

# Northumbria Research Link

Citation: Zhang, Lei, Pan, Cunhua, Wang, Yu, Ren, Hong and Wang, Kezhi (2022) Robust Beamforming Design for Intelligent Reflecting Surface Aided Cognitive Radio Systems with Imperfect Cascaded CSI. IEEE Transactions on Cognitive Communications and Networking, 8 (1). pp. 186-201. ISSN 2372-2045

Published by: IEEE

URL: <https://doi.org/10.1109/tccn.2021.3107510>  
<<https://doi.org/10.1109/tccn.2021.3107510>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/47097/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

# Robust Beamforming Design for Intelligent Reflecting Surface Aided Cognitive Radio Systems with Imperfect Cascaded CSI

Lei Zhang, *Member, IEEE*, Cunhua Pan, *Member, IEEE*,

Yu Wang, *Member, IEEE*, Hong Ren, *Member, IEEE*,

and Kezhi Wang, *Member, IEEE*

## Abstract

In this paper, intelligent reflecting surface (IRS) is introduced to enhance the network performance of cognitive radio (CR) systems. Specifically, we investigate robust beamforming design based on both bounded channel state information (CSI) error model and statistical CSI error model for primary user (PU)-related channels in IRS-aided CR systems. We jointly optimize the transmit precoding (TPC) at the secondary user (SU) transmitter (ST) and phase shifts at the IRS to minimize the ST's total transmit power subject to the quality of service of SUs, the limited interference imposed on the PU and unit-modulus of the reflective beamforming. The successive convex approximation (SCA) method, Schur's complement, General sign-definiteness principle, inverse Chi-square distribution and penalty convex-concave procedure are invoked for dealing with these intricate constraints. The non-convex optimization problems are transformed into several convex subproblems and efficient algorithms are proposed. Simulation results verify the efficiency of the proposed algorithms and reveal the impacts of CSI uncertainties on ST's minimum transmit power and feasibility rate of the optimization problems.

Lei Zhang, Yu Wang, are with the School of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou, Jiangsu, 213001 China, and with National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China, e-mail: {zhlei, yuwang\_edina}@jsut.edu.cn.

Cunhua Pan is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, U.K., e-mail: c.pan@qmul.ac.uk.

Hong Ren is with National Mobile Communications Research Laboratory, Southeast University, Nanjing, China, e-mail: hren@seu.edu.cn

Kezhi Wang is with Department of Computer and Information Sciences, Northumbria University, UK, e-mail: kezhi.wang@northumbria.ac.uk.

Simulation results also show that the number of transmit antennas at the ST and the number of phase shifts at the IRS should be carefully chosen to balance the channel realization feasibility rate and the total transmit power.

### Index Terms

Reconfigurable intelligent surface, intelligent reflecting surface, robust beamforming design, cognitive radio, imperfect channel state information.

## I. INTRODUCTION

As a revolutionary technique, intelligent reflecting surface (IRS) has received extensive attention from both academia and industry since it can enhance both the spectral and energy efficiency of the wireless communication systems through a preprogrammed controller [1]–[4]. IRS is equipped with a large number of elements made of special materials and functioned by adjusting the reflecting coefficients (i.e., phase or amplitude) of the incident radio-frequency (RF) wave and reflecting it passively. The signals reflected by IRS can be added with other signal paths either to increase the signal strength at the desired receiver, or to mitigate the co-channel interference at the unintended users. The existing contributions have demonstrated benefits brought by introducing an IRS into wireless communication systems. For instance, some certain performance metrics such as channel capacity [2], [3], [5]–[9], physical layer security rate [10]–[13], transmission latency and total transmit power [14], [15] are efficiently enhanced by jointly optimizing the active transmit precoding (TPC) at the base station (BS) and the passive beamforming at the IRS.

Another effective technology to enhance spectrum efficiency is cognitive radio (CR) [16]–[21]. In CR systems, the primary user (PU) is defined as a spectrum-licensed user who always has high priority to access the spectrum, while the secondary user (SU) is normally unlicensed but can be allowed to share the spectrum without causing harmful interference to the PU. However, the challenge of CR systems is that the performance improvements for the PU and the SU are conflicting [17], [18], [20]–[22]. In specific, increasing the transmit power at the SU's BS to enhance the signal strength will generate increased interference towards the PU. Fortunately, this bottleneck can be addressed by introducing an IRS into a CR system thanks to the IRS's reconfiguration function that can help enhance the desired signal strength of the

SU and mitigate the co-channel interference to the PU through jointly optimizing the TPC and passive beamforming [23]–[25].

However, the above papers [23]–[25] studied the transmission design based on perfect channel state information (CSI), which are challenging to realize in practice. The reason is that the channels between SUs and PUs are more difficult to estimate due to the uncooperative relationship between them. The channel estimation error is inevitable. For some users or devices which require strict quality of service (QoS) requirements, it is imperative to study the robust beamforming design for the IRS-aided CR systems to guarantee each user's QoS under arbitrary channel estimation error. In IRS-aided CR system, the sensed channels related to the PU can be divided into two branches. The first one is the direct channel spanning from the SU's BS to the PU (BS-PU). The second one consists of two IRS-related channels spanning from the SU's BS to the IRS (BS-IRS) and the IRS to the PU (IRS-PU). In fact, in general IRS-aided systems without CR, there were some contributions on the robust beamforming design based on the assumption that only the channel from the IRS to the user (IRS-user) was imperfectly estimated [7], [15], [26]. However, it is very challenging to estimate the BS-IRS channel and IRS-user channel independently since the IRS is passive and can neither send nor receive pilot symbols. Installing some active elements at the IRS will increase an undesired burden of the IRS due to the increased hardware and extra power cost. Additionally, the extra information exchange overhead is required to feed back from the IRS to the BS. Therefore, another approach to design the robust beamforming is based on the assumption that the two IRS-related channels are regarded as a cascaded BS-IRS-user channel, which is the product of the BS-IRS channel and the IRS-user channel. It is more cost-effective to estimate the cascaded BS-IRS-user channel since no active RF chains are required at the IRS. It has been verified that considering cascaded BS-IRS-user channel is sufficient for the beamforming design [27]–[30]. Based on the cascaded channel, there are some contributions on robust beamforming design in IRS-aided systems [31]. In [31], a framework of robust transmission beamforming was proposed based on imperfect cascaded IRS-related channels at the transmitter. The worst-case and outage probability robust beamforming designs were provided by minimizing the total transmit power. However, the above robust beamforming designs are not applicable for the CR network since the architecture and optimization problem of IRS-aided CR systems are completely different from traditional CR systems.

To the best of our knowledge, only a few contributions studied transmission beamforming

design based on imperfect PU-related channels for IRS-aided CR systems [32], [33]. The authors in [32] studied the robust beamforming design under bounded CSI uncertainty of PU-related channels in IRS-aided full-duplex CR systems with the aim of maximizing the system sum rate of SUs. However, they considered the imperfect IRS-PU channel instead of the imperfect cascaded BS-IRS-PU channel. In [33], the cascaded BS-IRS-PU channel was first considered to be imperfect with the assumption of the bounded CSI error model, based on which the authors proposed a robust beamforming design to maximize the single user's data rate. The simulation results therein showed that the SU's achievable rate is significantly improved in an IRS-aided multi-input single-output (MISO) CR system compared with a CR system without IRS. However, the above mentioned contributions that targeted at maximizing the capacity cannot guarantee the quality of service (QoS) requirements of each SU, which cannot be applied in some emerging applications with stringent QoS requirements such as video conferencing, autonomous vehicles, etc.

Against the above background, in this paper, we investigate robust beamforming design aiming at minimizing the total transmit power of SUs subject to each SU's QoS requirement and the interference limit imposed on the PU. The imperfect PU-related cascaded channel is considered to avoid the requirement of additional RF chains at the IRS. In contrast to the rate maximization problem in [32] and [33], the power minimization problem may be infeasible due to the conflicting constraints of SU's QoS requirements and PU's limited interference imposed by the SU. Different from [32] and [33] where only the bounded CSI error model is considered, in this paper, we additionally consider another type of CSI error, i.e., statistical CSI error, which is closely relevant to the channel estimation error. Specifically, the contributions of this paper can be summarized as follows:

- The robust beamforming design under both the bounded and statistical CSI error models are studied for the IRS-aided CR system. The TPC matrix at the SU transmitter (ST) and the reflective element diagonalized (RED) matrix of the IRS are jointly optimized to minimize the total transmit power at the ST subject to the unit-modulus constraint of reflective elements, the QoS requirement of each SU receiver (SR) and the interference limit imposed on the PU receiver (PR). The power minimization problem may be infeasible due to the conflicting constraints of SU's QoS requirements and PU's limited interference imposed by the SU. This motivate us to check the feasibility of optimization problems and analyze the feasibility for each channel realization. The block coordinate descent (BCD) algorithm

is employed to alternately optimize the TPC matrix and the RED matrix. Meanwhile, the feasibility checking for optimization problems is analyzed.

- For the bounded CSI error model, we consider the worst-case transmission beamforming design, where the channel error is characterized by two separated PU-related channels. In order to deal with these tricky constraints, we develop artful mathematical derivations. The successive convex approximation (SCA) method, Schur's complement, General sign-definiteness principle and penalty CCP method are adopted to transform the non-convex problems into second-order cone programming (SOCP) problems. This scheme is named as SCD scheme since which is based on the separating cascaded channel and direct channel.
- For the statistical CSI error model, the CSI error follows the circularly symmetric complex Gaussian (CSCG) distribution. This scheme named as STA also aims to minimize the total transmit power at the ST. The inverse Chi-square distribution is used to simplify the probabilistic constraint of the interference limit imposed on the PR. Finally, the problem for optimizing TPC matrix is transformed into an SDP problem and the problem for optimizing RED matrix is transformed into an SOCP problem.
- Some important results are obtained. With the assistance of the IRS, the number of phase shifts should be carefully chosen to obtain a tradeoff between the total minimum transmit power and the feasibility rate of the optimization problem. Moreover, improving the uncertainty level of the PU-related channels can reduce the ST's total transmit power, while a high uncertainty level will lead to a low probability for finding the optimal beamforming.

The remainder of this paper is organized as follows. Section II describes the system model and gives the problem formulation. Section III provides the SCD robust design based on bounded error model. The STA robust design based on statistical error model is provided in Section IV. The feasibility checking problems are analyzed in Section V. Section VI extends one PR scenario to multiple PRs scenario. Section VII shows the simulation results. Finally, we conclude this paper in Section VIII.

*Notations:* The symbols such as  $\mathbf{A}$  and  $\mathbf{a}$  are complex-valued matrix and vector, respectively.  $\mathbb{C}$  and  $\mathbb{C}^{a \times b}$  are complex value set and space of  $a \times b$  complex-valued matrix, respectively.  $\text{diag}\{\cdot\}$  is diagonalization operation.  $\mathbb{E}[\cdot]$  represents mathematical expectation.  $\mathbf{A}^*$ ,  $\mathbf{A}^T$  and  $\mathbf{A}^H$  mean the conjugate, transpose and Hermitian of  $\mathbf{A}$ .  $\mathcal{CN}(0, \sigma)$  is the CSCG distribution with zero mean and variance  $\sigma$ .  $|\cdot|$ ,  $\|\cdot\|_2$  and  $\|\cdot\|_F$  represent modulo, Euclidean norm and Frobenius norm operations, respectively.  $\mathbf{I}_a$  is the  $a \times a$  unitary matrix.  $\text{vec}(\mathbf{A})$ ,  $\text{Tr}(\mathbf{A})$  and  $\text{Re}(\mathbf{A})$  mean

TABLE I  
EXPLANATIONS OF ABBREVIATIONS.

Abbreviations	Explanations	Abbreviations	Explanations
TPC	Transmit precoding	RED	Reflective element diagonalize
ST	Secondary user's transmitter	SR	Secondary user's receiver
PR	Primary user's receiver	BCD	Block coordinate descent
SCD	Separating cascaded channel and direct channel	SDP	Semi-definite programming
CCP	Convex-concave procedure	STA	Statistical CSI error
IT	Interference temperature	SOCP	Second-order cone programming

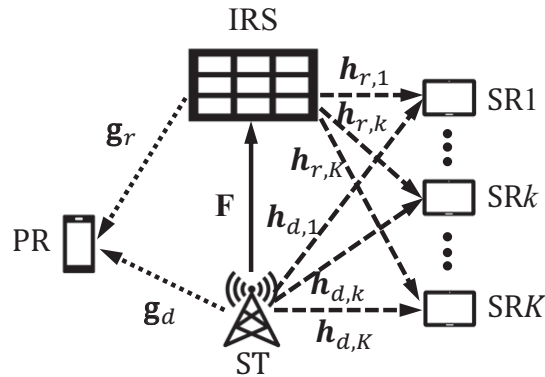


Fig. 1. IRS-aided CR.

vectorization, trace and extracting the real part of  $\mathbf{A}$ , respectively. There are lots of abbreviations in this paper. For the sake of readability, these abbreviations are listed in Table I.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

### A. Signal Transmission Model

In order to improve the performance of SUs, an IRS is deployed in the CR system shown in Fig. 1 where we consider the downlink MISO transmission. The system consists of one IRS, one PR, one ST and  $K$  SRs, where the superimposed signals of  $K$  SUs are transmitted from the ST. The ST is equipped with  $M_t$  transmit antennas and the PR (or each SR) is equipped with a single receive antenna. Each SR will receive the encoded signals via the IRS or directly from the ST and then decode its own signal. Similarly, the PR will also receive the interference signals via the IRS or from the ST directly. The IRS which is equipped with  $N$  reflective elements can receive the transmitted signals and passively reflect them without additional RF electric circuit.

Each reflective element is denoted by  $\phi_n = e^{j\theta_n}$ ,  $n \in \mathcal{N} = \{1, 2, \dots, N\}$ <sup>1</sup>, where  $j$  is the imaginary unit,  $\theta_n \in [0, 2\pi]$  is the phase shift of the  $n$ th element and thus  $\phi_n$  has unit modulus, i.e.,  $|\phi_n| = 1$ .  $\Phi = \text{diag}\{\phi_1, \phi_2, \dots, \phi_N\}$  is the RED matrix. By appropriately tuning the phase shifts of the reflective elements of the IRS, the interference imposed on the PR can be mitigated whereas the useful signal received at the SR can be strengthened [23]. Denote the channels ST-IRS, ST-PR and ST-SR  $k$  (SR  $k$  is the  $k$ th SR,  $k \in \mathcal{K} = \{1, 2, \dots, K\}$ ) by  $\mathbf{F} \in \mathbb{C}^{N \times M_t}$ ,  $\mathbf{g}_d \in \mathbb{C}^{M_t \times 1}$  and  $\mathbf{h}_{d,k} \in \mathbb{C}^{M_t \times 1}$ , respectively. The reflective channels IRS-PR and IRS-SR  $k$  are denoted by  $\mathbf{g}_r \in \mathbb{C}^{N \times 1}$  and  $\mathbf{h}_{r,k} \in \mathbb{C}^{N \times 1}$ , respectively. Note that the channels  $\mathbf{g}_d$  and  $\mathbf{g}_r$  are related to the PR. The PR-related channels  $\mathbf{g}_d$  and  $\mathbf{g}_r$  are imperfectly estimated, since the CSI of the PR-related channels, including those from the ST to the IRS and from the IRS to the PRs, is usually more difficult to acquire than that of SR-related channels due to the lack of cooperation between licensed users and unlicensed users in CR systems [34]. In this paper, we assumed that the underlay spectrum access scheme is adopted. Based on that, practical energy detection is assumed to be used for spectrum sensing individually by SRs. To alleviate the hidden terminal problem, cooperative spectrum sensing is proposed in which measurements of multiple spectrum sensors are combined to make a final decision on the existence of a primary signal.

The desired signal of SR  $k$  is denoted by  $s_k \in \mathbb{C}$  satisfying  $\mathbb{E}[s_k s_k^*] = 1$  and  $\mathbb{E}[s_i s_j^*] = 0$  ( $i \neq j$ ), which has a corresponding TPC vector  $\mathbf{w}_k \in \mathbb{C}^{M_t \times 1}$ . The TPC matrix is denoted by  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_K] \in \mathbb{C}^{M_t \times K}$ . Then, the transmit signal from the ST can be written as  $\mathbf{x} = \sum_{k=1}^K \mathbf{w}_k s_k$ . The received signal at SR  $k$  is  $y_k = (\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \Phi \mathbf{F}) \mathbf{x} + n_{s_k}$ , where  $n_{s_k} \sim \mathcal{CN}(0, \sigma_{s_k}^2)$  is the equivalent noise which captures the joint effect of the thermal noise and the received interference from the primary transmitters [32], [34]. The received interference signal at the PR is given by  $y_p = (\mathbf{g}_d^H + \mathbf{g}_r^H \Phi \mathbf{F}) \mathbf{x} + n_p$ , where  $n_p \sim \mathcal{CN}(0, \sigma_p^2)$ . Hence, the signal to interference plus noise ratio (SINR) of SR  $k$  is

$$SINR_k = \frac{|(\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \Phi \mathbf{F}) \mathbf{w}_k|^2}{\|(\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \Phi \mathbf{F}) \mathbf{W}_{-k}\|_2^2 + \sigma_{s_k}^2} = \frac{|(\mathbf{h}_{d,k}^H + \phi^H \text{diag}(\mathbf{h}_{r,k}^H) \mathbf{F}) \mathbf{w}_k|^2}{\|(\mathbf{h}_{d,k}^H + \phi^H \text{diag}(\mathbf{h}_{r,k}^H) \mathbf{F}) \mathbf{W}_{-k}\|_2^2 + \sigma_{s_k}^2}, \quad (1)$$

where  $\phi = [\phi_1, \phi_2, \dots, \phi_N]^T$ , and  $\mathbf{W}_{-k} = [\mathbf{w}_1, \dots, \mathbf{w}_{k-1}, \mathbf{w}_{k+1}, \dots, \mathbf{w}_K]$ .

The interference temperature (IT) imposed on the PR from the ST is

$$IT = \|(\mathbf{g}_d^H + \mathbf{g}_r^H \Phi \mathbf{F}) \mathbf{W}\|_2^2 = \|(\mathbf{g}_d^H + \phi^H \mathbf{G}_r) \mathbf{W}\|_2^2 = \|\tilde{\phi}^H \mathbf{G} \mathbf{W}\|_2^2, \quad (2)$$

<sup>1</sup>Here, the amplitude is set to 1 which is usually given to maximize the signal reflection strength.



where  $\mathbf{G}_r = \text{diag}(\mathbf{g}_r^H)\mathbf{F} \in \mathbb{C}^{N \times M_t}$  is regarded as the cascaded ST-IRS-PR channel,  $\mathbf{G} = [\mathbf{G}_r^H \ \mathbf{g}_d]^H \in \mathbb{C}^{(N+1) \times M_t}$  is an equivalent combined channel integrating  $\mathbf{g}_d$  and  $\mathbf{G}_r$ , and  $\tilde{\boldsymbol{\phi}} = [\phi_1, \phi_2, \dots, \phi_N, 1]^T \in \mathbb{C}^{(N+1) \times 1}$ .

### B. Channel Uncertainty Models

The channel uncertainty is caused by the imperfect estimation of the PR-related channels. If the direct channel  $\mathbf{g}_d$  and the cascaded ST-IRS-PR channel  $\mathbf{G}_r$  are separately estimated, the channels can be modeled as

$$\mathbf{g}_d = \hat{\mathbf{g}}_d + \Delta \mathbf{g}_d, \mathbf{G}_r = \hat{\mathbf{G}}_r + \Delta \mathbf{G}_r, \quad (3)$$

where  $\hat{\mathbf{g}}_d$  and  $\hat{\mathbf{G}}_r$  are estimated CSIs for the direct channel  $\mathbf{g}_d$  and cascaded channel  $\mathbf{G}_r$ , respectively.  $\Delta \mathbf{g}_d$  and  $\Delta \mathbf{G}_r$  are corresponding CSI errors. If we substitute the channel estimations of (3) into the equivalent combined channel, i.e.,  $\mathbf{G}$ , then we have

$$\mathbf{G} = \hat{\mathbf{G}} + \Delta \mathbf{G}, \quad (4)$$

where  $\hat{\mathbf{G}} = [\hat{\mathbf{G}}_r^H \ \hat{\mathbf{g}}_d]^H$  is regarded as the estimated combined CSI at the ST,  $\Delta \mathbf{G} = [\Delta \mathbf{G}_r^H \ \Delta \mathbf{g}_d]^H$  is regarded as the combined CSI error matrix. In this paper, two different types of error models are investigated to describe the above CSI errors.

1) *Bounded CSI Error Model*: In this model, the CSI errors of the direct channel and cascaded channel are assumed to be bounded in the region as follows:

$$\|\Delta \mathbf{g}_d\|_2 \leq \epsilon_d, \|\Delta \mathbf{G}_r\|_F \leq \epsilon_r, \quad (5)$$

where  $\epsilon_d$  and  $\epsilon_r$  are the radii of the bounded regions of CSI errors.

2) *Statistical CSI Error Model*: In this model, the CSI error vectors of  $\Delta \mathbf{G}_r$  and  $\Delta \mathbf{g}_d$  are assumed to follow the CSCG distributions with zero mean and covariance matrices of  $\boldsymbol{\Sigma}_{g_r}$  and  $\boldsymbol{\Sigma}_{g_d}$ , i.e.,

$$\Delta \mathbf{g}_d \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_{g_d}), \text{vec}(\Delta \mathbf{G}_r) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_{g_r}), \quad (6)$$

where  $\boldsymbol{\Sigma}_{g_d} \in \mathbb{C}^{M_t \times M_t}$  and  $\boldsymbol{\Sigma}_{g_r} \in \mathbb{C}^{NM_t \times NM_t}$  are positive semidefinite matrices. In addition,  $\Delta \mathbf{g}_d$  is independent of  $\text{vec}(\Delta \mathbf{G}_r)$ . According to the relationship of  $\Delta \mathbf{G} = [\Delta \mathbf{G}_r^H \ \Delta \mathbf{g}_d]^H$ , it can be verified that  $\Delta \mathbf{G}$  also follows the CSCG distribution with zero mean and covariance matrix of  $\boldsymbol{\Sigma}$ , i.e.,

$$\text{vec}(\Delta \mathbf{G}) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}), \quad (7)$$

where  $\Sigma = \begin{bmatrix} \Sigma_{g_r} & \mathbf{0} \\ \mathbf{0} & \Sigma_{g_d} \end{bmatrix}$ . That means the model of (6) is equivalent to that of (7). We will only focus on the latter one in this paper when we consider the statistical CSI error model.

The robust beamforming design will be first investigated under SCD scheme based on the bounded CSI error model shown in (5). It is regarded as the worst-case beamforming to guarantee the SINR and the IT requirements for any channel quality. Then we will investigate the robust beamforming design under STA scheme based on the statistical CSI error model shown in (7). This type of robust beamforming is designed to guarantee the outage probability requirement of the PU's transmission.

### III. SCD ROBUST DESIGN BASED ON BOUNDED ERROR MODEL

In this section, we investigate the SCD robust designing scheme based on the error model  $\|\Delta \mathbf{g}_d\|_2 \leq \epsilon_d$  and  $\|\Delta \mathbf{G}_r\|_F \leq \epsilon_r$ . The original main problem is firstly transformed into a deterministic problem by considering the worst-case IT constraint.

#### A. Optimization Problem for SCD Scheme

The problem with the aim of minimizing the ST's total transmit power by optimizing the TPC matrix  $\mathbf{W}$  and the RED matrix  $\Phi$  can be formulated as

$$\mathcal{P}1 \quad \min_{\mathbf{W}, \Phi} \quad \|\mathbf{W}\|_F^2 \quad (8a)$$

$$\text{s.t.} \quad \frac{|(\mathbf{h}_{d,k}^H + \Phi^H \text{diag}(\mathbf{h}_{r,k}^H) \mathbf{F}) \mathbf{w}_k|^2}{\|(\mathbf{h}_{d,k}^H + \Phi^H \text{diag}(\mathbf{h}_{r,k}^H) \mathbf{F}) \mathbf{W}_{-k}\|_2^2 + \sigma_{s_k}^2} \geq \gamma_k, \forall k \in \mathcal{K}, \quad (8b)$$

$$\|(\mathbf{g}_d^H + \Phi^H \mathbf{G}_r) \mathbf{W}\|_2^2 \leq \Gamma, \|\Delta \mathbf{g}_d\|_2 \leq \epsilon_d, \|\Delta \mathbf{G}_r\|_F \leq \epsilon_r, \quad (8c)$$

$$|\phi_n| = 1, \forall n \in \mathcal{N}. \quad (8d)$$

In this problem, the error of the direct channel ST-PR and the error of the cascaded channel ST-BS-IRS are separately considered. The SINR constraint (8b) is non-convex. Fortunately, we can find its successive convex approximation (SCA) version. Then, the real challenge is the bounded CSI error model of PU-related channels which makes the problem more complicated. In this section, we will employ Schur's complement and the General sign-definiteness principle to deal with the bounded CSI error.

By applying the Schur's complement lemma [35], the IT inequality constraint (8c) can be rewritten as the following matrix inequality

$$\begin{bmatrix} \Gamma & \mathbf{b}^H \\ \mathbf{b} & \mathbf{I} \end{bmatrix} \succeq \mathbf{0}, \quad (9)$$

where  $\mathbf{b} = [(\mathbf{g}_d^H + \phi^H \mathbf{G}_r) \mathbf{W}]^H$ . By substituting  $\mathbf{g}_d = \hat{\mathbf{g}}_d + \Delta \mathbf{g}_d$  and  $\mathbf{G}_r = \hat{\mathbf{G}}_r + \Delta \mathbf{G}_r$  into (9), we have

$$\begin{aligned} \begin{bmatrix} \Gamma & \hat{\mathbf{b}}^H \\ \hat{\mathbf{b}} & \mathbf{I} \end{bmatrix} &\succeq - \begin{bmatrix} \mathbf{0} \\ \mathbf{W}^H \end{bmatrix} \Delta \mathbf{G}_r^H [\phi \ \mathbf{0}] - \begin{bmatrix} \phi^H \\ \mathbf{0} \end{bmatrix} \Delta \mathbf{G}_r [\mathbf{0} \ \mathbf{W}] \\ &- \begin{bmatrix} \mathbf{0} \\ \mathbf{W}^H \end{bmatrix} [\Delta \mathbf{g}_d \ \mathbf{0}] - \begin{bmatrix} \Delta \mathbf{g}_d^H \\ \mathbf{0} \end{bmatrix} [\mathbf{0} \ \mathbf{W}], \end{aligned} \quad (10)$$

where  $\hat{\mathbf{b}} = [(\hat{\mathbf{g}}_d^H + \phi^H \hat{\mathbf{G}}_r) \mathbf{W}]^H$ .

*Lemma 1:* (General sign-definiteness principle) For a given set of matrices  $\{\mathbf{Z}, \mathbf{U}_i, \mathbf{V}_i, i = 1, \dots, P\}$ , where  $\mathbf{Z}$  is Hermitian matrix, the following inequality

$$\mathbf{Z} \succeq \sum_{i=1}^P (\mathbf{U}_i^H \mathbf{X}_i \mathbf{V}_i + \mathbf{V}_i^H \mathbf{X}_i^H \mathbf{U}_i), \|\mathbf{X}_i\| \leq \epsilon_i, \forall i, \quad (11)$$

holds if and only if there exist real values  $\rho_i \geq 0, \forall i$  such that

$$\begin{bmatrix} \mathbf{Z} - \sum_{i=1}^P \rho_i \mathbf{V}_i^H \mathbf{V}_i & -\epsilon_1 \mathbf{U}_1^H & \cdots & -\epsilon_P \mathbf{U}_P^H \\ -\epsilon_1 \mathbf{U}_1 & \rho_1 \mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ -\epsilon_P \mathbf{U}_P & \mathbf{0} & \cdots & \rho_P \mathbf{I} \end{bmatrix} \succeq \mathbf{0}. \quad (12)$$

*Proof:* Please refer to [36]. ■

By comparing (11) with (10), we set the following equalities

$$\mathbf{Z} = \begin{bmatrix} \Gamma & \hat{\mathbf{b}}^H \\ \hat{\mathbf{b}} & \mathbf{I} \end{bmatrix}, P = 2, \epsilon_1 = \epsilon_r, \epsilon_2 = \epsilon_d, \mathbf{U}_1^H = \mathbf{U}_2^H = - \begin{bmatrix} \mathbf{0} \\ \mathbf{W}^H \end{bmatrix},$$

$$\mathbf{X}_1 = \Delta \mathbf{G}_r^H, \mathbf{X}_2 = \Delta \mathbf{g}_d, \mathbf{V}_1 = [\phi \ \mathbf{0}], \mathbf{V}_2 = [1 \ \mathbf{0}].$$

According to Lemma 1, the equivalent form of the worst-case IT constraint (10) is

$$\begin{bmatrix} \Gamma - \rho_1 N - \rho_2 & \hat{\mathbf{b}}^H & \mathbf{0}_{1 \times M_t} & \mathbf{0}_{1 \times M_t} \\ \hat{\mathbf{b}} & \mathbf{I}_K & \epsilon_r \mathbf{W}^H & \epsilon_d \mathbf{W}^H \\ \mathbf{0}_{M_t \times 1} & \epsilon_r \mathbf{W} & \rho_1 \mathbf{I}_{M_t} & \mathbf{0} \\ \mathbf{0}_{M_t \times 1} & \epsilon_d \mathbf{W} & \mathbf{0} & \rho_2 \mathbf{I}_{M_t} \end{bmatrix} \succeq \mathbf{0}. \quad (13)$$

Therefore, problem  $\mathcal{P}1$  can be reformulated as

$$\mathcal{P}1' \quad \min_{\mathbf{W}, \Phi, \rho = \{\rho_1, \rho_2\}} \|\mathbf{W}\|_F^2 \quad \text{s.t.} \quad (8b), (13), (8d), \rho \geq 0. \quad (14)$$

Since  $\mathbf{W}$  and  $\Phi$  are coupled in problem  $\mathcal{P}1'$ , the BCD method is used to alternately optimize  $\mathbf{W}$  and  $\Phi$ .

### B. Optimizing $\mathbf{W}$ with Fixed $\Phi$ for SCD Scheme

For convenience, we firstly deal with SINR constraint (8b) when  $\Phi$  is fixed. Then we apply the convex optimization approach to solve this subproblem.

1) *Dealing with SINR constraint (8b)*: After denoting  $\hat{\mathbf{h}}_{\Phi, k}^H = \mathbf{h}_{d, k}^H + \phi^H \text{diag}(\mathbf{h}_{r, k}^H) \mathbf{F}$  and  $\hat{\mathbf{H}}_{\Phi, k} = \hat{\mathbf{h}}_{\Phi, k} \hat{\mathbf{h}}_{\Phi, k}^H$ , we have

$$|(\mathbf{h}_{d, k}^H + \phi^H \text{diag}(\mathbf{h}_{r, k}^H) \mathbf{F}) \mathbf{w}_k|^2 = \mathbf{w}_k^H \hat{\mathbf{H}}_{\Phi, k} \mathbf{w}_k, \quad (15)$$

$$\|(\mathbf{h}_{d, k}^H + \phi^H \text{diag}(\mathbf{h}_{r, k}^H) \mathbf{F}) \mathbf{W}_{-k}\|_2^2 = \hat{\mathbf{h}}_{\Phi, k}^H \mathbf{W}_{-k} \mathbf{W}_{-k}^H \hat{\mathbf{h}}_{\Phi, k} = \sum_{j \neq k}^K \mathbf{w}_j^H \hat{\mathbf{H}}_{\Phi, k} \mathbf{w}_j. \quad (16)$$

By substituting (15) and (16) into (8b), we can rewrite the SINR constraint (8b) as

$$\frac{\mathbf{w}_k^H \hat{\mathbf{H}}_{\Phi, k} \mathbf{w}_k}{\sum_{j \neq k}^K \mathbf{w}_j^H \hat{\mathbf{H}}_{\Phi, k} \mathbf{w}_j + \sigma_{s_k}^2} \geq \gamma_k, \forall k \in \mathcal{K}, \quad (17)$$

which is non-convex constraint. By using the first-order Taylor inequality,  $\mathbf{w}_k^H \hat{\mathbf{H}}_{\Phi, k} \mathbf{w}_k$  can be lower bounded linearly by  $2\text{Re}\{\mathbf{w}_k^{(t)H} \hat{\mathbf{H}}_{\Phi, k} \mathbf{w}_k\} - \mathbf{w}_k^{(t)H} \hat{\mathbf{H}}_{\Phi, k} \mathbf{w}_k^{(t)}$ , where  $\mathbf{w}_k^{(t)}$  is the optimal value of the TPC at the  $t$ th iteration. Then we can construct an SCA version of SINR constraint as

$$2\text{Re}(\mathbf{w}_k^{(t)H} \hat{\mathbf{H}}_{\Phi, k} \mathbf{w}_k) - \gamma_k \sum_{j \neq k}^K \mathbf{w}_j^H \hat{\mathbf{H}}_{\Phi, k} \mathbf{w}_j \geq \tilde{\gamma}_k, \quad (18)$$

where  $\tilde{\gamma}_k = \mathbf{w}_k^{(t)H} \hat{\mathbf{H}}_{\Phi, k} \mathbf{w}_k^{(t)} + \gamma_k \sigma_{s_k}^2$ .

2) *Subproblem to Optimize  $\mathbf{W}$* : When RED matrix is fixed, the SINR constraint (8b) can be replaced by (18), then the subproblem for optimizing  $\mathbf{W}$  is given by

$$\mathcal{P}1.1 \quad \min_{\{\mathbf{w}_k\}, \rho} \sum_{k=1}^K \mathbf{w}_k^H \mathbf{w}_k \quad \text{s.t.} \quad (18), (13), \rho \geq 0. \quad (19)$$

IT constraint (13) is a linear matrix inequality (LMI) when  $\Phi$  is fixed. This problem is an SOCP problem with respect to  $\{\mathbf{w}_k, \forall k\}$ ,  $\rho_1$  and  $\rho_2$ , which can be solved by using CVX.

### C. Optimizing $\Phi$ with Fixed $\mathbf{W}$ for SCD Scheme

When TPC matrix  $\mathbf{W}$  is given, the subproblem degenerates into a feasibility-check problem. We firstly transform the non-convex SINR constraint into a convex form. Then we reformulate the optimization problem to find the optimal  $\Phi$ .

1) *Dealing with SINR constraint (8b)*: Denoting  $\mathbf{H}_{r,k} = \text{diag}(\mathbf{h}_{r,k}^H)\mathbf{F}$ , the constraint (8b) can be rewritten as

$$\phi^H \Omega_k \phi - \gamma_k \phi^H \Omega_{-k} \phi + 2\text{Re}\{\omega_k^H \phi\} + \omega_k \geq \gamma_k \sigma_{s_k}^2, \forall k \in \mathcal{K}, \quad (20)$$

where

$\Omega_k = \mathbf{H}_{r,k} \mathbf{w}_k \mathbf{w}_k^H \mathbf{H}_{r,k}^H$ ,  $\omega_k = \mathbf{h}_{d,k}^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{d,k} - \gamma_k \mathbf{h}_{d,k}^H \sum_{j \neq k}^K (\mathbf{w}_j \mathbf{w}_j^H) \mathbf{h}_{d,k}$ ,  $\Omega_{-k} = \mathbf{H}_{r,k} \sum_{j \neq k}^K (\mathbf{w}_j \mathbf{w}_j^H) \mathbf{H}_{r,k}^H$ , and  $\omega_k = \mathbf{H}_{r,k} \mathbf{w}_k \mathbf{w}_k^H \mathbf{h}_{d,k} - \gamma_k \mathbf{H}_{r,k} \sum_{j \neq k}^K (\mathbf{w}_j \mathbf{w}_j^H) \mathbf{h}_{d,k}$ . As a constraint, (20) is still non-convex.

By using the first-order Taylor inequality,  $\phi^H \Omega_k \phi$  can be lower bounded linearly by  $2\text{Re}\{\phi^{(t)H} \Omega_k \phi\} - \phi^{(t)H} \Omega_k \phi^{(t)}$ . We can construct an SCA version of SINR constraint (20), which is given by

$$\gamma_k \phi^H \Omega_{-k} \phi - 2\text{Re}\{(\omega_k^H + \phi^{(t)H} \Omega_k) \phi\} \leq \bar{\gamma}_k, \quad (21)$$

where  $\phi^{(t)}$  is solution of the  $t$ th iteration and  $\bar{\gamma}_k = \omega_k - \phi^{(t)H} \Omega_k \phi^{(t)} - \gamma_k \sigma_{s_k}^2$ .

2) *Subproblem to Optimize  $\Phi$* : When  $\mathbf{W}$  is fixed, the SINR constraint can be approximated by (21). To deal with this feasibility-check problem, the slack variables  $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_K]$  are introduced to tighten the SINR constraints (20). Then the quadratic inequality (21) is modified as

$$\gamma_k \phi^H \Omega_{-k} \phi - 2\text{Re}\{(\omega_k^H + \phi^{(t)H} \Omega_k) \phi\} \leq \bar{\gamma}_k - \varphi_k, \forall k \in \mathcal{K}. \quad (22)$$

Therefore, for the feasibility-check problem, both the SINR constraint given by (22) and the IT constraint given by (13) are convex. Then the subproblem can be reformulated as

$$\mathcal{P}1.2 \quad \max_{\phi, \varphi} \quad \sum_{k=1}^K \varphi_k \quad \text{s.t.} \quad (22), (13), (8d), \varphi \geq 0. \quad (23)$$

The only non-convexity of problem  $\mathcal{P}1.2$  is from the unit-modulus constraint of  $\phi$  in (8d). Then we adopt the penalty CCP method to deal with this non-convex constraint [31]. According to the penalty CCP principle, the non-convex constraint (8d) can be first equivalently transformed

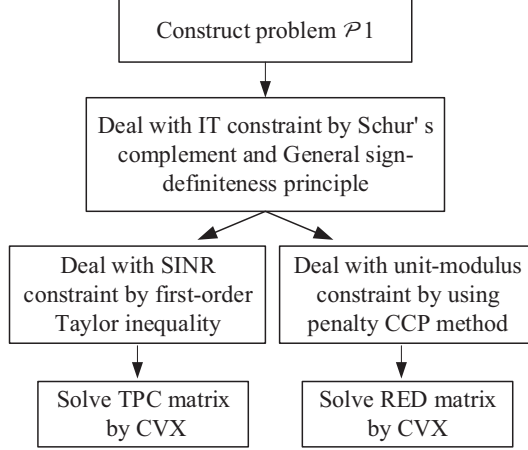


Fig. 2. The key steps and methodologies involved in SCD scheme.

into  $1 \leq |\phi_n|^2 \leq 1$ . The non-convex part can be linearized by  $|\phi_n^{(t)}|^2 - 2\text{Re}(\phi_n^* \phi_n^{(t)}) \leq -1$ . Then, we can reformulate Problem  $\mathcal{P}1.2$  as

$$\mathcal{P}1.2\_1 : \max_{\phi, \varphi, \tau} \sum_{k=1}^K \varphi_k - \kappa \sum_{n=1}^{2N} \tau_n \quad (24a)$$

$$\text{s.t.} \quad (22), (13), \varphi \geq 0, \tau \geq 0. \quad (24b)$$

$$|\phi_n^{(t)}|^2 - 2\text{Re}(\phi_n^* \phi_n^{(t)}) \leq \tau_n - 1, \forall n \in \mathcal{N}, \quad (24c)$$

$$|\phi_n|^2 \leq 1 + \tau_{N+n}, \forall n \in \mathcal{N}, \quad (24d)$$

where  $\tau = [\tau_1, \dots, \tau_{2N}]^T$  are the slack variables.  $\kappa$  is the penalty multiplier to scale the penalty item  $\sum_{n=1}^{2N} \tau_n$  which can control the feasibility of  $\phi$  combining with adjustable  $\kappa$ . This subproblem is an SOCP problem which can be solved by using CVX.

For the convenience of understanding, a flow chart is provided in Fig. 2 to summarize the key steps of the whole optimization processing for SCD scheme.

#### D. Overall Algorithm for SCD Scheme

1) *Convergence analysis:* The overall algorithm for the SCD scheme is provided in Algorithm 1.

The initialization value of  $\mathbf{w}_k^{(0)}$  and  $\phi^{(0)}$  can be obtained from the optimal solution of feasibility checking problem which will be discussed in Section V. In the iteration for updating  $\kappa$ , we set a sufficiently low value  $l_1$  to check whether  $\|\tau\|_1 < l_1$ , which is one condition to stop the

---

**Algorithm 1** Optimizing  $\mathbf{W}$  and  $\Phi$  for SCD
 

---

- 1: Initialize  $\mathbf{w}_k^{(0)}$  and  $\phi^{(0)}$  in feasible region for  $\forall k \in \mathcal{K}$ , set maximum iteration number  $t_{max}$ , target convergence accuracy  $\zeta$ ,  $l_\kappa > 1$ ,  $\kappa_{max}$  and  $t = 0$ .
  - 2: **Repeat**
  - 3: Update  $\mathbf{w}_k^{(t+1)}$  by solving problem  $\mathcal{P}1.1$ ;  $n = 0$ ,  $\phi^{(n)} = \phi^{(t)}$ ;
  - 4: **Repeat**
  - 5: Update  $\phi^{(n+1)}$  by solving problem  $\mathcal{P}1.2\_1$ ;
  - 6:  $\kappa^{(n+1)} = \max\{l_\kappa \kappa^{(n)}, \kappa_{max}\}$ ;
  - 7: **Until**  $\|\boldsymbol{\tau}\|_1 \leq l_1$  and  $\|\phi^{(n+1)} - \phi^{(n)}\|_1 \leq l_2$ ;
  - 8:  $\phi^{(t+1)} = \phi^{(n+1)}$ ;
  - 9: Calculate  $\|\mathbf{W}^{(t+1)}\|_{\mathbb{F}}^2$  from (8a) in problem  $\mathcal{P}1$ ;
  - 10: **Until**  $t > t_{max}$  or  $\frac{\|\mathbf{W}^{(t+1)}\|_{\mathbb{F}}^2 - \|\mathbf{W}^{(t)}\|_{\mathbb{F}}^2}{\|\mathbf{W}^{(t+1)}\|_{\mathbb{F}}^2} < \zeta$ .
- 

iteration.  $\|\boldsymbol{\tau}\|_1 < l_1$  satisfies the constraint (8d) in (23). Another iteration stopping condition is  $\|\phi^{(n+1)} - \phi^{(n)}\|_1 \leq l_2$ , where  $l_2$  is a low value. It can be verified that the sequences of the objective function value produced by Algorithm 1 are guaranteed to converge. For the first subproblem optimizing  $\mathbf{W}$ , we show that the solution sequence  $\{\mathbf{W}^{(t+1)}, t = 1, 2, \dots\}$  is feasible and the objective function value sequence  $\{f(\mathbf{W}^{(t)}) = \|\mathbf{W}^{(t)}\|_{\mathbb{F}}^2, t = 1, 2, \dots\}$  is monotonically decreasing. Define the functions  $g(\mathbf{w}_k) \triangleq \mathbf{w}_k^H \widehat{\mathbf{H}}_{\Phi, k} \mathbf{w}_k$  and  $\bar{g}(\mathbf{w}_k | \mathbf{w}_k^{(t)}) \triangleq -\mathbf{w}_k^{(t)H} \widehat{\mathbf{H}}_{\Phi, k} \mathbf{w}_k^{(t)} + 2\text{Re}\{\mathbf{w}_k^{(t)H} \widehat{\mathbf{H}}_{\Phi, k} \mathbf{w}_k\}$ . Assume that  $\mathbf{W}^{(t+1)}$  is a feasible solution of problem  $\mathcal{P}1.1$ , and thus  $\mathbf{W}^{(t+1)}$  satisfies the constraints (18) and (13). By substituting  $\mathbf{W}^{(t+1)}$  into (18), we have  $\frac{\bar{g}(\mathbf{w}_k^{(t+1)} | \mathbf{w}_k^{(t)})}{\sum_{j \neq k}^K \mathbf{w}_j^{(t+1)H} \widehat{\mathbf{H}}_{\Phi, k} \mathbf{w}_j^{(t+1)} + \sigma_{s_k}^2} \geq \gamma_k, \forall k$ . According to the first-order Taylor inequality, we have  $\bar{g}(\mathbf{w}_k^{(t+1)} | \mathbf{w}_k^{(t)}) \leq g(\mathbf{w}_k^{(t+1)})$ . Thus,  $\frac{g(\mathbf{w}_k^{(t+1)})}{\sum_{j \neq k}^K \mathbf{w}_j^{(t+1)H} \widehat{\mathbf{H}}_{\Phi, k} \mathbf{w}_j^{(t+1)} + \sigma_{s_k}^2} \geq \gamma_k$  holds, which means that the sequence  $\{\mathbf{W}^{(t+1)}, t = 1, 2, \dots\}$  satisfies the constraint (17) that is the SINR constraint (8b). As  $\mathbf{W}^{(t+1)}$  is the globally optimal solution of problem  $\mathcal{P}1.1$ ,  $f(\mathbf{W}^{(t+1)}) \leq f(\mathbf{W}^{(t)})$  holds. Thus, the objective function value sequence  $\{f(\mathbf{W}^{(t)}), t = 1, 2, \dots\}$  is guaranteed to converge. For the second subproblem of optimizing  $\Phi$ , the convergence can be easily verified by using the same analysis approach.

2) *Complexity analysis*: Since the proposed Algorithm 1 involves SOC, LMI and linear constraints which can be solved by a standard interior point method, the general expression

of the computational complexity of which is given by

$$\mathcal{O}\left(\left(\sum_J J_j b_j + 2I\right)^{0.5} n(n^2 + n \sum_J J_j b_j^2 + \sum_J J_j b_j^3 + n \sum_I I_i a_i^2)\right), \quad (25)$$

where we ignore the complexity of linear constraints,  $n$  is the number of variables,  $I_i$  is the number of SOC constraints with the size of  $a_i$  and  $\sum I_i = I$ ,  $J_j$  is the number of LMI constraints with size of  $b_j$  and  $\sum J_j = J$ . Based on the above general expression, the complexity for solving problem  $\mathcal{P}1.1$  is  $o_{\mathbf{w}}^{scd} = \mathcal{O}((2M_t + 3K + 1)^{0.5} M_t K (M_t K^2 + M_t K (2M_t + K + 1)^2 + (2M_t + K + 1)^3 + K^2 M_t^3))$ , the complexity of problem  $\mathcal{P}1.2$  is  $o_{\Phi}^{scd} = \mathcal{O}(\log(1/\min\{l_1, l_2\})(2N + 3K + 1)^{0.5} N(N(K + 1)^2 + N(K + 1)^3 + KN^3 + 2N^2))$ . Thus, the overall complexity for solving problem  $\mathcal{P}1$  is  $o_{\mathbf{w}}^{scd} + o_{\Phi}^{scd}$ .

#### IV. STA ROBUST DESIGN BASED ON STATISTICAL ERROR MODEL

In this section, we optimize the robust beamforming based on statistical CSI error model where the uncertain CSI  $\Delta \mathbf{G}$  satisfies the distribution as shown in (7), i.e.,  $\text{vec}(\Delta \mathbf{G}) \sim \mathcal{CN}(\mathbf{0}, \Sigma)$ . This statistical characteristic is implied in a probabilistic constraint. We first transform the probabilistic model into a tractable form by Chi-square distribution method.

##### A. Optimization Problem for STA Scheme

The power minimization problem considering statistical CSI error can be formulated as

$$\mathcal{P}2: \quad \min_{\mathbf{W}, \Phi} \quad \|\mathbf{W}\|_F^2 \quad (26a)$$

$$\text{s.t.} \quad \frac{|(\mathbf{h}_{d,k}^H + \Phi^H \text{diag}(\mathbf{h}_{r,k}^H) \mathbf{F}) \mathbf{w}_k|^2}{\|(\mathbf{h}_{d,k}^H + \Phi^H \text{diag}(\mathbf{h}_{r,k}^H) \mathbf{F}) \mathbf{W}_{-k}\|_2^2 + \sigma_{s_k}^2} \geq \gamma_k, \forall k \in \mathcal{K}, \quad (26b)$$

$$\Pr \left\{ \left\| (\tilde{\Phi}^H \hat{\mathbf{G}} + \tilde{\Phi}^H \Delta \mathbf{G}) \mathbf{W} \right\|_2^2 \leq \Gamma \right\} \geq 1 - \beta, \text{vec}(\Delta \mathbf{G}) \sim \mathcal{CN}(\mathbf{0}, \Sigma), \quad (26c)$$

$$|\phi_n| = 1, \forall n \in \mathcal{N}, \quad (26d)$$

where  $0 \leq \beta \leq 1$  is the outage probability when the IT imposed on the PR exceeds the threshold. The real challenge to solve this optimization problem is the statistical CSI error model of (26c) since the SINR constraint (26b) and the unit-modulus constraint (26d) of the phase shifts at the IRS can be handled in the same way as in problem  $\mathcal{P}1$ . Hence, in this section, we will first reformulate the probability constraint by using Chi-square distribution derivation method. The related result is shown in Proposition 1.



Firstly, we reformulate the IT inequality constraint. According to the triangle inequality  $\|(\tilde{\phi}^H \hat{\mathbf{G}} + \tilde{\phi}^H \Delta \mathbf{G}) \mathbf{W}\|_2^2 \leq \|\tilde{\phi}^H \hat{\mathbf{G}} \mathbf{W}\|_2^2 + \|\tilde{\phi}^H \Delta \mathbf{G} \mathbf{W}\|_2^2$ , the following probabilistic relationship holds

$$\Pr\{\|(\tilde{\phi}^H \hat{\mathbf{G}} + \tilde{\phi}^H \Delta \mathbf{G}) \mathbf{W}\|_2^2 \leq \Gamma\} \geq \Pr\{\|\tilde{\phi}^H \hat{\mathbf{G}} \mathbf{W}\|_2^2 + \|\tilde{\phi}^H \Delta \mathbf{G} \mathbf{W}\|_2^2 \leq \Gamma\}. \quad (27)$$

Thus, the IT inequality constraint can be approximated by

$$\Pr\{\|\tilde{\phi}^H \Delta \mathbf{G} \mathbf{W}\|_2^2 \leq \Gamma - \|\tilde{\phi}^H \hat{\mathbf{G}} \mathbf{W}\|_2^2\} \geq 1 - \beta. \quad (28)$$

Based on the above analysis, the following proposition can be obtained.

*Proposition 1:* Assume that  $\text{vec}(\Delta \mathbf{G}) \in \mathbb{C}^{(N+1)M_t \times 1}$  is a complex Gaussian vector satisfying  $\text{vec}(\Delta \mathbf{G}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$  as in (7). The sufficient condition for the probabilistic constraint (28) to hold is that

$$\vartheta F_{2(N+1)M_t}^{-1}(1 - \beta) \|\mathbf{W}\|_{\text{F}}^2 + \|\tilde{\phi}^H \hat{\mathbf{G}} \mathbf{W}\|_2^2 \leq \Gamma, \quad (29)$$

where  $\vartheta = (N + 1)\lambda_{\max}(\mathbf{\Sigma})$ ,  $\lambda_{\max}(\cdot)$  is the maximum eigenvalue and  $F_n^{-1}(\cdot)$  is the inverse Chi-square cumulative distribution function with  $n$  degrees of freedom.

*Proof:* Please see Appendix A. ■

Note that when  $\beta = 0$ , the constraint in (26c) means the worst-case for the statistical error model. If  $\beta = 1$ , the constraint in (26c) can be removed because the probability constraint is always satisfied. In this case, the PR cannot be protected.

By replacing the IT constraint (26c) with (29), problem  $\mathcal{P}2$  is reformulated as

$$\mathcal{P}2' : \min_{\mathbf{W}, \Phi} \|\mathbf{W}\|_{\text{F}}^2 \quad \text{s.t.} \quad (26b), (29), (26d). \quad (30)$$

This problem is also intractable to solve because  $\mathbf{W}$  and  $\Phi$  are coupled in (26b) and (29). The BCD algorithm is adopted to optimize these two variables alternately.

### B. Optimizing $\mathbf{W}$ with Fixed $\Phi$ for STA Scheme

When  $\Phi$  is fixed, the IT constraint (29) can be rewritten as

$$\sum_{k=1}^K \mathbf{w}_k^H (\vartheta F_{2(N+1)M_t}^{-1}(1 - \beta) \mathbf{I} + \tilde{\mathbf{G}}_{\tilde{\phi}}) \mathbf{w}_k \leq \Gamma. \quad (31)$$

The inequality constraint (17) in SCD scheme can be directly used to replace the SINR constraint (26b). Then the problem for optimizing  $\mathbf{W}$  is reformulated as

$$\mathcal{P}2.1 : \min_{\{\mathbf{w}_k\}} \sum_{k=1}^K \mathbf{w}_k^H \mathbf{w}_k \quad \text{s.t.} \quad (17), (31). \quad (32)$$

The SINR constraint (17) is non-convex. We can use SDR approach to solve this problem. Specifically, define  $\mathbf{S}_k = \mathbf{w}_k \mathbf{w}_k^H, \forall k \in \mathcal{K}$  with the constraints that  $\text{rank}(\mathbf{S}_k) = 1, \forall k \in \mathcal{K}$ . Then, problem  $\mathcal{P}2.1$  can be reformulated as

$$\mathcal{P}2.1\_1 : \quad \min_{\{\mathbf{S}_k \succeq \mathbf{0}\}_{k \in \mathcal{K}}} \sum_{k=1}^K \text{tr}(\mathbf{S}_k) \quad (33a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{S}_k \widehat{\mathbf{H}}_{\Phi,k}) \geq \gamma_k \left( \sum_{j \neq k}^K \text{tr}(\mathbf{S}_j \widehat{\mathbf{H}}_{\Phi,k}) + \sigma_{s_k}^2 \right), \forall k \in \mathcal{K}, \quad (33b)$$

$$\sum_{k=1}^K \text{tr} \left( \mathbf{S}_k \left( \vartheta F_{2(N+1)M_t}^{-1} (1 - \beta) \mathbf{I} + \tilde{\mathbf{G}}_{\tilde{\phi}} \right) \right) \leq \Gamma. \quad (33c)$$

$$\text{rank}(\mathbf{S}_k) = 1, \forall k \in \mathcal{K}. \quad (33d)$$

However, problem  $\mathcal{P}2.1\_1$  is still non-convex due to the rank-one constraint. We further relax this rank-one constraint and obtain the following optimization problem:

$$\mathcal{P}2.1\_2 : \quad \min_{\{\mathbf{S}_k \succeq \mathbf{0}\}_{k \in \mathcal{K}}} \sum_{k=1}^K \text{tr}(\mathbf{S}_k) \quad \text{s.t.} \quad (33b), (33c) \quad (34)$$

Obviously, problem  $\mathcal{P}2.1\_2$  is an SDP problem, which is convex and can be effectively solved by using the standard tools such as CVX.

In general, the optimal solution obtained from SDP problem may not satisfy the rank-one constraints. Fortunately, for problem  $\mathcal{P}2.1\_2$ , we can prove that the relaxation of the non-convex constraints is tight. Denote the optimal solution of  $\mathcal{P}2.1\_2$  as  $\mathbf{S}_k^*, \forall k \in \mathcal{K}$ , we have the following theorem.

*Theorem 1:* The optimal solution obtained from the SDP problem  $\mathcal{P}2.1\_2$  is guaranteed to satisfy the rank-one constraints, i.e.,  $\text{rank}(\mathbf{S}_k^*) = 1, \forall k \in \mathcal{K}$ .

*Proof:* Please see Appendix B. ■

Since the rank of  $\mathbf{S}_k, \forall k \in \mathcal{K}$  are equal to one, we can use the simple singular value decomposition operation to obtain the optimal beam-vector of problem  $\mathcal{P}2.1$ .

### C. Optimizing $\Phi$ with Fixed $\mathbf{W}$ for STA SCheme

When  $\mathbf{W}$  is fixed, this subproblem degenerates into a feasibility-check problem. The SINR constraint (21) can be directly used to replace (26b). The IT constraint (29) can be rewritten as

$$\sum_{k=1}^K \tilde{\phi}^H \widehat{\mathbf{G}} \mathbf{w}_k \mathbf{w}_k^H \widehat{\mathbf{G}}^H \tilde{\phi} \leq \Gamma - \vartheta F_{2(N+1)M_t}^{-1} (1 - \beta) \|\mathbf{W}\|_F^2. \quad (35)$$

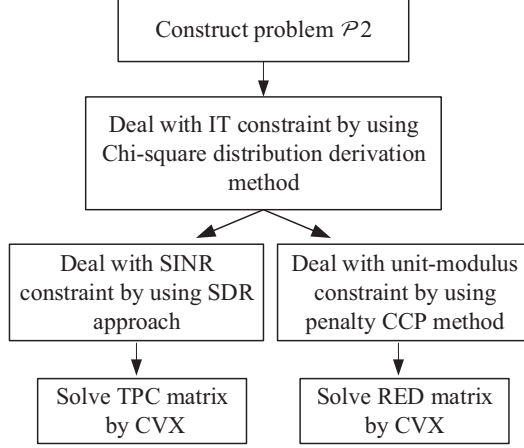


Fig. 3. The key steps and methodologies involved in STA scheme.

Let  $\mathbf{X} \triangleq \sum_{k=1}^K \widehat{\mathbf{G}} \mathbf{w}_k \mathbf{w}_k^H \widehat{\mathbf{G}}^H$  which is a Hermitian matrix. Extract the first  $N$  rows and  $N$  columns elements of  $\mathbf{X}$  as a sub-matrix  $\mathbf{B}$ . Denote the vector consisting of elements in the  $(N+1)$ th column from the first row to the  $N$ th row of  $\mathbf{X}$  by  $\mathbf{b}$ . Denote the vector consisting of elements in the  $(N+1)$ th row from the first column to the  $N$ th column of  $\mathbf{X}$  by  $\mathbf{c}$  and denote the  $(N+1)$ th row and  $(N+1)$ th column element by  $b_{N+1}$ . Thus, we have

$$\sum_{k=1}^K \tilde{\phi}^H \widehat{\mathbf{G}} \mathbf{w}_k \mathbf{w}_k^H \widehat{\mathbf{G}}^H \tilde{\phi} = \phi^H \mathbf{B} \phi + 2\text{Re}\{\mathbf{b}^H \phi\} + b_{N+1}. \quad (36)$$

Finally, the IT constraint (35) is reformulated as

$$\phi^H \mathbf{B} \phi + 2\text{Re}\{\mathbf{b}^H \phi\} \leq \Gamma - \vartheta F_{2(N+1)M_t}^{-1}(1-\beta) \|\mathbf{W}\|_{\text{F}}^2 - b_{N+1}. \quad (37)$$

By introducing the slack variables to the SINR constraint like the method used in problem  $\mathcal{P}1.2$ , the subproblem is

$$\mathcal{P}2.2 : \max_{\phi, \varphi} \sum_{k=1}^K \varphi_k \quad \text{s.t.} \quad (22), (37), (26d), \varphi \geq 0. \quad (38)$$

Problem  $\mathcal{P}2.2$  can be solved by using the same method with  $\mathcal{P}1.2$ , here we omit it for simplicity.

For the convenience of understanding, a flow chart is provided in Fig. 3 to summarize the key steps of the whole optimization processing for STA scheme.

#### D. Overall Algorithm for STA Scheme

The overall algorithm for solving problem  $\mathcal{P}2$  is given in Algorithm 2. The initialization value of  $\mathbf{w}_k^{(0)}$  and  $\phi^{(0)}$  can be obtained from the optimal solution of feasibility checking problem which

---

**Algorithm 2** Optimizing  $\mathbf{W}$  and  $\Phi$  for STA
 

---

- 1: Initialize  $\mathbf{w}_k^{(0)}$  and  $\phi^{(0)}$  in feasible region for  $\forall k \in \mathcal{K}$ , set maximum iteration number  $t_{max}$ , target convergence accuracy  $\zeta$  and  $t = 0$ .
  - 2: **Repeat**
  - 3: Calculate  $\mathbf{w}_k^{(t+1)}$  by solving problem  $\mathcal{P}2.1$ ;
  - 4: Calculate  $\phi^{(t+1)}$  by solving problem  $\mathcal{P}2.2$ ;
  - 5: Calculate  $\|\mathbf{W}^{(t+1)}\|_{\mathbb{F}}^2$  from (26a) in problem  $\mathcal{P}2$ ;
  - 6: **Until**  $t > t_{max}$  or  $\frac{\|\mathbf{W}^{(t+1)}\|_{\mathbb{F}}^2 - \|\mathbf{W}^{(t)}\|_{\mathbb{F}}^2}{\|\mathbf{W}^{(t+1)}\|_{\mathbb{F}}^2} < \zeta$ .
- 

will be discussed in Section V. Since the proposed Algorithm 2 involving SOC, LMI and linear constraints can be solved by a standard interior point method, the general expression of the computational complexity of which can be given as the equation (25). Based on that, the complexity of problem  $\mathcal{P}2.1$  is  $o_{\mathbf{w}}^{sta} = \mathcal{O}((K+1)^{0.5}(K(K+1)^2M_t^{4.5} + K^3M_t^{3.5})) + \mathcal{O}(M_t^3)$  where the second term  $\mathcal{O}(M_t^3)$  is produced by singular value decomposition operation of matrix solution. The complexity of problem  $\mathcal{P}2.2$  is  $o_{\Phi}^{sta} = \mathcal{O}(\log(1/\min\{l_1, l_2\})\sqrt{2}(K+N+1)^{0.5}[(K+1)N^4 + 2N^3])$ . Then the overall complexity of problem  $\mathcal{P}2$  is  $o_{\mathbf{w}}^{sta} + o_{\Phi}^{sta}$ . It can be also easily verified that the sequences of the objective function value produced by Algorithm 2 are guaranteed to converge.

## V. FEASIBILITY CHECKING FOR PROBLEMS $\mathcal{P}_1$ AND $\mathcal{P}_2$

Due to the conflicting SINR and IT constraints, problems  $\mathcal{P}1$  and  $\mathcal{P}2$  may be infeasible for SCD scheme and STA scheme, respectively. Hence, we have to first check whether problems  $\mathcal{P}1$  and  $\mathcal{P}2$  are feasible or not. To this end, we construct the following two feasibility checking problems.

### A. Feasibility Checking for Problem $\mathcal{P}1$

By minimizing the interference imposed on the PU, we can construct the feasibility checking problem for SCD scheme as

$$\overline{\mathcal{P}1} : \min_{\mathbf{W}, \Phi, \alpha} \alpha \quad (39a)$$

$$\text{s.t. } \|(\mathbf{g}_d^H + \phi^H \mathbf{G}_r) \mathbf{W}\|_2^2 \leq \alpha \Gamma, \|\Delta \mathbf{g}_d\|_2 \leq \epsilon_d, \|\Delta \mathbf{G}_r\|_{\mathbb{F}} \leq \epsilon_r, \quad (39b)$$

$$(8b), (8d) \quad (39c)$$

where  $\alpha$  is an additional variable introduced to act as the indicators of the feasibility of problem  $\mathcal{P}1$ . If the optimal solution of  $\alpha$  is smaller or equal to one, we claim the main problem is feasible. Otherwise, the main problem  $\mathcal{P}1$  is declared to be infeasible. By applying Schur's complement and Lemma 1, the IT constraint (39b) of problem  $\overline{\mathcal{P}1}$  can be transformed into the following LMI

$$\begin{bmatrix} \alpha\Gamma - \rho_1 N - \rho_2 & \widehat{\mathbf{b}}^H & \mathbf{0}_{1 \times M_t} & \mathbf{0}_{1 \times M_t} \\ \widehat{\mathbf{b}} & \mathbf{I}_K & \epsilon_r \mathbf{W}^H & \epsilon_d \mathbf{W}^H \\ \mathbf{0}_{M_t \times 1} & \epsilon_r \mathbf{W} & \rho_1 \mathbf{I}_{M_t} & \mathbf{0} \\ \mathbf{0}_{M_t \times 1} & \epsilon_d \mathbf{W} & \mathbf{0} & \rho_2 \mathbf{I}_{M_t} \end{bmatrix} \succeq \mathbf{0}. \quad (40)$$

Problem  $\overline{\mathcal{P}1}$  can be reformulated as

$$\overline{\mathcal{P}1'}: \min_{\mathbf{W}, \Phi, \rho, \alpha} \alpha \quad \text{s.t.} \quad (8b), (40), (8d), \rho \geq 0. \quad (41)$$

Since  $\mathbf{W}$  and  $\Phi$  are coupled, the BCD approach is used again to alternately solve this problem. Based on this idea, we present an iterative algorithm with two main steps.

In the first step, the TPC matrix  $\mathbf{W}$  is optimized with fixed RED matrix  $\Phi$ . Then the alternative problem becomes

$$\overline{\mathcal{P}1.1}: \min_{\mathbf{W}, \rho, \alpha} \alpha \quad \text{s.t.} \quad (18), (40), \rho \geq 0. \quad (42)$$

Problem  $\overline{\mathcal{P}1.1}$  is an SOCP problem which is convex and can be effectively solved by using CVX. The optimal solution  $\alpha$  obtained from this problem can be used to check the feasibility.

In the second step, the RED matrix is optimized when the TPC matrix,  $\rho$  and  $\alpha$  are given. Then the alternative problem becomes

$$\overline{\mathcal{P}1.2}: \max_{\Phi, \varphi} \sum_{k=1}^K \varphi_k \quad \text{s.t.} \quad (22), (40), (8d), \varphi \geq 0. \quad (43)$$

where  $\varphi = [\varphi_1, \varphi_2, \dots, \varphi_K]$  are slack variables introduced into the SINR constraint. This problem is non-convex due to the unit-modulus constraint (8d). Then we adopt the penalty CCP method to deal with this non-convex constraint. Problem  $\overline{\mathcal{P}1.2}$  is transformed into an SOCP problem. The problem transformation process is similar to problem  $\mathcal{P}1.2$ . Here we omit it for simplicity.

The feasibility checking algorithm for the main problem  $\mathcal{P}1$  is formally described in Algorithm 3.

---

**Algorithm 3** Feasibility Checking Algorithm for Problem  $\mathcal{P}1$ 


---

- 1: Initialize  $\mathbf{W}^{(0)}$ ,  $\Phi^{(0)}$ ,  $t_{max}$  and set  $t = 1$ .
  - 2: Given RED matrix  $\Phi^{(t-1)}$ , update TPC matrix by solving problem  $\overline{\mathcal{P}1.1}$ . Denote the solution by  $\alpha^*$ ,  $\mathbf{W}^*$ . Update  $\mathbf{W}^{(t)} = \mathbf{W}^*$ ,  $\alpha^{(t)} = \alpha^*$ .
  - 3: If  $\alpha^{(t)} \leq 1$ , declare the main problem  $\mathcal{P}1$  is feasible and output these feasible  $\mathbf{W}^{(t)}$  for the initialization of the main problem  $\mathcal{P}1$ , and terminate. If  $\alpha^{(t)} \geq 1$  and  $t \geq t_{max}$ , declare the main problem  $\mathcal{P}1$  is infeasible and terminate. Otherwise, continue.
  - 4: Given TPC matrix  $\mathbf{W}^{(t)}$ , update RED matrix by solving problem  $\overline{\mathcal{P}1.2}$ . Denote the solution by  $\Phi^*$ . Update  $\Phi^{(t)} = \Phi^*$ .
  - 5: Let  $t \leftarrow t + 1$  and go to step 2.
- 

### B. Feasibility Checking for Problem $\mathcal{P}2$

If we set an outage probability  $\beta$ , the objective function can be regarded as the interference imposed on the PU which causes communication interruption of PU with probability  $\beta$  where the uncertain CSI  $\Delta\mathbf{G}$  satisfies the distribution of  $\text{vec}(\Delta\mathbf{G}) \sim \mathcal{CN}(\mathbf{0}, \Sigma)$ . According to Proposition 1, the objective function can be given by  $\vartheta F_{2(N+1)M_t}^{-1}(1 - \beta) \|\mathbf{W}\|_F^2 + \|\tilde{\phi}^H \hat{\mathbf{G}} \mathbf{W}\|_2^2$ . Hence, the feasibility checking problem for the main problem  $\mathcal{P}2$  is

$$\overline{\mathcal{P}2} : \min_{\mathbf{W}, \Phi} \vartheta F_{2(N+1)M_t}^{-1}(1 - \beta) \|\mathbf{W}\|_F^2 + \|\tilde{\phi}^H \hat{\mathbf{G}} \mathbf{W}\|_2^2 \quad (44a)$$

$$\text{s.t.} \quad (26b), (26d). \quad (44b)$$

where (26b) is SINR constraint and (26d) is unit-modulus constraint of phase shift at the IRS. The feasibility can be checked by judging the optimal objective value. If the optimal objective value is lower than IT threshold  $\Gamma$ , problem  $\overline{\mathcal{P}2}$  is feasible. Otherwise, it is infeasible. Since  $\mathbf{W}$  and  $\Phi$  are coupled, we have to alternately solve this problem. We present the iterative algorithm with two main steps.

In the first step, TPC matrix is optimized with fixed RED matrix. When  $\Phi$  is fixed, the objective function (44a) can be rewritten as  $\sum_{k=1}^K \mathbf{w}_k^H (\vartheta F_{2(N+1)M_t}^{-1}(1 - \beta) \mathbf{I} + \tilde{\mathbf{G}}_{\tilde{\phi}}) \mathbf{w}_k$ . The SINR constraints can be replaced by (17). Then we can give the TPC matrix optimization problem by

$$\overline{\mathcal{P}2.1} : \min_{\mathbf{W}} \sum_{k=1}^K \mathbf{w}_k^H (\vartheta F_{2(N+1)M_t}^{-1}(1 - \beta) \mathbf{I} + \tilde{\mathbf{G}}_{\tilde{\phi}}) \mathbf{w}_k \quad \text{s.t.} \quad (17). \quad (45a)$$

This problem can be solved by using the same SDR approach with problem  $\mathcal{P}2.1$ . After defining  $\mathbf{S}_k = \mathbf{w}_k \mathbf{w}_k^H, \forall k \in \mathcal{K}$  with the constraints that  $\text{rank}(\mathbf{S}_k) = 1, \forall k \in \mathcal{K}$ , problem  $\overline{\mathcal{P}2.1}$  can be transformed into

$$\overline{\mathcal{P}2.1\_1} : \quad \min_{\{\mathbf{S}_k \succeq \mathbf{0}\}_{k \in \mathcal{K}}} \sum_{k=1}^K \text{tr} \left( \mathbf{S}_k \left( \vartheta F_{2(N+1)M_t}^{-1} (1 - \beta) \mathbf{I} + \tilde{\mathbf{G}}_{\tilde{\phi}} \right) \right) \quad (46a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{S}_k \hat{\mathbf{H}}_{\Phi, k}) \geq \gamma_k \left( \sum_{j \neq k}^K \text{tr}(\mathbf{S}_j \hat{\mathbf{H}}_{\Phi, k}) + \sigma_{s_k}^2 \right), \forall k \in \mathcal{K}, \quad (46b)$$

$$\text{rank}(\mathbf{S}_k) = 1, \forall k \in \mathcal{K}. \quad (46c)$$

We further relax this rank-one constraint and obtain an SDP optimization problem, which can be effectively solved by CVX. The solution of the SDP problem can be guaranteed to satisfy the rank-one constraints.

In the second step, RED matrix is optimized with fixed TPC matrix. When  $\mathbf{W}$  is given, the RED matrix optimization problem is given by

$$\overline{\mathcal{P}2.2} : \quad \min_{\phi} \quad \phi^H \mathbf{B} \phi + 2\text{Re}\{\mathbf{b}^H \phi\} + \vartheta F_{2(N+1)M_t}^{-1} (1 - \beta) \|\mathbf{W}\|_{\text{F}}^2 + b_{N+1} \quad (47a)$$

$$\text{s.t.} \quad \gamma_k \phi^H \boldsymbol{\Omega}_{-k} \phi - 2\text{Re}\{(\boldsymbol{\omega}_k^H + \phi^{(t)H} \boldsymbol{\Omega}_k) \phi\} \leq \bar{\gamma}_k - \varphi_k, \forall k \in \mathcal{K}, \quad (47b)$$

$$|\phi_n| = 1, \forall n \in \mathcal{N}, \quad (47c)$$

where  $\mathbf{B}$ ,  $\phi$  and  $b_{N+1}$  in objective function is obtained from (37). This optimization problem can be solved by using the same method with problem  $\overline{\mathcal{P}1.2}$  where the penalty CCP method is used to deal with unit-modulus constraint. The problem can be transformed into an SOCP problem solved by CVX.

The feasibility checking algorithm for the main problem  $\mathcal{P}2$  is formally described in Algorithm 4.

## VI. EXTENSION TO MULTIPLE PRS SCENARIO

In this section, we extend one PR scenario to multiple PRs scenario. Assume that  $L$  PRs randomly locate in a cell region in the IRS-aided CR system. The other assumptions remain the same as in Fig. 1. The system parameters correspond to the  $l$ th PR ( $l \in \mathcal{L} = \{1, 2, \dots, L\}$ ) are direct channel  $\mathbf{g}_{d,l}$ , reflect channel  $\mathbf{g}_{r,l}$  and IT threshold requirement  $\Gamma_l$ . By using  $\mathbf{g}_{d,l}$ ,  $\mathbf{g}_{r,l}$  and  $\Gamma_l$  in IT expressions and channel uncertainty model, we can rewrite the optimization problem for the multiple PRs scenario.

---

**Algorithm 4** Feasibility Checking Algorithm for Problem  $\mathcal{P}2$ 


---

- 1: Initialize  $\mathbf{W}^{(0)}$ ,  $\Phi^{(0)}$ ,  $t_{max}$  and set  $t = 1$ .
  - 2: Given RED matrix  $\Phi^{(t-1)}$ , update TPC matrix by solving problem  $\overline{\mathcal{P}2.1}$ . Denote the solution by  $\mathbf{W}^*$ . Update  $\mathbf{W}^{(t)} = \mathbf{W}^*$ .
  - 3: Given TPC matrix  $\mathbf{W}^{(t)}$ , update RED matrix by solving problem  $\overline{\mathcal{P}2.2}$ . Denote the solution by  $\Phi^*$ . Update  $\Phi^{(t)} = \Phi^*$ . Calculate the optimal value of objective function for problem  $\overline{\mathcal{P}2}$  and denote it as  $f^*$ .
  - 4: If  $f^* \leq \Gamma$ , declare the main problem  $\mathcal{P}2$  is feasible and output these feasible  $\mathbf{W}^{(t)}$  and  $\Phi^{(t)}$  for the initialization of the main problem  $\mathcal{P}2$ , and terminate. If  $f^* \geq \Gamma$  and  $t \geq t_{max}$ , declare the main problem  $\mathcal{P}2$  is infeasible and terminate. Otherwise, continue.
  - 5: Let  $t \leftarrow t + 1$  and go to step 2.
- 

Firstly, the IT imposed on PR  $l$  from the ST is given by

$$IT_l = \|\tilde{\phi}^H \mathbf{G}_l \mathbf{W}\|_2^2, \quad (48)$$

where  $\mathbf{G}_l = [\mathbf{G}_{r,l}^H \ \mathbf{g}_{d,l}]^H \in \mathbb{C}^{(N+1) \times M_t}$ , for  $\forall l \in \mathcal{L}$ , is an equivalent combined channel.  $\mathbf{G}_{r,l} = \text{diag}(\mathbf{g}_{r,l}^H) \mathbf{F} \in \mathbb{C}^{N \times M_t}$  is regarded as the cascaded ST-IRS-PR  $l$  channel.

Secondly, for the bounded CSI error model, the radii of the bounded regions of CSI errors for direct channel  $\mathbf{g}_{d,l}$  and cascaded channel  $\mathbf{G}_{r,l}$  can be denoted by  $\epsilon_{d,l}$  and  $\epsilon_{r,l}$ , respectively, for  $\forall l$ . Then, the radius of the bounded region of CSI error for the combined channel  $\mathbf{G}_l$  is denoted by  $\epsilon_l$ . For the statistical CSI error model, the distribution assumption in Eq. (6) is rewritten as

$$\Delta \mathbf{g}_{d,l} \sim \mathcal{CN}(\mathbf{0}, \Sigma_{g_{d,l}}) \quad \text{vec}(\Delta \mathbf{G}_{r,l}) \sim \mathcal{CN}(\mathbf{0}, \Sigma_{g_{r,l}}), \quad (49)$$

where  $\Sigma_{g_{d,l}}$  and  $\Sigma_{g_{r,l}}$  are covariance matrices for  $\forall l \in \mathcal{L}$ . Thus, the CSCG distribution for  $\Delta \mathbf{G}_l$  is given by

$$\text{vec}(\Delta \mathbf{G}_l) \sim \mathcal{CN}(\mathbf{0}, \Sigma_l), \forall l \in \mathcal{L}, \quad (50)$$

where  $\Sigma_l = \begin{bmatrix} \Sigma_{g_{r,l}} & \mathbf{0} \\ \mathbf{0} & \Sigma_{g_{d,l}} \end{bmatrix}$ ,  $\forall l \in \mathcal{L}$ .

Based on the above assumptions, for the optimization problems of SCD and STA schemes, only the IT constraints need to be modified into multi-PRs form. Hence, constraint (8c) in SCD optimization problem is rewritten as

$$\|(\mathbf{g}_{d,l}^H + \phi^H \mathbf{G}_{r,l}) \mathbf{W}\|_2^2 \leq \Gamma_l, \|\Delta \mathbf{g}_{d,l}\|_2 \leq \epsilon_{d,l}, \|\Delta \mathbf{G}_{r,l}\|_F \leq \epsilon_{r,l}. \quad (51)$$



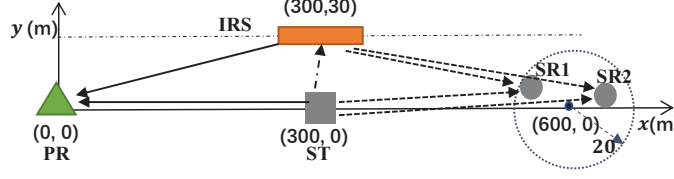


Fig. 4. The simulated scenario of IRS-aided CR.

Constraint (26c) in STA optimization problem is rewritten as

$$\Pr \left\{ \left\| (\tilde{\phi}^H \hat{\mathbf{G}}_l + \tilde{\phi}^H \Delta \mathbf{G}_l) \mathbf{W} \right\|_2^2 \leq \Gamma_l \right\} \geq 1 - \beta_l, \text{vec}(\Delta \mathbf{G}_l) \sim \mathcal{CN}(\mathbf{0}, \Sigma_l), \forall l \in \mathcal{L}. \quad (52)$$

where  $\beta_l$  is outage probability due to secondary interference for PR  $l$ . The methods and mathematical derivations in single-PR optimization problems are applicable in multi-PRs optimization problems. Here, we omit it for simplicity. In simulation results, we will provide the impacts of number of PRs on system performance.

## VII. SIMULATION RESULTS

In this section, we firstly provide simulation results of one PR scenario to study the impacts of parameters on the robust beamforming design in the IRS-aided MISO CR system. Then we expand the simulation to multiple PRs scenario. The simulation scenario is shown in Fig. 4 where the PR, ST and IRS are located at the same horizontal line with locations given by (0 m, 0 m), (300 m, 0) and (300 m, 30 m), respectively. There are  $K = 2$  SRs randomly distributed in the region of a cell. The channel models are assumed to include large-scale fading and small-scale fading. Denote the large-scale path loss in dB by  $P_L = P_{L_0} - 10\alpha \log_{10}(\frac{d}{d_0})$ , where  $\alpha$  is the path loss exponent and  $d$  is the transmission distance. The small-scale fading follows Rayleigh distribution. The minimum data rate requirement for SU is assumed to be the same at  $r_k = r, \forall k \in \mathcal{K}$ . The main parameters are collected in Table II.

The CSI error vectors  $\Delta \mathbf{g}_d$  and  $\text{vec}(\Delta \mathbf{G}_r)$  are assumed to be correlated for the STA scheme. The covariance matrices of  $\Sigma_{g_r}$  and  $\Sigma_{g_d}$  are defined as  $\Sigma_{g_r}^{1/2} = \sigma_{g_r} \mathbf{C}_{g_r}$  and  $\Sigma_{g_d}^{1/2} = \sigma_{g_d} \mathbf{C}_{g_d}$ .  $\mathbf{C}_{g_r} \in \mathbb{C}^{NM_t \times NM_t}$  and  $\mathbf{C}_{g_d} \in \mathbb{C}^{M_t \times M_t}$  are the correlation matrices, the elements of which are given by  $[\mathbf{C}_{g_r}]_{m,n}, [\mathbf{C}_{g_d}]_{m,n} = c_\eta^{|m-n|}$  and  $c_\eta$  is set as 0.9. To make the relative amount of correlated CSI errors comparable with the case of uncorrelated CSI errors, we assume that  $\sigma_{g_r}^2 = \delta_{g_r}^2 \left\| \text{vec}(\hat{\mathbf{G}}_r) \right\|_2^2 \left[ \frac{\|\mathbf{I}_{NM_t}\|_F}{\|\mathbf{C}_{g_r}^2\|_F} \right]$  and  $\sigma_{g_d}^2 = \delta_{g_d}^2 \left\| (\hat{\mathbf{g}}_d) \right\|_2^2 \left[ \frac{\|\mathbf{I}_{M_t}\|_F}{\|\mathbf{C}_{g_d}^2\|_F} \right]$ .  $\delta_{g_d} \in [0, 1)$  and  $\delta_{g_r} \in [0, 1)$  represent the channel uncertainty levels which measure the relative amount of CSI uncertainties, here we set  $\delta_{g_d} = \delta_{g_r} = \delta_g$ . The radii of the uncertainty regions are set as  $\epsilon_d = \sqrt{\frac{\sigma_{g_d}^2}{2} F_{2M_t}^{-1}(1 - \beta)}$

TABLE II  
SIMULATION PARAMETERS.

Parameter	Value
PR location	(0 m, 0 m)
IRS location	(300 m, 30 m)
ST location	(300 m, 0 m)
SR's cell location	(600 m, 0 m)
Cell radius	20 m
Pass loss	$P_{L_0} = 30$ dB
Reference distance	$d_0 = 1$ m
Path loss exponents of the IRS-related links	$\alpha_{\text{IRS}} = 2.2$
Path loss exponent of the ST-PR link or the ST-SR link	$\alpha_{\text{STU}} = 3.75$ [23], [31], [32]
Outage probability	$\beta = 0.05$
Noise power density	$-174$ dBm/Hz

and  $\epsilon_r = \sqrt{\frac{\sigma_{gr}^2}{2} F_{2NM_t}^{-1}(1 - \beta)}$ , respectively. According to [37], this bounded CSI error model provides a fair comparison between the performance of the worst-case robust design and the outage constrained robust design.

### A. The Convergence and Complexity of Algorithms

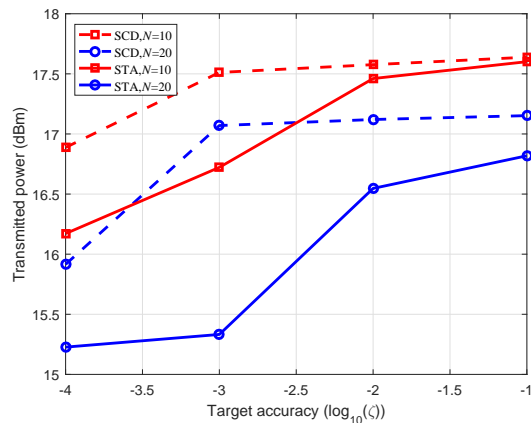
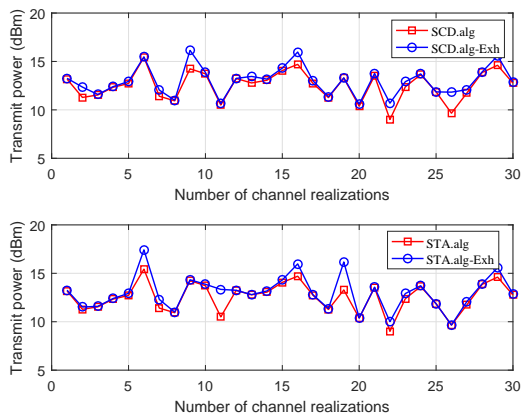


Fig. 5. The performance comparison of different initializations, with  $r = 0.5$ ,  $M_t = 4$ ,  $\Gamma = -60$  dBm,  $\beta = 0.05$  and  $M_t = 4$ ,  $\Gamma = -90$  dBm,  $\beta = 0.05$  and  $\delta_g = 0.05$ .

$\delta_g = 0.05$ .

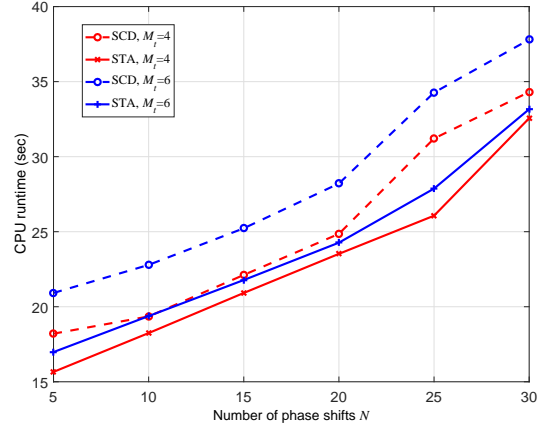
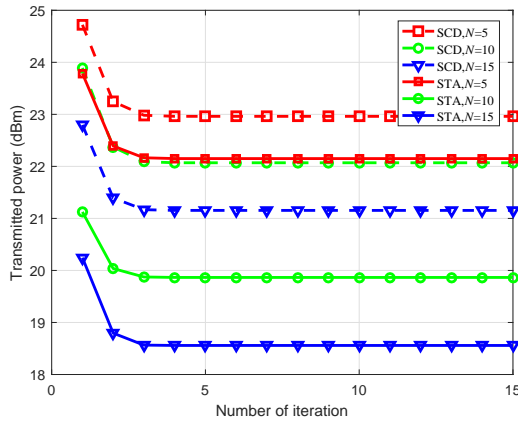


Fig. 7. Transmit power versus the number of iterations, with Fig. 8. CPU runtime versus  $N$ , with  $r = 1$ ,  $\Gamma = -70$  dBm,  $r = 2$ ,  $M_t = 4$ ,  $\Gamma = -90$  dBm,  $\beta = 0.05$  and  $\delta_g = 0.05$ .  $\beta = 0.05$  and  $\delta_g = 0.05$ .

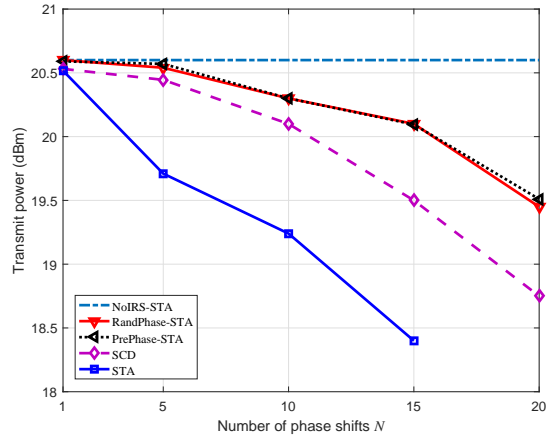
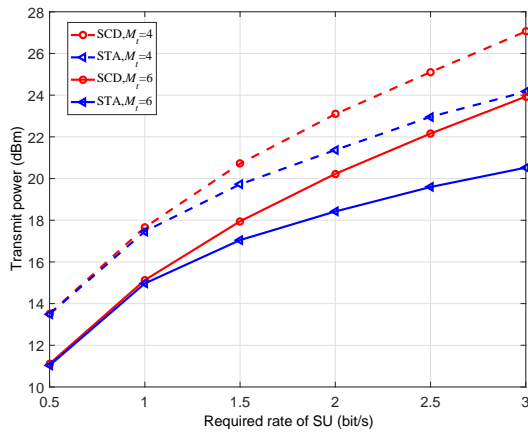


Fig. 9. Transmit power versus minimum required bit rate of Fig. 10. Comparison with benchmarks with  $r = 2$ ,  $M_t = 6$ , SU, with  $N = 6$ ,  $\Gamma = -80$  dBm,  $\beta = 0.05$  and  $\delta_g = 0.2$ .  $\Gamma = -80$  dBm,  $\beta = 0.05$  and  $\delta_g = 0.1$ .

Consider the fact of the nonconvexity of problems  $\mathcal{P}1$  and  $\mathcal{P}2$ , different initial points may result in different locally optimal solutions obtained by our proposed algorithms. By testing 30 randomly channel realizations, Fig. 5 illustrates the impact of the initializations on the performance of the proposed algorithms. The initializations of  $\mathbf{W}$  and  $\Phi$  for SCD.alg and STA.alg are determined by our proposed feasibility checking problems. SCD.alg-Exh and STA.alg-Exh refers to the best initial point of 1000 random initial points for each channel realization. It can be seen that the minimum transmit power obtained by SCD.alg (STA.alg) is almost the same as that

of SCD.alg-Exh (STA.alg-Exh), which implies that the initial point obtained from the feasibility checking problem is a good option for the initialization. According to the target convergence criteria  $\frac{\|\mathbf{W}^{(t+1)}\|_{\mathbb{F}}^2 - \|\mathbf{W}^{(t)}\|_{\mathbb{F}}^2}{\|\mathbf{W}^{(t+1)}\|_{\mathbb{F}}^2} < \zeta$  which is set in Algorithm 1 and Algorithm 2, Fig. 6 shows that the higher the requirement of convergence accuracy, the smaller the optimal target value. The x-axis values from  $-4$  to  $-1$  represent the accuracy from  $10^{-4}$  to  $10^{-1}$  which means that the accuracy requirement decreases. Given the convergence accuracy  $\zeta = 10^{-4}$ , Fig. 7 shows that both the proposed algorithms can converge within six iterations. Compared with STA algorithm, the SCD algorithm always converges at a higher transmit power when the number of phase shifts ranges from 5 to 15. Moreover, from Fig. 9, we can find that the transmit power increases with SUs' data rate requirements and the SCD algorithm yields a higher transmit power than STA algorithm when each SU's data rate is in a high value range for different numbers of the transmit antennas. This is due to the fact that the SCD scheme is the worst-case optimization which is more conservative than the statistical optimization and requires more transmit power to ensure that each SU's target rate requirement for the worst-case CSI error realization can be satisfied by the achievable rate.

For the computational complexity, Fig. 8 shows the average CPU running time versus the number of phase shifts  $N$  for both proposed algorithms when  $r = 1$ ,  $\Gamma = -70$  dBm,  $\beta = 0.05$  and  $\delta_g = 0.05$ . The results are obtained by using a computer with a 1.99 GHz i7-8550U CPU and 8 GB RAM. The STA algorithm requires much less CPU running time than that required by the SCD algorithm. This is due to fact that there are some large-dimensional LMIs that increase the computational complexity of the SCD algorithm. With the increase of the number of  $N$ , all the algorithms need more running time to obtain the optimal solutions. Fig. 7, Fig. 8 and Fig. 9 indicate that the STA algorithm outperforms the SCD algorithm both in terms of optimal value and computational complexity.

### B. Compared with Other Benchmarks

Fig. 10 compares the SCD and STA schemes with three benchmark schemes, named as NoIRS-STA, RandPhase-STA and Prephase-STA. For the NoIRS-STA scheme, there is no IRS deployed in the system that can be regarded as the traditional CR system. For the RandPhase-STA scheme, the phase of each reflection element is randomly and uniformly generated between 0 and  $2\pi$  for each subproblem of optimizing RED matrix. For the Prephase-STA scheme, all modulus values of the phase shifts are fixed to be one. In the above three benchmark schemes, TPC matrix is

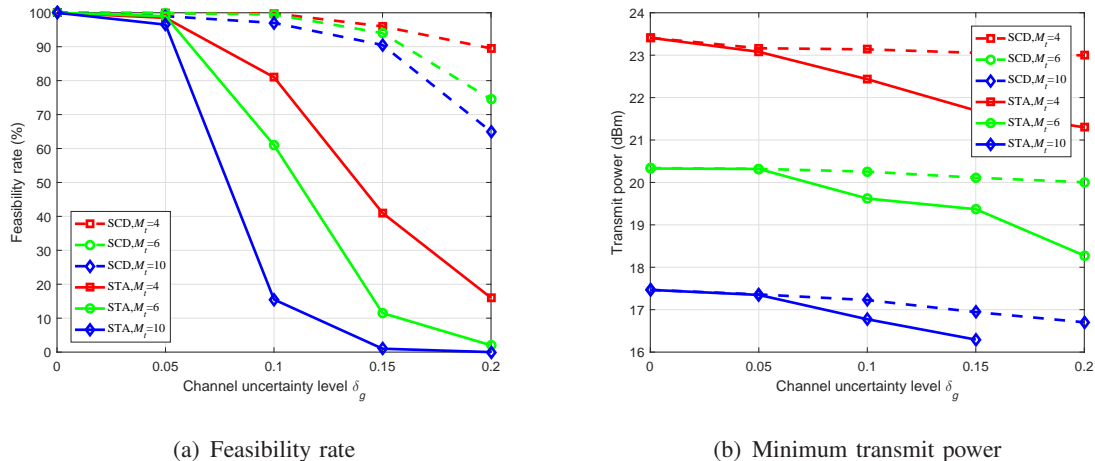


Fig. 11. Feasibility rate and minimum transmit power versus channel uncertainty level  $\delta_g$ .

optimized by the same method as used in Algorithm 2. Our proposed SCD and STA schemes outperform these three benchmark schemes. The NoIRS scheme is the worst one that verifies the benefit brought by introducing the IRS into CR systems. The Prephase and RandPhase schemes without phase shifts optimization have almost the same performance and are worse than SCD and STA schemes.

### C. Feasibility and Objective Evaluations

Fig. 11 shows the feasibility rate and minimum transmit power versus the channel uncertainty level with various numbers of transmit antennas when  $r = 2$ ,  $N = 6$ ,  $\Gamma = -80$  dBm and  $\beta = 0.05$ . The feasibility rate is defined as the ratio of the number of feasible channel to the total number of channel generations, where the feasible channel means there exists a solution for the optimization problem under this channel generation. From Fig. 11(a), we can find that the feasibility rate decreases with the increase of both  $\delta_g$  and the number of transmit antennas  $M_t$ . With the increase of  $\delta_g$ , the feasibility rate with large  $M_t$  will reduce to zero more rapidly than that with small  $M_t$ . From Fig. 11(b), the minimum transmit power will decrease with the increase of  $M_t$ , which is due to the following two reasons. The first one is that large  $M_t$  improves the degrees of freedom which can be exploited to optimize the active beamforming at the ST. The second is that the PU-related channels become worse with the increase of  $M_t$  and the IT imposed on PR becomes lower. The same results hold for  $\delta_g$ . With the increase of channel uncertainty, the PU-related channels become worse and more signal power is allocated

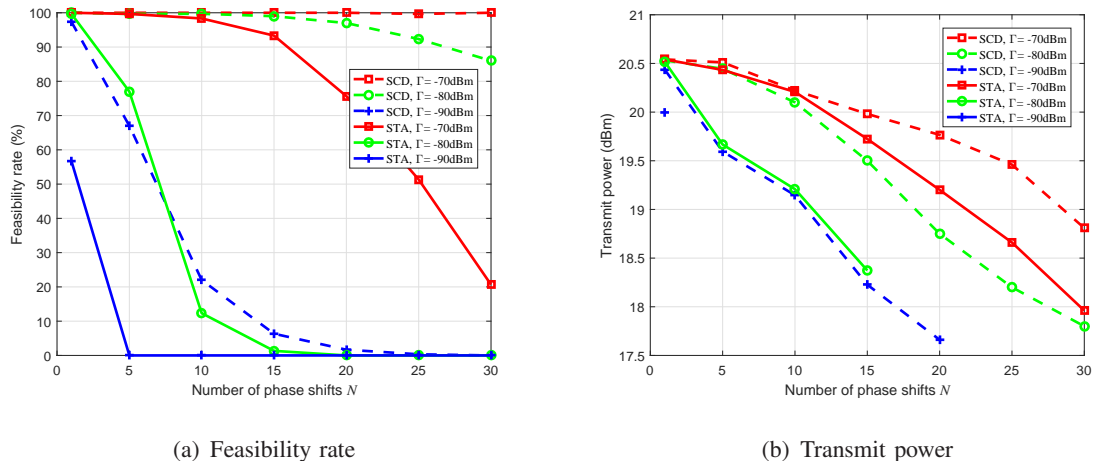


Fig. 12. Feasibility rate and minimum transmit power versus  $N$ .

to SUs, then the transmit power of ST can be reduced. Comparing Fig. 11(b) with Fig. 11(a), we can know that the values of the number of the transmit antennas and channel uncertainty level should be limited to achieve a good tradeoff between the feasibility rate and the ST's minimum transmit power in IRS-aided CR networks.

Note that, in Fig. 11(b), the blue solid line of STA doesn't complete the transmit power when  $M_t = 10$ . This is due to the fact that the corresponding feasibility rate shown in Fig. 11(a) becomes zero when the observation point in x axis is 0.2.

Fig. 12 shows the feasibility rate and the minimum transmit power versus the number of phase shifts  $N$  for various values of IT threshold  $\Gamma$  when  $r = 2$ ,  $M_t = 6$ ,  $\beta = 0.05$ ,  $\delta_g = 0.1$ . From Fig. 12(a), it is observed that both the feasibility rates of the SCD algorithm and the STA algorithm decrease with the increase of  $N$ . This is due to the fact that the cascaded ST-IRS-PR channel estimation error increases with  $N$ . Another phenomenon is that the feasibility rates of both the SCD and the STA algorithms decrease with  $\Gamma$ . The decrease of  $\Gamma$  means the feasible space about IT limitation requirement shrinks. Fig. 12(b) shows the minimum transmit power versus  $N$  for various values of  $\Gamma$ . The ST's minimum transmit power decreases with the increase of  $N$ . This is due to the fact that increasing  $N$  can enhance the reflective beamforming gain by optimizing the phase shift matrix. However, increasing  $N$  can also increase the PU-related channel estimation error, which will reduce the feasibility rate. Hence, the number of phase shifts should be carefully chosen especially when the IT threshold is lower than  $-80$  dBm. An extreme case can be found that when  $\Gamma = -90$  dBm, the optimization problem is not feasible

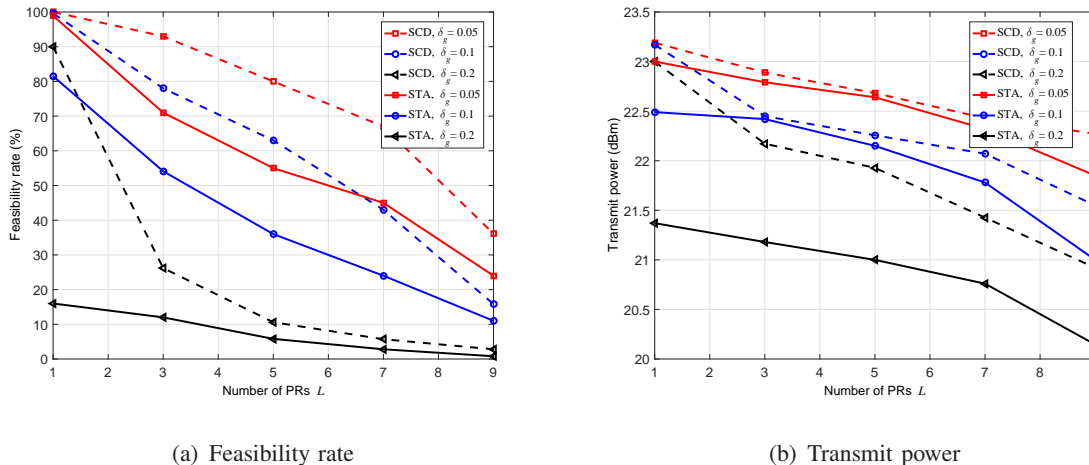


Fig. 13. Feasibility rate and minimum transmit power versus  $L$ .

at almost all range of phase shifts from 5 to 30. This results in no yield of minimum transmit power.

#### D. Multi-PR Scenario Evaluations

Fig. 13 shows the feasibility rate and the minimum transmit power versus the number of PRs  $L$  for various values of channel uncertainty level  $\delta_g$  when  $r = 2$ ,  $M_t = 4$ ,  $\beta = 0.05$ ,  $N = 6$ . The PRs are randomly located in a cell region centered at origin with radius 200 m and other scenario settings remain unchanged as shown in Fig. 4. For PR  $l$ ,  $\forall l \in \mathcal{L}$ , we set  $\delta_{g_{d,l}} = \delta_{g_{r,l}} = \delta_g$  and  $\Gamma_l = -80$  dBm. From Fig. 13(a), it is observed that both the feasibility rates of the SCD algorithm and the STA algorithm decrease with the increase of the number of PRs and channel uncertainty level. This is due to the fact that the feasibility region shrinks with the increase of the number of constraints brought by increasing  $L$  and  $\delta_g$ . Fig. 13(b) shows the minimum transmit power versus  $L$  for various values of  $\delta_g$ . For the same scheme, with the increase of  $\delta_g$ , the PU-related channels become worse and more signal power is allocated to SUs, then the transmit power of ST can be reduced. The ST's minimum transmit power decreases with the increase of  $L$ . The reason can be explained as follows. With the increase of the number of PRs, the number of PRs that are close to the ST will increase as well. In order to guarantee interference constraint imposed on PRs, the transmit power of ST should be reduced.

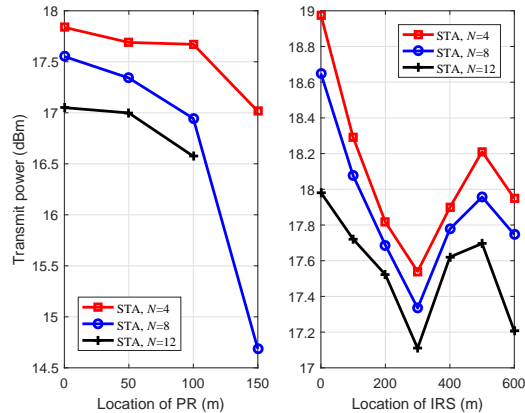


Fig. 14. Impacts of PR and IRS locations on minimum transmit power.

### E. Evaluations of Locations of PR and IRS

By moving PR from the original location (0 m, 0 m) to (150 m, 0 m) or IRS from (0 m, 30 m) to (600 m, 30 m) along the x-axis, respectively, we investigate the impact of PR location or IRS location on the performance of the network in Fig. 14, where we set  $r = 1$ ,  $M_t = 4$ ,  $\beta = 0.05$ ,  $\Gamma = -70$  dBm and  $\delta_g = 0.05$ . Fig. 14 indicates that when PR is closer to ST, the minimum transmit power is smaller. This is because that the closer the PR is to the ST, the greater the possibility of the ST interfering with the PR. The ST has to decrease the transmit power. However, this will result in reducing the number of SUs allowed to access the channel. This is illustrated by the black line in first sub-figure where the optimization problem has solution only when the location of PR is less than 150 m. Another result indicated by the second sub-figure is that the location of IRS can affect the minimum transmit power of ST. When the IRS is deployed near the SU (ST or SR), the transmit power can be lowest. This result can guide us where to deploy the IRS in IRS-aided CR systems.

## VIII. CONCLUSION

In this paper, we have investigated two types of CSI error models for the PU-related channels in IRS-aided CR networks. Two schemes, i.e., SCD and STA were proposed to jointly optimize the TPC matrix and phase shift matrix. Simulation results show that if the estimated CSI error of the PU-related channel is large, the optimization problem has a higher probability to be infeasible. Even though the CSI error is small, the number of transmit antennas at the ST and



the number of phase shifts of the IRS should be carefully chosen to balance the feasibility rate of the optimization problem and the total minimum transmit power. In order to achieve a certain feasibility rate for the SUs, the admissible access control scheme will be another research direction in our future work.

## APPENDIX A

### PROOF OF PROPOSITION 1

For the complex matrix  $\Delta\mathbf{G}$  where the vector  $\text{vec}(\Delta\mathbf{G})$  is a Gaussian distribution vector,  $\text{vec}(\Delta\mathbf{G})$  can be normalized as  $\text{vec}(\Delta\mathbf{G}) = \Sigma^{\frac{1}{2}}\mathbf{t}$ , where the vector follows the distribution of  $\mathbf{t} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ .

As  $\left\| \tilde{\phi}^H \Delta\mathbf{G}\mathbf{W} \right\|_2^2 \leq \lambda_{max}(\Sigma) \left\| \tilde{\phi}^H \right\|_2^2 \|\mathbf{W}\|_F^2 \|\mathbf{t}\|_2^2$ , the sufficient condition that the probability constraint in (28) holds is

$$\begin{aligned} & \Pr \left\{ (N+1)\lambda_{max}(\Sigma) \|\mathbf{W}\|_F^2 \|\mathbf{t}\|_2^2 \leq \Gamma - \left\| \tilde{\phi}^H \hat{\mathbf{G}}\mathbf{W} \right\|_2^2 \right\} \geq 1 - \beta \\ & \Leftrightarrow \Pr \left\{ \|\mathbf{t}\|_2^2 \leq \tilde{\Gamma} \right\} \geq 1 - \beta, \end{aligned} \quad (53)$$

where  $\tilde{\Gamma} = \frac{\Gamma - \left\| \tilde{\phi}^H \hat{\mathbf{G}}\mathbf{W} \right\|_2^2}{(N+1)\lambda_{max}(\Sigma)\|\mathbf{W}\|_F^2}$ .

Since  $\mathbf{t}$  follows a zero-mean complex Gaussian distribution,  $\|\mathbf{t}\|_2^2$  satisfies the chi-square distribution with  $2(N+1)M_t$  degrees of freedom, i.e.,  $\|\mathbf{t}\|_2^2 \sim \chi_{2(N+1)M_t}^2$ . We define the cumulative distribution function  $F_{2(N+1)M_t}(\tilde{\Gamma}) = \Pr \left\{ \|\mathbf{t}\|_2^2 \leq \tilde{\Gamma} \right\}$  and the inverse cumulative distribution function  $F_{2(N+1)M_t}^{-1}(\cdot)$ . Then we have

$$\Pr \left\{ \|\mathbf{t}\|_2^2 \leq \tilde{\Gamma} \right\} \geq 1 - \beta \Leftrightarrow F_{2(N+1)M_t}^{-1}(1 - \beta) \leq \tilde{\Gamma}. \quad (54)$$

By substituting  $\tilde{\Gamma}$  into (54), we obtain the result in (29). In practice, the inverse function of the central chi-square cumulative distribution function can be evaluated directly or be stored in a lookup table in practical implementation. The proof is completed.

## APPENDIX B

### PROOF OF THEOREM 1

Denote the collection of  $\{\mathbf{S}_k, \forall k \in \mathcal{K}\}$  as  $\mathbf{S}$  and define  $\boldsymbol{\lambda} = [\lambda_k, \forall k \in \mathcal{K}]$  and  $\mu$  as the non-negative Lagrangian multipliers associated with SINR constraints (33b) and IT constraint

(33c), respectively. Then, the Lagrangian function of problem  $\mathcal{P}2.1\_2$  is

$$\begin{aligned} \mathcal{L}(\mathbf{S}, \boldsymbol{\lambda}, \mu) &= \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{S}_k) + \sum_{k \in \mathcal{K}} \lambda_k \left( \gamma_k \left( \sum_{j \neq k, j \in \mathcal{K}} \text{tr}(\mathbf{S}_j \widehat{\mathbf{H}}_{\phi, k}) + \sigma_{s_k}^2 \right) - \text{tr}(\mathbf{S}_k \widehat{\mathbf{H}}_{\phi, k}) \right) \\ &+ \mu \left( \sum_{k=1}^K \text{tr} \left( \mathbf{S}_k \left( \vartheta F_{2(N+1)M_t}^{-1} (1 - \beta) \mathbf{I} + \widetilde{\mathbf{G}}_{\tilde{\phi}} \right) \right) - \Gamma \right) - \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{S}_k \mathbf{Z}_k), \end{aligned} \quad (55)$$

where  $\{\mathbf{Z}_k \succeq \mathbf{0}, \forall k \in \mathcal{K}\}$  denote the dual variable matrices associated with the semi-definite constraints on  $\{\mathbf{S}_k, \forall k \in \mathcal{K}\}$ .

As problem  $\mathcal{P}2.1\_2$  is a convex optimization problem, its globally optimal solution must satisfy its first-order optimality condition as:

$$\frac{\partial \mathcal{L}(\mathbf{S}, \boldsymbol{\lambda}, \mu)}{\partial \mathbf{S}_k} = \mathbf{I}_{M_t} - \lambda_k \widehat{\mathbf{H}}_{\phi, k} + \mu \left( \widetilde{\mathbf{G}}_{\tilde{\phi}} + \vartheta F_{2(N+1)M_t}^{-1} (1 - \beta) \mathbf{I}_{M_t} \right) - \mathbf{Z}_k = \mathbf{0}, \forall k \in \mathcal{K}. \quad (56)$$

Then,  $\mathbf{Z}_k$  can be represented as  $\mathbf{Z}_k = \mathbf{D} - \lambda_k \widehat{\mathbf{H}}_{\phi, k}$ , where  $\mathbf{D}$  is given by

$$\mathbf{D} = \left( \mu \vartheta F_{2(N+1)M_t}^{-1} (1 - \beta) + 1 \right) \mathbf{I}_{M_t} + \mu \widetilde{\mathbf{G}}_{\tilde{\phi}}. \quad (57)$$

According to the Karush-Kuhn-Tucker (KKT) condition, we have

$$\lambda_k \left[ \gamma_k \left( \sum_{j \neq k, j \in \mathcal{K}} \text{tr}(\mathbf{S}_j \widehat{\mathbf{H}}_{\phi, k}) + \sigma_{s_k}^2 \right) - \text{tr}(\mathbf{S}_k \widehat{\mathbf{H}}_{\phi, k}) \right] = 0, \forall k \in \mathcal{K}, \quad (58)$$

$$\mathbf{S}_k \mathbf{Z}_k = \mathbf{0}, \forall k \in \mathcal{K}. \quad (59)$$

Before proving the theorem, we first give the following lemma.

*Lemma 2:* The optimal Lagrangian dual multipliers  $\{\lambda_k, \forall k\}$  are positive, i.e.,  $\lambda_k > 0, \forall k$ .

*Proof:* This can be proved by using contradiction. Denote the optimal solution of problem  $\mathcal{P}2.1\_2$  as  $\{\mathbf{S}_k^*, \forall k\}$  and the corresponding Lagrangian multipliers as  $\{\lambda_k^*, \mu, \forall k\}$ . Assume there exists one  $\lambda_l^*$  that is zero. Then, according to KKT conditions in (58), we have

$$\gamma_l \left( \sum_{j \neq l, j \in \mathcal{K}} \text{tr}(\mathbf{S}_j^* \widehat{\mathbf{H}}_{\phi, l}) + \sigma_{s_l}^2 \right) < \text{tr}(\mathbf{S}_l^* \widehat{\mathbf{H}}_{\phi, l}). \quad (60)$$

Then, we can find a new precoding matrix  $\mathbf{S}_l^o = \rho \mathbf{S}_l^*$  with  $0 < \rho < 1$  such that

$$\gamma_l \left( \sum_{j \neq l, j \in \mathcal{K}} \text{tr}(\mathbf{S}_j^o \widehat{\mathbf{H}}_{\phi, l}) + \sigma_{s_l}^2 \right) \leq \text{tr}(\mathbf{S}_l^o \widehat{\mathbf{H}}_{\phi, l}) \quad (61)$$

holds. Then, we find a new feasible solution  $\{\mathbf{S}_k^*, k \neq l, \forall k, \mathbf{S}_l^o\}$  that yields a lower objective value than that of  $\{\mathbf{S}_k^*, \forall k\}$ , which contradicts that  $\{\mathbf{S}_k^*, \forall k\}$  is the optimal solution.  $\blacksquare$

The first term of  $\mathbf{D}$  in (57) is an identity matrix multiplied by a constant. In addition,  $\widetilde{\mathbf{G}}_{\tilde{\phi}}$  is positive definite matrix and the Lagrange multiplier  $\mu$  is nonnegative value. As a result,  $\mathbf{D}$  is

positive definite matrix with  $\text{rank}(\mathbf{D}) = M_t$ . Since  $\text{rank}(\mathbf{Z}_k) \geq \text{rank}(\mathbf{D}) - \text{rank}(\lambda_k \widehat{\mathbf{H}}_{\phi,k})$ , we obtain  $\text{rank}(\mathbf{Z}_k) \geq M_t - 1$ , where we use the fact that  $\text{rank}(\lambda_k \widehat{\mathbf{H}}_{\phi,k}) = 1$  since  $\widehat{\mathbf{H}}_{\Phi,k} = \widehat{\mathbf{h}}_{\Phi,k} \widehat{\mathbf{h}}_{\Phi,k}^H$  and  $\lambda_k$  is positive according to Lemma 2. Furthermore, according to (59), we have  $\text{rank}(\mathbf{S}_k) \leq M_t - \text{rank}(\mathbf{Z}_k)$ . Then, we have  $\text{rank}(\mathbf{S}_k) \leq 1$ . Obviously, the optimal precoder  $\mathbf{S}_k$  is not zero matrix with zero-rank. Hence, the optimal precoder  $\mathbf{S}_k$  has rank one. The proof is completed.

## REFERENCES

- [1] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2020.
- [2] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, 2019.8.
- [3] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surface." [Online]. Available: <https://arxiv.org/abs/1907.10864>
- [4] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, pp. 1–1, 2020.
- [5] 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, 2019.11.
- [6] C. Pan, H. Ren, K. Wang, M. Elkashlan, A. Nallanathan, J. Wang, and L. Hanzo, "Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer." [Online]. Available: <https://arxiv.org/abs/1908.04863>
- [7] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication." [Online]. Available: <https://arxiv.org/abs/1910.01573>
- [8] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast MISO communication systems." [Online]. Available: <https://arxiv.org/abs/1909.04606>
- [9] Y. Han, W. Tang, S. Jin, C. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8238–8242, 2019.
- [10] J. Chen, Y. Liang, Y. Pei, and H. Guo, "Intelligent reflecting surface: A programmable wireless environment for physical layer security," *IEEE Access*, vol. 7, pp. 82 599–82 612, 2019.7.
- [11] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1410–1414, 2019.10.
- [12] 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, 2019.9.
- [13] X. Yu, D. Xu, and R. Schober, "Enabling secure wireless communications via intelligent reflecting surfaces." [Online]. Available: <https://arxiv.org/abs/1904.09573>
- [14] T. Bai, C. Pan, Y. Deng, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Latency minimization for intelligent reflecting surface aided mobile edge computing." [Online]. Available: <https://arxiv.org/abs/1910.07990>
- [15] 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." [Online]. Available: <https://arxiv.org/abs/1911.06237>
- [16] W. Zhang, C. Wang, X. Ge, and Y. Chen, "Enhanced 5G cognitive radio networks based on spectrum sharing and spectrum aggregation," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6304–6316, 2018.

- [17] H. Xuemin, W. Jing, W. Cheng-Xiang, and S. Jianghong, "Cognitive radio in 5G: a perspective on energy-spectral efficiency trade-off," *IEEE Commun. Mag.*, vol. 52, no. 7, pp. 46–53, 2014.
- [18] H. Al-Hraishawi and G. A. A. Baduge, "Wireless energy harvesting in cognitive massive MIMO systems with underlay spectrum sharing," *IEEE Wireless Commun. Lett.*, vol. 6, no. 1, pp. 134–137, 2017.
- [19] F. Tseng, L. Chou, H. Chao, and J. Wang, "Ultra-dense small cell planning using cognitive radio network toward 5G," *IEEE Wireless Commun.*, vol. 22, no. 6, pp. 76–83, 2015.
- [20] D. Hamza, P. Ki-Hong, M. S. Alouini, and S. Aissa, "Throughput maximization for cognitive radio networks using active cooperation and superposition coding," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3322–3336, 2015.
- [21] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Opportunistic spectrum access via periodic channel sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 785–796, 2008.
- [22] M. Zhang and Y. Liu, "Secure beamforming for untrusted MISO cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4861–4872, 2018.
- [23] L. Zhang, Y. Wang, W. Tao, Z. Jia, T. Song, and C. Pan, "Intelligent reflecting surface aided MIMO cognitive radio systems," *IEEE Trans. Veh. Technol.*, vol. 20, no. 10, pp. 11 445–11 457, 2020.
- [24] A. U. Makarfi, R. Kharel, K. M. Rabie, O. Kaiwartya, X. Li, and D.-T. Do, "Reconfigurable intelligent surfaces based cognitive radio networks." [Online]. Available: <https://arxiv.org/abs/2003.07467>
- [25] J. He, K. Yu, Y. Zhou, and Y. Shi, "Reconfigurable intelligent surface enhanced cognitive radio networks." [Online]. Available: <https://arxiv.org/abs/2005.10995>
- [26] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces." [Online]. Available: <https://arxiv.org/abs/1912.0149>
- [27] Z. Zhou, N. Ge, Z. Wang, and L. Hanzo, "Joint transmit precoding and reconfigurable intelligent surface phase adjustment: A decomposition-aided channel estimation approach." [Online]. Available: <https://www.researchgate.net/publication/337824343>
- [28] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis." [Online]. Available: <https://arxiv.org/abs/1912.11783>
- [29] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation and joint beamforming for intelligent reflecting surface-assisted millimeter wave systems." [Online]. Available: <https://arxiv.org/abs/1911.07202>
- [30] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems." [Online]. Available: <https://arxiv.org/abs/1912.03619>
- [31] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels," *IEEE Trans. Signal Process.*, vol. 68, pp. 5092–5106, 2020.
- [32] D. Xu, X. Yu, Y. Sun, D. W. K. Ng, and R. Schober, "Resource allocation for IRS-assisted full-duplex cognitive radio systems," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7376–7394, 2020.
- [33] J. Yuan, Y.-C. Liang, J. Joung, G. Feng, and E. G. Larsson, "Intelligent reflecting surface-assisted cognitive radio system." [Online]. Available: <https://arxiv.org/abs/1912.10678>
- [34] Z. Lan, L. Ying-Chang, X. Yan, and H. V. Poor, "Robust cognitive beamforming with partial channel state information," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4143–4153, 2009.
- [35] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [36] 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, 2013.
- [37] 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, 2014.