Research Article

Joint Linear Processing for an Amplify-and-Forward MIMO Relay Channel with Imperfect Channel State Information

Batu K. Chalise[1] and Luc Vandendorpe[2]

[1] *Center for Advanced Communications, Villanova University, Villanova, PA 19085, USA*

[2] *Communication and Remote Sensing Laboratory, Universitè catholique de Louvain, Place du Levant, 2, 1348 Louvain la Neuve, Belgium*

Correspondence should be addressed to Batu K. Chalise, batu.chalise@villanova.edu

The problem of jointly optimizing the source precoder, relay transceiver, and destination equalizer has been considered in this paper for a multiple-input-multiple-output (MIMO) amplify-and-forward (AF) relay channel, where the channel estimates of all links are assumed to be imperfect. The considered joint optimization problem is nonconvex and does not offer closed-form solutions. However, it has been shown that the optimization of one variable when others are fixed is a convex optimization problem which can be efficiently solved using interior-point algorithms. In this context, an iterative technique with the guaranteed convergence has been proposed for the AF MIMO relay channel that includes the direct link. It has been also shown that, for the double-hop relay case without the receive-side antenna correlations in each hop, the global optimality can be confirmed since the structures of the source precoder, relay transceiver, and destination equalizer have closed forms and the remaining joint power allocation can be solved using Geometric Programming (GP) technique under high signal-to-noise ratio (SNR) approximation. In the latter case, the performance of the iterative technique and the GP method has been compared with simulations to ensure that the iterative approach gives reasonably good solutions with an acceptable complexity. Moreover, simulation results verify the robustness of the proposed design when compared to the nonrobust design that assumes estimated channels as actual channels.

1. Introduction

The application of relays for cooperative communications has received a lot of interest in recent years. It is well known that the channel impairments such as shadowing, multipath fading, distance-dependent path losses, and interference often degrade the link quality between the source and destination in a wireless network. If the link quality degrades severely, relays can be employed between the source and destination nodes for assisting the transmission of data from the source to destination [1]. In the literature, various types of cooperative communications such as amplify-and-forward (AF), decode-and-forward [1], coded-cooperation [2], and compress-and-forward [3] have been presented. In [4], the outage and ergodic capacities have been analyzed for a three-node network where one of the nodes relays the messages of another node towards the third one. Among several cooperation schemes [1–4], the AF scheme is more

attractive due to its simplicity since the relay simply forwards the signal and does not decode it. Recently, space-time coding strategies have been developed for relay networks [5]. In [6], the authors study distributed beamforming for a cooperative network which consists of a transmitter, a receiver, and an arbitrary number of relay nodes. The common things among aforementioned works are that the transmitter, receiver, and the relays are all single-antenna nodes and the channel state information (CSI) (either instantaneous or second-order statistics of the channel) is assumed to be error-free.

The performance of cooperative communications can be further enhanced by employing multiple-input-multiple-output (MIMO) relays [7]. The optimal designs of AF MIMO relays have been investigated in [8, 9] for point-to-point and in [10, 11] for point-to-multipoint communications assuming that the available CSI is perfect. The robust design of MIMO relay for multipoint-to-multipoint

communications has been solved in [12], where the sources and destinations are single antenna nodes. The optimal design of multiple AF MIMO relays in a point-to-point communication scenario has been considered in [13, 14] to minimize the mean-square error (MSE) and satisfy the quality of service (QoS) requirements. These works also assume perfect knowledge of CSI. Recently, the joint robust design of AF MIMO relay and destination equalizer has been investigated in [15] for a double-hop (without direct link) MIMO relay channel. To the best source of our knowledge, the joint optimization of the source precoder, MIMO relay, and the destination equalizer has not been considered in the literature for the case where the CSI is imperfect and the direct link is included. Although the path attenuation for the direct link is much larger than that for the link via relay, due to the fading of the wireless channels, there can be still a significant number of instantaneous channels during which the direct link is better than the relay link. As a result, we consider the direct link in our analysis and exploit the benefit provided by the relay channel in terms of diversity. Moreover, in practice, channel estimation is required to obtain the CSI, where the estimation errors are inevitable due to various factors such as the limited length of the training sequences and the time-varying nature of wireless channels. The performance degradation due to such estimation errors can be mitigated by using robust designs that take into account the possible estimation errors. As a result, robust methods are highly desired for practical applications. The robust techniques can be divided mainly into worst-case and stochastic approaches [16]. The worst-case approach [17, 18] considers that the errors belong to a predefined uncertainty region, where the objective is to optimize the worst system performance for any error in this region. The stochastic approach guarantees a certain system performance averaged over channel realizations [19]. The latter approach has been used in [20] to minimize the power of the transmit beamformer while satisfying the QoS requirements for all users. In the sequel, we use stochastic approach for the robust design.

In this paper, we deal with the joint robust design of source precoder, relay transceiver, and destination equalizer for an AF MIMO relay system where the CSI is considered to be imperfect at all nodes. A stochastic approach is employed in which the objective is to minimize the average sum mean-square error (If the channel estimation is perfect at the receiver, the minimum mean-square error (MMSE) matrix $\mathbf{X}$ can be related to the rate using the relation $r = -\log \det(\mathbf{X})$. However, if the receiver does not have perfect estimation of the channel, the relation between the rate and MMSE matrix is not straightforward. Consequently, for our current system model where both the estimates of *source-relay* and *relay-destination* channels are imperfect, deriving rate expression and solving the optimization problem based on that expression are still an open issue.) under the source and relay power constraints. The considered joint optimization is nonconvex and also does not lead to closed-form solutions. However, it has been shown that the optimization of one parameter when others are fixed is a quadratic convex-optimization problem that can be easily solved within the framework of

convex optimization techniques. We first propose an iterative approach both for the MIMO relay channels with and without the direct link. Although the iterative method guarantees fast convergence for the case with the direct link, the global optimality cannot be proven since the joint optimization problem is nonconvex. As a result, in the second part of this paper, we limit the joint optimization problem for the case without the direct link in which the *source-relay* and *relay-destination* MIMO channels have only transmit-side antenna correlations. In the latter case, it is shown that structures of the optimal source precoder and relay transceiver have closed forms, where the remaining joint power allocation problem can be approximately formulated into a Geometric Programming (GP) problem. With the help of computer simulations, we compare the solutions of the iterative technique and the GP approach under high signal-to-noise ratio (SNR) approximation for the case without the direct link. This comparison is helpful to conclude that the iterative approach gives reasonably good solutions with an acceptable complexity.

The remainder of this paper is organized as follows. The system model for MIMO relay channel is presented in Section 2. In Section 3, the iterative approach is described for jointly optimizing the source precoder, relay transceiver, and destination equalizer for the MIMO relay channel with the direct link. The closed-form solutions and the approximate GP problem formulation are provided in Section 4 for the MIMO relay channel without the direct link where single-side antenna correlations have been considered for *source-relay* and *relay-destination* channels. In Section 5, simulation results are presented to show the performance of the proposed robust and nonrobust methods, and in Section 6, conclusions are drawn.

Notations . Upper (lower) bold face letters will be used for matrices (vectors); $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $\mathrm{E}\{\cdot\}$, $\mathbf{I}_n$, and $\|\cdot\|$ denote conjugate, transpose, Hermitian transpose, mathematical expectation, $n \times n$ identity matrix, and Frobenius norm, respectively. $\mathrm{tr}(\cdot)$, $\mathrm{vec}(\cdot)$, $\mathcal{C}^{M \times M}$, $\otimes$ denote the matrix trace operator, vectorization operator, space of $M \times M$ matrices with complex entries, and the Kronecker product, respectively.

2. System Model

We consider a cooperative communication system that consists of a source, a relay, and a destination which are all multiantenna nodes. The block diagram is shown in Figure 1. Notice that the direct link between the source and destination is taken into account, so that the diversity order of the cooperative system can be maintained. The source has M antennas, the relay has N_{R} receiving antennas and N_{T} transmitting antennas, and the destination has N_{D} antennas. The relay protocol consists of two timeslots. In the first timeslot, the source sends a symbol vector to the destination and relay. The relay linearly processes the source symbol vector and sends it to the destination in the second timeslot. The source remains idle during the second timeslot. At the
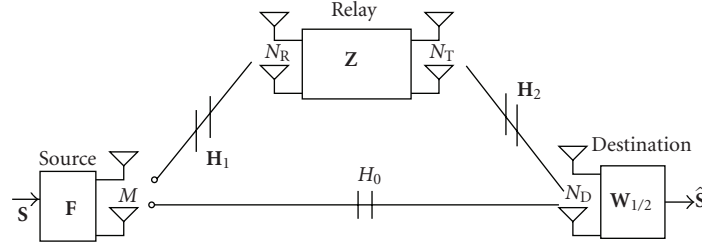
FIGURE 1: Cooperative MIMO relay channel.

end of two time slots, the destination linearly combines the symbol vectors received from the source and relay [1]. It is assumed that the estimates of the *source-relay*, *relay-destination*, and *source-destination* channels are available instead of their exact knowledge. The MIMO channels are considered to be spatially correlated block-fading frequency-flat Rayleigh channels. The signal received by the relay is given by

$$\mathbf{y}_\mathrm{r} = \mathbf{H}_1 \mathbf{F}\mathbf{s} + \mathbf{n}_\mathrm{r}, \tag{1}$$

where $\mathbf{s} \in \mathcal{C}^{N_S \times 1}$ is the complex source signal of length N_S, $\mathbf{F} \in \mathcal{C}^{M \times N_S}$ is the precoder that maps the $N_S \times 1$ symbol vector into M source antennas, $\mathbf{H}_1 \in \mathcal{C}^{N_R \times M}$ is the MIMO channel between the source and relay, and $\mathbf{n}_\mathrm{r} \in \mathcal{C}^{N_R \times 1}$ is the additive Gaussian noise vector at the relay. We also assume that the elements of $\mathbf{s}$ are statistically independent with the zero-mean and unit variance, that is, $\mathrm{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_{N_S}$. The precoder $\mathbf{F}$ at the source operates under the power constraint $P_S = \mathrm{tr}(\mathbf{F}\mathbf{F}^H) \leq P_S^{\max}$, where $P_S^{\max}$ is the maximum power of the source. We consider $\mathbf{n}_\mathrm{r} \sim \mathcal{N}_C(\mathbf{0}, \sigma_\mathrm{r}^2 \mathbf{I}_{N_R})$, that is, the entries of $\mathbf{n}_\mathrm{r}$ are zero-mean circularly symmetric complex Gaussian (ZMCSCG) with the variance σ_r^2. In order to ensure that the symbol $\mathbf{s}$ can be recovered at the destination, it is assumed that N_D, N_R, and N_T are greater than or equal to N_S. The signal received by the destination in first timeslot can be expressed as follows:

$$\mathbf{y}_{\mathrm{d},1} = \mathbf{H}_0 \mathbf{F}\mathbf{s} + \mathbf{n}_{\mathrm{d},1}, \tag{2}$$

where $\mathbf{H}_0 \in \mathcal{C}^{N_D \times M}$ is the MIMO channel between the source and destination and $\mathbf{n}_{\mathrm{d},1} \in \mathcal{C}^{N_D \times 1}$ is the additive Gaussian noise vector at the destination and follows $\mathbf{n}_{\mathrm{d},1} \sim \mathcal{N}_C(\mathbf{0}, \sigma_\mathrm{d}^2 \mathbf{I}_{N_D})$. The MIMO relay processes the signal $\mathbf{y}_\mathrm{r}$ using the linear operator $\mathbf{Z} \in \mathcal{C}^{N_T \times N_R}$ and forwards the following signal to the destination in the second timeslot:

$$\mathbf{y}_\mathrm{o} = \mathbf{Z}\mathbf{H}_1 \mathbf{F}\mathbf{s} + \mathbf{Z}\mathbf{n}_\mathrm{r}, \tag{3}$$

where the relay transceiver $\mathbf{Z}$ operates under the power constraint $P_R = \mathrm{E}\{\mathbf{y}_\mathrm{o}^H \mathbf{y}_\mathrm{o}\} \leq P_R^{\max}$ with total relay power of $P_R^{\max}$. The signal received by the destination in the second timeslot is

$$\mathbf{y}_{\mathrm{d},2} = \mathbf{H}_2 \mathbf{Z}\mathbf{H}_1 \mathbf{F}\mathbf{s} + \mathbf{H}_2 \mathbf{Z}\mathbf{n}_\mathrm{r} + \mathbf{n}_{\mathrm{d},2}, \tag{4}$$

where $\mathbf{H}_2 \in \mathcal{C}^{N_D \times N_T}$ is the MIMO channel between the relay and destination and $\mathbf{n}_{\mathrm{d},2}$ is the additive noise described

as $\mathbf{n}_{\mathrm{d},2} \sim \mathcal{N}_C(\mathbf{0}, \sigma_\mathrm{d}^2 \mathbf{I}_{N_D})$. The double-sided spatially correlated *source-relay*, *relay-destination* and *source-destination* are modelled according to Kronecker model as follows:

$$\mathbf{H}_1 = \mathbf{\Sigma}_1^{1/2} \mathbf{H}_1^\mathrm{w} \mathbf{\Psi}_1^{1/2},$$

$$\mathbf{H}_2 = \mathbf{\Sigma}_2^{1/2} \mathbf{H}_2^\mathrm{w} \mathbf{\Psi}_2^{1/2}, \tag{5}$$

$$\mathbf{H}_0 = \mathbf{\Sigma}_0^{1/2} \mathbf{H}_0^\mathrm{w} \mathbf{\Psi}_0^{1/2},$$

where $\mathbf{\Sigma}_1 \in \mathcal{C}^{N_R \times N_R}$, $\mathbf{\Sigma}_2 \in \mathcal{C}^{N_D \times N_D}$, and $\mathbf{\Sigma}_0 \in \mathcal{C}^{N_D \times N_D}$ are the receive-side spatial correlation matrices, and $\mathbf{\Psi}_1 \in \mathcal{C}^{M \times M}$, $\mathbf{\Psi}_2 \in \mathcal{C}^{N_T \times N_T}$, and $\mathbf{\Psi}_0 \in \mathcal{C}^{M \times M}$ are the transmit-side correlation matrices for the channels $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_0$, respectively. The elements of $\mathbf{H}_1^\mathrm{w}$, $\mathbf{H}_2^\mathrm{w}$, and $\mathbf{H}_0^\mathrm{w}$ are ZMCSCG random variables with the unit variance. Note that the transmit and receive spatial correlation matrices are positive semidefinite matrices and are a function of the antenna spacing, average direction of arrival/departure of the wavefronts at/from the transmitter/receiver, and the corresponding angular spread (see [21] and the references therein). The spatial correlation matrices represent the second-order statistics of the channels which vary slowly and can be precisely estimated. However, estimation of the fast fading parts $\mathbf{H}_1^\mathrm{w}$, $\mathbf{H}_2^\mathrm{w}$, and $\mathbf{H}_0^\mathrm{w}$ of the spatially correlated MIMO channels can lead to a significant amount of estimation error. For the linear minimum mean-square error (MMSE) estimation, we can write the following error model [22]:

$$\mathbf{H}_1^\mathrm{w} = \hat{\mathbf{H}}_1^\mathrm{w} + \mathbf{E}_1^\mathrm{w},$$

$$\mathbf{H}_2^\mathrm{w} = \hat{\mathbf{H}}_2^\mathrm{w} + \mathbf{E}_2^\mathrm{w}, \tag{6}$$

$$\mathbf{H}_0^\mathrm{w} = \hat{\mathbf{H}}_0^\mathrm{w} + \mathbf{E}_0^\mathrm{w},$$

where $\hat{\mathbf{H}}_1^\mathrm{w}$, $\hat{\mathbf{H}}_2^\mathrm{w}$, and $\hat{\mathbf{H}}_0^\mathrm{w}$ are the estimated CSI, and $\mathbf{E}_1^\mathrm{w}$, $\mathbf{E}_2^\mathrm{w}$, and $\mathbf{E}_0^\mathrm{w}$ are the corresponding channel estimation errors whose elements are ZMCCSG random variables with the variances $\sigma_{e,1}^2$, $\sigma_{e,2}^2$, and $\sigma_{e,0}^2$, respectively. Substituting (6) into (5), the error modelling for the actual channels $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_0$ can be simply given by

$$\mathbf{H}_1 = \mathbf{\Sigma}_1^{1/2} \hat{\mathbf{H}}_1^\mathrm{w} \mathbf{\Psi}_1^{1/2} + \mathbf{\Sigma}_1^{1/2} \mathbf{E}_1^\mathrm{w} \mathbf{\Psi}_1^{1/2} \triangleq \hat{\mathbf{H}}_1 + \mathbf{E}_1,$$

$$\mathbf{H}_2 = \mathbf{\Sigma}_2^{1/2} \hat{\mathbf{H}}_2^\mathrm{w} \mathbf{\Psi}_2^{1/2} + \mathbf{\Sigma}_2^{1/2} \mathbf{E}_2^\mathrm{w} \mathbf{\Psi}_2^{1/2} \triangleq \hat{\mathbf{H}}_2 + \mathbf{E}_2, \tag{7}$$

$$\mathbf{H}_0 = \mathbf{\Sigma}_0^{1/2} \hat{\mathbf{H}}_0^\mathrm{w} \mathbf{\Psi}_0^{1/2} + \mathbf{\Sigma}_0^{1/2} \mathbf{E}_0^\mathrm{w} \mathbf{\Psi}_0^{1/2} \triangleq \hat{\mathbf{H}}_0 + \mathbf{E}_0,$$

which shows that the errors $\mathbf{E}_1$, $\mathbf{E}_2$, and $\mathbf{E}_0$ are also double-sided correlated like the MIMO channels. The destination recovers the source signal $\mathbf{s}$ by linearly combining the signals $\mathbf{y}_{d,1}$ (2) and $\mathbf{y}_{d,2}$ (4) of two time slots as follows:

$$\hat{\mathbf{s}} = \mathbf{W}_1 \mathbf{y}_{d,1} + \mathbf{W}_2 \mathbf{y}_{d,2}, \tag{8}$$

where $\mathbf{W}_1, \mathbf{W}_2 \in \mathcal{C}^{N_S \times N_D}$ denote the linear operators for the signals received from the direct and relay links, respectively. The MSE between $\mathbf{s}$ and $\hat{\mathbf{s}}$ can be defined as follows:

$$\mathbf{M}(\mathbf{F}, \mathbf{Z}, \mathbf{W}_1, \mathbf{W}_2) = \mathrm{E}\left\{(\hat{\mathbf{s}} - \mathbf{s})(\hat{\mathbf{s}} - \mathbf{s})^H\right\} \tag{9}$$

where the MSE matrix depends on $\mathbf{F}$, $\mathbf{Z}$, $\mathbf{W}_1$ and $\mathbf{W}_2$ and the mathematical expectation is only taken with respect to noise and signal realizations. Considering that $\mathbf{n}_r$, $\mathbf{n}_{d,1}$, $\mathbf{n}_{d,2}$ and $\mathbf{s}$ are statistically independent and applying (2) and (4) into (8), we can write (9) as follows:

$$
\begin{aligned}
\mathbf{M}(\mathbf{F}, \mathbf{Z}, \mathbf{W}_1, \mathbf{W}_2) = \mathbf{W}_1 &\left[\underbrace{\mathbf{H}_0 \mathbf{F} \mathbf{F}^H \mathbf{H}_0^H + \sigma_{n_d}^2 \mathbf{I}_{N_D}}_{I} \right] \mathbf{W}_1^H \\
&- (\mathbf{W}_1 \mathbf{H}_0 + \mathbf{W}_2 \mathbf{H}_2 \mathbf{Z} \mathbf{H}_1)\mathbf{F} - (\mathbf{W}_1 \mathbf{H}_0 \mathbf{F})^H \\
&- (\mathbf{W}_2 \mathbf{H}_2 \mathbf{Z} \mathbf{H}_1 \mathbf{F})^H \\
&+ \left(\mathbf{W}_1 \mathbf{H}_0 \mathbf{F} \mathbf{F}^H \mathbf{H}_1^H \mathbf{Z}^H \mathbf{H}_2^H \mathbf{W}_2^H \right)^H \\
&+ \left(\mathbf{W}_1 \underbrace{\mathbf{H}_0 \mathbf{F} \mathbf{F}^H \mathbf{H}_1^H \mathbf{Z}^H \mathbf{H}_2^H}_{II} \mathbf{W}_2^H \right) \\
&+ \mathbf{W}_2 \left[\underbrace{\mathbf{H}_2 \mathbf{Z} \mathbf{H}_1 \mathbf{F} (\mathbf{H}_2 \mathbf{Z} \mathbf{H}_1 \mathbf{F})^H}_{III} \right. \\
&\left. + \sigma_{n_r}^2 \underbrace{\mathbf{H}_2 \mathbf{Z} (\mathbf{H}_2 \mathbf{Z})^H}_{IV} + \sigma_{n_d}^2 \mathbf{I}_{N_D} \right] \mathbf{W}_2^H + \mathbf{I}_{N_S}.
\end{aligned}
\tag{10}
$$

For the given channel estimates $\hat{\mathbf{H}}_1$, $\hat{\mathbf{H}}_2$, and $\hat{\mathbf{H}}_0$, the MSE matrix of (10) is random due to the random errors $\mathbf{E}_1$, $\mathbf{E}_2$, and $\mathbf{E}_0$. Since the exact errors are not known and only the covariance matrices of these errors are known, we need to derive the average MSE matrix. This can be done by averaging the MSE matrix (10) over the independent errors $\mathbf{E}_1$, $\mathbf{E}_2$, and $\mathbf{E}_0$. Hence, we can write

$$
\begin{aligned}
\mathrm{E}_{\mathbf{E}_0}\{I\} &= \hat{\mathbf{H}}_0 \mathbf{F} \mathbf{F}^H \hat{\mathbf{H}}_0^H + \mathrm{E}_{\mathbf{E}_0}\left\{\mathbf{E}_0 \mathbf{F} \mathbf{F}^H \mathbf{E}_0^H\right\} \\
&= \hat{\mathbf{H}}_0 \mathbf{F} \mathbf{F}^H \hat{\mathbf{H}}_0^H + \mathrm{tr}\left(\mathbf{F} \mathbf{F}^H \mathbf{\Psi}_0\right) \tilde{\mathbf{\Sigma}}_0,
\end{aligned}
\tag{11}
$$

where $\tilde{\mathbf{\Sigma}}_0 = \sigma_{e,0}^2 \mathbf{\Sigma}_0$, and we have applied (7) and used the facts that $\mathbf{E}_0$ is zero-mean and $\mathrm{E}_{\mathbf{X}}\{\mathbf{X} \mathbf{A} \mathbf{X}^H\} = \mathrm{tr}(\mathbf{A})\sigma_x^2 \mathbf{I}$ for $\mathbf{X} \sim$

$\mathcal{N}_C(0, \sigma_x^2 \mathbf{I})$. Similarly, since $\mathbf{E}_0$, $\mathbf{E}_1$, and $\mathbf{E}_2$ are independent, we can easily show that

$$
\begin{aligned}
&\mathrm{E}_{\mathbf{E}_0, \mathbf{E}_1, \mathbf{E}_2}\{II\} \\
&= \mathrm{E}_{\mathbf{E}_0, \mathbf{E}_1, \mathbf{E}_2}\left\{ \left(\hat{\mathbf{H}}_0 + \mathbf{E}_0\right) \mathbf{F} \mathbf{F}^H \left(\hat{\mathbf{H}}_1 + \mathbf{E}_1\right)^H \mathbf{Z}^H \left(\hat{\mathbf{H}}_2 + \mathbf{E}_2\right)^H \right\} \\
&= \hat{\mathbf{H}}_0 \mathbf{F} \mathbf{F}^H \hat{\mathbf{H}}_1^H \mathbf{Z}^H \hat{\mathbf{H}}_2^H.
\end{aligned}
\tag{12}
$$

Furthermore, we can write

$$\mathrm{E}_{\mathbf{E}_1, \mathbf{E}_2}\{III\} = \mathrm{E}_{\mathbf{E}_2}\left\{ \mathbf{H}_2 \mathbf{Z} \mathrm{E}_{\mathbf{E}_1}\left\{ \mathbf{H}_1 \mathbf{F} \mathbf{F}^H \mathbf{H}_1^H \right\} \mathbf{Z}^H \mathbf{H}_2^H \right\}, \tag{13}$$

where the inner expectation is

$$
\begin{aligned}
\mathrm{E}_{\mathbf{E}_1}\left\{ \mathbf{H}_1 \mathbf{F} \mathbf{F}^H \mathbf{H}_1^H \right\} &= \mathrm{E}_{\mathbf{E}_1}\left\{ \left(\hat{\mathbf{H}}_1 + \mathbf{E}_1\right)^H \mathbf{F} \mathbf{F}^H \left(\hat{\mathbf{H}}_1 + \mathbf{E}_1\right) \right\} \\
&= \underbrace{\hat{\mathbf{H}}_1 \mathbf{F} \mathbf{F}^H \hat{\mathbf{H}}_1^H + \mathrm{tr}\left(\mathbf{F} \mathbf{F}^H \mathbf{\Psi}_1\right) \tilde{\mathbf{\Sigma}}_1}_{\mathbf{A}}
\end{aligned}
\tag{14}
$$

and $\tilde{\mathbf{\Sigma}}_1 = \sigma_{e,1}^2 \mathbf{\Sigma}_1$. Substituting the result of (14) into (13), we have

$$
\begin{aligned}
\mathrm{E}_{\mathbf{E}_1, \mathbf{E}_2}\{III\} &= \mathrm{E}_{\mathbf{E}_2}\left\{ \mathbf{H}_2 \mathbf{Z} \mathbf{A} \mathbf{Z}^H \mathbf{H}_2^H \right\} \\
&= \hat{\mathbf{H}}_2 \mathbf{Z} \mathbf{A} \mathbf{Z}^H \hat{\mathbf{H}}_2^H + \mathrm{tr}\left(\mathbf{Z} \mathbf{A} \mathbf{Z}^H \mathbf{\Psi}_2\right) \tilde{\mathbf{\Sigma}}_2,
\end{aligned}
\tag{15}
$$

where $\tilde{\mathbf{\Sigma}}_2 = \sigma_{e,2}^2 \mathbf{\Sigma}_2$. Applying similar steps, we can also get

$$\mathrm{E}_{\mathbf{E}_2}\{IV\} = \hat{\mathbf{H}}_2 \mathbf{Z} \mathbf{Z}^H \hat{\mathbf{H}}_2^H + \mathrm{tr}\left(\mathbf{Z} \mathbf{Z}^H \mathbf{\Psi}_2\right) \tilde{\mathbf{\Sigma}}_2. \tag{16}$$

Using the results of (11) to (16), the average MSE matrix can be written as follows:

$$
\begin{aligned}
\overline{\mathbf{M}}(\mathbf{F}, \mathbf{Z}, \mathbf{W}_1, \mathbf{W}_2) &= \mathrm{E}_{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_0}\{\mathbf{M}(\mathbf{F}, \mathbf{Z}, \mathbf{W}_1, \mathbf{W}_2)\} \\
&= \mathbf{W}_1 \left[\underbrace{\hat{\mathbf{H}}_0 \mathbf{F} \mathbf{F}^H \hat{\mathbf{H}}_0^H + \mathrm{tr}\left(\mathbf{F} \mathbf{F}^H \mathbf{\Psi}_0\right) \tilde{\mathbf{\Sigma}}_0 + \sigma_{n_d}^2 \mathbf{I}_{N_D}}_{\mathbf{A}_m} \right] \mathbf{W}_1^H \\
&\quad + \mathbf{W}_1 \underbrace{\hat{\mathbf{H}}_0 \mathbf{F} \mathbf{F}^H \hat{\mathbf{H}}_1^H \mathbf{Z}^H \hat{\mathbf{H}}_2^H}_{\mathbf{B}_m} \mathbf{W}_2^H \\
&\quad + \left(\mathbf{W}_1 \hat{\mathbf{H}}_0 \mathbf{F} \mathbf{F}^H \hat{\mathbf{H}}_1^H \mathbf{Z}^H \hat{\mathbf{H}}_2^H \mathbf{W}_2^H \right)^H \\
&\quad + \mathbf{W}_2 \left[\underbrace{\hat{\mathbf{H}}_2 \mathbf{Z} \tilde{\mathbf{A}} \mathbf{Z}^H \hat{\mathbf{H}}_2^H + \mathrm{tr}\left(\mathbf{Z} \tilde{\mathbf{A}} \mathbf{Z}^H \mathbf{\Psi}_2\right) \tilde{\mathbf{\Sigma}}_2 + \sigma_{n_d}^2 \mathbf{I}_{N_D}}_{\mathbf{C}_m} \right] \mathbf{W}_2^H \\
&\quad - \left(\mathbf{W}_1 \hat{\mathbf{H}}_0 \mathbf{F} + \mathbf{W}_2 \underbrace{\hat{\mathbf{H}}_2 \mathbf{Z} \hat{\mathbf{H}}_1 \mathbf{F}}_{\mathbf{D}_m} \right) \\
&\quad - \left(\mathbf{W}_1 \hat{\mathbf{H}}_0 \mathbf{F} + \mathbf{W}_2 \hat{\mathbf{H}}_2 \mathbf{Z} \hat{\mathbf{H}}_1 \mathbf{F} \right)^H + \mathbf{I}_{N_S},
\end{aligned}
\tag{17}
$$

where $\widetilde{\mathbf{A}} = \mathbf{A} + \sigma_{n_r}^2 \mathbf{I}_{N_R}$. The instantaneous relay power can be obtained as follows:

$$P_R = \text{tr}\left(\text{E}\left\{\mathbf{y}_o \mathbf{y}_o^H\right\}\right) = \text{tr}\left(\mathbf{Z}\mathbf{H}_1\mathbf{F}\mathbf{F}^H\mathbf{H}_1^H\mathbf{Z}^H\right) + \sigma_{n_r}^2 \text{tr}\left(\mathbf{Z}\mathbf{Z}^H\right), \tag{18}$$

where expectation is taken w.r.t. noise and signal realizations. After including the estimation error $\mathbf{E}_1$ in (18), the relay power averaged over $\mathbf{E}_1$ can be expressed with the help of (14) as follows:

$$\overline{P}_R = \text{tr}\left(\mathbf{Z}\widetilde{\mathbf{A}}\mathbf{Z}^H\right). \tag{19}$$

3. Joint Optimization: Iterative Approach

The objective of joint optimization is to minimize the sum of the average MSE (17) under power constraints of the source and relay. This optimization problem can be expressed as follows:

$$\min_{\mathbf{F},\mathbf{Z},\mathbf{W}_1,\mathbf{W}_2} \overline{f}_{\text{mse}} = \text{tr}(\overline{\mathbf{M}}(\mathbf{F},\mathbf{Z},\mathbf{W}_1,\mathbf{W}_2)) \text{ s.t. } \text{tr}\left(\mathbf{F}\mathbf{F}^H\right) \le P_S^{\max},$$

$$\text{tr}\left(\mathbf{Z}\widetilde{\mathbf{A}}\mathbf{Z}^H\right) \le P_R^{\max}. \tag{20}$$

The constraints of the optimization problem (20) do not depend on $\mathbf{W}_1$ and $\mathbf{W}_2$. As a result, the optimal $\mathbf{W}_1$ and $\mathbf{W}_2$ can be easily obtained in terms of $\mathbf{F}$ and $\mathbf{Z}$. Unfortunately, after substituting such optimal $\mathbf{W}_1$ and $\mathbf{W}_2$ into the objective function of (20), the resulting objective function in terms of $\mathbf{F}$ and $\mathbf{Z}$ appears to be a nontractable nonconvex problem. This fact will be later shown in this section. The joint optimization problem (20) is a nonconvex problem and does not offer closed-form solutions. However, it can be easily observed that the considered problem is a convex problem over one optimization variable when others are fixed. Hence, we propose to solve this optimization problem using iterative technique, where each optimization variable is updated at a time considering others as fixed. The iterative algorithm may be implemented as follows. The destination estimates the *source-destination* and *relay-destination* channels and the relay estimates the *source-relay* channel, separately with the help of training sequence. The relay sends the estimated *source-relay* channel to the destination where the iterative algorithm is executed. The destination feedbacks optimally designed $\mathbf{F}$ and $\mathbf{Z}$ to the source and relay, respectively. The channel is considered to remain constant within a block but vary from one block to another, where the block consists of training signal and useful data (Notice that the adaptation of the source precoder and relay transceiver matrices in fast fading scenario can be impractical if the design is based on instantaneous channels [23]. Therefore, robust designs based on channel covariance information [20] can be appropriate for such a scenario.) .

Remark 1. It is worthwhile to mention here that the minimization of the sum of the source and relay powers under the MSE constraint can also be solved by using the iterative framework that we are proposing in the sequel. Moreover,

the quality of fairness approach such as minimizing the sum of the source and relay powers while fulfilling the SNR/MSE requirements of each symbol stream can also be handled by the proposed iterative method. For conciseness, the latter two methods are not considered in this paper.

After solving the first-order partial derivative of the objective function of (20) w.r.t to $\mathbf{W}_1$, we get

$$\mathbf{W}_1 = \left[\mathbf{F}^H\hat{\mathbf{H}}_0^H - \mathbf{W}_2\mathbf{B}_m^H\right]\mathbf{A}_m^{-1}. \tag{21}$$

Substituting (21) into the objective of (20), the latter can be expressed in terms of $\mathbf{W}_2$, $\mathbf{Z}$, and $\mathbf{F}$ as follows:

$$\begin{aligned} \text{tr}(\overline{\mathbf{M}}(\mathbf{F},\mathbf{Z},\mathbf{W}_2)) &= \text{tr}\Big(\mathbf{W}_2(\mathbf{C}_m - \mathbf{B}_m\mathbf{A}_m^{-1}\mathbf{B}_m)\mathbf{W}_2^H \\ &\quad - \mathbf{W}_2\mathbf{D}_m - \mathbf{D}_m^H\mathbf{W}_2^H\Big) + N_S \\ &\quad + \text{tr}\Big(\mathbf{W}_2\mathbf{B}_m^H\mathbf{A}_m^{-1}\hat{\mathbf{H}}_0\mathbf{F} + \mathbf{F}^H\hat{\mathbf{H}}_0^H\mathbf{A}_m^{-1} \\ &\quad \times \left(\mathbf{B}_m\mathbf{W}_2^H - \hat{\mathbf{H}}_0\mathbf{F}\right)\Big). \end{aligned} \tag{22}$$

Now, solving the derivative of (22) w.r.t to $\mathbf{W}_2$, we get the optimal $\mathbf{W}_2$ as follows:

$$\mathbf{W}_2 = \left(\mathbf{D}_m^H - \mathbf{F}^H\hat{\mathbf{H}}_0^H\mathbf{A}_m^{-1}\mathbf{B}_m\right)\left(\mathbf{C}_m - \mathbf{B}_m^H\mathbf{A}_m^{-1}\mathbf{B}_m\right)^{-1}. \tag{23}$$

The optimal $\mathbf{W}_1$ can be obtained by substituting $\mathbf{W}_2$ from (23) into (21). Using the results of (21) and (23) and then resubstituting $\mathbf{A}_m$, $\mathbf{B}_m$, and $\mathbf{C}_m$, (22) can be written in terms of $\mathbf{F}$ and $\mathbf{Z}$ as follows:

$$\begin{aligned} \text{tr}(\overline{\mathbf{M}}(\mathbf{F},\mathbf{Z})) &= \text{tr}\left(\mathbf{G}\right) - \text{tr}\Big(\hat{\mathbf{H}}_2\mathbf{Z}\hat{\mathbf{H}}_1\mathbf{F}\mathbf{G}\mathbf{G}\left(\hat{\mathbf{H}}_2\mathbf{Z}\hat{\mathbf{H}}_1\mathbf{F}\right)^H \\ &\quad \times \left[\hat{\mathbf{H}}_2\mathbf{Z}\hat{\mathbf{H}}_1\mathbf{F}\mathbf{G}\left(\hat{\mathbf{H}}_2\mathbf{Z}\hat{\mathbf{H}}_1\mathbf{F}\right)^H + \mathbf{\Gamma}\right]^{-1}\Big), \end{aligned} \tag{24}$$

where

$$\mathbf{G} = \mathbf{I}_{N_S} - \mathbf{F}^H\hat{\mathbf{H}}_0^H\left[\hat{\mathbf{H}}_0\mathbf{F}\mathbf{F}^H\hat{\mathbf{H}}_0^H + \underbrace{\text{tr}\left(\mathbf{F}\mathbf{F}^H\mathbf{\Psi}_0\right)\widetilde{\mathbf{\Sigma}}_0 + \sigma_{n_d}^2\mathbf{I}_{N_D}}_{\mathbf{\Sigma}_N}\right]^{-1}\hat{\mathbf{H}}_0\mathbf{F}$$

$$= \left[\underbrace{\mathbf{F}^H\hat{\mathbf{H}}_0^H\mathbf{\Sigma}_N^{-1}\hat{\mathbf{H}}_0\mathbf{F}}_{\mathbf{\Gamma}_t} + \mathbf{I}_{N_S}\right]^{-1},$$

$$\begin{aligned} \mathbf{\Gamma} &= \hat{\mathbf{H}}_2\mathbf{Z}\left(\widetilde{\mathbf{\Sigma}}_1 \text{tr}\left(\mathbf{F}\mathbf{F}^H\mathbf{\Psi}_1\right) + \sigma_{n_r}^2\mathbf{I}_{N_D}\right)\mathbf{Z}^H\hat{\mathbf{H}}_2^H \\ &\quad + \text{tr}\left(\mathbf{Z}\widetilde{\mathbf{A}}\mathbf{Z}^H\mathbf{\Psi}_2\right)\widetilde{\mathbf{\Sigma}}_2 + \sigma_{n_d}^2\mathbf{I}_{N_D}. \end{aligned} \tag{25}$$

It is interesting to observe that $\mathbf{G}$ is the MMSE matrix of the direct link, where the sum MMSE is simply given by $\overline{f}_{\text{mmse,DL}} = \text{tr}(\mathbf{G})$. The second equality for $\mathbf{G}$ in (25) is obtained by using the fact that $\mathbf{X}^H(\mathbf{X}\mathbf{X}^H + \mathbf{I})^{-1}\mathbf{X} =$

$\mathbf{I} - (\mathbf{X}^H \mathbf{X} + \mathbf{I})^{-1}$. Applying the same fact and after some manipulations, we get

$$
\begin{aligned}
&\mathrm{tr}\left(\overline{\mathbf{M}}(\mathbf{F}, \mathbf{Z})\right) \\
&= \mathrm{tr}\left(\mathbf{G}\left[\mathbf{I}_{N_\mathrm{S}} + \mathbf{G}^{H/2}\underbrace{\mathbf{F}^H \hat{\mathbf{H}}_1^H \mathbf{Z}^H \hat{\mathbf{H}}_2^H \mathbf{\Gamma}^{-1} \hat{\mathbf{H}}_2 \mathbf{Z} \hat{\mathbf{H}}_1 \mathbf{F}}_{\mathbf{Y}} \mathbf{G}^{1/2}\right]^{-1}\right) \\
&= \mathrm{tr}\left(\left[\mathbf{G}^{-1} + \mathbf{Y}\right]^{-1}\right),
\end{aligned}
$$
$$(26)$$

where second equality is obtained after simple steps using the fact that $\mathbf{G}$ is a positive definite square matrix. Notice that $\mathbf{Y}$ is only related to the MMSE of the double-hop channel, where the sum MMSE is $\overline{f}_{\mathrm{mmse,DH}} = \mathrm{tr}\left(\mathbf{I}_{N_\mathrm{S}} + \mathbf{Y}\right)^{-1}$. With these observations, we can formulate the following lemma:

Lemma 1. *The sum MMSE of the MIMO relay system with the direct link is upper bounded by the sum MMSE of the direct link and source-relay-destination link.*

Proof. This Lemma can be easily proven by using the properties of the positive (semi) definite matrices. Since $\mathbf{G}^{-1}$ is positive definite and $\mathbf{Y}$ and $\mathbf{\Gamma}_\mathrm{t}$ are positive semidefinite, we can show that

$$
\begin{aligned}
\mathbf{G}^{-1} + \mathbf{Y} \succeq \mathbf{G}^{-1} &\longrightarrow \left(\mathbf{G}^{-1} + \mathbf{Y}\right)^{-1} \preceq \mathbf{G} \longrightarrow \mathrm{tr}\left(\left(\mathbf{G}^{-1} + \mathbf{Y}\right)^{-1}\right) \\
&\leq \mathrm{tr}(\mathbf{G}) \triangleq \overline{f}_{\mathrm{mmse,DL}},
\end{aligned}
$$
$$(27)$$

$$
\begin{aligned}
\mathbf{G}^{-1} + \mathbf{Y} = \mathbf{\Gamma}_\mathrm{t} + \mathbf{I}_{N_\mathrm{S}} + \mathbf{Y} \succeq \mathbf{I}_{N_\mathrm{S}} + \mathbf{Y} &\longrightarrow \left(\mathbf{\Gamma}_\mathrm{t} + \left(\mathbf{I}_{N_\mathrm{S}} + \mathbf{Y}\right)\right)^{-1} \\
&\preceq \left(\mathbf{I}_{N_\mathrm{S}} + \mathbf{Y}\right)^{-1} \longrightarrow \\
\mathrm{tr}\left(\left(\mathbf{G}^{-1} + \mathbf{Y}\right)^{-1}\right) &\leq \mathrm{tr}\left(\left(\mathbf{I}_{N_\mathrm{S}} + \mathbf{Y}\right)^{-1}\right) \triangleq \overline{f}_{\mathrm{mmse,DH}}.
\end{aligned}
$$
$$(28)$$

The results of (27) and (28) prove the Lemma. $\square$

It can be seen that the minimization of (26) under source and relay power constraints is a nontractable problem. We have noticed that even in the case of the nonrobust design, such an objective is difficult to handle. This difficulty has motivated us to use the iterative optimization based on the MSE function (17) for which $\mathbf{W}_1$ and $\mathbf{W}_2$ have been already determined in terms of $\mathbf{Z}$ and $\mathbf{F}$ (see (21) and (23)). In the following, we show the optimizations over $\mathbf{Z}$ and $\mathbf{F}$ when other variables are fixed.

(1) Optimization over $\mathbf{Z}$. With some straightforward manipulations of (17) and using the fact that $\mathrm{tr}(\mathbf{X}\mathbf{X}^H) = \|\mathbf{X}\|^2$,

the average sum MSE (objective function of (20)) can be alternatively expressed as follows:

$$
\begin{aligned}
\overline{f}_{\mathrm{mse}} = &\left\|\left(\mathbf{W}_1 \hat{\mathbf{H}}_0 + \mathbf{W}_2 \hat{\mathbf{H}}_2 \mathbf{Z} \hat{\mathbf{H}}_1\right)\mathbf{F} - \mathbf{I}_{N_\mathrm{S}}\right\|^2 \\
&+ \mathrm{tr}\left(\mathbf{F}\mathbf{F}^H \mathbf{\Psi}_1\right)\mathrm{tr}\left(\mathbf{B}\mathbf{Z}\tilde{\mathbf{\Sigma}}_1 \mathbf{Z}^H\right) + \sigma_{n_\mathrm{r}}^2 \mathrm{tr}\left(\mathbf{Z}^H \mathbf{B} \mathbf{Z}\right) \\
&+ \mathrm{tr}\left(\mathbf{Z}\tilde{\mathbf{A}}\mathbf{Z}^H \mathbf{\Psi}_2\right)\mathrm{tr}\left(\mathbf{W}_2 \tilde{\mathbf{\Sigma}}_2 \mathbf{W}_2^H\right) \\
&+ \sigma_{n_\mathrm{d}}^2 \mathrm{tr}\left(\mathbf{W}_1 \mathbf{W}_1^H + \mathbf{W}_2 \mathbf{W}_2^H\right) \\
&+ \mathrm{tr}\left(\mathbf{W}_1 \tilde{\mathbf{\Sigma}}_0 \mathbf{W}_1^H\right)\mathrm{tr}\left(\mathbf{F}\mathbf{F}^H \mathbf{\Psi}_0\right),
\end{aligned}
$$
$$(29)$$

where $\mathbf{B} = \hat{\mathbf{H}}_2^H \mathbf{W}_2^H \mathbf{W}_2 \hat{\mathbf{H}}_2$. Applying the following results [24]:

$$
\begin{aligned}
\mathrm{vec}(\mathbf{X}\mathbf{W}\mathbf{Y}) &= \left(\mathbf{Y}^T \bigotimes \mathbf{X}\right)\mathrm{vec}(\mathbf{W}), \\
\mathrm{tr}\left(\mathbf{X}^H \mathbf{Y}\mathbf{X}\mathbf{W}\right) &= \mathrm{vec}(\mathbf{X})^H \left(\mathbf{W}^T \bigotimes \mathbf{Y}\right)\mathrm{vec}(\mathbf{X})
\end{aligned}
$$
$$(30)$$

and denoting $\mathbf{z}_\mathrm{L} \triangleq \mathrm{vec}(\mathbf{Z}) \in \mathcal{C}^{N_\mathrm{T} N_\mathrm{R} \times 1}$, we can write

$$
\begin{aligned}
\overline{f}_{\mathrm{mse}} = &\left\|\mathrm{vec}\left(\mathbf{W}_1 \hat{\mathbf{H}}_0 \mathbf{F}\right) + \left(\left(\hat{\mathbf{H}}_1 \mathbf{F}\right)^T \bigotimes \mathbf{W}_2 \hat{\mathbf{H}}_2\right)\mathbf{z}_\mathrm{L} - \mathrm{vec}(\mathbf{I}_{N_\mathrm{S}})\right\|^2 \\
&+ \mathbf{z}_\mathrm{L}^H \mathbf{D}\mathbf{z}_\mathrm{L} + s_3 + s_4,
\end{aligned}
$$
$$(31)$$

where

$$
\begin{aligned}
\mathbf{D} &= s_1\left(\tilde{\mathbf{\Sigma}}_1^T \bigotimes \mathbf{B}\right) + \sigma_{n_\mathrm{r}}^2\left(\mathbf{I}_{N_\mathrm{R}} \bigotimes \mathbf{B}\right) + s_2\left(\tilde{\mathbf{A}}^T \bigotimes \mathbf{\Psi}_2\right), \\
s_1 &\triangleq \mathrm{tr}\left(\mathbf{F}\mathbf{F}^H \mathbf{\Psi}_1\right), \qquad s_0 \triangleq \mathrm{tr}\left(\mathbf{F}\mathbf{F}^H \mathbf{\Psi}_0\right), \\
s_2 &\triangleq \mathrm{tr}\left(\mathbf{W}_2 \tilde{\mathbf{\Sigma}}_2 \mathbf{W}_2^H\right), \qquad s_3 \triangleq \sigma_{n_\mathrm{d}}^2 \mathrm{tr}\left(\mathbf{W}_2 \mathbf{W}_2^H\right), \\
s_4 &\triangleq \sigma_{n_\mathrm{d}}^2 \mathrm{tr}\left(\mathbf{W}_1 \mathbf{W}_1^H\right) + \mathrm{tr}\left(\mathbf{W}_1 \tilde{\mathbf{\Sigma}}_0 \mathbf{W}_1^H\right)s_0.
\end{aligned}
$$
$$(32)$$

The optimization problem w.r.t. $\mathbf{Z}$ can be thus written as follows:

$$
\begin{aligned}
\mathcal{P}_1 &: \min_{\mathbf{z}_\mathrm{L}} \overline{f}_{\mathrm{mse}} \quad \text{s.t.} \\
\overline{P}_\mathrm{R} &= \left\|\left[\tilde{\mathbf{A}}^T \otimes \mathbf{I}_{N_\mathrm{T}}\right]^{1/2}\mathbf{z}_\mathrm{L}\right\|^2 \leq P_\mathrm{R}^{\max}.
\end{aligned}
$$
$$(33)$$

Noting that $\mathbf{z}_\mathrm{L}^H \mathbf{D}\mathbf{z}_\mathrm{L} = \|\mathbf{D}^{1/2}\mathbf{z}_\mathrm{L}\|^2$, the optimization problem (33) can be written as follows:

$$
\begin{aligned}
\mathcal{P}_1 &: \min_{\mathbf{z}_\mathrm{L}, t_1, t_2} t_1^2 + t_2^2 \quad \text{s.t.} \\
&\left\|\mathrm{vec}\left(\mathbf{W}_1 \hat{\mathbf{H}}_0 \mathbf{F}\right) + \left(\left(\hat{\mathbf{H}}_1 \mathbf{F}\right)^T \bigotimes \mathbf{W}_2 \hat{\mathbf{H}}_2\right)\right. \\
&\left. \times \mathbf{z}_\mathrm{L} - \mathrm{vec}(\mathbf{I}_{N_\mathrm{S}})\right\| \leq t_1, \qquad \left\|\mathbf{D}^{1/2}\mathbf{z}_\mathrm{L}\right\| \leq t_2, \\
&\left\|\left[\tilde{\mathbf{A}}^T \bigotimes \mathbf{I}_{N_\mathrm{T}}\right]^{1/2}\mathbf{z}_\mathrm{L}\right\| \leq \sqrt{P_\mathrm{R}^{\max}}.
\end{aligned}
$$
$$(34)$$

Using the notation $\mathbf{t} \triangleq [t_1, t_2]^T$, the fact that $t_1^2 + t_2^2 = \mathbf{t}^T \mathbf{t}$ and introducing an auxiliary variable $\tilde{t} \geq 0$, the problem (34) can be formulated as the following standard convex optimization problem:

$$\mathcal{P}_1 : \min_{\mathbf{z}_L, \tilde{t}, \mathbf{t}} \tilde{t} \text{ s. t.}$$

$$\begin{bmatrix} \mathbf{I}_2 & \mathbf{t} \\ \mathbf{t}^T & \tilde{t} \end{bmatrix} \succeq 0,$$

$$\left\| \operatorname{vec}\left(\mathbf{W}_1 \hat{\mathbf{H}}_0 \mathbf{F}\right) + \left(\left(\hat{\mathbf{H}}_1 \mathbf{F}\right)^T \otimes \mathbf{W}_2 \hat{\mathbf{H}}_2\right) \mathbf{z}_L \right. \quad (35)$$

$$\left. - \operatorname{vec}(\mathbf{I}_{N_S})\right\| \leq \mathbf{a}^T \mathbf{t},$$

$$\left\| \mathbf{D}^{1/2} \mathbf{z}_L \right\| \leq \mathbf{b}^T \mathbf{t},$$

$$\left\| \left[\tilde{\mathbf{A}}^T \otimes \mathbf{I}_{N_T} \right]^{1/2} \mathbf{z}_L \right\| \leq \sqrt{P_R^{\max}},$$

where $\mathbf{a}^T = [1, 0]$, $\mathbf{b}^T = [0, 1]$, and the quadratic inequality constraint $\mathbf{t}^T \mathbf{t} \leq \tilde{t}$ is converted to a linear matrix inequality constraint (LMI) using Schur-Complement theorem [16].

Remark 2. Notice that when other variables are fixed, $\mathbf{Z}$ can be optimized by solving the Karush-Kuhn-Tucker (KKT) conditions, where the Lagrangian multiplier that arises due to the relay power constraint can be obtained by using the bisection algorithm like in [15]. However, in order to make the proposed iterative approach applicable for other related problems briefly discussed in the beginning of this section and also for the optimization over $\mathbf{F}$ in the sequel, we propose to formulate our optimization problem in the convex form $\mathcal{P}_1$, which has been proven to be computationally efficient and flexible to accommodate even a large number of convex constraints [16].

(2) *Optimization over* $\mathbf{F}$. First, we define the following scalars that do not depend on $\mathbf{F}$:

$$s_5 \triangleq \operatorname{tr}\left(\mathbf{B} \mathbf{Z} \tilde{\boldsymbol{\Sigma}}_1 \mathbf{Z}^H\right), \qquad s_6 \triangleq \operatorname{tr}\left(\mathbf{B} \mathbf{Z} \mathbf{Z}^H\right),$$
$$s_7 \triangleq \operatorname{tr}\left(\mathbf{Z} \mathbf{Z}^H \boldsymbol{\Psi}_2\right), \qquad s_8 \triangleq \operatorname{tr}\left(\mathbf{Z} \tilde{\boldsymbol{\Sigma}}_1 \mathbf{Z}^H \boldsymbol{\Psi}_2\right). \quad (36)$$

After some simple steps and again using the fact that $\operatorname{tr}(\mathbf{X} \mathbf{X}^H) = \|\mathbf{X}\|^2$ in (17), the average sum MSE (20) can also be expressed as follows:

$$\overline{f}_{\text{mse}} = \left\| \left(\mathbf{W}_1 \hat{\mathbf{H}}_0 + \mathbf{W}_2 \hat{\mathbf{H}}_2 \mathbf{Z} \hat{\mathbf{H}}_1\right) \mathbf{F} - \mathbf{I}_{N_S} \right\|^2$$
$$+ \tilde{s}_2 \operatorname{tr}\left(\mathbf{F} \mathbf{F}^H \boldsymbol{\Psi}_0\right) + \sigma_{n_r}^2 s_6 + s_3$$
$$\times s_2 \operatorname{tr}\left(\mathbf{F}^H \mathbf{E} \mathbf{F}\right) + (s_5 + s_2 s_8) \operatorname{tr}\left(\mathbf{F} \mathbf{F}^H \boldsymbol{\Psi}_1\right) \quad (37)$$
$$+ s_2 s_7 \sigma_{n_r}^2 + \sigma_{n_d}^2 \operatorname{tr}\left(\mathbf{W}_1 \mathbf{W}_1^H\right),$$

where $\tilde{s}_2 = \operatorname{tr}(\mathbf{W}_1 \tilde{\boldsymbol{\Sigma}}_0 \mathbf{W}_1^H)$ and $\mathbf{E} = \hat{\mathbf{H}}_1^H \mathbf{Z}^H \boldsymbol{\Psi}_2 \mathbf{Z} \hat{\mathbf{H}}_1$. Noting that $\operatorname{vec}(\mathbf{X} \mathbf{Y}) = (\mathbf{I} \otimes \mathbf{X}) \operatorname{vec}(\mathbf{Y})$ [24] and applying (30), we can write

$$\overline{f}_{\text{mse}} = \left\| \left(\mathbf{I} \bigotimes \left(\mathbf{W}_1 \hat{\mathbf{H}}_0 + \mathbf{W}_2 \hat{\mathbf{H}}_2 \mathbf{Z} \hat{\mathbf{H}}_1\right)\right) \mathbf{f}_L - \operatorname{vec}(\mathbf{I}_{N_S}) \right\|^2$$
$$+ s_2 s_7 \sigma_{n_r}^2 + \sigma_{n_r}^2 s_6$$
$$+ \mathbf{f}_L^H \left[s_2 \left(\mathbf{I}_{N_S} \bigotimes \mathbf{E}\right) + \tilde{s}_2 \left(\mathbf{I}_{N_S} \bigotimes \boldsymbol{\Psi}_0\right) + (s_5 + s_2 s_8) \right.$$
$$\left. \times \left(\mathbf{I}_{N_S} \bigotimes \boldsymbol{\Psi}_1\right) \right] \mathbf{f}_L + s_3 + \sigma_{n_d}^2 \operatorname{tr}\left(\mathbf{W}_1 \mathbf{W}_1^H\right), \quad (38)$$

where $\mathbf{f}_L = \operatorname{vec}(\mathbf{F}) \in \mathcal{C}^{M N_S \times 1}$. The relay power in terms of $\mathbf{f}_L$ can be expressed as follows:

$$\overline{P}_R = s_9 + \mathbf{f}_L^H \left[s_{10} \left(\mathbf{I}_{N_S} \bigotimes \boldsymbol{\Psi}_1\right) + \left(\mathbf{I}_{N_S} \bigotimes \mathbf{Q}\right) \right] \mathbf{f}_L, \quad (39)$$

where we have used $s_9 \triangleq \sigma_{n_r}^2 \operatorname{tr}(\mathbf{Z} \mathbf{Z}^H), s_{10} \triangleq \operatorname{tr}(\mathbf{Z} \tilde{\boldsymbol{\Sigma}}_1 \mathbf{Z}^H)$, and $\mathbf{Q} \triangleq \hat{\mathbf{H}}_1^H \mathbf{Z}^H \mathbf{Z} \hat{\mathbf{H}}_1$. The optimization problem w.r.t. $\mathbf{F}$ can be now formulated as follows:

$$\mathcal{P}_2 : \min_{\mathbf{f}_L} \overline{f}_{\text{mse}} \quad \text{s. t.}$$

$$\|\mathbf{f}_L\|^2 \leq P_S^{\max},$$

$$s_9 + \left\| \left[\mathbf{I}_{N_S} \bigotimes (s_{10} \boldsymbol{\Psi}_1 + \mathbf{Q}) \right]^{1/2} \mathbf{f}_L \right\|^2 \leq P_R^{\max}. \quad (40)$$

Finally, like in the case of the optimization over $\mathbf{Z}$, we can transform (40) into the following convex problem:

$$\mathcal{P}_2 : \min_{\mathbf{f}_L, \mathbf{t}, \tilde{t}} \tilde{t} \quad \text{s. t.}$$

$$\begin{bmatrix} \mathbf{I}_2 & \mathbf{t} \\ \mathbf{t}^T & \tilde{t} \end{bmatrix} \succeq 0,$$

$$\left\| \left(\mathbf{I}_{N_S} \bigotimes \left(\mathbf{W}_1 \hat{\mathbf{H}}_0 + \mathbf{W}_2 \hat{\mathbf{H}}_2 \mathbf{Z} \hat{\mathbf{H}}_1\right)\right) \mathbf{f}_L - \operatorname{vec}(\mathbf{I}_{N_S}) \right\|$$

$$\leq \mathbf{a}^T \mathbf{t}, \quad \left\| \mathbf{S}^{1/2} \mathbf{f}_L \right\| \leq \mathbf{b}^T \mathbf{t},$$

$$\|\mathbf{f}_L\| \leq \sqrt{P_S^{\max}},$$

$$\left\| \left[\mathbf{I}_{N_S} \bigotimes (s_{10} \boldsymbol{\Psi}_1 + \mathbf{Q}) \right]^{1/2} \mathbf{f}_L \right\| \leq \sqrt{P_R^{\max} - s_9}, \quad (41)$$

where $\mathbf{S} = [s_2(\mathbf{I}_{N_S} \otimes \mathbf{E}) + \tilde{s}_2(\mathbf{I}_{N_S} \otimes \boldsymbol{\Psi}_0) + (s_5 + s_2 s_8)(\mathbf{I}_{N_S} \otimes \boldsymbol{\Psi}_1)]$ and the LMI is equivalent to the quadratic inequality constraint like in the case of $\mathcal{P}_1$. Note that the optimization problems $\mathcal{P}_1$ and $\mathcal{P}_2$ can also be solved first by reformulating them as quadratic matrix programming [25] problems and then applying the semidefinite programming (SDP) relaxation technique. However, since our problems are already convex and second-order cone programming (SOCP) formulation is possible, applying SDP relaxation only increases the computational complexity as we know that the SDP has higher computational burden than the SOCP [26]. The iterative optimization technique can be now summarized as follows:

Algorithm 1. The iterative algorithm for joint optimization of $\mathbf{W}_1$, $\mathbf{W}_2$, $\mathbf{Z}$ and $\mathbf{F}$.

(1) Initialize the algorithm with $\mathbf{Z}_0$ and $\mathbf{F}_0$ such that the source and relay power constraints are satisfied.

    (a) **repeat**
    (b) Update $\mathbf{W}_2^i$ using (23).
    (c) Update $\mathbf{W}_1^i$ using (21) and (23).
    (d) Update $\mathbf{Z}_i$ by solving the convex problem $\mathcal{P}_1$.
    (e) Update $\mathbf{F}_i$ by solving the convex problem $\mathcal{P}_2$.
    (f) $i = i + 1$;

(2) **until** both $|\overline{f}_{\text{mse},i} - \overline{f}_{\text{mse},i+1}|$ is smaller than a threshold ϵ, where the index i denotes the ith iteration.

Remark 3. In the case of the relay channel without the direct link, the optimization problem over destination equalizer, $\mathbf{Z}$ and $\mathbf{F}$ can be iteratively solved by omitting all terms containing $\hat{\mathbf{H}}_0$ and $\mathbf{W}_1$ in (23), $\mathcal{P}_1$, and $\mathcal{P}_2$.

Remark 4. If the channel estimates are perfect, (21), (23), $\mathcal{P}_1$, and $\mathcal{P}_2$ can be changed to shorter forms by using the fact that $\sigma_{e,1}^2 = \sigma_{e,2}^2 = \sigma_{e,0}^2 = 0$. The resulting equations and optimization problems correspond to the perfect CSI case.

3.1. Computational Complexity.

The computational complexity of Algorithm 1 mainly depends on the work loads of the convex optimization problems $\mathcal{P}_1$ and $\mathcal{P}_2$ which consist of second-order cone (SOC) as well as SDP constraints. An enormous amount of effort will be required to compute the exact computational complexity of $\mathcal{P}_1$ and $\mathcal{P}_2$, and thus their exact complexity analysis is beyond the scope of this paper. However, using the results of [26], we determine the worst-case computational complexity of $\mathcal{P}_1$ and $\mathcal{P}_2$. In $\mathcal{P}_1$, there are 3 SOC constraints where the first SOC constraint has a dimension (also the size of the cone) with $2N_S^2 + 1$ real variables and the remaining two SOC constraints have the same size of $2N_T N_R + 1$. The SOC constraints in $\mathcal{P}_1$ consist of $2N_T N_R + 2$ real optimization variables. The only one SDP constraint of $\mathcal{P}_1$ has a size of 3 with 3 real optimization variables. Therefore, according to [26], the computational load of $\mathcal{P}_1$ per iteration is $\mathcal{O}((2N_T N_R + 2)^2(2N_S^2 + 4N_T N_R + 3) + 81)$. The number of iterations required can be upper bounded by $\mathcal{O}(2\sqrt{3})$. Hence, the overall worst-case complexity of $\mathcal{P}_1$ is $\mathcal{O}(2\sqrt{3}((2N_T N_R + 2)^2(2N_S^2 + 4N_T N_R + 3) + 81))$. Similarly, we can compute the worst-case computational load for $\mathcal{P}_2$. The first SOC constraint in $\mathcal{P}_2$ has a size of $2N_S^2 + 1$ while the other SOC constraints are of the same size of $2MN_S + 1$. The SOC constraints consist of $2MN_S + 2$ real optimization variables. Like in the case of $\mathcal{P}_1$, the single SDP constraint is of size 3 with 3 real optimization variables. Thus (see also [26]), we find that $\mathcal{P}_2$ has a work load of $\mathcal{O}((2MN_S + 2)^2(2N_S^2 + 6MN_S + 4) + 81)$ per iteration, where the required number of iterations is upper bounded by $\mathcal{O}(\sqrt{3} + 2)$. As a result, the worst-case complexity of $\mathcal{P}_2$ is $\mathcal{O}((\sqrt{3} + 2)((2MN_S + 2)^2(2N_S^2 + 6MN_S + 4) + 81))$. Notice that in practice the interior-point algorithms used for

solving $\mathcal{P}_1$ and $\mathcal{P}_2$ behave much better than predicted by the aforementioned worst-case analysis [27].

3.2. Convergence Analysis.

It can be shown that the proposed iterative method converges. It has been already discussed that the optimization problem is convex w.r.t. each variable when the others are fixed. For the given $\mathbf{F}$ and $\mathbf{Z}$, the solutions given by (21) and (23) correspond to the MMSE receiver. As a result, we have $\text{tr}(\overline{\mathbf{M}}(\mathbf{F}^i, \mathbf{Z}^i, \mathbf{W}_1^{i+1}, \mathbf{W}_2^{i+1})) \leq \text{tr}(\overline{\mathbf{M}}(\mathbf{F}^i, \mathbf{Z}^i, \mathbf{W}_1^i, \mathbf{W}_2^i))$. Similarly, for the given $\mathbf{W}_1$, $\mathbf{W}_2$, and $\mathbf{F}$, the optimization problem (20) is convex w.r.t. $\mathbf{Z}$ and, thus, the problem $\mathcal{P}_1$ provides optimal solution for $\mathbf{Z}$ which means that $\text{tr}(\overline{\mathbf{M}}(\mathbf{F}^i, \mathbf{Z}^{i+1}, \mathbf{W}_1^i, \mathbf{W}_2^i)) \leq \text{tr}(\overline{\mathbf{M}}(\mathbf{F}^i, \mathbf{Z}^i, \mathbf{W}_1^i, \mathbf{W}_2^i))$. Finally, for the given $\mathbf{W}_1$, $\mathbf{W}_2$, and $\mathbf{Z}$, problem (20) is convex w.r.t. $\mathbf{F}$ and, hence, the problem $\mathcal{P}_2$ gives optimal solution for $\mathbf{F}$ thereby confirming $\text{tr}(\overline{\mathbf{M}}(\mathbf{F}^{i+1}, \mathbf{Z}^i, \mathbf{W}_1^i, \mathbf{W}_2^i)) \leq \text{tr}(\overline{\mathbf{M}}(\mathbf{F}^i, \mathbf{Z}^i, \mathbf{W}_1^i, \mathbf{W}_2^i))$. Therefore, it can be found that with each update of $\mathbf{W}_1$, $\mathbf{W}_2$, $\mathbf{F}$, and $\mathbf{Z}$, the objective function decreases and the iterative method converges. It will be later shown via numerical results that the iterative algorithm gives satisfactory performance with acceptable convergence speed. However, the global optimality of the solutions of the iterative method for the relay channel with the direct link cannot be guaranteed as the joint problem is nonconvex. In the next section, the joint optimization problem will be restricted to a double-hop MIMO relay, where receive antenna correlations are assumed to be negligible for each hop. In this case, the optimization problem turns to the joint power allocation, for which the global optimality can be guaranteed under high-SNR approximation.

4. Joint Power Allocation with GP: Without Direct Link

Due to severe shadowing, hotspots, and so forth, the destination may be out of the coverage area of the source. In such a scenario, it is reasonable to consider that the direct link between the source and destination does not exist. In the latter case of the relay channel without the direct link, the sum MSE can be expressed in terms of $\mathbf{F}$ and $\mathbf{Z}$ (see also (28)) as follows:

$$\begin{aligned}
\overline{f}_{\text{mmse}} &= \overline{f}_{\text{mmse,DH}} \\
&= \text{tr}\left(\mathbf{I}_{N_S} + \mathbf{F}^H \hat{\mathbf{H}}_1^H \mathbf{Z}^H \hat{\mathbf{H}}_2^H \mathbf{\Gamma}^{-1} \hat{\mathbf{H}}_2 \mathbf{Z} \hat{\mathbf{H}}_1 \mathbf{F}\right)^{-1}.
\end{aligned} \tag{42}$$

The objective is to minimize the average sum MSE (42) under the constraints of source and relay powers. Thus, the optimization problem is

$$\min_{\mathbf{Z}, \mathbf{F}} \text{tr}(\mathbf{MMSE}(\mathbf{Z}, \mathbf{F})) \quad \text{s.t.} \quad \text{tr}\left(\mathbf{FF}^H\right) \leq P_S^{\max}, \tag{43}$$

$$\text{tr}\left(\mathbf{Z}\widetilde{\mathbf{A}}\mathbf{Z}^H\right) \leq P_R^{\max},$$

where the MMSE matrix is $\mathbf{MMSE}(\mathbf{Z}, \mathbf{F}) = [\mathbf{I}_{N_S} + \mathbf{F}^H \hat{\mathbf{H}}_1^H \mathbf{Z}^H \hat{\mathbf{H}}_2^H \mathbf{\Gamma}^{-1} \hat{\mathbf{H}}_2 \mathbf{Z} \hat{\mathbf{H}}_1 \mathbf{F}]^{-1}$. Let $\boldsymbol{\lambda}(\mathbf{MMSE}(\mathbf{Z}, \mathbf{F}))$ be the vector of eigenvalues and $\mathbf{d}(\mathbf{MMSE}(\mathbf{Z}, \mathbf{F}))$ be the vector of the diagonal elements of $\mathbf{MMSE}(\mathbf{Z}, \mathbf{F})$ in decreasing order. In

this case, $\mathbf{d}$ is said to be majorized by $\boldsymbol{\lambda}$, and the Schur-concave function can be defined as $f(\boldsymbol{\lambda}(\mathbf{MMSE}(\mathbf{Z}, \mathbf{F}))) \leq f(\mathbf{d}(\mathbf{MMSE}(\mathbf{Z}, \mathbf{F})))$, where $f(\mathbf{x})$ stands for the function of $\mathbf{x}$. It is clear that the minimum of this function is obtained if the diagonal elements of $\mathbf{MMSE}(\mathbf{Z}, \mathbf{F})$ become its eigenvalues [8]. This can occur when $\mathbf{MMSE}(\mathbf{Z}, \mathbf{F})$ is a diagonal matrix with its elements in decreasing order. Let the singular value decompositions of $\hat{\mathbf{H}}_1$ and $\hat{\mathbf{H}}_2$ be

$$\hat{\mathbf{H}}_1 = \mathbf{U}_1 \boldsymbol{\Lambda}_1^{1/2} \mathbf{V}_1^H, \qquad \hat{\mathbf{H}}_2 = \mathbf{U}_2 \boldsymbol{\Lambda}_2^{1/2} \mathbf{V}_2^H, \qquad (44)$$

where the diagonal elements of $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ are considered to be in the decreasing order. If these elements are not in decreasing order, some permutation matrices can be applied to make the diagonal elements of $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ in the decreasing order [8]. This means that, in (44), the permutation matrices are implicitly included. The closed-form expressions for the optimal $\mathbf{Z}$ and $\mathbf{F}$ are difficult to obtain in general. However, for the double-hop channels without receive-side antenna correlations, that is, when $\boldsymbol{\Sigma}_1 = \mathbf{I}_{N_R}$ and $\boldsymbol{\Sigma}_2 = \mathbf{I}_{N_D}$, the optimal $\mathbf{Z}$ and $\mathbf{F}$ that diagonalize the MMSE matrix can be given by

$$\mathbf{Z} = \mathbf{V}_2 \boldsymbol{\Lambda}_{\mathbf{Z}}^{1/2} \mathbf{U}_1^H, \qquad \mathbf{F} = \mathbf{V}_1 \boldsymbol{\Lambda}_{\mathbf{F}}^{1/2}, \qquad (45)$$

where $\boldsymbol{\Lambda}_{\mathbf{Z}}$ and $\boldsymbol{\Lambda}_{\mathbf{F}}$ are the diagonal matrices with the elements in decreasing order. Substituting (44) and (45) into the MMSE matrix and after some straightforward steps, we get

$$\overline{f}_{\text{mmse}} = \text{tr}\Big(\boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \boldsymbol{\Lambda}_1^{T/2} \boldsymbol{\Lambda}_{\mathbf{Z}}^{T/2} \boldsymbol{\Lambda}_2^{T/2} \boldsymbol{\Gamma}_{\mathbf{D}}^{-1} \times \Big(\boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \boldsymbol{\Lambda}_1^{T/2} \boldsymbol{\Lambda}_{\mathbf{Z}}^{T/2} \boldsymbol{\Lambda}_2^{T/2} \Big)^T + \mathbf{I}_{N_S} \Big)^{-1}, \qquad (46)$$

where

$$\boldsymbol{\Gamma}_{\mathbf{D}} = \Big(\sigma_{e,1}^2 \, \text{tr}\big(\widetilde{\mathbf{B}} \boldsymbol{\Lambda}_{\mathbf{F}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \big) + \sigma_{n_r}^2 \Big) \boldsymbol{\Lambda}_2^{1/2} \boldsymbol{\Lambda}_{\mathbf{Z}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{Z}}^{T/2} \boldsymbol{\Lambda}_2^{T/2}$$
$$+ \text{tr}\big(\widetilde{\mathbf{C}} \boldsymbol{\Lambda}_{\mathbf{Z}}^{1/2} \boldsymbol{\Lambda}_1^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \boldsymbol{\Lambda}_1^{T/2} \boldsymbol{\Lambda}_{\mathbf{Z}}^{T/2} \big) \sigma_{e,2}^2 \mathbf{I}_{N_D} + \sigma_{e,2}^2 \cdot$$
$$\text{tr}\big(\widetilde{\mathbf{C}} \boldsymbol{\Lambda}_{\mathbf{Z}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{Z}}^{T/2} \big) \Big[\sigma_{e,1}^2 \, \text{tr}\big(\widetilde{\mathbf{B}} \boldsymbol{\Lambda}_{\mathbf{F}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \big) + \sigma_{n_r}^2 \Big] \mathbf{I}_{N_D} + \sigma_{n_d}^2 \mathbf{I}_{N_D}. \qquad (47)$$

In (47), we use $\widetilde{\mathbf{B}} = \mathbf{V}_1^H \boldsymbol{\Psi}_1 \mathbf{V}_1$ and $\widetilde{\mathbf{C}} = \mathbf{V}_2^H \boldsymbol{\Psi}_2 \mathbf{V}_2$. Similarly, the source and relay powers become

$$P_S = \text{tr}\big(\boldsymbol{\Lambda}_{\mathbf{F}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \big),$$
$$\overline{P}_R = \text{tr}\big(\boldsymbol{\Lambda}_{\mathbf{Z}}^{1/2} \boldsymbol{\Lambda}_1^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \boldsymbol{\Lambda}_1^{T/2} \boldsymbol{\Lambda}_{\mathbf{Z}}^{T/2} \big) \qquad (48)$$
$$+ \text{tr}\big(\boldsymbol{\Lambda}_{\mathbf{Z}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{Z}}^{T/2} \big) \Big[\sigma_{e,1}^2 \, \text{tr}\big(\widetilde{\mathbf{B}} \boldsymbol{\Lambda}_{\mathbf{F}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \big) + \sigma_{n_r}^2 \Big].$$

Let $q = \min(M, N_S)$ and $v = \min(N_T, N_R)$, and $\{ \sqrt{\lambda_{\mathbf{F}}^j} \}_{j=1}^q$ and let $\{ \sqrt{\lambda_{\mathbf{Z}}^k} \}_{k=1}^v$ be the nonzero diagonal elements of $\boldsymbol{\Lambda}_F^{1/2}$

and $\boldsymbol{\Lambda}_Z^{1/2}$, respectively, in the decreasing order. For brevity, we also define

$$d \triangleq \Big(\sigma_{e,1}^2 \, \text{tr}\big(\widetilde{\mathbf{B}} \boldsymbol{\Lambda}_{\mathbf{F}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \big) + \sigma_{n_r}^2 \Big) \triangleq \Big(\sigma_{e,1}^2 \sum_{i=1}^q \widetilde{\mathbf{B}}_{i,i} \lambda_{\mathbf{F}}^i + \sigma_{n_r}^2 \Big)$$

$$e \triangleq \text{tr}\big(\widetilde{\mathbf{C}} \boldsymbol{\Lambda}_{\mathbf{Z}}^{1/2} \boldsymbol{\Lambda}_1^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \boldsymbol{\Lambda}_1^{T/2} \boldsymbol{\Lambda}_{\mathbf{Z}}^{T/2} \big) \sigma_{e,2}^2 \triangleq \sigma_{e,2}^2 \sum_{i=1}^p \widetilde{\mathbf{C}}_{i,i} \lambda_{\mathbf{Z}}^i \lambda_{\mathbf{F}}^i \lambda_1^i.$$

$$f \triangleq \sigma_{e,2}^2 \, \text{tr}\big(\widetilde{\mathbf{C}} \boldsymbol{\Lambda}_{\mathbf{Z}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{Z}}^{T/2} \big) \Big[\sigma_{e,1}^2 \, \text{tr}\big(\widetilde{\mathbf{B}} \boldsymbol{\Lambda}_{\mathbf{F}}^{1/2} \boldsymbol{\Lambda}_{\mathbf{F}}^{T/2} \big) + \sigma_{n_r}^2 \Big]$$

$$\triangleq \sigma_{e,2}^2 \sum_{i=1}^v \widetilde{\mathbf{C}}_{i,i} \lambda_{\mathbf{Z}}^i \Big[\sum_{i=1}^q \widetilde{\mathbf{B}}_{i,i} \lambda_{\mathbf{F}}^i \sigma_{e,1}^2 + \sigma_{n_r}^2 \Big], \qquad (49)$$

where $\widetilde{\mathbf{B}}_{i,i}$ and $\widetilde{\mathbf{C}}_{i,i}$ are the ith diagonal elements of $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{C}}$, respectively, and $p = \min(N_T, N_R, M, N_S)$. Since $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{C}}$ are positive semidefinite, we have $\widetilde{\mathbf{B}}_{i,i} \geq 0$ and $\widetilde{\mathbf{C}}_{i,i} \geq 0$ for all i. Using (47) and (49), the sum MMSE can be finally expressed as follows:

$$\overline{f}_{\text{mmse}} = \sum_{r=1}^R \frac{1}{1 + (\lambda_{\mathbf{F}}^r \lambda_{\mathbf{Z}}^r \lambda_1^r \lambda_2^r) / \big(d \lambda_{\mathbf{Z}}^r \lambda_2^r + e + f + \sigma_{n_d}^2 \big)}$$

$$\triangleq \sum_{r=1}^R \frac{1}{1 + \text{SNR}_r}, \qquad (50)$$

where $\text{SNR}_r = \lambda_F^r \lambda_{\mathbf{Z}}^r \lambda_1^r \lambda_2^r / (d \lambda_{\mathbf{Z}}^r \lambda_2^r + e + f + \sigma_{n_d}^2)$ is the SNR of each data stream provided that N_S is also smaller than or equal to M and $R = \min(N_T, N_R, M, N_S, \text{ and } N_D)$. It is interesting to note that when the channel estimates are perfect, that is, when $\sigma_{e,1}^2 = \sigma_{e,2}^2 = 0$, the sum MMSE of (50) reduces to the objective function of [28]. Applying (48), the joint power allocation problem can be formulated as follows:

$$\min_{\{ \lambda_{\mathbf{F}}^j \}_{j=1}^q, \{ \lambda_{\mathbf{Z}}^k \}_{k=1}^v} \overline{f}_{\text{mmse}} \quad \text{s.t.} \quad \sum_{j=1}^q \lambda_{\mathbf{F}}^j \leq P_S^{\max}$$

$$\sum_{m=1}^p \lambda_{\mathbf{Z}}^m \lambda_{\mathbf{F}}^m \lambda_1^m + d \sum_{k=1}^v \lambda_{\mathbf{Z}}^k \leq P_R^{\max}, \qquad (51)$$

which is a nonlinear and nonconvex problem. Since this problem is nonconvex, the global optimal solution is difficult to obtain. Considering the fact that the global optimal solutions of the problems similar to the nonrobust version of (51) can be obtained only with very high computational cost (see [29, 30]), the authors of [31] use an iterative waterfilling technique to solve the nonrobust form of (51). However, we have noticed that it is hard to solve (51) using the waterfilling method of [31]. The major difficulty arises from the fact that the first-order partial derivatives of the corresponding Lagrangian function w.r.t. $\lambda_{\mathbf{F}}^r$ for the fixed $\{ \lambda_{\mathbf{Z}}^r \}_{r=1}^R$ and w.r.t. $\lambda_{\mathbf{Z}}^r$ for the given $\{ \lambda_{\mathbf{F}}^r \}_{r=1}^R$ do not lead to equations that are decoupled in $\lambda_{\mathbf{F}}^r$ and $\lambda_{\mathbf{Z}}^r$, respectively. It is easier to see that this difficulty appears due to the reason that d, e, and f in (51) consist of not only $\lambda_{\mathbf{F}}^r$ and $\lambda_{\mathbf{Z}}^r$ but also $\lambda_{\mathbf{F}}^k$ and

$\lambda_{\mathbf{Z}}^k$, for all $k \in \{1, \ldots, R\}$. Furthermore, although iterative waterfilling method is computationally efficient, it does not guarantee the global optimal solution. In the following, we use an alternative approach based on GP technique. Note that the optimization problem (51) is not a GP problem but a signomial programming (SP) problem [16] which can be iteratively solved as a GP problem after approximating the required posynomial terms by monomial terms. It is known that the SPs do not guarantee global optimality and the computational cost is high. Thus, using the high-SNR approximation, (51) can be solved as a GP whose global optimality can be confirmed. In this regard, we have

$$\overline{f}_{\text{mmse}} \leq \sum_{r=1}^{R} \frac{d\lambda_{\mathbf{Z}}^r \lambda_2^r + e + f + \sigma_{n_{\mathrm{d}}}^2}{\lambda_{\mathbf{F}}^r \lambda_{\mathbf{Z}}^r \lambda_1^r \lambda_2^r}. \tag{52}$$

Using the upper bound (52), the power optimization problem can be expressed as follows:

$$\min_{\{\lambda_{\mathbf{F}}^j\}_{j=1}^q, \{\lambda_{\mathbf{Z}}^k\}_{k=1}^v} \sum_{r=1}^{R} t_r \quad \text{s.t.}$$

$$d\lambda_{\mathbf{Z}}^r \lambda_2^r + e + f + \sigma_{n_{\mathrm{d}}}^2 \leq t_r \lambda_{\mathbf{F}}^r \lambda_{\mathbf{Z}}^r \lambda_1^r \lambda_2^r, \quad \forall r$$

$$\sum_{j=1}^{q} \lambda_{\mathbf{F}}^j \leq P_{\mathrm{S}}^{\max},$$

$$\sum_{m=1}^{p} \lambda_{\mathbf{Z}}^m \lambda_{\mathbf{F}}^m \lambda_1^m + d\sum_{k=1}^{v} \lambda_{\mathbf{Z}}^k \leq P_{\mathrm{R}}^{\max}, \tag{53}$$

which is a GP problem that can be solved efficiently to guarantee the global optimality.

Remark 5. Notice that the joint power allocation problem is nonconvex even for the case without channel estimation errors [32]. In such a nonrobust design, it has been shown in [28] that the lower and upper bounds can be established for the MSE for each data stream. Unfortunately, this is not the case for the proposed robust design due to the fact that the terms e, f, and d are again functions of $\lambda_{\mathbf{Z}}^k$, for all $k = 1, \ldots, v$ and $\lambda_{\mathbf{F}}^j$, for all $j = 1, \ldots, q$. It is also worthwhile to note that several optimization problems which can be solved using the GP method and still provide solutions close to the optimal solution of the sum MSE minimization problem are; (a) maximization of the minimum of the SNRs of the data streams and (b) maximization of the geometric mean of the SNRs of the data streams, both under the source and relay power constraints.

5. Numerical Results and Discussions

In this section, the performance of the proposed methods will be investigated. The proposed robust designs are also compared with the nonrobust case, where the channel estimation errors are not taken into account. Notice that the *nonrobust* design corresponds to the so called *naïve design*, where the optimization problems of interest are

solved assuming that there are no errors in channel estimates (see also Remark 4) although in reality the estimates are erroneous. For all numerical simulations, we take $N_{\mathrm{S}} = 3$, $M = 4$, $N_{\mathrm{R}} = 3$, $N_{\mathrm{T}} = 4$, and $N_{\mathrm{D}} = 3$. The spatial covariance matrices for *source-relay*, *relay-destination*, and *source-destination* channels are modelled according to the widely used exponential correlation model. In our examples, we take

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_0 = \begin{bmatrix} 1 & \beta & \beta^2 \\ \beta & 1 & \beta \\ \beta^2 & \beta & 1 \end{bmatrix},$$

$$\boldsymbol{\Psi}_1 = \boldsymbol{\Psi}_2 = \boldsymbol{\Psi}_0 = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ \alpha & 1 & \alpha & \alpha^2 \\ \alpha^2 & \alpha & 1 & \alpha \\ \alpha^3 & \alpha^2 & \alpha & 1 \end{bmatrix}. \tag{54}$$

In all cases, the optimization problems $\mathcal{P}_1$, $\mathcal{P}_2$, and (53) are solved using the CVX software [33]. The SNRs for *source-relay*, *source-destination*, and *relay-destination* channels are defined as $\text{SNR}_{\text{sr}} = P_{\mathrm{S}}^{\max}/\sigma_{n_r}^2$, $\text{SNR}_{\text{sd}} = P_{\mathrm{S}}^{\max}/\sigma_{n_{\mathrm{d}}}^2$, and $\text{SNR}_{\text{rd}} = P_{\mathrm{R}}^{\max}/\sigma_{n_{\mathrm{d}}}^2$, respectively. Throughout all examples, we take $P_{\mathrm{S}}^{\max} = P_{\mathrm{R}}^{\max} = 0$ dBw and vary the values of $\sigma_{n_r}^2$ and $\sigma_{n_{\mathrm{d}}}^2$ to change SNR_{sr}, SNR_{sd}, and SNR_{rd}. The estimated channels are generated according to (7), so that the elements of actual channels $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_0$ have the variance of 1. For all results, we compute the average MSE by taking 200 realizations of the estimated channels.

The convergence behaviour of the proposed iterative method as a function of iteration index is illustrated in Figure 2 for the relay channel with the direct link. The parameters in this figure are set as $\alpha = 0.2$, $\beta = 0.1$, $\sigma_{e,1}^2 = \sigma_{e,2}^2 = 0.01$, and $\sigma_{e,0}^2 = 0.03$. We take three sets of SNRs as $\text{SNR}_{\text{sr}} = \text{SNR}_{\text{sd}} = \text{SNR}_{\text{rd}} = 10$ dB, $\text{SNR}_{\text{sr}} = \text{SNR}_{\text{sd}} = \text{SNR}_{\text{rd}} = 20$ dB and $\text{SNR}_{\text{sr}} = 10$ dB, and $\text{SNR}_{\text{sd}} = \text{SNR}_{\text{rd}} = 0$ dB. It can be seen from this figure that the iterative method converges in about 15 iterations. The convergence is faster for the lower values of the SNR. The effect of different initializations on the convergence behaviours of the iterative method is also displayed in Figure 2. The convergence speed for the cases where $\mathbf{F}$ and $\mathbf{Z}$ are initialized to randomly generated matrices of ZMCSCG random variables is similar to the cases where $\mathbf{F}$ and $\mathbf{Z}$ are initialized to the matrices proportional to identity (i.e., $\mathbf{F} \propto \mathbf{I}$ and $\mathbf{Z} \propto \mathbf{I}$). Moreover, it can be noticed from Figure 2 that different initializations lead to the similar solutions. In Figure 3, the performance of the iterative method as a function of the iteration index is illustrated for the relay channel without the direct link. We take $\alpha = 0.3$, $\beta = 0$, and $\sigma_{e,1}^2 = \sigma_{e,2}^2 = 0.01$ in this figure. As a reference, the performance of the GP problem which gives optimal solution under high-SNR assumption is also displayed in Figure 3. It can be noticed from Figure 3 that the difference between the solutions of the iterative method and GP method is negligible after 10–15 iterations.

The performance of the proposed iterative method for the MIMO relay channel with the direct link is shown in Figure 4 for different values of σ_e^2. The performance of the
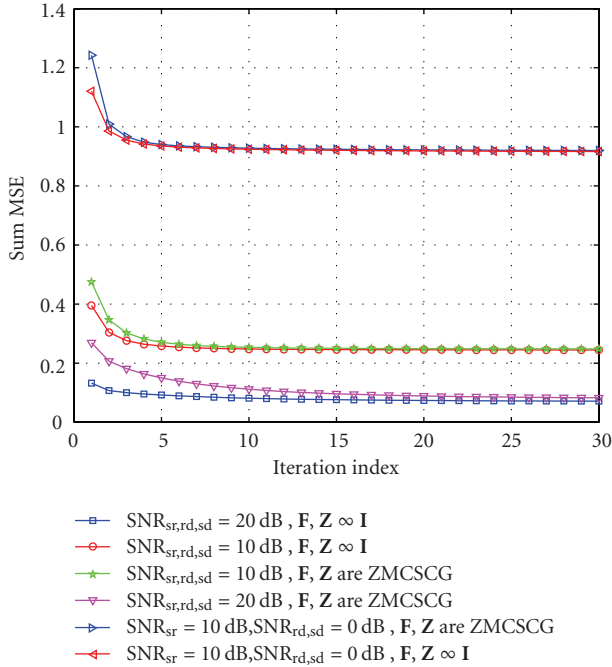
Figure 2: Convergence behaviour of the iterative approach for different initializations with the direct link ($\alpha = 0.2$, $\beta = 0.1$, $\sigma_{e,1}^2 = \sigma_{e,2}^2 = 0.01$, $\sigma_{e,0}^2 = 0.03$, and $P_S^{max} = P_R^{max} = 0$ dB).

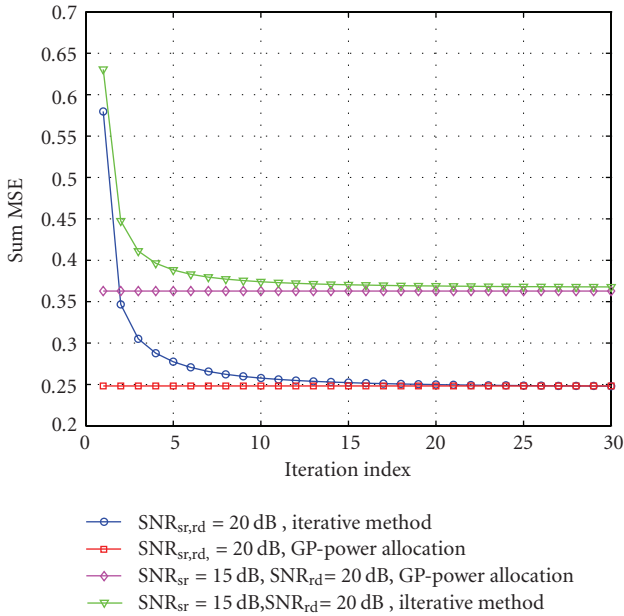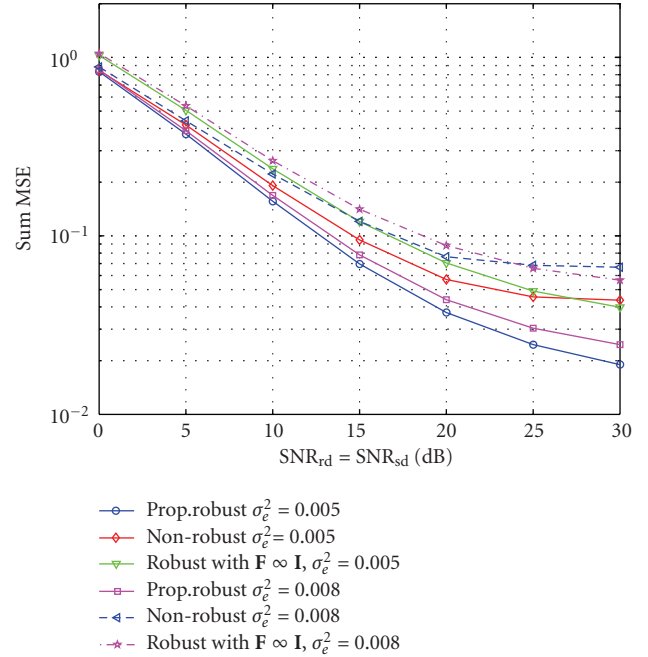Figure 4: Sum MSE as a function of SNR (SNR$_{rd}$ = SNR$_{sd}$) for the relay channel with the direct link ($\alpha = 0.6$, $\beta = 0.4$, $\sigma_{e,1}^2 = \sigma_{e,2}^2 = \sigma_{e,0}^2 = \sigma_e^2$, and $P_S^{max} = P_R^{max} = 0$ dB).

and SNR$_{rd}$ from 0 to 30 dB. The threshold ϵ for stopping the iterative process is set to $1e - 4$. It can be noticed from this figure that in all cases, the MSE decreases when the SNR increases and when the variance of the channel estimation error σ_e^2 decreases. Furthermore, the proposed robust design outperforms both the nonrobust method and the robust method with $\mathbf{F} \propto c\mathbf{I}$ [15]. In Figure 5, the performance between the proposed iterative method and the GP power allocation method for the relay channel without the direct link is compared. This figure displays the sum MSE as a function of the correlation coefficient α for different values of SNR$_{sd}$ and SNR$_{rd}$. We keep $\beta = 0$, $\sigma_e^2 = 0.01$ and $\epsilon = 1e - 4$ for Figure 5. It can be observed from this figure that the proposed robust methods significantly outperform the nonrobust method. Moreover, the performance gap between the iterative approach and the GP power allocation method is negligible for $\alpha \leq 0.4$. For $\alpha > 0.4$, the iterative approach outperforms the GP method. In all cases, the sum MSE increases with the increasing values of correlation coefficient.

6. Conclusions

The problem of jointly optimizing the source precoder, relay transceiver, and destination equalizer for a cooperative MIMO relay system has been treated in this paper. An iterative approach with the guaranteed convergence has been proposed to solve the nonconvex problem. It has been shown that for the case without the direct link where the two-hop MIMO channels have only transmit-side antenna correlations, the joint optimization turns to the joint source and relay power allocation problem which has been solved by
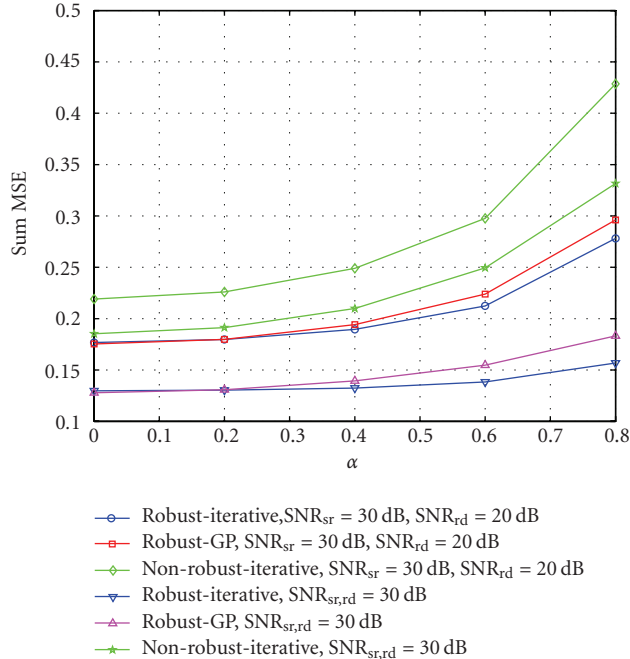

Figure 3: Convergence performance of the iterative approach and the GP power allocation method for the case without the direct link ($\alpha = 0.3$, $\beta = 0$, $\sigma_{e,1}^2 = \sigma_{e,2}^2 = 0.01$, and $P_S^{max} = P_R^{max} = 0$ dB).

nonrobust method which considers the estimated channels as actual channels and the performance of the robust method without a source precoder [15] (the case with $\mathbf{F} \propto c\mathbf{I}$ where c is the positive scaling factor chosen for satisfying the source power constraint) are also displayed in this figure. We keep $\alpha = 0.4$, $\beta = 0.6$, and SNR$_{sr} = 30$ dB and change SNR$_{sd}$

FIGURE 5: Sum MSE as a function of α for the relay channel without the direct link ($\beta = 0$, $\sigma_{e,1}^2 = \sigma_{e,1}^2 = 0.01$, and $P_S^{\max} = P_R^{\max} = 0$ dB).

using GP technique under high-SNR approximation. Simulation results confirm the efficiency and good performance of the iterative approach for the MIMO relay channel with and without the direct link. Furthermore, the proposed robust methods significantly outperform the nonrobust method when the channel estimates are imperfect.

Acknowledgments

References

[1] J. N. Laneman and G. W. Wornell, "Exploiting distributed spatial diversity in wireless networks," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, Monticello, Ill, USA, October 2000.

[2] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: space-time transmission and iterative decoding," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 362–371, 2004.

[3] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, 2005.

[4] R. U. Nabar, H. Bölcskei, and F. W. Kneubühler, "Fading relay channels: performance limits and space-time signal design," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1099–1109, 2004.

[5] Y. Jing and H. Jafarkhani, "Using orthogonal and quasi-orthogonal designs in wireless relay networks," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4106–4118, 2007.

[6] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z.-Q. Luo, "Distributed beamforming for relay networks based on second-order statistics of the channel state information," *IEEE Transactions on Signal Processing*, vol. 56, no. 9, pp. 4306–4316, 2008.

[7] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 29–43, 2005.

[8] O. Muñoz-Medina, J. Vidal, and A. Agustín, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Transactions on Signal Processing*, vol. 55, pp. 2593–2604, 2007.

[9] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1398–1406, 2007.

[10] C.-B. Chae, T. Tang, R. W. Heath Jr., and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 727–738, 2008.

[11] R. Zhang, C. C. Chai, and Y.-C. Liang, "Joint beamforming and power control for multiantenna relay broadcast channel with QoS constraints," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 726–737, 2009.

[12] B. K. Chalise and L. Vandendorpe, "MIMO relay design for multipoint-to-multipoint communications with imperfect channel state information," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2785–2796, 2009.

[13] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 5062–5073, 2008.

[14] W. Guan, H. Luo, and W. Chen, "Linear relaying scheme for MIMO relay system with QoS requirements," *IEEE Signal Processing Letters*, vol. 15, pp. 697–700, 2008.

[15] X. Chengweng, M. Shaodan, and W. Yik-Chung, "Robust joint design of linear precoder and destination equalizer for dual-hop amplifyand-forward MIMO relay systems," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2273–2283, 2010.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[17] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 313–324, 2003.

[18] A. Pascual-Iserte, D. Pérez Palomar, A. I. Pérez-Neira, and M. Á. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 346–360, 2006.

[19] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Transactions on Signal Processing*, vol. 56, pp. 3678–3689, 2008.

[20] B. K. Chalise, S. Shahbazpanahi, A. Czylwik, and A. B. Gershman, "Robust downlink beamforming based on outage probability specifications," *IEEE Transactions on Wireless Communications*, vol. 6, no. 10, pp. 3498–3503, 2007.

[21] J. Luo, J. R. Zeidler, and S. McLaughlin, "Performance analysis of compact antenna arrays with MRC in correlated Nakagami fading channels," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 1, pp. 267–277, 2001.

[22] T. Yoo and A. Goldsmith, "Capacity of fading MIMO channels with channel estimation error," in *Proceedings of the IEEE International Conference on Communications (ICC '04)*, pp. 808–813, Paris, France, June 2004.

[23] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1673–1690, 2002.

[24] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[25] A. Beck, "Quadratic matrix programming," *SIAM Journal on Optimization*, vol. 17, no. 4, pp. 1224–1238, 2007.

[26] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and Its Applications*, vol. 284, no. 1–3, pp. 193–228, 1998.

[27] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Transactions on Signal Processing*, vol. 53, no. 1, pp. 168–181, 2005.

[28] C. Li, X. Wang, L. Yang, and W.-P. Zhu, "A joint source and relay power allocation scheme for a class of MIMO relay systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4852–4860, 2009.

[29] W. Zhang, U. Mitra, and M. Chiang, "Optimization of amplify-and-forward multicarrier two-hop transmission," submitted to *IEEE Transactions on Communications*.

[30] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 922–933, 2006.

[31] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4837–4851, 2009.

[32] I. Hammerstrom and A. Wittneben, "Power allocation for amplify and forward MIMO-OFDM relay links," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1774–1786, 2007.

[33] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," June 2009, http://stanford.edu/~boyd/cvx.