (i.i.d.) complex Gaussian random variables

with mean zero and variance σ_{n}^2, where $\mathrm{n} \in \{\mathrm{r}, \mathrm{d}\}$ indicates the noise introduced at the relay or at the destination.

Remark *Note that the interference term in (3) does not appear in the received signal of the single-user MIMO relay system considered in [19] or in the multicasting MIMO relay system considered in [20, 21]. Hence the subsequent analyses remain considerably simpler in [19–21], whereas we need to deal with this troublesome interference term in this paper.*

Considering the input-output relationship at the relay node given in (2), the average transmit power consumed by the MIMO relay node is defined as

$$\mathrm{tr}\big(\mathrm{E}\{\mathbf{x}_{\mathrm{r}}\mathbf{x}_{\mathrm{r}}^H\}\big) = \mathrm{tr}\big(\mathbf{F}\boldsymbol{\Psi}\mathbf{F}^H\big), \qquad (5)$$

where $\mathrm{tr}(\cdot)$ denotes trace of a matrix, $\mathrm{E}\{\cdot\}$ indicates statistical expectation, and $\boldsymbol{\Psi} \triangleq \mathrm{E}\{\mathbf{y}_{\mathrm{r}}\mathbf{y}_{\mathrm{r}}^H\} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{B}_k \mathbf{B}_k^H \mathbf{H}_k^H + \sigma_{\mathrm{r}}^2 \mathbf{I}_{N_{\mathrm{r}}}$ represents the covariance matrix of the signal vector received at the relay node.

For signal detection, linear receivers are used at the destination nodes for simplicity reasons. Denoting $\mathbf{W}_k$ as the $N_{\mathrm{d},k} \times N_{\mathrm{b},k}$ receiver matrix used by the kth destination node, the corresponding estimated signal vector $\hat{\mathbf{s}}_k$ can be written as

$$\hat{\mathbf{s}}_k = \mathbf{W}_k^H \mathbf{y}_{\mathrm{d},k}, \text{ for } k = 1, \ldots, K, \qquad (6)$$

where $(\cdot)^H$ indicates the conjugate transpose (Hermitian) of a matrix (vector). Thus the MSE of signal estimation at the kth receiver can be expressed as

$$
\begin{aligned}
E_k &= \mathrm{tr}\left(\mathbf{E}_k \triangleq \mathrm{E}\left[\big(\hat{\mathbf{s}}_k - \mathbf{s}_k\big)\big(\hat{\mathbf{s}}_k - \mathbf{s}_k\big)^H\right]\right), \\
&= \mathrm{tr}\begin{pmatrix} \mathbf{I}_{N_{\mathrm{b}},k} - \mathbf{W}_k^H \mathbf{G}_k \mathbf{F} \mathbf{H}_k \mathbf{B}_k - \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{F}^H \mathbf{G}_k^H \mathbf{W}_k \\ + \sum_{j=1}^{K} \mathbf{W}_k^H \mathbf{G}_k \mathbf{F} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H \mathbf{F}^H \mathbf{G}_k^H \mathbf{W}_k \\ + \sigma_{\mathrm{r}}^2 \mathbf{W}_k^H \mathbf{G}_k \mathbf{F} \mathbf{F}^H \mathbf{G}_k^H \mathbf{W}_k + \sigma_{\mathrm{d}}^2 \mathbf{W}_k^H \mathbf{W}_k \end{pmatrix}, \\
&= \mathrm{tr}\left(\big(\mathbf{W}_k^H \bar{\mathbf{H}}_k - \mathbf{I}_{N_{\mathrm{b}},k}\big)\big(\mathbf{W}_k^H \bar{\mathbf{H}}_k - \mathbf{I}_{N_{\mathrm{b}},k}\big)^H + \mathbf{W}_k^H \bar{\mathbf{C}}_k \mathbf{W}_k\right), \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{for } k = 1, \ldots, K, \quad (7)
\end{aligned}
$$

where $\mathbf{E}_k$ denotes the error covariance matrix at the kth receiver, and

$$\bar{\mathbf{C}}_k \triangleq \sum_{\substack{j=1 \\ j \neq k}}^{K} \mathbf{G}_k \mathbf{F} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H \mathbf{F}^H \mathbf{G}_k^H + \sigma_{\mathrm{r}}^2 \mathbf{G}_k \mathbf{F} \mathbf{F}^H \mathbf{G}_k^H + \sigma_{\mathrm{d}}^2 \mathbf{I}_{N_{\mathrm{d}}}. \qquad (8)$$

is the combined interference and noise covariance matrix.

In the following subsections, we develop optimization approaches that minimize the worst-user MSE among all the receivers subject to source and relay power constraints.

3.1 Problem formulation

In this section, we formulate the joint source and relay precoding optimization problem for MIMO interference systems. Our aim is to minimize the maximal MSE among all the source-destination pairs yet satisfying the transmit power constraints at the source as well as the relay nodes. To fulfill this aim, the following joint optimization problem is formulated:

$$\min_{\{\mathbf{B}_k\},\mathbf{F},\{\mathbf{W}_k\}} \quad \max_k \quad E_k \tag{9a}$$

$$\text{s.t.} \quad \text{tr}\left(\mathbf{F}\mathbf{\Psi}\mathbf{F}^H\right) \leq P_\text{r} \tag{9b}$$

$$\text{tr}\left(\mathbf{B}_k\mathbf{B}_k^H\right) \leq P_{\text{s},k}, \quad \text{for } k = 1,\ldots,K \tag{9c}$$

where (9b) and (9c), respectively, constrains the transmit power at the relay node and the kth transmitter to $P_\text{r} > 0$, $P_{\text{s},k} > 0$. Our next endeavor is to develop optimal solutions for this problem. Note that the problem is strictly non-convex with matrix variables appearing in quadratic form, and hence any closed-form solution is intractable. Therefore, we first resort to developing an iterative algorithm for the problem and then propose a sub-optimal solution which has lower computational complexity.

3.2 Iterative joint transceiver optimization

In this subsection, we investigate the non-convex source, relay, and destination filter design problem in an alternating fashion. We tend to optimize one group of variables while fixing the others. Given source and relay matrices $\{\mathbf{B}_k\}$, $\mathbf{F}$, the optimal receiver matrices $\{\mathbf{W}_k\}$ are obtained through solving the unconstrained optimization problem of $\min_{\mathbf{W}_k} E_k$, since E_k does not depend on $\mathbf{W}_j$, for $j \neq k$, and $\mathbf{W}_k$ does not appear in constraints (9b) and (9c). Using

the matrix derivative formulas, the gradient $\nabla_{\mathbf{W}_k^H}\left(\text{tr}\left(\mathbf{E}_k\right)\right)$ can be written as

$$\nabla_{\mathbf{W}_k^H}\left(\text{tr}\left(\mathbf{E}_k\right)\right) = -\mathbf{G}_k\mathbf{F}\mathbf{H}_k\mathbf{B}_k + \sum_{j=1}^K \mathbf{G}_k\mathbf{F}\mathbf{H}_j\mathbf{B}_j\mathbf{B}_j^H\mathbf{H}_j^H\mathbf{F}^H\mathbf{G}_k^H\mathbf{W}_k$$
$$+ \sigma_\text{r}^2\mathbf{G}_k\mathbf{F}\mathbf{F}^H\mathbf{G}_k^H\mathbf{W}_k + \sigma_\text{d}^2\mathbf{W}_k, \quad \text{for } k = 1,\ldots,K. \tag{10}$$

Equating $\nabla_{\mathbf{W}_k^H}\left(\text{tr}\left(\mathbf{E}_k\right)\right) = \mathbf{0}$ yields the linear MMSE receive filter given by

$$\mathbf{W}_k = \left(\sum_{j=1}^K \mathbf{G}_k\mathbf{F}\mathbf{H}_j\mathbf{B}_j\mathbf{B}_j^H\mathbf{H}_j^H\mathbf{F}^H\mathbf{G}_k^H + \sigma_\text{r}^2\mathbf{G}_k\mathbf{F}\mathbf{F}^H\mathbf{G}_k^H + \sigma_\text{d}^2\mathbf{I}_{N_{\text{d},k}}\right)^{-1}$$
$$\times \mathbf{G}_k\mathbf{F}\mathbf{H}_k\mathbf{B}_k \tag{11}$$

where $(\cdot)^{-1}$ indicates the inversion operation of a matrix.

Then for given source and receiver matrices $\{\mathbf{B}_k\}$ and $\{\mathbf{W}_k\}$, the relay precoding matrix $\mathbf{F}$ optimization problem can be formulated as

$$\min_{\mathbf{F}} \quad \max_k \quad E_k \tag{12a}$$

$$\text{s.t.} \quad \text{tr}\left(\mathbf{F}\mathbf{\Psi}\mathbf{F}^H\right) \leq P_\text{r}. \tag{12b}$$

Note that (12) is non-convex with a matrix variable since $\mathbf{F}$ appears in quadratic form in the objective function as well as in the constraint. However, we can reformulate this problem as an SDP using Schur complement [22] as follows. By introducing a matrix $\mathbf{\Xi}_k$ we conclude from the second equation in (7) that the k-th link MSE will be upper-bounded if

$$- \mathbf{W}_k^H\mathbf{G}_k\mathbf{F}\mathbf{H}_k\mathbf{B}_k - \mathbf{B}_k^H\mathbf{H}_k^H\mathbf{F}^H\mathbf{G}_k^H\mathbf{W}_k$$
$$+ \mathbf{W}_k^H\mathbf{G}_k\mathbf{F}\mathbf{\Psi}\mathbf{F}^H\mathbf{G}_k^H\mathbf{W}_k \preceq \mathbf{\Xi}_k. \tag{13}$$

In the above inequality, $\mathbf{A} \preceq \mathbf{B}$ indicates that the matrix $\mathbf{B} - \mathbf{A}$ is positive semidefinite (PSD). Now, by introducing a matrix $\mathbf{\Phi}$ such that $\mathbf{F}\mathbf{\Psi}\mathbf{F}^H \preceq \mathbf{\Phi}$, and a scaler variable τ_r, the relay optimization problem (12) can be transformed to

$$\min_{\tau_\text{r},\mathbf{F},\{\mathbf{\Xi}_k\},\mathbf{\Phi}} \quad \tau_\text{r} \tag{14a}$$

$$\text{s.t.} \quad \text{tr}\left(\mathbf{\Xi}_k\right) + \text{tr}\left(\mathbf{W}_k^H\mathbf{W}_k\right) + N_{\text{b},k} \leq \tau_\text{r}, \quad \text{for } k = 1,\ldots,K, \tag{14b}$$

$$\begin{bmatrix} \mathbf{\Xi}_k + \mathbf{W}_k^H\mathbf{G}_k\mathbf{F}\mathbf{H}_k\mathbf{B}_k + \mathbf{B}_k^H\mathbf{H}_k^H\mathbf{F}^H\mathbf{G}_k^H\mathbf{W}_k & \mathbf{W}_k^H\mathbf{G}_k\mathbf{F} \\ \mathbf{F}^H\mathbf{G}_k^H\mathbf{W}_k & \mathbf{\Psi}^{-1} \end{bmatrix} \succeq \mathbf{0}, \text{for } k = 1,\ldots,K, \tag{14c}$$

$$\begin{bmatrix} \mathbf{\Phi} & \mathbf{F} \\ \mathbf{F}^H & \mathbf{\Psi}^{-1} \end{bmatrix} \succeq \mathbf{0} \tag{14d}$$

$$\text{tr}\left(\mathbf{\Phi}\right) \leq P_\text{r} \tag{14e}$$

where we have used the Schur complement to obtain (14c) and (14d). Note that the problem (14) is an SDP problem which is convex and can, as a result, be efficiently solved using interior-point based solvers [23] at a maximal complexity order of $\mathcal{O}\big((K+2N_r^2+\sum_{k=1}^{K}N_{b,k}^2+2)^{3.5}\big)$ [24]. However, the actual complexity is usually much less in many practical cases. Interested readers are referred to [24] for a detailed analysis of the computational complexity based on interior-point methods.

Finally, we optimize the source matrices $\{\mathbf{B}_k\}$ using the relay matrix $\mathbf{F}$ and the receiver matrices $\{\mathbf{W}_k\}$ known from the previous steps. Let us define $\tilde{\mathbf{H}}_{k,j} \triangleq \mathbf{W}_k^H \mathbf{G}_k \mathbf{F} \mathbf{H}_j$. Applying the matrix identity $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \mathrm{vec}(\mathbf{B})$, we can rewrite E_k in (7) as

$$E_k = \sum_{j=1}^{K} \mathbf{b}_j^H \left(\mathbf{I}_{N_{b,j}} \otimes \left(\tilde{\mathbf{H}}_{k,j}^H \tilde{\mathbf{H}}_{k,j}\right)\right) \mathbf{b}_j - \left(\mathrm{vec}\left(\tilde{\mathbf{H}}_{k,j}\right)\right)^T \mathbf{b}_k$$
$$- \mathbf{b}_k^H \mathrm{vec}\left(\tilde{\mathbf{H}}_{k,j}^H\right) + \theta_k, \tag{15}$$

where the vector $\mathbf{b}_k \triangleq \mathrm{vec}(\mathbf{B}_k)$ is created by stacking all the columns of the matrix $\mathbf{B}_k$ on top of each other, $\theta_k \triangleq \mathrm{tr}(\sigma_r^2 \mathbf{W}_k^H \mathbf{G}_k \mathbf{F} \mathbf{F}^H \mathbf{G}_k^H \mathbf{W}_k + \sigma_d^2 \mathbf{W}_k^H \mathbf{W}_k) + N_{b,k}$, and $\otimes$ indicates matrix Kronecker product. Let us now denote

$$\begin{cases} \tilde{\mathbf{G}}_k \triangleq \mathrm{bd}\left(\mathbf{I}_{N_{b,1}} \otimes \left(\tilde{\mathbf{H}}_{k,1}^H \tilde{\mathbf{H}}_{k,1}\right), \ldots, \mathbf{I}_{N_{b,K}} \otimes \left(\tilde{\mathbf{H}}_{k,K}^H \tilde{\mathbf{H}}_{k,K}\right)\right), \\ \mathbf{c}_k \triangleq \left[\left(\mathrm{vec}\left(\tilde{\mathbf{C}}_{k,1}\right)\right)^T, \ldots, \left(\mathrm{vec}\left(\tilde{\mathbf{C}}_{k,K}\right)\right)^T\right]^T, \\ \mathbf{b} \triangleq \left[\mathbf{b}_1^T, \ldots, \mathbf{b}_K^T\right]^T, \end{cases} \tag{16}$$

where $\mathrm{bd}(\cdot)$ constructs a block-diagonal matrix taking the parameter matrices as the diagonal blocks, $\tilde{\mathbf{C}}_{k,k} = \tilde{\mathbf{H}}_{k,k}$ and $\tilde{\mathbf{C}}_{k,j} = \mathbf{0}_{N_{b,k} \times N_{s,j}}$, if $j \neq k$. The MSE in (15) can be rewritten as

$$E_k = \mathbf{b}^H \tilde{\mathbf{G}}_k \mathbf{b} - \mathbf{c}_k^H \mathbf{b} - \mathbf{b}^H \mathbf{c}_k + \theta_k. \tag{17}$$

By introducing $\mathbf{M}_k \triangleq \mathbf{F} \mathbf{H}_k$, the power constraints in (9b) can be rewritten as

$$\mathbf{b}^H \mathbf{M} \mathbf{b} \leq \bar{P}_r, \ \text{for } k = 1, \ldots, K, \tag{18}$$

where $\mathbf{M} \triangleq \mathrm{bd}\big(\mathbf{I}_{N_{b,1}} \otimes (\mathbf{M}_1^H \mathbf{M}_1), \ldots, \mathbf{I}_{N_{b,K}} \otimes (\mathbf{M}_K^H \mathbf{M}_K)\big)$, and $\bar{P}_r = P_r - \sigma_r^2 \mathrm{tr}(\mathbf{F}\mathbf{F}^H)$. Using (17) and (18), problem (9) can be written as

$$\min_{\mathbf{b}} \ \max_{k} \ \mathbf{b}^H \tilde{\mathbf{G}}_k \mathbf{b} - \mathbf{c}_k^H \mathbf{b} - \mathbf{b}^H \mathbf{c}_k + \theta_k \tag{19a}$$
$$\text{s.t.} \ \mathbf{b}^H \mathbf{M} \mathbf{b} \leq \bar{P}_r \tag{19b}$$
$$\mathbf{b}^H \mathcal{I}_k \mathbf{b} \leq P_{s,k}, \ \text{for } k = 1, \ldots, K, \tag{19c}$$

where $\mathcal{I} \triangleq \mathrm{bd}(\mathcal{I}_{k1}, \ldots, \mathcal{I}_{kk}, \ldots, \mathcal{I}_{kK})$ with $\mathcal{I}_{kk} = \mathbf{I}_{N_{s,k}N_{b,k}}$ and $\mathcal{I}_{kj} = \mathbf{0}$, if $j \neq k$. Problem (19) is a standard quadratically-constrained quadratic program

(QCQP) which can be solved using off-the-shelf convex optimization toolboxes [23]. In the following, we also provide an SDP formulation of problem (19):

$$\min_{t_s, \mathbf{b}} \ \tau_s \tag{20a}$$
$$\text{s.t.} \ \begin{pmatrix} \tau_s - \theta_k + \mathbf{c}_k^H \mathbf{b} + \mathbf{b}^H \mathbf{c}_k & \mathbf{b}^H \\ \mathbf{b} & \tilde{\mathbf{G}}_k^{-1} \end{pmatrix} \succeq \mathbf{0},$$
$$\text{for } k = 1, \ldots, K, \tag{20b}$$
$$\begin{pmatrix} \bar{P}_r & \mathbf{b}^H \\ \mathbf{b} & \mathbf{M}^{-1} \end{pmatrix} \succeq \mathbf{0}, \tag{20c}$$
$$\begin{pmatrix} P_{s,k} & \mathbf{b}^H \mathcal{I}_k^{\frac{1}{2}} \\ \mathcal{I}_k^{\frac{1}{2}} \mathbf{b} & \mathbf{I}_p \end{pmatrix} \succeq \mathbf{0}, \ \text{for } k = 1, \ldots, K, \tag{20d}$$

where τ_s is a slack variable and $p \triangleq \sum_{k=1}^{K} N_{s,k} N_{b,k}$. The problem (20) can be solved at a maximal complexity order of $\mathcal{O}\big((\sum_{k=1}^{K} N_{b,k}^2 + 1)^{3.5}\big)$ [24]. The proposed iterative optimization technique for solving the original problem (9) is summarized in Table 1.

Since in each step of the iterative algorithm we solve a convex subproblem to update one set of variables, the conditional update of each set will either decrease or maintain the objective function (9a). From this observation, a monotonic convergence of the iterative algorithm follows. However, the overall computational complexity of the iterative algorithm increases as the multiple of the number of iterations required until convergence. Thus the complexity of the iterative algorithms is often reasonably high. Note that the sum-MSE based iterative algorithms proposed in [8–10] have similar complexity orders. Hence in the following subsection, we contrive an algorithm for the joint optimization problem such that the computational overhead is substantially reduced.

3.3 Simplified joint optimization algorithm
In the previous subsection, we optimized the source, relay, and receiver matrices in an alternating fashion. Here, we propose a simplified approach to solve problem (9) using the error covariance matrix decomposition technique. The following theorem paves the foundation of the simplified algorithm.

Table 1 Iterative solution of problem (9)

1	Randomly initialize $\mathbf{F}$ and $\{\mathbf{B}_k\}$ such that the constraints (9b) and (9c) are satisfied.
2	Repeat
	(a) Obtain $\{\mathbf{W}_k\}$ as defined in (11) using known $\{\mathbf{B}_k\}$ and $\mathbf{F}$.
	(b) Solve the subproblem (14) to update $\mathbf{F}$ using fixed $\{\mathbf{W}_k\}$ and $\{\mathbf{B}_k\}$.
	(c) Update $\{\mathbf{B}_k\}$ through solving the subproblem (20) using $\mathbf{F}$ and $\{\mathbf{W}_k\}$ known from the previous steps.
3	Until convergence.

Theorem 1 *For given* $\{\mathbf{B}_k\}$ *and* $\{\mathbf{W}_k\}$, *the optimum relaying matrix* $\mathbf{F}$ *for minimizing the worst-user MSE has the form:*

$$\mathbf{F} = \sum_{k=1}^{K} \mathbf{T}_k \mathbf{D}_k^H = \mathbf{T}\mathbf{D}^H, \qquad (21)$$

where $\mathbf{T} \triangleq [\mathbf{T}_1, \ldots, \mathbf{T}_K]$ *and* $\mathbf{D} \triangleq [\mathbf{D}_1, \ldots, \mathbf{D}_K]$ *with* $\mathbf{T}_k$ *and* $\mathbf{D}_k$, *respectively, defined as*

$$\mathbf{T}_k \triangleq \lambda_{e,k} \left(\sum_{i=1}^{K} \lambda_{e,i} \mathbf{G}_i^H \mathbf{W}_i \mathbf{W}_i^H \mathbf{G}_i + \lambda_r \mathbf{I}_{N_r} \right)^{-1} \mathbf{G}_k^H \mathbf{W}_k \quad (22)$$

and

$$\mathbf{D}_k \triangleq \left(\sum_{j=1}^{K} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H + \sigma_r^2 \mathbf{I}_{N_r} \right)^{-1} \mathbf{H}_k \mathbf{B}_k, \qquad (23)$$

λ_r *and* $\lambda_{e,k}$, $\forall k$, *are the corresponding Lagrange multipliers as defined in Appendix* 1.

Proof See Appendix 1. □

Note that $\mathbf{D}_k = \left(\mathbf{H}_k \mathbf{B}_k \mathbf{B}_k^H \mathbf{H}_k^H + \sum_{\substack{j=1 \\ j \neq k}}^{K} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H + \sigma_r^2 \mathbf{I}_{N_r} \right)^{-1} \mathbf{H}_k \mathbf{B}_k$ can be regarded as the MMSE receive filter of the first-hop MIMO channel for the kth transmitter's signal received at the relay node given by (1).

The implication of the structure of the relay amplifying matrix in the proposed simplified design can be observed while applying the following theorem.

Theorem 2 *The MSE term appearing in* (9a) *can be equivalently decomposed into*

$$E_k = tr \left(\mathbf{I}_{N_b,k} + \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}_{\bar{k}}^{-1} \mathbf{H}_k \mathbf{B}_k \right)^{-1} \qquad (24)$$

$$+ tr \left(\left(\mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}^{-1} \mathbf{H}_k \mathbf{B}_k \right)^{-1} + \tilde{\mathbf{T}}^H \mathbf{G}_k^H \mathbf{G}_k \tilde{\mathbf{T}} \right)^{-1},$$

where $\mathbf{\Psi}_{\bar{k}} \triangleq \mathbf{\Psi} - \mathbf{H}_k \mathbf{B}_k \mathbf{B}_k^H \mathbf{H}_k^H = \sum_{\substack{j=1 \\ j \neq k}}^{K} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H + \sigma_n^2 \mathbf{I}_{N_r}$ *and* $\tilde{\mathbf{T}}$ *is defined in Appendix* 2.

Proof See Appendix 2. □

Even given the structure, an analytical optimal solution to the joint optimization problem is still difficult to obtain due to the cross-link interference from the relay node to the destination nodes. Therefore, we resort to develop an efficient suboptimal solution. The following proposition provides the foundation of the proposed simplified suboptimal solution.

Proposition 1 *In the practically reasonably high SNR regime, the term* $\mathbf{B}_k^H \mathbf{H}_k^H \times \mathbf{\Psi}^{-1} \mathbf{H}_k \mathbf{B}_k$ *in* (24) *can be approximated as* $\mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}^{-1} \mathbf{H}_k \mathbf{B}_k \approx \mathbf{I}_{N_b,k}$.

Proof See Appendix 3. □

The result in *Proposition* 1 is guided by the observation that the eigenvalues of $\mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}^{-1} \mathbf{H}_k \mathbf{B}_k$ approach unity with increasing first-hop SNR. It will be demonstrated in Section 5 through numerical simulations that such an approximation results in negligible performance loss while reducing the computational complexity significantly. Applying *Proposition* 1, the transmit power of the relay node defined in (5) can be expressed as $\operatorname{tr}\left(\mathbf{F}\mathbf{\Psi}\mathbf{F}^H\right) = \operatorname{tr}(\tilde{\mathbf{T}}\mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}^{-1} \mathbf{H}_k \mathbf{B}_k \tilde{\mathbf{T}}^H) = \operatorname{tr}(\tilde{\mathbf{T}}\tilde{\mathbf{T}}^H)$. Therefore, problem (9) can be approximated as

$$\min_{\{\mathbf{B}_k\},\{\mathbf{W}_k\},\tilde{\mathbf{T}}} \quad \max_k \operatorname{tr}\left(\mathbf{I}_{N_b,k} + \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}_{\bar{k}}^{-1} \mathbf{H}_k \mathbf{B}_k \right)^{-1}$$

$$+ \operatorname{tr}\left(\mathbf{I}_{N_b,k} + \tilde{\mathbf{T}}^H \mathbf{G}_k^H \mathbf{G}_k \tilde{\mathbf{T}} \right)^{-1} \quad (25a)$$

$$\text{s.t.} \quad \operatorname{tr}\left(\mathbf{B}_k \mathbf{B}_k^H \right) \leq P_{s,k}, \text{ for } k = 1, \ldots, K, \quad (25b)$$

$$\operatorname{tr}\left(\tilde{\mathbf{T}}\tilde{\mathbf{T}}^H \right) \leq P_r. \quad (25c)$$

Note that the optimal receiver matrices $\{\mathbf{W}_k\}$ can be obtained as in (11). Interestingly, the source and relay optimization variables $\{\mathbf{B}_k\}$ and $\tilde{\mathbf{T}}$ are separable both in the objective function as well as in the constraints in problem (25). Therefore, applying the results from *Theorem* 2 and *Proposition* 1, we can decompose the problem (25) into the following source precoding matrices optimization problem:

$$\min_{\{\mathbf{B}_k\}} \quad \max_k \operatorname{tr}\left(\mathbf{I}_{N_b,k} + \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}_{\bar{k}}^{-1} \mathbf{H}_k \mathbf{B}_k \right)^{-1} \quad (26a)$$

$$\text{s.t.} \quad \operatorname{tr}(\mathbf{B}_k \mathbf{B}_k^H) \leq P_{s,k}, \text{ for } k = 1, \ldots, K, \quad (26b)$$

and the relay amplifying matrix optimization problem:

$$\min_{\tilde{\mathbf{T}}} \quad \max_k \quad \operatorname{tr}\left(\left[\mathbf{I}_{N_b,k} + \tilde{\mathbf{T}}^H \mathbf{G}_k^H \mathbf{G}_k \tilde{\mathbf{T}} \right]^{-1} \right) \quad (27a)$$

$$\text{s.t.} \quad \operatorname{tr}\left(\tilde{\mathbf{T}}\tilde{\mathbf{T}}^H \right) \leq P_r. \quad (27b)$$

Note that the objective function in (26a) can be interpreted as the MSE of the kth transmitter's signal vector $\mathbf{s}_k$. In particular, the equivalent received signal for the kth transmitter's signal in the first hop received at the relay node is given by $\mathbf{y}_r^{(k)} = \mathbf{H}_k \mathbf{B}_k \mathbf{s}_k + \sum_{j \neq k}^{K} \mathbf{H}_j \mathbf{B}_j \mathbf{s}_j + \mathbf{n}_r$, treating other users' signals as noise. As such, the corresponding MMSE receiver is given by $\mathbf{D}_k$ in (23). Thus the MSE expression in (26a) actually represents the equivalent first-hop MSE of the kth transmitter's signal $\mathbf{s}_k$.

Given the corresponding MMSE receiver $\mathbf{D}_k$, (26a) can be rewritten as

$$
\begin{aligned}
E_{\mathrm{s},k} &\triangleq \operatorname{tr}\left(\mathbf{D}_k^H\left(\mathbf{\Psi}+\sigma_{\mathrm{r}}^2\mathbf{I}_{N_{\mathrm{r}}}\right)\mathbf{D}_k - \mathbf{D}_k^H\mathbf{H}_k\mathbf{B}_k - \mathbf{B}_k^H\mathbf{H}_k^H\mathbf{D}_k + \mathbf{I}_{N_{\mathrm{b},k}}\right) \\
&= \operatorname{tr}\left(\left(\mathbf{D}_k^H\mathbf{H}\mathbf{\Upsilon}_k\mathbf{B} - \mathbf{\Omega}_k\right)\left(\mathbf{D}_k^H\mathbf{H}\mathbf{\Upsilon}_k\mathbf{B} - \mathbf{\Omega}_k\right)^H + \sigma_{\mathrm{r}}^2\mathbf{D}_k^H\mathbf{D}_k\right) \\
&= \left\|\operatorname{vec}\left(\mathbf{D}_k^H\mathbf{H}\mathbf{\Upsilon}_k\mathbf{B} - \mathbf{\Omega}_k\right)\right\|_2^2 + \sigma_{\mathrm{r}}^2\operatorname{tr}\left(\mathbf{D}_k^H\mathbf{D}_k\right) \qquad (28) \\
&= \left\|\begin{bmatrix} \omega_k \\ \left(\mathbf{I}_{N_{\mathrm{r}}}\otimes\mathbf{D}_k^H\mathbf{H}\mathbf{\Upsilon}_k\right)\operatorname{vec}(\mathbf{B}) - \operatorname{vec}(\mathbf{\Omega}_k) \end{bmatrix}\right\|_2^2,
\end{aligned}
$$

where $\omega_k \triangleq \sigma_{\mathrm{r}}\sqrt{\operatorname{tr}(\mathbf{D}_k^H\mathbf{D}_k)}$ and $\mathbf{\Upsilon}_k \triangleq [\mathbf{\Upsilon}_{k1},\ldots,\mathbf{\Upsilon}_{kk},\ldots, \mathbf{\Upsilon}_{kK}]$ with $\mathbf{\Upsilon}_{kk}=\mathbf{I}_{N_{\mathrm{r}}}$ and $\mathbf{\Upsilon}_{kj}=\mathbf{0}$, if $j\neq k$. Introducing an auxiliary variable t_{s}, problem (26) can be rewritten as the following second-order cone program (SOCP):

$$
\min_{\{\mathbf{B}_k\},t_{\mathrm{s}}} \quad t_{\mathrm{s}} \qquad (29a)
$$

$$
\text{s.t.} \quad \left\|\begin{bmatrix} \omega_k \\ \left(\mathbf{I}_{N_{\mathrm{r}}}\otimes\mathbf{D}_k^H\mathbf{H}\mathbf{\Upsilon}_k\right)\operatorname{vec}(\mathbf{B}) - \operatorname{vec}(\mathbf{\Omega}_k) \end{bmatrix}\right\|_2 \leq t_{\mathrm{s}},
$$
$$
\text{for } k=1,\ldots,K, \qquad (29b)
$$
$$
\|\operatorname{vec}(\mathbf{B}_k)\|_2 \leq \sqrt{P_{\mathrm{s},k}}, \quad \text{for } k=1,\ldots,K, \qquad (29c)
$$

which can be efficiently solved by standard optimization packages at a complexity order of $\mathcal{O}\big((\sum_{k=1}^K N_{\mathrm{b},k}^2 + 1)^3\big)$ [24]. Thus, we can update $\{\mathbf{D}_k\}$ and $\{\mathbf{B}_k\}$ in an alternating fashion.

Regarding the relay amplifying matrix optimization, by introducing $\tilde{\mathbf{T}}^H\tilde{\mathbf{T}} \triangleq \mathbf{Q}$, the relay matrix optimization problem (27) can be equivalently transformed to

$$
\min_{\mathbf{Q}\succeq 0} \ \max_{k} \ \operatorname{tr}\left(\left[\mathbf{I}_{N_{\mathrm{d},k}}+\mathbf{G}_k\mathbf{Q}\mathbf{G}_k^H\right]^{-1}\right)+N_{\mathrm{b},k}-N_{\mathrm{d},k} \ (30a)
$$
$$
\text{s.t.} \quad \operatorname{tr}(\mathbf{Q})\leq P_{\mathrm{r}}. \qquad (30b)
$$

Let us now introduce a matrix variable $\mathbf{Y}_k \succeq \left(\mathbf{I}_{N_{\mathrm{d},k}}+\mathbf{G}_k\mathbf{Q}\mathbf{G}_k^H\right)^{-1}$, and a scalar variable t_{r}. Using these variables, the relay optimization problem (30) can be equivalently rewritten as the following SDP:

$$
\min_{t_{\mathrm{r}},\mathbf{Q},\{\mathbf{Y}_k\}} \quad t_{\mathrm{r}} \qquad (31a)
$$
$$
\text{s.t.} \quad \operatorname{tr}(\mathbf{Y}_k)\leq t_{\mathrm{r}}, \ \text{for } k=1,\ldots,K, \qquad (31b)
$$
$$
\operatorname{tr}(\mathbf{Q})\leq P_{\mathrm{r}}, \qquad (31c)
$$
$$
\begin{pmatrix} \mathbf{Y}_k & \mathbf{I}_{N_{\mathrm{d}},k}, \\ \mathbf{I}_{N_{\mathrm{d}},k} & \mathbf{I}_{N_{\mathrm{d}},k}+\mathbf{G}_k\mathbf{Q}\mathbf{G}_k^H \end{pmatrix}\succeq 0,
$$
$$
\text{for } k=1,\ldots,K, \qquad (31d)
$$
$$
t_{\mathrm{r}}\geq 0, \qquad (31e)
$$
$$
\mathbf{Q}\succeq 0. \qquad (31f)
$$

Problem (31) is convex and the globally optimal solution can be easily obtained [23]. The complexity order of solving problem (31) is at most $\mathcal{O}\big((\sum_{k=1}^K N_{\mathrm{b},k}^2 + \sum_{k=1}^K N_{\mathrm{d},k}^2 + K+2)^{3.5}\big)$ [24]. Note that in the simplified algorithm, only the source matrices are obtained in an alternating fashion.

The overall joint optimization procedure is summarized in Table 2.

4 Two-way relaying

Two-way relaying is being considered as a promising technique for future generation wireless systems since two-way relaying can significantly improve spectral efficiency. Hence, in this section, we consider two-way relaying in an interference MIMO relay system where each pair of users transmit signals to each other through the assisting relay node. The information exchange in the two-way relay channel is accomplished in two time slots: MAC phase and the BC phase. During the MAC phase, all the users simultaneously send their messages to the relay node. Thus the signal vector received at the relay node during the MAC phase can be expressed as

$$
\mathbf{y}_{\mathrm{r}} = \sum_{k=1}^{2K}\mathbf{H}_k\mathbf{B}_k\mathbf{s}_k + \mathbf{n}_{\mathrm{r}}, \qquad (32)
$$

where $\mathbf{H}_{K+k} \triangleq \mathbf{G}_k^T$ for $k=1,\ldots,K$ and $\mathbf{n}_{\mathrm{r}}$ is the $N_{\mathrm{r}}\times 1$ AWGN vector received at the relay node.

Upon receiving $\mathbf{y}_{\mathrm{r}}$, the relay node linearly precodes the signal vector by an $N_{\mathrm{r}}\times N_{\mathrm{r}}$ amplifying matrix $\mathbf{F}$ and transmits the $N_{\mathrm{r}}\times 1$ precoded signal vector $\mathbf{x}_{\mathrm{r}}$ in the MAC phase:

$$
\mathbf{x}_{\mathrm{r}} = \mathbf{F}\mathbf{y}_{\mathrm{r}}. \qquad (33)
$$

The received signal at the kth user in the BC phase is given by

$$
\begin{aligned}
\mathbf{y}_k &= \mathbf{H}_k^T\mathbf{x}_{\mathrm{r}} + \mathbf{n}_{\mathrm{d},k} \qquad (34) \\
&= \mathbf{H}_k^T\mathbf{F}\mathbf{H}_{\bar{k}}\mathbf{B}_{\bar{k}}\mathbf{s}_{\bar{k}} + \mathbf{H}_k^T\mathbf{F}\left(\sum_{\substack{j=1 \\ j\neq\bar{k}}}^{2K}\mathbf{H}_j\mathbf{B}_j\mathbf{s}_j + \mathbf{n}_{\mathrm{r}}\right) + \mathbf{n}_{\mathrm{d},k}, \\
&\qquad\qquad\qquad\qquad\qquad \text{for } k=1,\ldots,2K,
\end{aligned}
$$

where we have defined $\bar{k}$ as the index of user k's partner (e.g., $\bar{1}=K+1,\overline{K+1}=1$), $\mathbf{n}_{\mathrm{d},k}$ is the $N_{\mathrm{d},k}\times 1$ AWGN vector at the kth destination node. As in the case of the one-way relaying system, all noises are assumed to be i.i.d. complex Gaussian random variables with mean zero and variance σ_{n}^2.

Table 2 Proposed simplified algorithm for solving problem (9)

1	Initialize $\mathbf{B}_k, \forall k$, satisfying the constraints (29c).
2	Repeat
	(a) Update $\mathbf{D}_k, \forall k$, as in (23).
	(b) Update $\mathbf{B}_k, \forall k$, through solving the subproblem (29).
3	Until convergence.
4	Solve the subproblem (31) to obtain $\mathbf{Q}$.

Since the transmitting node k knows its own signal vector $\mathbf{s}_k$ and the full CSI of the corresponding source-destination link $\mathbf{H}_k^T \mathbf{F} \mathbf{H}_k \mathbf{B}_k$, each transmitter can completely cancel the self-interference component in (34). Thus, the effective received signal vector at the kth receiving node is given by

$$\mathbf{y}_k = \mathbf{H}_k^T \mathbf{F} \mathbf{H}_{\bar{k}} \mathbf{B}_{\bar{k}} \mathbf{s}_{\bar{k}} + \mathbf{H}_k^T \mathbf{F} \left(\sum_{j \neq k, \bar{k}}^{2K} \mathbf{H}_j \mathbf{B}_j \mathbf{s}_j + \mathbf{n}_r \right) + \mathbf{n}_{d,k}, \quad (35)$$

$$= \bar{\mathbf{H}}_k \mathbf{s}_{\bar{k}} + \bar{\mathbf{n}}_{d,k}, \ \ \text{for } k = 1, \dots, 2K. \quad (36)$$

Using (33), the transmission power required at the relay node can be defined as

$$\mathrm{tr} \left(\mathrm{E} \{ \mathbf{x}_r \mathbf{x}_r^H \} \right) = \mathrm{tr} \left(\mathbf{F} \mathbf{\Psi} \mathbf{F}^H \right), \quad (37)$$

where $\mathbf{\Psi} \triangleq \mathrm{E}\{ \mathbf{y}_r \mathbf{y}_r^H \} = \sum_{k=1}^{2K} \mathbf{H}_k \mathbf{B}_k \mathbf{B}_k^H \mathbf{H}_k^H + \sigma_r^2 \mathbf{I}_{N_r}$ is the covariance matrix of the signal received at the relay node from all the transmitters. Furthermore, the MSE of the estimated signal using an $N_d \times N_b$ linear weight matrix $\mathbf{W}_k$ at the kth receiving node can be expressed as

$$E_k = \mathrm{tr} \begin{pmatrix} \mathbf{I}_{N_s,k} - \mathbf{W}_k^H \mathbf{H}_k^T \mathbf{F} \mathbf{H}_{\bar{k}} \mathbf{B}_{\bar{k}} - \mathbf{B}_{\bar{k}}^H \mathbf{H}_{\bar{k}}^H \mathbf{F}^H \mathbf{H}_k^* \mathbf{W}_k \\ + \sum_{\substack{j=1 \\ j \neq k}}^{2K} \mathbf{W}_k^H \mathbf{H}_k^T \mathbf{F} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H \mathbf{F}^H \mathbf{H}_k^* \mathbf{W}_k \\ + \sigma_r^2 \mathbf{W}_k^H \mathbf{H}_k^T \mathbf{F} \mathbf{F}^H \mathbf{H}_k^* \mathbf{W}_k + \sigma_d^2 \mathbf{W}_k^H \mathbf{W}_k \end{pmatrix},$$
$$\text{for } k = 1, \dots, 2K. \quad (38)$$

Similar to the case of one-way relaying, the problem of optimizing the transmit, relay, and receive matrices for the two-way scenario can be formulated as

$$\min_{\{\mathbf{B}_k\}, \mathbf{F}, \{\mathbf{W}_k\}} \ \max_k \ E_k \quad (39a)$$

$$\text{s.t.} \ \ \mathrm{tr} \left(\mathbf{F} \mathbf{\Psi} \mathbf{F}^H \right) \leq P_r \quad (39b)$$

$$\mathrm{tr}(\mathbf{B}_k \mathbf{B}_k^H) \leq P_{s,k}, \ \ \text{for } k = 1, \dots, 2K, \quad (39c)$$

where (39b) and (39c) indicates the corresponding transmit power constraints.

4.1 Iterative joint transceiver optimization
Similar to the one-way relaying scenario, it can be shown that the transmitter, relay, and receiver matrices can be optimized in an alternating fashion through solving convex sub-problems. In each iteration of the algorithm, the receiver weight matrices are updated as follows:

$$\mathbf{W}_k = \left(\sum_{\substack{j=1 \\ j \neq k}}^{2K} \mathbf{G}_k \mathbf{F} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H \mathbf{F}^H \mathbf{G}_k^H + \sigma_r^2 \mathbf{G}_k \mathbf{F} \mathbf{F}^H \mathbf{G}_k^H + \sigma_d^2 \mathbf{I}_{N_d} \right)^{-1}$$
$$\times \mathbf{G}_k \mathbf{F} \mathbf{H}_k \mathbf{B}_k, \ \ \text{for } k = 1, \dots, 2K. \quad (40)$$

The relay beamforming matrix $\mathbf{F}$ is optimized through solving the following SDP problem:

$$\min_{\tau_r, \mathbf{F}, \{\mathbf{\Xi}_k\}, \mathbf{\Phi}} \ \tau_r \quad (41a)$$

$$\text{s.t.} \ \ \mathrm{tr} \left(\mathbf{\Xi}_k \right) + \mathrm{tr} \left(\mathbf{F}_k^H \mathbf{W}_k \right) \leq \tau_r, \quad (41b)$$

$$\begin{bmatrix} \mathbf{\Xi}_k + \mathbf{W}_k^H \mathbf{H}_k^T \mathbf{F} \mathbf{H}_{\bar{k}} \mathbf{B}_{\bar{k}} + \mathbf{B}_{\bar{k}}^H \mathbf{H}_{\bar{k}}^H \mathbf{F}^H \mathbf{H}_k^* \mathbf{W}_k & \mathbf{W}_k^H \mathbf{H}_k^T \mathbf{F} \\ \mathbf{F}^H \mathbf{H}_k^* \mathbf{W}_k & \mathbf{\Psi}_{\bar{k}}^{-1} \end{bmatrix} \succeq \mathbf{0},$$
$$\text{for } k = 1, \dots, 2K, \quad (41c)$$

$$\begin{bmatrix} \mathbf{\Phi} & \mathbf{F} \\ \mathbf{F}^H & \mathbf{\Psi}^{-1} \end{bmatrix} \succeq \mathbf{0}, \quad (41d)$$

$$\mathrm{tr} \left(\mathbf{\Phi} \right) \leq P_r, \quad (41e)$$

where we have defined

$$\begin{cases} \mathbf{F} \mathbf{\Psi} \mathbf{F}^H \preceq \mathbf{\Phi}, \\ -\mathbf{W}_k^H \mathbf{H}_k^T \mathbf{F} \mathbf{H}_{\bar{k}} \mathbf{B}_{\bar{k}} - \mathbf{B}_{\bar{k}}^H \mathbf{H}_{\bar{k}}^H \mathbf{F}^H \mathbf{H}_k^* \mathbf{W}_k + \mathbf{W}_k^H \mathbf{H}_k^T \mathbf{F} \mathbf{\Psi}_{\bar{k}} \mathbf{F}^H \mathbf{H}_k^* \mathbf{W}_k \preceq \mathbf{\Xi}_k. \end{cases}$$
$$(42)$$

Finally, the optimal source precoding matrices are obtained by solving

$$\min_{t_s, \mathbf{b}} \ \tau_s \quad (43a)$$

$$\text{s.t.} \ \ \begin{pmatrix} \tau_s - \theta_k + \mathbf{c}_k^H \mathbf{b} + \mathbf{b}^H \mathbf{c}_k & \mathbf{b}^H \\ \mathbf{b} & \tilde{\mathbf{G}}_k^{-1} \end{pmatrix} \succeq \mathbf{0},$$
$$\text{for } k = 1, \dots, 2K, \quad (43b)$$

$$\begin{pmatrix} \bar{P}_r & \mathbf{b}^H \\ \mathbf{b} & \mathbf{M}^{-1} \end{pmatrix} \succeq \mathbf{0}, \quad (43c)$$

$$\begin{pmatrix} P_{s,k} & \mathbf{b}^H \mathcal{I}_k^{\frac{1}{2}} \\ \mathcal{I}_k^{\frac{1}{2}} \mathbf{b} & \mathbf{I}_p \end{pmatrix} \succeq \mathbf{0}, \ \ \text{for } k = 1, \dots, 2K, \quad (43d)$$

where

$$\theta_k \triangleq \mathrm{tr} \left(\sigma_r^2 \mathbf{W}_k^H \mathbf{G}_k \mathbf{F} \mathbf{F}^H \mathbf{G}_k^H \mathbf{W}_k + \sigma_d^2 \mathbf{W}_k^H \mathbf{W}_k \right)$$
$$+ N_{b,k}, \ \ \text{for } k = 1, \dots, 2K, \quad (44a)$$

$$\tilde{\mathbf{G}}_k \triangleq \mathrm{bd} \left(\mathbf{I}_{N_{b,1}} \otimes \left(\tilde{\mathbf{H}}_{k,1}^H \tilde{\mathbf{H}}_{k,1} \right), \cdots, \mathbf{I}_{N_{b,2K}} \right.$$
$$\left. \otimes \left(\tilde{\mathbf{H}}_{k,2K}^H \tilde{\mathbf{H}}_{k,2K} \right) \right), \text{for } k = 1, \dots, 2K, \quad (44b)$$

$$\mathbf{c}_k \triangleq \left[\left(\mathrm{vec} \left(\tilde{\mathbf{C}}_{k,1} \right) \right)^T, \dots, \left(\mathrm{vec} \left(\tilde{\mathbf{C}}_{k,2K} \right) \right)^T \right]^T, \quad (44c)$$

$$\tilde{\mathbf{C}}_{k,k} = \tilde{\mathbf{H}}_{k,k}, \quad (44d)$$

$$\tilde{\mathbf{C}}_{k,j} = \mathbf{0}_{N_{b,k} \times N_{s,j}}, \ \ \text{for } j \neq k, \quad (44e)$$

$$\mathbf{b} \triangleq \left[\mathbf{b}_1^T, \dots, \mathbf{b}_{2K}^T \right]^T, \quad (44f)$$

$$\mathbf{M} \triangleq \mathrm{bd} \left(\mathbf{I}_{N_{b,1}} \otimes \left(\mathbf{M}_1^H \mathbf{M}_1 \right), \dots, \mathbf{I}_{N_{b,K}} \otimes \left(\mathbf{M}_{2K}^H \mathbf{M}_{2K} \right) \right), \quad (44g)$$

$$p \triangleq \sum_{k=1}^{2K} N_{s,k} N_{b,k}. \quad (44h)$$

4.2 Simplified non-iterative approach

Assuming moderate SNR in the MAC phase, it can be shown, similar to the one-way relaying case, that the generic structure of the relay matrix $\mathbf{F}$ is defined as $\mathbf{F} = \mathbf{TD}^H$. Using this particular structure of $\mathbf{F}$, the MSE at the kth receiver can be equivalently decomposed into two parts as shown below:

$$E_k = \text{tr}\left(\mathbf{I}_{N_b,k} + \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}_{\bar{k}}^{-1} \mathbf{H}_k \mathbf{B}_k\right)^{-1}$$
$$+ \text{tr}\left(\left(\mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}_{\bar{k}}^{-1} \mathbf{H}_k \mathbf{B}_k\right)^{-1} + \tilde{\mathbf{T}}^H \mathbf{H}_{\bar{k}}^* \mathbf{H}_{\bar{k}}^T \tilde{\mathbf{T}}\right)^{-1}. \quad (45)$$

Accordingly, the joint precoding design problem (25) can be decomposed into two sub-problems, namely, the source precoding matrices optimization problem:

$$\min_{\{\mathbf{B}_k\}} \; \max_k \; \text{tr}\left(\mathbf{I}_{N_b,k} + \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}_{\bar{k}}^{-1} \mathbf{H}_k \mathbf{B}_k\right)^{-1} \quad (46a)$$

$$\text{s.t.} \quad \text{tr}\left(\mathbf{B}_k \mathbf{B}_k^H\right) \le P_{s,k}, \; \text{for } k = 1, \dots, K, \quad (46b)$$

and the relay beamforming matrix optimization problem:

$$\min_{\tilde{\mathbf{T}}} \; \max_k \; \text{tr}\left(\left[\mathbf{I}_{N_b,k} + \tilde{\mathbf{T}}^H \mathbf{H}_{\bar{k}}^* \mathbf{H}_{\bar{k}}^T \tilde{\mathbf{T}}\right]^{-1}\right) \quad (47a)$$

$$\text{s.t.} \quad \text{tr}\left(\tilde{\mathbf{T}} \tilde{\mathbf{T}}^H\right) \le P_r, \quad (47b)$$

which can be solved following the similar approach as for the one-way relaying scenario.

5 Numerical simulations

In this section, we analyze the performance of the proposed one- and two-way MIMO relay interference system optimization algorithms through numerical examples. For simplicity, we assume that the source and the destination nodes are equipped with N_s and N_d antennas each, respectively, and $P_{s,k} = P_s$, $\forall k$. We simulated a flat Rayleigh fading environment such that the channel matrices have zero-mean entries with variances $1/N_s$ for $\mathbf{H}_k$, $\forall k$, and $1/N_r$ for $\mathbf{G}_k$, $\forall k$. All the simulation results were obtained by averaging over 500 independent channel realizations.

The performance of the proposed min-max MSE algorithms have been compared with that of the naive AF (NAF) algorithm in terms of both MSE and bit error rate (BER). The NAF algorithm is a simple baseline scheme that forwards the signals at the transmitters and the relay node assigning equal power to each data stream. In particular, the source and the relay matrices, in their simplest forms, in the NAF scheme are defined as

$$\begin{cases} \mathbf{B}_k = \sqrt{P_s/N_s}\, \mathbf{I}_{N_s}, \; \text{for } k = 1, \dots, K, \\ \mathbf{F} = \sqrt{P_r/\text{tr}(\mathbf{\Psi})}\, \mathbf{I}_{N_r}. \end{cases} \quad (48)$$

In the first example, we compare the performance of the proposed min-max MSE-based one-way algorithms with

that of the sum-MSE minimization algorithm in [8] as well as the NAF approach in terms of the MSE normalized by the number of data streams (NMSE) with $K = 3$, $N_s = 3$, $N_r = 9$, and $N_d = 3$. Figure 2 shows the NMSE performance of the algorithms versus transmit power P_s with fixed $P_r = 20$ dB. Note that for the proposed simplified *non-iterative* algorithm, we plot the NMSE of the user with the worst channel (Worst) as well as the average per-stream MSE of all the users (Avg.). On the other hand, for the rest of the algorithms, the worst-user NMSE has been plotted. The results clearly indicate that the proposed joint optimization algorithms consistently yield better performance compared to the existing schemes. It can also be revealed that the proposed iterative algorithm has the best MSE performance compared to the other approaches over the entire P_s range. It is no surprise that the NAF algorithm yields much higher MSE compared to the other schemes since the NAF algorithm performs no optimization operation. Most importantly, the iterative sum-MSE minimization algorithm in [8] always penalizes the user with the worst channel condition.

Since the NAF algorithm does not allocate the transmit power optimally and equally divides the power among multiple data streams instead, the inter-stream interference and the inter-user interference increase significantly at higher transmit power. Hence, the MSE of the NAF algorithm does not improve notably at higher transmit power.

Further analysis of the results in Fig. 2 reveals that the proposed simplified algorithm yields the worst-user MSE performance which is comparable to that of the iterative algorithm, even at low P_s region. This observation illustrates that the approximation made in the simplified algorithm encounters negligible performance loss compared to the iterative optimal design. On the other hand, the computational complexity of the proposed simplified optimization is less than that of even one iteration of the
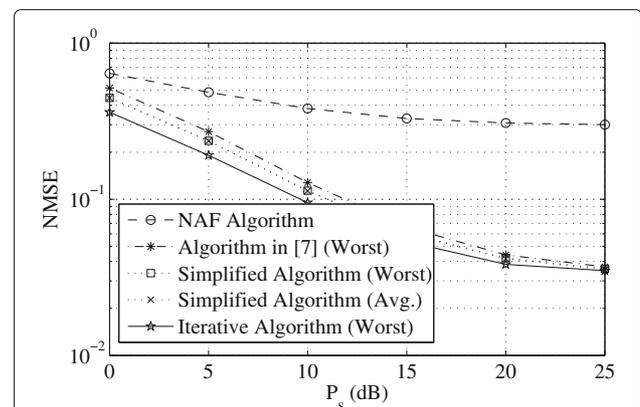
Fig. 2 Example 1: normalized MSE versus P_s. $K = 3$, $N_s = 3$, $N_r = 9$, $N_d = 3$, $P_r = 20$ dB

Table 3 Iterations required till convergence in the proposed algorithm

P_s (dB)	0	5	10	15	20	25
Iterations	3	3	3	4	5	5

Fig. 4 Example 3: MSE versus P_s in two-way relaying. Varying number of links, $N_s = 2, N_r = KN_s, N_d = 6, P_r = 20$ dB

iterative design, making it much more attractive for practical interference MIMO relay systems. The number of iterations required for convergence up to 10^{-3} in terms of MSE in a random channel realization for the iterative algorithm are listed in Table 3.

In the next example, we focus on the proposed simplified optimization scheme and compare its performance with that of the proposed iterative approach and the NAF algorithm in terms of BER. Quadrature phase-shift keying (QPSK) signal constellations were assumed to modulate the transmitted signals and maximum-likelihood detection is applied at the receivers. We set $K = 3, N_s = 2, N_r = 6, N_d = 3$, and transmit $1000N_s$ randomly generated bits from each transmitter in each channel realization. The BER performance of the algorithms are shown in Fig. 3 versus P_s with $P_r = 20$ dB. As we can see, the proposed simplified algorithm yields a much lower BER compared to the conventional NAF scheme. Compared with the iterative approach the simplified algorithm has much lower computational task at the cost of marginal performance loss.

In the last couple of examples, we analyze the performance of the two-way MIMO relaying scheme. The NMSE performance of the two-way relaying algorithms is shown for different number of communication links K in Fig. 4. This time we set $N_s = 2, N_r = KN_s$, and $N_d = 6$ to plot the NMSE of the proposed algorithms versus P_s with $P_r = 20$ dB. It can be clearly seen from Fig. 4 that as the number of links increases, the worst-user

MSE keeps increasing. This is due to the additional cross-link interferences generated by the increased number of active users.

In Fig. 5, the BER performance of the proposed two-way relaying algorithms has been compared with the sum-MSE-based algorithms originally proposed for one-way relaying in [8–10]. QPSK signal constellations were assumed to modulate the transmitted signals. We set $N_s = 2, K = 3, N_r = KN_s, N_d = 6, P_r = 20$ dB, and transmit $1000N_s$ randomly generated bits from each transmitter in each channel realization. Most importantly, the iterative sum-MSE minimization algorithms in [8–10] always penalize the user with the worst channel condition in the two-way relaying system.

6 Conclusions

We considered a two-hop interference MIMO relay system and developed schemes to minimize the worst-user

Fig. 3 Example 2: BER versus P_s. $K = 3, N_s = 2, N_r = 6, N_d = 3$, $P_r = 20$ dB
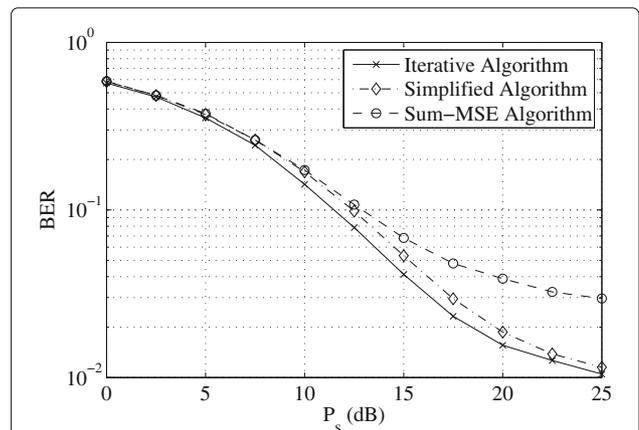
Fig. 5 Example 4: BER versus P_s in two-way relaying for different algorithms, $N_s = 2, K = 3, N_r = KN_s, N_d = 6, P_r = 20$ dB

MSE of signal estimation for both one- and two-way relaying schemes. At first, we proposed an iterative solution for both relaying schemes by solving several convex subproblems alternatingly and in an iterative fashion. Then to reduce the computational overhead of the optimization approach, we develop a simplified non-iterative algorithm using the error covariance matrix decomposition technique based on the high SNR assumption. Simulation results have illustrated that the proposed simplified approach performs nearly as well as the iterative approach, while offering significant reduction in computational complexity.

Endnote
[1] The min-max MSE criterion is considered by many to be more desirable than the min-sum MSE criterion in [8–10] because fairness is imposed and weaker users are not being sacrificed for the minimization of the sum.

Appendix 1: Proof of Theorem 1
For given $\{\mathbf{B}_k\}$ and $\{\mathbf{W}_k\}$, problem (9) reduces to

$$\min_{\mathbf{F}} \quad \tau \tag{49a}$$

$$\text{s.t.} \quad E_k \leq \tau, \ \text{for } k = 1, \ldots, K, \tag{49b}$$

$$\text{tr}\left(\mathbf{F}\boldsymbol{\Psi}\mathbf{F}^H\right) \leq P_{\mathrm{r}}. \tag{49c}$$

The Lagrangian function of problem (49) can be written as

$$
\begin{aligned}
&\mathcal{L}\left(\mathbf{F}, \{\lambda_{\mathrm{s},k}\}, \lambda_{\mathrm{r}}\right) \\
&= \tau + \sum_{k=1}^{K} \lambda_{\mathrm{e},k} \text{tr}\left(\begin{array}{c} \mathbf{I}_{N_{\mathrm{s}},k} - 2\text{Re}\left(\mathbf{B}_k^H \mathbf{H}_k^H \mathbf{F}^H \mathbf{G}_k^H \mathbf{W}_k\right) \\ + \sum_{j=1}^{K} \mathbf{W}_k^H \mathbf{G}_k \mathbf{F} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H \mathbf{F}^H \mathbf{G}_k^H \mathbf{W}_k \\ + \sigma_{\mathrm{r}}^2 \mathbf{W}_k^H \mathbf{G}_k \mathbf{F} \mathbf{F}^H \mathbf{G}_k^H \mathbf{W}_k + \sigma_{\mathrm{d}}^2 \mathbf{W}_k^H \mathbf{W}_k - \tau \end{array}\right) \\
&\quad + \lambda_{\mathrm{r}}\left(\text{tr}\left(\mathbf{F}\left(\sum_{k=1}^{K} \mathbf{H}_k \mathbf{B}_k \mathbf{B}_k^H \mathbf{H}_k^H + \sigma_{\mathrm{r}}^2 \mathbf{I}_{N_{\mathrm{r}}}\right) \mathbf{F}^H\right) - P_{\mathrm{r}}\right). \tag{50}
\end{aligned}
$$

The derivative of the Lagrangian function over $\mathbf{F}^H$ is given by

$$
\frac{\partial \mathcal{L}}{\partial \mathbf{F}^H} = \sum_{k=1}^{K} \lambda_{\mathrm{e},k}\left(-\mathbf{G}_k^H \mathbf{W}_k \mathbf{B}_k^H \mathbf{H}_k^H + \sum_{j=1}^{K} \mathbf{G}_k^H \mathbf{W}_k \mathbf{W}_k^H \mathbf{G}_k \mathbf{F} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H \right.
$$
$$
\left. + \sigma_{\mathrm{r}}^2 \mathbf{G}_k^H \mathbf{W}_k \mathbf{W}_k^H \mathbf{G}_k \mathbf{F}\right) + \lambda_{\mathrm{r}} \mathbf{F}\left(\sum_{k=1}^{K} \mathbf{H}_k \mathbf{B}_k \mathbf{B}_k^H \mathbf{H}_k^H + \sigma_{\mathrm{r}}^2 \mathbf{I}_{N_{\mathrm{r}}}\right). \tag{51}
$$

Rearranging the terms in (51), $\frac{\partial \mathcal{L}}{\partial \mathbf{F}^H}$ can be expressed as

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \mathbf{F}^H} &= \sum_{k=1}^{K} -\lambda_{\mathrm{e},k} \mathbf{G}_k^H \mathbf{W}_k \mathbf{B}_k^H \mathbf{H}_k^H \\
&\quad + \left(\sum_{i=1}^{K} \lambda_{\mathrm{e},i} \mathbf{G}_i^H \mathbf{W}_i \mathbf{W}_i^H \mathbf{G}_i + \lambda_{\mathrm{r}} \mathbf{I}_{N_{\mathrm{r}}}\right) \mathbf{F} \\
&\quad \times \left(\sum_{j=1}^{K} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H + \sigma_{\mathrm{r}}^2 \mathbf{I}_{N_{\mathrm{r}}}\right). \tag{52}
\end{aligned}
$$

Equating $\frac{\partial \mathcal{L}}{\partial \mathbf{F}^*} = 0$, we have the optimal relay filter given by

$$\mathbf{F} = \sum_{k=1}^{K} \mathbf{T}_k \mathbf{D}_k^H \tag{53}$$

with

$$
\begin{cases}
\mathbf{T}_k \triangleq \lambda_{\mathrm{e},k}\left(\sum_{i=1}^{K} \lambda_{\mathrm{e},i} \mathbf{G}_i^H \mathbf{W}_i \mathbf{W}_i^H \mathbf{G}_i + \lambda_{\mathrm{r}} \mathbf{I}_{N_{\mathrm{r}}}\right)^{-1} \mathbf{G}_k^H \mathbf{W}_k, \\
\mathbf{D}_k \triangleq \left(\sum_{j=1}^{K} \mathbf{H}_j \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_j^H + \sigma_{\mathrm{r}}^2 \mathbf{I}_{N_{\mathrm{r}}}\right)^{-1} \mathbf{H}_k \mathbf{B}_k.
\end{cases} \tag{54}
$$

Denoting $\mathbf{T} \triangleq [\mathbf{T}_1 \cdots \mathbf{T}_K]$ and $\mathbf{D} \triangleq [\mathbf{D}_1 \cdots \mathbf{D}_K]$, $\mathbf{F}$ can be expressed as $\mathbf{F} = \mathbf{T}\mathbf{D}^H$. $\square$

Appendix 2: Proof of Theorem 2
The MSE in (9a) can be rewritten as

$$
E_k = \left[\mathbf{I}_{N_{\mathrm{s}},k} + \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{F}^H \mathbf{G}_k^H \bar{\mathbf{C}}_k^{-1} \mathbf{G}_k \mathbf{F} \mathbf{H}_k \mathbf{B}_k\right]^{-1} \tag{55}
$$
$$
= \text{tr}\left(\mathbf{I}_{N_{\mathrm{s}},k} - \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{F}^H \mathbf{G}_k^H \left(\mathbf{G}_k \mathbf{F} \mathbf{H}_k \mathbf{B}_k \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{F}^H \mathbf{G}_k^H\right.\right.
$$
$$
\left.\left. + \bar{\mathbf{C}}_k\right)^{-1} \mathbf{G}_k \mathbf{F} \mathbf{H}_k \mathbf{B}_k\right) \tag{56}
$$
$$
= \text{tr}\left(\mathbf{I}_{N_{\mathrm{s}},k} - \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{F}^H \mathbf{G}_k^H \left(\mathbf{G}_k \mathbf{F} \boldsymbol{\Psi} \mathbf{F}^H \mathbf{G}_k^H\right.\right.
$$
$$
\left.\left. + \sigma_{\mathrm{d}}^2 \mathbf{I}_{N_{\mathrm{d}},k}\right)^{-1} \mathbf{G}_k \mathbf{F} \mathbf{H}_k \mathbf{B}_k\right) \tag{57}
$$
$$
= \text{tr}\left(\mathbf{I}_{N_{\mathrm{s}},k} - \mathbf{B}_k^H \mathbf{H}_k^H \left[\boldsymbol{\Psi}^{-1} - \left(\boldsymbol{\Psi} \mathbf{F}^H \mathbf{G}_k^H \mathbf{G}_k \mathbf{F} \boldsymbol{\Psi}\right.\right.\right.
$$
$$
\left.\left.\left. + \boldsymbol{\Psi}\right)^{-1}\right] \mathbf{H}_k \mathbf{B}_k\right) \tag{58}
$$
$$
= \text{tr}\left(\mathbf{I}_{N_{\mathrm{s}},k} + \mathbf{B}_k^H \mathbf{H}_k^H \boldsymbol{\Psi}_k^{-1} \mathbf{H}_k \mathbf{B}_k\right)^{-1}
$$
$$
+ \text{tr}\left(\mathbf{B}_k^H \mathbf{H}_k^H \left(\boldsymbol{\Psi} \mathbf{F}^H \mathbf{G}_k^H \mathbf{G}_k \mathbf{F} \boldsymbol{\Psi} + \boldsymbol{\Psi}\right)^{-1} \times \mathbf{H}_k \mathbf{B}_k\right), \tag{59}
$$

where we used matrix inversion lemma $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1}\right)^{-1}\mathbf{DA}^{-1}$ to obtain (56) and the first term in (59) whereas the matrix identity $\mathbf{B}^H(\mathbf{BCB}^H + \mathbf{I})^{-1}\mathbf{B} = \mathbf{C}^{-1} - (\mathbf{CB}^H\mathbf{BC} + \mathbf{C})^{-1}$ is used to obtain (58) in the above derivation. Note that the first term

in (59) is irrelevant to $\mathbf{F}$. Hence for given source matrices, the problem of optimizing $\mathbf{F}$ can be simplified as

$$\min_{\mathbf{F}} \quad \mathrm{tr}\left(\mathbf{B}_k^H \mathbf{H}_k^H \left(\mathbf{\Psi}\mathbf{F}^H\mathbf{G}_k^H\mathbf{G}_k\mathbf{F}\mathbf{\Psi} + \mathbf{\Psi}\right)^{-1} \mathbf{H}_k\mathbf{B}_k\right) \tag{60a}$$

$$\text{s.t.} \quad \mathrm{tr}\left(\mathbf{F}\mathbf{\Psi}\mathbf{F}^H\right) \leq P_{\mathrm{r}}. \tag{60b}$$

By introducing $\tilde{\mathbf{F}} = \mathbf{F}\mathbf{\Psi}^{\frac{1}{2}}$, problem (60) can be rewritten as

$$\min_{\tilde{\mathbf{F}}} \quad \mathrm{tr}\left(\mathbf{B}_k^H \mathbf{H}_k^H \mathbf{\Psi}^{-\frac{1}{2}} \left(\tilde{\mathbf{F}}^H\mathbf{G}_k^H\mathbf{G}_k\tilde{\mathbf{F}} + \mathbf{I}_{N_{\mathrm{r}}}\right)^{-1} \mathbf{\Psi}^{-\frac{1}{2}} \mathbf{H}_k\mathbf{B}_k\right) \tag{61a}$$

$$\text{s.t.} \quad \mathrm{tr}\left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H\right) \leq P_{\mathrm{r}}. \tag{61b}$$

Let us write the eigenvalue decomposition (EVD) $\mathbf{G}_k^H\mathbf{G}_k = \mathbf{V}_{\mathrm{g}}\mathbf{\Lambda}_{\mathrm{g}}\mathbf{V}_{\mathrm{g}}^H$ and the singular value decomposition (SVD) $\mathbf{\Psi}^{-\frac{1}{2}}\mathbf{H}_k\mathbf{B}_k = \mathbf{U}_\psi \mathbf{\Lambda}_\psi \mathbf{V}_\psi^H$. The following lemma defines the optimal $\tilde{\mathbf{F}}$.

Lemma 1 ([25] Lemma 2) *For matrices* $\mathbf{A}, \bar{\mathbf{T}}, \mathbf{H}$ *of dimensions* $m \times n$, $l \times m$, *and* $k \times l$, *respectively, with* $k, l, m \geq n$, $r \triangleq \mathrm{rank}(\mathbf{H}) \geq n$ *and* $\mathrm{rank}(\bar{\mathbf{T}}) = n$, *the solution to the optimization problem*

$$\min_{\bar{\mathbf{T}}} \quad tr\left(\mathbf{A}^H \left(\bar{\mathbf{T}}^H\mathbf{H}^H\mathbf{H}\bar{\mathbf{T}} + \mathbf{I}_m\right)^{-1} \mathbf{A}\right) \tag{62a}$$

$$s.t. \quad tr\left(\bar{\mathbf{T}}\bar{\mathbf{T}}^H\right) \leq p, \tag{62b}$$

is given by $\bar{\mathbf{T}} = \tilde{\mathbf{V}}_{\mathrm{h}}\mathbf{\Lambda}_{\mathrm{T}}\mathbf{U}_{\mathrm{a}}^H$ *in terms of the SVD of* $\bar{\mathbf{T}}$. *Here* $\mathbf{H} = \mathbf{U}_{\mathrm{h}}\mathbf{\Sigma}_{\mathrm{h}}\mathbf{V}_{\mathrm{h}}^H$ *and* $\mathbf{A} = \mathbf{U}_{\mathrm{a}}\mathbf{\Sigma}_{\mathrm{a}}\mathbf{V}_{\mathrm{a}}^H$ *are the SVDs of* $\mathbf{H}$ *and* $\mathbf{A}$, *respectively, with the diagonal elements of* $\mathbf{\Sigma}_{\mathrm{h}}$ *and* $\mathbf{\Sigma}_{\mathrm{a}}$ *sorted in a decreasing order, and* $\tilde{\mathbf{V}}_{\mathrm{h}}$ *contains the leftmost* n *columns of* $\mathbf{V}_{\mathrm{h}}$.

According to Lemma 1, the optimal $\tilde{\mathbf{F}}$ in (61) has the SVD $\tilde{\mathbf{F}} = \tilde{\mathbf{V}}_{\mathrm{g}}\mathbf{\Lambda}_{\mathrm{f}}\mathbf{U}_\psi^H$ where $\tilde{\mathbf{V}}_{\mathrm{g}}$ contains the left-most columns of $\mathbf{V}_{\mathrm{g}}$ corresponding to the non-zero eigenvalues. Then after some simple manipulations, $\tilde{\mathbf{F}}$ can be rewritten as $\tilde{\mathbf{F}} = \tilde{\mathbf{V}}_{\mathrm{g}}\mathbf{\Lambda}_{\mathrm{f}}\mathbf{\Lambda}_\psi^{-1}\mathbf{V}_\psi^H\mathbf{V}_\psi\mathbf{\Lambda}_\psi\mathbf{U}_\psi^H = \tilde{\mathbf{T}}\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-\frac{1}{2}}$ where $\tilde{\mathbf{T}} \triangleq \tilde{\mathbf{V}}_{\mathrm{g}}\mathbf{\Lambda}_{\mathrm{f}}\mathbf{\Lambda}_\psi^{-1}\mathbf{V}_\psi^H$. Hence $\mathbf{F}$ can be expressed as $\mathbf{F} = \tilde{\mathbf{T}}\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}$. Interestingly, $\mathbf{F} = \tilde{\mathbf{T}}\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}$ can be expressed as $\mathbf{F} = \tilde{\mathbf{T}}\tilde{\mathbf{D}}^H$, which is structurally identical to the one defined in Theorem 1.

Applying this structure of the relay matrix, the second term in (59) can be written as

$$\mathrm{tr}\left(\mathbf{B}_k^H\mathbf{H}_k^H\left(\mathbf{\Psi}\mathbf{F}^H\mathbf{G}_k^H\mathbf{G}_k\mathbf{F}\mathbf{\Psi} + \mathbf{\Psi}\right)^{-1}\mathbf{H}_k\mathbf{B}_k\right)$$

$$= \mathrm{tr}\left(\mathbf{B}_k^H\mathbf{H}_k^H\left(\mathbf{\Psi}\mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k\tilde{\mathbf{T}}^H\mathbf{G}_k^H\mathbf{G}_k\tilde{\mathbf{T}}\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\mathbf{\Psi} + \mathbf{\Psi}\right)^{-1}\mathbf{H}_k\mathbf{B}_k\right)$$

$$= \mathrm{tr}\left(\mathbf{B}_k^H\mathbf{H}_k^H\left(\mathbf{\Psi}^{-1} - \mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k\left(\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k\right.\right.\right.$$
$$\left.\left.\left. + \left(\tilde{\mathbf{T}}^H\mathbf{G}_k^H\mathbf{G}_k\tilde{\mathbf{T}}\right)^{-1}\right)^{-1}\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\right)\mathbf{H}_k\mathbf{B}_k\right)$$

$$= \mathrm{tr}\left(\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k - \mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k\left(\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k\right.\right.$$
$$\left.\left. + \left(\tilde{\mathbf{T}}^H\mathbf{G}_k^H\mathbf{G}_k\tilde{\mathbf{T}}\right)^{-1}\right)^{-1}\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k\right)$$

$$= \mathrm{tr}\left(\left(\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k\right)^{-1} + \tilde{\mathbf{T}}^H\mathbf{G}_k^H\mathbf{G}_k\tilde{\mathbf{T}}\right)^{-1}. \tag{63}$$

Thus the MSE in (9a) can be expressed as the sum of two MSEs given by

$$E_k = \mathrm{tr}\left(\mathbf{I}_{N_{\mathrm{s}},k} + \mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}_{\bar{k}}^{-1}\mathbf{H}_k\mathbf{B}_k\right)^{-1} \tag{64}$$
$$+ \mathrm{tr}\left(\left(\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k\right)^{-1} + \tilde{\mathbf{T}}^H\mathbf{G}_k^H\mathbf{G}_k\tilde{\mathbf{T}}\right)^{-1}.$$

$\square$

Appendix 3: Proof of Proposition 1
Assuming that the first-hop SNR is reasonably high, it emerges that $\sum_{j=1}^{K}\mathbf{H}_j\mathbf{B}_j\mathbf{B}_j^H\mathbf{H}_j^H \gg \sigma_{\mathrm{r}}^2\mathbf{I}_{N_{\mathrm{r}}}$ where $\mathbf{A} \gg \mathbf{B}$ effectively means that the eigenvalues of $\mathbf{A} - \mathbf{B}$ are much greater than zero. Hence,

$$\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k = \mathbf{B}_k^H\mathbf{H}_k^H\left(\sum_{j=1}^{K}\mathbf{H}_j\mathbf{B}_j\mathbf{B}_j^H\mathbf{H}_j^H + \sigma_{\mathrm{r}}^2\mathbf{I}_{N_{\mathrm{r}}}\right)^{-1}\mathbf{H}_k\mathbf{B}_k$$

$$\approx \mathbf{B}_k^H\mathbf{H}_k^H\left(\sum_{j=1}^{K}\mathbf{H}_j\mathbf{B}_j\mathbf{B}_j^H\mathbf{H}_j^H\right)^{-1}\mathbf{H}_k\mathbf{B}_k. \tag{65}$$

Let $\mathbf{U}_k\mathbf{\Lambda}_k\mathbf{U}_k^H$ be the EVD of $\mathbf{H}_k\mathbf{B}_k\mathbf{B}_k^H\mathbf{H}_k^H$. Without loss of generality, we express $\mathbf{U}_k = \left[\mathbf{U}_k^{(\bar{0})} \; \mathbf{U}_k^{(0)}\right]$ and $\mathbf{\Lambda}_k = \begin{bmatrix} \mathbf{\Lambda}_k^{(\bar{0})} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$, where $\mathbf{U}_k^{(\bar{0})}$ and $\mathbf{U}_k^{(0)}$ contain the eigenvectors corresponding to the non-zero and zero eigenvalues, respectively, in $\mathbf{U}_k$ while $\mathbf{\Lambda}_k^{(\bar{0})}$ is an $N_{\mathrm{b},k} \times N_{\mathrm{b},k}$ diagonal matrix containing the non-zero eigenvalues as the main

diagonal. Thus $\mathbf{H}_k\mathbf{B}_k = \mathbf{U}_k\bar{\mathbf{\Lambda}}_k^{(\bar{0})}$ where $\bar{\mathbf{\Lambda}}_k^{(\bar{0})} = \begin{bmatrix} \mathbf{\Lambda}_k^{(\bar{0})\frac{1}{2}} \\ \mathbf{0} \end{bmatrix}$.
Similarly, we obtain the following EVD

$$
\sum_{\substack{j=1 \\ k \neq k}}^{K} \mathbf{H}_j\mathbf{B}_j\mathbf{B}_j^H\mathbf{H}_j^H = \mathbf{U}_{\bar{k}}\mathbf{\Lambda}_{\bar{k}}\mathbf{U}_{\bar{k}}^H
$$

$$
= \begin{bmatrix} \mathbf{U}_{\bar{k}}^{(\bar{0})} & \mathbf{U}_{\bar{k}}^{(0)} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_{\bar{k}}^{(\bar{0})} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\bar{k}}^{(\bar{0})} & \mathbf{U}_{\bar{k}}^{(0)} \end{bmatrix}^H \tag{66}
$$

$$
= \begin{bmatrix} \mathbf{U}_{\bar{k}}^{(0)} & \mathbf{U}_{\bar{k}}^{(\bar{0})} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{\bar{k}}^{(\bar{0})} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\bar{k}}^{(0)} & \mathbf{U}_{\bar{k}}^{(\bar{0})} \end{bmatrix}^H . \tag{67}
$$

Substituting $\mathbf{H}_k\mathbf{B}_k$ in (65) with $\mathbf{H}_k\mathbf{B}_k = \mathbf{U}_k\bar{\mathbf{\Lambda}}_k^{(\bar{0})}$, we obtain

$$
\mathbf{B}_k^H\mathbf{H}_k^H \left(\sum_{j=1}^{K} \mathbf{H}_j\mathbf{B}_j\mathbf{B}_j^H\mathbf{H}_j^H \right)^{-1} \mathbf{H}_k\mathbf{B}_k
$$

$$
= \bar{\mathbf{\Lambda}}_k^{(\bar{0})H} \left(\mathbf{\Lambda}_k + \mathbf{U}_k^H\mathbf{U}_{\bar{k}}\mathbf{\Lambda}_{\bar{k}}\mathbf{U}_{\bar{k}}^H\mathbf{U}_k \right)^{-1} \bar{\mathbf{\Lambda}}_k^{(\bar{0})} . \tag{68}
$$

Now we rewrite $\mathbf{U}_k^H\mathbf{U}_{\bar{k}}$ as

$$
\mathbf{U}_k^H\mathbf{U}_{\bar{k}} = \begin{bmatrix} \mathbf{U}_k^{(\bar{0})} & \mathbf{U}_k^{(0)} \end{bmatrix}^H \begin{bmatrix} \mathbf{U}_{\bar{k}}^{(0)} & \mathbf{U}_{\bar{k}}^{(\bar{0})} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{U}}_k^{(0)} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{U}}_k^{(\bar{0})} \end{bmatrix} , \tag{69}
$$

where $\bar{\mathbf{U}}_k^{(0)}$ and $\bar{\mathbf{U}}_k^{(\bar{0})}$ are $N_{\mathrm{b},k} \times N_{\mathrm{b},k}$ and $(N_{\mathrm{r}} - N_{\mathrm{b},k}) \times (N_{\mathrm{r}} - N_{\mathrm{b},k})$ unitary matrices, respectively. As a consequence, we obtain

$$
\mathbf{U}_k^H\mathbf{U}_{\bar{k}}\mathbf{\Lambda}_{\bar{k}}\mathbf{U}_{\bar{k}}^H\mathbf{U}_k = \mathbf{U}_k^H \begin{bmatrix} \mathbf{U}_{\bar{k}}^{(0)} & \mathbf{U}_{\bar{k}}^{(\bar{0})} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{\bar{k}}^{(\bar{0})} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\bar{k}}^{(0)} & \mathbf{U}_{\bar{k}}^{(\bar{0})} \end{bmatrix}^H \mathbf{U}_k
$$

$$
= \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{U}}_k^{(\bar{0})}\mathbf{\Lambda}_{\bar{k}}^{(\bar{0})}\bar{\mathbf{U}}_k^{(\bar{0})H} \end{bmatrix} . \tag{70}
$$

Using the identity $\mathbf{U}^{-1} = \mathbf{U}^H$ for a unitary matrix $\mathbf{U}$, we obtain

$$
\left(\mathbf{\Lambda}_k + \mathbf{U}_k^H\mathbf{U}_{\bar{k}}\mathbf{\Lambda}_{\bar{k}}\mathbf{U}_{\bar{k}}^H\mathbf{U}_k \right)^{-1} = \begin{bmatrix} \mathbf{\Lambda}_k^{(\bar{0})-1} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{U}}_k^{(\bar{0})}\mathbf{\Lambda}_{\bar{k}}^{(\bar{0})-1}\bar{\mathbf{U}}_k^{(\bar{0})H} \end{bmatrix} . \tag{71}
$$

Substituting (71) into (68), we obtain

$$
\mathbf{B}_k^H\mathbf{H}_k^H \left(\sum_{j=1}^{K} \mathbf{H}_j\mathbf{B}_j\mathbf{B}_j^H\mathbf{H}_j^H \right)^{-1} \mathbf{H}_k\mathbf{B}_k
$$

$$
= \begin{bmatrix} \mathbf{\Lambda}_k^{(\bar{0})\frac{1}{2}H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_k^{(\bar{0})-1} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{U}}_k^{(\bar{0})}\mathbf{\Lambda}_{\bar{k}}^{(\bar{0})-1}\bar{\mathbf{U}}_k^{(\bar{0})H} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_k^{(\bar{0})\frac{1}{2}} \\ \mathbf{0} \end{bmatrix}
$$

$$
= \mathbf{\Lambda}_k^{(\bar{0})\frac{1}{2}H}\mathbf{\Lambda}_k^{(\bar{0})-1}\mathbf{\Lambda}_k^{(\bar{0})\frac{1}{2}} = \mathbf{I}_{N_{\mathrm{b},k}} . \tag{72}
$$

Thus for high first-hop SNR, $\mathbf{B}_k^H\mathbf{H}_k^H\mathbf{\Psi}^{-1}\mathbf{H}_k\mathbf{B}_k$ can be approximated as $\mathbf{I}_{N_{\mathrm{b},k}}$. □

Authors' contributions
MRAK formulated the optimization problems, designed the proposed solutions, performed numerical simulations, and prepared the initial draft as well as the revision. K-KW managed funding for the research, modified solution pattern, verified mathematical derivations, checked and analyzed the simulation results, and improved the writeup. Both authors read and approved the final manuscript.

Authors' information
Muhammad R. A. Khandaker received the B.Sc. degree (Hons.) in computer science and engineering from Jahangirnagar University, Dhaka, Bangladesh, in 2006, the M.Sc. degree in telecommunications engineering from East West University, Dhaka, in 2007, and the Ph.D. degree in electrical and computer engineering from Curtin University, Australia, in 2013. He was a Junior Hardware Design Engineer with Visual Magic Corporation Ltd., in 2005. He joined the Department of Computer Science and Engineering, IBAIS University, Dhaka, in 2006, as a Lecturer. In 2007, he joined the Department of Information and Communication Technology, Mawlana Bhasani Science and Technology University, as a Lecturer. He joined the Institute of Information Technology, Jahangirnagar University, Dhaka, in 2008, as a Lecturer. Since 2013, he has been a Postdoctoral Research Associate with the Department of Electronic and Electrical Engineering, University College London, UK. He received the Curtin International Postgraduate Research Scholarship for his Ph.D. studies in 2009. He also received the Best Paper Award at the 16th Asia-Pacific Conference on Communications, Auckland, New Zealand, 2010. Kai-Kit Wong received the B.Eng, M.Phil., and Ph.D. degrees from the Hong Kong University of Science and Technology, Hong Kong, in 1996, 1998, and 2001, respectively, all in electrical and electronic engineering. He is currently a Professor of Wireless Communications with the Department of Electronic and Electrical Engineer- ing, University College London, U.K. Prior to this, he took up faculty and visiting positions at the University of Hong Kong, Lucent Technologies, Bell-Labs, Holmdel, NJ, U.S., the Smart Antennas Research Group of Stanford University, and the Department of Engineering, the University of Hull, U.K. He is a Fellow of IET. He serves on the Editorial Board of the IEEE WIRELESS COMMUNICATIONS LETTERS, the IEEE COMMUNICATIONS LETTERS, and the IEEE ComSoc/KICS Journal of Communications and Networks. He also served as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2005 to 2011 and the IEEE SIGNAL PROCESSING LETTERS from 2009 to 2012.

Competing interests
The authors declare that they have no competing interests.

References
1. D Tse, P Viswanath, *Fundamentals of Wireless Communication*. (Cambridge University Press, Cambridge, 2005)
2. F Rashid-Farrokhi, L Tassiulas, KJR Liu, Joint optimal power control and beamforming in wireless networks using antenna arrays. IEEE Trans. Commun. **46**, 1313–24 (1998)
3. J-H Chang, L Tassiulas, F Rashid-Farrokhi, Joint transmitter receiver diversity for efficient space division multiaccess. IEEE Trans. Wireless Commun. **1**, 16–27 (2002)
4. S Fazeli-Dehkordy, S Shahbazpanahi, S Gazor, Multiple peer-to-peer communications using a network of relays. IEEE Trans. Signal Process. **57**, 3053–62 (2009)
5. BK Chalise, L Vandendorpe, Optimization of MIMO relays for multipoint-to-multipoint communications: Nonrobust and robust designs. IEEE Trans. Signal Process. **58**, 6355–68 (2010)
6. MRA Khandaker, Y Rong, Joint transceiver optimization for multiuser MIMO relay communication systems. IEEE Trans. Signal Process. **60**, 5977–86 (2012)

7.   MRA Khandaker, Y Rong, Interference MIMO relay channel: Joint power control and transceiver-relay beamforming. IEEE Trans. Signal Process. **60**, 6509–18 (2012)

8.   KX Nguyen, Y Rong, in *Proc. Int. Symposium Inf. Theory Its Applications*. Joint Source and Relay Matrices Optimization for Interference MIMO Relay Systems (IEEE, Melbourne, 2014)

9.   KX Nguyen, Y Rong, S Nordholm, MMSE-based Joint Source and Relay Optimization for Interference MIMO Relay Systems. EURASIP J. Wireless Commun. Netw. **73** (2015)

10. KX Nguyen, Y Rong, S Nordholm, MMSE-based transceiver design algorithms for interference MIMO relay systems. IEEE Trans. Wireless Commun. **14**, 6414–6424 (2015)

11. KX Nguyen, Y Rong, S Nordholm, Simplified MMSE precoding design in interference two-way MIMO relay systems. IEEE Signal Process. Lett. **23**, 262–266 (2016)

12. B Rankov, A Wittneben, in *Proc. IEEE ISIT*. Achievable Rate Regions for the Two-way Relay Channel (IEEE, Seattle, 2006), pp. 1668–1672

13. K-J Lee, H Sung, E Park, I Lee, Joint optimization for one and two-way MIMO AF multiple-relay systems. IEEE Trans. Wireless Commun. **9**, 3671–3681 (2010)

14. T Cui, F Gao, T Ho, A Nallanathan, in *Proc. IEEE ICC*. Distributed Space-time Coding for Two-way Wireless Relay Networks (IEEE, Beijing, 2008), pp. 3888–3892

15. S Avestimehr, A Khajehnejad, A Sezgin, B Hassibi, in *Proc. IEEE ISIT*. Capacity Region of the Deterministic Multi-pair Bi-directional Relay Network (IEEE, 2009)

16. K Lee, N Lee, I Lee, Achievable degrees of freedom on MIMO two-way relay interference channels. IEEE Trans. Wireless Commun. **12**, 1472–80 (2013)

17. AC Cirik, R Wang, Y Hua, M Latva-aho, Weighted sum-rate maximization for full-duplex MIMO interference channels. IEEE Trans. Commn. **63**, 801–15 (2015)

18. C Song, K-J Lee, I Lee, MMSE based transceiver designs in closed-loop non-regenerative MIMO relaying systems. IEEE Trans. Wireless Commun. **9**, 2310–9 (2010)

19. Y Huang, L Yang, M Bengtsson, B Ottersten, A limited feedback joint precoding for amplify-and-forward relaying. IEEE Trans. S. **58**, 1347–57 (2010)

20. MRA Khandaker, Y Rong, Precoding design for MIMO relay multicasting. IEEE Trans. Wireless Commun. **12**, 3544–55 (2013)

21. MRA Khandaker, Y Rong, Transceiver optimization for multi-hop MIMO relay multicasting from multiple sources. IEEE Trans. Wireless Commun. **13**, 5162–72 (2014)

22. RA Horn, CR Johnson, *Matrix Analysis*. (Cambridge Univ. Press, Cambridge, 1985)

23. M Grant, S Boyd, CVX: Matlab software for disciplined convex programming (web page and software) (2010). http://cvxr.com/cvx

24. Y Nesterov, A Nemirovski, *Interior Point Polynomial Algorithms in Convex Programming*. (Philadelphia, SIAM, 1994)

25. Y Rong, Simplified algorithms for optimizing multiuser multi-hop MIMO relay systems. IEEE Trans. Commun. **59**, 2896–2904 (2011)