A Framework of Robust Transmission Design for IRS-Aided MISO Communications With Imperfect Cascaded Channels

Gui Zhou, Cunhua Pan, Hong Ren, Kezhi Wang, and Arumugam Nallanathan, *Fellow, IEEE*

***Abstract*—Intelligent reflection surface (IRS) has recently been recognized as a promising technique to enhance the performance of wireless systems due to its ability of reconfiguring the signal propagation environment. However, the perfect channel state information (CSI) is challenging to obtain at the base station (BS) due to the lack of radio frequency (RF) chains at the IRS. Since most of the existing channel estimation methods were developed to acquire the cascaded BS-IRS-user channels, this paper is the first work to study the robust beamforming based on the imperfect cascaded BS-IRS-user channels at the transmitter (CBIUT). Specifically, the transmit power minimization problems are formulated subject to the worst-case rate constraints under the bounded CSI error model, and the rate outage probability constraints under the statistical CSI error model, respectively. After approximating the worst-case rate constraints by using the S-procedure and the rate outage probability constraints by using the Bernstein-type inequality, the reformulated problems can be efficiently solved. Numerical results show that the negative impact of the CBIUT error on the system performance is greater than that of the direct CSI error.**

***Index Terms*—Intelligent reflecting surface (IRS), reconfigurable intelligent surface (RIS), robust design, imperfect channel state information (CSI), cascaded BS-IRS-user channels.**

I. Introduction

INTELLIGENT reflecting surface (IRS), which is also known as reconfigurable intelligent surface (RIS) or large intelligent surface (LIS), has emerged as a promising technique to enhance the spectral and energy efficiency of the wireless networks [1]–[3], thanks to its artificial planar passive radio array structure which is cost-effective and energy-efficient. More explicitly, each passive element on the IRS is capable of reconfiguring the channels between the BS and users constructively or destructively by imposing an independent phase shift to the incident signal. The existing literature on IRS-aided wireless communications has demonstrated that IRS is an enabler for enhancing the spectral and energy efficiency through jointly optimizing the active beamforming at the BS and the passive beamforming at the IRS [4]–[11]. However, the algorithms developed in the above contributions were based on the assumption of perfect channel state information at the transmitter (CSIT).

Unfortunately, it is challenging to estimate the channels for the IRS-aided wireless systems, since IRS is passive and can neither send nor receive pilot symbols. In IRS-aided communication systems, there are two types of channels: the direct channel spanning from the BS to the user, and the IRS-related channels. The direct channel can be readily estimated by using conventional channel estimation methods such as the least square algorithm. Hence, most of the existing contributions focused on the channel estimation for the IRS-related channels, which are composed of the channel from the BS to the IRS (BS-IRS channel), and those from the IRS to the users (IRS-user channels).

In general, there are two main approaches to estimate the IRS-related channels. The first approach is to directly estimate the IRS-related channels, i.e., estimate BS-IRS channel and IRS-user channels separately [12]. Specifically, in [12], some active channel elements are installed at the IRS to estimate the individual channels. This method, however, has several drawbacks. The active elements may increase the hardware cost and consume extra power, which causes unaffordable burden on the IRS. In addition, the channel information estimated at the IRS needs to be fed back to the BS, which increases the information exchange overhead.

Fortunately, it is observed that the cascaded BS-IRS-user channels, which are the product of the BS-IRS channel and the IRS-user channels, are sufficient for the joint active and passive beamforming design [8]–[11]. As a result, most of the existing contributions focused on the second approach, i.e., the cascaded channel estimation [13]–[16]. Specifically, the channel estimation of the cascaded channel has been investigated both in the single-user multiple-input multiple-output (SU-MIMO) system [13] and the multi-user multiple-input single-output (MU-MISO) system [14]. However, the pilot overhead of the estimation methods in [13], [14] is prohibitively high, which scales up with the number of reflection elements. In order to reduce the pilot overhead, the authors in [15] exploited the sparse property of the channel matrix and proposed a channel estimation method based on compressed sensing technique. Furthermore, another sparsity representation of the cascaded channel has been found

Manuscript received February 17, 2020; revised July 19, 2020; accepted August 23, 2020. Date of publication August 28, 2020; date of current version September 15, 2020. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Athanasios A. Rontogiannis. This work was supported by Grant EP/R006466/1. *(Corresponding author: Cunhua Pan.)*

Gui Zhou, Cunhua Pan, Hong Ren, and Arumugam Nallanathan are with the School of Electronic Engineering, and Computer Science, Queen Mary University of London, London E1 4NS, U.K. (e-mail: g.zhou@qmul.ac.uk; c.pan@qmul.ac.uk; h.ren@qmul.ac.uk; a.nallanathan@qmul.ac.uk).

Kezhi Wang is with the Department of Computer, and Information Sciences, Northumbria University, Newcastle Upon Tyne NE1 8QH, U.K. (e-mail: kezhi.wang@northumbria.ac.uk).

Digital Object Identifier 10.1109/TSP.2020.3019666

in [16] by using the fact that the height of the BS and the IRS are often the same.

All the above-mentioned literature [4]–[11] did not consider the transmission design by taking into account the channel estimation error. Due to the inevitable channel estimation error, it will induce system performance loss if naively treating the estimated channels as perfect ones. Hence, it is imperative to design robust transmission strategies for the IRS-aided wireless communication systems. To the best of our knowledge, there are only a few contributions in this area [17], [18]. Specifically, in [17], we first proposed a worst-case robust design algorithm by assuming that the BS only knew the imperfect IRS-user channels in a MU-MISO wireless system. Then, the authors in [18] further proposed a robust secure transmission strategy by also applying the worst-case optimization method when the channels from the IRS to the eavesdroppers were imperfect. However, to implement the above robust design algorithms in [17] and [18], one should rely on the first channel estimation approach, where the BS-IRS channels and IRS-user channels should be independently estimated. This is difficult to achieve since several active elements should be installed at the IRS.

Against the above background, this paper studies the robust transmission design based on the imperfect cascaded BS-IRS-user channels at the transmitter (CBIUT). Specifically, we aim to design a robust active and passive beamforming scheme to minimize the total transmit power under both the bounded CSI error model and the statistical CSI error model. Unfortunately, the robust beamforming algorithms developed in [17] and [18] are not applicable for the imperfect CBIUT case. Hence, the contributions of this work are summarized as follows:

- To the best of our knowledge, this is the first work to study the robust transmission design based on imperfect cascaded BS-IRS-user channels, which is more practical than the previous works in which imperfect IRS-user channels were considered. In addition, we consider the robust transmission design under two channel error models: the bounded CSI error model and the statistical CSI error model. However, both [17] and [18] only considered the bounded CSI error model.
- For the bounded CSI error, we formulate worst-case robust beamforming design problems that minimize the transmit power subject to unit modulus of the reflection beamforming and the worst-case QoS constraints with imperfect CBIUT. The worst-case robust design can guarantee that the achievable rate of each user is no less than its minimum rate requirement for all possible channel error realizations. To address this non-convex problem, S-procedure is firstly adopted to approximate the semi-infinite inequality constraints. Then, under the alternate optimization (AO) framework, the precoder is updated in an second-order cone programming (SOCP) and the reflection beamforming is updated by using the penalty convex-concave procedure (CCP).
- For the statistical CSI error model, we aim to minimize the transmit power subject to unit-modulus constraints and the rate outage probability constraints. Here, the rate outage probability constraints represent the probability that the achievable rate of each user being below its minimum rate requirment needs to be less than a predetermined probability. By applying the Bernstein-Type Inequality, the safe approximation of the rate outage probability is obtained to make the original problem tractable. Then, the precoder and the reflection beamforming are optimized by using the semidefinite relaxation (SDR) and penalty CCP techniques respectively in an iterative manner.
- We demonstrate through numerical results that the robust beamforming under the statistical CSI error model can achieve superior system performance in terms of the minimum transmit power, convergence speed and complexity, than that under the bounded CSI error model. In addition, it is observed that the level of the CBIUT error plays an important role in the IRS-aided systems. Specifically, when the CBIUT error is small, the total transmit power decreases with the number of the reflection elements due to the increased beamforming gain. However, when the CBIUT error is large, the transmit power increases with the number of the reflection elements due to the increased channel estimation error. Hence, whether to deploy the IRS in wireless communication systems depends on the level of the CBIUT error.

The remainder of this paper is organized as follows. Section II introduces the system model and the CSI error models. Worst-case robust design problems are formulated and solved in Section III. Section IV further investigates the outage constrained robust design problems. Section V compares the computational complexity of the developed robust design methods. Finally, Section VI and Section VII show the numerical results and conclusions, respectively.

Notations: The following mathematical notations and symbols are used throughout this paper. Vectors and matrices are denoted by boldface lowercase letters and boldface uppercase letters, respectively. The symbols $\mathbf{X}^*$, $\mathbf{X}^{\mathrm{T}}$, $\mathbf{X}^{\mathrm{H}}$, and $||\mathbf{X}||_F$ denote the conjugate, transpose, Hermitian (conjugate transpose), Frobenius norm of matrix $\mathbf{X}$, respectively. The symbol $||\mathbf{x}||_2$ denotes 2-norm of vector $\mathbf{x}$. The symbols $\mathrm{Tr}\{\cdot\}, \mathrm{Re}\{\cdot\}, |\cdot|, \lambda(\cdot)$, and $\angle(\cdot)$ denote the trace, real part, modulus, eigenvalue, and angle of a complex number, respectively. $\mathrm{diag}(\mathbf{x})$ is a diagonal matrix with the entries of $\mathbf{x}$ on its main diagonal. $[\mathbf{x}]_m$ means the m^{th} element of the vector $\mathbf{x}$. The Kronecker product between two matrices $\mathbf{X}$ and $\mathbf{Y}$ is denoted by $\mathbf{X} \otimes \mathbf{Y}$. $\mathbf{X} \succeq \mathbf{Y}$means that $\mathbf{X} - \mathbf{Y}$ is positive semidefinite. Additionally, the symbol $\mathbb{C}$ denotes complex field, $\mathbb{R}$ represents real field, and $j \triangleq \sqrt{-1}$ is the imaginary unit.

II. System Model

In this section, we first introduce the system model of the IRS-aided MISO downlink communication system, and then discuss the channel uncertainty scenarios as well as the CSI error models.

A. Signal Transmission Model

As shown in Fig. 1, we consider an IRS-aided MISO broadcast (BC) communication system, which consists of one multi-antenna BS, K single-antenna users and one IRS. It is assumed

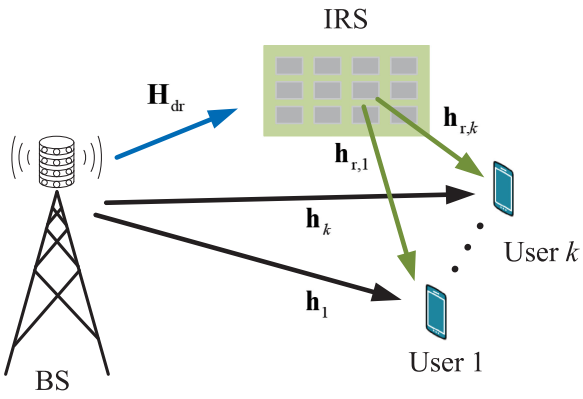


Fig. 1. An IRS-aided multi-user communication system.

that the BS is equipped with N active antennas, and transmits K Gaussian data symbols denoted by $\mathbf{s} = [s_1, \ldots, s_K]^{\rm T} \in \mathbb{C}^{K\times 1}$ to all the users, where $\mathbb{E}[\mathbf{s}\mathbf{s}^{\rm H}] = \mathbf{I}$. IRS with M programmable phase shifters is deployed to enhance the system performance. Therefore, by defining the set of users as $\mathcal{K} = \{1, 2, \ldots, K\}$, the received baseband signal of users is given by

$$y_k = (\mathbf{h}_k^{\rm H} + \mathbf{h}_{{\rm r},k}^{\rm H}\mathbf{E}\mathbf{H}_{\rm dr})\mathbf{F}\mathbf{s} + n_k, \forall k \in \mathcal{K}. \tag{1}$$

Here, $\mathbf{F} = [\mathbf{f}_1, \ldots, \mathbf{f}_K] \in \mathbb{C}^{N\times K}$ is the precoder matrix, in which $\mathbf{f}_k$ is the precoding vector associated with user k. Then, the transmit power at the BS is $\mathbb{E}\{{\rm Tr}[\mathbf{F}\mathbf{s}\mathbf{s}^{\rm H}\mathbf{F}^{\rm H}]\} = ||\mathbf{F}||_F^2$. n_k is the additive white Gaussian noise (AWGN) at user k, with zero mean and noise variance σ_k^2, i.e., $n_k \sim \mathcal{CN}(0, \sigma_k^2)$. The reflection beamforming of the IRS is a diagonal matrix $\mathbf{E} = \sqrt{\iota}{\rm diag}(e_1, \ldots, e_M) \in \mathbb{C}^{M\times M}$, of which has unit-modulus phase shifts, i.e., $|e_m|^2 = 1$. $0 \le \iota \le 1$ indicates the reflection efficiency and the power loss of reflection operation usually comes from multiple reflections of signals. Here, we assume that only the first-order reflection on the IRS is considered and set $\iota = 1$. It is assumed that the phase shifts of the IRS are calculated by the BS and then fed back to the IRS controller through dedicated feedback channels [4], [5]. In addition, the channel vectors spanning from the BS to user k and from the IRS to user k are denoted by $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$ and $\mathbf{h}_{{\rm r},k} \in \mathbb{C}^{M\times 1}$, respectively. The channel matrix between the BS and the IRS is represented by $\mathbf{H}_{\rm dr} \in \mathbb{C}^{M\times N}$.

Denote by $\mathbf{G}_k = {\rm diag}(\mathbf{h}_{{\rm r},k}^{\rm H})\mathbf{H}_{\rm dr}$ the cascaded channel from the BS to user k via the IRS, by $\mathbf{e} = [e_1, \ldots, e_M]^{\rm T} \in \mathbb{C}^{M\times 1}$ the vector containing diagonal elements of matrix $\mathbf{E}$, and by $\beta_k = ||(\mathbf{h}_k^{\rm H} + \mathbf{e}^{\rm H}\mathbf{G}_k)\mathbf{F}_{-k}||_2^2 + \sigma_k^2$ the interference-plus-noises (INs) power of user k, where $\mathbf{F}_{-k} = [\mathbf{f}_1, \ldots, \mathbf{f}_{k-1}, \mathbf{f}_{k+1}, \ldots, \mathbf{f}_K]$. Then, the achievable data rate (bit/s/Hz) at user k is given by

$$\mathcal{R}_k(\mathbf{F}, \mathbf{e}) = \log_2\left(1 + \frac{1}{\beta_k}\left|\left(\mathbf{h}_k^{\rm H} + \mathbf{e}^{\rm H}\mathbf{G}_k\right)\mathbf{f}_k\right|^2\right). \tag{2}$$

B. Two Scenarios and CSI Error Models

In the IRS-aided communication system, there are two types of channels: the direct channel $\mathbf{h}_k$, and the cascaded BS-IRS-user channel $\mathbf{G}_k$. The system performance of the IRS-aided communication system is highly affected by the accuracy of the direct channel state information at the transmitter (DCSIT) and the CBIUT. In the following, we first introduce two scenarios of the channel uncertainties and then two types of CSI error models.

1) Scenario 1: Partial Channel Uncertainty (PCU): In IRS-aided communications, the CBIUT is much more challenging to obtain than the DCSIT due to the passive features of the IRS. Hence, in this scenario, we assume that the DCSIT is perfect, while the CBIUT is imperfect. The CBIUT can be represented as

$$\mathbf{G}_k = \widehat{\mathbf{G}}_k + \triangle\mathbf{G}_k, \forall k \in \mathcal{K}, \tag{3}$$

where $\widehat{\mathbf{G}}_k$ is the estimated cascaded CSI known at the BS, $\triangle\mathbf{G}_k$ is the unknown CBIUT error.

2) Scenario 2: Full Channel Uncertainty (FCU): In complex electromagnetic environment, the accurate DCSIT is also challenging to obtain. In this scenario, we assume both the DCSIT and the CBIUT are imperfect. In addition to the CBIUT error model in (3), the direct channel is expressed as

$$\mathbf{h}_k = \widehat{\mathbf{h}}_k + \triangle\mathbf{h}_k, \forall k \in \mathcal{K}, \tag{4}$$

where $\widehat{\mathbf{h}}_k$ is the estimated DCSIT known at the BS and $\triangle\mathbf{h}_k$ is the unknown DCSIT error.

In this work, we investigate two types of robust beamforming design for IRS-aided MISO communication systems depending on the CSI error models.

3) Error Model 1: Bounded CSI Error Model: Specifically, one is the worst-case robust beamforming design subject to the bounded CSI error model, i.e.,

$$\|\triangle\mathbf{G}_k\|_F \le \xi_{{\rm g},k}, \|\triangle\mathbf{h}_k\|_2 \le \xi_{{\rm h},k}, \forall k \in \mathcal{K}, \tag{5}$$

where $\xi_{{\rm g},k}$ and $\xi_{{\rm h},k}$ are the radii of the uncertainty regions known at the BS. This CSI error model characterizes the channel quantization error which naturally belongs to a bounded region [19]. For example, in the frequency division duplex (FDD) setting, the receiver estimates the downlink channel and then feeds the rate-limited quantized CSI back to the transmitter. Then, the acquired CSI is plagued by quantization errors.

4) Error Model 2: Statistical CSI Error Model: The other is the outage-constrained robust beamforming design associated with the statistical CSI error model, in which each CSI error vector is assumed to follow the circularly symmetric complex Gaussian (CSCG) distribution, i.e.,

$${\rm vec}(\triangle\mathbf{G}_k) \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_{{\rm g},k}), \mathbf{\Sigma}_{{\rm g},k} \succeq \mathbf{0}, \forall k \in \mathcal{K}, \tag{6a}$$

$$\triangle\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_{{\rm h},k}), \mathbf{\Sigma}_{{\rm h},k} \succeq \mathbf{0}, \forall k \in \mathcal{K}, \tag{6b}$$

where $\mathbf{\Sigma}_{{\rm g},k} \in \mathbb{C}^{MN\times MN}$ and $\mathbf{\Sigma}_{{\rm h},k} \in \mathbb{C}^{N\times N}$ are positive semidefinite error covariance matrices. In this case, the CSI imperfection is caused by the channel estimation error [20]. For example, in the time division duplex (TDD) setting, noise and limited training will cause the uplink channel estimation error. The conventional MMSE method is generally adopted to estimate the cascaded channel, and thus the channel estimation generally follows the CSCG distribution.

In the following, we first consider the first type of robust beamforming design based on the bounded CSI error model. Then, we deal with the second one based on the statistical CSI error model.

III. Worst-Case Robust Beamforming Design

In this section, the worst-case robust beamforming design is considered under the bounded CSI error model. We aim to minimize the total transmit power of the BS by the joint design of the precoder matrix $\mathbf{F}$ and reflection beamforming vector $\mathbf{e}$ under the unit-modulus constraints and the worst-case QoS constraints, i.e., ensuring the achievable rate of each user to be above a threshold for all possible channel error realizations. In order to solve the non-convex robust design problem with semi-infinite inequality constraints and coupled variables, an AO algorithm is proposed based on S-Procedure, SOCP and penalty CCP [21].

First, two useful lemmas about multiple complex valued uncertainties are formally introduced as follows, which will be used in the later derivations.

Lemma 1: (General S-Procedure [22]) Define the quadratic functions of the variable $\mathbf{x} \in \mathbb{C}^{n\times 1}$:

$$f_i(\mathbf{x}) = \mathbf{x}^{\mathrm{H}}\mathbf{W}_i\mathbf{x} + 2\mathrm{Re}\left\{\mathbf{w}_i^{\mathrm{H}}\mathbf{x}\right\} + w_i,\ i = 0, \ldots, P,$$

where $\mathbf{W}_i = \mathbf{W}_i^{\mathrm{H}}$. The condition $\{f_i(\mathbf{x}) \geq 0\}_{i=1}^{P} \Rightarrow f_0(\mathbf{x}) \geq 0$ holds if and only if there exist $\forall i, \varpi_i \geq 0$ such that

$$\begin{bmatrix} \mathbf{W}_0 & \mathbf{w}_0 \\ \mathbf{w}_0^{\mathrm{H}} & w_0 \end{bmatrix} - \sum_{i=1}^{P} \varpi_i \begin{bmatrix} \mathbf{W}_i & \mathbf{w}_i \\ \mathbf{w}_i^{\mathrm{H}} & w_i \end{bmatrix} \succeq \mathbf{0}.$$

Lemma 2: (General sign-definiteness [23]) For a given set of matrices $\mathbf{W} = \mathbf{W}^{\mathrm{H}}$, $\{\mathbf{Y}_i, \mathbf{Z}_i\}_{i=1}^{P}$, the following linear matrix inequality (LMI) satisfies

$$\mathbf{W} \succeq \sum_{i=1}^{P} \left(\mathbf{Y}_i^{\mathrm{H}}\mathbf{X}_i\mathbf{Z}_i + \mathbf{Z}_i^{\mathrm{H}}\mathbf{X}_i^{\mathrm{H}}\mathbf{Y}_i\right), \forall i, ||\mathbf{X}_i||_F \leq \xi_i,$$

if and only if there exist real numbers $\forall i, \mu_i \geq 0$ such that

$$\begin{bmatrix} \mathbf{W} - \sum_{i=1}^{P} \mu_i \mathbf{Z}_i^{\mathrm{H}}\mathbf{Z}_i & -\xi_1\mathbf{Y}_1^{\mathrm{H}} & \cdots & -\xi_P\mathbf{Y}_P^{\mathrm{H}} \\ -\xi_1\mathbf{Y}_1 & \mu_1\mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ -\xi_P\mathbf{Y}_P & \mathbf{0} & \cdots & \mu_P\mathbf{I} \end{bmatrix} \succeq \mathbf{0}.$$

It is noted that Lemma 2 can be proved by applying Lemma 1 and the detailed proof is given in [24].

A. Scenario 1: Partial Channel Uncertainty

In this subsection, we design the robust beamforming for the IRS-aided communication system under Scenario 1 with perfect DCSIT and imperfect CBIUT. This problem is simpler than the one with full channel uncertainty and the algorithm developed for Scenario 1 has lower complexity than that for Scenario 2. Mathematically, let $\mathcal{E}_k^{partial} \triangleq \{\forall \| \triangle \mathbf{G}_k\|_F \leq \xi_{\mathrm{g},k}\}$ and denote by $\mathcal{M} = \{1, 2, \ldots, M\}$ the set of reflection elements, the worst-case transmit power minimization problem is formulated as

$$\min_{\mathbf{F},\mathbf{e}} \ ||\mathbf{F}||_F^2 \tag{7a}$$

$$\text{s.t. } \mathcal{R}_k(\mathbf{F}, \mathbf{e}) \geq R_k, \mathcal{E}_k^{partial}, \forall k \in \mathcal{K} \tag{7b}$$

$$|e_m|^2 = 1, \forall m \in \mathcal{M}. \tag{7c}$$

Here, R_k is the target rate of user k. Constraints (7b) are the worst-case QoS requirements for the users, while constraints (7c) correspond to the unit-modulus requirements of the reflection elements at the IRS.

To start with, the non-convexity of constraints (7b) can be addressed by firstly treating the INs power $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_K]^{\mathrm{T}}$ as auxiliary variables. Hence, constraints (7b) are reformulated as

$$\left|\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k\right)\mathbf{f}_k\right|^2 \geq \beta_k(2^{R_k} - 1), \mathcal{E}_k^{partial}, \forall k \in \mathcal{K}, \tag{8}$$

$$\left\|\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k\right)\mathbf{F}_{-k}\right\|_2^2 + \sigma_k^2 \leq \beta_k, \mathcal{E}_k^{partial}, \forall k \in \mathcal{K}. \tag{9}$$

Constraints (8) and (9) are termed as the worst-case useful signal power constraints and the worst-case INs power constraints, respectively.

Then, the non-convex semi-infinite inequality constraints (8) are handled by firstly approximating the non-convex parts and then dealing with the semi-infinite inequalities by using the S-Procedure. Specifically, the following lemma shows the linear approximation of the useful signal power in (8).

Lemma 3: Substituting $\mathbf{G}_k = \widehat{\mathbf{G}}_k + \triangle\mathbf{G}_k$ into the useful signal power in (8) and let $\mathbf{f}_k^{(n)}$ and $\mathbf{e}^{(n)}$ be the optimal solutions obtained at iteration n, then $|[\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}(\widehat{\mathbf{G}}_k + \triangle\mathbf{G}_k)]\mathbf{f}_k|^2$ is linearly approximated by its lower bound at $(\mathbf{f}_k^{(n)}, \mathbf{e}^{(n)})$ as follows

$$\mathrm{vec}^{\mathrm{T}}(\triangle\mathbf{G}_k)\mathbf{A}_k\mathrm{vec}(\triangle\mathbf{G}_k^*) + 2\mathrm{Re}\left\{\mathbf{a}_k^{\mathrm{T}}\mathrm{vec}(\triangle\mathbf{G}_k^*)\right\} + a_k, \tag{10}$$

where

$$\begin{aligned}
\mathbf{A}_k &= \mathbf{f}_k\mathbf{f}_k^{(n),\mathrm{H}} \otimes \mathbf{e}^*\mathbf{e}^{(n),\mathrm{T}} + \mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}} \otimes \mathbf{e}^{(n),*}\mathbf{e}^{\mathrm{T}} \\
&\quad - (\mathbf{f}_k^{(n)}\mathbf{f}_k^{(n),\mathrm{H}} \otimes \mathbf{e}^{(n),*}\mathbf{e}^{(n),\mathrm{T}}), \\
\mathbf{a}_k &= \mathrm{vec}(\mathbf{e}\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}) \\
&\quad + \mathrm{vec}(\mathbf{e}^{(n)}\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{f}_k\mathbf{f}_k^{(n),\mathrm{H}}) \\
&\quad - \mathrm{vec}(\mathbf{e}^{(n)}\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{f}_k^{(n)}\mathbf{f}_k^{(n),\mathrm{H}}), \\
a_k &= 2\mathrm{Re}\left\{\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}\left(\mathbf{h}_k + \widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}\right)\right\} \\
&\quad - \left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{f}_k^{(n)}\mathbf{f}_k^{(n),\mathrm{H}}\left(\mathbf{h}_k + \widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}^{(n)}\right).
\end{aligned}$$

Proof: Please refer to Appendix A. ■

By replacing the useful signal power in (8) with its linear approximation (10), constraints (8) are reformulated as

$$\begin{aligned}
&\mathrm{vec}^{\mathrm{T}}(\triangle\mathbf{G}_k)\mathbf{A}_k\mathrm{vec}(\triangle\mathbf{G}_k^*) + 2\mathrm{Re}\left\{\mathbf{a}_k^{\mathrm{T}}\mathrm{vec}(\triangle\mathbf{G}_k^*)\right\} + a_k \\
&\geq \beta_k(2^{R_k} - 1), \mathcal{E}_k^{partial}, \forall k \in \mathcal{K}.
\end{aligned} \tag{11}$$

Lemma 1 is then used to tackle the CSI uncertainty in the above constraints. Specifically, constraint corresponding to each user in (11) can be recast by setting the parameters in Lemma 1 as follows

$$\begin{aligned}
&P = 1,\ \mathbf{W}_0 = \mathbf{A}_k,\ \mathbf{w}_0 = \mathbf{a}_k,\ w_0 = a_k - \beta_k(2^{R_k} - 1), \\
&\mathbf{x} = \mathrm{vec}(\triangle\mathbf{G}_k^*), \qquad \mathbf{W}_1 = -\mathbf{I},\ w_1 = \xi_k^2.
\end{aligned}$$

Then, (11) is transformed into the following equivalent LMIs as

$$\begin{bmatrix} \varpi_{\mathrm{g},k}\mathbf{I}_{MN}+\mathbf{A}_k & \mathbf{a}_k \\ \mathbf{a}_k^{\mathrm{T}} & C_k^{partial} \end{bmatrix} \succeq \mathbf{0}, \forall k \in \mathcal{K}, \tag{12}$$

where $\boldsymbol{\varpi}_{\mathrm{g}} = [\varpi_{\mathrm{g},1}, \ldots, \varpi_{\mathrm{g},K}]^{\mathrm{T}} \geq 0$ are slack variables and $C_k^{partial} = a_k - \beta_k(2^{R_k}-1) - \varpi_{\mathrm{g},k}\xi_k^2$.

Next, we consider the uncertainty in $\{\triangle\mathbf{G}_k\}_{\forall k\in\mathcal{K}}$ of (9). Specifically, we firstly adopt Schur's complement Lemma [25] to equivalently recast the INs power inequalities in (9) into matrix inequalities as follows

$$\begin{bmatrix} \beta_k - \sigma_k^2 & \mathbf{t}_k^{\mathrm{H}} \\ \mathbf{t}_k & \mathbf{I} \end{bmatrix} \succeq \mathbf{0}, \forall k \in \mathcal{K}, \tag{13}$$

where $\mathbf{t}_k = ((\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k)\mathbf{F}_{-k})^{\mathrm{H}}$. By using $\mathbf{G}_k = \widehat{\mathbf{G}}_k + \triangle\mathbf{G}_k$, (13) is then rewritten as

$$\begin{aligned} \begin{bmatrix} \beta_k - \sigma_k^2 & \widehat{\mathbf{t}}_k^{\mathrm{H}} \\ \widehat{\mathbf{t}}_k & \mathbf{I} \end{bmatrix} \succeq & -\begin{bmatrix} \mathbf{0} \\ \mathbf{F}_{-k}^{\mathrm{H}} \end{bmatrix} \triangle\mathbf{G}_k^{\mathrm{H}} \begin{bmatrix} \mathbf{e} & \mathbf{0} \end{bmatrix} \\ & -\begin{bmatrix} \mathbf{e}^{\mathrm{H}} \\ \mathbf{0} \end{bmatrix} \triangle\mathbf{G}_k \begin{bmatrix} \mathbf{0} & \mathbf{F}_{-k} \end{bmatrix}, \forall k \in \mathcal{K}, \end{aligned} \tag{14}$$

where $\widehat{\mathbf{t}}_k = ((\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{F}_{-k})^{\mathrm{H}}$.

In order to use Lemma 2, we choose the following parameters (It is noted that the subscript i in Lemma 2 has been ignored since $P = 1$.) for each constraint in (14) as

$$\mathbf{W} = \begin{bmatrix} \beta_k - \sigma_k^2 & \widehat{\mathbf{t}}_k^{\mathrm{H}} \\ \widehat{\mathbf{t}}_k & \mathbf{I} \end{bmatrix}, \mathbf{Y} = -\begin{bmatrix} \mathbf{0} & \mathbf{F}_{-k} \end{bmatrix},$$
$$\mathbf{Z} = \begin{bmatrix} \mathbf{e} & \mathbf{0} \end{bmatrix}, \qquad \mathbf{X} = \triangle\mathbf{G}_k^{\mathrm{H}}.$$

Then, the equivalent LMIs of the worst-case INs power constraints (9) are given by

$$\begin{bmatrix} \beta_k - \sigma_k^2 - \mu_{\mathrm{g},k}M & \widehat{\mathbf{t}}_k^{\mathrm{H}} & \mathbf{0}_{1\times N} \\ \widehat{\mathbf{t}}_k & \mathbf{I}_{(K-1)} & \xi_{\mathrm{g},k}\mathbf{F}_{-k}^{\mathrm{H}} \\ \mathbf{0}_{N\times 1} & \xi_{\mathrm{g},k}\mathbf{F}_{-k} & \mu_{\mathrm{g},k}\mathbf{I}_N \end{bmatrix} \succeq \mathbf{0}, \forall k \in \mathcal{K}, \tag{15}$$

where $\boldsymbol{\mu}_{\mathrm{g}} = [\mu_{\mathrm{g},1}, \ldots, \mu_{\mathrm{g},K}]^{\mathrm{T}} \geq 0$ are slack variables.

Based on the above discussions, Problem (7) is approximately rewritten as

$$\min_{\mathbf{F},\mathbf{e},\boldsymbol{\beta},\boldsymbol{\varpi}_{\mathrm{g}},\boldsymbol{\mu}_{\mathrm{g}}} \quad ||\mathbf{F}||_F^2 \tag{16a}$$
$$\text{s.t.} \quad (12), (15), (7c), \tag{16b}$$
$$\boldsymbol{\varpi}_{\mathrm{g}} \geq 0, \boldsymbol{\mu}_{\mathrm{g}} \geq 0. \tag{16c}$$

This problem is still non-convex and difficult to optimize $\mathbf{F}$ and $\mathbf{e}$ simultaneously since $\mathbf{F}$ and $\mathbf{e}$ are coupled in $\mathbf{A}_k$, $\mathbf{a}_k$ and $\widehat{\mathbf{t}}_k$. In the following, we adopt the AO method to optimize $\mathbf{F}$ and $\mathbf{e}$ sequentially in an iterative manner. In particular, we minimize the transmit power by first fixing the reflection beamforming $\mathbf{e}$ so that the problem reduces to a convex one with respect to $\mathbf{F}$. CVX tool [26] is adopted to solve the resulting convex problem. Precoder $\mathbf{F}$ is then fixed and the resulting non-convex problem of $\mathbf{e}$ is handled under the penalty CCP method. Specifically, for given $\mathbf{e}$, the subproblem of $\mathbf{F}$ is given by

$$\mathbf{F}^{(n+1)} = \arg \min_{\mathbf{F},\boldsymbol{\beta},\boldsymbol{\varpi}_{\mathrm{g}},\boldsymbol{\mu}_{\mathrm{g}}} \quad ||\mathbf{F}||_F^2 \tag{17a}$$
$$\text{s.t.} \quad (12), (15), (16c), \tag{17b}$$

where $\mathbf{F}^{(n+1)}$ is the optimal solution obtained in the $(n+1)$-th iteration. Problem (17) is a semidefinite program (SDP) and can be solved by the CVX tool.

Then, for given $\mathbf{F}$, the subproblem of $\mathbf{e}$ is a feasibility-check problem. According to [17], [27] and in order to improve the converged solution in the optimization of $\mathbf{e}$, the useful signal power inequalities in (8) are modified by introducing slack variables $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_K]^{\mathrm{T}} \geq 0$ and recast as

$$\left|\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k\right)\mathbf{f}_k\right|^2 \geq \beta_k(2^{R_k}-1) + \alpha_k, \forall k \in \mathcal{K}. \tag{18}$$

Subsequently, the LMIs (12) are modified as

$$\begin{bmatrix} \varpi_{\mathrm{g},k}\mathbf{I}_{MN}+\mathbf{A}_k & \mathbf{a}_k \\ \mathbf{a}_k^{\mathrm{T}} & C_k^{partial} - \alpha_k \end{bmatrix} \succeq \mathbf{0}, \forall k \in \mathcal{K}. \tag{19}$$

In addition, we note that only the submatrix of $K \times K$ in the upper left corner of (15) depends on $\mathbf{e}$, so the dimension of the LMIs (15) can be reduced from $(K+N)\times(K+N)$ to $K \times K$ as

$$\begin{bmatrix} \beta_k - \sigma_k^2 - \mu_{\mathrm{g},k}M & \widehat{\mathbf{t}}_k^{\mathrm{H}} \\ \widehat{\mathbf{t}}_k & \mathbf{I}_{(K-1)} \end{bmatrix} \succeq \mathbf{0}, \forall k \in \mathcal{K}. \tag{20}$$

Combining (19) and (20), the sub-problem of $\mathbf{e}$ can be formulated as

$$\max_{\boldsymbol{\alpha},\mathbf{e},\boldsymbol{\beta},\boldsymbol{\varpi}_{\mathrm{g}},\boldsymbol{\mu}_{\mathrm{g}}} \quad \sum_{k=1}^{K}\alpha_k \tag{21a}$$
$$\text{s.t.} \quad (19), (20), (7c), (16c), \tag{21b}$$
$$\boldsymbol{\alpha} \geq 0. \tag{21c}$$

Note that the solution of Problem (21) can yield a lower objective value compared with Problem (17), the explanation of which can be found in [27].

We note that the above problem is still non-convex due to the unit-modulus constraints. As in our previous work [17], we here adopt the penalty CCP [21] to deal with the non-convex constraints. Following the penalty CCP framwork, the constraints (7c) are firstly equivalently rewritten as $1 \leq |e_m|^2 \leq 1, \forall m \in \mathcal{M}$. The non-convex parts of the resulting constraints are then linearized by $|e_m^{[t]}|^2 - 2\mathrm{Re}(e_m^* e_m^{[t]}) \leq -1, \forall m \in \mathcal{M}$, at fixed $e_m^{[t]}$. We finally have the following convex subproblem of $\mathbf{e}$ as

$$\max_{\substack{\mathbf{e},\boldsymbol{\alpha},\mathbf{b},\\ \boldsymbol{\beta},\boldsymbol{\varpi}_{\mathrm{g}},\boldsymbol{\mu}_{\mathrm{g}}}} \quad \sum_{k=1}^{K}\alpha_k - \lambda^{[t]}\sum_{m=1}^{2M} b_m \tag{22a}$$
$$\text{s.t.} \quad (19), (20), (16c), (21c), \tag{22b}$$
$$|e_m^{[t]}|^2 - 2\mathrm{Re}(e_m^* e_m^{[t]}) \leq b_m - 1, \forall m \in \mathcal{M} \tag{22c}$$
$$|e_m|^2 \leq 1 + b_{M+m}, \forall m \in \mathcal{M} \tag{22d}$$
$$\mathbf{b} \geq 0, \tag{22e}$$

Algorithm 1: Penalty CCP Optimization for Reflection Beamforming Optimization.

Initialize: Initialize $\mathbf{e}^{[0]}$, $\gamma^{[0]} > 1$, and set $t = 0$.
repeat
if $t < T_{max}$ **then**
  Update $\mathbf{e}^{[t+1]}$ from Problem (22);
  $\lambda^{[t+1]} = \min\{\gamma\lambda^{[t]}, \lambda_{max}\}$;
  $t = t + 1$;
else
  Initialize with a new random $\mathbf{e}^{[0]}$, set up $\gamma^{[0]} > 1$ again, and set $t = 0$.
end if
until $||\mathbf{b}||_1 \leq \chi$ and $||\mathbf{e}^{[t]} - \mathbf{e}^{[t-1]}||_1 \leq \nu$.
Output $\mathbf{e}^{(n+1)} = \mathbf{e}^{[t]}$.

where $\mathbf{b} = [b_1, \ldots, b_{2\,M}]^{\mathrm{T}}$ are slack variables imposed over the equivalent linear constraints of the unit-modulus constraints, and $||\mathbf{b}||_1$ is the penalty term in the objective function. $||\mathbf{b}||_1$ is scaled by the regularization factor $\lambda^{[t]}$ to control the feasibility of the constraints.

Problem (22) is an SDP and can be solved by the CVX tool. The steps of finding a feasible solution of $\mathbf{e}$ to Problem (21) is summarized in Algorithm 1. We remark that: *a)* When χ is sufficiently low, constraints (7c) in the original Problem (21) is guaranteed by $||\mathbf{b}||_1 \leq \chi$; *b)* The maximum value λ_{max} is imposed to avoid a numerical problem, that is, a feasible solution satisfying $||\mathbf{b}||_1 \leq \chi$ may not be found when the iteration converges to the stopping criteria $||\mathbf{e}^{[t]} - \mathbf{e}^{[t-1]}||_1 \leq \nu$ with the increase of $\lambda^{[t]}$; *c)* Stopping criteria $||\mathbf{e}^{[t]} - \mathbf{e}^{[t-1]}||_1 \leq \nu$ controls the convergence of Algorithm 1; *d)* As mentioned in [21], a feasible solution to Problem (22) may not be feasible for Problem (21). Hence, the feasibility of Problem (21) is guaranteed by imposing a maximum number of iterations T_{max} and, in case it is reached, we restart the iteration based on a new initial point.

Finally, under the AO framework, Problem (16) is solved by solving Problems (17) and (21) in an iterative manner. We remark that the fixed point $\mathbf{e}^{[t]}$ in constraint (22c) is updated iteratively in Algorithm 1, which is the same as $\lambda^{[t]}$. While fixed point $\mathbf{e}^{(n)}$ in constraint (19) is updated iteratively in the outer AO framework.

B. Scenario 2: Full Channel Uncertainty

In this subsection, we extend the robust beamforming design in the previous subsection to the case where both the DCSIT and CBIUT are imperfect. By considering the full channel uncertainty in (3) and (4) and denoting $\mathcal{E}_k^{full} \triangleq \{\forall|| \triangle \mathbf{h}_k||_2 \leq \xi_{\mathrm{h},k}, \forall|| \triangle \mathbf{G}_k||_F \leq \xi_{\mathrm{g},k}\}$, constraints (7b) can be extended to

$$\mathcal{R}_k\left(\mathbf{F}, \mathbf{e}\right) \geq R_k, \mathcal{E}_k^{full}, \forall k \in \mathcal{K}, \tag{23}$$

which is then equivalent to

$$\left|\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k\right)\mathbf{f}_k\right|^2 \geq \beta_k(2^{R_k} - 1), \mathcal{E}_k^{full}, \forall k \in \mathcal{K}, \tag{24}$$

$$\left\|\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k\right)\mathbf{F}_{-k}\right\|_2^2 + \sigma_k^2 \leq \beta_k, \mathcal{E}_k^{full}, \forall k \in \mathcal{K}. \tag{25}$$

The above non-convex semi-infinite inequality constraints can be addressed in the same way as Scenario 1. In particular, the linear approximation of the useful signal power in (24) is given in the following lemma.

Lemma 4: Let $\mathbf{f}_k^{(n)}$ and $\mathbf{e}^{(n)}$ be the optimal solutions obtained at iteration n, and by inserting $\mathbf{h}_k = \widehat{\mathbf{h}}_k + \triangle\mathbf{h}_k$ and $\mathbf{G}_k = \widehat{\mathbf{G}}_k + \triangle\mathbf{G}_k$ into the useful signal power in (24), then the resulting $|[(\widehat{\mathbf{h}}_k + \triangle\mathbf{h}_k)^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}(\widehat{\mathbf{G}}_k + \triangle\mathbf{G}_k)]\mathbf{f}_k|^2$ is lower bounded linearly at $(\mathbf{f}_k^{(n)}, \mathbf{e}^{(n)})$ as follows

$$\mathbf{x}_k^{\mathrm{H}}\widetilde{\mathbf{A}}_k\mathbf{x}_k + 2\mathrm{Re}\left\{\widetilde{\mathbf{a}}_k^{\mathrm{H}}\mathbf{x}_k\right\} + \widetilde{a}_k, \tag{26}$$

where

$$\widetilde{\mathbf{A}}_k = \mathbf{D}_k + \mathbf{D}_k^{\mathrm{H}} - \mathbf{Z}_k,$$

$$\mathbf{D}_k = \begin{bmatrix} \mathbf{f_k^{(n)}} \\ \mathbf{f}_k^{(n)} \otimes \mathbf{e}^{(\mathrm{n}),*} \end{bmatrix}\left[\mathbf{f}_k^{\mathrm{H}}\ \mathbf{f}_k^{\mathrm{H}} \otimes \mathbf{e}^{\mathrm{T}}\right],$$

$$\mathbf{Z}_k = \begin{bmatrix} \mathbf{f_k^{(n)}} \\ \mathbf{f}_k^{(n)} \otimes \mathbf{e}^{(\mathrm{n}),*} \end{bmatrix}\left[\mathbf{f}_k^{(n),\mathrm{H}}\ \mathbf{f}_k^{(n),\mathrm{H}} \otimes \mathbf{e}^{(n),\mathrm{T}}\right],$$

$$\widetilde{\mathbf{a}}_k = \mathbf{d}_{1,k} + \mathbf{d}_{2,k} - \mathbf{z}_k,$$

$$\mathbf{d}_{1,k} = \begin{bmatrix} \mathbf{f_k f_k^{(n),H}}\left(\widehat{\mathbf{h}}_\mathbf{k} + \widehat{\mathbf{G}}_\mathbf{k}^{\mathrm{H}}\mathbf{e^{(n)}}\right) \\ \mathrm{vec}^*(\mathbf{e}\left(\widehat{\mathbf{h}}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}) \end{bmatrix},$$

$$\mathbf{d}_{2,k} = \begin{bmatrix} \mathbf{f_k^{(n)} f_k^{H}}\left(\widehat{\mathbf{h}}_\mathbf{k} + \widehat{\mathbf{G}}_\mathbf{k}^{\mathrm{H}}\mathbf{e}\right) \\ \mathrm{vec}^*(\mathbf{e}^{(n)}\left(\widehat{\mathbf{h}}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{f}_k\mathbf{f}_k^{(n),\mathrm{H}}) \end{bmatrix},$$

$$\mathbf{z}_k = \begin{bmatrix} \mathbf{f_k^{(n)} f_k^{(n),H}}\left(\widehat{\mathbf{h}}_\mathbf{k} + \widehat{\mathbf{G}}_\mathbf{k}^{\mathrm{H}}\mathbf{e^{(n)}}\right) \\ \mathrm{vec}^*(\mathbf{e}^{(n)}\left(\widehat{\mathbf{h}}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{f}_k^{(n)}\mathbf{f}_k^{(n),\mathrm{H}}) \end{bmatrix},$$

$$\widetilde{a}_k = 2\mathrm{Re}\left\{d_k\right\} - z_k,$$

$$d_k = \left(\widehat{\mathbf{h}}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}\left(\widehat{\mathbf{h}}_k + \widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}\right),$$

$$z_k = \left(\widehat{\mathbf{h}}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{f}_k^{(n)}\mathbf{f}_k^{(n),\mathrm{H}}\left(\widehat{\mathbf{h}}_k + \widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}^{(n)}\right),$$

$$\mathbf{x}_k = \left[\triangle\mathbf{h}_k^{\mathrm{H}}\ \mathrm{vec}^{\mathrm{H}}(\triangle\mathbf{G}_k^*)\right]^{\mathrm{H}}.$$

Proof: Please refer to Appendix B. ■

Based on Lemma 4, constraints (24) are equivalently rewritten as

$$\begin{aligned}&\mathbf{x}_k^{\mathrm{H}}\widetilde{\mathbf{A}}_k\mathbf{x}_k + 2\mathrm{Re}\left\{\widetilde{\mathbf{a}}_k^{\mathrm{H}}\mathbf{x}_k\right\} + \widetilde{a}_k \\ &\geq \beta_k(2^{R_k} - 1), \mathcal{E}_k^{full}, \forall k \in \mathcal{K}.\end{aligned} \tag{27}$$

Before applying Lemma 1, it is beneficial to express $\mathcal{E}_k^{full}$ in terms of the following quadratic expressions as

$$\mathcal{E}_k^{full} \triangleq \begin{cases} \mathbf{x}_k^{\mathrm{H}}\begin{bmatrix}\mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{x}_k - \xi_{\mathrm{h},k}^2 \leq 0, \\ \mathbf{x}_k^{\mathrm{H}}\begin{bmatrix}\mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{MN}\end{bmatrix}\mathbf{x}_k - \xi_{\mathrm{g},k}^2 \leq 0. \end{cases}$$

Therefore, after introducing $\boldsymbol{\varpi}_{\mathrm{h}} = [\varpi_{\mathrm{h},1}, \ldots, \varpi_{\mathrm{h},K}]^{\mathrm{T}} \geq 0$ and $\boldsymbol{\varpi}_{\mathrm{g}} = [\varpi_{\mathrm{g},1}, \ldots, \varpi_{\mathrm{g},K}]^{\mathrm{T}} \geq 0$ as slack variables, constraints (24) can be transformed by Lemma 1 into the following equivalent LMIs as

$$\left[\begin{array}{cc} \widetilde{\mathbf{A}}_k + \left[\begin{array}{cc} \varpi_{\mathrm{h},k}\mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \varpi_{\mathrm{g},k}\mathbf{I}_{MN} \end{array}\right] & \widetilde{\mathbf{a}}_k \\ \widetilde{\mathbf{a}}_k^{\mathrm{H}} & C_k^{full} \end{array}\right] \succeq \mathbf{0}, \forall k \in \mathcal{K}, \quad (28)$$

where $C_k^{full} = \widetilde{a}_k - \beta_k(2^{R_k} - 1) - \varpi_{\mathrm{h},k}\xi_{\mathrm{h},k}^2 - \varpi_{\mathrm{g},k}\xi_{\mathrm{g},k}^2$.

Next, by inserting $\mathbf{h}_k = \widehat{\mathbf{h}}_k + \triangle\mathbf{h}_k$ and $\mathbf{G}_k = \widehat{\mathbf{G}}_k + \triangle\mathbf{G}_k$ into the equivalent matrix inequality of the INs power in (13), we have

$$\begin{aligned} \mathbf{0} \preceq & \left[\begin{array}{cc} \beta_k - \sigma_k^2 & \widetilde{\mathbf{t}}_k^{\mathrm{H}} \\ \widetilde{\mathbf{t}}_k & \mathbf{I} \end{array}\right] \\ & + \left[\begin{array}{cc} 0 & \left(\triangle\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\triangle\mathbf{G}_k\right)\mathbf{F}_{-k} \\ \mathbf{F}_{-k}^{\mathrm{H}}\left(\triangle\mathbf{h}_k + \triangle\mathbf{G}_k^{\mathrm{H}}\mathbf{e}\right) & \mathbf{0} \end{array}\right] \\ & \preceq \left[\begin{array}{c} \mathbf{0} \\ \mathbf{F}_{-k}^{\mathrm{H}} \end{array}\right]\left[\triangle\mathbf{h}_{\mathrm{r},k}\ \mathbf{0}\right] + \left[\begin{array}{c} \triangle\mathbf{h}_{\mathrm{r},k}^{\mathrm{H}} \\ \mathbf{0} \end{array}\right]\left[\mathbf{0}\ \mathbf{F}_{-k}\right] \\ & + \left[\begin{array}{c} \mathbf{0} \\ \mathbf{F}_{-k}^{\mathrm{H}} \end{array}\right]\triangle\mathbf{G}_k^{\mathrm{H}}\left[\mathbf{e}\ \mathbf{0}\right] + \left[\begin{array}{c} \mathbf{e}^{\mathrm{H}} \\ \mathbf{0} \end{array}\right]\triangle\mathbf{G}_k\left[\mathbf{0}\ \mathbf{F}_{-k}\right] \\ & + \left[\begin{array}{cc} \beta_k - \sigma_k^2 & \widetilde{\mathbf{t}}_k^{\mathrm{H}} \\ \widetilde{\mathbf{t}}_k & \mathbf{I} \end{array}\right], \end{aligned} \quad (29)$$

where $\widetilde{\mathbf{t}}_k = ((\widehat{\mathbf{h}}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{F}_{-k})^{\mathrm{H}}$.

Applying Lemma 2 and defining slack variables $\boldsymbol{\mu}_{\mathrm{g}} = [\mu_{\mathrm{g},1}, \ldots, \mu_{\mathrm{g},K}]^{\mathrm{T}} \geq 0$ and $\boldsymbol{\mu}_{\mathrm{h}} = [\mu_{\mathrm{h},1}, \ldots, \mu_{\mathrm{h},K}]^{\mathrm{T}} \geq 0$, the equivalent LMIs of the worst-case INs power constraints (25) are given as

$$\left[\begin{array}{cccc} Temp_k & \widetilde{\mathbf{t}}_k^{\mathrm{H}} & \mathbf{0}_{1\times N} & \mathbf{0}_{1\times N} \\ \widetilde{\mathbf{t}}_k & \mathbf{I}_{(K-1)} & \xi_{\mathrm{g},k}\mathbf{F}_{-k}^{\mathrm{H}} & \xi_{\mathrm{h},k}\mathbf{F}_{-k}^{\mathrm{H}} \\ \mathbf{0}_{N\times 1} & \xi_{\mathrm{g},k}\mathbf{F}_{-k} & \mu_{\mathrm{g},k}\mathbf{I}_N & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times 1} & \xi_{\mathrm{h},k}\mathbf{F}_{-k} & \mathbf{0}_{N\times N} & \mu_{\mathrm{h},k}\mathbf{I}_N \end{array}\right] \succeq \mathbf{0}, \forall k \in \mathcal{K}, \quad (30)$$

where $Temp_k = \beta_k - \sigma_k^2 - \mu_{\mathrm{g},k}M - \mu_{\mathrm{h},k}$.

With (28) and (30), the worst-case robust beamforming design problem under full channel uncertainty can be formulated as

$$\min_{\substack{\mathbf{F},\mathbf{e},\boldsymbol{\beta},\boldsymbol{\varpi}_{\mathrm{g}},\\ \boldsymbol{\varpi}_{\mathrm{h}},\boldsymbol{\mu}_{\mathrm{g}},\boldsymbol{\mu}_{\mathrm{h}}}} \quad ||\mathbf{F}||_F^2 \quad (31a)$$

$$\text{s.t.} \quad (28), (30), (7c), \quad (31b)$$

$$\boldsymbol{\varpi}_{\mathrm{g}} \geq 0, \boldsymbol{\varpi}_{\mathrm{h}} \geq 0, \boldsymbol{\mu}_{\mathrm{g}} \geq 0, \boldsymbol{\mu}_{\mathrm{h}} \geq 0. \quad (31c)$$

Problem (31) is again non-convex and has coupled variables, which can be solved similarly to Problem (16) and thus omitted for simplicity.

IV. Outage Constrained Robust Beamforming Design

In general, the channel estimation error follows the Gaussian distribution [14]. Hence, it is unbounded. The above bounded channel model may not be able to characterize the practical channel error model. As a result, in this section, we consider the statistical CSI error model. Specifically, by defining the maximum data rate outage probabilities $\rho_1, \ldots, \rho_K \in (0, 1]$, the transmit power minimization problem is formulated as

$$\min_{\mathbf{F},\mathbf{e}} \quad ||\mathbf{F}||_F^2 \quad (32a)$$

$$\text{s.t.} \quad \Pr\{\mathcal{R}_k(\mathbf{F}, \mathbf{e}) \geq R_k\} \geq 1 - \rho_k, \forall k \in \mathcal{K} \quad (32b)$$

$$|e_m|^2 = 1, 1 \leq m \leq M. \quad (32c)$$

The rate outage constraints (32b) guarantee that the probability of each user that can successfully decode its message at a data rate of R_k is no less than $1 - \rho_k$.

The outage constrained robust beamforming design problem in (32) is computationally intractable due to the fact that the rate outage probability constraints (32b) have no simple closed-form expressions [28]. In order to solve Problem (32), a safe approximation based on Bernstein-type inequality is given in the following lemma.

Lemma 5: (Bernstein-Type Inequality: Lemma 1 in [28]) Assume $f(\mathbf{x}) = \mathbf{x}^{\mathrm{H}}\mathbf{U}\mathbf{x} + 2\mathrm{Re}\{\mathbf{u}^{\mathrm{H}}\mathbf{x}\} + u$, where $\mathbf{U} \in \mathbb{H}^{n\times n}$, $\mathbf{u} \in \mathbb{C}^{n\times 1}$, $u \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{C}^{n\times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Then for any $\rho \in [0, 1]$, the following approximation holds:

$$\Pr\{\mathbf{x}^{\mathrm{H}}\mathbf{U}\mathbf{x} + 2\mathrm{Re}\{\mathbf{u}^{\mathrm{H}}\mathbf{x}\} + u \geq 0\} \geq 1 - \rho \quad (33a)$$

$$\Rightarrow \mathrm{Tr}\{\mathbf{U}\} - \sqrt{2\ln(1/\rho)}x + \ln(\rho)\lambda_{\max}^{+}(-\mathbf{U}) + u \geq 0 \quad (33b)$$

$$\Rightarrow \left\{\begin{array}{c} \mathrm{Tr}\{\mathbf{U}\} - \sqrt{2\ln(1/\rho)}x + \ln(\rho)y + u \geq 0 \\ \sqrt{||\mathbf{U}||_F^2 + 2||\mathbf{u}||^2} \leq x \\ y\mathbf{I} + \mathbf{U} \succeq \mathbf{0}, y \geq 0, \end{array}\right. \quad (33c)$$

where $\lambda_{\max}^{+}(-\mathbf{U}) = \max(\lambda_{\max}(-\mathbf{U}), 0)$. x and y are slack variables.

Please refer to [28] for the proof of Lemma 5.

In the following subsections, we first design the relatively simple robust beamforming under the partial channel uncertainty, and then extend it to the full channel uncertainty case.

A. Scenario 1: Partial Channel Uncertainty

Before the derivations, the rate outage probability of user k in (32b) is rewritten as

$$\begin{aligned} & \Pr\left\{\log_2\left(1 + \frac{\left|\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k\right)\mathbf{f}_k\right|^2}{\left\|\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k\right)\mathbf{F}_{-k}\right\|_2^2 + \sigma_k^2}\right) \geq R_k\right\} \\ & = \Pr\left\{\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k\right)\boldsymbol{\Phi}_k\left(\mathbf{h}_k + \mathbf{G}_k^{\mathrm{H}}\mathbf{e}\right) - \sigma_k^2 \geq 0\right\}, \end{aligned} \quad (34)$$

where $\boldsymbol{\Phi}_k = \mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}/(2^{R_k} - 1) - \mathbf{F}_{-k}\mathbf{F}_{-k}^{\mathrm{H}}$.

For the convenience of derivations, we assume that $\boldsymbol{\Sigma}_{\mathrm{g},k} = \varepsilon_{\mathrm{g},k}^2\mathbf{I}$, then the RCSIT error in (6) can be rewritten as $\mathrm{vec}(\triangle\mathbf{G}_k) = \varepsilon_{\mathrm{g},k}\mathbf{i}_{\mathrm{g},k}$ where $\mathbf{i}_{\mathrm{g},k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Defining $\mathbf{E} = \mathbf{e}\mathbf{e}^{\mathrm{H}}$, the rate outage probability (34) is reformulated in (35) at the top of the next page. Therefore, the rate outage constraints (32b) are given as

$$\begin{aligned} & \Pr\left\{\mathbf{i}_{\mathrm{g},k}^{\mathrm{H}}\mathbf{U}_k\mathbf{i}_{\mathrm{g},k} + 2\mathrm{Re}\{\mathbf{u}_k^{\mathrm{H}}\mathbf{i}_{\mathrm{g},k}\} + u_k \geq 0\right\} \\ & \quad \geq 1 - \rho_k, \forall k \in \mathcal{K}, \end{aligned} \quad (36)$$

$$\begin{aligned}
&\Pr\left\{\left(\mathbf{h}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}(\widehat{\mathbf{G}}_k+\triangle\mathbf{G}_k)\right)\boldsymbol{\Phi}_k\left(\mathbf{h}_k+(\widehat{\mathbf{G}}_k+\triangle\mathbf{G}_k)^{\mathrm{H}}\mathbf{e}\right)-\sigma_k^2\geq 0\right\}\\
&=\Pr\Big\{\mathrm{vec}^{\mathrm{H}}(\triangle\mathbf{G}_k)(\boldsymbol{\Phi}_k^{\mathrm{T}}\otimes\mathbf{E})\mathrm{vec}(\triangle\mathbf{G}_k)+2\mathrm{Re}\{\mathrm{vec}^{\mathrm{H}}((\mathbf{e}\mathbf{h}_k^{\mathrm{H}}+\mathbf{E}\widehat{\mathbf{G}}_k)\boldsymbol{\Phi}_k)\mathrm{vec}(\triangle\mathbf{G}_k)\}\\
&\qquad+(\mathbf{h}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\boldsymbol{\Phi}_k(\mathbf{h}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e})-\sigma_k^2\geq 0\Big\}\\
&=\Pr\left\{\varepsilon_{\mathrm{g},k}^2\mathbf{i}_{\mathrm{g},k}^{\mathrm{H}}(\boldsymbol{\Phi}_k^{\mathrm{T}}\otimes\mathbf{E})\mathbf{i}_{\mathrm{g},k}+2\mathrm{Re}\{\varepsilon_{\mathrm{g},k}\mathrm{vec}^{\mathrm{H}}((\mathbf{e}\mathbf{h}_k^{\mathrm{H}}+\mathbf{E}\widehat{\mathbf{G}}_k)\boldsymbol{\Phi}_k)\mathbf{i}_{\mathrm{g},k}\}+(\mathbf{h}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\boldsymbol{\Phi}_k(\mathbf{h}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e})-\sigma_k^2\geq 0\right\}. \quad (35)
\end{aligned}$$

where

$$\mathbf{U}_k=\varepsilon_{\mathrm{g},k}^2(\boldsymbol{\Phi}_k^{\mathrm{T}}\otimes\mathbf{E}), \quad (37a)$$

$$\mathbf{u}_k=\varepsilon_{\mathrm{g},k}\mathrm{vec}((\mathbf{e}\mathbf{h}_k^{\mathrm{H}}+\mathbf{E}\widehat{\mathbf{G}}_k)\boldsymbol{\Phi}_k^{\mathrm{H}}), \quad (37b)$$

$$u_k=(\mathbf{h}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\boldsymbol{\Phi}_k(\mathbf{h}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e})-\sigma_k^2. \quad (37c)$$

Applying Lemma 5 and introducing auxiliary variables $\mathbf{x}=[x_1,\ldots,x_K]^{\mathrm{T}}$ and $\mathbf{y}=[y_1,\ldots,y_K]^{\mathrm{T}}$, rate outage constraint of user k in (36) is transformed into the deterministic form as

$$\mathrm{Tr}\left\{\mathbf{U}_k\right\}-\sqrt{2\ln(1/\rho_k)}x_k+\ln(\rho_k)y_k+u_k\geq 0, \quad (38a)$$

$$\sqrt{||\mathbf{U}_k||_F^2+2||\mathbf{u}_k||^2}\leq x_k, \quad (38b)$$

$$y_k\mathbf{I}+\mathbf{U}_k\succeq\mathbf{0}, y_k\geq 0. \quad (38c)$$

(38) can be further simplified by some mathematical transformations as follows

$$\begin{aligned}
\mathrm{Tr}\left\{\mathbf{U}_k\right\}&=\varepsilon_{\mathrm{g},k}^2\mathrm{Tr}\left\{\boldsymbol{\Phi}_k^{\mathrm{T}}\otimes\mathbf{E}\right\}=\varepsilon_{\mathrm{g},k}^2\mathrm{Tr}\left\{\boldsymbol{\Phi}_k\right\}\mathrm{Tr}\left\{\mathbf{E}\right\}\\
&=\varepsilon_{\mathrm{g},k}^2M\mathrm{Tr}\left\{\boldsymbol{\Phi}_k\right\}, \quad (39a)\\
||\mathbf{U}_k||_F^2&=\varepsilon_{\mathrm{g},k}^4||(\boldsymbol{\Phi}_k^{\mathrm{T}}\otimes\mathbf{E})||_F^2=\varepsilon_{\mathrm{g},k}^4||\boldsymbol{\Phi}_k||_F^2||\mathbf{E}||_F^2\\
&=\varepsilon_{\mathrm{g},k}^4M^2||\boldsymbol{\Phi}_k||_F^2, \quad (39b)\\
||\mathbf{u}_k||^2&=\varepsilon_{\mathrm{g},k}^2||\mathrm{vec}((\mathbf{e}\mathbf{h}_k^{\mathrm{H}}+\mathbf{E}\widehat{\mathbf{G}}_k)\boldsymbol{\Phi}_k^{\mathrm{H}})||^2\\
&=\varepsilon_{\mathrm{g},k}^2M||\left(\mathbf{h}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k\right)\boldsymbol{\Phi}_k||_2^2, \quad (39c)\\
\lambda(\mathbf{U}_k)&=\lambda(\varepsilon_{\mathrm{g},k}^2(\boldsymbol{\Phi}_k^{\mathrm{T}}\otimes\mathbf{E}))=\varepsilon_{\mathrm{g},k}^2\lambda(\boldsymbol{\Phi}_k^{\mathrm{T}}\otimes\mathbf{E})\\
&=\varepsilon_{\mathrm{g},k}^2\lambda(\boldsymbol{\Phi}_k)\lambda(\mathbf{E})=\varepsilon_{\mathrm{g},k}^2M\lambda(\boldsymbol{\Phi}_k). \quad (39d)
\end{aligned}$$

Operation $\lambda(\mathbf{X})$ means the eigenvalues of $\mathbf{X}$. (39a) and (39b) are from [P76 in [29]], (39d) is from [P421 in [29]].

Therefore, according to Lemma 5 and equation (39), the approximation problem of Problem (32) can be given as

$$\min_{\mathbf{F},\mathbf{e},\mathbf{x},\mathbf{y}}\ ||\mathbf{F}||_F^2 \quad (40a)$$

$$\text{s.t. } \varepsilon_{\mathrm{g},k}^2M\mathrm{Tr}\left\{\boldsymbol{\Phi}_k\right\}-\sqrt{2\ln(1/\rho_k)}x_k-\ln(1/\rho_k)y_k+u_k\geq 0,\forall k\in\mathcal{K} \quad (40b)$$

$$\left\|\begin{array}{c}\varepsilon_{\mathrm{g},k}^2M\mathrm{vec}(\boldsymbol{\Phi}_k)\\ \sqrt{2M}\varepsilon_{\mathrm{g},k}\boldsymbol{\Phi}_k\left(\mathbf{h}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}\right)\end{array}\right\|\leq x_k,\forall k\in\mathcal{K} \quad (40c)$$

$$y_k\mathbf{I}+\varepsilon_{\mathrm{g},k}^2M\boldsymbol{\Phi}_k\succeq\mathbf{0}, y_k\geq 0,\forall k\in\mathcal{K} \quad (40d)$$

$$|e_m|^2=1,\forall m\in\mathcal{M}. \quad (40e)$$

Problem (40) is still difficult to solve because constraints (40c) are non-convex and have coupled variables $\mathbf{F}$ and $\mathbf{e}$. We use AO method to update $\mathbf{F}$ and $\mathbf{e}$ in an iterative manner. More specifically, by first fixing $\mathbf{e}$, the non-convex problem in $\mathbf{F}$ at hand is relaxed by employing the SDR technique [30] and solved by CVX. $\mathbf{F}$ is then fixed and the resulting non-convex problem of $\mathbf{e}$ is also handled under the penalty CCP method.

For fixed $\mathbf{e}$, let $\boldsymbol{\Phi}_k=\boldsymbol{\Gamma}_k/(2^{R_k}-1)-\sum_{i=1,i\neq k}^K\boldsymbol{\Gamma}_i$ where $\boldsymbol{\Gamma}_k=\mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}$, Problem (40) corresponding to $\mathbf{F}$ is rewritten as

$$\min_{\boldsymbol{\Gamma},\mathbf{x},\mathbf{y}}\ \sum_{k=1}^K\mathrm{Tr}\left\{\boldsymbol{\Gamma}_k\right\} \quad (41a)$$

$$\text{s.t. } \varepsilon_{\mathrm{g},k}^2M\mathrm{Tr}\left\{\boldsymbol{\Phi}_k\right\}-\sqrt{2\ln(1/\rho_k)}x_k-\ln(1/\rho_k)y_k+u_k\geq 0,\forall k\in\mathcal{K} \quad (41b)$$

$$\left\|\begin{array}{c}\varepsilon_{\mathrm{g},k}^2M\mathrm{vec}(\boldsymbol{\Phi}_k)\\ \sqrt{2M}\varepsilon_{\mathrm{g},k}\boldsymbol{\Phi}_k\left(\mathbf{h}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}\right)\end{array}\right\|\leq x_k,\forall k\in\mathcal{K} \quad (41c)$$

$$y_k\mathbf{I}+\varepsilon_{\mathrm{g},k}^2M\boldsymbol{\Phi}_k\succeq\mathbf{0}, y_k\geq 0,\forall k\in\mathcal{K} \quad (41d)$$

$$\boldsymbol{\Gamma}_k\succeq\mathbf{0},\forall k\in\mathcal{K} \quad (41e)$$

$$\mathrm{rank}(\boldsymbol{\Gamma}_k)=1,\forall k\in\mathcal{K}, \quad (41f)$$

where $\boldsymbol{\Gamma}=[\boldsymbol{\Gamma}_1,\ldots,\boldsymbol{\Gamma}_K]$. Problem (41) can be solved by adopting the SDR technique, i.e., removing $\mathrm{rank}(\boldsymbol{\Gamma}_k)=1,\forall k\in\mathcal{K}$ from the problem formulation, the resulting convex SDP problem is then efficiently solved by the CVX tools. The following theorem reveals the tightness of the SDR.

Theorem 1: If the relaxed version of Problem (41) is feasible, then there always exists a feasible solution, denoted as $\boldsymbol{\Gamma}^\star=[\boldsymbol{\Gamma}_1^\star,\ldots,\boldsymbol{\Gamma}_K^\star]$, satisfying $\mathrm{rank}(\boldsymbol{\Gamma}_k^\star)=1,\forall k\in\mathcal{K}$.

Proof: Please refer to Appendix C. ■

Remark 1: Numerical results show that, the optimal $\boldsymbol{\Gamma}_k^\star$ is usually of rank one before we construct the rank-1 solution mentioned in Appendix C. The optimal $\mathbf{f}_k$ can be obtained from $\boldsymbol{\Gamma}_k^\star$ by using eigenvalue decomposition.

We now consider the subproblem of $\mathbf{e}$ with fixed $\mathbf{F}$. With the same purpose of (18), we introduce slack variables $\boldsymbol{\alpha}=[\alpha_1,\ldots,\alpha_K]^{\mathrm{T}}$ to the rate outage probability in (34) and have

$$\Pr\left\{\left(\mathbf{h}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\mathbf{G}_k\right)\boldsymbol{\Phi}_k\left(\mathbf{h}_k+\mathbf{G}_k^{\mathrm{H}}\mathbf{e}\right)-\sigma_k^2-\alpha_k\geq 0\right\}. \quad (42)$$

Then, (37c) is also modified as follows

$$u_k^{\mathbf{e}}=(\mathbf{h}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\boldsymbol{\Phi}_k(\mathbf{h}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e})-\sigma_k^2-\alpha_k. \quad (43)$$

We note that (43) is non-concave in $\mathbf{e}$ due to the fact that $\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k\mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}/(2^{R_k}-1)$ in $\widehat{\mathbf{G}}_k\mathbf{\Phi}_k\widehat{\mathbf{G}}_k^{\mathrm{H}}$ is convex. By using the first-order Taylor inequality given in Appendix A, $\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k\mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}/(2^{R_k}-1)$ can be lower bounded linearly by $u^{\mathbf{e}}_{\mathrm{linear},k}=(2\mathrm{Re}\{\mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k\mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}\}-\mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k\mathbf{f}_k\mathbf{f}_k^{\mathrm{H}}\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}^{(n)})/(2^{R_k}-1)$. We then construct an equivalent concave version of (43), which is given as

$$u_k^{\mathbf{e}}=u^{\mathbf{e}}_{\mathrm{linear},k}-\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k\mathbf{F}_{-k}\mathbf{F}_{-k}^{\mathrm{H}}\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}+2\mathrm{Re}\{\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k\mathbf{\Phi}_k\mathbf{h}_k\}+\mathbf{h}_k^{\mathrm{H}}\mathbf{\Phi}_k\mathbf{h}_k-\sigma_k^2-\alpha_k+M\mathrm{const}\mathrm{E}_k. \tag{44}$$

In addition, constraints (40d) are independent of $\mathbf{e}$ and transformed from $\lambda^{+}_{\max}(-\mathbf{U})$ in Lemma 5, we then can have $y_k=\max(\lambda_{\max}(-\varepsilon^2_{\mathrm{g},k}M\mathbf{\Phi}_k),0),\forall k\in\mathcal{K}$. With $\boldsymbol{\alpha}$ and (44), the subproblem of (40) corresponding to $\mathbf{e}$ is formulated as

$$\max_{\mathbf{e},\boldsymbol{\alpha},\mathbf{x},\mathbf{y}}\ \sum_{k=1}^{K}\alpha_k \tag{45a}$$

$$\text{s.t. } \varepsilon^2_{\mathrm{g},k}M\mathrm{Tr}\{\mathbf{\Phi}_k\}-\sqrt{2\ln(1/\rho_k)}x_k-\ln(1/\rho_k)y_k+u_k^{\mathbf{e}}\geq 0,\forall k\in\mathcal{K} \tag{45b}$$

$$\left\|\begin{array}{c}\varepsilon^2_{\mathrm{g},k}M||\mathbf{\Phi}_k||_F\\ \sqrt{2\,M}\varepsilon_{\mathrm{g},k}\mathbf{\Phi}_k\left(\mathbf{h}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}\right)\end{array}\right\|\leq x_k,\forall k\in\mathcal{K} \tag{45c}$$

$$\boldsymbol{\alpha}\geq 0 \tag{45d}$$

$$|e_m|^2=1,\forall m\in\mathcal{M}. \tag{45e}$$

The non-convex constraints (45e) in Problem (45) is solved by using the same techniques as those used for solving Problem (21), then the resulting approximation problem for Problem (45) can be solved by using Algorithm 1.

B. Scenario 2: Full Channel Uncertainty

In this subsection, we extend the outage constrained robust beamforming design from the partial channel uncertainty to the case where all the channels are imperfect at the BS. By considering the full statistical CSI error in (6), (34) is then formulated as

$$\begin{aligned}&\Pr\Big\{\left(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{\Phi}_k\left(\widehat{\mathbf{h}}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}\right)\\&+2\mathrm{Re}\left\{\left(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k\right)\mathbf{\Phi}_k\left(\triangle\mathbf{h}_k+\triangle\mathbf{G}_k^{\mathrm{H}}\mathbf{e}\right)\right\}-\sigma_k^2\\&+\left(\triangle\mathbf{h}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\triangle\mathbf{G}_k\right)\mathbf{\Phi}_k\left(\triangle\mathbf{h}_k+\triangle\mathbf{G}_k^{\mathrm{H}}\mathbf{e}\right)\geq 0\Big\}.\end{aligned} \tag{46}$$

Assuming that $\mathbf{\Sigma}_{\mathrm{h},k}=\varepsilon^2_{\mathrm{h},k}\mathbf{I}$, then the DCSIT can be expressed as $\triangle\mathbf{h}_k=\varepsilon_{\mathrm{h},k}\mathbf{i}_{\mathrm{h},k}$ where $\mathbf{i}_{\mathrm{h},k}\sim\mathcal{CN}(\mathbf{0},\mathbf{I})$. The second term inside (46) is rewritten as

$$\begin{aligned}&2\mathrm{Re}\Big\{(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{\Phi}_k\triangle\mathbf{h}_k+\mathrm{vec}^{\mathrm{T}}(\mathbf{e}(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{\Phi}_k)\mathrm{vec}(\triangle\mathbf{G}_k^{*})\Big\}\\&=2\mathrm{Re}\Big\{\varepsilon_{\mathrm{h},k}(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{\Phi}_k\mathbf{i}_{\mathrm{h},k}+\varepsilon_{\mathrm{g},k}\mathrm{vec}^{\mathrm{T}}(\mathbf{e}(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{\Phi}_k)\mathbf{i}_{\mathrm{g},k}^{*}\Big\}\\&=2\mathrm{Re}\left\{\widetilde{\mathbf{u}}_k^{\mathrm{H}}\widetilde{\mathbf{i}}_k\right\},\end{aligned}$$

where $\widetilde{\mathbf{i}}_k=[\mathbf{i}_{\mathrm{h},k}^{\mathrm{H}}\ \mathbf{i}_{\mathrm{g},k}^{\mathrm{T}}]^{\mathrm{H}}$ and

$$\widetilde{\mathbf{u}}_k=\begin{bmatrix}\varepsilon_{\mathrm{h},k}\mathbf{\Phi}_k(\widehat{\mathbf{h}}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e})\\ \varepsilon_{\mathrm{g},k}\mathrm{vec}^{*}(\mathbf{e}(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{\Phi}_k)\end{bmatrix}.$$

The fourth term on the left hand side of (46) is rewritten as

$$\begin{aligned}&\triangle\mathbf{h}_k^{\mathrm{H}}\mathbf{\Phi}_k\triangle\mathbf{h}_k+2\mathrm{Re}\left\{\mathbf{e}^{\mathrm{H}}\triangle\mathbf{G}_k\mathbf{\Phi}_k\triangle\mathbf{h}_k\right\}+\mathbf{e}^{\mathrm{H}}\triangle\mathbf{G}_k\mathbf{\Phi}_k\triangle\mathbf{G}_k^{\mathrm{H}}\mathbf{e}\\&=\varepsilon^2_{\mathrm{h},k}\mathbf{i}_{\mathrm{h},k}^{\mathrm{H}}\mathbf{\Phi}_k\mathbf{i}_{\mathrm{h},k}+2\mathrm{Re}\left\{\triangle\mathbf{h}_k^{\mathrm{H}}(\mathbf{\Phi}_k\otimes\mathbf{e}^{\mathrm{T}})\mathrm{vec}(\triangle\mathbf{G}_k^{*})\right\}+\mathrm{vec}^{\mathrm{T}}(\triangle\mathbf{G}_k)(\mathbf{\Phi}_k\otimes\mathbf{E}^{\mathrm{T}})\mathrm{vec}(\triangle\mathbf{G}_k^{*})\\&=\varepsilon^2_{\mathrm{h},k}\mathbf{i}_{\mathrm{h},k}^{\mathrm{H}}\mathbf{\Phi}_k\mathbf{i}_{\mathrm{h},k}+2\mathrm{Re}\left\{\varepsilon_{\mathrm{h},k}\varepsilon_{\mathrm{g},k}\mathbf{i}_{\mathrm{h},k}^{\mathrm{H}}(\mathbf{\Phi}_k\otimes\mathbf{e}^{\mathrm{T}})\mathbf{i}_{\mathrm{g},k}^{*}\right\}+\varepsilon^2_{\mathrm{g},k}\mathbf{i}_{\mathrm{g},k}^{\mathrm{T}}(\mathbf{\Phi}_k\otimes\mathbf{E}^{\mathrm{T}})\mathbf{i}_{\mathrm{g},k}^{*}\\&=\widetilde{\mathbf{i}}_k^{\mathrm{H}}\widetilde{\mathbf{U}}_k\widetilde{\mathbf{i}}_k,\end{aligned}$$

where

$$\widetilde{\mathbf{U}}_k=\begin{bmatrix}\mathbf{\Sigma}_{\mathrm{h},k}^{1/2}\mathbf{\Phi}_k\mathbf{\Sigma}_{\mathrm{h},k}^{1/2} & \varepsilon_{\mathrm{h},k}\varepsilon_{\mathrm{g},k}(\mathbf{\Phi}_k\otimes\mathbf{e}^{\mathrm{T}})\\ \varepsilon_{\mathrm{h},k}\varepsilon_{\mathrm{g},k}(\mathbf{\Phi}_k\otimes\mathbf{e}^{*}) & \varepsilon^2_{\mathrm{g},k}(\mathbf{\Phi}_k\otimes\mathbf{E}^{\mathrm{T}})\end{bmatrix}.$$

Denote $\widetilde{u}_k=(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{\Phi}_k(\widehat{\mathbf{h}}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e})-\sigma_k^2$, the rate outage constraint (46) is then equivalent to

$$\Pr\left\{\widetilde{\mathbf{i}}_k^{\mathrm{H}}\widetilde{\mathbf{U}}_k\widetilde{\mathbf{i}}_k+2\mathrm{Re}\left\{\widetilde{\mathbf{u}}_k^{\mathrm{H}}\widetilde{\mathbf{i}}_k\right\}+\widetilde{u}_k\geq 0\right\}\geq 1-\rho_k. \tag{47}$$

Combining Lemma 5 and new auxiliary variables $\widetilde{\mathbf{x}}=[\widetilde{x}_1,\ldots,\widetilde{x}_K]^{\mathrm{T}}$ and $\widetilde{\mathbf{y}}=[\widetilde{y}_1,\ldots,\widetilde{y}_K]^{\mathrm{T}}$, the approximation of the data rate outage constraint of user k in (47) is given by

$$\mathrm{Tr}\left\{\widetilde{\mathbf{U}}_k\right\}-\sqrt{2\ln(1/\rho_k)}\widetilde{x}_k+\ln(\rho_k)\widetilde{y}_k+\widetilde{u}_k\geq 0, \tag{48a}$$

$$\sqrt{||\widetilde{\mathbf{U}}_k||_F^2+2||\widetilde{\mathbf{u}}_k||^2}\leq\widetilde{x}_k, \tag{48b}$$

$$\widetilde{y}_k\mathbf{I}+\widetilde{\mathbf{U}}_k\succeq\mathbf{0},\widetilde{y}_k\geq 0. \tag{48c}$$

We simplify some terms in (48) as follows:

$$\begin{aligned}\mathrm{Tr}\left\{\widetilde{\mathbf{U}}_k\right\}&=\mathrm{Tr}\left\{\begin{bmatrix}\varepsilon_{\mathrm{h},k}\mathbf{\Phi}_k^{1/2}\\ \varepsilon_{\mathrm{g},k}(\mathbf{\Phi}_k^{1/2}\otimes\mathbf{e}^{*})\end{bmatrix}\bullet\left[\varepsilon_{\mathrm{h},k}\mathbf{\Phi}_k^{1/2}\ \varepsilon_{\mathrm{g},k}(\mathbf{\Phi}_k^{1/2}\otimes\mathbf{e}^{\mathrm{T}})\right]\right\}\\&=(\varepsilon^2_{\mathrm{h},k}+\varepsilon^2_{\mathrm{g},k}M)\mathrm{Tr}\{\mathbf{\Phi}_k\},\end{aligned} \tag{49a}$$

$$||\widetilde{\mathbf{U}}_k||_F^2=(\varepsilon^2_{\mathrm{h},k}+\varepsilon^2_{\mathrm{g},k}M)^2||\mathbf{\Phi}_k||_F^2, \tag{49b}$$

$$||\widetilde{\mathbf{u}}_k||^2=(\varepsilon^2_{\mathrm{h},k}+\varepsilon^2_{\mathrm{g},k}M)||(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{\Phi}_k||_2^2, \tag{49c}$$

$$\widetilde{y}_k\mathbf{I}+\widetilde{\mathbf{U}}_k\succeq\mathbf{0}\Longrightarrow\widetilde{y}_k\mathbf{I}+(\varepsilon^2_{\mathrm{h},k}+\varepsilon^2_{\mathrm{g},k}M)\mathbf{\Phi}_k\succeq\mathbf{0}. \tag{49d}$$

The derivations of (49) are similar to (39).

Based on the above results, Problem (32) with imperfect DCSIT and imperfect CBIUT is given by

$$\min_{\mathbf{F},\mathbf{e},\widetilde{\mathbf{x}},\widetilde{\mathbf{y}}} \quad ||\mathbf{F}||_F^2 \tag{50a}$$

$$\text{s.t.} \quad (\varepsilon_{\mathrm{h},k}^2+\varepsilon_{\mathrm{g},k}^2 M)\mathrm{Tr}\left\{\mathbf{\Phi}_k\right\} - \sqrt{2\ln(1/\rho_k)}x_k - \ln(1/\rho_k)y_k + \widetilde{u}_k \geq 0, \forall k \in \mathcal{K} \tag{50b}$$

$$\left\| \frac{(\varepsilon_{\mathrm{h},k}^2+\varepsilon_{\mathrm{g},k}^2 M)\mathrm{vec}(\mathbf{\Phi}_k)}{\sqrt{2(\varepsilon_{\mathrm{h},k}^2+\varepsilon_{\mathrm{g},k}^2 M)\mathbf{\Phi}_k\left(\widehat{\mathbf{h}}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e}\right)}} \right\| \leq \widetilde{x}_k, \tag{50c}$$

$$\forall k \in \mathcal{K} \tag{50d}$$

$$\widetilde{y}_k\mathbf{I} + (\varepsilon_{\mathrm{h},k}^2+\varepsilon_{\mathrm{g},k}^2 M)\mathbf{\Phi}_k \succeq \mathbf{0}, \widetilde{y}_k \geq 0, \forall k \in \mathcal{K} \tag{50e}$$

$$|e_m|^2 = 1, \forall m \in \mathcal{M}. \tag{50f}$$

Comparing Problem (50) with Problem (40), we find that the former can be obtained from the latter by replacing $\varepsilon_{\mathrm{g},k}^2 M$ with $\varepsilon_{\mathrm{h},k}^2+\varepsilon_{\mathrm{g},k}^2 M$ and $\mathbf{h}_k$ with $\widehat{\mathbf{h}}_k$. Therefore, Problem (50) can be solved by using the same techniques as those used to solve Problem (40). In addition, when M is large, the impact of imperfect CBIUT dominates the performance of the system, which will be illustrated in the numerical results later. Thus, it is significant to investigate the robust beamforming in an IRS-aided system in which there are a large number of reflection elements with high channel estimation error.

V. Computational Complexity

In this section, we analyze the computational complexity of the proposed robust transmission design methods. Since all the resulting convex problems involving LMI, second-order cone (SOC) constraints and linear constraints that can be solved by a standard interior point method [31], we can compare the computational complexity of different methods in terms of their worst-case runtime, the general expression (we ignore the complexity of the linear constraints) of which is given by

$$\mathcal{O}\left(\left(\sum_{j=1}^{J} b_j + 2I\right)^{1/2} n\left(n^2 + n\underbrace{\sum_{j=1}^{J} b_j^2 + \sum_{j=1}^{J} b^3}_{\text{due to LMI}} + \underbrace{n\sum_{i=1}^{I} a_i^2}_{\text{due to SOC}}\right)\right),$$

where n is the number of variables, J is the number of LMIs of size b_j, and I is the number of SOC of size a_i. Based on the above expression, we provide the computational complexity per iteration of the proposed methods as follows:

1) *PCU-bounded method* denotes the worst-case beamforming design method under Scenario 1. The approximate complexity of Problem (17) is $o_{\mathbf{F}} = \mathcal{O}([K(MN+K+N+1)]^{1/2}n_1[n_1^2 + n_1K((MN+1)^2+(K+N)^2) + K((MN+1)^3+(K+N)^3)])$ where $n_1 = NK$, and that of Problem (22) is $o_{\mathbf{e}} = \mathcal{O}([K(MN+1+K)+2M]^{1/2}n_2[n_2^2 + n_2K((MN+1)^2+K^2) + K((MN+1)^3+K^3)+n_2M])$ where $n_2 = M$. Finally, the approximate complexity of PCU-bounded method per iteration is $o_{\mathbf{F}} + o_{\mathbf{e}}$.
2) *FCU-bounded method* denotes the worst-case beamforming design method under Scenario 2. The approximate complexity of Problem (31) is $o_{\mathbf{F}} + o_{\mathbf{e}}$, where $o_{\mathbf{F}} = \mathcal{O}([K(MN+3N+K+1)]^{1/2}n_1[n_1^2 + n_1K((MN+N+1)^2+(K+2N)^2) + K((MN+N+1)^3+(K+2N)^2)])$ with $n_1 = NK$, and $o_{\mathbf{e}} = \mathcal{O}([K(MN+1+K)+2M]^{1/2}n_2[n_2^2 + n_2K((MN+1)^2+K^2) + K((MN+1)^3+K^3)+n_2M])$ with $n_2 = M$.
3) *PCU-statistic method* denotes the outage constrained beamforming design method under Scenario 1. The approximate complexity of Problem (41) is $o_{\mathbf{F}} = \mathcal{O}([2K(N+1)]^{1/2}n_1[n_1^2 + 2n_1KN^2 + 2KN^3 + nKN^2(N+1)^2])$ where $n_1 = NK$, and that of Problem (45) is $o_{\mathbf{e}} = \mathcal{O}([4K+2M]^{1/2}n_2[n_2^2 + n_2(K(M^2+(N+1)^2)+M)])$ where $n_2 = M$. Finally, the approximate complexity of PCU-statistic method per iteration is $o_{\mathbf{F}} + o_{\mathbf{e}}$.
4) *FCU-statistic method* denotes the outage constrained beamforming design method under Scenario 2. Here, the approximate complexity per iteration is the same with the PCU-statistic method since they only have some different coefficients.

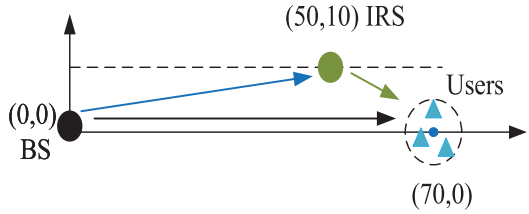


Fig. 2. The simulated system setup.

VI. Numerical Results and Discussions

In this section, we provide numerical results to evaluate the performance of our proposed algorithms. The simulated system setup of our consider network is shown in Fig. 2, in which we assume that the BS is located at (0 m, 0 m) and the IRS is placed at (50 m, 10 m). K users are randomly and uniformly distributed in a circle centered at (70 m, 0 m) with radius of 5 m. The channel models, i.e., $\{\mathbf{h}_k, \mathbf{G}_k\}_{\forall k\in\mathcal{K}}$, are assumed to include large-scale fading and small-scale fading. The large-scale fading model is expressed as $\mathrm{PL} = -\mathrm{PL}_0 - 10\alpha\log_{10}(d)$ dB, where α is the path loss exponent and d is the link distance in meters. PL_0 denotes the pathloss at the distance of 1 meter, which is set as 40 dB based on the 3GPP UMi model [32] with 3.5 GHz carrier frequency (i.e., carrier frequency of 5 G in China). The small-scale fading in $\{\mathbf{h}_k, \mathbf{G}_k\}_{\forall k\in\mathcal{K}}$ is assumed to be Rayleigh fading distribution. For the statistical CSI error model, the variance of $\mathrm{vec}(\triangle\mathbf{G}_k)$ and $\triangle\mathbf{h}_k$ are defined as $\varepsilon_{\mathrm{g},k}^2 = \delta_{\mathrm{g}}^2||\mathrm{vec}(\widehat{\mathbf{G}}_k)||_2^2$ and $\varepsilon_{\mathrm{h},k}^2 = \delta_{\mathrm{h}}^2||\widehat{\mathbf{h}}_k||_2^2$, respectively. $\delta_{\mathrm{g}} \in [0,1)$ and $\delta_{\mathrm{h}} \in [0,1)$ measure the relative amount of CSI uncertainties. For the bounded

TABLE I
SYSTEM PARAMETERS

Path loss exponents of BS-user link	$\alpha_{\mathrm{BU}} = 4$
Path loss exponents of BS-IRS link	$\alpha_{\mathrm{BI}} = 2.2$
Path loss exponents of IRS-user link	$\alpha_{\mathrm{IU}} = 2$
Noise power	$\sigma_1^2 = ... = \sigma_K^2 = -80$ dBm
Convergence tolerance	10^{-4}
Maximun outage probabilities	$\rho_1 = ... = \rho_K = \rho = 0.05$

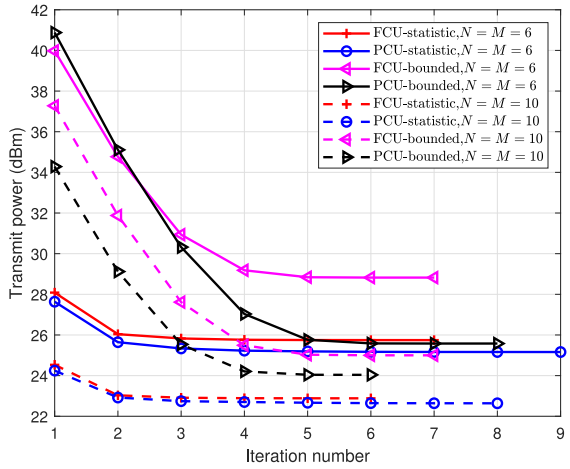


Fig. 3. Transmit power versus the iteration numbers of different algorithms, when $K = 3$ and $\{\delta_{\mathrm{g}}, \delta_{\mathrm{h}}\} = \{0.01, 0.02\}$.

CSI error model, the radii of the uncertainty regions are set as

$$\xi_{\mathrm{g},k} = \sqrt{\frac{\varepsilon_{\mathrm{g},k}^2}{2} F_{2MN}^{-1}(1-\rho)},$$

and

$$\xi_{\mathrm{h},k} = \sqrt{\frac{\varepsilon_{\mathrm{h},k}^2}{2} F_{2N}^{-1}(1-\rho)},$$

where $F_{2MN}^{-1}(\cdot)$ and $F_{2N}^{-1}(\cdot)$ denote the inverse cumulative distribution function (CDF) of the Chi-square distribution with degrees of freedom equal to $2MN$ and $2N$, respectively. According to [28], the above bounded CSI error model provides a fair comparison between the performance of the worst-case robust design and the outage constrained robust design. In addition, the target rates of all users are assumed to be the same, i.e., $R_1 = \cdots = R_K = R$ and the fixed simulation settings for our simulations are given in Table I.

Fig. 3 illustrates the convergence behavior of the proposed four algorithms. Here, the minimum rate is set as $R = 2$ bit/s/Hz, and the channel uncertainty levels are chosen as $\{\delta_{\mathrm{g}}, \delta_{\mathrm{h}}\} = \{0.01, 0.02\}$. It is observed that all algorithms converge rapidly and 10 iterations are sufficient for the algorithms to converge. It also shows that the convergence speed increases with the number of antennas. In addition, the algorithms under the statistical error model converge faster than those under the bounded error model.

Fig. 4 compares the average (central processing unit) CPU running time of the proposed algorithms versus the numbers of antenna elements at the BS and/or reflection elements at the IRS. The results are obtained by using a computer with

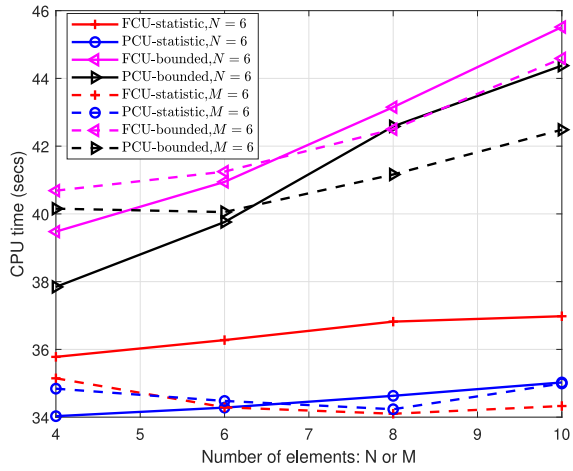


Fig. 4. Average CPU time versus the number of antenna elements at the IRS M and at the BS N, when $K = 2$ and $\{\delta_{\mathrm{g}}, \delta_{\mathrm{h}}\} = \{0.01, 0.02\}$.

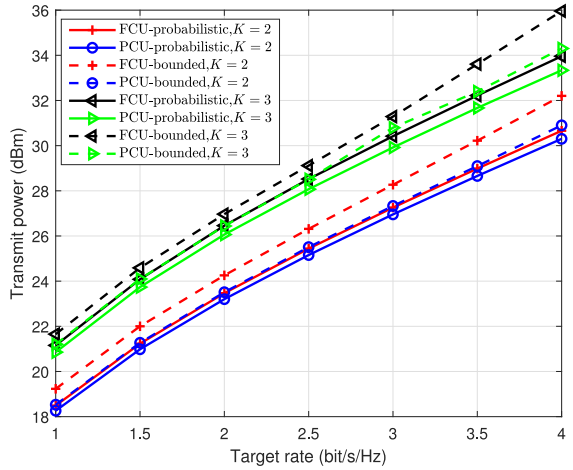
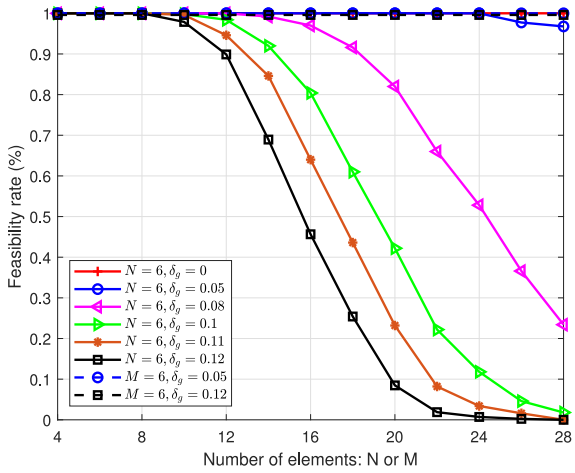


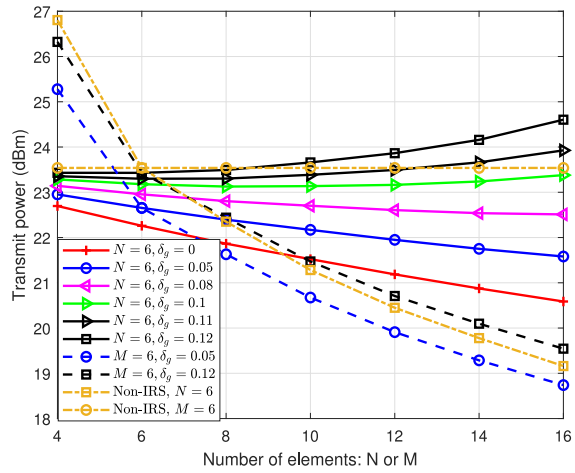
Fig. 5. Transmit power versus the target rate R under $N = M = 6$ and $\{\delta_{\mathrm{g}}, \delta_{\mathrm{h}}\} = \{0.01, 0.02\}$.

a 1.99 GHz i7-8550 U CPU and 16 GB RAM. Here, we set $K = 2$, $R = 2$ bit/s/Hz, and $\{\delta_{\mathrm{g}}, \delta_{\mathrm{h}}\} = \{0.01, 0.02\}$. Firstly, it is observed that the robust algorithms under the statistical CSI error model require much less CPU running time than those under the bounded CSI error model. This is due to the fact that there are some large-dimensional LMIs that increase the computational complexity of the worst-case algorithms. Secondly, the FCU-bounded algorithm requires more CPU time than the PCU-bounded algorithm because the DCSIT error $\triangle \mathbf{h}_k$ increases the dimension of the LMIs. Thirdly, when $M = 6$, the CPU running time of the outage constrained algorithm under two scenarios is similar due to the fact that no additional complexity is introduced by considering the additional DCSIT error.

Fig. 5 shows the minimum transmit power of the IRS-aided communication system versus the target rate requirements of users under various CSI error models. Some system parameters are set as $N = M = 6$, $K = \{2, 3\}$, $\{\delta_{\mathrm{g}}, \delta_{\mathrm{h}}\} = \{0.01, 0.02\}$. It is seen that the minimum transmit power increases with the target rate for both channel uncertainty scenarios and both CSI error

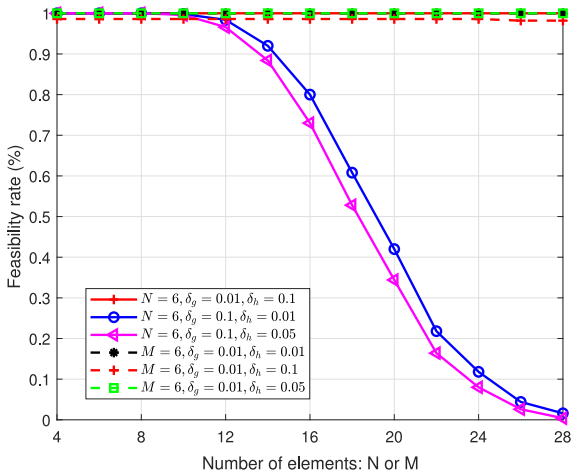


(a) Feasibility rate

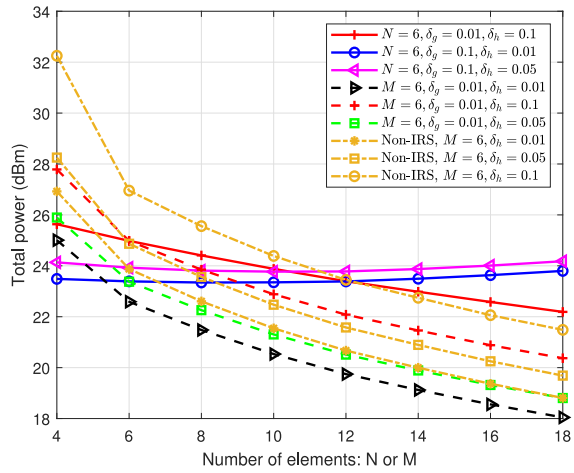


(b) Transmit power

Fig. 6. Feasibility rate and transmit power versus the number of antenna elements under the PCU scenario, when $K = 2$.



(a) Feasibility rate



(b) Transmit power

Fig. 7. Feasibility rate and transmit power versus the number of antenna elements under the FCU scenario, when $K = 2$.

models. In addition, it is also observed that the minimum transmit power of the worst-case robust design algorithms is larger than that of the outage constrained robust design algorithms. This is due to the fact that the worst-case optimization is the most conservative robust design, which requires more transmit power with the aim of ensuring that the achievable rate of each user meets the target rate requirement for the worst-case CSI error realization.

In the following, we study the impact of the accuracy of the CSI on the system performance. We adopt outage constrained robust beamforming design algorithms since the computational complexity of the worst-case robust beamforming design algorithms is unacceptable at large numbers of antennas.

Fig. 6 shows the feasibility rate and the minimum transmit power versus N or M when only the CBIUT is imperfect, i.e., $\delta_{\mathrm{h}} = 0$. We assume there are $K = 2$ users with $R = 2$ bit/s/Hz. The feasibility rate is defined as the ratio of the number of feasible channel realizations to the total number of channel realizations, where the feasible channel realization means that there exists a feasible solution to the outage constrained problem in (32) with this channel realization. An interesting phenomenon can be observed from Fig. 6 (a). When fixing the number of transmit antennas N, the feasibility rate decreases rapidly with the number of phase shifters at a high level of channel uncertainty ($\delta_{\mathrm{g}} \geq 0.08$). By contrast, when fixing the number of phase shifters M the feasibility rate keeps stable for different numbers of antennas even at a high level of channel uncertainty.

Based on the observations of Fig. 6(a), we further examine the minimum transmit power consumption of different channel uncertainty levels in Fig. 6(b) with a benchmark scheme without IRS. Fig. 6(b) is generated based on the channel realizations for which the feasible solutions can be obtained at $N = 16$ or $M = 16$.

We first study the case with fixed number of transmit antennas $N = 6$. In Fig. 6(b), the case with $\delta_{\mathrm{g}} = 0$ can be regarded as the perfect CBIUT case, and its minimum transmit power decreases with the number of the reflection elements. This trend is consistent with that of Fig. 4 in [27]. The minimum transmit

power consumption values under small values of δ_{g}, e.g., $\delta_{\mathrm{g}} = \{0.05, 0.08\}$, also decrease with the number of the reflection elements, and are higher than that of the perfect CBIUT case. The reason is that the BS needs to consume more power to compensate for the rate loss caused by the CBIUT error. However, when δ_{g} increases to 0.1 or larger, transmit power consumption starts to increase with the number of reflection elements. The reason is that increasing the number of reflection elements cannot only reduce the transmit power due to its increased beamforming gain, but also increase the channel estimation error that more transmit power is required to compensate for the channel errors. Therefore, when the CBIUT error is small, the benefits brought by the increase of M, outweighs its drawbacks, and vice versa. As a result, the number of IRS reflection elements should be carefully chosen, and the accuracy of the CBIUT estimation is crucial to reap the benefits offered by the IRS.

On the other hand, for the case with a fixed number of reflection elements, the transmit power consumption values decrease with the number of antennas at the BS even when the CBIUT error is high as $\delta_{\mathrm{g}} = 0.12$. The reason is that when the number of antennas is large, more degrees of freedom can be exploited to optimize the active beamforming vector at the BS to compensate for the channel estimation error. Finally, compared with the system without IRS, the IRS may lose its performance gain advantage under high CBIUT error.

Fig. 7 shows the feasibility rate and the minimum transmit power versus M or N when both the DCSIT and the CBIUT are imperfect. The simulation parameters are the same as those in Fig. 6. Fig. 7(a) shows that when δ_{g} is low, the feasibility rates achieved by various cases are always high. In addition, from Fig. 7(b), we find that the increase of the number of antennas at the BS is effective in reducing the transmit power consumption, which is not affected by the DCSIT error δ_{h} (see curves $M = 6, \delta_{\mathrm{g}} = 0.01, \delta_{\mathrm{h}} = \{0.01, 0.05, 0.1\}$).

VII. Conclusion

In this work, we investigated robust beamforming designs under imperfect CBIUT for the IRS-aided MU-MISO system. Our aim was to minimize the transmit power subject to the worst-case rate constraints under the bounded CSI error model and the rate outage probability constraints under the statistical CSI error model. The CSI uncertainties under the bounded CSI error model were addressed by applying the S-procedure, and those under the statistical CSI error model were tackled by using the Bernstein-Type Inequality. The reformulated problems were efficiently solved under the AO framework. It is shown that the performance in terms of the minimum achievable transmit power, convergence and complexity under the statistical CSI error model is higher than that under the bounded CSI error model. The number of elements on the IRS may have a negative impact on system performance when the CBIUT error is large. This conclusion provides an engineering insight for the careful selection of the size of the IRS. In the end, this work provides a framework of robust transmission design in a simple single-cell multiuser scenario. The more complicated scenarios, such as the IRSs-assisted full-duplex communication systems, IRS-aided energy efficiency systems and IRS-aided physical layer security systems, will be studied as our future work. Furthermore, the robustness of the IRS in millimeter wave system under a geometric channel model is also worth studying.

Appendix A
The Proof of Lemma 3

Let x be a complex scalar variable, we have the first-order Taylor inequality

$$|x|^2 \geq 2\mathrm{Re}\left\{x^{*,(n)}x\right\} - x^{*,(n)}x^{(n)}, \tag{51}$$

for any fixed point $x^{(n)}$. By replacing x and $x^{(n)}$ in (51) with $(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k)\mathbf{f}_k$ and $(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\mathbf{G}_k)\mathbf{f}_k^{(n)}$, respectively, we have

$$\begin{aligned}
&\left|\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k\right)\mathbf{f}_k\right|^2 \\
&\geq 2\mathrm{Re}\left\{\left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\mathbf{G}_k\right)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}\left(\mathbf{h}_k + \mathbf{G}_k^{\mathrm{H}}\mathbf{e}\right)\right\} \\
&\quad - \left(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\mathbf{G}_k\right)\mathbf{f}_k^{(n)}\mathbf{f}_k^{(n),\mathrm{H}}\left(\mathbf{h}_k + \mathbf{G}_k^{\mathrm{H}}\mathbf{e}^{(n)}\right).
\end{aligned} \tag{52}$$

By plugging $\mathbf{G}_k = \widehat{\mathbf{G}}_k + \triangle\mathbf{G}_k$ into the right hand side of (52) and expanding it by using mathematical transformations, i.e., $\mathrm{Tr}(\mathbf{A}^{\mathrm{H}}\mathbf{B}) = \mathrm{vec}^{\mathrm{H}}(\mathbf{A})\mathrm{vec}(\mathbf{B})$ and $\mathrm{Tr}(\mathbf{ABCD}) = (\mathrm{vec}^{\mathrm{T}}(\mathbf{D}))^{\mathrm{T}}(\mathbf{C}^{\mathrm{T}} \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$ [29], we can obtained (10).

Hence, the proof is completed.

Appendix B
The Proof of Lemma 4

The lower bound of (26) can also be derived from (52) under the full channel uncertainty. In particular, we insert $\mathbf{h}_k = \widehat{\mathbf{h}}_k + \triangle\mathbf{h}_k$ and $\mathbf{G}_k = \widehat{\mathbf{G}}_k + \triangle\mathbf{G}_k$ into the first term on the right hand side of (52), and then get (53) at the top of the next page.

With the similar mathematical transformations, the remaining two terms on the right hand side of (52) under the full channel uncertainty can be expressed as

$$\begin{aligned}
&(\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{\mathrm{H}}\mathbf{G}_k)\mathbf{f}_k\mathbf{f}_k^{(n),\mathrm{H}}(\mathbf{h}_k + \mathbf{G}_k^{\mathrm{H}}\mathbf{e}^{(n)}) \\
&= \widetilde{\mathbf{i}}_k^{\mathrm{H}}\mathbf{D}_k^{\mathrm{H}}\widetilde{\mathbf{i}}_k + \mathbf{d}_{2,k}^{\mathrm{H}}\widetilde{\mathbf{i}}_k + \widetilde{\mathbf{i}}_k^{\mathrm{H}}\mathbf{d}_{1,k} + d_k^* \\
&\quad + (\mathbf{h}_k^{\mathrm{H}} + \mathbf{e}^{(n),\mathrm{H}}\mathbf{G}_k)\mathbf{f}_k^{(n)}\mathbf{f}_k^{(n),\mathrm{H}}(\mathbf{h}_k + \mathbf{G}_k^{\mathrm{H}}\mathbf{e}^{(n)}) \\
&= \widetilde{\mathbf{i}}_k^{\mathrm{H}}\mathbf{Z}_k\widetilde{\mathbf{i}}_k + \mathbf{z}_k^{\mathrm{H}}\widetilde{\mathbf{i}}_k + \widetilde{\mathbf{i}}_k^{\mathrm{H}}\mathbf{z}_k + z_k.
\end{aligned} \tag{54}$$

Hence, the proof is completed.

Appendix C
The Proof of Theorem 1

Denote by $\widehat{\mathbf{\Gamma}}^\star = [\widehat{\mathbf{\Gamma}}_1^\star, \ldots, \widehat{\mathbf{\Gamma}}_K^\star]$ the optimal solution of the relaxed version of Problem (41) and define the projection matrices as $\mathbf{P}_k = \widehat{\mathbf{\Gamma}}_k^{\star\frac{1}{2}}\widehat{\mathbf{h}}_k\widehat{\mathbf{h}}_k^{\mathrm{H}}\widehat{\mathbf{\Gamma}}_k^{\star\frac{1}{2}}/||\widehat{\mathbf{\Gamma}}_k^{\star\frac{1}{2}}\widehat{\mathbf{h}}_k||^2, \forall k \in \mathcal{K}$, where $\widehat{\mathbf{h}}_k = (\mathbf{h}_k + \mathbf{G}_k^{\mathrm{H}}\mathbf{e})$. Then, we construct a rank-one solution $\widetilde{\mathbf{\Gamma}}^\star = [\widetilde{\mathbf{\Gamma}}_1^\star, \ldots, \widetilde{\mathbf{\Gamma}}_K^\star]$, each sub-matrix of which is given by

$$\widetilde{\mathbf{\Gamma}}_k^\star = \widehat{\mathbf{\Gamma}}_k^{\star\frac{1}{2}}\mathbf{P}_k\widehat{\mathbf{\Gamma}}_k^{\star\frac{1}{2}}. \tag{55}$$

$$
\begin{aligned}
&\left[(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\triangle\mathbf{h}_k^{\mathrm{H}})+\mathbf{e}^{(n),\mathrm{H}}(\widehat{\mathbf{G}}_k+\triangle\mathbf{G}_k)\right]\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}\left[(\widehat{\mathbf{h}}_k+\triangle\mathbf{h}_k)+(\widehat{\mathbf{G}}_k^{\mathrm{H}}+\triangle\mathbf{G}_k^{\mathrm{H}})\mathbf{e}\right]\\
&=(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}(\widehat{\mathbf{h}}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e})+(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}(\triangle\mathbf{h}_k+\triangle\mathbf{G}_k^{\mathrm{H}}\mathbf{e})\\
&\quad+(\triangle\mathbf{h}_k^{\mathrm{H}}+\mathbf{e}^{(n),\mathrm{H}}\triangle\,\mathbf{G}_k)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}(\widehat{\mathbf{h}}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e})+(\triangle\mathbf{h}_k^{\mathrm{H}}+\mathbf{e}^{(n),\mathrm{H}}\triangle\,\mathbf{G}_k)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}(\triangle\mathbf{h}_k+\triangle\mathbf{G}_k^{\mathrm{H}}\mathbf{e})\\
&=(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}(\widehat{\mathbf{h}}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e})+(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}\triangle\,\mathbf{h}_k\\
&\quad+\mathrm{vec}^{\mathrm{H}}(\triangle\mathbf{G}_k)\mathrm{vec}(\mathbf{e}(\widehat{\mathbf{h}}_k^{\mathrm{H}}+\mathbf{e}^{(n),\mathrm{H}}\widehat{\mathbf{G}}_k)\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}})+\triangle\mathbf{h}_k^{\mathrm{H}}\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}(\widehat{\mathbf{h}}_k+\widehat{\mathbf{G}}_k^{\mathrm{H}}\mathbf{e})+\triangle\mathbf{h}_k^{\mathrm{H}}\mathbf{f}_k^{(n)}\mathbf{f}_k^{\mathrm{H}}\triangle\,\mathbf{h}_k\\
&\quad+\mathrm{vec}^{\mathrm{H}}(\mathbf{e}^{(n)}(\widehat{\mathbf{h}}_k+\mathbf{e}^{\mathrm{H}}\widehat{\mathbf{G}}_k^{\mathrm{H}})\mathbf{f}_k\mathbf{f}_k^{(n),\mathrm{H}})\mathrm{vec}(\triangle\mathbf{G}_k)\ +\mathrm{vec}^{\mathrm{H}}(\triangle\mathbf{G}_k)(\mathbf{f}_k^{*}\mathbf{f}_k^{(n),\mathrm{T}}\otimes\mathbf{e})\triangle\mathbf{h}_k^{*}\\
&\quad+\triangle\mathbf{h}_k^{\mathrm{T}}(\mathbf{f}_k^{*}\mathbf{f}_k^{(n),\mathrm{T}}\otimes\mathbf{e}^{(n),\mathrm{H}})\mathrm{vec}(\triangle\mathbf{G}_k)+\mathrm{vec}^{\mathrm{H}}(\triangle\mathbf{G}_k)(\mathbf{f}_k^{*}\mathbf{f}_k^{(n),\mathrm{T}}\otimes\mathbf{e}\mathbf{e}^{(n),\mathrm{H}})\mathrm{vec}(\triangle\mathbf{G}_k)\\
&=\widetilde{\mathbf{i}}_k^{\mathrm{H}}\mathbf{D}_k\widetilde{\mathbf{i}}_k+\mathbf{d}_{1,k}^{\mathrm{H}}\widetilde{\mathbf{i}}_k+\widetilde{\mathbf{i}}_k^{\mathrm{H}}\mathbf{d}_{2,k}+d_k. \qquad (53)
\end{aligned}
$$

Firstly, we check the objective value of Problem (41) with solution $\widetilde{\mathbf{\Gamma}}^{\star}$:

$$
\begin{aligned}
&\sum_{k=1}^{K}\mathrm{Tr}\left\{\widetilde{\mathbf{\Gamma}}_k^{\star}\right\}-\sum_{k=1}^{K}\mathrm{Tr}\left\{\widehat{\mathbf{\Gamma}}_k^{\star}\right\}\quad=\sum_{k=1}^{K}\mathrm{Tr}\left\{\widehat{\mathbf{\Gamma}}_k^{\star\frac{1}{2}}(\mathbf{P}_k-\mathbf{I})\widehat{\mathbf{\Gamma}}_k^{\star\frac{1}{2}}\right\}\\
&\le 0, \qquad (56)
\end{aligned}
$$

which means that the objective value acheived by using the solution $\widetilde{\mathbf{\Gamma}}^{\star}$ is no more than that generated from the optimal solution $\widehat{\mathbf{\Gamma}}^{\star}$.

Then, since it is computationally intractable to check whether the constructed solution satisfies the constraints (41b)–(41e) directly, we instead consider the constraint (32b) in the original Problem (32). Specifically, from (34), we have

$$
\frac{\widehat{\mathbf{h}}_k^{\mathrm{H}}\widetilde{\mathbf{\Gamma}}_k^{\star}\widehat{\mathbf{h}}_k}{(2^{R_k}-1)}=\frac{|\widehat{\mathbf{h}}_k^{\mathrm{H}}\widehat{\mathbf{\Gamma}}_k^{\star}\widehat{\mathbf{h}}_k|^2}{||\widehat{\mathbf{\Gamma}}_k^{\star\frac{1}{2}}\widehat{\mathbf{h}}_k||^2(2^{R_k}-1)}=\frac{\widehat{\mathbf{h}}_k^{\mathrm{H}}\widehat{\mathbf{\Gamma}}_k^{\star}\widehat{\mathbf{h}}_k}{(2^{R_k}-1)}, \qquad (57)
$$

as well as

$$
\begin{aligned}
\widehat{\mathbf{h}}_k^{\mathrm{H}}\widetilde{\mathbf{\Gamma}}_i^{\star}\widehat{\mathbf{h}}_k&=\widehat{\mathbf{h}}_i^{\mathrm{H}}\widehat{\mathbf{\Gamma}}_i^{\star\frac{1}{2}}\frac{\widehat{\mathbf{\Gamma}}_i^{\star\frac{1}{2}}\widehat{\mathbf{h}}_k\widehat{\mathbf{h}}_k^{\mathrm{H}}\widehat{\mathbf{\Gamma}}_i^{\star\frac{1}{2}}}{||\widehat{\mathbf{\Gamma}}_i^{\star\frac{1}{2}}\widehat{\mathbf{h}}_i||^2}\widehat{\mathbf{\Gamma}}_i^{\star\frac{1}{2}}\widehat{\mathbf{h}}_i\\
&\le\lambda_{max}\left(\widehat{\mathbf{\Gamma}}_i^{\star\frac{1}{2}}\widehat{\mathbf{h}}_k\widehat{\mathbf{h}}_k^{\mathrm{H}}\widehat{\mathbf{\Gamma}}_i^{\star\frac{1}{2}}\right)=\widehat{\mathbf{h}}_k^{\mathrm{H}}\widehat{\mathbf{\Gamma}}_i^{\star}\widehat{\mathbf{h}}_k. \qquad (58)
\end{aligned}
$$

Combining (57) with (58), we have

$$
\begin{aligned}
&\widehat{\mathbf{h}}_k^{\mathrm{H}}[\widetilde{\mathbf{\Gamma}}_k^{\star}/(2^{R_k}-1)-\sum_{i\neq k}^{K}\widetilde{\mathbf{\Gamma}}_k^{\star}]\widehat{\mathbf{h}}_k\\
&\ge\widehat{\mathbf{h}}_k^{\mathrm{H}}[\widehat{\mathbf{\Gamma}}_k^{\star}/(2^{R_k}-1)-\sum_{i\neq k}^{K}\widehat{\mathbf{\Gamma}}_i^{\star}]\widehat{\mathbf{h}}_k, \qquad (59)
\end{aligned}
$$

which implies that the constructed solution $\widetilde{\mathbf{\Gamma}}^{\star}$ satisfies constraint (32b) and then satisfies constraints (41b)–(41e).

With (56) and (59), we conclude that $\widetilde{\mathbf{\Gamma}}^{\star}$ is also a feasible solution of the relaxed version of Problem (41) with rank one.

Hence, the proof is completed.

REFERENCES

[1] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.

[2] M. Di Renzo *et al.*, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 798–807, 2020.

[3] X. Yuan, Y.-J. Zhang, Y. Shi, W. Yan, and H. Liu, "Reconfigurable-intelligent-surface empowered wireless communications: Challenges and opportunities," 2020. [Online]. Available: https://arxiv.org/abs/2001.00364

[4] C. Pan *et al.*, "Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1719–1734, Aug. 2020.

[5] C. Pan *et al.*, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.

[6] T. Bai, C. Pan, Y. Deng, M. Elkashlan, and A. Nallanathan, "Latency minimization for intelligent reflecting surface aided mobile edge computing," *IEEE J. Sel. Areas Commun.*, to be published.

[7] H. Han *et al.*, "Intelligent reflecting surface aided network: Power control for physical-layer broadcasting," 2019. [Online]. Available: https://arxiv.org/abs/1910.14383

[8] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast MISO communication systems," *IEEE Trans. Signal Process.*, vol. 68, pp. 3236–3251, 2020.

[9] X. Yu, D. Xu, and R. Schober, "Enabling secure wireless communications via intelligent reflecting surfaces," in *Proc. IEEE GLOBECOM*, Dec. 2019, pp. 1–6.

[10] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1488–1492, Sep. 2019.

[11] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1823–1838, Aug. 2020.

[12] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," 2019. [Online]. Available: https://arxiv.org/abs/1904.10136

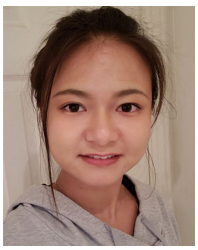
[13] Z. Zhou, N. Ge, Z. Wang, and L. Hanzo, "Joint transmit precoding and reconfigurable intelligent surface phase adjustment: A decomposition-aided channel estimation approach," 2019. [Online]. Available: https://www.researchgate.net/publication/337824343

[14] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, to be published.

[15] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation and joint beamforming for intelligent reflecting surface-assisted millimeter wave systems," 2019. [Online]. Available: https://arxiv.org/abs/1911.07202

[16] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems," 2019. [Online]. Available: https://arxiv.org/abs/1912.03619

[17] G. Zhou, C. Pan, H. Ren, K. Wang, M. Di Renzo, and A. Nallanathan, "Robust beamforming design for intelligent reflecting surface aided MISO communication systems," *IEEE Wireless Commun. Lett.*, to be published.
[18] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," *IEEE J. Sel. Areas Commun.*, to be published.
[19] M. Botros and T. N. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service constraints," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 714–724, Dec. 2007.
[20] J. Zhang, M. Kountouris, J. G. Andrews, and R. W. Heath, "Multimode transmission for the MIMO broadcast channel with imperfect channel state information," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 803–814, Mar. 2011.
[21] T. Lipp and S. Boyd, "Variations and extension of the convex-concave procedure," *Optim. Eng.*, vol. 17, no. 2, pp. 263–287, 2016.
[22] S. Boyd, L. G. El, E. Ferron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA, USA: SIAM, 1994.
[23] I. R. Petersen, "A stabilization algorithm for a class of uncertain linear systems," *Syst. Control Lett.*, vol. 17, no. 2, pp. 351–357, 1987.
[24] E. A. Gharavol and E. G. Larsson, "The sign-definiteness lemma and its applications to robust transceiver optimization for multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 238–252, Jan. 2013.
[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[26] M. Grant and S. Boyd, "CVX: MATLAB software for disciplined convex programming," Dec. 2018. [Online]. Available: https://cvxr.com/cvx
[27] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.
[28] K. Wang, A. M. So, T. Chang, W. Ma, and C. Chi, "Outage constrained robust transmit optimization for multiuser miso downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.
[29] X.-D. Zhang, *Matrix Analysis and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2017.
[30] Z. Luo, W. Ma, A. M. 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.
[31] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia, PA, USA: SIAM, 2001.
[32] 3GPP, "Technical specification group radio access network; study on 3D channel model for LTE (release 12)," *TR 36.873 V12.7.0*, Dec. 2017.



Gui Zhou received the B.S. and M.E. degrees from the School of Information and Electronics, Beijing Institute of Technology, Beijing, China, in 2015 and 2019, respectively. She is currently pursuing the Ph.D. degree at the School of electronic Engineering and Computer Science, Queen Mary University of London, U.K. Her major research interests include intelligent reflection surface (IRS) and signal processing.

Cunhua Pan received the B.S. and Ph.D. degrees from the School of Information Science and Engineering, Southeast University, Nanjing, China, in 2010 and 2015, respectively. From 2015 to 2016, he was a Research Associate at the University of Kent, U.K. He held a post-doctoral position at Queen Mary University of London, U.K., from 2016 and 2019, where he is currently a Lecturer.

His research interests mainly include intelligent reflection surface (IRS), machine learning, UAV, Internet of Things, and mobile edge computing. He serves as a TPC member for numerous conferences, such as ICC and GLOBECOM, and the Student Travel Grant Chair for ICC 2019. He also serves as an Editor of IEEE WIRELESS COMMUNICATION LETTERS and IEEE ACCESS.

Hong Ren received the B.S. degree in electrical engineering from Southwest Jiaotong University, Chengdu, China, in 2011, and the M.S. and Ph.D. degrees in electrical engineering from Southeast University, Nanjing, China, in 2014 and 2018, respectively. From 2016 to 2018, she was a Visiting Student with the School of Electronics and Computer Science, University of Southampton, U.K. She is currently a Postdoctoral Scholar with the School of Electronic Engineering and Computer Science, Queen Mary University of London, U.K. Her research interests lie in the areas of communication and signal processing, including green communication systems, cooperative transmission, and cross layer transmission optimization.

Kezhi Wang received the B.E. and M.E. degrees in School of Automation from Chongqing University, China, in 2008 and 2011, respectively. He received the Ph.D. degree in Engineering from the University of Warwick, U.K. in 2015. He was a Senior Research Officer in University of Essex, U.K. Currently, he is a Senior Lecturer with Department of Computer and Information Sciences at Northumbria University, U.K. His research interests include wireless communications and machine learning.

Arumugam Nallanathan (Fellow, IEEE) received the B.Sc. (Eng.) degree in electrical and electronic engineering from the University of Peradeniya, Sri Lanka, in 1991, the Certificate of Postgraduate Studies (CPGS) from the University of Cambridge in 1994, and the Ph.D. degree in electrical and electronic engineering from the University of Hong Kong in 2000.

He is Professor of Wireless Communications and Head of the Communication Systems Research (CSR) group in the School of Electronic Engineering and Computer Science at Queen Mary University of London since September 2017. He was with the Department of Informatics at King's College London from December 2007 to August 2017, where he was a Professor of Wireless Communications from April 2013 to August 2017 and a Visiting Professor from September 2017. He was an Assistant Professor in the Department of Electrical and Computer Engineering, National University of Singapore from August 2000 to December 2007. His research interests include Artificial Intelligence for Wireless Systems, Beyond 5G Wireless Networks, Internet of Things (IoT) and Molecular Communications. He published nearly 500 technical papers in scientific journals and international conferences. He is a co-recipient of the best paper awards presented at the IEEE International Conference on Communications 2016 (ICC'2016), IEEE Global Communications Conference 2017 (GLOBECOM'2017) and IEEE Vehicular Technology Conference 2018 (VTC'2018). He is an IEEE Distinguished Lecturer. He has been selected as a Web of Science Highly Cited Researcher in 2016.

He is an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS and Senior Editor for IEEE WIRELESS COMMUNICATIONS LETTERS. He was an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2006–2011), IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY (2006–2017) and IEEE SIGNAL PROCESSING LETTERS. He served as the Chair for the Signal Processing and Communication Electronics Technical Committee of IEEE Communications Society and Technical Program Chair and member of Technical Program Committees in numerous IEEE conferences. He received the IEEE Communications Society SPCE outstanding service award 2012 and IEEE Communications Society RCC outstanding service award 2014.