Joint Transceiver Optimization for MIMO Multiuser Relaying Networks with Channel Uncertainties

Jiaxin Yang and Benoit Champagne
Department of Electrical and Computer Engineering, McGill University, Montreal, Quebec, Canada
Email: jiaxin.yang@mail.mcgill.ca; benoit.champagne@mcgill.ca

***Abstract*—This paper addresses the problem of amplify-and-forward (AF) relaying for multiple-input multiple-output (MIMO) multiuser relay networks, where each source transmits multiple data streams to its corresponding destination with the assistance of multiple relays. Assuming only imperfect channel state information (CSI) of all the source-relay and relay-destination links, we propose a robust approach to jointly design the source and relay precoders and the receive filters, in which the worst per-stream mean square error (MSE) is minimized subject to source and relay power constraints. The channel uncertainties are assumed to be Gaussian distributed and the well-known Kronecker model is employed to characterize the spatial correlations in the proposed design. The resultant optimization problem is nonconvex and therefore, an algorithmic solution with proven convergence is proposed by resorting to the iterative block coordinate update approach along with matrix transformation and convex conic optimization techniques. Simulation results show that the proposed joint transceiver design can achieve an improved robustness against the channel uncertainties when compared to the non-robust approaches.**

I. Introduction

Multiple-input multiple-output (MIMO) amplify-and-forward (AF) relaying designed for multiuser networks has attracted considerable interest in the context of wireless standards such as 3GPP LTE-Advanced. The relay precoder optimization has been extensively studied in a single-antenna multiuser framework, where each source/destination is equipped with a single antenna, under different design criteria; see, e.g., [1], [2] and references therein.

However, upcoming wireless standards target mobile broadband services with enhanced data rate and quality-of-service (QoS), which leads to a strong interest in studying cooperative relaying in a MIMO multiuser framework where multiple antennas are employed at the sources, relays and destinations. The joint transceiver optimization in this scenario, although more challenging than the relay precoder design in the single-antenna case, can provide further performance leverage. In the literature, the authors in [3] consider the joint transceiver design to minimize certain global objective functions, such as the sum power of interference received at all the destinations and the sum mean square error (MSE) of all the estimated data streams, by adopting the *alternating minimization* approach, which, however, does not take into account the weighted fairness of each stream. The authors in [4] consider minimizing the total source and relay power subject to a minimum signal-to-noise-plus-interference ratio (SINR) requirement for each source-destination link. To this end, a two-level iterative algorithm is proposed which involves semidefinite relaxation (SDR); however, the number of data streams transmitted by each source in this work is limited to one.

The efficacy of joint transceiver design in [3], [4] relies on the assumption of perfect channel state information (CSI) of all the source-relay and relay-destination links. In practice, acquiring perfect CSI at a central processing node is quite challenging. This is primarily due to the combined effects of various sources of imperfection, e.g., channel estimation errors, limited quantized feedback, and feedback delays. The performance of these previous design approaches therefore will substantially degrade in the presence of CSI errors.

In this paper, we study robust joint transceiver design for a general MIMO multiuser relay network in the presence of statistical CSI errors. Our objective, in contrast to the prior works, is to minimize the worst weighted per-stream MSE subject to source and relay power constraints, thus ensuring weighted fairness among multiple data streams. The resultant optimization problem being nonconvex, we propose an algorithmic solution by resorting to the block coordinate update approach, which iteratively solves for the optimal source and relay precoders, and receive filters until convergence, which is guaranteed in this case. The superiority of our proposed robust design approach is demonstrated by numerical simulations.

The remainder of the paper is organized as follows. Section II introduces the system model and outlines the robust joint transceiver design problem. The iterative algorithm proposed to solve the nonconvex optimization problem is exposed in Section III. Numerical results and discussions are reported in Section IV. Finally, Section V presents the conclusions.

II. System Model and Problem Formulation

We consider a MIMO multiuser relay network where M AF relay nodes assist the one-way communication between K source/destination pairs, as depicted in Fig. 1, with all the nodes equipped with multiple antennas. Specifically, the k^{th} source and destination, respectively, employ $N_{\mathrm{S},k}$ and $N_{\mathrm{D},k}$ antennas, for $k \in \mathcal{K} \triangleq \{1, 2, \cdots, K\}$, while the m^{th} relay employs $N_{\mathrm{R},m}$ antennas, for $m \in \mathcal{M} \triangleq \{1, \cdots, M\}$. All the relays work under the half-duplex AF protocol in which the

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and InterDigital Canada.

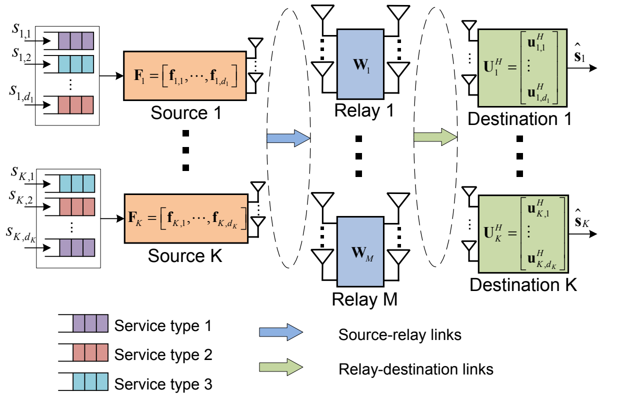

Fig. 1. Multiuser MIMO relay network with each source transmitting multiple data streams to its corresponding destination.

data transmission from the sources to their destinations is completed in two stages. In the first stage, all the sources transmit their signals to the relays simultaneously while in the second stage, the relays apply linear processing to the received signals and forward the resultant to all the destinations. We assume that no direct links between the sources and destinations are available due to the severe attenuation.

A narrowband flat-fading radio propagation model is considered where we denote by $\mathbf{H}_{m,k} \in \mathbb{C}^{N_{\mathrm{R},m} \times N_{\mathrm{S},k}}$ the channel matrix between the k^{th} source and the m^{th} relay, and by $\mathbf{G}_{k,m} \in \mathbb{C}^{N_{\mathrm{D},k} \times N_{\mathrm{R},m}}$ the channel matrix between the m^{th} relay and the k^{th} destination. Let $\mathbf{s}_k \triangleq [s_{k,1}, \cdots, s_{k,d_k}]^T$ denote the information symbols to be transmitted by the k^{th} source at a given time instance, where $d_k \leq \min\{N_{\mathrm{S},k}, N_{\mathrm{D},k}\}$ is the number of independent data streams. The symbols are modeled as independent random variables with zero mean and unit variance, i.e., $\mathbb{E}\left\{\mathbf{s}_k\mathbf{s}_k^H\right\} = \mathbf{I}_{d_k}$. The k^{th} source applies a linear precoding matrix $\mathbf{F}_k \in \mathbb{C}^{N_{\mathrm{S},k} \times d_k}$ to map the information symbols $\mathbf{s}_k$ to its $N_{\mathrm{S},k}$ transmit antennas. The transmit power is thus given by $\mathrm{Tr}\left(\mathbf{F}_k\mathbf{F}_k^H\right) \leq P_{\mathrm{S},k}^{\max}$, where $P_{\mathrm{S},k}^{\max}$ is the maximum affordable power of the k^{th} source. Let $\mathbf{n}_{\mathrm{R},m} \in \mathbb{C}^{N_{\mathrm{R},m} \times 1}$ be the spatially white, additive noise vector at the m^{th} relay, with zero mean and covariance matrix $\mathbb{E}\left\{\mathbf{n}_{\mathrm{R},m}\mathbf{n}_{\mathrm{R},m}^H\right\} = \sigma_{\mathrm{R},m}^2\mathbf{I}_{N_{\mathrm{R},m}}$. After the first stage of transmission, the received signal at the m^{th} relay is given by

$$\mathbf{z}_{\mathrm{R},m} = \sum_{k=1}^{K} \mathbf{H}_{m,k}\mathbf{F}_k\mathbf{s}_k + \mathbf{n}_{\mathrm{R},m}. \tag{1}$$

Each relay then applies a linear precoding matrix $\mathbf{W}_m \in \mathbb{C}^{N_{\mathrm{R},m} \times N_{\mathrm{R},m}}$ to $\mathbf{z}_{\mathrm{R},m}$ and forward the resulting signal

$$\mathbf{r}_{\mathrm{R},m} = \mathbf{W}_m\mathbf{z}_{\mathrm{R},m} = \sum_{k=1}^{K} \mathbf{W}_m\mathbf{H}_{m,k}\mathbf{F}_k\mathbf{s}_k + \mathbf{W}_m\mathbf{n}_{\mathrm{R},m} \tag{2}$$

to the destinations with power

$$P_{\mathrm{R},m} = \sum_{k=1}^{K} \mathrm{Tr}\left(\mathbf{W}_m\mathbf{H}_{m,k}\mathbf{F}_k\mathbf{F}_k^H\mathbf{H}_{m,k}^H\mathbf{W}_m^H\right) + \sigma_{\mathrm{R},m}^2\,\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\right). \tag{3}$$

Let $\mathbf{n}_{\mathrm{D},k}$ denote the spatially white, additive noise vector at the k^{th} destination with zero mean and covariance matrix $\mathbb{E}\left\{\mathbf{n}_{\mathrm{D},k}\mathbf{n}_{\mathrm{D},k}^H\right\} = \sigma_{\mathrm{D},k}^2\mathbf{I}_{N_{\mathrm{D},k}}$. The k^{th} destination observes the following signal after the second stage of transmission

$$\mathbf{y}_k = \sum_{q=1}^{K}\sum_{m=1}^{M} \mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,q}\mathbf{F}_q\mathbf{s}_q + \sum_{m=1}^{M} \mathbf{G}_{k,m}\mathbf{W}_m\mathbf{n}_{\mathrm{R},m} + \mathbf{n}_{\mathrm{D},k} \tag{4}$$

where subscript q is now used to index the sources. To estimate the l^{th} data stream from its corresponding source, the k^{th} destination applies a linear beamforming vector $\mathbf{u}_{k,l}$ to the received signal, thus forming a receive filtering matrix $\mathbf{U}_k = [\mathbf{u}_{k,1}, \cdots, \mathbf{u}_{k,d_k}] \in \mathbb{C}^{N_{\mathrm{D},k} \times d_k}$. The estimated information symbols can then be given by

$$\hat{\mathbf{s}}_k = \sum_{q=1}^{K}\sum_{m=1}^{M} \mathbf{U}_k^H\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,q}\mathbf{F}_q\mathbf{s}_q + \sum_{m=1}^{M} \mathbf{U}_k^H\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{n}_{\mathrm{R},m} + \mathbf{U}_k^H\mathbf{n}_{\mathrm{D},k}. \tag{5}$$

We adopt the MSE as the QoS metric for each estimated data stream. The major advantage of using MSE is to make the design problem tractable, which has been well justified in the relay precoder design literature (see e.g., [5] and references therein). The MSE of the l^{th} estimated data stream for $l \in \mathcal{D}_k \triangleq \{1, \cdots, d_k\}$ at the k^{th} destination is defined as

$$\varepsilon_{k,l} = \mathbb{E}\left\{|\hat{s}_{k,l} - s_{k,l}|^2\right\}. \tag{6}$$

Using (5), we can obtain

$$\varepsilon_{k,l} = \left\|\mathbf{u}_{k,l}^H\sum_{m=1}^{M}\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,k}\mathbf{F}_k - \mathbf{e}_{k,l}^T\right\|^2 + \sum_{q=1,q\neq k}^{K}\left\|\mathbf{u}_{k,l}^H\sum_{m=1}^{M}\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,q}\mathbf{F}_q\right\|^2 + \sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\left\|\mathbf{u}_{k,l}^H\mathbf{G}_{k,m}\mathbf{W}_m\right\|^2 + \sigma_{\mathrm{D},k}^2\left\|\mathbf{u}_{k,l}\right\|^2 \tag{7}$$

where $\mathbf{e}_{k,l}$ is a $d_k \times 1$ vector with all zero entries but the l^{th} entry, which is equal to one.

In typical relay scenarios, the CSI of both source-relay and relay-destination links available at the central processing node is often imperfect due to channel estimation errors, limited quantized feedback, and feedback delays. To model the CSI errors, consider expressing the true but unknown channels as

$$\mathbf{H}_{m,k} = \hat{\mathbf{H}}_{m,k} + \Delta\mathbf{H}_{m,k},\ \ \mathbf{G}_{k,m} = \hat{\mathbf{G}}_{k,m} + \Delta\mathbf{G}_{k,m} \tag{9}$$

where $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$, respectively, denote the estimated source-relay and relay-destination channels, while $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ capture the corresponding *channel uncertainties*. A statistical uncertainty model is considered here where the elements of $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ are zero-mean Gaussian random variables. Specially, using the Kronecker model [6], they can be in general written as

$$\Delta\mathbf{H}_{m,k} = \mathbf{\Sigma}_{\mathrm{H}_{m,k}}^{1/2} \Delta\mathbf{H}_{m,k}^{\mathrm{W}} \mathbf{\Psi}_{\mathrm{H}_{m,k}}^{1/2} \quad (10)$$

$$\Delta\mathbf{G}_{k,m} = \mathbf{\Sigma}_{\mathrm{G}_{k,m}}^{1/2} \Delta\mathbf{G}_{k,m}^{\mathrm{W}} \mathbf{\Psi}_{\mathrm{G}_{k,m}}^{1/2} \quad (11)$$

where $\mathbf{\Sigma}_{\mathrm{H}_{m,k}}$, $\mathbf{\Sigma}_{\mathrm{G}_{k,m}}$ are the row correlation matrices and $\mathbf{\Psi}_{\mathrm{H}_{m,k}}$, $\mathbf{\Psi}_{\mathrm{G}_{k,m}}$ are the column correlation matrices, all being positive definite. The entries of $\Delta\mathbf{H}_{m,k}^{\mathrm{W}}$ and $\Delta\mathbf{G}_{k,m}^{\mathrm{W}}$ are independently and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance.

In this work, we assume that the magnitudes of the CSI errors are significantly less than those of the channel estimates, and therefore, the third- and higher-order terms in $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ are neglected in the subsequent analysis. Substituting (9) into (7), and applying the latter assumption, the MSE incorporating the CSI errors can be written as in (8), shown at the bottom of this page where the following notations are introduced to simply the exposition: $\boldsymbol{\mathcal{G}}_{k,m} \triangleq \hat{\mathbf{G}}_{k,m}\mathbf{W}_m$, $\boldsymbol{\mathcal{W}}_{m,k} \triangleq \mathbf{W}_m\hat{\mathbf{H}}_{m,k}$ and $\boldsymbol{\mathcal{T}}_{k,q} \triangleq \sum_{m=1}^{M} \hat{\mathbf{G}}_{k,m}\mathbf{W}_m\hat{\mathbf{H}}_{m,q}\mathbf{F}_q$.

We observe that the MSE $\varepsilon_{k,l}$ now becomes uncertain in the CSI errors $\Delta\mathbf{H}_{m,k}$, $\forall(m,k) \in \mathcal{M}\times\mathcal{K}$, and $\Delta\mathbf{G}_{k,m}$, $\forall m \in \mathcal{M}$. Therefore, we shall next formulate the robust design problem, which aims to guarantee satisfactory per-stream QoS in the presence of such channel uncertainties.

For notational convenience, we define $\mathbf{F} \triangleq (\mathbf{F}_1, \cdots, \mathbf{F}_K)$, $\mathbf{W} \triangleq (\mathbf{W}_1, \cdots, \mathbf{W}_M)$ and $\mathbf{U} \triangleq (\mathbf{U}_1, \cdots, \mathbf{U}_K)$, which compactly represent the corresponding design variables. We aim to find $\{\mathbf{F}, \mathbf{W}, \mathbf{U}\}$ that jointly minimize the worst weighted per-stream MSE subject to the source and relay power constraints. Mathematically, the problem can be formulated as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{\forall k\in\mathcal{K}, l\in\mathcal{D}_k} \kappa_{k,l}\overline{\varepsilon}_{k,l} \quad (12a)$$

$$\text{s.t.} \quad \overline{P}_{\mathrm{R},m} \le \rho_m P_{\mathrm{R}},\ \forall m \in \mathcal{M} \quad (12b)$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right) \le P_{\mathrm{S},k}^{\max},\ \forall k \in \mathcal{K} \quad (12c)$$

where

$$\overline{\varepsilon}_{k,l} \triangleq \mathbb{E}\left\{\varepsilon_{k,l}\right\},\ \overline{P}_{\mathrm{R},m} \triangleq \mathbb{E}\left\{P_{\mathrm{R},m}\right\} \quad (13)$$

respectively denote the averaged MSE and relay power, $\{\kappa_{k,l} > 0 : \forall k \in \mathcal{K}, l \in \mathcal{D}_k\}$ is a set of weights that are assigned to different data streams, P_{R} is a common maximum allowed power for all the relays, and $\{\rho_m > 0 : \forall m \in \mathcal{M}\}$ is a set of coefficients used to control the power of each relay.

The above optimization problem is nonconvex and in general NP-hard; therefore, finding a global optimum is computationally expensive, if at all possible. In fact, the large number of design variables involved in the optimization makes even the exhaustive search challenging. To overcome these difficulties, it is more meaningful to employ instead a computationally tractable algorithm which leads to a local optimum of the design problem, as presented below.

III. Robust Joint Transceiver Optimization

In this section, we propose an algorithmic solution to the problem (12), which is robust to the statistical CSI errors. To further exploit the structure of (12), we need to compute the averaged MSE and relay power in (13). First, exploiting the independence of $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$, the averaged MSE over the channel uncertainties can be given by (14), which is shown on top of the next page, where

$$\mathbf{R}_k = \sum_{q=1,q\neq k}^{K} \boldsymbol{\mathcal{T}}_{k,q}\boldsymbol{\mathcal{T}}_{k,q}^H + \sum_{m=1}^{M} \sigma_{\mathrm{R},m}^2 \boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\mathcal{G}}_{k,m}^H + \sigma_{\mathrm{D},k}^2\mathbf{I}_{d_k}. \quad (15)$$

To compute the expectations in (14), we shall make use of the following lemma.

Lemma 1: Consider a matrix-variate Gaussian distributed $\Delta\mathbf{X}$, i.e., $\mathbf{\Delta X} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Sigma}_{\mathrm{X}} \otimes \mathbf{\Psi}_{\mathrm{X}}^T\right)$, which can be expressed as $\Delta\mathbf{X} = \mathbf{\Sigma}_{\mathrm{X}}^{1/2}\Delta\mathbf{X}^{\mathrm{W}}\mathbf{\Psi}_{\mathrm{X}}^{1/2}$, where the elements of $\Delta\mathbf{X}^{\mathrm{W}}$ are i.i.d. zero-mean complex Gaussian distributed with unit variance. Given any Hermitian matrix $\mathbf{R}$, the following identity holds: $\mathbb{E}_{\Delta\mathbf{X}}\left\{\Delta\mathbf{X}\mathbf{R}\Delta\mathbf{X}^H\right\} = \mathrm{Tr}\left(\mathbf{R}\mathbf{\Psi}_{\mathrm{X}}\right)\mathbf{\Sigma}_{\mathrm{X}}$.

Then applying Lemma 1 to $\mathcal{I}_1$ in (14), we can obtain

$$\begin{aligned}\mathcal{I}_1 &= \mathbf{u}_{k,l}^H\,\mathbb{E}\left\{\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\Delta\mathbf{G}_{k,m}^H\right\}\mathbf{u}_{k,l}\\ &= \mathbf{u}_{k,l}^H\,\mathrm{Tr}\left(\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\mathbf{\Psi}_{\mathrm{G}_{k,m}}\right)\mathbf{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l} \quad (16)\end{aligned}$$

Similarly, $\mathcal{I}_2$ and $\mathcal{I}_3$ can be simplified as

$$\mathcal{I}_2 = \mathbf{u}_{k,l}^H\,\mathrm{Tr}\left(\mathbf{F}_q\mathbf{F}_q^H\mathbf{\Psi}_{\mathrm{H}_{m,q}}\right)\boldsymbol{\mathcal{G}}_{k,m}\mathbf{\Sigma}_{\mathrm{H}_{m,q}}\boldsymbol{\mathcal{G}}_{k,m}^H\mathbf{u}_{k,l} \quad (17)$$

$$\mathcal{I}_3 = \sigma_{\mathrm{R},m}^2\mathbf{u}_{k,l}^H\,\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\mathbf{\Psi}_{\mathrm{G}_{k,m}}\right)\mathbf{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}. \quad (18)$$

Based on (16)–(18), the averaged MSE (14) is equivalent to

$$\begin{aligned}\overline{\varepsilon}_{k,l} = &\ \mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^H\mathbf{u}_{k,l} - 2\Re\left\{\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k}\mathbf{e}_{k,l}\right\}\\ &+ \mathbf{u}_{k,l}^H\mathbf{R}_k\mathbf{u}_{k,l} + 1 + \mathbf{u}_{k,l}^H\mathbf{\Omega}_k\mathbf{u}_{k,l} \quad (19)\end{aligned}$$

where

$$\begin{aligned}\mathbf{\Omega}_k = &\sum_{q=1}^{K}\sum_{m=1}^{M}\Big(\mathrm{Tr}\left(\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\mathbf{\Psi}_{\mathrm{G}_{k,m}}\right)\mathbf{\Sigma}_{\mathrm{G}_{k,m}}\\ &+ \mathrm{Tr}\left(\mathbf{F}_q\mathbf{F}_q^H\mathbf{\Psi}_{\mathrm{H}_{m,q}}\right)\boldsymbol{\mathcal{G}}_{k,m}\mathbf{\Sigma}_{\mathrm{H}_{m,q}}\boldsymbol{\mathcal{G}}_{k,m}^H\Big)\\ &+ \sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\,\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\mathbf{\Psi}_{\mathrm{G}_{k,m}}\right)\mathbf{\Sigma}_{\mathrm{G}_{k,m}}. \quad (20)\end{aligned}$$

$$\begin{aligned}\varepsilon_{k,l} = &\left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k} + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,k}\mathbf{F}_k + + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,k}\mathbf{F} - \mathbf{e}_{k,l}^T\right\|^2 + \sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m} + \mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\mathbf{W}_m\right\|^2\\ &+ \sum_{q=1,q\neq k}^{K}\left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,q} + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,q}\mathbf{F}_q\right\|^2 + \sigma_{\mathrm{D},k}^2\left\|\mathbf{u}_{k,l}\right\|^2. \quad (8)\end{aligned}$$

$$\overline{\varepsilon}_{k,l} = \mathbf{u}_{k,l}^H \left(\boldsymbol{\mathcal{T}}_{k,k} \boldsymbol{\mathcal{T}}_{k,k}^H + \mathbf{R}_k \right) \mathbf{u}_{k,l} - 2\Re\left\{ \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{T}}_{k,k} \mathbf{e}_{k,l} \right\} + 1 + \sum_{q=1}^{K} \sum_{m=1}^{M} \underbrace{\mathbb{E}\left\{ \mathbf{u}_{k,l}^H \Delta\mathbf{G}_{k,m} \boldsymbol{\mathcal{W}}_{m,q} \mathbf{F}_q \mathbf{F}_q^H \boldsymbol{\mathcal{W}}_{m,q}^H \Delta\mathbf{G}_{k,m}^H \mathbf{u}_{k,l} \right\}}_{\mathcal{I}_1}$$

$$+ \sum_{q=1}^{K} \sum_{m=1}^{M} \underbrace{\mathbb{E}\left\{ \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_{k,m} \Delta\mathbf{H}_{m,q} \mathbf{F}_q \mathbf{F}_q^H \Delta\mathbf{H}_{m,q}^H \boldsymbol{\mathcal{G}}_{k,m}^H \mathbf{u}_{k,l} \right\}}_{\mathcal{I}_2} + \sum_{m=1}^{M} \underbrace{\sigma_{\mathrm{R},m}^2 \mathbb{E}\left\{ \mathbf{u}_{k,l}^H \Delta\mathbf{G}_{k,m} \mathbf{W}_m \mathbf{W}_m^H \Delta\mathbf{G}_{k,m}^H \mathbf{u}_{k,l} \right\}}_{\mathcal{I}_3}. \quad (14)$$

After some careful inspection, we find that $\overline{\varepsilon}_{k,l}$ is convex with respect to each of $\mathbf{F}$, $\mathbf{W}$ and $\mathbf{U}$, although not jointly convex in all the design variables.

The averaged relay power $\overline{P}_{\mathrm{R},m}$ can also be obtained based on Lemma 1, specifically:

$$\overline{P}_{\mathrm{R},m} = \sum_{k=1}^{K} \Bigg(\mathrm{Tr}(\mathbf{F}_k^H \hat{\mathbf{H}}_{m,k}^H \mathbf{W}_m^H \mathbf{W}_m \hat{\mathbf{H}}_{m,k} \mathbf{F}_k) + \mathrm{Tr}\left(\mathbf{F}_k \mathbf{F}_k^H \boldsymbol{\Psi}_{\mathrm{H}_{m,k}} \right) \mathrm{Tr}\left(\mathbf{W}_m^H \mathbf{W}_m \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}} \right) \Bigg) + \sigma_{\mathrm{R},m}^2 \mathrm{Tr}\left(\mathbf{W}_m \mathbf{W}_m^H \right). \quad (21)$$

The convexity of $\overline{P}_{\mathrm{R},m}$ in each of $\mathbf{F}$ and $\mathbf{W}$ is immediate.

Note that the inner point-wise maximization in (12a) preserves the partial convexity of $\overline{\varepsilon}_{k,l}$. Hence, due to the partial convexity of (12a) and (12b), problem (12) has a so-called *block multi-convex* structure, i.e., the problem is convex in each block of variables, although in general not jointly convex in all the variables. Motivated by this property, we propose an algorithmic solution for the joint transceiver optimization based on the *block coordinate update approach*, which updates the three blocks of design variables one at a time while fixing the values of the remaining blocks. In this way, the original problem is decomposed into three sub-problems, where each amounts to updating $\mathbf{F}$, $\mathbf{W}$ and $\mathbf{U}$ separately. We iterate through the solutions to each three sub-problems in a Gauss-Seidel manner, and repeat the procedure until a termination criterion is reached. Next, we derive the solution to each of these sub-problems.

1) Receive Filter Design: It can be observed in (14) that $\overline{\varepsilon}_{k,l}$ only depends on the corresponding beamforming vector $\mathbf{u}_{k,l}$, while the constraints (12b) and (12c) do not involve $\mathbf{u}_{k,l}$. Therefore, for fixed $\mathbf{F}$ and $\mathbf{W}$, the optimal $\mathbf{u}_{k,l}$ for $k \in \mathcal{K}$ and $l \in \mathcal{D}_k$ in (12a) can be obtained independently and in parallel for different values of (k, l) by solving the following equation:

$$\frac{\partial \overline{\varepsilon}_{k,l}}{\partial \mathbf{u}_{k,l}^*} = \mathbf{0}. \quad (22)$$

The resulting optimal solution of (22) is the Wiener filter

$$\mathbf{u}_{k,l} = \left(\boldsymbol{\mathcal{T}}_{k,k} \boldsymbol{\mathcal{T}}_{k,k}^H + \mathbf{R}_k + \boldsymbol{\Omega}_k \right)^{-1} \boldsymbol{\mathcal{T}}_{k,k} \mathbf{e}_{k,l}. \quad (23)$$

2) Source Precoder Design: We then solve for the source precoders $\mathbf{F}$ while keeping $\mathbf{W}$ and $\mathbf{U}$ fixed. For better exposition of our solution, and after some matrix manipulations, we can rewrite (12) explicitly in terms of $\mathbf{F}$, as shown in (24) on top of the next page, where

$$a_3^{k,l} \triangleq \mathbf{u}_{k,l}^H \Big[\sum_{m=1}^{M} \sigma_{\mathrm{R},m}^2 \Big(\mathrm{Tr}\left(\mathbf{W}_m \mathbf{W}_m^H \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \right) \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} + \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\mathcal{G}}_{k,m}^H \Big) + \sigma_{\mathrm{D},k}^2 \mathbf{I}_{N_{\mathrm{D},k}} \Big] \mathbf{u}_{k,l} + 1 \quad (25)$$

$\mathbf{E}_{k,l} \triangleq \mathbf{e}_{k,l} \mathbf{e}_{k,l}^T$ and $\eta_{\mathrm{R},m} \triangleq \rho_m P_{\mathrm{R}} - \sigma_{\mathrm{R},m}^2 \mathrm{Tr}\left(\mathbf{W}_m \mathbf{W}_m^H \right)$. The solution to the problem (24) is not straightforward and we therefore need transform it into a more tractable form. To this end, we introduce new variables $\mathbf{f}_k \triangleq \mathrm{vec}\left(\mathbf{F}_k \right) \in \mathbb{C}^{N_{\mathrm{S},k}^2 \times 1}$ and define the following matrices and vectors that are independent of $\mathbf{f}_k, \forall k \in \mathcal{K}$

$$\mathbf{A}_{1,q}^{k,l} \triangleq \sum_{m=1}^{M} \mathbf{I}_{d_k} \otimes \Big(\sum_{n=1}^{M} \boldsymbol{\mathcal{W}}_{m,q}^H \boldsymbol{\mathcal{U}}_{k,m}^H \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n} \boldsymbol{\mathcal{W}}_{n,q} + \mathrm{Tr}\left(\mathbf{u}_{k,l}^H \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l} \right) \boldsymbol{\mathcal{W}}_{m,k}^H \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \boldsymbol{\mathcal{W}}_{m,k} + \mathrm{Tr}\left(\mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}} \boldsymbol{\mathcal{G}}_{k,m}^H \mathbf{u}_{k,l} \right) \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \Big) \quad (26)$$

$$\mathbf{a}_2^{k,l} \triangleq \mathrm{vec}\left(\sum_{m=1}^{M} \boldsymbol{\mathcal{W}}_{m,k}^H \boldsymbol{\mathcal{U}}_{k,m}^H \mathbf{E}_{k,l} \right) \quad (27)$$

$$\mathbf{A}_{4,k}^{m} \triangleq \mathbf{I}_{d_k} \otimes \Big(\boldsymbol{\mathcal{W}}_{m,k}^H \boldsymbol{\mathcal{W}}_{m,k} + \mathrm{Tr}\left(\mathbf{W}_m^H \mathbf{W}_m \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}} \right) \boldsymbol{\Psi}_{\mathrm{H}_{m,k}} \Big) \quad (28)$$

where $\boldsymbol{\mathcal{U}}_{k,m} \triangleq \mathbf{U}_k^H \hat{\mathbf{G}}_{k,m}$. It is not difficult to verify that $\mathbf{A}_{1,q}^{k,l}$ and $\mathbf{A}_{4,k}^{m}$ are positive definite matrices. Then we use the following identities, $\mathrm{Tr}\left(\mathbf{A}^H \mathbf{B} \mathbf{A} \right) = \mathrm{vec}\left(\mathbf{A} \right)^H \left(\mathbf{I} \otimes \mathbf{B} \right) \mathrm{vec}\left(\mathbf{A} \right)$ and $\mathrm{Tr}\left(\mathbf{A}^H \mathbf{B} \right) = \mathrm{vec}\left(\mathbf{B} \right)^H \mathrm{vec}\left(\mathbf{A} \right)$ to transform both the objective and the constraints in (24) into quadratic expressions of $\mathbf{f}_k$, and finally obtain the equivalent formulation

$$\min_{\mathbf{f}_1, \cdots, \mathbf{f}_K, t} \quad t \quad (29a)$$

$$\mathrm{s.t.} \quad \sum_{q=1}^{K} \mathbf{f}_q^H \mathbf{A}_{1,q}^{k,l} \mathbf{f}_q - 2\Re\left\{ \mathbf{f}_k^H \mathbf{a}_2^{k,l} \right\} + a_3^{k,l} \leq \frac{t}{\kappa_{k,l}} \quad \forall k \in \mathcal{K}, l \in \mathcal{D}_k \quad (29b)$$

$$\sum_{k=1}^{K} \mathbf{f}_k^H \mathbf{A}_{4,k}^{m} \mathbf{f}_k \leq \eta_{\mathrm{R},m}, \quad \forall m \in \mathcal{M} \quad (29c)$$

$$\mathbf{f}_k^H \mathbf{f}_k \leq P_{\mathrm{S},k}^{\max}, \quad \forall k \in \mathcal{K} \quad (29d)$$

where t is an auxiliary variable. Problem (29) by definition is a convex separable inhomogeneous quadratic-constrained

$$\begin{aligned}
\min_{\mathbf{F}} \max_{\forall k\in\mathcal{K},l\in\mathcal{D}_k} \kappa_{k,l} \sum_{q=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^H \boldsymbol{\mathcal{W}}_{m,q}^H \boldsymbol{\mathcal{U}}_{k,m}^H \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n} \boldsymbol{\mathcal{W}}_{n,q} \mathbf{F}_q\right) - \sum_{m=1}^{M} 2\Re\left\{\mathrm{Tr}\left(\mathbf{E}_{k,l}\boldsymbol{\mathcal{U}}_{k,m}\boldsymbol{\mathcal{W}}_{m,k}\mathbf{F}_k\right)\right\} + a_3^{k,l} \\
+ \sum_{q=1}^{K}\sum_{m=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^H \boldsymbol{\mathcal{W}}_{m,k}^H \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \boldsymbol{\mathcal{W}}_{m,k}\mathbf{F}_q\right) \mathrm{Tr}\left(\mathbf{u}_{k,l}^H \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l}\right) + \sum_{q=1}^{K}\sum_{m=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^H \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \mathbf{F}_q\right) \mathrm{Tr}\left(\mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}} \boldsymbol{\mathcal{G}}_{k,m}^H \mathbf{u}_{k,l}\right) \\
\text{s.t.} \quad \sum_{k=1}^{K} \mathrm{Tr}\left(\mathbf{F}_k^H\left(\hat{\mathbf{H}}_{m,k}^H \mathbf{W}_m^H \mathbf{W}_m \hat{\mathbf{H}}_{m,k} + \mathrm{Tr}\left(\mathbf{W}_m^H\mathbf{W}_m\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}\right)\boldsymbol{\Psi}_{\mathrm{H}_{m,k}}\right)\mathbf{F}_k\right) \le \eta_{\mathrm{R},m},\ \forall m\in\mathcal{M} \\
\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right) \le P_{\mathrm{S},k}^{\max},\ \forall k\in\mathcal{K} \qquad (24)
\end{aligned}$$

linear program (QCLP). This class of optimization problems can be equivalently recast as a standard second-order cone program (SOCP), which can be efficiently solved by a generic external optimization tool, e.g., `SeduMi` [7]. Due to the space limitation, the details of the transformation to the SOCP are omitted.

3) Relay Precoder Design: To solve for the relay precoders, we concatenate the design variable into a single column vector as

$$\mathbf{w} \triangleq \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_M \end{bmatrix} = \begin{bmatrix} \mathrm{vec}\left(\mathbf{W}_1\right) \\ \vdots \\ \mathrm{vec}\left(\mathbf{W}_M\right) \end{bmatrix}. \qquad (30)$$

For notational convenience, we also introduce $\boldsymbol{\mathcal{H}}_{m,k} \triangleq \hat{\mathbf{H}}_{m,k}\mathbf{F}_k$ and define the following intermediate quantities, which are independent of $\mathbf{w}$:

$$\left[\mathbf{B}_1^{k,l}\right]_{m,n} = \sum_{q=1}^{K}\left[\left(\boldsymbol{\mathcal{H}}_{m,q}^*\boldsymbol{\mathcal{H}}_{n,q}^T\right)\otimes\left(\boldsymbol{\mathcal{U}}_{k,m}^H\mathbf{E}_{k,l}\boldsymbol{\mathcal{U}}_{k,n}\right)\right] \qquad (31)$$

$$\mathbf{b}_{2,m}^{k,l} \triangleq \mathrm{vec}\left(\boldsymbol{\mathcal{U}}_{k,m}^H\mathbf{E}_{k,l}\boldsymbol{\mathcal{H}}_{m,k}^H\right) \qquad (32)$$

$$\begin{aligned}
\mathbf{B}_{3,m}^{k,l} \triangleq \sum_{q=1}^{K}\Big[&\mathrm{Tr}\left(\mathbf{u}_{k,l}^H\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}\right)\boldsymbol{\mathcal{H}}_{m,q}^*\boldsymbol{\mathcal{H}}_{m,q}^T\otimes\boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \\
&+\mathrm{Tr}\left(\mathbf{F}_q^H\boldsymbol{\Psi}_{\mathrm{H}_{m,q}}\mathbf{F}_q\right)\boldsymbol{\Sigma}_{\mathrm{H}_{m,q}}^T\otimes\boldsymbol{\mathcal{U}}_{k,m}^H\mathbf{E}_{k,l}\boldsymbol{\mathcal{U}}_{k,m}\Big] \\
&+\sigma_{\mathrm{R},m}^2\,\mathrm{Tr}\left(\mathbf{u}_{k,l}^H\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}\right)\mathbf{I}_{N_{\mathrm{R},m}}\otimes\boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \\
&+\sigma_{\mathrm{R},m}^2\mathbf{I}_{N_{\mathrm{R},m}}\otimes\left(\boldsymbol{\mathcal{U}}_{k,m}^H\mathbf{E}_{k,l}\boldsymbol{\mathcal{U}}_{k,m}\right) \qquad (33)
\end{aligned}$$

$$b_4^{k,l} \triangleq \sigma_{\mathrm{D},k}^2\left\|\mathbf{u}_{k,l}\right\|^2 + 1 \qquad (34)$$

$$\begin{aligned}
\mathbf{B}_{5,m} \triangleq \Big[\sigma_{\mathrm{R},m}^2\mathbf{I}_{N_{\mathrm{R},m}} + \sum_{k=1}^{K}\big(&\boldsymbol{\mathcal{H}}_{m,k}^*\boldsymbol{\mathcal{H}}_{m,k}^T \\
&+\mathrm{Tr}\left(\mathbf{F}_k\mathbf{F}_k^H\boldsymbol{\Psi}_{\mathrm{H}_{m,k}}\right)\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}^T\big)\Big]\otimes\mathbf{I}_{N_{\mathrm{R},m}} \qquad (35)
\end{aligned}$$

where $\mathbf{B}_1^{k,l}$ is a block matrix with the $(m,n)^{\text{th}}$ block as defined above.

Using the identities $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\mathbf{C}\mathbf{D}^H\right) = \mathrm{vec}\left(\mathbf{A}\right)^H\left(\mathbf{D}^T\otimes\mathbf{B}\right)\mathrm{vec}\left(\mathbf{C}\right)$, $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\mathbf{A}\right) = \mathrm{vec}\left(\mathbf{A}\right)^H\left(\mathbf{I}\otimes\mathbf{B}\right)\mathrm{vec}\left(\mathbf{A}\right)$ and $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\right) = \mathrm{vec}\left(\mathbf{B}\right)^H\mathrm{vec}\left(\mathbf{A}\right)$, we can formulate the optimization problem as follows:

$$\min_{\mathbf{w},t} \quad t \qquad (36a)$$

$$\begin{aligned}
\text{s.t.} \quad &\mathbf{w}^H\mathbf{B}_1^{k,l}\mathbf{w} - \sum_{m=1}^{M}2\Re\left\{\mathbf{w}_m^H\mathbf{b}_{2,m}^{k,l}\right\} \\
&+\sum_{m=1}^{M}\mathbf{w}_m^H\mathbf{B}_{3,m}^{k,l}\mathbf{w}_m + b_4^{k,l} \le \frac{t}{\kappa_{k,l}},\ \forall l\in\mathcal{D}_k, k\in\mathcal{K} \qquad (36b)
\end{aligned}$$

$$\mathbf{w}_m^H\mathbf{B}_{5,m}\mathbf{w}_m \le \rho_m P_{\mathrm{R}},\ \forall m\in\mathcal{M}. \qquad (36c)$$

Since $\mathbf{B}_1^{k,l}$, $\mathbf{B}_{3,m}^{k,l}$, and $\mathbf{B}_{5,m}$ are all positive definite matrices, (36) is also a convex separable inhomogeneous QCLP.

The overall algorithm, with the following convergence property, iteratively solves sub-problems (22), (29) and (36) until a termination criterion is reached.

Proposition 1: The objective (12a) is monotonically non-increasing after each sub-problem is solved, and lower bounded by zero. Therefore, the iterative algorithm is convergent.

IV. Simulation Results

This section presents Monte Carlo simulation results to verify the robustness of the proposed algorithm against the channel uncertainties. In all the experiments, we consider a scenario with $K = 2$ source/destination pairs and $M = 2$ relays, where all nodes are equipped with 3 antennas, i.e., $N_{\mathrm{S},k} = N_{\mathrm{R},m} = N_{\mathrm{D},k} = 3$, $\forall k, m$. Each source transmits 2 independent data streams to its corresponding destination, i.e., $d_k = 2$, and equal weights $\kappa_{k,l}$ $\forall k, l$ are assigned to all the data streams. Equal noise variance $\sigma_{\mathrm{D},k}^2 = \sigma_{\mathrm{R},m}^2$ and unit maximum source and relay power $P_{\mathrm{S},k}^{\max} = \rho_m P_{\mathrm{R}}^{\max} = 1$, $\forall k, m$ are assumed. The channels are assumed to be flat-fading, with the coefficients given by zero-mean unit-variance complex circular Gaussian random variables. The signal-to-noise ratios (SNRs) at the relays and the destinations are defined as $\mathrm{SNR}_{\mathrm{R},m} \triangleq \frac{P_{\mathrm{S}}^{\max}}{N_{\mathrm{R},m}\sigma_{\mathrm{R},m}^2}$ and $\mathrm{SNR}_{\mathrm{D},k} \triangleq \frac{P_{\mathrm{R}}^{\max}}{N_{\mathrm{D},k}\sigma_{\mathrm{D},k}^2}$, respectively. The channel correlation matrices in (10) and (11) are obtained by the widely-employed exponential model [8]. Specifically, they are given by $\left[\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}\right]_{i,j} = \left[\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\right]_{i,j} = \alpha^{|i-j|}$ and $\boldsymbol{\Psi}_{\mathrm{H}_{m,k}} = \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} = \sigma_e^2\beta^{|i-j|}$, $i, j \in \{1, 2, 3\}$, where α and β are the correlation coefficients, and σ_e^2 denotes the variance of the CSI errors. The available chan-

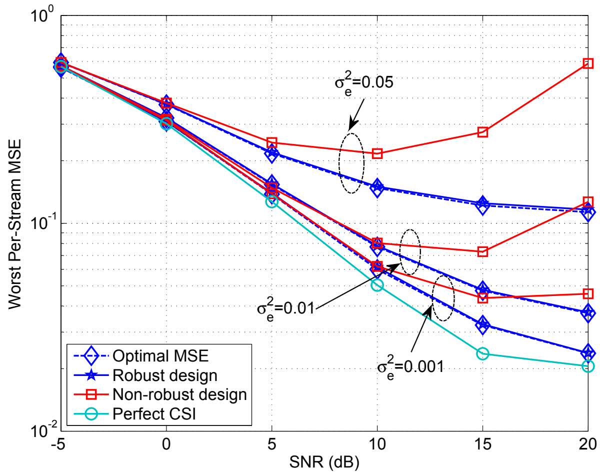

Fig. 2. Worst per-stream MSE of different approaches versus SNR under different values of CSI error variance (Channel correlation factors $\alpha = \beta = 0.5$; $\mathrm{SNR}_{\mathrm{R},m} = \mathrm{SNR}_{\mathrm{D},k} = \mathrm{SNR}$).

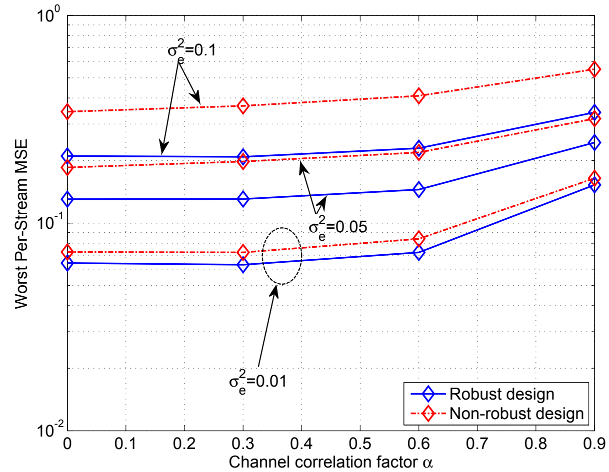

Fig. 3. Worst per-stream MSE of different approaches versus channel correlation factor α under different values of CSI error variance (β is fixed to 0.45; $\mathrm{SNR}_{\mathrm{R},m} = \mathrm{SNR}_{\mathrm{D},k} = 10\mathrm{dB}$).

nel estimates, i.e., $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$ are generated according to $\hat{\mathbf{H}}_{m,k} \sim \mathcal{CN}\left(\mathbf{0}_{N_{\mathrm{R},m}\times N_{\mathrm{S},k}}, \frac{1-\sigma_e^2}{\sigma_e^2}\mathbf{\Sigma}_{\mathrm{H}_{m,k}} \otimes \mathbf{\Psi}_{\mathrm{H}_{m,k}}^T\right)$ and $\hat{\mathbf{G}}_{k,m} \sim \mathcal{CN}\left(\mathbf{0}_{N_{\mathrm{D},k}\times N_{\mathrm{R},m}}, \frac{1-\sigma_e^2}{\sigma_e^2}\mathbf{\Sigma}_{\mathrm{G}_{k,m}} \otimes \mathbf{\Psi}_{\mathrm{G}_{k,m}}^T\right)$, respectively, such that the elements of the true channels have unit variances. We compare the proposed robust design approach in Section III with i) the non-robust approach, obtained by neglecting the effects of CSI errors in the robust approach in Section III, ii) the perfect CSI scenario, obtained by using the true channels in place of the estimated channels in the robust approach in Section III and dropping the CSI errors, and iii) the theoretical MSE given by the optimal objective value of (12) after optimization (labelled as "optimal MSE").

In Fig. 2, the worst per-stream MSE is shown as a function of the SNR with different variances of CSI errors. It is observed that the proposed algorithm achieves better robustness against the CSI errors than the non-robust design approach in all cases. Especially, the performance gains are significant in median and high SNR regime. For the non-robust design, degradations are observed because the MSE discovered at high SNRs is dominated by the interference, rather than the noise. Therefore, the relays are confined to relatively low transmit power in order to control the interference. We also observe that the "optimal MSE" and our simulation results tally well, which justifies the approximations invoked in calculating the MSE in (8). We thereby conclude that the proposed transceiver design is capable of achieving an improved robustness against CSI errors when compared to the non-robust approach.

The effects of channel correlation on the MSE performance of different approaches are investigated in Fig. 3. It can be observed that the robust design shows consistent performance gains over its non-robust counterpart with different α and σ_e^2. The discrepancies between the two approaches become less significant with an increase in α because the achievable *spatial multiplexing* is reduced by a higher channel correlation, and therefore the robust design can only attain a limited performance improvement.

V. Conclusion

This paper studied the joint transceiver optimization in multiuser MIMO relay networks in the presence of statistical CSI errors. A robust design approach was formulated to minimize the worst per-stream MSE subject to source and relay transmit power constraints. To overcome the non-convexity of the problem, we then proposed an algorithmic solution by adopting an iterative block coordinate update approach, where the original problem is decomposed into three convex sub-problems, and solved iteratively until the objective value converges. Results of simulation experiments over correlated Rayleigh flat-fading channels confirm the improved robustness of the proposed robust design approach against the CSI errors.

References

[1] Y. Liu and A. P. Petropulu, "On the sumrate of amplify-and-forward relay networks with multiple source-destination pairs," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3732–3742, Nov. 2011.

[2] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO-relay networks with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5473–5484, Nov. 2011.

[3] K. T. Truong, P. Sartori, and R. W. Heath, "Cooperative algorithms for MIMO amplify-and-forward relay networks," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1272–1287, Mar. 2013.

[4] M. R. A. Khandaker and Y. Rong, "Interference MIMO relay channel: Joint power control and transceiver-relay beamforming," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6509–6518, Dec. 2012.

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

[6] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[7] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11, no. 1-4, pp. 625–653, Jan. 1999.

[8] L. Musavian, M. R. Nakhi, M. Dohler, and A. H. Aghvami, "Effect of channel uncertainty on the mutual information of MIMO fading channels," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2798–2806, Sep. 2007.