Joint Optimization of Transceiver Matrices for MIMO-Aided Multiuser AF Relay Networks: Improving the QoS in the Presence of CSI Errors

Jiaxin Yang, *Student Member, IEEE*, Benoit Champagne, *Senior Member, IEEE*, Yulong Zou, *Senior Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***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 realistic imperfect channel state information (CSI) of all the source–relay and relay–destination links, we propose a robust optimization framework for the joint design of the source transmit precoders (TPCs), relay AF matrices and receive filters. Specifically, two well–known CSI error models are considered, namely, the *statistical* and the *norm-bounded* error models. We commence by considering the problem of minimizing the maximum per-stream mean square error (MSE) subject to the source and relay power constraints (min–max problem). Then, the statistically robust and worst-case robust versions of this problem, which take into account the statistical and norm-bounded CSI errors, respectively, are formulated. Both of the resultant optimization problems are nonconvex (semi-infinite in the worst-case robust design). Therefore, algorithmic solutions having proven convergence and tractable complexity are proposed by resorting to the iterative block coordinate update approach along with matrix transformation and convex conic optimization techniques. We then consider the problem of minimizing the maximum per-relay power subject to the quality-of-service (QoS) constraints for each stream and the source power constraints (QoS problem). Specifically, an efficient initial feasibility search algorithm is proposed based on the relationship between the feasibility check and the min–max problems. Our simulation results show that the proposed joint transceiver design is capable of achieving improved robustness against different types of CSI errors when compared with non-robust approaches.**

***Index Terms*—Amplify-and-forward (AF) relaying, channel state information (CSI) error, convex optimization, multiple-input multiple-output (MIMO), multiuser, robust transceiver design.**

Manuscript received May 29, 2014; revised December 24, 2014; accepted February 16, 2015. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and in part by InterDigital Canada. The review of this paper was coordinated by Prof. C.-X. Wang.

J. Yang and B. Champagne are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: jiaxin.yang@mail.mcgill.ca; benoit.champagne@mcgill.ca).

Y. Zou is with the School of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: yulong.zou@njupt.edu.cn).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2015.2410759

I. Introduction

Cooperative relaying [1] is capable of improving the communication link between the source and destination nodes, in the context of wireless standards such as those of the Long-Term Evolution Advanced [2], Worldwide Interoperability for Microwave Access (WiMAX) [3], and fifth-generation networks [4]. Relaying strategies may be classified as amplify-and-forward (AF) and decode-and-forward (DF) techniques. The AF relaying technique imposes lower signal processing complexity and latency; therefore, it is preferred in many operational applications [5] and is the focus of our attention in this paper.

Recently, multiple-input–multiple-output (MIMO) AF relaying designed for multiuser networks has attracted considerable interest [6]–[11]. In typical wireless multiuser networks, the amount of spectral resources available to each user decreases with an increase in the density of users sharing the channel, hence imposing a degradation on the quality of service (QoS) of each user. MIMO AF relaying is emerging as a promising technique of mitigating this fundamental limitation. By exploiting the so-called *distributed spatial multiplexing* [5] at the multiantenna assisted relays, it allows multiple source/destination pairs to communicate concurrently at an acceptable QoS over the same physical channel [5]. The relay matrix optimization has been extensively studied in a single-antenna assisted multiuser framework, under different design criteria (see, e.g., [6]–[10]), where each source/destination is equipped with a single antenna. In general, finding the optimal relay matrix in these design approaches is deemed challenging because the resultant optimization problems are typically nonconvex. Hence, existing algorithms have relied on convex approximation techniques, e.g., semi-definite relaxation (SDR) [9], [10] and second-order cone programming (SOCP) approximation [7], [8], in order to obtain approximate solutions to the original design problems.

Again, the given contributions focus on single-antenna multiuser networks. However, wireless standards aim for the promotion of mobile broadband multimedia services with an enhanced data rate and QoS, where parallel streams corresponding to different service types can be transmitted simultaneously by each source using multiple antennas [11]. This aspiration has led to a strong interest in the study of cooperative relaying in a MIMO multiuser framework, where multiple antennas are employed by all the sources ($\mathtt{S}$), relays ($\mathtt{R}$), and destinations (D). The joint transceiver design[1] is more challenging than the relay matrix design of the single-antenna scenario, but it provides further performance benefits. Prior contributions [6]–[10], [12], [13] are therefore not readily extendable to this more general case. At the time of this writing, the literature of the joint transceiver design for MIMO multiuser relaying networks is still limited. To be specific, in [14], global objective functions such as the sum power of the interference received at all the destinations and the sum mean square error (MSE) of all the estimated data streams are minimized by adopting the *alternating minimization* approach of [15], where only a single design variable is updated at each iteration based on the SDR technique of [16]. However, the use of global objective functions is not readily applicable to multimedia applications supporting several types of services, each characterized by a specific QoS requirement. To overcome this problem, in [17], the objective of minimizing the total source and relay power subject to a minimum signal-to-noise-plus-interference ratio (SINR) requirement for each S−D link is considered. To this end, a two-level iterative algorithm is proposed, which also involves SDR. Since the main goal of [17] was that of achieving a high *spatial diversity* gain to improve the attainable transmission integrity, the number of data streams transmitted by each source in this setting is limited to one [17].

The efficacy of the joint transceiver design in [14] and [17] relies on the idealized simplifying assumption of perfect channel state information (CSI) for all the S−R and R−D links. In practice, acquiring perfect or even accurate channel estimates at a central processing node is quite challenging. This is primarily due to the combined effects of various sources of imperfections, such as the affordable channel estimation complexities and the limited quantized feedback and feedback delays [18], [19]. The performance of the previous methods may hence be substantially degraded in the presence of realistic CSI errors. In view of this, robust transceiver designs, which explicitly take into account the effects of CSI errors, are highly desirable. Depending on the assumptions concerning the CSI errors, robust designs fall into two major categories, namely, *statistically* robust [18] and *worst-case* robust designs [19]. The former class models the CSI errors as random variables with certain statistical distributions (e.g., Gaussian distributions), and robustness is achieved by optimizing the average performance over all the CSI error realizations; the latter family assumes that the CSI errors belong to some predefined bounded uncertainty regions, such as norm-bounded regions, and optimizes the worst-case performance for all the possible CSI errors within the region.

As a further contribution, we study the joint transceiver design in a more general MIMO multiuser relay network, where multiple S−D pairs communicate with the assistance of multiple relays, and each source transmits multiple parallel data streams to its corresponding destination. Assuming realistic imperfect CSI for all the S−R and R−D links, we propose a new robust optimization framework for minimizing the maximum per-stream MSE subject to the source and relay power constraints, which is termed as the *min–max* problem. In the proposed framework, we aim for solving both the *statistically* robust and *worst-case* robust versions of the min–max problem, which take into account either the statistical CSI errors or the norm-bounded CSI errors, respectively, while maintaining tractable computational complexity. Furthermore, to strictly satisfy the QoS specifications of all the data streams, we subsequently consider the problem of minimizing the maximum per-relay power, subject to the QoS constraints of all the data streams and to the source power constraints, which is referred to as the *QoS* problem. Against this background, the main contributions of this paper are threefold.

- With the statistically robust min–max problem for the *joint* transceiver design being nonconvex, an algorithmic solution having proven convergence is proposed by invoking the iterative *block coordinate update approach* of [20] while relying on both matrix transformation and convex conic optimization techniques. The proposed iterative algorithm successively solves in a circular manner three subproblems corresponding to the source transmit precoders (TPCs), relay AF matrices, and receive filters, respectively. We show that the receive filter subproblem yields a closed-form solution, whereas the other two subproblems can be transformed to convex quadratically constrained linear programs (QCLPs). Then, each QCLP can subsequently be reformulated as a efficiently solvable SOCP.
- The worst-case robust min–max problem is both nonconvex and *semi-infinite*. To overcome these challenges, we first present a generalized version of the so-called $\mathcal{S}$ lemma given in [21], based on which each subproblem can be exactly reformulated as a semi-definite program (SDP) with only linear matrix inequality (LMI) constraints. This results in an iterative algorithmic solution involving several SDPs.
- The QoS-based transceiver optimization is more challenging than that of the min–max problem because it is difficult to find a feasible initialization. Hence, our major contribution here is to propose an efficient procedure for finding a feasible starting point for the iterative QoS-based optimization algorithm, provided that there exits one; otherwise, the procedure also returns a certificate of infeasibility.

The remainder of this paper is organized as follows. Section II introduces our system model and the modeling of CSI errors. The robust joint transceiver design problems are also formulated here. In Sections III and IV, iterative algorithms are proposed for solving the min–max problem both under the statistical and the norm-bounded CSI error models, respectively. The QoS problem is dealt with in Section V. Our numerical results are reported in Section VI. This paper is then concluded in Section VII.

Notations: Boldface uppercase (lowercase) letters represent matrices (vectors), and normal letters denote scalars. $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^{-1}$ denote the conjugate, transpose, Hermitian transpose, and inverse, respectively. $\|\cdot\|$ corresponds to the Euclidean norm of a vector, whereas $\|\cdot\|_F$ and $\|\cdot\|_S$ denote the

[1] We use "transceiver design" to collectively denote the design of the source TPCs, relay AF matrices, and receive filters.

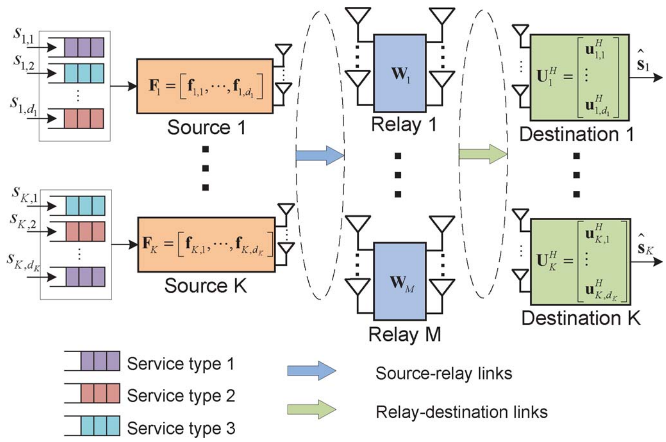

Fig. 1. MIMO multiuser multirelay one-way network with each source transmitting multiple data streams to its corresponding destination.

Frobenius norm and the spectral norm of a matrix, respectively. Furthermore, $\mathrm{Tr}(\cdot)$, $\mathrm{vec}(\cdot)$, and $\otimes$ denote the matrix trace, the vectorization, and the Kronecker product, respectively. $\mathbb{R}^{M\times N}$ and $\mathbb{C}^{M\times N}$ denote the spaces of $M\times N$ matrices with real and complex entries, respectively. $\mathbf{I}_N$ represents the $N\times N$ identity matrix. $\mathbb{E}\{\cdot\}$ denotes the statistical expectation. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a scalar, respectively.

II. System Model and Problem Formulation

We consider a MIMO multiuser relaying network, where M AF relay nodes assist the one-way communication between K S$-$D pairs, as shown in Fig. 1, where all the nodes are equipped with multiple antennas. Specifically, the kth S and D, respectively, employ $N_{\mathrm{S},k}$ and $N_{\mathrm{D},k}$ antennas for $k\in\mathcal{K}\triangleq\{1,2,\ldots,K\}$, whereas the mth R employs $N_{\mathrm{R},m}$ antennas for $m\in\mathcal{M}\triangleq\{1,\ldots,M\}$. All the relays operate under the half-duplex AF protocol, where the 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 concurrently, whereas in the second stage, the relays apply linear processing to the received signals and forward the resultant signals to all the destinations. We assume that no direct links are available between the sources and destinations due to the severe attenuation.

A narrow-band flat-fading radio propagation model is considered, where we denote the channel matrix between the kth S and the mth R by $\mathbf{H}_{m,k}\in\mathbb{C}^{N_{\mathrm{R},m}\times N_{\mathrm{S},k}}$, and the channel matrix between the mth R and the kth D by $\mathbf{G}_{k,m}\in\mathbb{C}^{N_{\mathrm{D},k}\times N_{\mathrm{R},m}}$. Let $\mathbf{s}_k\triangleq[s_{k,1},\ldots,s_{k,d_k}]^T$ denote the information symbols to be transmitted by the kth S at a given time instant, where $d_k\le\min\{N_{\mathrm{S},k},N_{\mathrm{D},k}\}$ is the number of independent data streams. The symbols are modeled as independent random variables with a zero mean and unit variance; hence, $\mathbb{E}\{\mathbf{s}_k\mathbf{s}_k^H\}=\mathbf{I}_{d_k}$. The kth S applies a linear vector of $\mathbf{f}_{k,l}\in\mathbb{C}^{N_{\mathrm{S},k}\times 1}$ for mapping the lth data stream to its $N_{\mathrm{S},k}$ antennas for $l\in\mathcal{D}_k\triangleq\{1,\ldots,d_k\}$, thus forming a linear TPC of $\mathbf{F}_k=[\mathbf{f}_{k,1},\ldots,\mathbf{f}_{k,d_k}]\in\mathbb{C}^{N_{\mathrm{S},k}\times d_k}$. The transmit power is thus given by $\mathrm{Tr}(\mathbf{F}_k\mathbf{F}_k^H)\le P_{\mathrm{S},k}^{\max}$, where $P_{\mathrm{S},k}^{\max}$ is the maximum affordable power of the kth S. Let $\mathbf{n}_{\mathrm{R},m}\in\mathbb{C}^{N_{\mathrm{R},m}\times 1}$ be the spatially white additive noise vector at the mth R, with a zero mean and covariance matrix of $\mathbb{E}\{\mathbf{n}_{\mathrm{R},m}\mathbf{n}_{\mathrm{R},m}^H\}=\sigma_{\mathrm{R},m}^2\mathbf{I}_{N_{\mathrm{R},m}}$.

After the first stage of transmission, the signal received at the mth R 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 R applies a linear matrix $\mathbf{W}_m\in\mathbb{C}^{N_{\mathrm{R},m}\times N_{\mathrm{R},m}}$ to $\mathbf{z}_{\mathrm{R},m}$ and forwards the resultant 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 all the destinations at a power of

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

Let $\mathbf{n}_{\mathrm{D},k}$ denote the spatially white additive noise vector at the kth D with a zero mean and covariance matrix of $\mathbb{E}\{\mathbf{n}_{\mathrm{D},k}\mathbf{n}_{\mathrm{D},k}^H\}=\sigma_{\mathrm{D},k}^2\mathbf{I}_{N_{\mathrm{D},k}}$. The kth D 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 for indexing the sources. To estimate the lth data stream received from its corresponding source, the kth D applies a linear vector $\mathbf{u}_{k,l}$ to the received signal, thus forming a receive filter $\mathbf{U}_k=[\mathbf{u}_{k,1},\ldots,\mathbf{u}_{k,d_k}]\in\mathbb{C}^{N_{\mathrm{D},k}\times d_k}$. Specifically, the estimated information symbols are given by $\hat{s}_{k,l}=\mathbf{u}_{k,l}^H\mathbf{y}_k$, which can be expressed as

$$\begin{aligned}\hat{s}_{k,l}=&\underbrace{\mathbf{u}_{k,l}^H\sum_{m=1}^{M}\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,k}\mathbf{f}_{k,l}s_{k,l}}_{\text{desired data stream}}\\&+\underbrace{\mathbf{u}_{k,l}^H\sum_{m=1}^{M}\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,k}\sum_{p=1,p\neq l}^{d_k}\mathbf{f}_{k,p}s_{k,p}}_{\text{interstream interference}}\\&+\underbrace{\sum_{q=1,q\neq k}^{K}\mathbf{u}_{k,l}^H\sum_{m=1}^{M}\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,q}\mathbf{F}_q\mathbf{s}_q}_{\text{interuser interference}}\\&+\underbrace{\sum_{m=1}^{M}\mathbf{u}_{k,l}^H\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{n}_{\mathrm{R},m}}_{\text{enhanced noise from relays}}+\underbrace{\mathbf{u}_{k,l}^H\mathbf{n}_{\mathrm{D},k}}_{\text{receiver noise}}.\end{aligned} \tag{5}$$

Throughout this paper, we also make the following common assumptions concerning the statistical properties of the signals.

A1) The information symbols transmitted from different S are uncorrelated, i.e., we have $\mathbb{E}\{\mathbf{s}_k\mathbf{s}_m^H\}=\mathbf{0}\ \forall k,m\in\mathcal{K}$ and $k\neq m$.

A2) The information symbols $\mathbf{s}_k$, the relay noise $\mathbf{n}_{\mathrm{R},m}$, and the receiver noise $\mathbf{n}_{\mathrm{D},l}$ are mutually statistically independent $\forall k, l \in \mathcal{K}$ and $m \in \mathcal{M}$.

A. QoS Metric

We adopt the MSE as the QoS metric for each estimated data stream. The major advantage of using the MSE is to make our design problem tractable, which has been well justified in the AF relay matrix design literature [22], [23] and in the references therein. In fact, the links between the MSE and other classic criteria such as the bit error rate (BER) and the SINR have been well established in [22], [24]. Specifically, it has been shown that an improvement in MSE will naturally lead to a reduced BER.

The MSE of the lth estimated data stream received at the kth D is defined as

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

Substituting (5) into (6), and using assumptions A1 and A2, we obtain

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

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

B. CSI Error Model

In typical relaying scenarios, the CSI of both the S−R and R−D links, which is available at the central processing node, is contaminated by channel estimation errors and by the quantized feedback, and is outdated due to feedback delays. To model these CSI errors, let us characterize the true but unknown channels as

$$\mathbf{H}_{m,k} = \hat{\mathbf{H}}_{m,k} + \Delta\mathbf{H}_{m,k}, \mathbf{G}_{k,m} = \hat{\mathbf{G}}_{k,m} + \Delta\mathbf{G}_{k,m} \tag{8}$$

where $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$, respectively, denote the estimated S−R and R−D channels, whereas $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ capture the corresponding *channel uncertainties* [8], [9]. In what follows, we consider two popular techniques of modeling the channel uncertainties.

1) Statistical Error Model: In this model, we assume that the elements of $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ are zero-mean complex Gaussian random variables. Specifically, based on the Kronecker model [18], [25], they can, in general, be written as

$$\Delta\mathbf{H}_{m,k} = \mathbf{\Sigma}_{\mathrm{H}_{m,k}}^{1/2}\Delta\mathbf{H}_{m,k}^{\mathrm{W}}\mathbf{\Psi}_{\mathrm{H}_{m,k}}^{1/2} \tag{9}$$

$$\Delta\mathbf{G}_{k,m} = \mathbf{\Sigma}_{\mathrm{G}_{k,m}}^{1/2}\Delta\mathbf{G}_{k,m}^{\mathrm{W}}\mathbf{\Psi}_{\mathrm{G}_{k,m}}^{1/2} \tag{10}$$

TABLE I
EQUIVALENT NOTATIONS USED IN THE SUBSEQUENT ANALYSIS

Notations	Definitions
$\boldsymbol{\mathcal{G}}_{k,m}$	$\hat{\mathbf{G}}_{k,m}\mathbf{W}_m$
$\boldsymbol{\mathcal{W}}_{m,k}$	$\mathbf{W}_m\hat{\mathbf{H}}_{m,k}$
$\boldsymbol{\mathcal{U}}_{k,m}$	$\mathbf{U}_k^H\hat{\mathbf{G}}_{k,m}$
$\boldsymbol{\mathcal{H}}_{m,k}$	$\hat{\mathbf{H}}_{m,k}\mathbf{F}_k$
$\boldsymbol{\mathcal{T}}_{k,q}$	$\sum_{m=1}^{M}\hat{\mathbf{G}}_{k,m}\mathbf{W}_m\hat{\mathbf{H}}_{m,q}\mathbf{F}_q$

where $\mathbf{\Sigma}_{\mathrm{H}_{m,k}}$ and $\mathbf{\Sigma}_{\mathrm{G}_{k,m}}$ are the row correlation matrices, whereas $\mathbf{\Psi}_{\mathrm{H}_{m,k}}$ and $\mathbf{\Psi}_{\mathrm{G}_{k,m}}$ 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 a zero mean and unit variance.[2] This model is suitable when the CSI errors are dominated by the channel estimation errors.

2) Norm-Bounded Error Model: When the CSI is subject to quantization errors due to the limited-rate feedback, it can no longer be accurately characterized by the given statistical model. Instead, $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ are considered to assume values from the following norm-bounded sets [19]:

$$\mathcal{H}_{m,k} \triangleq \{\Delta\mathbf{H}_{m,k} : \|\Delta\mathbf{H}_{m,k}\|_F \leq \eta_{m,k}\} \tag{11}$$

$$\mathcal{G}_{k,m} \triangleq \{\Delta\mathbf{G}_{k,m} : \|\Delta\mathbf{G}_{k,m}\|_F \leq \xi_{k,m}\} \tag{12}$$

where $\eta_{m,k} > 0$ and $\xi_{k,m} > 0$ specify the radii of the uncertainty regions, thus reflecting the degree of uncertainties. The benefits of such an error model have been well justified in the literature of robust relay optimization (see, e.g., [8], [9], and [26]). The determination of the radii of the uncertainty regions has also been discussed in [19].

Throughout this paper, we assume that the magnitudes of the CSI errors are significantly lower than those of the channel estimates; therefore, the third- and higher-order terms in $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ are neglected in our subsequent analysis. We also introduce in Table I some useful notations to simplify our exposition.

Substituting (8) into (7) and applying the aforementioned assumptions, the per-stream MSE in the presence of CSI errors can be expressed as

$$\begin{aligned}&\varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right) \\ &\approx \left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k} + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,k}\mathbf{F}_k\right. \\ &\quad \left. + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,k}\mathbf{F}_k - \mathbf{e}_{k,l}^T\right\|^2 + \sigma_{\mathrm{D},k}^2\left\|\mathbf{u}_{k,l}\right\|^2 \\ &\quad + \sum_{q=1,q\neq k}^{K}\left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,q} + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\right. \\ &\qquad\qquad \left. + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,q}\mathbf{F}_q\right\|^2 \\ &\quad + \sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m} + \mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\mathbf{W}_m\right\|^2. \end{aligned} \tag{13}$$

[2]The superscript "W" simply refers to the spatially white or uncorrelated nature of these random variables.

We now observe that the per-stream MSE becomes uncertain in $\Delta\mathbf{H}_{m,k}\ \forall(m,k)\in\mathcal{M}\times\mathcal{K}$ and $\Delta\mathbf{G}_{k,m}\ \forall m\in\mathcal{M}$. Therefore, we introduce the following compact notations for convenience:

$$\Delta\mathbf{G}_k \triangleq (\Delta\mathbf{G}_{k,1},\ldots,\Delta\mathbf{G}_{k,M}) \in \mathcal{G}_k \triangleq \mathcal{G}_{k,1}\times\cdots\times\mathcal{G}_{k,M}$$

$$\Delta\mathbf{H} \triangleq (\Delta\mathbf{H}_{1,1},\ldots,\Delta\mathbf{H}_{M,K}) \in \mathcal{H} \triangleq \mathcal{H}_{1,1}\times\cdots\times\mathcal{H}_{M,K}.$$

For subsequent derivations, the dependence of $\varepsilon_{k,l}$ on $\Delta\mathbf{H}$ and $\Delta\mathbf{G}_k$ is made explicit in (13).

The kth relay's transmit power in the presence of CSI errors can also be explicitly expressed as $P_{\mathrm{R},m}(\Delta\mathbf{H}_m)$, where $\Delta\mathbf{H}_m \triangleq (\Delta\mathbf{H}_{m,1},\ldots,\Delta\mathbf{H}_{m,K}) \in \mathcal{H}_m \triangleq \mathcal{H}_{m,1}\times\cdots\times\mathcal{H}_{m,K}$.

C. *Problem Formulation*

In contrast to the prior advances [6]–[8], [14], [22] found in the relay optimization literature, where certain global objective functions are minimized subject to power constraints at the sources and relays, we formulate the following robust design problems under the explicit consideration of QoS. Let us commence by introducing the following unified operation:

$$\mathcal{U}\{f(\Delta\mathbf{X})\} = \begin{cases} \mathbb{E}_{\Delta\mathbf{X}} f(\Delta\mathbf{X}), & \Delta\mathbf{X} \text{ is random} \\ \max\limits_{\Delta\mathbf{X}\in\mathcal{X}} f(\Delta\mathbf{X}), & \Delta\mathbf{X} \text{ is deterministic} \end{cases} \tag{14}$$

where $\Delta\mathbf{X}\in\mathbb{C}^{M\times N}$ and $f(\cdot):\mathbb{C}^{M\times N}\to\mathbb{R}$. Depending on the specific assumptions concerning $\Delta\mathbf{X}, \mathcal{U}\{\cdot\}$ either computes the expectation of $f(\Delta\mathbf{X})$ over the ensemble of realizations $\Delta\mathbf{X}$ or maximizes $f(\Delta\mathbf{X})$ for all $\Delta\mathbf{X}$ within some bounded set $\mathcal{X}$. This notation will be useful and convenient for characterizing the per-stream MSE of (13) and the relay's power $P_{\mathrm{R},m}(\Delta\mathbf{H}_m)$ for different types of CSI errors in a unified form in our subsequent analysis.

1) Min–Max Problem: For notational convenience, we define $\mathbf{F}\triangleq(\mathbf{F}_1,\ldots,\mathbf{F}_K)$, $\mathbf{W}\triangleq(\mathbf{W}_1,\ldots,\mathbf{W}_M)$, and $\mathbf{U}\triangleq(\mathbf{U}_1,\ldots,\mathbf{U}_K)$, which collects the corresponding design variables. In this problem, we jointly design $\{\mathbf{F},\mathbf{W},\mathbf{U}\}$ with the goal of minimizing the maximum per-stream MSE subject to the source and relay power constraints. This problem pertains to the design of energy-efficient relay networks, where there is a strict constraint on the affordable power consumption. Based on the notation in (14), it can be expressed in the following unified form, which is denoted $\mathcal{M}(P_{\mathrm{R}})$:

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{\forall k\in\mathcal{K},l\in\mathcal{D}_k} \kappa_{k,l}\,\mathcal{U}\{\varepsilon_{k,l}(\Delta\mathbf{H},\Delta\mathbf{G}_k)\} \tag{15a}$$

$$\text{s.t.}\quad \mathcal{U}\{P_{\mathrm{R},m}(\Delta\mathbf{H}_m)\} \le \rho_m P_{\mathrm{R}} \quad \forall m\in\mathcal{M} \tag{15b}$$

$$\mathrm{Tr}(\mathbf{F}_k^H\mathbf{F}_k) \le P_{\mathrm{S},k}^{\max} \quad \forall k\in\mathcal{K} \tag{15c}$$

where $\{\kappa_{k,l}>0:\forall k\in\mathcal{K}, l\in\mathcal{D}_k\}$ is a set of weights assigned to the different data streams for maintaining fairness among them, P_{R} is the common maximum affordable transmit power of all the relays, and $\{\rho_m>0:\forall m\in\mathcal{M}\}$ is a set of coefficients specifying the individual power of each relay.

2) QoS Problem: The second strategy, which serves as a complement to the given min–max problem, aims for minimizing the maximum per-relay power, while strictly satisfying the QoS constraints for all the data streams and all the source power constraints.[3] Specifically, this problem, which is denoted $\mathcal{Q}(\gamma)$, can be formulated as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{m\in\mathcal{M}} \quad \frac{1}{\rho_m}\mathcal{U}\{P_{\mathrm{R},m}(\Delta\mathbf{H}_m)\} \tag{16a}$$

$$\text{s.t.}\quad \mathcal{U}\{\varepsilon_{k,l}(\Delta\mathbf{H},\Delta\mathbf{G}_k)\} \le \frac{\gamma}{\kappa_{k,l}} \quad \forall k\in\mathcal{K}, l\in\mathcal{D}_k \tag{16b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right) \le P_{\mathrm{S},k}^{\max} \quad \forall k\in\mathcal{K} \tag{16c}$$

where γ denotes a common QoS target for all the data streams.

The following remark is of interest.

Remark 1: The major difference between the min–max and QoS problems is that solving the QoS problem is not always feasible. This is because the per-stream MSE imposed by the interstream and interuser interference [cf. (13)] cannot be made arbitrarily small by simply increasing the transmit power. By contrast, solving the min–max problem is always feasible since it relies on its "*best effort*" to improve the QoS for all the data streams at limited power consumption. Both problem formulations are nonconvex and in general NP-hard. These issues motivate the pursuit of a tractable but suboptimal solution to the design problems considered.

III. Statistically Robust Transceiver Design for the Min–Max Problem

Here, we propose an algorithmic solution to the min–max problem of (15) in the presence of the statistical CSI errors of Section II-B1. The corresponding statistically robust version of (15) can be formulated as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{\forall k\in\mathcal{K},l\in\mathcal{D}_k} \kappa_{k,l}\bar{\varepsilon}_{k,l} \tag{17a}$$

$$\text{s.t.}\quad \overline{P}_{\mathrm{R},m} \le \rho_m P_{\mathrm{R}} \quad \forall m\in\mathcal{M} \tag{17b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right) \le P_{\mathrm{S},k}^{\max} \quad \forall k\in\mathcal{K} \tag{17c}$$

where we have

$$\bar{\varepsilon}_{k,l} \triangleq \mathbb{E}_{\Delta\mathbf{H},\Delta\mathbf{G}_k}\{\varepsilon_{k,l}(\Delta\mathbf{H},\Delta\mathbf{G}_k)\}$$

$$\overline{P}_{\mathrm{R},m} \triangleq \mathbb{E}_{\Delta\mathbf{H}_m}\{P_{\mathrm{R},m}(\Delta\mathbf{H}_m)\}. \tag{18}$$

To further exploit the structure of (17), we have to compute the expectations in (18), which we refer to as the averaged MSE and relay power, respectively. By exploiting the independence

[3]In fact, the min–max problem $\mathcal{M}(P_{\mathrm{R}})$ and the QoS problem $\mathcal{Q}(\gamma)$ are the so-called *inverse problems*, i.e., we have $\gamma=\mathcal{M}[\mathcal{Q}(\gamma)]$ and $P_{\mathrm{R}}=\mathcal{Q}[\mathcal{M}(P_{\mathrm{R}})]$. The proof follows a similar argument to that of [27, Th. 3]. However, as shown in the subsequent analysis, the proposed algorithm cannot guarantee finding the global optimum of the design problems. Therefore, monotonic convergence cannot be guaranteed, which is formally stated as $P_{\mathrm{R}}\ge P'_{\mathrm{R}} \nRightarrow \mathcal{M}(P_{\mathrm{R}})\le\mathcal{M}(P'_{\mathrm{R}})$ and $\gamma\ge\gamma' \nRightarrow \mathcal{Q}(\gamma)\le\mathcal{Q}(\gamma')$. Due to the lack of the monotonicity, a 1-D binary search algorithm is unable to solve $\mathcal{Q}(\gamma)$ via a sequence of $\mathcal{M}(P_{\mathrm{R}})$ evaluations. Consequently, a formal inverse problem definition is not stated in this paper.

of $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ in (13), the per-stream MSE averaged over the channel uncertainties can be expanded as

$$\begin{aligned}\overline{\varepsilon}_{k,l} =& \mathbf{u}_{k,l}^H\left(\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^H + \mathbf{R}_k\right)\mathbf{u}_{k,l} - 2\Re\left\{\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k}\mathbf{e}_{k,l}\right\} + 1\\ &+\sum_{q=1}^{K}\sum_{m=1}^{M}\underbrace{\mathbb{E}\left\{\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\Delta\mathbf{G}_{k,m}^H\mathbf{u}_{k,l}\right\}}_{\mathcal{I}_1}\\ &+\sum_{q=1}^{K}\sum_{m=1}^{M}\underbrace{\mathbb{E}\left\{\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\Delta}\mathbf{H}_{m,q}^H\boldsymbol{\mathcal{G}}_{k,m}^H\mathbf{u}_{k,l}\right\}}_{\mathcal{I}_2}\\ &+\sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\underbrace{\mathbb{E}\left\{\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{W}_m^H\Delta\mathbf{G}_{k,m}^H\mathbf{u}_{k,l}\right\}}_{\mathcal{I}_3} \quad (19)\end{aligned}$$

where we have

$$\mathbf{R}_k = \sum_{q=1,q\neq k}^{K}\boldsymbol{\mathcal{T}}_{k,q}\boldsymbol{\mathcal{T}}_{k,q}^H + \sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\mathcal{G}}_{k,m}^H + \sigma_{\mathrm{D},k}^2\mathbf{I}_{d_k}. \quad (20)$$

To compute the expectations in (19), we rely on the results of [28, (10)] to obtain

$$\begin{aligned}\mathcal{I}_1 &= \mathbf{u}_{k,l}^H\mathbb{E}\left\{\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\Delta\mathbf{G}_{k,m}^H\right\}\mathbf{u}_{k,l}\\ &= \mathrm{Tr}\left(\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\mathbf{u}_{k,l}^H\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}. \quad (21)\end{aligned}$$

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

$$\mathcal{I}_2 = \mathrm{Tr}\left(\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\Psi}_{\mathrm{H}_{m,q}}\right)\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\Sigma}_{\mathrm{H}_{m,q}}\boldsymbol{\mathcal{G}}_{k,m}^H\mathbf{u}_{k,l} \quad (22)$$

$$\mathcal{I}_3 = \mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\mathbf{u}_{k,l}^H\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}. \quad (23)$$

Based on (21)–(23), the averaged MSE in (19) is therefore equivalent to

$$\begin{aligned}\overline{\varepsilon}_{k,l} =& \mathbf{u}_{k,l}^H\left(\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^H + \mathbf{R}_k + \boldsymbol{\Omega}_k\right)\mathbf{u}_{k,l}\\ &- 2\Re\left\{\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k}\mathbf{e}_{k,l}\right\} + 1 \quad (24)\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{\Omega}_k =& \sum_{q=1}^{K}\sum_{m=1}^{M}\Big(\mathrm{Tr}\left(\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\\ &+ \mathrm{Tr}\left(\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\Psi}_{\mathrm{H}_{m,q}}\right)\boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\Sigma}_{\mathrm{H}_{m,q}}\boldsymbol{\mathcal{G}}_{k,m}^H\Big)\\ &+\sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}. \quad (25)\end{aligned}$$

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

The averaged relay power $\overline{P}_{\mathrm{R},m}$ can be derived as

$$\begin{aligned}\overline{P}_{\mathrm{R},m} =& \sum_{k=1}^{K}\Big(\mathrm{Tr}\big(\mathbf{F}_k^H\hat{\mathbf{H}}_{m,k}^H\mathbf{W}_m^H\mathbf{W}_m\hat{\mathbf{H}}_{m,k}\mathbf{F}_k\big)\\ &+\mathrm{Tr}\left(\mathbf{F}_k\mathbf{F}_k^H\boldsymbol{\Psi}_{\mathrm{H}_{m,k}}\right)\mathrm{Tr}\left(\mathbf{W}_m^H\mathbf{W}_m\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}\right)\Big)\\ &+\sigma_{\mathrm{R},m}^2\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\right) \quad (26)\end{aligned}$$

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

A. *Iterative Joint Transceiver Optimization*

It is worthwhile noting that the inner pointwise maximization in (17a) preserves the partial convexity of $\overline{\varepsilon}_{k,l}$. Substituting (24) and (26) back into (17), the latter is shown to possess a so-called *block multiconvex* structure [20], which implies that the problem is convex in each block of variables, although in general not jointly convex in all the variables.

Motivated by the given 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 associated with the remaining blocks. In this way, three subproblems can be derived from (17), with each updating $\mathbf{F}$, $\mathbf{W}$, and $\mathbf{U}$, respectively. Each subproblem can be transformed into a *convex* one, which is computationally much simpler than directly finding the optimal solution to the original joint problem (if at all possible). Since solving for each block at the current iteration depends on the values of the other blocks gleaned from the previous iteration, this method in effect can be recognized as a joint optimization approach in terms of both the underlying theory [15], [20] and the related applications [14], [17]. We now proceed by analyzing each of these subproblems.

1) Receive Filter Design: It can be observed in (19) that $\overline{\varepsilon}_{k,l}$ in (17a) only depends on the corresponding linear vector $\mathbf{u}_{k,l}$, whereas the constraints (17b) and (17c) do not involve $\mathbf{u}_{k,l}$. Hence, for a fixed $\mathbf{F}$ and $\mathbf{W}$, the optimal $\mathbf{u}_{k,l}$ can be obtained independently and in parallel for different (k,l) values by equating the following complex gradient to zero:

$$\nabla_{\mathbf{u}_{k,l}^*}\overline{\varepsilon}_{k,l} = \mathbf{0}. \quad (27)$$

The resultant optimal solution of (27) is the Wiener filter, i.e.,

$$\mathbf{u}_{k,l} = \left(\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^H + \mathbf{R}_k + \boldsymbol{\Omega}_k\right)^{-1}\boldsymbol{\mathcal{T}}_{k,k}\mathbf{e}_{k,l}. \quad (28)$$

2) Source TPC Design: We then solve our problem for the TPC $\mathbf{F}$, while keeping $\mathbf{W}$ and $\mathbf{U}$ fixed. For better exposition of our solution, we can rewrite (17) after some matrix manipulations, explicitly in terms of $\mathbf{F}$ as given in (29), shown at the bottom of the next page, where $\mathbf{E}_{k,l}\triangleq\mathbf{e}_{k,l}\mathbf{e}_{k,l}^T$, $\eta_{\mathrm{R},m}\triangleq\rho_m P_{\mathrm{R}} - \sigma_{\mathrm{R},m}^2\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\right)$, and

$$\begin{aligned}a_3^{k,l} \triangleq& \mathbf{u}_{k,l}^H\Bigg[\sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\Big(\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\\ &+\boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\mathcal{G}}_{k,m}^H\Big) + \sigma_{\mathrm{D},k}^2\mathbf{I}_{N_{\mathrm{D},k}}\Bigg]\mathbf{u}_{k,l} + 1. \quad (30)\end{aligned}$$

The solution to the problem (29) is not straightforward; hence, we transform it into a more tractable form. To this end, we introduce the new variables of $\mathbf{f}_k \triangleq \mathrm{vec}\left(\mathbf{F}_k\right) \in \mathbb{C}^{N_{\mathrm{S},k} d_k \times 1}$ $\forall k \in \mathcal{K}$ and define the following quantities that are independent of $\mathbf{f}_k$ $\forall k \in \mathcal{K}$:

$$\mathbf{A}_{1,q}^{k,l} \triangleq \sum_{m=1}^{M} \mathbf{I}_{d_k} \otimes \Bigg(\sum_{n=1}^{M} \boldsymbol{\mathcal{W}}_{m,q}^{H} \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n} \boldsymbol{\mathcal{W}}_{n,q} + \mathrm{Tr}\left(\mathbf{u}_{k,l}^{H} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l}\right) \boldsymbol{\mathcal{W}}_{m,k}^{H} \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \boldsymbol{\mathcal{W}}_{m,k} + \mathrm{Tr}\left(\mathbf{u}_{k,l}^{H} \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}} \boldsymbol{\mathcal{G}}_{k,m}^{H} \mathbf{u}_{k,l}\right) \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \Bigg) \tag{31}$$

$$\mathbf{a}_{2}^{k,l} = \mathrm{vec}\left(\sum_{m=1}^{M} \boldsymbol{\mathcal{W}}_{m,k}^{H} \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \right) \tag{32}$$

$$\mathbf{A}_{4,k}^{m} = \mathbf{I}_{d_k} \otimes \left(\boldsymbol{\mathcal{W}}_{m,k}^{H} \boldsymbol{\mathcal{W}}_{m,k} + \mathrm{Tr}\left(\mathbf{W}_{m}^{H} \mathbf{W}_{m} \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}} \right) \boldsymbol{\Psi}_{\mathrm{H}_{m,k}} \right). \tag{33}$$

It may be readily verified that $\mathbf{A}_{1,q}^{k,l}$ and $\mathbf{A}_{4,k}^{m}$ are positive definite matrices. Then, we invoke the following identities, i.e., $\mathrm{Tr}\left(\mathbf{A}^{H}\mathbf{B}\mathbf{A}\right) = \mathrm{vec}\left(\mathbf{A}\right)^{H} \left(\mathbf{I} \otimes \mathbf{B}\right) \mathrm{vec}\left(\mathbf{A}\right)$ and $\mathrm{Tr}\left(\mathbf{A}^{H}\mathbf{B}\right) = \mathrm{vec}\left(\mathbf{B}\right)^{H} \mathrm{vec}\left(\mathbf{A}\right)$, for transforming both the objective (29a) and the constraints (29b)–(29c) into quadratic expressions of $\mathbf{f}_k$, and finally reach the following equivalent formulation:

$$\min_{\mathbf{f}_1,\ldots,\mathbf{f}_K,t} \quad t \tag{34a}$$

$$\text{s.t.} \quad \sum_{q=1}^{K} \mathbf{f}_q^{H} \mathbf{A}_{1,q}^{k,l} \mathbf{f}_q - 2\Re\left\{ \mathbf{f}_k^{H} \mathbf{a}_2^{k,l} \right\} + a_3^{k,l} \leq \frac{t}{\kappa_{k,l}} \quad \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{34b}$$

$$\sum_{k=1}^{K} \mathbf{f}_k^{H} \mathbf{A}_{4,k}^{m} \mathbf{f}_k \leq \eta_{\mathrm{R},m} \quad \forall m \in \mathcal{M} \tag{34c}$$

$$\mathbf{f}_k^{H} \mathbf{f}_k \leq P_{\mathrm{S},k}^{\max} \quad \forall k \in \mathcal{K} \tag{34d}$$

where t is an auxiliary variable. Problem (34) by definition is a convex separable inhomogeneous QCLP [16]. This class of optimization problems can be handled by the recently developed parser/solvers, such as `CVX` [29] where the built-in parser is capable of verifying the convexity of the optimization problem (in user-specified forms) and then, of automatically transforming it into a standard form; the latter may then be forwarded to external optimization solvers, such as `SeduMi` [30] and `MOSEK` [31]. To gain further insights into this procedure, we show in Appendix A that the problem (34) can be equivalently transformed into a standard SOCP that is directly solvable by a generic external optimization solver based on the interior-point method. Therefore, the SOCP form bypasses the tedious translation by the parser/solvers for every problem instance in real-time computation.

3) Relay AF Matrix Design: To solve for the relay AF matrices, we follow a similar procedure to that used for the source TPC design. However, here we introduce a new variable, which vertically concatenates all the vectorized relay AF matrices, yielding

$$\mathbf{w} \triangleq \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_M \end{bmatrix} \triangleq \begin{bmatrix} \mathrm{vec}\left(\mathbf{W}_1\right) \\ \vdots \\ \mathrm{vec}\left(\mathbf{W}_M\right) \end{bmatrix} \tag{35}$$

along with the following quantities, which are independent of $\mathbf{w}$:

$$\left[\mathbf{B}_1^{k,l}\right]_{m,n} = \sum_{q=1}^{K} \left[\left(\boldsymbol{\mathcal{H}}_{m,q}^{*} \boldsymbol{\mathcal{H}}_{n,q}^{T} \right) \otimes \left(\boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n} \right) \right] \tag{36}$$

$$\mathbf{b}_{2,m}^{k,l} \triangleq \mathrm{vec}\left(\boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{H}}_{m,k}^{H} \right) \tag{37}$$

$$\mathbf{B}_{3,m}^{k,l} \triangleq \sum_{q=1}^{K} \Big[\mathrm{Tr}\left(\mathbf{u}_{k,l}^{H} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l}\right) \boldsymbol{\mathcal{H}}_{m,q}^{*} \boldsymbol{\mathcal{H}}_{m,q}^{T} \otimes \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} + \mathrm{Tr}\left(\mathbf{F}_q^{H} \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \mathbf{F}_q \right) \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}}^{T} \otimes \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,m} \Big] + \sigma_{\mathrm{R},m}^{2} \mathrm{Tr}\left(\mathbf{u}_{k,l}^{H} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l} \right) \mathbf{I}_{N_{\mathrm{R},m}} \otimes \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} + \sigma_{\mathrm{R},m}^{2} \mathbf{I}_{N_{\mathrm{R},m}} \otimes \left(\boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,m} \right) \tag{38}$$

$$b_4^{k,l} \triangleq \sigma_{\mathrm{D},k}^{2} \left\| \mathbf{u}_{k,l} \right\|^{2} + 1 \tag{39}$$

$$\mathbf{B}_{5,m} \triangleq \Bigg[\sigma_{\mathrm{R},m}^{2} \mathbf{I}_{N_{\mathrm{R},m}} + \sum_{k=1}^{K} \Big(\boldsymbol{\mathcal{H}}_{m,k}^{*} \boldsymbol{\mathcal{H}}_{m,k}^{T} + \mathrm{Tr}\left(\mathbf{F}_k \mathbf{F}_k^{H} \boldsymbol{\Psi}_{\mathrm{H}_{m,k}} \right) \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}^{T} \Big) \Bigg] \otimes \mathbf{I}_{N_{\mathrm{R},m}} \tag{40}$$

$$\min_{\mathbf{F}} \max_{\forall k \in \mathcal{K}, l \in \mathcal{D}_k} \kappa_{k,l} \Bigg\{ \sum_{q=1}^{K} \sum_{m=1}^{M} \sum_{n=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^{H} \boldsymbol{\mathcal{W}}_{m,q}^{H} \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n} \boldsymbol{\mathcal{W}}_{n,q} \mathbf{F}_q \right) - \sum_{m=1}^{M} 2\Re\left\{ \mathrm{Tr}\left(\mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,m} \boldsymbol{\mathcal{W}}_{m,k} \mathbf{F}_k \right) \right\} + a_3^{k,l} + \sum_{q=1}^{K} \sum_{m=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^{H} \boldsymbol{\mathcal{W}}_{m,k}^{H} \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \boldsymbol{\mathcal{W}}_{m,k} \mathbf{F}_q \right) \mathrm{Tr}\left(\mathbf{u}_{k,l}^{H} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l} \right) + \sum_{q=1}^{K} \sum_{m=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^{H} \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \mathbf{F}_q \right) \mathrm{Tr}\left(\mathbf{u}_{k,l}^{H} \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}} \boldsymbol{\mathcal{G}}_{k,m}^{H} \mathbf{u}_{k,l} \right) \Bigg\} \tag{29a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \mathrm{Tr}\left(\mathbf{F}_k^{H} \left(\hat{\mathbf{H}}_{m,k}^{H} \mathbf{W}_m^{H} \mathbf{W}_m \hat{\mathbf{H}}_{m,k} + \mathrm{Tr}\left(\mathbf{W}_m^{H} \mathbf{W}_m \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}} \right) \boldsymbol{\Psi}_{\mathrm{H}_{m,k}} \right) \mathbf{F}_k \right) \leq \eta_{\mathrm{R},m}, \quad \forall m \in \mathcal{M} \tag{29b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^{H} \mathbf{F}_k \right) \leq P_{\mathrm{S},k}^{\max}, \quad \forall k \in \mathcal{K} \tag{29c}$$

where $\mathbf{B}_1^{k,l}$ is a block matrix with its (m,n)th block defined earlier. Then, using the identities $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\mathbf{C}\mathbf{D}^H\right) = \mathrm{vec}\left(\mathbf{A}\right)^H \left(\mathbf{D}^T \otimes \mathbf{B}\right)\mathrm{vec}(\mathbf{C})$, $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\mathbf{A}\right) = \mathrm{vec}(\mathbf{A})^H \left(\mathbf{I} \otimes \mathbf{B}\right) \mathrm{vec}\left(\mathbf{A}\right)$, and $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\right) = \mathrm{vec}\left(\mathbf{B}\right)^H \mathrm{vect}\left(\mathbf{A}\right)$, we can formulate the following optimization problem:

$$\min_{\mathbf{w},t} \quad t \tag{41a}$$

$$\text{s.t.} \quad \mathbf{w}^H\mathbf{B}_1^{k,l}\mathbf{w} - \sum_{m=1}^{M} 2\Re\left\{\mathbf{w}_m^H\mathbf{b}_{2,m}^{k,l}\right\} + \sum_{m=1}^{M} \mathbf{w}_m^H\mathbf{B}_{3,m}^{k,l}\mathbf{w}_m + b_4^{k,l} \le \frac{t}{\kappa_{k,l}} \; \forall l \in \mathcal{D}_k, k \in \mathcal{K} \tag{41b}$$

$$\mathbf{w}_m^H\mathbf{B}_{5,m}\mathbf{w}_m \le \rho_m P_{\mathrm{R}} \quad \forall m \in \mathcal{M}. \tag{41c}$$

It may be readily shown that $\mathbf{B}_1^{k,l}$, $\mathbf{B}_{3,m}^{k,l}$, and $\mathbf{B}_{5,m}$ are all positive definite matrices and that (41) is also a convex separable inhomogeneous QCLP. Using a similar approach to the one derived in Appendix A, the SOCP formulation of (41) can readily be obtained. The details of the transformation are therefore omitted for brevity.

B. Algorithm and Properties

We assume that there exists a central processing node, which, upon collecting the channel estimates $\{\hat{\mathbf{H}}_{m,k}, \hat{\mathbf{G}}_{k,m} \, \forall m \in \mathcal{M}, k \in \mathcal{K}\}$ and the covariance matrices of the CSI errors $\{\mathbf{\Sigma}_{\mathrm{H}_{m,k}}, \mathbf{\Sigma}_{\mathrm{G}_{k,m}}, \mathbf{\Psi}_{\mathrm{H}_{m,k}}, \mathbf{\Psi}_{\mathrm{G}_{k,m}} \, \forall m \in \mathcal{M}, k \in \mathcal{K}\}$, optimizes all the design variables and sends them back to the corresponding nodes. The iterative procedure listed in Algorithm 1 therefore should be implemented in a centralized manner, where $\{\mathbf{F}^{(i)}, \mathbf{W}^{(i)}, \mathbf{U}^{(i)}\}$ and $t^{(i)}$ represent the set of design variables and the objective value in (17a), respectively, at the ith iteration. A simple termination criterion can be $|t^{(i)} - t^{(i-1)}| < \epsilon$, where $\epsilon > 0$ is a predefined threshold. In the following, we shall analyze both the convergence properties and the complexity of the proposed algorithm.

1) Convergence: Provided that there is a feasible initialization for Algorithm 1, the solution to each subproblem is globally optimal. As a result, the sequence of the objective values in (17a) is monotonically nonincreasing as the iteration index i increases. Since the maximum per-stream MSE is bounded from below (at least) by zero, the sequence of the objective values must converge by invoking the monotonic convergence theorem.

2) Complexity: When the number of antennas at the sources and relays, i.e., $N_{\mathrm{S},k}$ and $N_{\mathrm{R},m}$, have the same order of magnitude, the complexity of Algorithm 1 is dominated by the SOCP of (62), which is detailed in Appendix A, as it involves all the constraints of the original problem (17). To simplify the complexity analysis, we assume that $N_{\mathrm{S},k} = N_{\mathrm{S}}$, and $d_k = d \; \forall k \in \mathcal{K}$. In (62), the total number of design variables is $N_{\mathrm{total}} = N_{\mathrm{S}}^2 K + 1 + K^2 d + KM$. The size of the second-order cones (SOCs) in the constraints (62b)–(62g) is given by $(N_{\mathrm{S}}^2 + 1)dK(K-1)$, $(N_{\mathrm{S}}^2 + 1)dK$, $(K+2)dK$, $(N_{\mathrm{S}}^2 + 1)KM$, $(K+1)M$, and $(N_{\mathrm{S}}^2 + 1)K$, respectively. Therefore, the total dimension of all the SOCs in these constraints can be shown to be $D_{\mathrm{SOCP}} = \mathcal{O}(N_{\mathrm{S}}^2 dK^2 + N_{\mathrm{S}}^2 MK)$. It has been shown in [32] that problem (62) can be solved most efficiently using the primal–dual interior-point method at *worst-case* complexity on the order of $\mathcal{O}(N_{\mathrm{total}}^2 D)$ if no special structure in the problem data is exploited. The computational complexity of Algorithm 1 is therefore on the order of $\mathcal{O}(N_{\mathrm{S}}^6)$, $\mathcal{O}(K^6)$, and $\mathcal{O}(M^3)$ in the individual parameters N_{S}, K and M, respectively. In practice, however, we find that the matrices $\mathbf{A}_{1,q}^{k,l}$ and $\mathbf{A}_{4,k}^{m}$ in (31) and (33), respectively, exhibit a significant level of sparsity, which allows solving the SOCP more efficiently. In our simulations, we therefore measured the CPU time required for solving (62) for different values of N_{S}, K, and M (the results are not reported due to the space limitation) and found that the orders of complexity obtained empirically are significantly lower than those of the given worst-case analysis. Empirically, we found these to be around $\mathcal{O}(N_{\mathrm{S}}^{1.6})$, $\mathcal{O}(K^{1.7})$, and $\mathcal{O}(M^{1.3})$.

Algorithm 1 Iterative Algorithm for Statistically Robust Min–Max Problem

Initialization:
1: Set the iteration index $i = 0$, $\mathbf{F}_k^{(0)} = \sqrt{P_{\mathrm{S},k}^{\max}} \mathbf{I}_{N_{\mathrm{S},k} \times d_k}$, $\forall k \in \mathcal{K}$ and $\mathbf{W}_m^{(0)} = \sqrt{\frac{\rho_m P_{\mathrm{R}}}{\mathrm{Tr}(\mathbf{B}_{5,m})}} \mathbf{I}_{N_{\mathrm{R},m}}, \forall m \in \mathcal{M}$
2: **repeat**
3: Compute $\mathbf{u}_{k,l}^{(i+1)} \; \forall k \in \mathcal{K}, l \in \mathcal{D}_k$, using the Wiener filter (28) in parallel;
4: Compute $\mathbf{F}_k^{(i+1)} \; \forall k \in \mathcal{K}$ by solving the SOCP (62);
5: Compute $\mathbf{W}_m^{(i+1)} \; \forall m \in \mathcal{M}$ by solving the SOCP (41);
6: $i \leftarrow i + 1$;
7: **until** $|t^{(i)} - t^{(i-1)}| < \epsilon$

IV. Worst-Case Robust Transceiver Design for the Min–Max Problem

Here, we consider the joint transceiver design problem under min–max formulation of (15) and the norm-bounded CSI error model of Section II-B2. To this end, based on the notation in (14), we explicitly rewrite this problem as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \max_{\substack{\forall k \in \mathcal{K}, l \in \mathcal{D}_k, \\ \forall \Delta\mathbf{H} \in \mathcal{H}, \Delta\mathbf{G}_k \in \mathcal{G}_k}} \kappa_{k,l}\varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right) \tag{42a}$$

$$\text{s.t.} \quad P_{\mathrm{R},m}\left(\Delta\mathbf{H}_m\right) \le \rho_m P_{\mathrm{R}} \, \forall m \in \mathcal{M}, \Delta\mathbf{H}_m \in \mathcal{H}_m \tag{42b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right) \le P_{\mathrm{S},k}^{\max} \, \forall k \in \mathcal{K} \tag{42c}$$

whose epigraph form can be expressed as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad t \tag{43a}$$

$$\text{s.t.} \quad \varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right) \le \frac{t}{\kappa_{k,l}} \, \forall k \in \mathcal{K}, l \in \mathcal{D}_k, \; \Delta\mathbf{H} \in \mathcal{H}, \Delta\mathbf{G}_k \in \mathcal{G}_k \tag{43b}$$

$$P_{\mathrm{R},m}\left(\Delta\mathbf{H}_m\right) \le \rho_m P_{\mathrm{R}} \, \forall m \in \mathcal{M}, \Delta\mathbf{H}_m \in \mathcal{H}_m \tag{43c}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right) \le P_{\mathrm{S},k}^{\max} \, \forall k \in \mathcal{K} \tag{43d}$$

where t is an auxiliary variable. As compared with the statistically robust version of (17), problem (43) now encounters two major challenges, namely the nonconvexity and the *semi-infinite* nature of the constraints (43b) and (43c), which render the optimization problem mathematically intractable. In what follows, we derive a solution to address these calamities.

A. Iterative Joint Transceiver Optimization

To overcome the first difficulty, we still rely on the iterative block coordinate update approach described in Section III; however, the three resultant subproblems are *semi-infinite* due to the continuous but bounded channel uncertainties in (43b) and (43c). To handle the semi-infiniteness, an equivalent reformulation of these constraints as LMI will be derived by using certain matrix transformation techniques and by exploiting an extended version of the $\mathcal{S}$-lemma of [21]. In turn, such LMI will convert each of the subproblems into an equivalent SDP [33] efficiently solvable by interior-point methods [34].

1) Receive Filter Design: In this subproblem, we have to minimize t in (43a) with respect to $\mathbf{u}_{k,l}$ subject to the constraint (43b). To transform this constraint into an equivalent LMI, the following lemma is presented, which is an extended version of the one in [21].

Lemma 1 (Extension of $\mathcal{S}$-lemma [21]): Let $\mathbf{A}(\mathbf{x}) = \mathbf{A}^H(\mathbf{x})$, $\boldsymbol{\Sigma}(\mathbf{x}) = \boldsymbol{\Sigma}^H(\mathbf{x})$, $\{\mathbf{D}_k(\mathbf{x})\}_{k=1}^N$, and $\{\mathbf{B}_k\}_{k=1}^N$ be matrices with appropriate dimensions, where $\mathbf{A}(\mathbf{x})$, $\boldsymbol{\Sigma}(\mathbf{x})$, and $\{\mathbf{D}_k(\mathbf{x})\}_{k=1}^N$ are affine functions of $\mathbf{x}$. The following *semi-infinite* matrix inequality:

$$\left(\mathbf{A}(\mathbf{x}) + \sum_{k=1}^{N} \mathbf{B}_k^H \mathbf{C}_k \mathbf{D}_k(\mathbf{x})\right) \times \left(\mathbf{A}(\mathbf{x}) + \sum_{k=1}^{N} \mathbf{B}_k^H \mathbf{C}_k \mathbf{D}_k(\mathbf{x})\right)^H \preceq \boldsymbol{\Sigma}(\mathbf{x}) \quad (44)$$

holds for all $\|\mathbf{C}_k\|_S \leq \rho_k, k = 1, \ldots, N$ if and only if there exist nonnegative scalars $\tau_1, \ldots, \tau_N$ satisfying (45), shown at the bottom of the page.

A simplified version of Lemma 1, which considers only a single uncertainty block, i.e., $N = 1$, can be traced back to [35], whereas a further related corollary is derived in [21, Proposition 2]. Lemma 1 extends this result to the case of multiple uncertainty blocks, i.e., $K > 1$; the proof which follows similar steps as in [21] is omitted owing to the space limitation.

Upon using Lemma 1, the constraint (43b) can equivalently be reformulated as follows.

Proposition 1: There exist nonnegative values of $\boldsymbol{\tau}_{k,l}^{\mathrm{G}} \in \mathbb{R}^{M\times 1}$ and $\boldsymbol{\tau}_{k,l}^{\mathrm{H}} \in \mathbb{R}^{KM\times 1}$ capable of ensuring that the semi-infinite constraint (43b) is equivalent to the matrix inequality in (46), shown at the bottom of the page, where we have $N_{\mathrm{R}} \triangleq \sum_{m=1}^{M} N_{\mathrm{R},m}$, $N_{\mathrm{S}} \triangleq \sum_{k=1}^{K} N_{\mathrm{S},k}$, and the operator $(*)$ denotes the Khatri–Rao product (blockwise Kronecker product) [36]. In (46), $\overline{\boldsymbol{\Theta}}_{k,l}$ and $\overline{\boldsymbol{\Phi}}_{k,l}$ are defined as

$$\overline{\boldsymbol{\Theta}}_{k,l} \triangleq \begin{bmatrix} \xi_{k,1}\boldsymbol{\Theta}_1^{k,l} \\ \vdots \\ \xi_{k,M}\boldsymbol{\Theta}_M^{k,l} \end{bmatrix}, \overline{\boldsymbol{\Phi}}_{k,l} \triangleq \begin{bmatrix} \eta_{1,1}\boldsymbol{\Phi}_{1,1}^{k,l} \\ \vdots \\ \eta_{M,K}\boldsymbol{\Phi}_{M,K}^{k,l} \end{bmatrix} \quad (47)$$

whereas $\boldsymbol{\Theta}_{k,l}$, $\boldsymbol{\Phi}_{k,l}$, and $\boldsymbol{\theta}_{k,l}$ are defined in (71) of Appendix B.

Proof: See Appendix B. ■

Using (46), the subproblem formulated for $\mathbf{u}_{k,l}$ can be equivalently recast as

$$\min_{t,\mathbf{u}_{k,l},\boldsymbol{\tau}_{k,l}^{\mathrm{g}},\boldsymbol{\tau}_{k,l}^{\mathrm{h}}} \ t \quad \text{s.t.} \quad \mathbf{Q}_{k,l} \succeq \mathbf{0}. \quad (48)$$

With fixed $\mathbf{F}$ and $\mathbf{W}$, (46) depends affinely on the design variables $\{t, \mathbf{u}_{k,l}, \boldsymbol{\tau}_{k,l}^{\mathrm{g}}, \boldsymbol{\tau}_{k,l}^{\mathrm{h}}\}$. Therefore, (48) is a convex SDP of the LMI form [33], which is efficiently solvable by existing optimization tools based on the interior-point method. Since the $\mathbf{u}_{k,l}$ for different values of (k,l) are independent of each other, they can be updated in parallel by solving (48) for different k and l.

2) Source TPC Design: We now have to solve problem (43) for $\mathbf{F}$ by fixing $\mathbf{U}$ and $\mathbf{W}$. The solution is formulated in the following proposition.

$$\begin{bmatrix} \boldsymbol{\Sigma}(\mathbf{x}) - \sum_{k=1}^{N} \tau_k \mathbf{B}_k^H \mathbf{B}_k & \mathbf{A}(\mathbf{x}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{A}^H(\mathbf{x}) & \mathbf{I} & \rho_1 \mathbf{D}_1^H(\mathbf{x}) & \cdots & \rho_N \mathbf{D}_N^H(\mathbf{x}) \\ \mathbf{0} & \rho_1 \mathbf{D}_1(\mathbf{x}) & \tau_1 \mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \rho_N \mathbf{D}_N(\mathbf{x}) & \mathbf{0} & \cdots & \tau_N \mathbf{I} \end{bmatrix} \succeq \mathbf{0} \quad (45)$$

$$\mathbf{Q}_{k,l} \triangleq \begin{bmatrix} \frac{t}{\kappa_{k,l}} - \mathbf{1}^T \boldsymbol{\tau}_{k,l}^{\mathrm{G}} - \mathbf{1}^T \boldsymbol{\tau}_{k,l}^{\mathrm{H}} & \boldsymbol{\theta}_{k,l} & \mathbf{0}_{1\times N_{\mathrm{D},k}N_{\mathrm{R}}} & \mathbf{0}_{1\times N_{\mathrm{S}}N_{\mathrm{R}}} \\ \boldsymbol{\theta}_{k,l}^H & \mathbf{I}_{d+N_{\mathrm{R}}+N_{\mathrm{D},k}} & \overline{\boldsymbol{\Theta}}_{k,l}^H & \overline{\boldsymbol{\Phi}}_{k,l}^H \\ \mathbf{0}_{N_{\mathrm{D},k}N_{\mathrm{R}}\times 1} & \overline{\boldsymbol{\Theta}}_{k,l} & \operatorname{diag}\left(\boldsymbol{\tau}_{k,l}^{\mathrm{G}}\right) * \mathbf{I}_{N_{\mathrm{D},k}N_{\mathrm{R}}} & \mathbf{0}_{N_{\mathrm{D},k}N_{\mathrm{R}}\times N_{\mathrm{S}}N_{\mathrm{R}}} \\ \mathbf{0}_{N_{\mathrm{S}}N_{\mathrm{R}}\times 1} & \overline{\boldsymbol{\Phi}}_{k,l} & \mathbf{0}_{N_{\mathrm{S}}N_{\mathrm{R}}\times N_{\mathrm{D},k}N_{\mathrm{R}}} & \operatorname{diag}\left(\boldsymbol{\tau}_{k,l}^{\mathrm{H}}\right) * \mathbf{I}_{N_{\mathrm{S}}N_{\mathrm{R}}} \end{bmatrix} \succeq \mathbf{0} \quad (46)$$

Proposition 2: The subproblem of optimizing the TPCs $\mathbf{F}$ can be formulated as the following SDP:

$$\min_{t,\mathbf{F},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l},\boldsymbol{\tau}^{\mathrm{p}}_{m}} \quad t \tag{49a}$$

$$\text{s.t.} \quad \mathbf{Q}_{k,l} \succeq \mathbf{0} \, \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{49b}$$

$$\mathbf{P}_m \succeq \mathbf{0} \, \forall m \in \mathcal{M} \tag{49c}$$

$$\begin{bmatrix} P^{\max}_{\mathrm{S},k} & \mathbf{f}^H_k \\ \mathbf{f}_k & \mathbf{I}_{N_{\mathrm{S},k} d_k} \end{bmatrix} \succeq \mathbf{0} \, \forall k \in \mathcal{K} \tag{49d}$$

where we have

$$\mathbf{P}_m \triangleq \begin{bmatrix} \rho_m P_{\mathrm{R}} - \mathbf{1}^T \boldsymbol{\tau}^{\mathrm{p}}_m & \mathbf{t}^H_m & \mathbf{0}_{1\times N_{\mathrm{S}} N_{\mathrm{R},m}} \\ \mathbf{t}_m & \mathbf{I} & \overline{\mathbf{T}}_m \\ \mathbf{0}_{N_{\mathrm{S}} N_{\mathrm{R},m}\times 1} & \overline{\mathbf{T}}^H_m & \mathrm{diag}\left(\boldsymbol{\tau}^{\mathrm{p}}_m\right) * \mathbf{I} \end{bmatrix} \succeq \mathbf{0} \tag{50}$$

with $\boldsymbol{\tau}^{\mathrm{p}}_m \in \mathbb{R}^{K\times 1}$, $\overline{\mathbf{T}}_m(\mathbf{F}) \triangleq \left[\mathbf{T}^T_{m,1}, \ldots, \mathbf{T}^T_{m,K}\right]^T$, and

$$\mathbf{t}_m \triangleq \begin{bmatrix} \mathrm{vec}\left(\mathbf{W}_m \hat{\mathbf{H}}_{m,k} \mathbf{F}_1\right) \\ \vdots \\ \mathrm{vec}\left(\mathbf{W}_m \hat{\mathbf{H}}_{m,K} \mathbf{F}_K\right) \\ \sigma_{\mathrm{R},m} \mathrm{vec}\left(\mathbf{W}_m\right) \end{bmatrix} \tag{51}$$

$$\mathbf{T}_{m,k} \triangleq \begin{bmatrix} \mathbf{0}_{\sum_{q=1}^{k-1} d_q N_{\mathrm{R},m} \times N_{\mathrm{S},k} N_{\mathrm{R},m}} \\ \mathbf{F}^T_k \otimes \mathbf{W}_m \\ \mathbf{0}_{\left(\sum_{q=k+1}^{K} d_q N_{\mathrm{R},m} + N^2_{\mathrm{R},m}\right)\times N_{\mathrm{S},k} N_{\mathrm{R},m}} \end{bmatrix}. \tag{52}$$

Proof: Since $\mathbf{F}$ is involved in all the constraints of the original problem (43), in the following, we will transform each of these constraints into tractable forms.

First, note that (43b) has already been reformulated as (46), which is a trilinear function of $\mathbf{F}$, $\mathbf{W}$, and $\mathbf{U}$. By fixing the values of $\mathbf{W}$ and $\mathbf{U}$, it essentially becomes an LMI in $\mathbf{F}$.

Then, to deal with the semi-infinite constraint of the relay power (43c), we can express $P_{\mathrm{R},m}$ as follows based on the definitions in (51):

$$P_{\mathrm{R},m} = \left\| \mathbf{t}_m + \sum_{k=1}^{K} \mathbf{T}_{m,k} \mathbf{h}_{m,k} \right\|^2. \tag{53}$$

Substituting (53) into (43c) and again applying Lemma 1, (43c) can be equivalently recast as the matrix inequality (49c), whose left-hand side is bilinear in $\mathbf{W}_m$ and $\mathbf{F}$, which is an LMI in $\mathbf{F}$ when $\mathbf{W}_m$ is fixed.

Finally, (43d) can be expressed as $\|\mathbf{f}_k\|^2 \le P^{\max}_{\mathrm{S},k}$, which can be equivalently recast as (49d) by using the Schur complement rule of [33]. The SDP form (49) is then readily obtained. ■

3) Relay AF Matrix Design: Since the constraint (49d) is independent of the relay AF matrices $\mathbf{W}$, this subproblem is equivalent to

$$\min_{t,\mathbf{W},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l},\boldsymbol{\tau}^{\mathrm{p}}_{m}} \quad t \quad \text{s.t.} \quad (49\text{b}), (49\text{c}). \tag{54}$$

The given problem becomes a standard SDP in $\mathbf{W}$ by noting that $\mathbf{Q}_{k,l}$ and $\mathbf{P}_m$ in (49b) and (49c), respectively, are LMIs in $\mathbf{W}$, provided that the other design variables are kept fixed.

The convergence analysis of the overall iterative algorithm, which solves problems (48), (49), and (54) with the aid of the block coordinate approach, is similar to that in Section III-B and therefore omitted for brevity. One slight difference from Algorithm 1 is that we initialize $\mathbf{F}^{(0)}_k = \sqrt{P^{\max}_{\mathrm{S},k}} \mathbf{I}_{N_{\mathrm{S},k}\times d_k} \, \forall k \in \mathcal{K}$ and $\mathbf{U}^{(0)}_k = \mathbf{I}_{d_k \times N_{\mathrm{S},k}} \forall k \in \mathcal{K}$, and the iterative algorithm will start by solving for the optimal $\mathbf{W}^{(1)}_m$. Solving (49) imposes a worst-case complexity on the order of $\mathcal{O}(N^2_{\mathrm{total}} D_{\mathrm{SDP}})$, where D_{SDP} represents the total dimensionality of the semi-definite cones in constraints (49b)–(49d). Comparing the SDP formulation of (49) derived for the norm-bounded CSI errors and the SOCP formulation in (62) deduced for the statistical CSI errors, the total dimensionality of (49) is seen to be significantly larger than that of (62).

V. Transceiver Design for the Quality-of-Service Problem

Here, we turn our attention to the joint transceiver design for the QoS problem (16). Following the same approaches as in Sections III and IV, the solution to the QoS problem can also be obtained by adopting the block coordinate update method. Since the derivations of the corresponding subproblems and algorithms are similar to those in Sections III and IV deduced for the min–max problem, we hereby only present the main results.

A. QoS Problem Under Statistical CSI Errors

1) Receive Filter Design: An optimal $\mathbf{u}_{k,l}$ can be obtained by minimizing $\bar{\varepsilon}_{k,l}(\Delta\mathbf{H}, \Delta\mathbf{G}_k)$ with respect to $\mathbf{u}_{k,l}$, which yields exactly the same solution as the Wiener filter in (28).

2) Source TPC Design: The specific subproblem of finding the optimal $\mathbf{F}$ can be solved by the following QCLP:

$$\min_{\mathbf{F},t} \quad t \tag{55a}$$

$$\text{s.t.} \quad \sum_{q=1}^{K} \mathbf{f}^H_q \mathbf{A}^{k,l}_{1,q} \mathbf{f}_q - 2\Re\left\{\mathbf{f}^H_k \mathbf{a}^{k,l}_2\right\} + a^{k,l}_3 \le \frac{\gamma}{\kappa_{k,l}} \quad \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{55b}$$

$$\sum_{k=1}^{K} \mathbf{f}^H_k \mathbf{A}^m_{4,k} \mathbf{f}_k \le \eta'_{\mathrm{R},m} \quad \forall m \in \mathcal{M} \tag{55c}$$

$$\mathrm{Tr}(\mathbf{F}^H_k \mathbf{F}_k) \le P^{\max}_{\mathrm{S},k} \quad \forall k \in \mathcal{K} \tag{55d}$$

where $\eta'_{\mathrm{R},m} \triangleq \rho_m t' - \sigma^2_{\mathrm{R},m} \mathrm{Tr}(\mathbf{W}_m \mathbf{W}^H_m)$.

3) Relay AF Matrix Design: The optimal $\mathbf{W}$ can be found by solving

$$\min_{\mathbf{w},t} \quad t \tag{56a}$$

$$\text{s.t.} \quad \mathbf{w}^H \mathbf{B}^{k,l}_1 \mathbf{w} - \sum_{m=1}^{M} 2\Re\left\{\mathbf{w}^H_m \mathbf{b}^{k,l}_{2,m}\right\} + \sum_{m=1}^{M} \mathbf{w}^H_m \mathbf{B}^{k,l}_{3,m} \mathbf{w}_m + b^{k,l}_4 \le \frac{\gamma}{\kappa_{k,l}} \, \forall k, l \tag{56b}$$

$$\mathbf{w}^H_m \mathbf{B}_{5,m} \mathbf{w}_m \le \rho_m t \quad \forall m \in \mathcal{M}. \tag{56c}$$

B. QoS Problem under Norm-Bounded CSI Errors

1) Receive Filter Design: The optimal $\mathbf{u}_{k,l}$ can be obtained from (48).

2) Source TPC Design: The optimal $\mathbf{F}$ can be obtained as the solution to the following SDP:

$$\min_{t,\mathbf{F},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l},\boldsymbol{\tau}^{\mathrm{p}}_{m}} \quad t \tag{57a}$$

$$\text{s.t.} \quad \mathbf{Q}'_{k,l} \succeq \mathbf{0} \quad \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{57b}$$

$$\mathbf{P}'_m \succeq \mathbf{0} \quad \forall m \in \mathcal{M} \tag{57c}$$

$$\begin{bmatrix} P^{\max}_{\mathrm{S},k} & \mathbf{f}_k^H \\ \mathbf{f}_k & \mathbf{I}_{N_{\mathrm{S},k} d_k} \end{bmatrix} \succeq \mathbf{0} \quad \forall k \in \mathcal{K} \tag{57d}$$

where $\mathbf{Q}'_{k,l}$ is obtained from $\mathbf{Q}_{k,l}$ in (46) upon replacing t by γ in the top-left entry (1,1). Similarly, $\mathbf{P}'_m$ can be obtained by substituting P_{R} with t in the (1,1)th entry of $\mathbf{P}_m$ in (50).

3) Relay AF Matrix Design: The optimal relay AF matrices are obtained by solving

$$\min_{t,\mathbf{W},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l}} \quad t \quad \text{s.t.} \quad (57\text{b}), (57\text{c}). \tag{58}$$

C. Initial Feasibility Search Algorithm

An important aspect of solving the given QoS problem is to find a feasible initial point. Indeed, it has been observed that, if the iterative algorithm is initialized with a random (possibly infeasible) point, the algorithm may fail at the first iteration. Finding a feasible initial point of a nonconvex problem, such as our QoS problem (16), is in general NP-hard. All these considerations motivate the study of an efficient initial feasibility search algorithm, which finds a reasonably "good" starting point for the QoS problem of (16).

Motivated by the "phase I" approach in general optimization theory [33], we formulate the feasibility check problem for the QoS problem as follows:

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad s \tag{59a}$$

$$\text{s.t.} \quad \kappa_{k,l} \mathcal{U}\left\{\varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right)\right\} \le s \ \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{59b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H \mathbf{F}_k\right) \le P^{\max}_{\mathrm{S},k} \quad \forall k \in \mathcal{K} \tag{59c}$$

where s is a slack variable, which represents an abstract measure for the violation of the constraint (16b). The given problem can be solved iteratively using the block coordinate approach until the objective value s converges or the maximum affordable number of iterations is reached. If, at the $(n+1)^{\mathrm{st}}$ iteration, $s^{(n+1)}$ meets the QoS target γ, then the procedure successfully finds a feasible initial point; otherwise, we claim that the QoS problem is infeasible. In this case, it is necessary to adjust γ or to drop the services of certain users by incorporating an admission control procedure, which, however, is beyond the scope of this paper.

Interestingly, (59) can be reformulated as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{\forall k \in \mathcal{K}, l \in \mathcal{D}_k} \kappa_{k,l} \mathcal{U}\left\{\varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right)\right\} \tag{60a}$$

$$\text{s.t.} \quad \mathcal{U}\left\{P_{\mathrm{R},m}\left(\Delta\mathbf{H}_m\right)\right\} \le \rho_m P^{\infty}_{\mathrm{R}} \quad \forall m \in \mathcal{M} \tag{60b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H \mathbf{F}_k\right) \le P^{\max}_{\mathrm{S},k} \quad \forall k \in \mathcal{K} \tag{60c}$$

where we have $P^{\infty}_{\mathrm{R}} \to \infty$, which is equivalent to removing the constraint on the relay's transmit power. In fact, (60) becomes exactly the same as the min–max problem of (15) upon setting $P_{\mathrm{R}} = P^{\infty}_{\mathrm{R}}$. We therefore propose an efficient iterative feasibility search algorithm, which is listed as Algorithm 2, based on the connection between the feasibility check and the min–max problems.

Algorithm 2 Iterative Initial Feasibility Search Algorithm for the QoS problems

---

1: **repeat**
2: Solve one cycle of the problem (60) and denote the current objective value by $\hat{\gamma}^{(i+1)}$;
3: Verify if $\hat{\gamma}^{(i+1)} \le \gamma$, and if so, stop the algorithm;
4: $i \leftarrow i + 1$;
5: **until** Termination criterion is satisfied, e.g., $|\hat{\gamma}^{(i)} - \hat{\gamma}^{(i-1)}| \le \epsilon$; or the maximum allowed number of iteration is reached.

Based on the definition of $\mathcal{U}\{\cdot\}$ in (14), Algorithm 2 is applicable to the QoS problems associated with both types of CSI errors considered. Furthermore, Algorithm 2 indeed provides a feasible initial point for the QoS problem if it exists. Otherwise, it provides a certificate of infeasibility if $\hat{\gamma}^{(i+1)} > \gamma$ after a few iterations. Then, the QoS problem is deemed infeasible in this case, and the admission control procedure may deny the access of certain users.

VI. Simulation Experiments and Discussions

This section presents our Monte Carlo simulation results for verifying the resilience of the proposed transceiver optimization algorithms against CSI errors. In all simulations, we assume that there are $K = 2$ S−D pairs, which communicate with the assistance of $M = 2$ relays. Each node is equipped with $N_{\mathrm{S},k} = N_{\mathrm{R},m} = N_{\mathrm{D},k} = 3$ antennas $\forall\, k \in \mathcal{K}, m \in \mathcal{M}$. Each source transmits 2 independent quadrature phase-shift keying (QPSK) modulated data streams to its corresponding destination, i.e., $d_k = 2 \ \forall\, k \in \mathcal{K}$. Equal noise variances of $\sigma^2_{\mathrm{D},k} = \sigma^2_{\mathrm{R},m}$ are assumed. The maximum source and relay transmit power is normalized to one, i.e., we have $P^{\max}_{\mathrm{S},k} = 1 \ \forall\, k \in \mathcal{K}$ and $\rho_m P_{\mathrm{R}} = 1,\ \forall\, m \in \mathcal{M}$. Equal weights of $\kappa_{k,l}$ are assigned to the different data streams, unless otherwise stated. The channels are assumed to be flat fading, with the coefficients given by i.i.d. zero-mean unit-variance complex Gaussian random variables. The signal-to-noise ratios (SNRs) at the relays and the destinations are defined as $\mathrm{SNR}_{\mathrm{R},m} \triangleq P^{\max}_{\mathrm{S}} / |N_{\mathrm{R},m} \sigma^2_{\mathrm{R},m}|$ and $\mathrm{SNR}_{\mathrm{D},k} \triangleq P^{\max}_{\mathrm{R}} / |N_{\mathrm{D},k} \sigma^2_{\mathrm{D},k}|$, respectively. The optimization solver MOSEK [31] is used for solving each optimization problem.

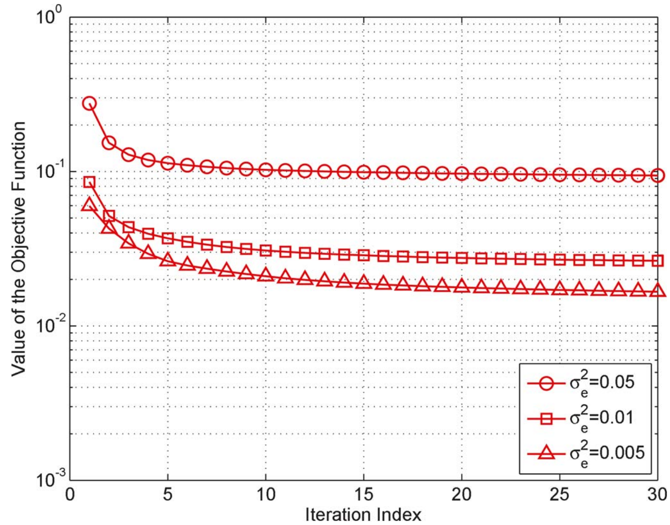

Fig. 2. Convergence behavior of the proposed iterative algorithm with statistical CSI errors.

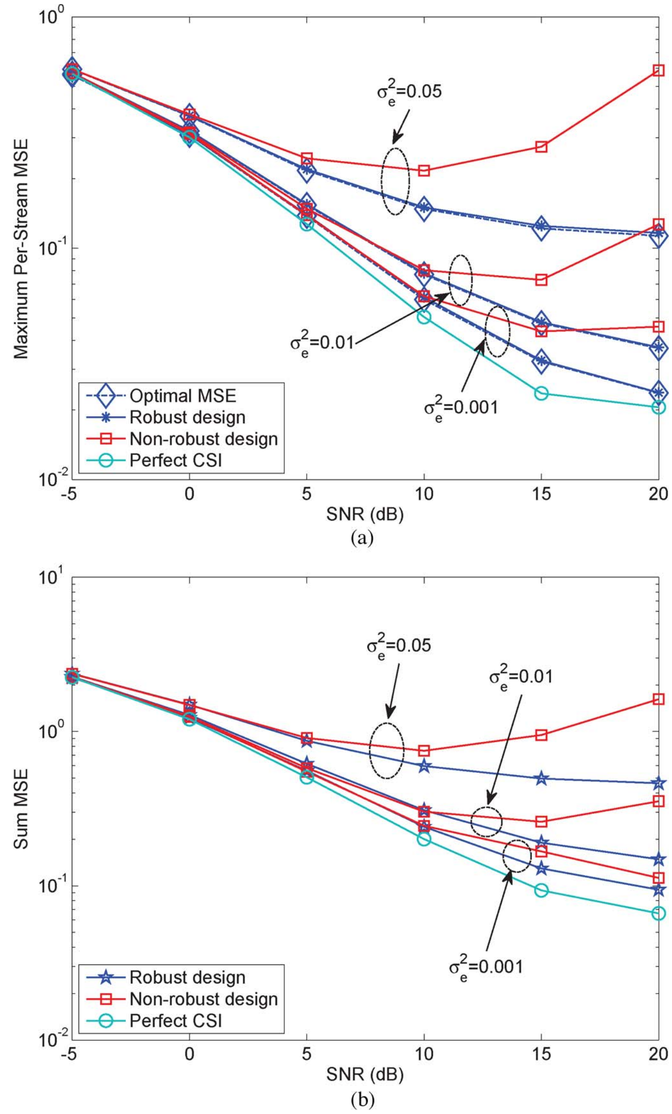

Fig. 3. MSE performance of different design approaches versus SNR. (a) Maximum per-stream MSE. (b) Sum MSE ($\mathrm{SNR}_{\mathrm{R},m} = \mathrm{SNR}_{\mathrm{D},k} = \mathrm{SNR}$, $\alpha = \beta = 0.5$).

A. Performance Evaluation Under Statistical CSI Errors

We first evaluate the performance of the iterative algorithm proposed in Section III under statistical CSI errors. The channel correlation matrices in (9) and (10) are obtained by the widely employed exponential model of [37]. Specifically, their entries are given by $[\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}]_{i,j} = [\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}]_{i,j} = \alpha^{|i-j|}$ and $[\boldsymbol{\Psi}_{\mathrm{H}_{m,k}}]_{i,j} = [\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}]_{i,j} = \sigma_e^2 \beta^{|i-j|}, i, j \in \{1, 2, 3\}$, where α and β are the correlation coefficients, and σ_e^2 denotes the variance of the CSI errors. The available channel estimates $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$ are generated according to $\hat{\mathbf{H}}_{m,k} \sim \mathcal{CN}(\mathbf{0}_{N_{\mathrm{R},m} \times N_{\mathrm{S},k}}, ((1-\sigma_e^2)/\sigma_e^2)\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}} \otimes \boldsymbol{\Psi}^T_{\mathrm{H}_{m,k}})$ and $\hat{\mathbf{G}}_{k,m} \sim \mathcal{CN}(\mathbf{0}_{N_{\mathrm{D},k} \times N_{\mathrm{R},m}}, ((1-\sigma_e^2)/\sigma_e^2)\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \otimes \boldsymbol{\Psi}^T_{\mathrm{G}_{k,m}})$, respectively, such that the entries of the true channel matrices have unit variances. We compare the robust transceiver design proposed in Algorithm 1 to the 1) nonrobust design, which differs from the robust design in that it assumes $\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}} = \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} = \mathbf{0}$ and $\boldsymbol{\Psi}_{\mathrm{H}_{m,k}} = \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} = \mathbf{0}$, i.e., it neglects the effects of the CSI errors; 2) perfect CSI case, where the true channel matrices $\mathbf{H}_{m,k}$ and $\mathbf{G}_{k,m}$ are used instead of the estimates $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$ in Algorithm 1 and where there are no CSI errors, i.e., we have $\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}} = \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} = \mathbf{0}$ and $\boldsymbol{\Psi}_{\mathrm{H}_{m,k}} = \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} = \mathbf{0}$. The curves labeled "optimal MSE" correspond to the value of the objective function in (17a) after optimization by Algorithm 1. In all the simulation figures, the MSEs of the different approaches are calculated by averaging the squared error between the transmitted and estimated experimental data symbols over 1000 independent CSI error realizations and 10 000 QPSK symbols for each realization.

As a prelude to the presentation of our main simulation results in the following, the convergence behavior of Algorithm 1 is presented for different CSI error variances, It can be observed in Fig, 2 that in all cases, the proposed algorithm can converge within a reasonable number of iterations, Therefore, in our experimental work, we set the number of iterations to a fixed value of 5, and the resultant performance gains will be discussed in the following.

1) Experiment A.1 (MSE Performance): In Fig. 3(a), the maximum per-stream MSE among all the data streams is shown as a function of the SNR for different values of CSI error variance. It is observed that the proposed robust design approach achieves better resilience against the CSI errors than the non-robust design approach. The performance gains become more evident in the medium-to-high SNR range. For the nonrobust design, degradations are observed because the MSE obtained at high SNRs is dominated by the interference, rather than by the noise. Therefore, the relays are confined to relatively low transmit power in order to control the interference. This, in turn, leads to performance degradation imposed by the CSI errors. In contrast, the proposed robust design is capable of compensating for the extra interference imposed by the CSI errors, thereby demonstrating its superiority over its nonrobust counterpart. Furthermore, we observe that the "Optimal MSE" and our simulation results tally well, which justifies the approximations invoked in calculating the per-stream MSE in (13). In addition

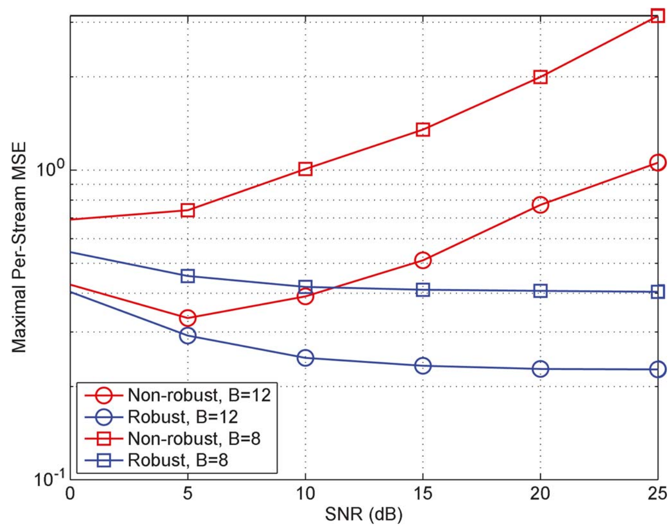

Fig. 4. Per-stream MSE performance with the optimized codebook based on the GLA-VQ. ($B = 8$ corresponds to $\sigma_e^2 = 0.334$, and $B = 12$ corresponds to $\sigma_e^2 = 0.175$.)

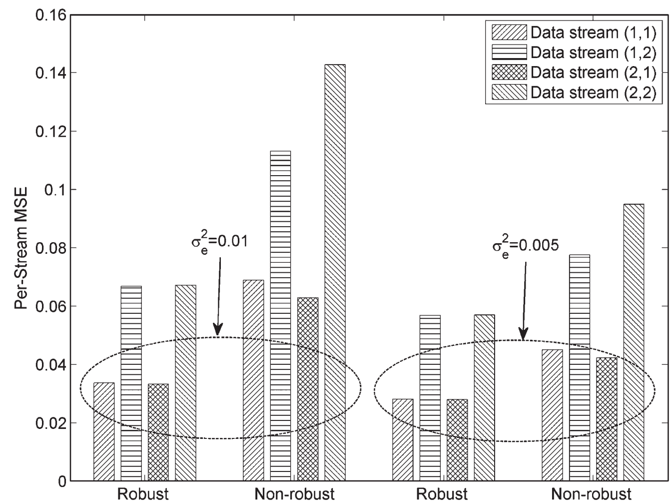

Fig. 5. Comparison of the per-stream MSEs of the robust and nonrobust design approaches ($\mathrm{SNR}_{\mathrm{R},m} = \mathrm{SNR}_{\mathrm{D},k} = 15$ dB, and $\alpha = \beta = 0.5$).

to the per-stream performance, the overall system performance[4] quantified in terms of the sum MSE of different approaches is examined in Fig. 3(b), where a similar trend to that of Fig. 3(a) can be observed.

The MSE performance associated with a limited number of feedback bits is also studied. To this end, we assume that each user is equipped with a codebook that is optimized using the generalized Lloyd algorithm of vector quantization (GLA-VQ) [38]. Each user then quantizes the channel vector, and the corresponding codebook index is fed back to the central processing unit. The results presented in Fig. 4 show that the proposed algorithm significantly outperformed the nonrobust one for the different number of quantization bits considered.

2) Experiment A.2 (Data Stream Fairness): Next, we examine the accuracy of the proposed robust design in providing weighted fairness for the different data streams. To this end, we set the weights for the different data streams to be $\kappa_{1,1} = \kappa_{2,1} = 1/3$ and $\kappa_{1,2} : \kappa_{2,2} = 1/6$. Fig. 5 shows the MSE of each data stream for different values of the error variance. Comparing the two methods, the robust design approach results in significantly better weighted fairness than the nonrobust one. In particular, the MSEs obtained are strictly inversely proportional to the predefined weights. This feature is particularly desirable for multimedia communications, where the streams corresponding to different service types may have different priorities.

3) Experiment A.3 (Effects of Channel Correlation): The effects of channel correlations on the MSE performance of the different approaches are investigated in Fig. 6. It can be observed that the performance of all the approaches is degraded as the correlation factor α increases. While the robust design

[4]Note that the objective of portraying the sum MSE performance is to validate whether the proposed robust design approach can also achieve a performance gain over the nonrobust approach in terms of its overall performance. In fact, the sum MSE performance can be optimized by solving a design problem with the sum MSE being the objective function.

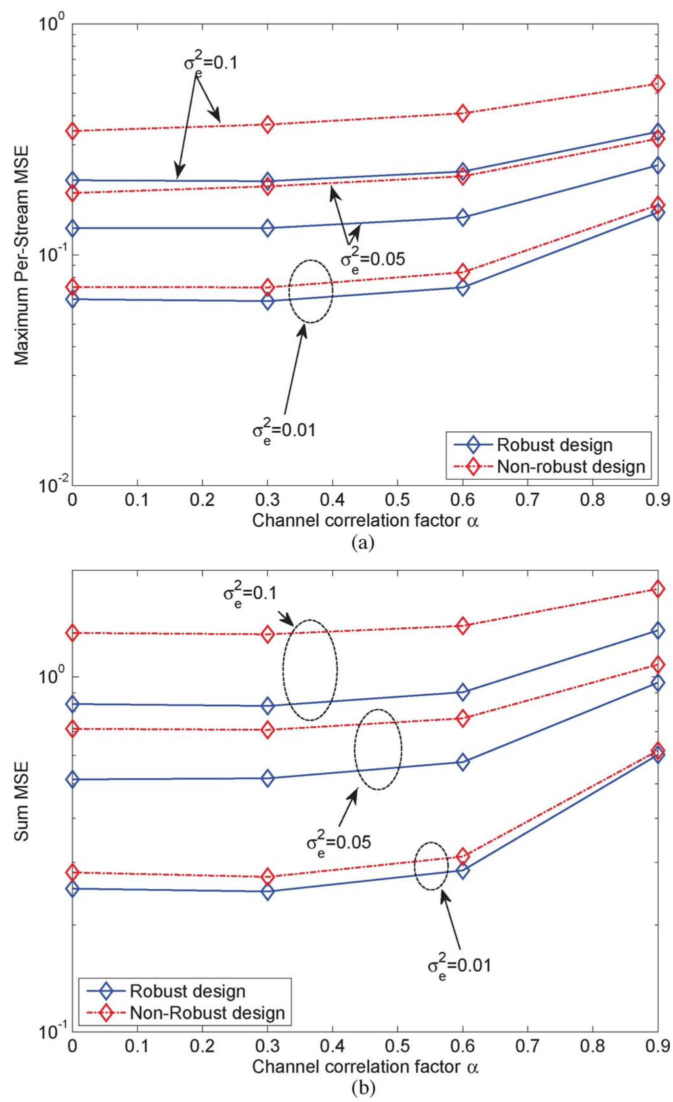

Fig. 6. MSE performance of different design approaches versus correlation factor of the source–relay channels. (a) Per-stream MSE. (b) Sum MSE ($\mathrm{SNR}_{\mathrm{R},m} = \mathrm{SNR}_{\mathrm{D},k} = 10$ dB, and $\beta = 0.45$).

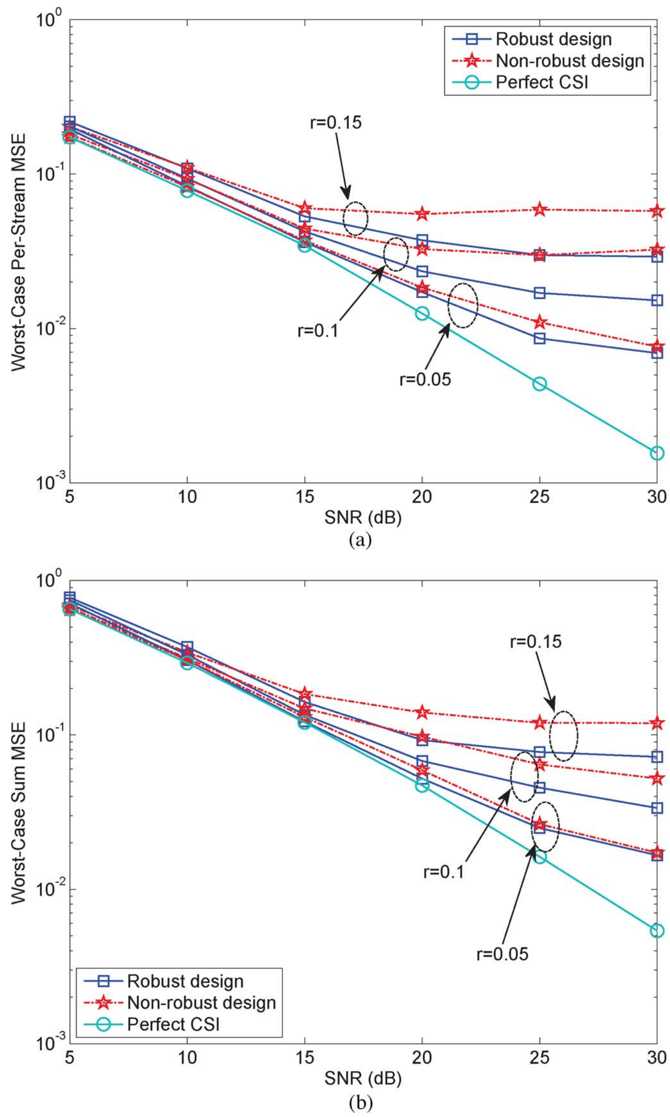

Fig. 7. MSE performance of different design approaches versus SNR. (a) Worst-case per-stream MSE. (b) Worst-case sum MSE.

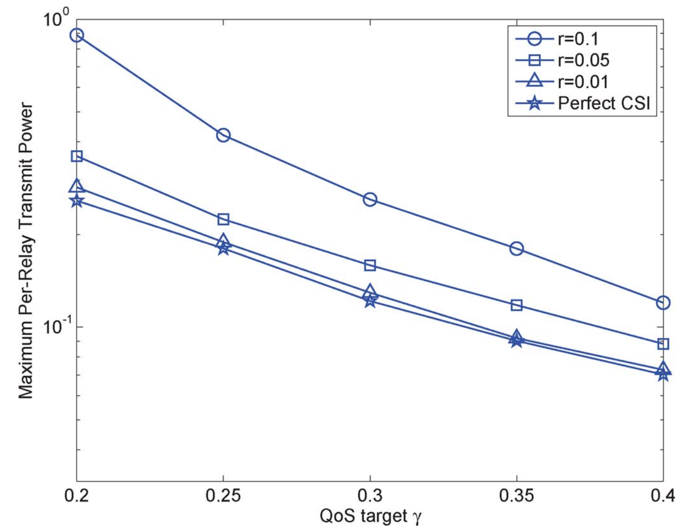

Fig. 8. Maximum relay transmit power versus QoS targets with different uncertainty sizes of the CSI errors.

shows consistent performance gains over its nonrobust one associated with different α and σ_e^2, the discrepancies between the two approaches tend to become less significant with an increase in α. This is because the achievable *spatial multiplexing* gain is reduced by a higher channel correlation; therefore, the robust design can only attain a limited performance improvement in the presence of high channel correlations.

B. Performance Evaluation Under Norm-Bounded CSI Errors

Here, we evaluate the performance of the proposed worst case design approach in Section V for the min–max problem under norm-bounded CSI errors. Similar to that given earlier, we compare the proposed robust design approach both to the nonrobust approach and to the perfect CSI scenario. We note that the power of each relay is a function of $\Delta\mathbf{H}_m$. According to the worst-case robust design philosophy, the maximum relay transmit power has to be bounded by the power budget, whereas the average relay transmit power may become significantly lower than that of the nonrobust design. To facilitate a fair comparison of the different approaches, we therefore assume the absence of CSI errors for the $\mathtt{S}-\mathtt{R}$ links, i.e., we have $\Delta\mathbf{H}_{m,k} = \mathbf{0}$. For the $\mathtt{R}-\mathtt{D}$ links, we consider the uncertainty regions with equal radius, i.e., we have $\xi_{k,m} = r\ \forall k \in \mathcal{K}, m \in \mathcal{M}$. To determine the worst-case per-stream MSE, we generate 5000 independent realizations of the CSI errors. For each realization, we evaluate the maximum per-stream MSE averaged over 1000 QPSK symbols and random Gaussian noise. Then, the worst-case per-stream MSE is obtained by selecting the largest one among all the realizations.

1) Experiment B.1 (MSE Performance): The worst-case per-stream MSE and the worst-case sum MSE are reported in Fig. 7 as a function of the SNR. Three sizes of the uncertainty region are considered, i.e., $r = 0.05$, $r = 0.1$, and $r = 0.15$. Focusing on the first case, it can be seen that the performance achieved by our robust design approach first monotonically decreases as the SNR increases and then subsequently remains approximately constant at high-SNR values. This is primarily because, at low SNR, the main source of error in the estimation of the data streams is the channel noise. At high SNR, the channel noise is no longer a concern, and the MSE is dominated by the CSI errors. Observe also in Fig. 7 that for $r = 0.1$ and $r = 0.15$, the MSE is clearly higher, although it presents a similar trend to the case of $r = 0.5$. The performance gain achieved by the robust design also becomes more noticeable for these larger sizes of the uncertainty regions.

2) Experiment B.2 (Relay Power Consumption): Next, we investigate the performance of the approach proposed in Section VI for the QoS problem under the norm-bounded CSI errors. The maximum per-relay transmit power is plotted in Fig. 8 as a function of the QoS target γ for different sizes of uncertainty regions. As expected, it can be observed that the relay power for all cases decreases as the QoS target is relaxed. An important observation from this figure is that, when the size of uncertainty region is large, the required relay transmit power becomes significantly higher than the perfect CSI case. From an

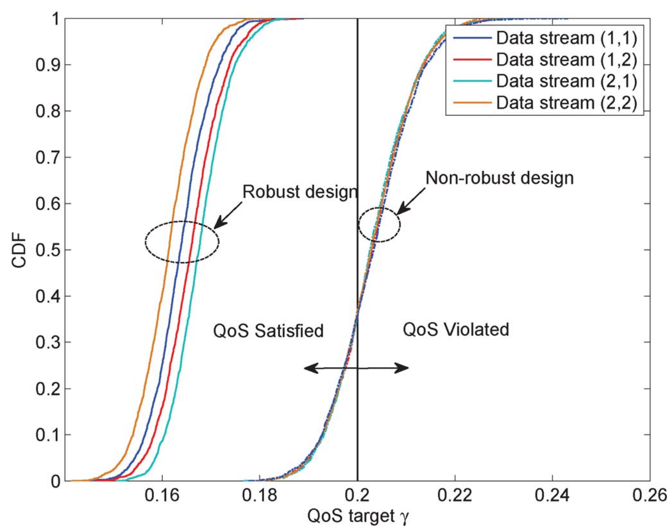

Fig. 9. CDFs of per-stream MSEs using the robust and nonrobust approaches for SNR = 5 dB.

energy-efficient design perspective, this is not desirable, which motivates the consideration of the min–max design in such applications.

3) Experiment B.3 (CDF of Per-stream MSE): Finally, we evaluate how consistently the QoS constraints of all the data streams can be satisfied by the proposed design approach for the QoS problem. In this experiment, the CSI errors of both the S−R and R−D links are taken into consideration and generated according to the i.i.d. zero-mean complex Gaussian distribution with a variance of $\sigma_e^2 = 0.001$. Then, the probability that the CSI errors are bounded by the predefined radius r can be formulated as [9, Sec. IV-C]

$$\Pr\left\{\|\mathbf{h}_{m,k}\|^2 \le r^2\right\} = \Pr\left\{\|\mathbf{g}_{k,m}\|^2 \le r^2\right\} = \frac{1}{\Gamma\left(\frac{N^2}{2}\right)}\gamma\left(\frac{N^2}{2}, \frac{r^2}{\sigma_e^2}\right) \quad (61)$$

where $\Gamma(\cdot)$ and $\gamma(\cdot,\cdot)$, respectively, denote the complete and lower incomplete Gamma functions. Given the required bounding probability of, e.g., 90% in the simulation, the radius r can be numerically determined from (61). Fig. 9 shows the cumulative distribution functions (cdfs) of the MSE of each data stream using both the robust and nonrobust design methods. As expected, the proposed robust method ensures that the MSE of each data stream never exceeds the QoS target shown as the vertical black solid line in Fig. 9. By contrast, for the nonrobust design, the MSE frequently violates the QoS target, namely for more than 60% of the realizations. Based on these observations, we conclude that the proposed robust design approach outperforms its nonrobust counterpart in satisfying the QoS constraints for all the data streams.

VII. Conclusion

Jointly optimized source TPCs, AF relay matrices, and receive filters were designed by considering two different types of objective functions with specific QoS consideration in the presence of CSI errors in both the S−R and R−D links. To this end, a pair of practical CSI error models, namely, the statistical and the norm-bounded models were considered. Accordingly, the robust transceiver design approach was formulated to minimize the maximum per-stream MSE subject to the source and relay power constraints (min–max problem). To solve the nonconvex optimization problems formulated, an iterative solution based on the block coordinate update algorithm was proposed, which involves a sequence of convex conic optimization problems. The proposed algorithm generated a convergent sequence of objective function values. The problem of relay power minimization subject to specific QoS constraints and to source power constraints was also studied. An efficient feasibility search algorithm was proposed by studying the link between the feasibility check and the min–max problems. Our simulation results demonstrate a significant enhancement in the performance of the proposed robust approaches over the conventional nonrobust approaches.

Appendix A
Transformation of (34) Into a standard Second-Order Cone Programming

By exploiting the separable structure of (34) and the properties of quadratic terms, the problem can be recast as

$$\min_{\substack{t,\{\mathbf{f}_k\},\\ \{\boldsymbol{\lambda}^{k,l}\},\{\boldsymbol{\theta}^m\}}} \quad t \quad (62a)$$

$$\text{s.t.} \quad \left\|\left(\mathbf{A}_{1,q}^{k,l}\right)^{1/2}\mathbf{f}_q\right\| \le \lambda_q^{k,l}$$
$$\forall q,k \in \mathcal{K}, q \ne k, l \in \mathcal{D}_k \quad (62b)$$

$$\left\|\left(\mathbf{A}_{1,k}^{k,l}\right)^{1/2}\mathbf{f}_k - \left(\mathbf{A}_{1,k}^{k,l}\right)^{-1/2}\mathbf{a}_2^{k,l}\right\| \le \lambda_k^{k,l}$$
$$\forall k \in \mathcal{K}, l \in \mathcal{D}_k \quad (62c)$$

$$\|\boldsymbol{\lambda}^{k,l}\|^2 - \left(\mathbf{a}_2^{k,l}\right)^H\left(\mathbf{A}_{1,k}^{k,l}\right)^{-1}\mathbf{a}_2^{k,l} + a_3^{k,l} \le \frac{t}{\kappa_{k,l}}$$
$$\forall k \in \mathcal{K}, l \in \mathcal{D}_k \quad (62d)$$

$$\left\|\left(\mathbf{A}_{4,k}^m\right)^{1/2}\mathbf{f}_k\right\| \le \theta_k^m \ \forall k \in \mathcal{K}, m \in \mathcal{M} \quad (62e)$$

$$\|\boldsymbol{\theta}^m\| \le \sqrt{\eta_{\mathrm{R},m}} \ \forall m \in \mathcal{M} \quad (62f)$$

$$\|\mathbf{f}_k\| \le \sqrt{P_{\mathrm{S},k}^{\max}} \ \forall k \in \mathcal{K} \quad (62g)$$

where $\boldsymbol{\lambda}^{k,l} = [\lambda_1^{k,l}, \ldots, \lambda_K^{k,l}]^T$, $\boldsymbol{\theta}^m = [\theta_1^m, \ldots, \theta_K^m]^T$, and t are auxiliary variables. The main difficulty in solving this problem is with (62d), which is a so-called *hyperbolic constraint* [32], whereas the remaining constraints are already in the form of SOC.

To tackle (62d), we observe that, for any $\mathbf{x}$ and $y, z \le 0$, the following equation holds:

$$\|\mathbf{x}\|^2 \le yz \Longleftrightarrow \left\|\begin{bmatrix} 2\mathbf{x} \\ y - z \end{bmatrix}\right\| \le y + z. \quad (63)$$

We can apply (63) to transform (62d) into

$$\left\|\begin{bmatrix} 2\boldsymbol{\lambda}^{k,l} \\ \frac{t}{\kappa_{k,l}} + \left(\mathbf{a}_2^{k,l}\right)^H \left(\mathbf{A}_{1,k}^{k,l}\right)^{-1} \mathbf{a}_2^{k,l} - a_3^{k,l} - 1 \end{bmatrix}\right\| \le \frac{t}{\kappa_{k,l}} + \left(\mathbf{a}_2^{k,l}\right)^H \left(\mathbf{A}_{1,k}^{k,l}\right)^{-1} \mathbf{a}_2^{k,l} - a_3^{k,l} + 1. \quad (64)$$

Therefore, substituting (62d) by (64), we can see that (62) is in the form of a standard SOCP.

Appendix B
Proof of Proposition 1

First, we define $\boldsymbol{\mathcal{T}}_k \triangleq [\boldsymbol{\mathcal{T}}_{k,1}, \ldots, \boldsymbol{\mathcal{T}}_{k,K}]$ and $\boldsymbol{\mathcal{G}}_k \triangleq [\sigma_{\mathrm{R},1}\boldsymbol{\mathcal{G}}_{k,1}, \ldots, \sigma_{\mathrm{R},M}\boldsymbol{\mathcal{G}}_{k,M}]$. We exploit the fact that, for any vectors $\{\mathbf{a}_k\}_{k=1}^N$, the following identity holds:

$$\sum_{k=1}^{N} \|\mathbf{a}_k\|^2 = \left\|\left[\mathbf{a}_1^T, \ldots, \mathbf{a}_N^T\right]\right\|^2. \quad (65)$$

The per-stream MSE (13) can be subsequently expressed as

$$\begin{aligned} \varepsilon_{k,l} = \Bigg\| & \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{T}}_k + \sum_{m=1}^{M} \mathbf{u}_{k,l}^H \Delta\mathbf{G}_{k,m} \left[\boldsymbol{\mathcal{W}}_{m,1}\mathbf{F}_1, \ldots, \boldsymbol{\mathcal{W}}_{m,K}\mathbf{F}_K\right] \\ & + \sum_{q=1}^{K}\sum_{m=1}^{M} \Big[\mathbf{0}_{1\times\sum_{t=1}^{q} d_t}, \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_{k,m} \\ & \qquad \times \Delta\mathbf{H}_{m,q}\mathbf{F}_q, \mathbf{0}_{1\times\sum_{q+1}^{K} d_t}\Big] \Bigg\|^2 \\ & + \Bigg\| \sum_{m=1}^{M} \Big[\mathbf{0}_{1\times\sum_{p=1}^{m-1} N_{\mathrm{R},p}}, \mathbf{u}_{k,l}^H \Delta\mathbf{G}_{k,m}\mathbf{W}_m, \\ & \qquad \mathbf{0}_{1\times\sum_{p=m+1}^{M} N_{\mathrm{R},p}}\Big] \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_k \Bigg\|^2 + \sigma_{\mathrm{D},k}^2 \|\mathbf{u}_{k,l}^H. \end{aligned} \quad (66)$$

Upon applying the identity $\mathrm{vec}^T(\mathbf{ABC}) = \mathrm{vec}(\mathbf{B})^T(\mathbf{C} \otimes \mathbf{A}^T)$ to (66), we arrive at

$$\begin{aligned} \varepsilon_{k,l} = & \left\| \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{T}}_k - \bar{\mathbf{e}}_{k,l}^T + \sum_{m=1}^{M} \mathbf{g}_{k,m}^T \mathbf{C}_{1,m}^{k,l} + \sum_{m,q} \mathbf{h}_{m,q}^T \mathbf{D}_{m,q}^{k,l} \right\|^2 \\ & + \left\| \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_k + \sum_{m=1}^{M} \mathbf{g}_{k,m}^T \mathbf{C}_{2,m}^{k,l} \right\|^2 + \left\| \sigma_{\mathrm{D},k} \mathbf{u}_{k,l}^H \right\|^2 \end{aligned} \quad (67)$$

where $\mathbf{h}_{m,k} \triangleq \mathrm{vec}(\Delta\mathbf{H}_{m,k})$ and $\mathbf{g}_{k,m} \triangleq \mathrm{vec}(\Delta\mathbf{G}_{k,m})$ denote the vectorized CSI errors, $\bar{\mathbf{e}}_{k,l} \triangleq [\mathbf{0}_{1\times\sum_{t=1}^{k-1} d_t}, \mathbf{e}_{k,l}^T, \mathbf{0}_{1\times\sum_{t=k+1}^{K} d_t}]^T$, and the following matrices have also been introduced:

$$\mathbf{C}_{1,m}^{k,l} \triangleq \left[(\boldsymbol{\mathcal{W}}_{m,1}\mathbf{F}_1) \otimes \mathbf{u}_{k,l}^*, \ldots, (\boldsymbol{\mathcal{W}}_{m,K}\mathbf{F}_K) \otimes \mathbf{u}_{k,l}^*\right] \quad (68)$$

$$\mathbf{C}_{2,m}^{k,l} \triangleq \Big[\mathbf{0}_{N_{\mathrm{D},k}N_{\mathrm{R},m}\times\sum_{p=1}^{m-1} N_{\mathrm{R},p}}, \mathbf{W}_m \otimes \mathbf{u}_{k,l}^* \; \mathbf{0}_{N_{\mathrm{D},k}N_{\mathrm{R},m}\times\sum_{p=m+1}^{M} N_{\mathrm{R},p}}\Big] \quad (69)$$

$$\mathbf{D}_{m,q}^{k,l} \triangleq \Big[\mathbf{0}_{N_{\mathrm{S},q}N_{\mathrm{R},m}\times\sum_{t=1}^{q-1} d_t}, \mathbf{F}_q \otimes \left(\boldsymbol{\mathcal{G}}_{k,m}^T \mathbf{u}_{k,l}^*\right) \; \mathbf{0}_{N_{\mathrm{S},q}N_{\mathrm{R},m}\times\sum_{t=q+1}^{K} d_t}\Big]. \quad (70)$$

Again, by exploiting the property in (65), we can write (67) in a more compact form as follows:

$$\begin{aligned} \varepsilon_{k,l} = \Bigg\| & \underbrace{\left[\mathbf{u}_{k,l}^H \boldsymbol{\mathcal{T}}_k - \bar{\mathbf{e}}_{k,l}, \mathbf{u}_{k,l}^H \overline{\boldsymbol{\mathcal{G}}}_k, \sigma_{\mathrm{D},k}\mathbf{u}_{k,l}^H\right]}_{\boldsymbol{\theta}_{k,l}} \\ & + \sum_{m=1}^{M} \mathbf{g}_{k,m}^T \underbrace{\left[\mathbf{C}_{1,m}^{k,l}, \mathbf{C}_{2,m}^{k,l}, \mathbf{0}_{N_{\mathrm{D},k}N_{\mathrm{R},m}\times N_{\mathrm{D},k}}\right]}_{\boldsymbol{\Theta}_m^{k,l}} \\ & + \sum_{m=1}^{M}\sum_{q=1}^{K} \mathbf{h}_{m,q}^T \underbrace{\left[\mathbf{D}_{m,q}^{k,l}, \mathbf{0}_{N_{\mathrm{R},m}N_{\mathrm{S},q}\times N_{\mathrm{R}}+N_{\mathrm{D},k}}\right]}_{\boldsymbol{\Phi}_{m,q}^{k,l}} \Bigg\|^2. \end{aligned} \quad (71)$$

Substituting (71) into (43b), we can express (43b) as

$$\begin{aligned} & \left(\boldsymbol{\theta}_{k,l} + \sum_{m=1}^{M} \mathbf{g}_{k,m}^T \boldsymbol{\Theta}_m^{k,l} + \sum_{m=1}^{M}\sum_{q=1}^{K} \mathbf{h}_{m,q}^T \boldsymbol{\Phi}_{m,q}^{k,l}\right) \\ & \times \left(\boldsymbol{\theta}_{k,l} + \sum_{m=1}^{M} \mathbf{g}_{k,m}^T \boldsymbol{\Theta}_m^{k,l} + \sum_{m=1}^{M}\sum_{q=1}^{K} \mathbf{h}_{m,q}^T \boldsymbol{\Phi}_{m,q}^{k,l}\right)^H \le t \end{aligned} \quad (72)$$

where the uncertain blocks $\mathbf{h}_{m,k}$ and $\mathbf{g}_{k,m}$ should satisfy $\|\mathbf{h}_{m,k}\|_S = \|\mathbf{h}_{m,k}\| \le \xi_{m,k}$ and $\|\mathbf{g}_{k,m}\|_S = \|\mathbf{g}_{k,m}\| \le \eta_{k,m}$, respectively. Through a direct application of Lemma 1, (72) can readily be recast as (46) where the nonnegativity of $\boldsymbol{\tau}_{k,l}^{\mathrm{G}}$ and $\boldsymbol{\tau}_{k,l}^{\mathrm{H}}$ has been implicitly included in the positive semi-definite nature of $\mathbf{Q}_{k,l}$.

Acknowledgment

The authors would like to thank the Editor and the anonymous reviewers for their valuable comments, which have helped to improve the quality and the presentation of this paper.

References

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

[2] S. W. Peters, A. Y. Panah, K. T. Truong, and R. W. Heath, "Relay architectures for 3GPP LTE-advanced," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 1, Jul. 2009, Art. ID. 618 787.

[3] C. Nie, P. Liu, T. Korakis, E. Erkip, and S. S. Panwar, "Cooperative relaying in next-generation mobile WiMAX networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 299–305, Mar. 2013.

[4] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution towards 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Commun.*, vol. 21, no. 3, pp. 118–127, Jun. 2014.

[5] S. Berger, M. Kuhn, A. Wittneben, T. Unger, and A. Klein, "Recent advances in amplify-and-forward two-hop relaying," *IEEE Commun. Mag.*, vol. 47, no. 7, pp. 50–56, Jul. 2009.

[6] 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.

[7] A. El-Keyi and B. Champagne, "Adaptive linearly constrained minimum variance beamforming for multiuser cooperative relaying using the Kalman filter," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 641–651, Feb. 2010.

[8] 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.

[9] B. K. Chalise and L. Vandendorpe, "MIMO relay design for multipoint-to-multipoint communications with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2785–2796, Jul. 2009.
[10] B. K. Chalise and L. Vandendorpe, "Optimization of MIMO relays for multipoint-to-multipoint communications: Nonrobust and robust designs," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6355–6368, Dec. 2010.
[11] J. Lee, J.-K. Han, and J. Zhang, "MIMO technologies in 3GPP LTE and LTE-advanced," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 1, Jul. 2009, Art. ID. 302092.
[12] M. Fadel, A. El-Keyi, and A. Sultan, "QOS-constrained multiuser peer-to-peer amplify-and-forward relay beamforming," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1397–1408, Mar. 2012.
[13] W. Liu, L.-L. Yang, and L. Hanzo, "SVD-assisted multiuser transmitter and multiuser detector design for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 1016–1021, Feb. 2009.
[14] 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.
[15] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Singapore: Athena Scientific, 1999.
[16] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
[17] 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.
[18] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3678–3689, Aug. 2008.
[19] A. Pascual-Iserte, D. P. Palomar, A. I. Perez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 346–360, Jan. 2006.
[20] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM J. Imag. Sci.*, vol. 6, no. 3, pp. 1758–1789, Sep. 2013.
[21] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 168–181, Jan. 2005.
[22] 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.
[23] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "On the design of amplify-and-forward MIMO-OFDM relay systems with QoS requirements specified as Schur-convex functions of the MSEs," *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1871–1877, May 2013.
[24] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.
[25] 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.
[26] D. Zheng, J. Liu, K.-K. Wong, H. Chen, and L. Chen, "Robust peer-to-peer collaborative-relay beamforming with ellipsoidal CSI uncertainties," *IEEE Signal Process. Lett.*, vol. 16, no. 4, pp. 442–445, Apr. 2012.
[27] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2005.
[28] C. Xing, S. Ma, and Y.-C. Wu, "Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2273–2283, Apr. 2010.
[29] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, Version 2.0 beta," Sep. 2013. [Online]. Available: http://cvxr.com/cvx
[30] 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.
[31] E. D. Andersen and K. D. Andersen, *MOSEK Modeling Manual*, Aug. 2013. [Online]. Available: http://mosek.com
[32] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Appl.*, vol. 284, no. 1–3, pp. 193–228, Nov. 1998.
[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U. K.: Cambridge Univ. Press, 2004.
[34] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Time Methods in Convex Programming: Theory and Applications*. Philadelphia, PA, USA: SIAM, 1994.
[35] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA, USA: SIAM, 1994.
[36] S. Liu, "Matrix results on the Khatri-Rao and Tracy-Singh products," *Linear Algebra Appl.*, vol. 289, no. 1–3, pp. 267–277, Mar. 1999.
[37] 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.
[38] A. D. Dabbagh and D. J. Love, "Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch," *IEEE Trans. Wireless Commun.*, vol. 56, no. 11, pp. 1859–1868, Nov. 2008.

Jiaxin Yang (S'11) received the B.Eng. degree in information engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and the M.E.Sc. degree in electrical and computer engineering from the University of Western Ontario, London, ON, Canada, in 2011. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC.

His research interests include optimization theory, statistical signal processing, detection and estimation, and the applications thereof in wireless communications such as multiple-input–multiple-output systems, cooperative communications, and physical-layer security.

Mr. Yang has received several award and scholarships, including the McGill Engineering Doctoral Award, the Graduate Excellence Fellowship, the International Differential Tuition Fee Waivers, the FQRNT International Internship Scholarship, and the PERSWADE Ph.D. Scholarship.

Benoit Champagne (S'87–M'89–SM'03) received the B.Ing. degree in engineering physics from École Polytechnique de Montréal, Montreal, QC, Canada, in 1983; the M.Sc. degree in physics from Université de Montréal, in 1985; and the Ph.D. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1990.

From 1990 to 1999, he was an Assistant and then Associate Professor with INRS-Telecommunications, Université du Quebec, Montréal. Since 1999, he has been with McGill University, Montreal, where he is currently a Full Professor with the Department of Electrical and Computer Engineering. From 2004 to 2007, he also served as an Associate Chair of Graduate Studies with the Department of Electrical and Computer Engineering, McGill University. He is the author or coauthor of more than 200 referred publications. His research has been funded by the Natural Sciences and Engineering Research Council of Canada, the "Fonds de Recherche sur la Nature et les Technologies" from the Government of Quebec, and some major industrial sponsors, including Nortel Networks, Bell Canada, InterDigital, and Microsemi. His research interests include the study of advanced algorithms for the processing of information bearing signals by digital means and several topics in statistical signal processing, such as detection and estimation, sensor array processing, adaptive filtering, and applications thereof to broadband communications and audio processing.

Dr. Champagne has also served on the Technical Committees of several international conferences in the fields of communications and signal processing. In particular, he served as the Cochair of the Wide Area Cellular Communications Track for the IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (Toronto, September 2011), as the Cochair of the Antenna and Propagation Track for the IEEE Vehicular Technology Conference-Fall, (Los Angeles, CA, USA, September 2004), and as the Registration Chair for the IEEE International Conference on Acoustics, Speech and Signal Processing (Montreal, May 2004). He has served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the *EURASIP Journal on Applied Signal Processing*.

Yulong Zou (M'12–SM'13) received the B.Eng. degree in information engineering and the Ph.D. degree signal and information processing from Nanjing University of Posts and Telecommunications (NUPT), Nanjing, China, in July 2006 and July 2012, respectively, and the Ph.D. degree in electrical engineering from Stevens Institute of Technology, Hoboken, NJ, USA, in May 2012.

He is currently a Full Professor and a Doctoral Supervisor with NUPT. His research interests include wireless communications and signal processing, cooperative communications, cognitive radio, wireless security, and energy-efficient communications.

Dr. Zou has acted as a Symposium Chair, a Session Chair, and a Technical Program Committee member for a number of IEEE sponsored conferences, including the IEEE Wireless Communications and Networking Conference, IEEE Global Communications Conference, IEEE International Conference on Communications, IEEE Vehicular Technology Conference, and the International Conference on Communications in China. He served as the Lead Guest Editor for a special issue on security challenges and issues in cognitive radio networks of the *EURASIP Journal on Advances in Signal Processing*. He currently serves as an Editor for the IEEE COMMUNICATIONS SURVEYS & TUTORIALS, IEEE COMMUNICATIONS LETTERS, the *EURASIP Journal on Advances in Signal Processing*, and the *KSII Transactions on Internet and Information Systems*. He also served as the Lead Guest Editor for a special issue on security and reliability challenges in industrial wireless sensor networks of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.

Lajos Hanzo (M'91–SM'92–F'04) received the M.S. degree in electronics and the Ph.D. degree from Budapest University of Technology and Economics (formerly, Technical University of Budapest), Budapest, Hungary, in 1976 and 1983, respectively; the D.Sc. degree from the University of Southampton, Southampton, U.K., in 2004; and the Doctor Honoris Causa degree from Budapest University of Technology and Economics in 2009.

During his 38-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, where he holds the Chair in Telecommunications. He is currently directing a 100-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council of U.K., the European Research Council's Advanced Fellow Grant, and the Royal Society Wolfson Research Merit Award. During 2008–2012, he was a Chaired Professor with Tsinghua University, Beijing, China. He is an enthusiastic supporter of industrial and academic liaison and offers a range of industrial courses. He has successfully supervised more than 80 Ph.D. students, coauthored 20 John Wiley/IEEE Press books on mobile radio communications, totaling in excess of 10 000 pages, and has published more than 1400 research entries on IEEE Xplore. He has more than 20 000 citations. His research is funded by the European Research Council's Senior Research Fellow Grant.

Dr. Hanzo is a Fellow of the Royal Academy of Engineering, The Institution of Engineering and Technology, and the European Association for Signal Processing. He is also a Governor of the IEEE Vehicular Technology Society. He has served as the Technical Program Committee Chair and the General Chair of IEEE conferences, has presented keynote lectures, and has received a number of distinctions. During 2008–2012, he was the Editor-in-Chief of the IEEE Press.

AUTHOR QUERY

NO QUERY.

Joint Optimization of Transceiver Matrices for MIMO-Aided Multiuser AF Relay Networks: Improving the QoS in the Presence of CSI Errors

Jiaxin Yang, *Student Member, IEEE*, Benoit Champagne, *Senior Member, IEEE*, Yulong Zou, *Senior Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***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 realistic imperfect channel state information (CSI) of all the source–relay and relay–destination links, we propose a robust optimization framework for the joint design of the source transmit precoders (TPCs), relay AF matrices and receive filters. Specifically, two well–known CSI error models are considered, namely, the *statistical* and the *norm-bounded* error models. We commence by considering the problem of minimizing the maximum per-stream mean square error (MSE) subject to the source and relay power constraints (min–max problem). Then, the statistically robust and worst-case robust versions of this problem, which take into account the statistical and norm-bounded CSI errors, respectively, are formulated. Both of the resultant optimization problems are nonconvex (semi-infinite in the worst-case robust design). Therefore, algorithmic solutions having proven convergence and tractable complexity are proposed by resorting to the iterative block coordinate update approach along with matrix transformation and convex conic optimization techniques. We then consider the problem of minimizing the maximum per-relay power subject to the quality-of-service (QoS) constraints for each stream and the source power constraints (QoS problem). Specifically, an efficient initial feasibility search algorithm is proposed based on the relationship between the feasibility check and the min–max problems. Our simulation results show that the proposed joint transceiver design is capable of achieving improved robustness against different types of CSI errors when compared with non-robust approaches.**

***Index Terms*—Amplify-and-forward (AF) relaying, channel state information (CSI) error, convex optimization, multiple-input multiple-output (MIMO), multiuser, robust transceiver design.**

Manuscript received May 29, 2014; revised December 24, 2014; accepted February 16, 2015. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and in part by InterDigital Canada. The review of this paper was coordinated by Prof. C.-X. Wang.

J. Yang and B. Champagne are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: jiaxin.yang@mail.mcgill.ca; benoit.champagne@mcgill.ca).

Y. Zou is with the School of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: yulong.zou@njupt.edu.cn).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2015.2410759

I. Introduction

COOPERATIVE relaying [1] is capable of improving the communication link between the source and destination nodes, in the context of wireless standards such as those of the Long-Term Evolution Advanced [2], Worldwide Interoperability for Microwave Access (WiMAX) [3], and fifth-generation networks [4]. Relaying strategies may be classified as amplify-and-forward (AF) and decode-and-forward (DF) techniques. The AF relaying technique imposes lower signal processing complexity and latency; therefore, it is preferred in many operational applications [5] and is the focus of our attention in this paper.

Recently, multiple-input–multiple-output (MIMO) AF relaying designed for multiuser networks has attracted considerable interest [6]–[11]. In typical wireless multiuser networks, the amount of spectral resources available to each user decreases with an increase in the density of users sharing the channel, hence imposing a degradation on the quality of service (QoS) of each user. MIMO AF relaying is emerging as a promising technique of mitigating this fundamental limitation. By exploiting the so-called *distributed spatial multiplexing* [5] at the multiantenna assisted relays, it allows multiple source/destination pairs to communicate concurrently at an acceptable QoS over the same physical channel [5]. The relay matrix optimization has been extensively studied in a single-antenna assisted multiuser framework, under different design criteria (see, e.g., [6]–[10]), where each source/destination is equipped with a single antenna. In general, finding the optimal relay matrix in these design approaches is deemed challenging because the resultant optimization problems are typically nonconvex. Hence, existing algorithms have relied on convex approximation techniques, e.g., semi-definite relaxation (SDR) [9], [10] and second-order cone programming (SOCP) approximation [7], [8], in order to obtain approximate solutions to the original design problems.

Again, the given contributions focus on single-antenna multiuser networks. However, wireless standards aim for the promotion of mobile broadband multimedia services with an enhanced data rate and QoS, where parallel streams corresponding to different service types can be transmitted simultaneously by each source using multiple antennas [11]. This aspiration has led to a strong interest in the study of cooperative relaying in a MIMO multiuser framework, where multiple antennas are employed by all the sources ($\mathtt{S}$), relays ($\mathtt{R}$), and destinations (D). The joint transceiver design[1] is more challenging than the relay matrix design of the single-antenna scenario, but it provides further performance benefits. Prior contributions [6]–[10], [12], [13] are therefore not readily extendable to this more general case. At the time of this writing, the literature of the joint transceiver design for MIMO multiuser relaying networks is still limited. To be specific, in [14], global objective functions such as the sum power of the interference received at all the destinations and the sum mean square error (MSE) of all the estimated data streams are minimized by adopting the *alternating minimization* approach of [15], where only a single design variable is updated at each iteration based on the SDR technique of [16]. However, the use of global objective functions is not readily applicable to multimedia applications supporting several types of services, each characterized by a specific QoS requirement. To overcome this problem, in [17], the objective of minimizing the total source and relay power subject to a minimum signal-to-noise-plus-interference ratio (SINR) requirement for each S−D link is considered. To this end, a two-level iterative algorithm is proposed, which also involves SDR. Since the main goal of [17] was that of achieving a high *spatial diversity* gain to improve the attainable transmission integrity, the number of data streams transmitted by each source in this setting is limited to one [17].

The efficacy of the joint transceiver design in [14] and [17] relies on the idealized simplifying assumption of perfect channel state information (CSI) for all the S−R and R−D links. In practice, acquiring perfect or even accurate channel estimates at a central processing node is quite challenging. This is primarily due to the combined effects of various sources of imperfections, such as the affordable channel estimation complexities and the limited quantized feedback and feedback delays [18], [19]. The performance of the previous methods may hence be substantially degraded in the presence of realistic CSI errors. In view of this, robust transceiver designs, which explicitly take into account the effects of CSI errors, are highly desirable. Depending on the assumptions concerning the CSI errors, robust designs fall into two major categories, namely, *statistically* robust [18] and *worst-case* robust designs [19]. The former class models the CSI errors as random variables with certain statistical distributions (e.g., Gaussian distributions), and robustness is achieved by optimizing the average performance over all the CSI error realizations; the latter family assumes that the CSI errors belong to some predefined bounded uncertainty regions, such as norm-bounded regions, and optimizes the worst-case performance for all the possible CSI errors within the region.

As a further contribution, we study the joint transceiver design in a more general MIMO multiuser relay network, where multiple S−D pairs communicate with the assistance of multiple relays, and each source transmits multiple parallel data streams to its corresponding destination. Assuming realistic imperfect CSI for all the S−R and R−D links, we propose a new robust optimization framework for minimizing the maximum per-stream MSE subject to the source and relay power constraints, which is termed as the *min–max* problem. In the proposed framework, we aim for solving both the *statistically* robust and *worst-case* robust versions of the min–max problem, which take into account either the statistical CSI errors or the norm-bounded CSI errors, respectively, while maintaining tractable computational complexity. Furthermore, to strictly satisfy the QoS specifications of all the data streams, we subsequently consider the problem of minimizing the maximum per-relay power, subject to the QoS constraints of all the data streams and to the source power constraints, which is referred to as the *QoS* problem. Against this background, the main contributions of this paper are threefold.

- With the statistically robust min–max problem for the *joint* transceiver design being nonconvex, an algorithmic solution having proven convergence is proposed by invoking the iterative *block coordinate update approach* of [20] while relying on both matrix transformation and convex conic optimization techniques. The proposed iterative algorithm successively solves in a circular manner three subproblems corresponding to the source transmit precoders (TPCs), relay AF matrices, and receive filters, respectively. We show that the receive filter subproblem yields a closed-form solution, whereas the other two subproblems can be transformed to convex quadratically constrained linear programs (QCLPs). Then, each QCLP can subsequently be reformulated as a efficiently solvable SOCP.
- The worst-case robust min–max problem is both nonconvex and *semi-infinite*. To overcome these challenges, we first present a generalized version of the so-called $\mathcal{S}$ lemma given in [21], based on which each subproblem can be exactly reformulated as a semi-definite program (SDP) with only linear matrix inequality (LMI) constraints. This results in an iterative algorithmic solution involving several SDPs.
- The QoS-based transceiver optimization is more challenging than that of the min–max problem because it is difficult to find a feasible initialization. Hence, our major contribution here is to propose an efficient procedure for finding a feasible starting point for the iterative QoS-based optimization algorithm, provided that there exits one; otherwise, the procedure also returns a certificate of infeasibility.

The remainder of this paper is organized as follows. Section II introduces our system model and the modeling of CSI errors. The robust joint transceiver design problems are also formulated here. In Sections III and IV, iterative algorithms are proposed for solving the min–max problem both under the statistical and the norm-bounded CSI error models, respectively. The QoS problem is dealt with in Section V. Our numerical results are reported in Section VI. This paper is then concluded in Section VII.

Notations: Boldface uppercase (lowercase) letters represent matrices (vectors), and normal letters denote scalars. $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^{-1}$ denote the conjugate, transpose, Hermitian transpose, and inverse, respectively. $\|\cdot\|$ corresponds to the Euclidean norm of a vector, whereas $\|\cdot\|_F$ and $\|\cdot\|_S$ denote the

[1]We use "transceiver design" to collectively denote the design of the source TPCs, relay AF matrices, and receive filters.

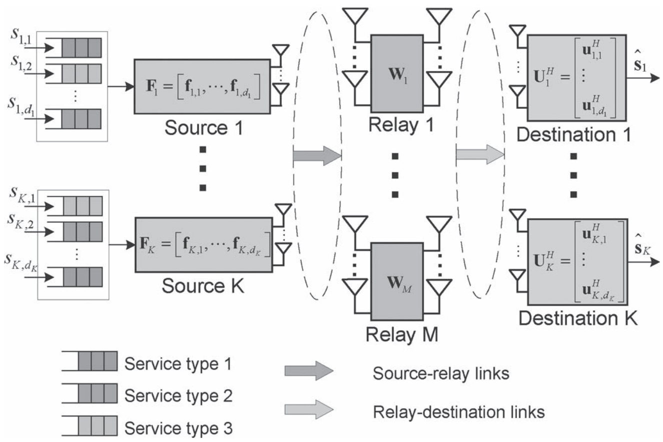

Fig. 1. MIMO multiuser multirelay one-way network with each source transmitting multiple data streams to its corresponding destination.

Frobenius norm and the spectral norm of a matrix, respectively. Furthermore, $\mathrm{Tr}(\cdot)$, $\mathrm{vec}(\cdot)$, and $\otimes$ denote the matrix trace, the vectorization, and the Kronecker product, respectively. $\mathbb{R}^{M\times N}$ and $\mathbb{C}^{M\times N}$ denote the spaces of $M\times N$ matrices with real and complex entries, respectively. $\mathbf{I}_N$ represents the $N\times N$ identity matrix. $\mathbb{E}\{\cdot\}$ denotes the statistical expectation. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a scalar, respectively.

II. System Model and Problem Formulation

We consider a MIMO multiuser relaying network, where M AF relay nodes assist the one-way communication between K S$-$D pairs, as shown in Fig. 1, where all the nodes are equipped with multiple antennas. Specifically, the kth S and D, respectively, employ $N_{\mathrm{S},k}$ and $N_{\mathrm{D},k}$ antennas for $k\in\mathcal{K}\triangleq\{1,2,\ldots,K\}$, whereas the mth R employs $N_{\mathrm{R},m}$ antennas for $m\in\mathcal{M}\triangleq\{1,\ldots,M\}$. All the relays operate under the half-duplex AF protocol, where the 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 concurrently, whereas in the second stage, the relays apply linear processing to the received signals and forward the resultant signals to all the destinations. We assume that no direct links are available between the sources and destinations due to the severe attenuation.

A narrow-band flat-fading radio propagation model is considered, where we denote the channel matrix between the kth S and the mth R by $\mathbf{H}_{m,k}\in\mathbb{C}^{N_{\mathrm{R},m}\times N_{\mathrm{S},k}}$, and the channel matrix between the mth R and the kth D by $\mathbf{G}_{k,m}\in\mathbb{C}^{N_{\mathrm{D},k}\times N_{\mathrm{R},m}}$. Let $\mathbf{s}_k\triangleq[s_{k,1},\ldots,s_{k,d_k}]^T$ denote the information symbols to be transmitted by the kth S at a given time instant, where $d_k\le\min\{N_{\mathrm{S},k},N_{\mathrm{D},k}\}$ is the number of independent data streams. The symbols are modeled as independent random variables with a zero mean and unit variance; hence, $\mathbb{E}\{\mathbf{s}_k\mathbf{s}_k^H\}=\mathbf{I}_{d_k}$. The kth S applies a linear vector of $\mathbf{f}_{k,l}\in\mathbb{C}^{N_{\mathrm{S},k}\times 1}$ for mapping the lth data stream to its $N_{\mathrm{S},k}$ antennas for $l\in\mathcal{D}_k\triangleq\{1,\ldots,d_k\}$, thus forming a linear TPC of $\mathbf{F}_k=[\mathbf{f}_{k,1},\ldots,\mathbf{f}_{k,d_k}]\in\mathbb{C}^{N_{\mathrm{S},k}\times d_k}$. The transmit power is thus given by $\mathrm{Tr}(\mathbf{F}_k\mathbf{F}_k^H)\le P_{\mathrm{S},k}^{\max}$, where $P_{\mathrm{S},k}^{\max}$ is the maximum affordable power of the kth S. Let $\mathbf{n}_{\mathrm{R},m}\in\mathbb{C}^{N_{\mathrm{R},m}\times 1}$ be the spatially white additive noise vector at the mth R, with a zero mean and covariance matrix of $\mathbb{E}\{\mathbf{n}_{\mathrm{R},m}\mathbf{n}_{\mathrm{R},m}^H\}=\sigma_{\mathrm{R},m}^2\mathbf{I}_{N_{\mathrm{R},m}}$.

After the first stage of transmission, the signal received at the mth R 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 R applies a linear matrix $\mathbf{W}_m\in\mathbb{C}^{N_{\mathrm{R},m}\times N_{\mathrm{R},m}}$ to $\mathbf{z}_{\mathrm{R},m}$ and forwards the resultant 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 all the destinations at a power of

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

Let $\mathbf{n}_{\mathrm{D},k}$ denote the spatially white additive noise vector at the kth D with a zero mean and covariance matrix of $\mathbb{E}\{\mathbf{n}_{\mathrm{D},k}\mathbf{n}_{\mathrm{D},k}^H\}=\sigma_{\mathrm{D},k}^2\mathbf{I}_{N_{\mathrm{D},k}}$. The kth D 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 for indexing the sources. To estimate the lth data stream received from its corresponding source, the kth D applies a linear vector $\mathbf{u}_{k,l}$ to the received signal, thus forming a receive filter $\mathbf{U}_k=[\mathbf{u}_{k,1},\ldots,\mathbf{u}_{k,d_k}]\in\mathbb{C}^{N_{\mathrm{D},k}\times d_k}$. Specifically, the estimated information symbols are given by $\hat{s}_{k,l}=\mathbf{u}_{k,l}^H\mathbf{y}_k$, which can be expressed as

$$\begin{aligned}\hat{s}_{k,l}=&\underbrace{\mathbf{u}_{k,l}^H\sum_{m=1}^{M}\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,k}\mathbf{f}_{k,l}s_{k,l}}_{\text{desired data stream}}\\&+\underbrace{\mathbf{u}_{k,l}^H\sum_{m=1}^{M}\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,k}\sum_{p=1,p\neq l}^{d_k}\mathbf{f}_{k,p}s_{k,p}}_{\text{interstream interference}}\\&+\underbrace{\sum_{q=1,q\neq k}^{K}\mathbf{u}_{k,l}^H\sum_{m=1}^{M}\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,q}\mathbf{F}_q\mathbf{s}_q}_{\text{interuser interference}}\\&+\underbrace{\sum_{m=1}^{M}\mathbf{u}_{k,l}^H\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{n}_{\mathrm{R},m}}_{\text{enhanced noise from relays}}+\underbrace{\mathbf{u}_{k,l}^H\mathbf{n}_{\mathrm{D},k}}_{\text{receiver noise}}.\end{aligned} \tag{5}$$

Throughout this paper, we also make the following common assumptions concerning the statistical properties of the signals.

A1) The information symbols transmitted from different S are uncorrelated, i.e., we have $\mathbb{E}\{\mathbf{s}_k\mathbf{s}_m^H\}=\mathbf{0}\ \forall k,m\in\mathcal{K}$ and $k\neq m$.

A2) The information symbols $\mathbf{s}_k$, the relay noise $\mathbf{n}_{\mathrm{R},m}$, and the receiver noise $\mathbf{n}_{\mathrm{D},l}$ are mutually statistically independent $\forall k, l \in \mathcal{K}$ and $m \in \mathcal{M}$.

A. QoS Metric

We adopt the MSE as the QoS metric for each estimated data stream. The major advantage of using the MSE is to make our design problem tractable, which has been well justified in the AF relay matrix design literature [22], [23] and in the references therein. In fact, the links between the MSE and other classic criteria such as the bit error rate (BER) and the SINR have been well established in [22], [24]. Specifically, it has been shown that an improvement in MSE will naturally lead to a reduced BER.

The MSE of the lth estimated data stream received at the kth $\mathtt{D}$ is defined as

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

Substituting (5) into (6), and using assumptions A1 and A2, we obtain

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

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

B. CSI Error Model

In typical relaying scenarios, the CSI of both the $\mathtt{S}-\mathtt{R}$ and $\mathtt{R}-\mathtt{D}$ links, which is available at the central processing node, is contaminated by channel estimation errors and by the quantized feedback, and is outdated due to feedback delays. To model these CSI errors, let us characterize the true but unknown channels as

$$\mathbf{H}_{m,k} = \hat{\mathbf{H}}_{m,k} + \Delta\mathbf{H}_{m,k}, \mathbf{G}_{k,m} = \hat{\mathbf{G}}_{k,m} + \Delta\mathbf{G}_{k,m} \tag{8}$$

where $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$, respectively, denote the estimated $\mathtt{S}-\mathtt{R}$ and $\mathtt{R}-\mathtt{D}$ channels, whereas $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ capture the corresponding *channel uncertainties* [8], [9]. In what follows, we consider two popular techniques of modeling the channel uncertainties.

1) Statistical Error Model: In this model, we assume that the elements of $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ are zero-mean complex Gaussian random variables. Specifically, based on the Kronecker model [18], [25], they can, in general, be written as

$$\Delta\mathbf{H}_{m,k} = \mathbf{\Sigma}_{\mathrm{H}_{m,k}}^{1/2}\Delta\mathbf{H}_{m,k}^{\mathrm{W}}\mathbf{\Psi}_{\mathrm{H}_{m,k}}^{1/2} \tag{9}$$

$$\Delta\mathbf{G}_{k,m} = \mathbf{\Sigma}_{\mathrm{G}_{k,m}}^{1/2}\Delta\mathbf{G}_{k,m}^{\mathrm{W}}\mathbf{\Psi}_{\mathrm{G}_{k,m}}^{1/2} \tag{10}$$

TABLE I
EQUIVALENT NOTATIONS USED IN THE SUBSEQUENT ANALYSIS

Notations	Definitions
$\boldsymbol{\mathcal{G}}_{k,m}$	$\hat{\mathbf{G}}_{k,m}\mathbf{W}_m$
$\boldsymbol{\mathcal{W}}_{m,k}$	$\mathbf{W}_m\hat{\mathbf{H}}_{m,k}$
$\boldsymbol{\mathcal{U}}_{k,m}$	$\mathbf{U}_k^H\hat{\mathbf{G}}_{k,m}$
$\boldsymbol{\mathcal{H}}_{m,k}$	$\hat{\mathbf{H}}_{m,k}\mathbf{F}_k$
$\boldsymbol{\mathcal{T}}_{k,q}$	$\sum_{m=1}^{M}\hat{\mathbf{G}}_{k,m}\mathbf{W}_m\hat{\mathbf{H}}_{m,q}\mathbf{F}_q$

where $\mathbf{\Sigma}_{\mathrm{H}_{m,k}}$ and $\mathbf{\Sigma}_{\mathrm{G}_{k,m}}$ are the row correlation matrices, whereas $\mathbf{\Psi}_{\mathrm{H}_{m,k}}$ and $\mathbf{\Psi}_{\mathrm{G}_{k,m}}$ 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 a zero mean and unit variance.[2] This model is suitable when the CSI errors are dominated by the channel estimation errors.

2) Norm-Bounded Error Model: When the CSI is subject to quantization errors due to the limited-rate feedback, it can no longer be accurately characterized by the given statistical model. Instead, $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ are considered to assume values from the following norm-bounded sets [19]:

$$\mathcal{H}_{m,k} \triangleq \{\Delta\mathbf{H}_{m,k} : \|\Delta\mathbf{H}_{m,k}\|_F \leq \eta_{m,k}\} \tag{11}$$

$$\mathcal{G}_{k,m} \triangleq \{\Delta\mathbf{G}_{k,m} : \|\Delta\mathbf{G}_{k,m}\|_F \leq \xi_{k,m}\} \tag{12}$$

where $\eta_{m,k} > 0$ and $\xi_{k,m} > 0$ specify the radii of the uncertainty regions, thus reflecting the degree of uncertainties. The benefits of such an error model have been well justified in the literature of robust relay optimization (see, e.g., [8], [9], and [26]). The determination of the radii of the uncertainty regions has also been discussed in [19].

Throughout this paper, we assume that the magnitudes of the CSI errors are significantly lower than those of the channel estimates; therefore, the third- and higher-order terms in $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ are neglected in our subsequent analysis. We also introduce in Table I some useful notations to simplify our exposition.

Substituting (8) into (7) and applying the aforementioned assumptions, the per-stream MSE in the presence of CSI errors can be expressed as

$$\begin{aligned}&\varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right) \\ &\approx \left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k} + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,k}\mathbf{F}_k \right.\\ &\quad\left. + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,k}\mathbf{F}_k - \mathbf{e}_{k,l}^T\right\|^2 + \sigma_{\mathrm{D},k}^2\left\|\mathbf{u}_{k,l}\right\|^2 \\ &\quad + \sum_{q=1,q\neq k}^{K}\left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,q} + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q \right.\\ &\quad\quad\left. + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,q}\mathbf{F}_q\right\|^2 \\ &\quad + \sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m} + \mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\mathbf{W}_m\right\|^2. \end{aligned} \tag{13}$$

[2]The superscript "W" simply refers to the spatially white or uncorrelated nature of these random variables.

We now observe that the per-stream MSE becomes uncertain in $\Delta\mathbf{H}_{m,k}\ \forall(m,k)\in\mathcal{M}\times\mathcal{K}$ and $\Delta\mathbf{G}_{k,m}\ \forall m\in\mathcal{M}$. Therefore, we introduce the following compact notations for convenience:

$$\Delta\mathbf{G}_k \triangleq (\Delta\mathbf{G}_{k,1},\ldots,\Delta\mathbf{G}_{k,M}) \in \mathcal{G}_k \triangleq \mathcal{G}_{k,1}\times\cdots\times\mathcal{G}_{k,M}$$
$$\Delta\mathbf{H} \triangleq (\Delta\mathbf{H}_{1,1},\ldots,\Delta\mathbf{H}_{M,K}) \in \mathcal{H} \triangleq \mathcal{H}_{1,1}\times\cdots\times\mathcal{H}_{M,K}.$$

For subsequent derivations, the dependence of $\varepsilon_{k,l}$ on $\Delta\mathbf{H}$ and $\Delta\mathbf{G}_k$ is made explicit in (13).

The kth relay's transmit power in the presence of CSI errors can also be explicitly expressed as $P_{\mathrm{R},m}(\Delta\mathbf{H}_m)$, where $\Delta\mathbf{H}_m \triangleq (\Delta\mathbf{H}_{m,1},\ldots,\Delta\mathbf{H}_{m,K}) \in \mathcal{H}_m \triangleq \mathcal{H}_{m,1}\times\cdots\times\mathcal{H}_{m,K}$.

C. Problem Formulation

In contrast to the prior advances [6]–[8], [14], [22] found in the relay optimization literature, where certain global objective functions are minimized subject to power constraints at the sources and relays, we formulate the following robust design problems under the explicit consideration of QoS. Let us commence by introducing the following unified operation:

$$\mathcal{U}\{f(\Delta\mathbf{X})\} = \begin{cases} \mathbb{E}_{\Delta\mathbf{X}} f(\Delta\mathbf{X}), & \Delta\mathbf{X} \text{ is random} \\ \max\limits_{\Delta\mathbf{X}\in\mathcal{X}} f(\Delta\mathbf{X}), & \Delta\mathbf{X} \text{ is deterministic} \end{cases} \tag{14}$$

where $\Delta\mathbf{X}\in\mathbb{C}^{M\times N}$ and $f(\cdot):\mathbb{C}^{M\times N}\to\mathbb{R}$. Depending on the specific assumptions concerning $\Delta\mathbf{X}, \mathcal{U}\{\cdot\}$ either computes the expectation of $f(\Delta\mathbf{X})$ over the ensemble of realizations $\Delta\mathbf{X}$ or maximizes $f(\Delta\mathbf{X})$ for all $\Delta\mathbf{X}$ within some bounded set $\mathcal{X}$. This notation will be useful and convenient for characterizing the per-stream MSE of (13) and the relay's power $P_{\mathrm{R},m}(\Delta\mathbf{H}_m)$ for different types of CSI errors in a unified form in our subsequent analysis.

1) Min–Max Problem: For notational convenience, we define $\mathbf{F}\triangleq(\mathbf{F}_1,\ldots,\mathbf{F}_K)$, $\mathbf{W}\triangleq(\mathbf{W}_1,\ldots,\mathbf{W}_M)$, and $\mathbf{U}\triangleq(\mathbf{U}_1,\ldots,\mathbf{U}_K)$, which collects the corresponding design variables. In this problem, we jointly design $\{\mathbf{F},\mathbf{W},\mathbf{U}\}$ with the goal of minimizing the maximum per-stream MSE subject to the source and relay power constraints. This problem pertains to the design of energy-efficient relay networks, where there is a strict constraint on the affordable power consumption. Based on the notation in (14), it can be expressed in the following unified form, which is denoted $\mathcal{M}(P_{\mathrm{R}})$:

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \max_{\forall k\in\mathcal{K},l\in\mathcal{D}_k} \kappa_{k,l}\mathcal{U}\{\varepsilon_{k,l}(\Delta\mathbf{H},\Delta\mathbf{G}_k)\} \tag{15a}$$
$$\text{s.t.}\quad \mathcal{U}\{P_{\mathrm{R},m}(\Delta\mathbf{H}_m)\}\le\rho_m P_{\mathrm{R}}\quad \forall m\in\mathcal{M} \tag{15b}$$
$$\mathrm{Tr}(\mathbf{F}_k^H\mathbf{F}_k)\le P_{\mathrm{S},k}^{\max}\quad \forall k\in\mathcal{K} \tag{15c}$$

where $\{\kappa_{k,l}>0:\forall k\in\mathcal{K}, l\in\mathcal{D}_k\}$ is a set of weights assigned to the different data streams for maintaining fairness among them, P_{R} is the common maximum affordable transmit power of all the relays, and $\{\rho_m>0:\forall m\in\mathcal{M}\}$ is a set of coefficients specifying the individual power of each relay.

2) QoS Problem: The second strategy, which serves as a complement to the given min–max problem, aims for minimizing the maximum per-relay power, while strictly satisfying the QoS constraints for all the data streams and all the source power constraints.[3] Specifically, this problem, which is denoted $\mathcal{Q}(\gamma)$, can be formulated as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \max_{m\in\mathcal{M}} \frac{1}{\rho_m}\mathcal{U}\{P_{\mathrm{R},m}(\Delta\mathbf{H}_m)\} \tag{16a}$$
$$\text{s.t.}\quad \mathcal{U}\{\varepsilon_{k,l}(\Delta\mathbf{H},\Delta\mathbf{G}_k)\}\le\frac{\gamma}{\kappa_{k,l}}\quad \forall k\in\mathcal{K}, l\in\mathcal{D}_k \tag{16b}$$
$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right)\le P_{\mathrm{S},k}^{\max}\quad \forall k\in\mathcal{K} \tag{16c}$$

where γ denotes a common QoS target for all the data streams.

The following remark is of interest.

Remark 1: The major difference between the min–max and QoS problems is that solving the QoS problem is not always feasible. This is because the per-stream MSE imposed by the interstream and interuser interference [cf. (13)] cannot be made arbitrarily small by simply increasing the transmit power. By contrast, solving the min–max problem is always feasible since it relies on its "*best effort*" to improve the QoS for all the data streams at limited power consumption. Both problem formulations are nonconvex and in general NP-hard. These issues motivate the pursuit of a tractable but suboptimal solution to the design problems considered.

III. Statistically Robust Transceiver Design for the Min–Max Problem

Here, we propose an algorithmic solution to the min–max problem of (15) in the presence of the statistical CSI errors of Section II-B1. The corresponding statistically robust version of (15) can be formulated as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \max_{\forall k\in\mathcal{K},l\in\mathcal{D}_k} \kappa_{k,l}\bar{\varepsilon}_{k,l} \tag{17a}$$
$$\text{s.t.}\quad \overline{P}_{\mathrm{R},m}\le\rho_m P_{\mathrm{R}}\quad \forall m\in\mathcal{M} \tag{17b}$$
$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right)\le P_{\mathrm{S},k}^{\max}\quad \forall k\in\mathcal{K} \tag{17c}$$

where we have

$$\bar{\varepsilon}_{k,l}\triangleq\mathbb{E}_{\Delta\mathbf{H},\Delta\mathbf{G}_k}\{\varepsilon_{k,l}(\Delta\mathbf{H},\Delta\mathbf{G}_k)\}$$
$$\overline{P}_{\mathrm{R},m}\triangleq\mathbb{E}_{\Delta\mathbf{H}_m}\{P_{\mathrm{R},m}(\Delta\mathbf{H}_m)\}. \tag{18}$$

To further exploit the structure of (17), we have to compute the expectations in (18), which we refer to as the averaged MSE and relay power, respectively. By exploiting the independence

[3] In fact, the min–max problem $\mathcal{M}(P_{\mathrm{R}})$ and the QoS problem $\mathcal{Q}(\gamma)$ are the so-called *inverse problems*, i.e., we have $\gamma=\mathcal{M}[\mathcal{Q}(\gamma)]$ and $P_{\mathrm{R}}=\mathcal{Q}[\mathcal{M}(P_{\mathrm{R}})]$. The proof follows a similar argument to that of [27, Th. 3]. However, as shown in the subsequent analysis, the proposed algorithm cannot guarantee finding the global optimum of the design problems. Therefore, monotonic convergence cannot be guaranteed, which is formally stated as $P_{\mathrm{R}}\ge P'_{\mathrm{R}}\nRightarrow\mathcal{M}(P_{\mathrm{R}})\le\mathcal{M}(P'_{\mathrm{R}})$ and $\gamma\ge\gamma'\nRightarrow\mathcal{Q}(\gamma)\le\mathcal{Q}(\gamma')$. Due to the lack of the monotonicity, a 1-D binary search algorithm is unable to solve $\mathcal{Q}(\gamma)$ via a sequence of $\mathcal{M}(P_{\mathrm{R}})$ evaluations. Consequently, a formal inverse problem definition is not stated in this paper.

of $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ in (13), the per-stream MSE averaged over the channel uncertainties can be expanded as

$$\begin{aligned}\bar{\varepsilon}_{k,l} &= \mathbf{u}_{k,l}^H\left(\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^H+\mathbf{R}_k\right)\mathbf{u}_{k,l}-2\Re\left\{\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k}\mathbf{e}_{k,l}\right\}+1\\ &\quad+\sum_{q=1}^{K}\sum_{m=1}^{M}\underbrace{\mathbb{E}\left\{\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\Delta\mathbf{G}_{k,m}^H\mathbf{u}_{k,l}\right\}}_{\mathcal{I}_1}\\ &\quad+\sum_{q=1}^{K}\sum_{m=1}^{M}\underbrace{\mathbb{E}\left\{\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\Delta}\mathbf{H}_{m,q}^H\boldsymbol{\mathcal{G}}_{k,m}^H\mathbf{u}_{k,l}\right\}}_{\mathcal{I}_2}\\ &\quad+\sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\underbrace{\mathbb{E}\left\{\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{W}_m^H\Delta\mathbf{G}_{k,m}^H\mathbf{u}_{k,l}\right\}}_{\mathcal{I}_3} \qquad (19)\end{aligned}$$

where we have

$$\mathbf{R}_k=\sum_{q=1,q\neq k}^{K}\boldsymbol{\mathcal{T}}_{k,q}\boldsymbol{\mathcal{T}}_{k,q}^H+\sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\mathcal{G}}_{k,m}^H+\sigma_{\mathrm{D},k}^2\mathbf{I}_{d_k}. \qquad (20)$$

To compute the expectations in (19), we rely on the results of [28, (10)] to obtain

$$\begin{aligned}\mathcal{I}_1 &= \mathbf{u}_{k,l}^H\mathbb{E}\left\{\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\Delta\mathbf{G}_{k,m}^H\right\}\mathbf{u}_{k,l}\\ &= \mathrm{Tr}\left(\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\mathbf{u}_{k,l}^H\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}. \qquad (21)\end{aligned}$$

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

$$\mathcal{I}_2=\mathrm{Tr}\left(\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\Psi}_{\mathrm{H}_{m,q}}\right)\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\Sigma}_{\mathrm{H}_{m,q}}\boldsymbol{\mathcal{G}}_{k,m}^H\mathbf{u}_{k,l} \qquad (22)$$

$$\mathcal{I}_3=\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\mathbf{u}_{k,l}^H\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}. \qquad (23)$$

Based on (21)–(23), the averaged MSE in (19) is therefore equivalent to

$$\begin{aligned}\bar{\varepsilon}_{k,l} &= \mathbf{u}_{k,l}^H\left(\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^H+\mathbf{R}_k+\boldsymbol{\Omega}_k\right)\mathbf{u}_{k,l}\\ &\quad-2\Re\left\{\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k}\mathbf{e}_{k,l}\right\}+1 \qquad (24)\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{\Omega}_k &= \sum_{q=1}^{K}\sum_{m=1}^{M}\Big(\mathrm{Tr}\left(\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\\ &\qquad+\mathrm{Tr}\left(\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\Psi}_{\mathrm{H}_{m,q}}\right)\boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\Sigma}_{\mathrm{H}_{m,q}}\boldsymbol{\mathcal{G}}_{k,m}^H\Big)\\ &\quad+\sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}. \qquad (25)\end{aligned}$$

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

The averaged relay power $\overline{P}_{\mathrm{R},m}$ can be derived as

$$\begin{aligned}\overline{P}_{\mathrm{R},m} &= \sum_{k=1}^{K}\Big(\mathrm{Tr}\big(\mathbf{F}_k^H\hat{\mathbf{H}}_{m,k}^H\mathbf{W}_m^H\mathbf{W}_m\hat{\mathbf{H}}_{m,k}\mathbf{F}_k\big)\\ &\qquad+\mathrm{Tr}\left(\mathbf{F}_k\mathbf{F}_k^H\boldsymbol{\Psi}_{\mathrm{H}_{m,k}}\right)\mathrm{Tr}\left(\mathbf{W}_m^H\mathbf{W}_m\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}\right)\Big)\\ &\quad+\sigma_{\mathrm{R},m}^2\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\right) \qquad (26)\end{aligned}$$

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

A. *Iterative Joint Transceiver Optimization*

It is worthwhile noting that the inner pointwise maximization in (17a) preserves the partial convexity of $\bar{\varepsilon}_{k,l}$. Substituting (24) and (26) back into (17), the latter is shown to possess a so-called *block multiconvex* structure [20], which implies that the problem is convex in each block of variables, although in general not jointly convex in all the variables.

Motivated by the given 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 associated with the remaining blocks. In this way, three subproblems can be derived from (17), with each updating $\mathbf{F}$, $\mathbf{W}$, and $\mathbf{U}$, respectively. Each subproblem can be transformed into a *convex* one, which is computationally much simpler than directly finding the optimal solution to the original joint problem (if at all possible). Since solving for each block at the current iteration depends on the values of the other blocks gleaned from the previous iteration, this method in effect can be recognized as a joint optimization approach in terms of both the underlying theory [15], [20] and the related applications [14], [17]. We now proceed by analyzing each of these subproblems.

1) Receive Filter Design: It can be observed in (19) that $\bar{\varepsilon}_{k,l}$ in (17a) only depends on the corresponding linear vector $\mathbf{u}_{k,l}$, whereas the constraints (17b) and (17c) do not involve $\mathbf{u}_{k,l}$. Hence, for a fixed $\mathbf{F}$ and $\mathbf{W}$, the optimal $\mathbf{u}_{k,l}$ can be obtained independently and in parallel for different (k,l) values by equating the following complex gradient to zero:

$$\nabla_{\mathbf{u}_{k,l}^*}\bar{\varepsilon}_{k,l}=\mathbf{0}. \qquad (27)$$

The resultant optimal solution of (27) is the Wiener filter, i.e.,

$$\mathbf{u}_{k,l}=\left(\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^H+\mathbf{R}_k+\boldsymbol{\Omega}_k\right)^{-1}\boldsymbol{\mathcal{T}}_{k,k}\mathbf{e}_{k,l}. \qquad (28)$$

2) Source TPC Design: We then solve our problem for the TPC $\mathbf{F}$, while keeping $\mathbf{W}$ and $\mathbf{U}$ fixed. For better exposition of our solution, we can rewrite (17) after some matrix manipulations, explicitly in terms of $\mathbf{F}$ as given in (29), shown at the bottom of the next page, where $\mathbf{E}_{k,l}\triangleq\mathbf{e}_{k,l}\mathbf{e}_{k,l}^T$, $\eta_{\mathrm{R},m}\triangleq\rho_m P_{\mathrm{R}}-\sigma_{\mathrm{R},m}^2\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\right)$, and

$$\begin{aligned}a_3^{k,l} &\triangleq \mathbf{u}_{k,l}^H\Bigg[\sum_{m=1}^{M}\sigma_{\mathrm{R},m}^2\Big(\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\\ &\qquad+\boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\mathcal{G}}_{k,m}^H\Big)+\sigma_{\mathrm{D},k}^2\mathbf{I}_{N_{\mathrm{D},k}}\Bigg]\mathbf{u}_{k,l}+1. \qquad (30)\end{aligned}$$

The solution to the problem (29) is not straightforward; hence, we transform it into a more tractable form. To this end, we introduce the new variables of $\mathbf{f}_k \triangleq \mathrm{vec}\left(\mathbf{F}_k\right) \in \mathbb{C}^{N_{\mathrm{S},k} d_k \times 1}$ $\forall k \in \mathcal{K}$ and define the following quantities that are independent of $\mathbf{f}_k$ $\forall k \in \mathcal{K}$:

$$\begin{aligned}\mathbf{A}_{1,q}^{k,l} \triangleq \sum_{m=1}^{M} \mathbf{I}_{d_k} \otimes \Bigg(& \sum_{n=1}^{M} \boldsymbol{\mathcal{W}}_{m,q}^{H} \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n} \boldsymbol{\mathcal{W}}_{n,q} \\ & + \mathrm{Tr}\big(\mathbf{u}_{k,l}^{H} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l}\big) \boldsymbol{\mathcal{W}}_{m,k}^{H} \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \boldsymbol{\mathcal{W}}_{m,k} \\ & + \mathrm{Tr}\left(\mathbf{u}_{k,l}^{H} \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}} \boldsymbol{\mathcal{G}}_{k,m}^{H} \mathbf{u}_{k,l}\right) \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \Bigg)\end{aligned} \tag{31}$$

$$\mathbf{a}_{2}^{k,l} = \mathrm{vec}\left(\sum_{m=1}^{M} \boldsymbol{\mathcal{W}}_{m,k}^{H} \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \right) \tag{32}$$

$$\mathbf{A}_{4,k}^{m} = \mathbf{I}_{d_k} \otimes \left(\boldsymbol{\mathcal{W}}_{m,k}^{H} \boldsymbol{\mathcal{W}}_{m,k} + \mathrm{Tr}\left(\mathbf{W}_{m}^{H} \mathbf{W}_{m} \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}\right) \boldsymbol{\Psi}_{\mathrm{H}_{m,k}} \right). \tag{33}$$

It may be readily verified that $\mathbf{A}_{1,q}^{k,l}$ and $\mathbf{A}_{4,k}^{m}$ are positive definite matrices. Then, we invoke the following identities, i.e., $\mathrm{Tr}\left(\mathbf{A}^H \mathbf{B} \mathbf{A}\right) = \mathrm{vec}\left(\mathbf{A}\right)^H \left(\mathbf{I} \otimes \mathbf{B}\right) \mathrm{vec}\left(\mathbf{A}\right)$ and $\mathrm{Tr}\left(\mathbf{A}^H \mathbf{B}\right) = \mathrm{vec}\left(\mathbf{B}\right)^H \mathrm{vec}\left(\mathbf{A}\right)$, for transforming both the objective (29a) and the constraints (29b)–(29c) into quadratic expressions of $\mathbf{f}_k$, and finally reach the following equivalent formulation:

$$\min_{\mathbf{f}_1,\ldots,\mathbf{f}_K,t} \quad t \tag{34a}$$

$$\text{s.t.} \quad \sum_{q=1}^{K} \mathbf{f}_q^H \mathbf{A}_{1,q}^{k,l} \mathbf{f}_q - 2\Re\left\{\mathbf{f}_k^H \mathbf{a}_2^{k,l}\right\} + a_3^{k,l} \leq \frac{t}{\kappa_{k,l}} \quad \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{34b}$$

$$\sum_{k=1}^{K} \mathbf{f}_k^H \mathbf{A}_{4,k}^{m} \mathbf{f}_k \leq \eta_{\mathrm{R},m} \quad \forall m \in \mathcal{M} \tag{34c}$$

$$\mathbf{f}_k^H \mathbf{f}_k \leq P_{\mathrm{S},k}^{\max} \quad \forall k \in \mathcal{K} \tag{34d}$$

where t is an auxiliary variable. Problem (34) by definition is a convex separable inhomogeneous QCLP [16]. This class of optimization problems can be handled by the recently developed parser/solvers, such as `CVX` [29] where the built-in parser is capable of verifying the convexity of the optimization problem (in user-specified forms) and then, of automatically transforming it into a standard form; the latter may then be forwarded to external optimization solvers, such as `SeduMi` [30] and `MOSEK` [31]. To gain further insights into this procedure, we show in Appendix A that the problem (34) can be equivalently transformed into a standard SOCP that is directly solvable by a generic external optimization solver based on the interior-point method. Therefore, the SOCP form bypasses the tedious translation by the parser/solvers for every problem instance in real-time computation.

3) Relay AF Matrix Design: To solve for the relay AF matrices, we follow a similar procedure to that used for the source TPC design. However, here we introduce a new variable, which vertically concatenates all the vectorized relay AF matrices, yielding

$$\mathbf{w} \triangleq \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_M \end{bmatrix} \triangleq \begin{bmatrix} \mathrm{vec}\left(\mathbf{W}_1\right) \\ \vdots \\ \mathrm{vec}\left(\mathbf{W}_M\right) \end{bmatrix} \tag{35}$$

along with the following quantities, which are independent of $\mathbf{w}$:

$$\left[\mathbf{B}_1^{k,l}\right]_{m,n} = \sum_{q=1}^{K} \left[\left(\boldsymbol{\mathcal{H}}_{m,q}^{*} \boldsymbol{\mathcal{H}}_{n,q}^{T}\right) \otimes \left(\boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n}\right)\right] \tag{36}$$

$$\mathbf{b}_{2,m}^{k,l} \triangleq \mathrm{vec}\left(\boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{H}}_{m,k}^{H}\right) \tag{37}$$

$$\begin{aligned}\mathbf{B}_{3,m}^{k,l} \triangleq \sum_{q=1}^{K} \Big[& \mathrm{Tr}\big(\mathbf{u}_{k,l}^{H} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l}\big) \boldsymbol{\mathcal{H}}_{m,q}^{*} \boldsymbol{\mathcal{H}}_{m,q}^{T} \otimes \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \\ & + \mathrm{Tr}\big(\mathbf{F}_q^H \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \mathbf{F}_q\big) \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}}^{T} \otimes \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,m} \Big] \\ & + \sigma_{\mathrm{R},m}^2 \mathrm{Tr}\left(\mathbf{u}_{k,l}^{H} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l}\right) \mathbf{I}_{N_{\mathrm{R},m}} \otimes \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \\ & + \sigma_{\mathrm{R},m}^2 \mathbf{I}_{N_{\mathrm{R},m}} \otimes \left(\boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,m}\right)\end{aligned} \tag{38}$$

$$b_4^{k,l} \triangleq \sigma_{\mathrm{D},k}^2 \left\|\mathbf{u}_{k,l}\right\|^2 + 1 \tag{39}$$

$$\begin{aligned}\mathbf{B}_{5,m} \triangleq \Bigg[& \sigma_{\mathrm{R},m}^2 \mathbf{I}_{N_{\mathrm{R},m}} + \sum_{k=1}^{K} \left(\boldsymbol{\mathcal{H}}_{m,k}^{*} \boldsymbol{\mathcal{H}}_{m,k}^{T}\right. \\ & \left. + \mathrm{Tr}\left(\mathbf{F}_k \mathbf{F}_k^H \boldsymbol{\Psi}_{\mathrm{H}_{m,k}}\right) \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}^{T}\right) \Bigg] \otimes \mathbf{I}_{N_{\mathrm{R},m}}\end{aligned} \tag{40}$$

$$\begin{aligned}\min_{\mathbf{F}} \quad \max_{\forall k \in \mathcal{K}, l \in \mathcal{D}_k} \kappa_{k,l} \Bigg\{ & \sum_{q=1}^{K} \sum_{m=1}^{M} \sum_{n=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^H \boldsymbol{\mathcal{W}}_{m,q}^{H} \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n} \boldsymbol{\mathcal{W}}_{n,q} \mathbf{F}_q\right) - \sum_{m=1}^{M} 2\Re\left\{\mathrm{Tr}\left(\mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,m} \boldsymbol{\mathcal{W}}_{m,k} \mathbf{F}_k\right)\right\} + a_3^{k,l} \\ & + \sum_{q=1}^{K} \sum_{m=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^H \boldsymbol{\mathcal{W}}_{m,k}^{H} \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \boldsymbol{\mathcal{W}}_{m,k} \mathbf{F}_q\right) \mathrm{Tr}\left(\mathbf{u}_{k,l}^{H} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l}\right) \\ & + \sum_{q=1}^{K} \sum_{m=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^H \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \mathbf{F}_q\right) \mathrm{Tr}\left(\mathbf{u}_{k,l}^{H} \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}} \boldsymbol{\mathcal{G}}_{k,m}^{H} \mathbf{u}_{k,l}\right) \Bigg\}\end{aligned} \tag{29a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \mathrm{Tr}\left(\mathbf{F}_k^H \left(\hat{\mathbf{H}}_{m,k}^{H} \mathbf{W}_m^H \mathbf{W}_m \hat{\mathbf{H}}_{m,k} + \mathrm{Tr}\left(\mathbf{W}_m^H \mathbf{W}_m \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}\right) \boldsymbol{\Psi}_{\mathrm{H}_{m,k}} \right) \mathbf{F}_k \right) \leq \eta_{\mathrm{R},m}, \quad \forall m \in \mathcal{M} \tag{29b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H \mathbf{F}_k\right) \leq P_{\mathrm{S},k}^{\max}, \quad \forall k \in \mathcal{K} \tag{29c}$$

where $\mathbf{B}_1^{k,l}$ is a block matrix with its (m,n)th block defined earlier. Then, using the identities $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\mathbf{C}\mathbf{D}^H\right) = \mathrm{vec}\left(\mathbf{A}\right)^H\left(\mathbf{D}^T\otimes\mathbf{B}\right)\mathrm{vec}(\mathbf{C})$, $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\mathbf{A}\right)=\mathrm{vec}(\mathbf{A})^H\left(\mathbf{I}\otimes\mathbf{B}\right)\mathrm{vec}\left(\mathbf{A}\right)$, and $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\right) = \mathrm{vec}\left(\mathbf{B}\right)^H\mathrm{vect}\left(\mathbf{A}\right)$, we can formulate the following optimization problem:

$$\min_{\mathbf{w},t} \quad t \tag{41a}$$

$$\text{s.t.} \quad \mathbf{w}^H\mathbf{B}_1^{k,l}\mathbf{w} - \sum_{m=1}^{M} 2\Re\left\{\mathbf{w}_m^H\mathbf{b}_{2,m}^{k,l}\right\} + \sum_{m=1}^{M}\mathbf{w}_m^H\mathbf{B}_{3,m}^{k,l}\mathbf{w}_m + b_4^{k,l} \leq \frac{t}{\kappa_{k,l}} \ \forall l\in\mathcal{D}_k, k\in\mathcal{K} \tag{41b}$$

$$\mathbf{w}_m^H\mathbf{B}_{5,m}\mathbf{w}_m \leq \rho_m P_{\mathrm{R}} \quad \forall m\in\mathcal{M}. \tag{41c}$$

It may be readily shown that $\mathbf{B}_1^{k,l}$, $\mathbf{B}_{3,m}^{k,l}$, and $\mathbf{B}_{5,m}$ are all positive definite matrices and that (41) is also a convex separable inhomogeneous QCLP. Using a similar approach to the one derived in Appendix A, the SOCP formulation of (41) can readily be obtained. The details of the transformation are therefore omitted for brevity.

B. Algorithm and Properties

We assume that there exists a central processing node, which, upon collecting the channel estimates $\{\hat{\mathbf{H}}_{m,k}, \hat{\mathbf{G}}_{k,m}\ \forall m\in\mathcal{M}, k\in\mathcal{K}\}$ and the covariance matrices of the CSI errors $\{\mathbf{\Sigma}_{\mathrm{H}_{m,k}}, \mathbf{\Sigma}_{\mathrm{G}_{k,m}}, \mathbf{\Psi}_{\mathrm{H}_{m,k}}, \mathbf{\Psi}_{\mathrm{G}_{k,m}}\ \forall m\in\mathcal{M}, k\in\mathcal{K}\}$, optimizes all the design variables and sends them back to the corresponding nodes. The iterative procedure listed in Algorithm 1 therefore should be implemented in a centralized manner, where $\{\mathbf{F}^{(i)}, \mathbf{W}^{(i)}, \mathbf{U}^{(i)}\}$ and $t^{(i)}$ represent the set of design variables and the objective value in (17a), respectively, at the ith iteration. A simple termination criterion can be $|t^{(i)} - t^{(i-1)}| < \epsilon$, where $\epsilon > 0$ is a predefined threshold. In the following, we shall analyze both the convergence properties and the complexity of the proposed algorithm.

1) Convergence: Provided that there is a feasible initialization for Algorithm 1, the solution to each subproblem is globally optimal. As a result, the sequence of the objective values in (17a) is monotonically nonincreasing as the iteration index i increases. Since the maximum per-stream MSE is bounded from below (at least) by zero, the sequence of the objective values must converge by invoking the monotonic convergence theorem.

2) Complexity: When the number of antennas at the sources and relays, i.e., $N_{\mathrm{S},k}$ and $N_{\mathrm{R},m}$, have the same order of magnitude, the complexity of Algorithm 1 is dominated by the SOCP of (62), which is detailed in Appendix A, as it involves all the constraints of the original problem (17). To simplify the complexity analysis, we assume that $N_{\mathrm{S},k} = N_{\mathrm{S}}$, and $d_k = d\ \forall k\in\mathcal{K}$. In (62), the total number of design variables is $N_{\mathrm{total}} = N_{\mathrm{S}}^2K + 1 + K^2d + KM$. The size of the second-order cones (SOCs) in the constraints (62b)–(62g) is given by $(N_{\mathrm{S}}^2+1)dK(K-1)$, $(N_{\mathrm{S}}^2+1)dK$, $(K+2)dK$, $(N_{\mathrm{S}}^2+1)KM$, $(K+1)M$, and $(N_{\mathrm{S}}^2+1)K$, respectively. Therefore, the total dimension of all the SOCs in these constraints can be shown to be $D_{\mathrm{SOCP}} = \mathcal{O}(N_{\mathrm{S}}^2dK^2 + N_{\mathrm{S}}^2MK)$. It has been shown in [32] that problem (62) can be solved most efficiently using the primal–dual interior-point method at *worst-case* complexity on the order of $\mathcal{O}(N_{\mathrm{total}}^2D)$ if no special structure in the problem data is exploited. The computational complexity of Algorithm 1 is therefore on the order of $\mathcal{O}(N_{\mathrm{S}}^6)$, $\mathcal{O}(K^6)$, and $\mathcal{O}(M^3)$ in the individual parameters N_{S}, K and M, respectively. In practice, however, we find that the matrices $\mathbf{A}_{1,q}^{k,l}$ and $\mathbf{A}_{4,k}^{m}$ in (31) and (33), respectively, exhibit a significant level of sparsity, which allows solving the SOCP more efficiently. In our simulations, we therefore measured the CPU time required for solving (62) for different values of N_{S}, K, and M (the results are not reported due to the space limitation) and found that the orders of complexity obtained empirically are significantly lower than those of the given worst-case analysis. Empirically, we found these to be around $\mathcal{O}(N_{\mathrm{S}}^{1.6})$, $\mathcal{O}(K^{1.7})$, and $\mathcal{O}(M^{1.3})$.

Algorithm 1 Iterative Algorithm for Statistically Robust Min–Max Problem

---

Initialization:

1: Set the iteration index $i = 0$, $\mathbf{F}_k^{(0)} = \sqrt{P_{\mathrm{S},k}^{\max}}\mathbf{I}_{N_{\mathrm{S},k}\times d_k}$, $\forall k\in\mathcal{K}$ and $\mathbf{W}_m^{(0)} = \sqrt{\frac{\rho_m P_{\mathrm{R}}}{\mathrm{Tr}(\mathbf{B}_{5,m})}}\mathbf{I}_{N_{\mathrm{R},m}}, \forall m\in\mathcal{M}$

2: **repeat**

3: Compute $\mathbf{u}_{k,l}^{(i+1)}\ \forall k\in\mathcal{K}, l\in\mathcal{D}_k$, using the Wiener filter (28) in parallel;

4: Compute $\mathbf{F}_k^{(i+1)}\ \forall k\in\mathcal{K}$ by solving the SOCP (62);

5: Compute $\mathbf{W}_m^{(i+1)}\ \forall m\in\mathcal{M}$ by solving the SOCP (41);

6: $i \leftarrow i+1$;

7: **until** $|t^{(i)} - t^{(i-1)}| < \epsilon$

IV. Worst-Case Robust Transceiver Design for the Min–Max Problem

Here, we consider the joint transceiver design problem under min–max formulation of (15) and the norm-bounded CSI error model of Section II-B2. To this end, based on the notation in (14), we explicitly rewrite this problem as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \max_{\substack{\forall k\in\mathcal{K}, l\in\mathcal{D}_k,\\ \forall\Delta\mathbf{H}\in\mathcal{H}, \Delta\mathbf{G}_k\in\mathcal{G}_k}} \kappa_{k,l}\varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right) \tag{42a}$$

$$\text{s.t.} \quad P_{\mathrm{R},m}\left(\Delta\mathbf{H}_m\right) \leq \rho_m P_{\mathrm{R}}\ \forall m\in\mathcal{M}, \Delta\mathbf{H}_m\in\mathcal{H}_m \tag{42b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right) \leq P_{\mathrm{S},k}^{\max}\ \forall k\in\mathcal{K} \tag{42c}$$

whose epigraph form can be expressed as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad t \tag{43a}$$

$$\text{s.t.} \quad \varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right) \leq \frac{t}{\kappa_{k,l}}\ \forall k\in\mathcal{K}, l\in\mathcal{D}_k, \Delta\mathbf{H}\in\mathcal{H}, \Delta\mathbf{G}_k\in\mathcal{G}_k \tag{43b}$$

$$P_{\mathrm{R},m}\left(\Delta\mathbf{H}_m\right) \leq \rho_m P_{\mathrm{R}}\ \forall m\in\mathcal{M}, \Delta\mathbf{H}_m\in\mathcal{H}_m \tag{43c}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right) \leq P_{\mathrm{S},k}^{\max}\ \forall k\in\mathcal{K} \tag{43d}$$

where t is an auxiliary variable. As compared with the statistically robust version of (17), problem (43) now encounters two major challenges, namely the nonconvexity and the *semi-infinite* nature of the constraints (43b) and (43c), which render the optimization problem mathematically intractable. In what follows, we derive a solution to address these calamities.

A. Iterative Joint Transceiver Optimization

To overcome the first difficulty, we still rely on the iterative block coordinate update approach described in Section III; however, the three resultant subproblems are *semi-infinite* due to the continuous but bounded channel uncertainties in (43b) and (43c). To handle the semi-infiniteness, an equivalent reformulation of these constraints as LMI will be derived by using certain matrix transformation techniques and by exploiting an extended version of the $\mathcal{S}$-lemma of [21]. In turn, such LMI will convert each of the subproblems into an equivalent SDP [33] efficiently solvable by interior-point methods [34].

1) Receive Filter Design: In this subproblem, we have to minimize t in (43a) with respect to $\mathbf{u}_{k,l}$ subject to the constraint (43b). To transform this constraint into an equivalent LMI, the following lemma is presented, which is an extended version of the one in [21].

Lemma 1 (Extension of $\mathcal{S}$-lemma [21]): Let $\mathbf{A}(\mathbf{x}) = \mathbf{A}^H(\mathbf{x})$, $\boldsymbol{\Sigma}(\mathbf{x}) = \boldsymbol{\Sigma}^H(\mathbf{x})$, $\{\mathbf{D}_k(\mathbf{x})\}_{k=1}^N$, and $\{\mathbf{B}_k\}_{k=1}^N$ be matrices with appropriate dimensions, where $\mathbf{A}(\mathbf{x})$, $\boldsymbol{\Sigma}(\mathbf{x})$, and $\{\mathbf{D}_k(\mathbf{x})\}_{k=1}^N$ are affine functions of $\mathbf{x}$. The following *semi-infinite* matrix inequality:

$$\left(\mathbf{A}(\mathbf{x}) + \sum_{k=1}^{N} \mathbf{B}_k^H \mathbf{C}_k \mathbf{D}_k(\mathbf{x})\right) \times \left(\mathbf{A}(\mathbf{x}) + \sum_{k=1}^{N} \mathbf{B}_k^H \mathbf{C}_k \mathbf{D}_k(\mathbf{x})\right)^H \preceq \boldsymbol{\Sigma}(\mathbf{x}) \quad (44)$$

holds for all $\|\mathbf{C}_k\|_S \leq \rho_k, k = 1, \ldots, N$ if and only if there exist nonnegative scalars $\tau_1, \ldots, \tau_N$ satisfying (45), shown at the bottom of the page.

A simplified version of Lemma 1, which considers only a single uncertainty block, i.e., $N = 1$, can be traced back to [35], whereas a further related corollary is derived in [21, Proposition 2]. Lemma 1 extends this result to the case of multiple uncertainty blocks, i.e., $K > 1$; the proof which follows similar steps as in [21] is omitted owing to the space limitation.

Upon using Lemma 1, the constraint (43b) can equivalently be reformulated as follows.

Proposition 1: There exist nonnegative values of $\boldsymbol{\tau}_{k,l}^{\mathrm{G}} \in \mathbb{R}^{M\times 1}$ and $\boldsymbol{\tau}_{k,l}^{\mathrm{H}} \in \mathbb{R}^{KM\times 1}$ capable of ensuring that the semi-infinite constraint (43b) is equivalent to the matrix inequality in (46), shown at the bottom of the page, where we have $N_{\mathrm{R}} \triangleq \sum_{m=1}^{M} N_{\mathrm{R},m}$, $N_{\mathrm{S}} \triangleq \sum_{k=1}^{K} N_{\mathrm{S},k}$, and the operator $(*)$ denotes the Khatri–Rao product (blockwise Kronecker product) [36]. In (46), $\overline{\boldsymbol{\Theta}}_{k,l}$ and $\overline{\boldsymbol{\Phi}}_{k,l}$ are defined as

$$\overline{\boldsymbol{\Theta}}_{k,l} \triangleq \begin{bmatrix} \xi_{k,1}\boldsymbol{\Theta}_1^{k,l} \\ \vdots \\ \xi_{k,M}\boldsymbol{\Theta}_M^{k,l} \end{bmatrix}, \overline{\boldsymbol{\Phi}}_{k,l} \triangleq \begin{bmatrix} \eta_{1,1}\boldsymbol{\Phi}_{1,1}^{k,l} \\ \vdots \\ \eta_{M,K}\boldsymbol{\Phi}_{M,K}^{k,l} \end{bmatrix} \quad (47)$$

whereas $\boldsymbol{\Theta}_{k,l}$, $\boldsymbol{\Phi}_{k,l}$, and $\boldsymbol{\theta}_{k,l}$ are defined in (71) of Appendix B.

Proof: See Appendix B. ∎

Using (46), the subproblem formulated for $\mathbf{u}_{k,l}$ can be equivalently recast as

$$\min_{t,\mathbf{u}_{k,l},\boldsymbol{\tau}_{k,l}^{\mathrm{g}},\boldsymbol{\tau}_{k,l}^{\mathrm{h}}} \ t \quad \text{s.t.} \quad \mathbf{Q}_{k,l} \succeq \mathbf{0}. \quad (48)$$

With fixed $\mathbf{F}$ and $\mathbf{W}$, (46) depends affinely on the design variables $\{t, \mathbf{u}_{k,l}, \boldsymbol{\tau}_{k,l}^{\mathrm{g}}, \boldsymbol{\tau}_{k,l}^{\mathrm{h}}\}$. Therefore, (48) is a convex SDP of the LMI form [33], which is efficiently solvable by existing optimization tools based on the interior-point method. Since the $\mathbf{u}_{k,l}$ for different values of (k, l) are independent of each other, they can be updated in parallel by solving (48) for different k and l.

2) Source TPC Design: We now have to solve problem (43) for $\mathbf{F}$ by fixing $\mathbf{U}$ and $\mathbf{W}$. The solution is formulated in the following proposition.

$$\begin{bmatrix} \boldsymbol{\Sigma}(\mathbf{x}) - \sum_{k=1}^{N} \tau_k \mathbf{B}_k^H \mathbf{B}_k & \mathbf{A}(\mathbf{x}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{A}^H(\mathbf{x}) & \mathbf{I} & \rho_1 \mathbf{D}_1^H(\mathbf{x}) & \cdots & \rho_N \mathbf{D}_N^H(\mathbf{x}) \\ \mathbf{0} & \rho_1 \mathbf{D}_1(\mathbf{x}) & \tau_1 \mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \rho_N \mathbf{D}_N(\mathbf{x}) & \mathbf{0} & \cdots & \tau_N \mathbf{I} \end{bmatrix} \succeq \mathbf{0} \quad (45)$$

$$\mathbf{Q}_{k,l} \triangleq \begin{bmatrix} \frac{t}{\kappa_{k,l}} - \mathbf{1}^T\boldsymbol{\tau}_{k,l}^{\mathrm{G}} - \mathbf{1}^T\boldsymbol{\tau}_{k,l}^{\mathrm{H}} & \boldsymbol{\theta}_{k,l} & \mathbf{0}_{1\times N_{\mathrm{D},k}N_{\mathrm{R}}} & \mathbf{0}_{1\times N_{\mathrm{S}}N_{\mathrm{R}}} \\ \boldsymbol{\theta}_{k,l}^H & \mathbf{I}_{d+N_{\mathrm{R}}+N_{\mathrm{D},k}} & \overline{\boldsymbol{\Theta}}_{k,l}^H & \overline{\boldsymbol{\Phi}}_{k,l}^H \\ \mathbf{0}_{N_{\mathrm{D},k}N_{\mathrm{R}}\times 1} & \overline{\boldsymbol{\Theta}}_{k,l} & \mathrm{diag}\left(\boldsymbol{\tau}_{k,l}^{\mathrm{G}}\right) * \mathbf{I}_{N_{\mathrm{D},k}N_{\mathrm{R}}} & \mathbf{0}_{N_{\mathrm{D},k}N_{\mathrm{R}}\times N_{\mathrm{S}}N_{\mathrm{R}}} \\ \mathbf{0}_{N_{\mathrm{S}}N_{\mathrm{R}}\times 1} & \overline{\boldsymbol{\Phi}}_{k,l} & \mathbf{0}_{N_{\mathrm{S}}N_{\mathrm{R}}\times N_{\mathrm{D},k}N_{\mathrm{R}}} & \mathrm{diag}\left(\boldsymbol{\tau}_{k,l}^{\mathrm{H}}\right) * \mathbf{I}_{N_{\mathrm{S}}N_{\mathrm{R}}} \end{bmatrix} \succeq \mathbf{0} \quad (46)$$

Proposition 2: The subproblem of optimizing the TPCs $\mathbf{F}$ can be formulated as the following SDP:

$$\min_{t,\mathbf{F},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l},\boldsymbol{\tau}^{\mathrm{p}}_{m}} \quad t \tag{49a}$$

$$\text{s.t.} \quad \mathbf{Q}_{k,l} \succeq \mathbf{0}\ \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{49b}$$

$$\mathbf{P}_m \succeq \mathbf{0}\ \forall m \in \mathcal{M} \tag{49c}$$

$$\begin{bmatrix} P^{\max}_{\mathrm{S},k} & \mathbf{f}_k^H \\ \mathbf{f}_k & \mathbf{I}_{N_{\mathrm{S},k} d_k} \end{bmatrix} \succeq \mathbf{0}\ \forall k \in \mathcal{K} \tag{49d}$$

where we have

$$\mathbf{P}_m \triangleq \begin{bmatrix} \rho_m P_{\mathrm{R}} - \mathbf{1}^T \boldsymbol{\tau}^{\mathrm{p}}_m & \mathbf{t}_m^H & \mathbf{0}_{1\times N_{\mathrm{S}} N_{\mathrm{R},m}} \\ \mathbf{t}_m & \mathbf{I} & \overline{\mathbf{T}}_m \\ \mathbf{0}_{N_{\mathrm{S}} N_{\mathrm{R},m}\times 1} & \overline{\mathbf{T}}_m^H & \mathrm{diag}\left(\boldsymbol{\tau}^{\mathrm{p}}_m\right) * \mathbf{I} \end{bmatrix} \succeq \mathbf{0} \tag{50}$$

with $\boldsymbol{\tau}^{\mathrm{p}}_m \in \mathbb{R}^{K\times 1}$, $\overline{\mathbf{T}}_m(\mathbf{F}) \triangleq \left[\mathbf{T}_{m,1}^T, \ldots, \mathbf{T}_{m,K}^T\right]^T$, and

$$\mathbf{t}_m \triangleq \begin{bmatrix} \mathrm{vec}\left(\mathbf{W}_m \hat{\mathbf{H}}_{m,k} \mathbf{F}_1\right) \\ \vdots \\ \mathrm{vec}\left(\mathbf{W}_m \hat{\mathbf{H}}_{m,K} \mathbf{F}_K\right) \\ \sigma_{\mathrm{R},m} \mathrm{vec}\left(\mathbf{W}_m\right) \end{bmatrix} \tag{51}$$

$$\mathbf{T}_{m,k} \triangleq \begin{bmatrix} \mathbf{0}_{\sum_{q=1}^{k-1} d_q N_{\mathrm{R},m} \times N_{\mathrm{S},k} N_{\mathrm{R},m}} \\ \mathbf{F}_k^T \otimes \mathbf{W}_m \\ \mathbf{0}_{\left(\sum_{q=k+1}^{K} d_q N_{\mathrm{R},m} + N_{\mathrm{R},m}^2\right) \times N_{\mathrm{S},k} N_{\mathrm{R},m}} \end{bmatrix}. \tag{52}$$

Proof: Since $\mathbf{F}$ is involved in all the constraints of the original problem (43), in the following, we will transform each of these constraints into tractable forms.

First, note that (43b) has already been reformulated as (46), which is a trilinear function of $\mathbf{F}$, $\mathbf{W}$, and $\mathbf{U}$. By fixing the values of $\mathbf{W}$ and $\mathbf{U}$, it essentially becomes an LMI in $\mathbf{F}$.

Then, to deal with the semi-infinite constraint of the relay power (43c), we can express $P_{\mathrm{R},m}$ as follows based on the definitions in (51):

$$P_{\mathrm{R},m} = \left\| \mathbf{t}_m + \sum_{k=1}^{K} \mathbf{T}_{m,k} \mathbf{h}_{m,k} \right\|^2. \tag{53}$$

Substituting (53) into (43c) and again applying Lemma 1, (43c) can be equivalently recast as the matrix inequality (49c), whose left-hand side is bilinear in $\mathbf{W}_m$ and $\mathbf{F}$, which is an LMI in $\mathbf{F}$ when $\mathbf{W}_m$ is fixed.

Finally, (43d) can be expressed as $\|\mathbf{f}_k\|^2 \leq P^{\max}_{\mathrm{S},k}$, which can be equivalently recast as (49d) by using the Schur complement rule of [33]. The SDP form (49) is then readily obtained. ■

3) Relay AF Matrix Design: Since the constraint (49d) is independent of the relay AF matrices $\mathbf{W}$, this subproblem is equivalent to

$$\min_{t,\mathbf{W},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l},\boldsymbol{\tau}^{\mathrm{p}}_{m}} \quad t \quad \text{s.t.} \quad (49\text{b}), (49\text{c}). \tag{54}$$

The given problem becomes a standard SDP in $\mathbf{W}$ by noting that $\mathbf{Q}_{k,l}$ and $\mathbf{P}_m$ in (49b) and (49c), respectively, are LMIs in $\mathbf{W}$, provided that the other design variables are kept fixed.

The convergence analysis of the overall iterative algorithm, which solves problems (48), (49), and (54) with the aid of the block coordinate approach, is similar to that in Section III-B and therefore omitted for brevity. One slight difference from Algorithm 1 is that we initialize $\mathbf{F}_k^{(0)} = \sqrt{P^{\max}_{\mathrm{S},k}} \mathbf{I}_{N_{\mathrm{S},k}\times d_k}\ \forall k \in \mathcal{K}$ and $\mathbf{U}_k^{(0)} = \mathbf{I}_{d_k \times N_{\mathrm{S},k}} \forall k \in \mathcal{K}$, and the iterative algorithm will start by solving for the optimal $\mathbf{W}_m^{(1)}$. Solving (49) imposes a worst-case complexity on the order of $\mathcal{O}(N_{\text{total}}^2 D_{\mathrm{SDP}})$, where D_{SDP} represents the total dimensionality of the semi-definite cones in constraints (49b)–(49d). Comparing the SDP formulation of (49) derived for the norm-bounded CSI errors and the SOCP formulation in (62) deduced for the statistical CSI errors, the total dimensionality of (49) is seen to be significantly larger than that of (62).

V. Transceiver Design for the Quality-of-Service Problem

Here, we turn our attention to the joint transceiver design for the QoS problem (16). Following the same approaches as in Sections III and IV, the solution to the QoS problem can also be obtained by adopting the block coordinate update method. Since the derivations of the corresponding subproblems and algorithms are similar to those in Sections III and IV deduced for the min–max problem, we hereby only present the main results.

A. QoS Problem Under Statistical CSI Errors

1) Receive Filter Design: An optimal $\mathbf{u}_{k,l}$ can be obtained by minimizing $\bar{\varepsilon}_{k,l}(\Delta\mathbf{H}, \Delta\mathbf{G}_k)$ with respect to $\mathbf{u}_{k,l}$, which yields exactly the same solution as the Wiener filter in (28).

2) Source TPC Design: The specific subproblem of finding the optimal $\mathbf{F}$ can be solved by the following QCLP:

$$\min_{\mathbf{F},t} \quad t \tag{55a}$$

$$\text{s.t.} \quad \sum_{q=1}^{K} \mathbf{f}_q^H \mathbf{A}_{1,q}^{k,l} \mathbf{f}_q - 2\Re\left\{\mathbf{f}_k^H \mathbf{a}_2^{k,l}\right\} + a_3^{k,l} \leq \frac{\gamma}{\kappa_{k,l}} \quad \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{55b}$$

$$\sum_{k=1}^{K} \mathbf{f}_k^H \mathbf{A}_{4,k}^m \mathbf{f}_k \leq \eta'_{\mathrm{R},m} \quad \forall m \in \mathcal{M} \tag{55c}$$

$$\mathrm{Tr}(\mathbf{F}_k^H \mathbf{F}_k) \leq P^{\max}_{\mathrm{S},k} \quad \forall k \in \mathcal{K} \tag{55d}$$

where $\eta'_{\mathrm{R},m} \triangleq \rho_m t' - \sigma^2_{\mathrm{R},m} \mathrm{Tr}(\mathbf{W}_m \mathbf{W}_m^H)$.

3) Relay AF Matrix Design: The optimal $\mathbf{W}$ can be found by solving

$$\min_{\mathbf{w},t} \quad t \tag{56a}$$

$$\text{s.t.} \quad \mathbf{w}^H \mathbf{B}_1^{k,l} \mathbf{w} - \sum_{m=1}^{M} 2\Re\left\{\mathbf{w}_m^H \mathbf{b}_{2,m}^{k,l}\right\} + \sum_{m=1}^{M} \mathbf{w}_m^H \mathbf{B}_{3,m}^{k,l} \mathbf{w}_m + b_4^{k,l} \leq \frac{\gamma}{\kappa_{k,l}}\ \forall k, l \tag{56b}$$

$$\mathbf{w}_m^H \mathbf{B}_{5,m} \mathbf{w}_m \leq \rho_m t \quad \forall m \in \mathcal{M}. \tag{56c}$$

B. QoS Problem under Norm-Bounded CSI Errors

1) Receive Filter Design: The optimal $\mathbf{u}_{k,l}$ can be obtained from (48).

2) Source TPC Design: The optimal $\mathbf{F}$ can be obtained as the solution to the following SDP:

$$\min_{t,\mathbf{F},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l},\boldsymbol{\tau}^{\mathrm{p}}_{m}} \quad t \tag{57a}$$

$$\text{s.t.} \quad \mathbf{Q}'_{k,l} \succeq \mathbf{0} \quad \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{57b}$$

$$\mathbf{P}'_m \succeq \mathbf{0} \quad \forall m \in \mathcal{M} \tag{57c}$$

$$\begin{bmatrix} P^{\max}_{\mathrm{S},k} & \mathbf{f}_k^H \\ \mathbf{f}_k & \mathbf{I}_{N_{\mathrm{S},k} d_k} \end{bmatrix} \succeq \mathbf{0} \quad \forall k \in \mathcal{K} \tag{57d}$$

where $\mathbf{Q}'_{k,l}$ is obtained from $\mathbf{Q}_{k,l}$ in (46) upon replacing t by γ in the top-left entry (1,1). Similarly, $\mathbf{P}'_m$ can be obtained by substituting P_{R} with t in the (1,1)th entry of $\mathbf{P}_m$ in (50).

3) Relay AF Matrix Design: The optimal relay AF matrices are obtained by solving

$$\min_{t,\mathbf{W},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l}} \quad t \quad \text{s.t.} \quad (57\text{b}), (57\text{c}). \tag{58}$$

C. Initial Feasibility Search Algorithm

An important aspect of solving the given QoS problem is to find a feasible initial point. Indeed, it has been observed that, if the iterative algorithm is initialized with a random (possibly infeasible) point, the algorithm may fail at the first iteration. Finding a feasible initial point of a nonconvex problem, such as our QoS problem (16), is in general NP-hard. All these considerations motivate the study of an efficient initial feasibility search algorithm, which finds a reasonably "good" starting point for the QoS problem of (16).

Motivated by the "phase I" approach in general optimization theory [33], we formulate the feasibility check problem for the QoS problem as follows:

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad s \tag{59a}$$

$$\text{s.t.} \quad \kappa_{k,l}\mathcal{U}\left\{\varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right)\right\} \le s \;\forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{59b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H \mathbf{F}_k\right) \le P^{\max}_{\mathrm{S},k} \quad \forall k \in \mathcal{K} \tag{59c}$$

where s is a slack variable, which represents an abstract measure for the violation of the constraint (16b). The given problem can be solved iteratively using the block coordinate approach until the objective value s converges or the maximum affordable number of iterations is reached. If, at the $(n+1)^{\mathrm{st}}$ iteration, $s^{(n+1)}$ meets the QoS target γ, then the procedure successfully finds a feasible initial point; otherwise, we claim that the QoS problem is infeasible. In this case, it is necessary to adjust γ or to drop the services of certain users by incorporating an admission control procedure, which, however, is beyond the scope of this paper.

Interestingly, (59) can be reformulated as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{\forall k \in \mathcal{K}, l \in \mathcal{D}_k} \kappa_{k,l}\mathcal{U}\left\{\varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right)\right\} \tag{60a}$$

$$\text{s.t.} \quad \mathcal{U}\left\{P_{\mathrm{R},m}\left(\Delta\mathbf{H}_m\right)\right\} \le \rho_m P^{\infty}_{\mathrm{R}} \quad \forall m \in \mathcal{M} \tag{60b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H \mathbf{F}_k\right) \le P^{\max}_{\mathrm{S},k} \quad \forall k \in \mathcal{K} \tag{60c}$$

where we have $P^{\infty}_{\mathrm{R}} \to \infty$, which is equivalent to removing the constraint on the relay's transmit power. In fact, (60) becomes exactly the same as the min–max problem of (15) upon setting $P_{\mathrm{R}} = P^{\infty}_{\mathrm{R}}$. We therefore propose an efficient iterative feasibility search algorithm, which is listed as Algorithm 2, based on the connection between the feasibility check and the min–max problems.

Algorithm 2 Iterative Initial Feasibility Search Algorithm for the QoS problems

1: **repeat**
2: Solve one cycle of the problem (60) and denote the current objective value by $\hat{\gamma}^{(i+1)}$;
3: Verify if $\hat{\gamma}^{(i+1)} \le \gamma$, and if so, stop the algorithm;
4: $i \leftarrow i + 1$;
5: **until** Termination criterion is satisfied, e.g., $|\hat{\gamma}^{(i)} - \hat{\gamma}^{(i-1)}| \le \epsilon$; or the maximum allowed number of iteration is reached.

Based on the definition of $\mathcal{U}\{\cdot\}$ in (14), Algorithm 2 is applicable to the QoS problems associated with both types of CSI errors considered. Furthermore, Algorithm 2 indeed provides a feasible initial point for the QoS problem if it exists. Otherwise, it provides a certificate of infeasibility if $\hat{\gamma}^{(i+1)} > \gamma$ after a few iterations. Then, the QoS problem is deemed infeasible in this case, and the admission control procedure may deny the access of certain users.

VI. Simulation Experiments and Discussions

This section presents our Monte Carlo simulation results for verifying the resilience of the proposed transceiver optimization algorithms against CSI errors. In all simulations, we assume that there are $K = 2$ S−D pairs, which communicate with the assistance of $M = 2$ relays. Each node is equipped with $N_{\mathrm{S},k} = N_{\mathrm{R},m} = N_{\mathrm{D},k} = 3$ antennas $\forall\, k \in \mathcal{K}, m \in \mathcal{M}$. Each source transmits 2 independent quadrature phase-shift keying (QPSK) modulated data streams to its corresponding destination, i.e., $d_k = 2 \;\forall\, k \in \mathcal{K}$. Equal noise variances of $\sigma^2_{\mathrm{D},k} = \sigma^2_{\mathrm{R},m}$ are assumed. The maximum source and relay transmit power is normalized to one, i.e., we have $P^{\max}_{\mathrm{S},k} = 1 \;\forall\, k \in \mathcal{K}$ and $\rho_m P_{\mathrm{R}} = 1,\; \forall\, m \in \mathcal{M}$. Equal weights of $\kappa_{k,l}$ are assigned to the different data streams, unless otherwise stated. The channels are assumed to be flat fading, with the coefficients given by i.i.d. zero-mean unit-variance complex Gaussian random variables. The signal-to-noise ratios (SNRs) at the relays and the destinations are defined as $\mathrm{SNR}_{\mathrm{R},m} \triangleq P^{\max}_{\mathrm{S}}/|N_{\mathrm{R},m}\sigma^2_{\mathrm{R},m}|$ and $\mathrm{SNR}_{\mathrm{D},k} \triangleq P^{\max}_{\mathrm{R}}/|N_{\mathrm{D},k}\sigma^2_{\mathrm{D},k}|$, respectively. The optimization solver MOSEK [31] is used for solving each optimization problem.

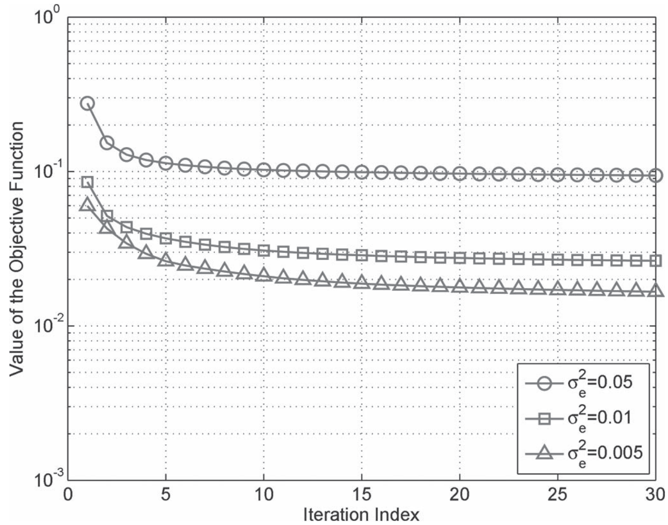

Fig. 2. Convergence behavior of the proposed iterative algorithm with statistical CSI errors.

A. Performance Evaluation Under Statistical CSI Errors

We first evaluate the performance of the iterative algorithm proposed in Section III under statistical CSI errors. The channel correlation matrices in (9) and (10) are obtained by the widely employed exponential model of [37]. Specifically, their entries are given by $[\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}]_{i,j} = [\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}]_{i,j} = \alpha^{|i-j|}$ and $[\boldsymbol{\Psi}_{\mathrm{H}_{m,k}}]_{i,j} = [\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}]_{i,j} = \sigma_e^2 \beta^{|i-j|}, i, j \in \{1, 2, 3\}$, where α and β are the correlation coefficients, and σ_e^2 denotes the variance of the CSI errors. The available channel estimates $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$ are generated according to $\hat{\mathbf{H}}_{m,k} \sim \mathcal{CN}(\mathbf{0}_{N_{\mathrm{R},m} \times N_{\mathrm{S},k}}, ((1-\sigma_e^2)/\sigma_e^2)\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}} \otimes \boldsymbol{\Psi}^T_{\mathrm{H}_{m,k}})$ and $\hat{\mathbf{G}}_{k,m} \sim \mathcal{CN}(\mathbf{0}_{N_{\mathrm{D},k} \times N_{\mathrm{R},m}}, ((1-\sigma_e^2)\,/\,\sigma_e^2)\,\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \otimes \boldsymbol{\Psi}^T_{\mathrm{G}_{k,m}})$, respectively, such that the entries of the true channel matrices have unit variances. We compare the robust transceiver design proposed in Algorithm 1 to the 1) nonrobust design, which differs from the robust design in that it assumes $\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}} = \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} = \mathbf{0}$ and $\boldsymbol{\Psi}_{\mathrm{H}_{m,k}} = \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} = \mathbf{0}$, i.e., it neglects the effects of the CSI errors; 2) perfect CSI case, where the true channel matrices $\mathbf{H}_{m,k}$ and $\mathbf{G}_{k,m}$ are used instead of the estimates $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$ in Algorithm 1 and where there are no CSI errors, i.e., we have $\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}} = \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} = \mathbf{0}$ and $\boldsymbol{\Psi}_{\mathrm{H}_{m,k}} = \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} = \mathbf{0}$. The curves labeled "optimal MSE" correspond to the value of the objective function in (17a) after optimization by Algorithm 1. In all the simulation figures, the MSEs of the different approaches are calculated by averaging the squared error between the transmitted and estimated experimental data symbols over 1000 independent CSI error realizations and 10 000 QPSK symbols for each realization.

As a prelude to the presentation of our main simulation results in the following, the convergence behavior of Algorithm 1 is presented for different CSI error variances, It can be observed in Fig, 2 that in all cases, the proposed algorithm can converge within a reasonable number of iterations, Therefore, in our experimental work, we set the number of iterations to a fixed value of 5, and the resultant performance gains will be discussed in the following.

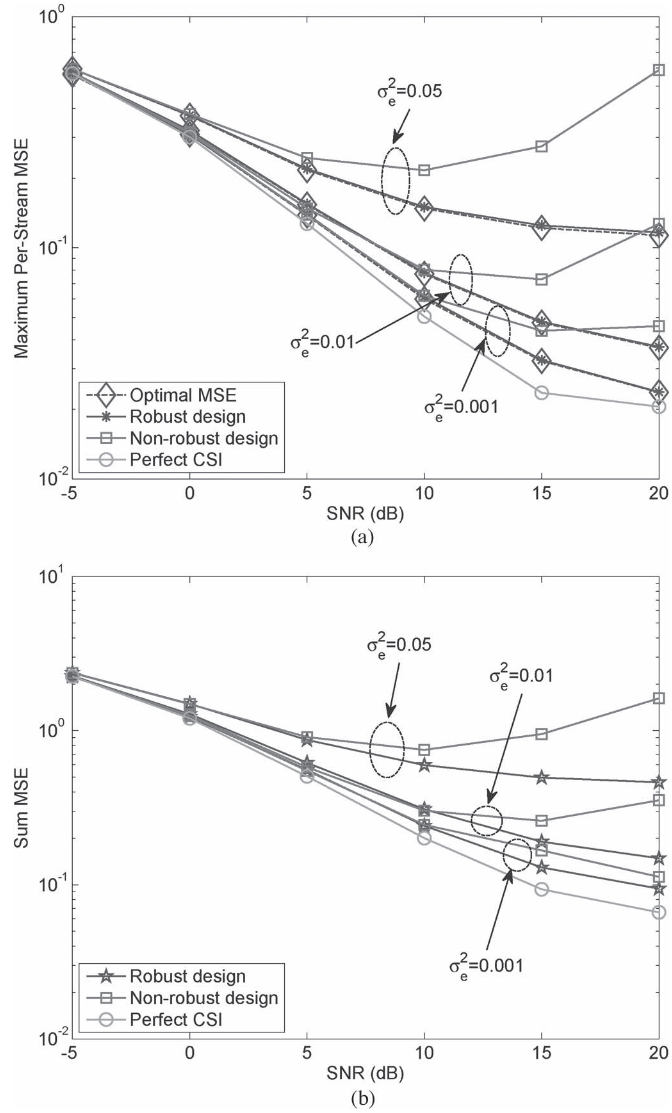

Fig. 3. MSE performance of different design approaches versus SNR. (a) Maximum per-stream MSE. (b) Sum MSE ($\mathrm{SNR}_{\mathrm{R},m} = \mathrm{SNR}_{\mathrm{D},k} = \mathrm{SNR}$, $\alpha = \beta = 0.5$).

1) Experiment A.1 (MSE Performance): In Fig. 3(a), the maximum per-stream MSE among all the data streams is shown as a function of the SNR for different values of CSI error variance. It is observed that the proposed robust design approach achieves better resilience against the CSI errors than the non-robust design approach. The performance gains become more evident in the medium-to-high SNR range. For the nonrobust design, degradations are observed because the MSE obtained at high SNRs is dominated by the interference, rather than by the noise. Therefore, the relays are confined to relatively low transmit power in order to control the interference. This, in turn, leads to performance degradation imposed by the CSI errors. In contrast, the proposed robust design is capable of compensating for the extra interference imposed by the CSI errors, thereby demonstrating its superiority over its nonrobust counterpart. Furthermore, we observe that the "Optimal MSE" and our simulation results tally well, which justifies the approximations invoked in calculating the per-stream MSE in (13). In addition

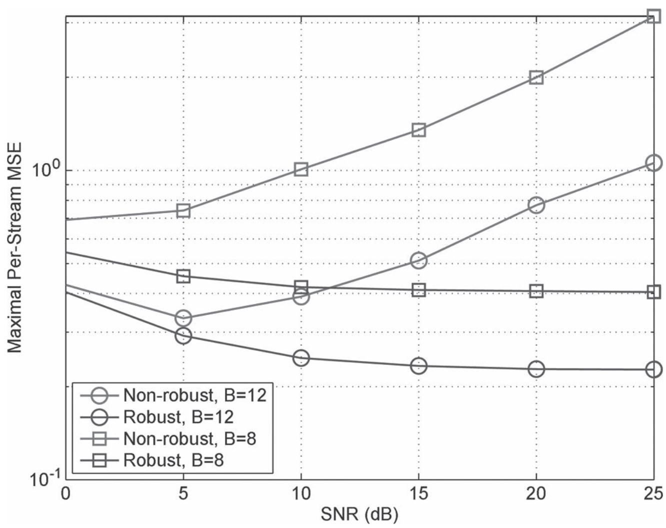

Fig. 4. Per-stream MSE performance with the optimized codebook based on the GLA-VQ. ($B = 8$ corresponds to $\sigma_e^2 = 0.334$, and $B = 12$ corresponds to $\sigma_e^2 = 0.175$.)

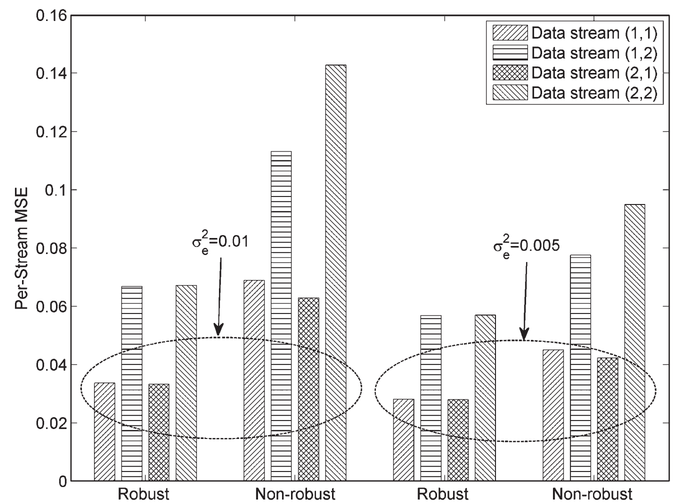

Fig. 5. Comparison of the per-stream MSEs of the robust and nonrobust design approaches ($\mathrm{SNR}_{\mathrm{R},m} = \mathrm{SNR}_{\mathrm{D},k} = 15$ dB, and $\alpha = \beta = 0.5$).

to the per-stream performance, the overall system performance[4] quantified in terms of the sum MSE of different approaches is examined in Fig. 3(b), where a similar trend to that of Fig. 3(a) can be observed.

The MSE performance associated with a limited number of feedback bits is also studied. To this end, we assume that each user is equipped with a codebook that is optimized using the generalized Lloyd algorithm of vector quantization (GLA-VQ) [38]. Each user then quantizes the channel vector, and the corresponding codebook index is fed back to the central processing unit. The results presented in Fig. 4 show that the proposed algorithm significantly outperformed the nonrobust one for the different number of quantization bits considered.

2) Experiment A.2 (Data Stream Fairness): Next, we examine the accuracy of the proposed robust design in providing weighted fairness for the different data streams. To this end, we set the weights for the different data streams to be $\kappa_{1,1} = \kappa_{2,1} = 1/3$ and $\kappa_{1,2} : \kappa_{2,2} = 1/6$. Fig. 5 shows the MSE of each data stream for different values of the error variance. Comparing the two methods, the robust design approach results in significantly better weighted fairness than the nonrobust one. In particular, the MSEs obtained are strictly inversely proportional to the predefined weights. This feature is particularly desirable for multimedia communications, where the streams corresponding to different service types may have different priorities.

3) Experiment A.3 (Effects of Channel Correlation): The effects of channel correlations on the MSE performance of the different approaches are investigated in Fig. 6. It can be observed that the performance of all the approaches is degraded as the correlation factor α increases. While the robust design

[4]Note that the objective of portraying the sum MSE performance is to validate whether the proposed robust design approach can also achieve a performance gain over the nonrobust approach in terms of its overall performance. In fact, the sum MSE performance can be optimized by solving a design problem with the sum MSE being the objective function.

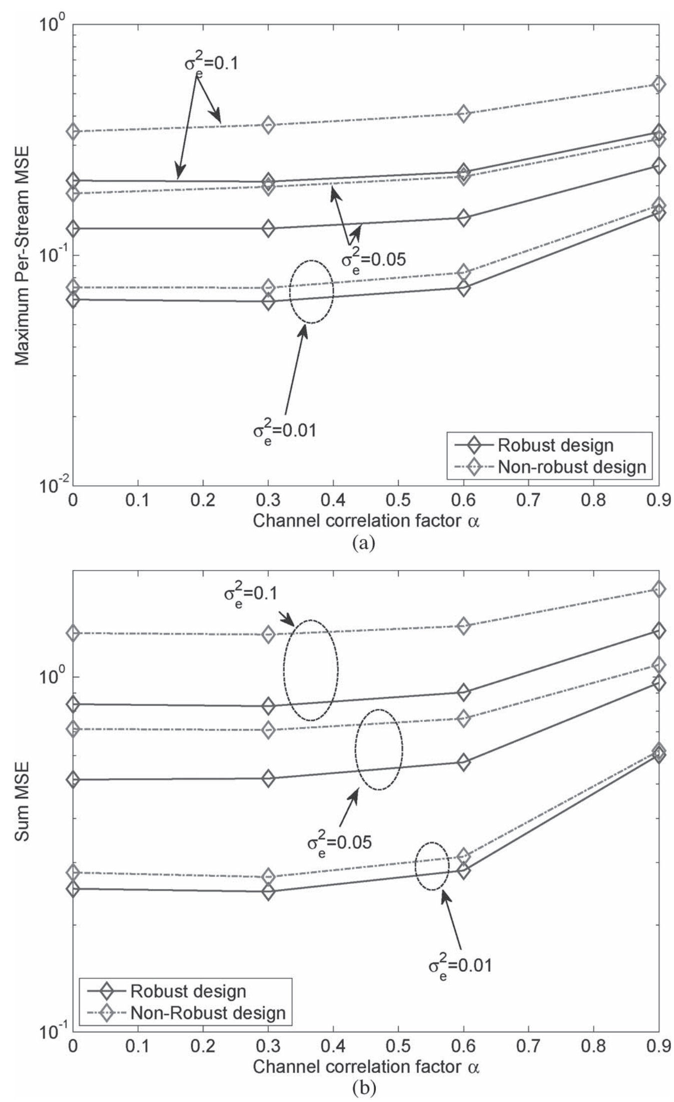

Fig. 6. MSE performance of different design approaches versus correlation factor of the source–relay channels. (a) Per-stream MSE. (b) Sum MSE ($\mathrm{SNR}_{\mathrm{R},m} = \mathrm{SNR}_{\mathrm{D},k} = 10$ dB, and $\beta = 0.45$).

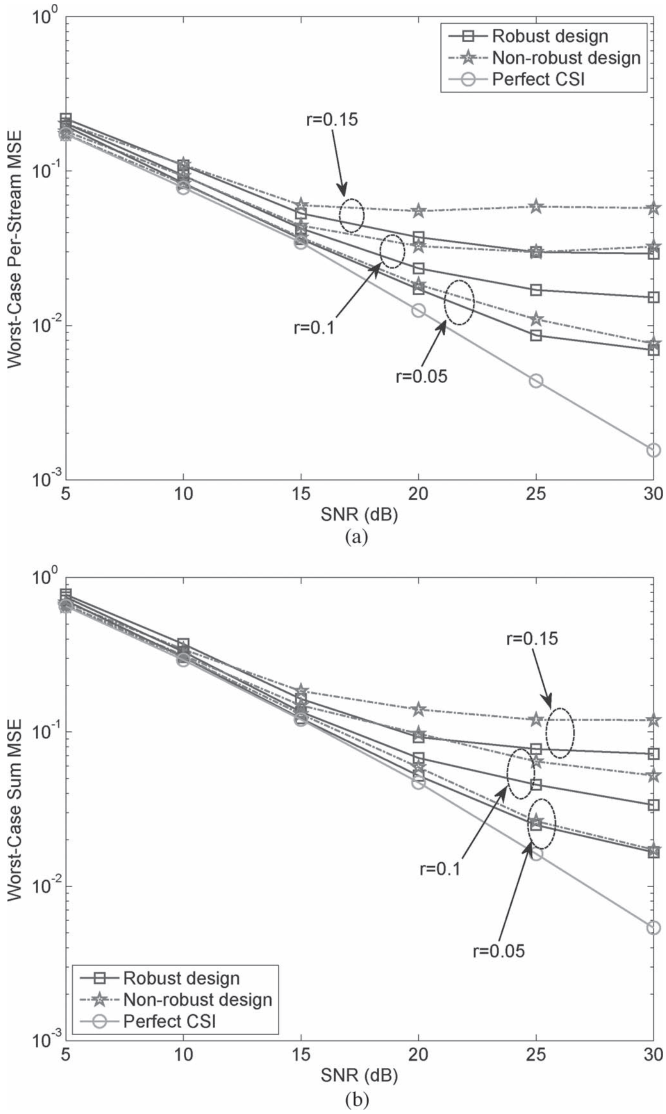

Fig. 7. MSE performance of different design approaches versus SNR. (a) Worst-case per-stream MSE. (b) Worst-case sum MSE.

shows consistent performance gains over its nonrobust one associated with different α and σ_e^2, the discrepancies between the two approaches tend to become less significant with an increase in α. This is because the achievable *spatial multiplexing* gain is reduced by a higher channel correlation; therefore, the robust design can only attain a limited performance improvement in the presence of high channel correlations.

B. Performance Evaluation Under Norm-Bounded CSI Errors

Here, we evaluate the performance of the proposed worst case design approach in Section V for the min–max problem under norm-bounded CSI errors. Similar to that given earlier, we compare the proposed robust design approach both to the nonrobust approach and to the perfect CSI scenario. We note that the power of each relay is a function of $\Delta\mathbf{H}_m$. According to the worst-case robust design philosophy, the maximum relay transmit power has to be bounded by the power budget, whereas the average relay transmit power may become significantly

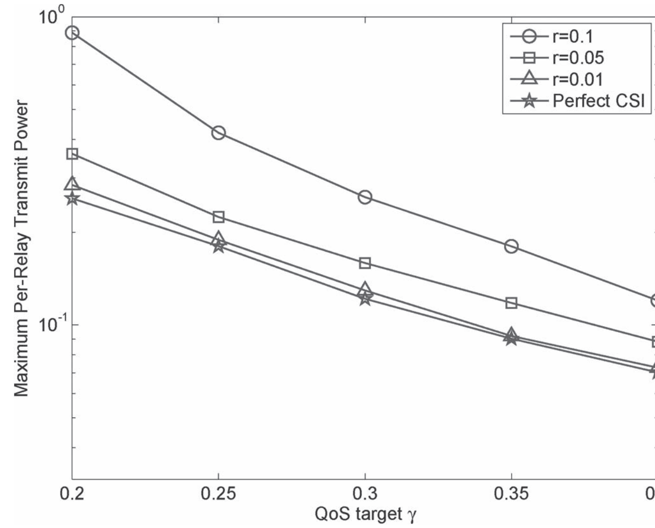

Fig. 8. Maximum relay transmit power versus QoS targets with different uncertainty sizes of the CSI errors.

lower than that of the nonrobust design. To facilitate a fair comparison of the different approaches, we therefore assume the absence of CSI errors for the $\mathtt{S}-\mathtt{R}$ links, i.e., we have $\Delta\mathbf{H}_{m,k} = \mathbf{0}$. For the $\mathtt{R}-\mathtt{D}$ links, we consider the uncertainty regions with equal radius, i.e., we have $\xi_{k,m} = r\ \forall k \in \mathcal{K}, m \in \mathcal{M}$. To determine the worst-case per-stream MSE, we generate 5000 independent realizations of the CSI errors. For each realization, we evaluate the maximum per-stream MSE averaged over 1000 QPSK symbols and random Gaussian noise. Then, the worst-case per-stream MSE is obtained by selecting the largest one among all the realizations.

1) Experiment B.1 (MSE Performance): The worst-case per-stream MSE and the worst-case sum MSE are reported in Fig. 7 as a function of the SNR. Three sizes of the uncertainty region are considered, i.e., $r = 0.05$, $r = 0.1$, and $r = 0.15$. Focusing on the first case, it can be seen that the performance achieved by our robust design approach first monotonically decreases as the SNR increases and then subsequently remains approximately constant at high-SNR values. This is primarily because, at low SNR, the main source of error in the estimation of the data streams is the channel noise. At high SNR, the channel noise is no longer a concern, and the MSE is dominated by the CSI errors. Observe also in Fig. 7 that for $r = 0.1$ and $r = 0.15$, the MSE is clearly higher, although it presents a similar trend to the case of $r = 0.5$. The performance gain achieved by the robust design also becomes more noticeable for these larger sizes of the uncertainty regions.

2) Experiment B.2 (Relay Power Consumption): Next, we investigate the performance of the approach proposed in Section VI for the QoS problem under the norm-bounded CSI errors. The maximum per-relay transmit power is plotted in Fig. 8 as a function of the QoS target γ for different sizes of uncertainty regions. As expected, it can be observed that the relay power for all cases decreases as the QoS target is relaxed. An important observation from this figure is that, when the size of uncertainty region is large, the required relay transmit power becomes significantly higher than the perfect CSI case. From an

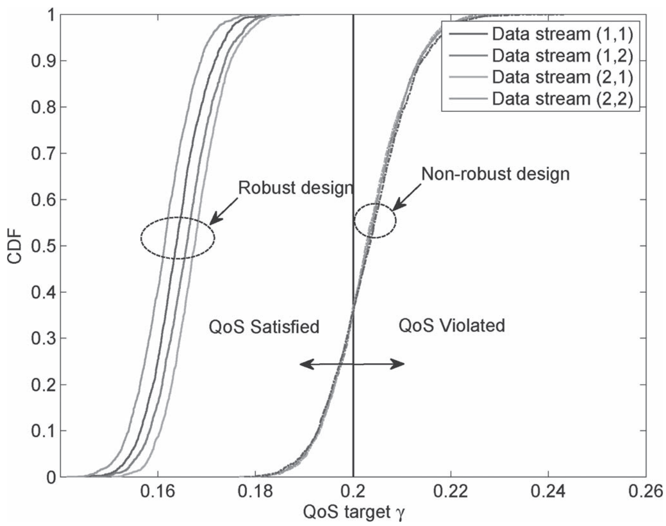

Fig. 9. CDFs of per-stream MSEs using the robust and nonrobust approaches for SNR = 5 dB.

energy-efficient design perspective, this is not desirable, which motivates the consideration of the min–max design in such applications.

3) Experiment B.3 (CDF of Per-stream MSE): Finally, we evaluate how consistently the QoS constraints of all the data streams can be satisfied by the proposed design approach for the QoS problem. In this experiment, the CSI errors of both the S−R and R−D links are taken into consideration and generated according to the i.i.d. zero-mean complex Gaussian distribution with a variance of $\sigma_e^2 = 0.001$. Then, the probability that the CSI errors are bounded by the predefined radius r can be formulated as [9, Sec. IV-C]

$$\Pr\left\{\|\mathbf{h}_{m,k}\|^2 \le r^2\right\} = \Pr\left\{\|\mathbf{g}_{k,m}\|^2 \le r^2\right\} = \frac{1}{\Gamma\left(\frac{N^2}{2}\right)}\gamma\left(\frac{N^2}{2}, \frac{r^2}{\sigma_e^2}\right) \tag{61}$$

where $\Gamma(\cdot)$ and $\gamma(\cdot,\cdot)$, respectively, denote the complete and lower incomplete Gamma functions. Given the required bounding probability of, e.g., 90% in the simulation, the radius r can be numerically determined from (61). Fig. 9 shows the cumulative distribution functions (cdfs) of the MSE of each data stream using both the robust and nonrobust design methods. As expected, the proposed robust method ensures that the MSE of each data stream never exceeds the QoS target shown as the vertical black solid line in Fig. 9. By contrast, for the nonrobust design, the MSE frequently violates the QoS target, namely for more than 60% of the realizations. Based on these observations, we conclude that the proposed robust design approach outperforms its nonrobust counterpart in satisfying the QoS constraints for all the data streams.

VII. Conclusion

Jointly optimized source TPCs, AF relay matrices, and receive filters were designed by considering two different types of objective functions with specific QoS consideration in the presence of CSI errors in both the S−R and R−D links. To this end, a pair of practical CSI error models, namely, the statistical and the norm-bounded models were considered. Accordingly, the robust transceiver design approach was formulated to minimize the maximum per-stream MSE subject to the source and relay power constraints (min–max problem). To solve the nonconvex optimization problems formulated, an iterative solution based on the block coordinate update algorithm was proposed, which involves a sequence of convex conic optimization problems. The proposed algorithm generated a convergent sequence of objective function values. The problem of relay power minimization subject to specific QoS constraints and to source power constraints was also studied. An efficient feasibility search algorithm was proposed by studying the link between the feasibility check and the min–max problems. Our simulation results demonstrate a significant enhancement in the performance of the proposed robust approaches over the conventional nonrobust approaches.

Appendix A
Transformation of (34) Into a Standard Second-Order Cone Programming

By exploiting the separable structure of (34) and the properties of quadratic terms, the problem can be recast as

$$\min_{\substack{t,\{\mathbf{f}_k\},\\ \{\boldsymbol{\lambda}^{k,l}\},\{\boldsymbol{\theta}^m\}}} \quad t \tag{62a}$$

$$\text{s.t.} \quad \left\|\left(\mathbf{A}_{1,q}^{k,l}\right)^{1/2}\mathbf{f}_q\right\| \le \lambda_q^{k,l} \quad \forall q,k\in\mathcal{K}, q\ne k, l\in\mathcal{D}_k \tag{62b}$$

$$\left\|\left(\mathbf{A}_{1,k}^{k,l}\right)^{1/2}\mathbf{f}_k - \left(\mathbf{A}_{1,k}^{k,l}\right)^{-1/2}\mathbf{a}_2^{k,l}\right\| \le \lambda_k^{k,l} \quad \forall k\in\mathcal{K}, l\in\mathcal{D}_k \tag{62c}$$

$$\|\boldsymbol{\lambda}^{k,l}\|^2 - \left(\mathbf{a}_2^{k,l}\right)^H\left(\mathbf{A}_{1,k}^{k,l}\right)^{-1}\mathbf{a}_2^{k,l} + a_3^{k,l} \le \frac{t}{\kappa_{k,l}} \quad \forall k\in\mathcal{K}, l\in\mathcal{D}_k \tag{62d}$$

$$\left\|\left(\mathbf{A}_{4,k}^m\right)^{1/2}\mathbf{f}_k\right\| \le \theta_k^m \ \forall k\in\mathcal{K}, m\in\mathcal{M} \tag{62e}$$

$$\|\boldsymbol{\theta}^m\| \le \sqrt{\eta_{\mathrm{R},m}} \ \forall m\in\mathcal{M} \tag{62f}$$

$$\|\mathbf{f}_k\| \le \sqrt{P_{\mathrm{S},k}^{\max}} \ \forall k\in\mathcal{K} \tag{62g}$$

where $\boldsymbol{\lambda}^{k,l} = [\lambda_1^{k,l},\ldots,\lambda_K^{k,l}]^T$, $\boldsymbol{\theta}^m = [\theta_1^m,\ldots,\theta_K^m]^T$, and t are auxiliary variables. The main difficulty in solving this problem is with (62d), which is a so-called *hyperbolic constraint* [32], whereas the remaining constraints are already in the form of SOC.

To tackle (62d), we observe that, for any $\mathbf{x}$ and $y,z\le 0$, the following equation holds:

$$\|\mathbf{x}\|^2 \le yz \Longleftrightarrow \left\|\begin{bmatrix} 2\mathbf{x} \\ y-z \end{bmatrix}\right\| \le y+z. \tag{63}$$

We can apply (63) to transform (62d) into

$$\left\|\begin{bmatrix} 2\boldsymbol{\lambda}^{k,l} \\ \frac{t}{\kappa_{k,l}} + \left(\mathbf{a}_2^{k,l}\right)^H \left(\mathbf{A}_{1,k}^{k,l}\right)^{-1} \mathbf{a}_2^{k,l} - a_3^{k,l} - 1 \end{bmatrix}\right\| \le \frac{t}{\kappa_{k,l}} + \left(\mathbf{a}_2^{k,l}\right)^H \left(\mathbf{A}_{1,k}^{k,l}\right)^{-1} \mathbf{a}_2^{k,l} - a_3^{k,l} + 1. \quad (64)$$

Therefore, substituting (62d) by (64), we can see that (62) is in the form of a standard SOCP.

APPENDIX B
PROOF OF PROPOSITION 1

First, we define $\boldsymbol{\mathcal{T}}_k \triangleq [\boldsymbol{\mathcal{T}}_{k,1}, \ldots, \boldsymbol{\mathcal{T}}_{k,K}]$ and $\boldsymbol{\mathcal{G}}_k \triangleq [\sigma_{\mathrm{R},1}\boldsymbol{\mathcal{G}}_{k,1}, \ldots, \sigma_{\mathrm{R},M}\boldsymbol{\mathcal{G}}_{k,M}]$. We exploit the fact that, for any vectors $\{\mathbf{a}_k\}_{k=1}^N$, the following identity holds:

$$\sum_{k=1}^{N} \|\mathbf{a}_k\|^2 = \left\| \left[\mathbf{a}_1^T, \ldots, \mathbf{a}_N^T\right] \right\|^2. \quad (65)$$

The per-stream MSE (13) can be subsequently expressed as

$$\begin{aligned} \varepsilon_{k,l} = & \Bigg\| \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{T}}_k + \sum_{m=1}^{M} \mathbf{u}_{k,l}^H \Delta\mathbf{G}_{k,m} \left[\boldsymbol{\mathcal{W}}_{m,1}\mathbf{F}_1, \ldots, \boldsymbol{\mathcal{W}}_{m,K}\mathbf{F}_K\right] \\ & + \sum_{q=1}^{K}\sum_{m=1}^{M} \Bigg[\mathbf{0}_{1\times\sum_{t=1}^{q} d_t}, \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_{k,m} \\ & \times \Delta\mathbf{H}_{m,q}\mathbf{F}_q, \mathbf{0}_{1\times\sum_{q+1}^{K} d_t}\Bigg] \Bigg\|^2 \\ & + \Bigg\| \sum_{m=1}^{M} \Bigg[\mathbf{0}_{1\times\sum_{p=1}^{m-1} N_{\mathrm{R},p}}, \mathbf{u}_{k,l}^H \Delta\mathbf{G}_{k,m}\mathbf{W}_m, \\ & \mathbf{0}_{1\times\sum_{p=m+1}^{M} N_{\mathrm{R},p}}\Bigg] \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_k \Bigg\|^2 + \sigma_{\mathrm{D},k}^2 \|\mathbf{u}_{k,l}^H. \end{aligned} \quad (66)$$

Upon applying the identity $\mathrm{vec}^T(\mathbf{ABC}) = \mathrm{vec}(\mathbf{B})^T(\mathbf{C} \otimes \mathbf{A}^T)$ to (66), we arrive at

$$\begin{aligned} \varepsilon_{k,l} = & \left\| \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{T}}_k - \bar{\mathbf{e}}_{k,l}^T + \sum_{m=1}^{M} \mathbf{g}_{k,m}^T \mathbf{C}_{1,m}^{k,l} + \sum_{m,q} \mathbf{h}_{m,q}^T \mathbf{D}_{m,q}^{k,l} \right\|^2 \\ & + \left\| \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_k + \sum_{m=1}^{M} \mathbf{g}_{k,m}^T \mathbf{C}_{2,m}^{k,l} \right\|^2 + \left\| \sigma_{\mathrm{D},k} \mathbf{u}_{k,l}^H \right\|^2 \end{aligned} \quad (67)$$

where $\mathbf{h}_{m,k} \triangleq \mathrm{vec}(\Delta\mathbf{H}_{m,k})$ and $\mathbf{g}_{k,m} \triangleq \mathrm{vec}(\Delta\mathbf{G}_{k,m})$ denote the vectorized CSI errors, $\bar{\mathbf{e}}_{k,l} \triangleq [\mathbf{0}_{1\times\sum_{t=1}^{k-1} d_t}, \mathbf{e}_{k,l}^T, \mathbf{0}_{1\times\sum_{t=k+1}^{K} d_t}]^T$, and the following matrices have also been introduced:

$$\mathbf{C}_{1,m}^{k,l} \triangleq \left[(\boldsymbol{\mathcal{W}}_{m,1}\mathbf{F}_1) \otimes \mathbf{u}_{k,l}^*, \ldots, (\boldsymbol{\mathcal{W}}_{m,K}\mathbf{F}_K) \otimes \mathbf{u}_{k,l}^*\right] \quad (68)$$

$$\mathbf{C}_{2,m}^{k,l} \triangleq \Bigg[\mathbf{0}_{N_{\mathrm{D},k}N_{\mathrm{R},m}\times\sum_{p=1}^{m-1} N_{\mathrm{R},p}}, \mathbf{W}_m \otimes \mathbf{u}_{k,l}^* \quad \mathbf{0}_{N_{\mathrm{D},k}N_{\mathrm{R},m}\times\sum_{p=m+1}^{M} N_{\mathrm{R},p}}\Bigg] \quad (69)$$

$$\mathbf{D}_{m,q}^{k,l} \triangleq \Bigg[\mathbf{0}_{N_{\mathrm{S},q}N_{\mathrm{R},m}\times\sum_{t=1}^{q-1} d_t}, \mathbf{F}_q \otimes \left(\boldsymbol{\mathcal{G}}_{k,m}^T \mathbf{u}_{k,l}^*\right) \quad \mathbf{0}_{N_{\mathrm{S},q}N_{\mathrm{R},m}\times\sum_{t=q+1}^{K} d_t}\Bigg]. \quad (70)$$

Again, by exploiting the property in (65), we can write (67) in a more compact form as follows:

$$\begin{aligned} \varepsilon_{k,l} = & \Bigg\| \underbrace{\left[\mathbf{u}_{k,l}^H \boldsymbol{\mathcal{T}}_k - \bar{\mathbf{e}}_{k,l}, \mathbf{u}_{k,l}^H \overline{\boldsymbol{\mathcal{G}}}_k, \sigma_{\mathrm{D},k}\mathbf{u}_{k,l}^H\right]}_{\boldsymbol{\theta}_{k,l}} \\ & + \sum_{m=1}^{M} \mathbf{g}_{k,m}^T \underbrace{\left[\mathbf{C}_{1,m}^{k,l}, \mathbf{C}_{2,m}^{k,l}, \mathbf{0}_{N_{\mathrm{D},k}N_{\mathrm{R},m}\times N_{\mathrm{D},k}}\right]}_{\boldsymbol{\Theta}_m^{k,l}} \\ & + \sum_{m=1}^{M}\sum_{q=1}^{K} \mathbf{h}_{m,q}^T \underbrace{\left[\mathbf{D}_{m,q}^{k,l}, \mathbf{0}_{N_{\mathrm{R},m}N_{\mathrm{S},q}\times N_{\mathrm{R}}+N_{\mathrm{D},k}}\right]}_{\boldsymbol{\Phi}_{m,q}^{k,l}} \Bigg\|^2. \end{aligned} \quad (71)$$

Substituting (71) into (43b), we can express (43b) as

$$\left(\boldsymbol{\theta}_{k,l} + \sum_{m=1}^{M} \mathbf{g}_{k,m}^T \boldsymbol{\Theta}_m^{k,l} + \sum_{m=1}^{M}\sum_{q=1}^{K} \mathbf{h}_{m,q}^T \boldsymbol{\Phi}_{m,q}^{k,l}\right) \times \left(\boldsymbol{\theta}_{k,l} + \sum_{m=1}^{M} \mathbf{g}_{k,m}^T \boldsymbol{\Theta}_m^{k,l} + \sum_{m=1}^{M}\sum_{q=1}^{K} \mathbf{h}_{m,q}^T \boldsymbol{\Phi}_{m,q}^{k,l}\right)^H \le t \quad (72)$$

where the uncertain blocks $\mathbf{h}_{m,k}$ and $\mathbf{g}_{k,m}$ should satisfy $\|\mathbf{h}_{m,k}\|_S = \|\mathbf{h}_{m,k}\| \le \xi_{m,k}$ and $\|\mathbf{g}_{k,m}\|_S = \|\mathbf{g}_{k,m}\| \le \eta_{k,m}$, respectively. Through a direct application of Lemma 1, (72) can readily be recast as (46) where the nonnegativity of $\boldsymbol{\tau}_{k,l}^{\mathrm{G}}$ and $\boldsymbol{\tau}_{k,l}^{\mathrm{H}}$ has been implicitly included in the positive semi-definite nature of $\mathbf{Q}_{k,l}$.

ACKNOWLEDGMENT

The authors would like to thank the Editor and the anonymous reviewers for their valuable comments, which have helped to improve the quality and the presentation of this paper.

REFERENCES

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

[2] S. W. Peters, A. Y. Panah, K. T. Truong, and R. W. Heath, "Relay architectures for 3GPP LTE-advanced," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 1, Jul. 2009, Art. ID. 618 787.

[3] C. Nie, P. Liu, T. Korakis, E. Erkip, and S. S. Panwar, "Cooperative relaying in next-generation mobile WiMAX networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 299–305, Mar. 2013.

[4] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution towards 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Commun.*, vol. 21, no. 3, pp. 118–127, Jun. 2014.

[5] S. Berger, M. Kuhn, A. Wittneben, T. Unger, and A. Klein, "Recent advances in amplify-and-forward two-hop relaying," *IEEE Commun. Mag.*, vol. 47, no. 7, pp. 50–56, Jul. 2009.

[6] 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.

[7] A. El-Keyi and B. Champagne, "Adaptive linearly constrained minimum variance beamforming for multiuser cooperative relaying using the Kalman filter," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 641–651, Feb. 2010.

[8] 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.

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

[10] B. K. Chalise and L. Vandendorpe, "Optimization of MIMO relays for multipoint-to-multipoint communications: Nonrobust and robust designs," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6355–6368, Dec. 2010.

[11] J. Lee, J.-K. Han, and J. Zhang, "MIMO technologies in 3GPP LTE and LTE-advanced," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 1, Jul. 2009, Art. ID. 302092.

[12] M. Fadel, A. El-Keyi, and A. Sultan, "QOS-constrained multiuser peer-to-peer amplify-and-forward relay beamforming," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1397–1408, Mar. 2012.

[13] W. Liu, L.-L. Yang, and L. Hanzo, "SVD-assisted multiuser transmitter and multiuser detector design for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 1016–1021, Feb. 2009.

[14] 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.

[15] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Singapore: Athena Scientific, 1999.

[16] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[17] 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.

[18] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3678–3689, Aug. 2008.

[19] A. Pascual-Iserte, D. P. Palomar, A. I. Perez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 346–360, Jan. 2006.

[20] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM J. Imag. Sci.*, vol. 6, no. 3, pp. 1758–1789, Sep. 2013.

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

[22] 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.

[23] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "On the design of amplify-and-forward MIMO-OFDM relay systems with QoS requirements specified as Schur-convex functions of the MSEs," *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1871–1877, May 2013.

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

[25] 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.

[26] D. Zheng, J. Liu, K.-K. Wong, H. Chen, and L. Chen, "Robust peer-to-peer collaborative-relay beamforming with ellipsoidal CSI uncertainties," *IEEE Signal Process. Lett.*, vol. 16, no. 4, pp. 442–445, Apr. 2012.

[27] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2005.

[28] C. Xing, S. Ma, and Y.-C. Wu, "Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2273–2283, Apr. 2010.

[29] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, Version 2.0 beta," Sep. 2013. [Online]. Available: http://cvxr.com/cvx

[30] 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.

[31] E. D. Andersen and K. D. Andersen, *MOSEK Modeling Manual*, Aug. 2013. [Online]. Available: http://mosek.com

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

[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U. K.: Cambridge Univ. Press, 2004.

[34] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Time Methods in Convex Programming: Theory and Applications*. Philadelphia, PA, USA: SIAM, 1994.

[35] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA, USA: SIAM, 1994.

[36] S. Liu, "Matrix results on the Khatri-Rao and Tracy-Singh products," *Linear Algebra Appl.*, vol. 289, no. 1–3, pp. 267–277, Mar. 1999.

[37] 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.

[38] A. D. Dabbagh and D. J. Love, "Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch," *IEEE Trans. Wireless Commun.*, vol. 56, no. 11, pp. 1859–1868, Nov. 2008.

Jiaxin Yang (S'11) received the B.Eng. degree in information engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and the M.E.Sc. degree in electrical and computer engineering from the University of Western Ontario, London, ON, Canada, in 2011. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC.

His research interests include optimization theory, statistical signal processing, detection and estimation, and the applications thereof in wireless communications such as multiple-input–multiple-output systems, cooperative communications, and physical-layer security.

Mr. Yang has received several award and scholarships, including the McGill Engineering Doctoral Award, the Graduate Excellence Fellowship, the International Differential Tuition Fee Waivers, the FQRNT International Internship Scholarship, and the PERSWADE Ph.D. Scholarship.

Benoit Champagne (S'87–M'89–SM'03) received the B.Ing. degree in engineering physics from École Polytechnique de Montréal, Montreal, QC, Canada, in 1983; the M.Sc. degree in physics from Université de Montréal, in 1985; and the Ph.D. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1990.

From 1990 to 1999, he was an Assistant and then Associate Professor with INRS-Telecommunications, Université du Quebec, Montréal. Since 1999, he has been with McGill University, Montreal, where he is currently a Full Professor with the Department of Electrical and Computer Engineering. From 2004 to 2007, he also served as an Associate Chair of Graduate Studies with the Department of Electrical and Computer Engineering, McGill University. He is the author or coauthor of more than 200 referred publications. His research has been funded by the Natural Sciences and Engineering Research Council of Canada, the "Fonds de Recherche sur la Nature et les Technologies" from the Government of Quebec, and some major industrial sponsors, including Nortel Networks, Bell Canada, InterDigital, and Microsemi. His research interests include the study of advanced algorithms for the processing of information bearing signals by digital means and several topics in statistical signal processing, such as detection and estimation, sensor array processing, adaptive filtering, and applications thereof to broadband communications and audio processing.

Dr. Champagne has also served on the Technical Committees of several international conferences in the fields of communications and signal processing. In particular, he served as the Cochair of the Wide Area Cellular Communications Track for the IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (Toronto, September 2011), as the Cochair of the Antenna and Propagation Track for the IEEE Vehicular Technology Conference-Fall, (Los Angeles, CA, USA, September 2004), and as the Registration Chair for the IEEE International Conference on Acoustics, Speech and Signal Processing (Montreal, May 2004). He has served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the *EURASIP Journal on Applied Signal Processing*.

Yulong Zou (M'12–SM'13) received the B.Eng. degree in information engineering and the Ph.D. degree signal and information processing from Nanjing University of Posts and Telecommunications (NUPT), Nanjing, China, in July 2006 and July 2012, respectively, and the Ph.D. degree in electrical engineering from Stevens Institute of Technology, Hoboken, NJ, USA, in May 2012.

He is currently a Full Professor and a Doctoral Supervisor with NUPT. His research interests include wireless communications and signal processing, cooperative communications, cognitive radio, wireless security, and energy-efficient communications.

Dr. Zou has acted as a Symposium Chair, a Session Chair, and a Technical Program Committee member for a number of IEEE sponsored conferences, including the IEEE Wireless Communications and Networking Conference, IEEE Global Communications Conference, IEEE International Conference on Communications, IEEE Vehicular Technology Conference, and the International Conference on Communications in China. He served as the Lead Guest Editor for a special issue on security challenges and issues in cognitive radio networks of the *EURASIP Journal on Advances in Signal Processing*. He currently serves as an Editor for the IEEE COMMUNICATIONS SURVEYS & TUTORIALS, IEEE COMMUNICATIONS LETTERS, the *EURASIP Journal on Advances in Signal Processing*, and the *KSII Transactions on Internet and Information Systems*. He also served as the Lead Guest Editor for a special issue on security and reliability challenges in industrial wireless sensor networks of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.

Lajos Hanzo (M'91–SM'92–F'04) received the M.S. degree in electronics and the Ph.D. degree from Budapest University of Technology and Economics (formerly, Technical University of Budapest), Budapest, Hungary, in 1976 and 1983, respectively; the D.Sc. degree from the University of Southampton, Southampton, U.K., in 2004; and the Doctor Honoris Causa degree from Budapest University of Technology and Economics in 2009.

During his 38-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, where he holds the Chair in Telecommunications. He is currently directing a 100-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council of U.K., the European Research Council's Advanced Fellow Grant, and the Royal Society Wolfson Research Merit Award. During 2008–2012, he was a Chaired Professor with Tsinghua University, Beijing, China. He is an enthusiastic supporter of industrial and academic liaison and offers a range of industrial courses. He has successfully supervised more than 80 Ph.D. students, coauthored 20 John Wiley/IEEE Press books on mobile radio communications, totaling in excess of 10 000 pages, and has published more than 1400 research entries on IEEE Xplore. He has more than 20 000 citations. His research is funded by the European Research Council's Senior Research Fellow Grant.

Dr. Hanzo is a Fellow of the Royal Academy of Engineering, The Institution of Engineering and Technology, and the European Association for Signal Processing. He is also a Governor of the IEEE Vehicular Technology Society. He has served as the Technical Program Committee Chair and the General Chair of IEEE conferences, has presented keynote lectures, and has received a number of distinctions. During 2008–2012, he was the Editor-in-Chief of the IEEE Press.

AUTHOR QUERY

NO QUERY.