Novel Relax-and-Retract Algorithm for Intelligent Reflecting Surface Design

Xin He, Lei Huang, *Senior Member, IEEE*, and Jiangzhou Wang, *Fellow, IEEE*

***Abstract*—In this paper, a novel relax-and-retract algorithm is proposed to tackle the nonconvex unit-modulus constraint in the joint intelligent reflecting surface (IRS) and multiuser multiple-input multiple-output (MU-MIMO) transceiver design. The conventional method to tackle the unit-modulus constraint is semidefinite relaxation (SDR), and its computational complexity is large. By using the symbol detection mean square error (MSE) as the quality of service (QoS), the proposed relax-and-retract approach enables us to get convex quadratically constrained quadratic programming (QCQP) subproblems, which have a much lower computational complexity than the conventional SDR approach. Simulation results show that the proposed relax-and-retract approach has excellent performance in terms of computational complexity, while the transmit power and the unit-modulus hardware implementation are the same as those of the SDR approach.**

***Index Terms*—Constant-modulus constraint, convex QCQP, intelligent reflecting surface, MU-MIMO transceiver design, relax-and-retract algorithm, SDR, unit-modulus constraint.**

I. Introduction

Intelligent reflecting surface (IRS) has a large number of low-cost passive reflecting elements with adjustable phase shifts and tuneable modulus, thus collaboratively changing the reflected signal propagation and significantly enhancing the performance of wireless communications [1]. Compared to the relay system, IRS is passive with full-duplex mode, and has no self-interference and noise amplification [2]. The massive multiple-input multiple-output (MIMO) relay system is involved with costly radio frequency (RF) chains, and its active antennas' power consumption is very high, while the IRS does not need any RF chains and its power consumption is negligible [1].

The IRS has been applied in many wireless communication systems, e.g., the IRS was proposed in the single user communication system [3], [4], the multiuser multiantenna system [5], the orthogonal frequency division multiplexing system [6], the wireless power transfer system [7] and the secure wireless communication system [8]–[10]. In order to simplify the IRS implementation, with signal-to-interference-plus-noise-ratio (SINR) as the quality of service (QoS), the modulus of the IRS is usually fixed as one [1], [6] and only the phase of the IRS system can be changed. Therefore, the major challenge of the IRS design lies in the nonconvex unit-modulus constraint, and the leading approach to solve the unit-modulus problem is semidefinite relaxation (SDR) [1], [6], [8]–[10]. However, the computational complexity of the SDR approach is high, and the simple method only exists for special cases, e.g., the single user situation [1], [6]. To tackle the SDR complexity issue, by using linear approximations, the majorization-minimization (MM) algorithm, complex circle manifold (CCM) method, and successive convex approximation (SCA) method were proposed [7], [11]. Owing to their linear convergence rate, the iteration number is generally large. With sum-rate as the objective function, the penalty dual decomposition (PDD) method was proposed in [12]. Although the subproblem of the PDD method is simple, the inner and outer loops and the non-monotonic convergence behavior make the iteration number large.

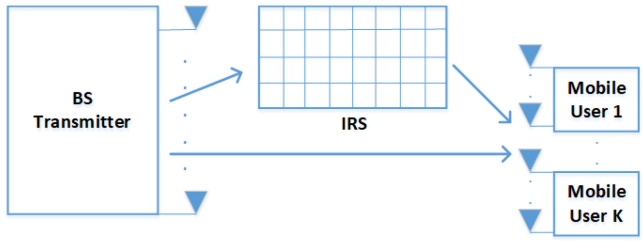

Fig. 1. IRS-aided MU-MIMO communication system.

In this work, instead of the SDR approach, a novel relax-and-retract algorithm is proposed to tackle the nonconvex unit-modulus constraint in the joint IRS and MU-MIMO transceiver design. By using the symbol detection mean square error (MSE) as the QoS criterion [13], the proposed relax-and-retract approach ensures that its subproblems are convex quadratically constrained quadratic programming (QCQP) subproblems, which have a much lower computational complexity than the SDR approach. Since the convex MSE function does not have a linear upper bound, the MM algorithm, the CCM method, and the SCA method cannot be applied to the proposed problem. Simulation results show that the proposed relax-and-retract approach has excellent performance in terms of computational complexity, while the transmit power and the unit-modulus hardware implementation are the same as those of the SDR approach.

Notation: In this paper, $\mathbb{E}(\cdot)$, $(\cdot)^T$, and $(\cdot)^H$ denote statistical expectation, transposition and Hermitian, respectively. The $\|\cdot\|_2$ denotes the norm of a vector, $\mathrm{Tr}(\cdot)$ and $\|\cdot\|_F$ stand for the trace and Frobenius norm of a matrix, respectively. The $\mathrm{vec}(\cdot)$ stands for the vectorization, $\mathrm{diag}(\mathbf{A})$ is a vector with its elements being from the diagonal elements of the square matrix $\mathbf{A}$, and $\mathrm{Diag}(\mathbf{a})$ is a diagonal matrix with its diagonal elements being from the elements of the vector $\mathbf{a}$. For a vector $\boldsymbol{\theta}$, $\boldsymbol{\theta}(n)$ is the n-th element of $\boldsymbol{\theta}$. The $\mathbf{I}_K$ is a $K \times K$ identity matrix. The operator $\oplus_{i=1}^{2}\mathbf{A}_i = \left[\begin{smallmatrix}\mathbf{A}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{A}_2\end{smallmatrix}\right]$ is adopted.

II. System Model

The IRS-aided MU-MIMO transceiver design in the downlink communication system is illustrated in Fig. 1. The base station (BS) has N transmit antennas, the IRS has N_R reflecting elements, and K active users with the k-th user equipped with M_k antennas. The $N_R \times N$ channel from the BS to the IRS is denoted as $\mathbf{H}$, the $M_k \times N_R$ channel

Manuscript received October 18, 2020; revised December 27, 2020; accepted January 21, 2021. Date of publication January 27, 2021; date of current version March 10, 2021. This work was supported in part by the National Science Fund for Distinguished Young Scholars under Grant 61925108, and in part by the National Natural Science Foundation of China under Grants 61807018, U1713217, and U1913221. The review of this article was coordinated by Dr. B. Di. *(Corresponding author: Lei Huang.)*

Xin He is with the College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, China, and also with the Department of Computer and Information Engineering, Hanshan Normal University, Chaozhou 521041, China (e-mail: hx66hx@gmail.com).

Lei Huang is with the Shenzhen Key Laboratory of Advanced Navigation Technology and Guangdong Key Laboratory of Intelligent Information Processing, College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, China (e-mail: lhuang@szu.edu.cn).

Jiangzhou Wang is with the School of Engineering and Digital Arts, University of Kent, Canterbury, Kent CT2 7NZ, U.K. (e-mail: j.z.wang@kent.ac.uk).

Digital Object Identifier 10.1109/TVT.2021.3054516

from the IRS to the k-th user is represented as $\mathbf{H}_{r,k}$, the $M_k \times N$ channel $\mathbf{H}_{d,k}$ is from the BS directly to the k-th user. The L_k independent data streams are transmitted to the k-th user and $\sum_{k=1}^{K} L_k = L$. To guarantee data recovery, it is required that $L_k \leq M_k$ and $L \leq N$.

Let $\mathbf{G}$ be the $N \times L$ precoding matrix at BS, $\mathbf{F}_k$ be the $L_k \times M_k$ receiver of the k-th user, and $\mathbf{\Theta} = \mathrm{Diag}(\boldsymbol{\theta})$ be the $N_R \times N_R$ processing matrix at the IRS. At the k-th mobile user, the received signal is

$$\mathbf{y}_k = (\mathbf{H}_{r,k} \underbrace{\mathrm{Diag}(\boldsymbol{\theta})}_{\mathbf{\Theta}} \mathbf{H} + \mathbf{H}_{d,k})\mathbf{G}\mathbf{x} + \mathbf{n}_k, \tag{1}$$

where $\mathbf{x}$ is the transmitted symbol with $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}_L$, and the noise $\mathbf{n}_k$ is circularly-symmetric complex Gaussian i.e., $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_k)$. The noise covariance matrix $\mathbf{\Sigma}_k \succeq 0$ has to be bounded with $\mathrm{Tr}(\mathbf{\Sigma}_k) \leq c_k$, where c_k is a positive finite constant. Owing to the complexity limitation of hardware implementation in IRS design, the constraint on $\boldsymbol{\theta}$ is,

$$|\boldsymbol{\theta}(n)| = 1, \forall n \in [1, N_R], \tag{2}$$

where the unit-modulus constraint only involves phase rotation.

At the k-th mobile user's receiver, an $L_k \times M_k$ receiver $\mathbf{F}_k$ is used to filter the received signal $\mathbf{y}_k$. Then the recovered $L_k \times 1$ data stream is

$$\hat{\mathbf{x}}_k = \mathbf{F}_k(\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k})\mathbf{G}\mathbf{x} + \mathbf{F}_k\mathbf{n}_k. \tag{3}$$

Since the transmitted data $\mathbf{x}$ are independent of the noise $\{\mathbf{n}_k\}_{k=1}^K$, the symbol detection MSE of the k-th user is

$$\mathrm{MSE}_k = \mathbb{E}_{\mathbf{x},\mathbf{n}_k}[\mathrm{Tr}\{(\mathbf{D}_k\mathbf{x} - \hat{\mathbf{x}}_k)(\mathbf{D}_k\mathbf{x} - \hat{\mathbf{x}}_k)^H\}], \tag{4}$$

$$= \|\mathbf{F}_k(\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k})\mathbf{G} - \mathbf{D}_k\|_F^2 + \mathrm{Tr}(\mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k^H), \tag{5}$$

where the $L_k \times L$ matrix $\mathbf{D}_k = [\mathbf{0}_{L_k \times \sum_{i=1}^{k-1} L_i} \ \mathbf{I}_{L_k} \ \mathbf{0}_{L_k \times \sum_{i=k+1}^{K} L_i}]$ is the k-th user's data selection matrix. Note that $\oplus_{k=1}^K \mathbf{I}_{L_k} = \mathbf{I}_L$.

III. Joint IRS and MU-MIMO Transceiver Design

The joint optimization of the IRS processing matrix $\mathbf{\Theta}$ and the MU-MIMO transceiver $(\mathbf{G}, \{\mathbf{F}_k\}_{k=1}^K)$ is presented as follows. The joint design is to minimize the transmit power at the BS, subject to K users' MSE requirements and the constraint on $\boldsymbol{\theta}$, i.e.,

$$\begin{aligned} &\min_{\mathbf{G},\mathbf{\Theta},\{\mathbf{F}_k\}_{k=1}^K} \ \|\mathbf{G}\|_F^2 \\ &\text{s.t. } \|\mathbf{F}_k(\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k})\mathbf{G} - \mathbf{D}_k\|_F^2 + \mathrm{Tr}(\mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k^H) \\ &\quad \leq \varepsilon_k, \ \forall k \in [1, K], \\ &\quad |\boldsymbol{\theta}(n)| = 1, \ \forall n \in [1, N_R], \end{aligned} \tag{6}$$

where ε_k is the k-th user's MSE target. Since the relationship between MSE and SINR is $\mathrm{MSE} = 1/(1 + \mathrm{SINR})$ [13], the MSE requirements also guarantee the SINR performances of all users.

In order to determine whether problem (6) is feasible, the following definition is presented,

$$\mathbf{H}_0 = \begin{bmatrix} \mathbf{H}_{r,1}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,1} \\ \cdots \\ \mathbf{H}_{r,K}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,K} \end{bmatrix}. \tag{7}$$

Subsequently, we have following conclusions.

Proposition 1: If there is a $\{\boldsymbol{\theta} | \{|\boldsymbol{\theta}(n)| = 1\}_{n=1}^{N_R}\}$ such that $\mathrm{rank}(\mathbf{H}_0) \geq L$ and $\{\mathrm{rank}(\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k}) \geq L_k\}_{k=1}^K$, then problem (6) is feasible.

Proof: Let us fix the equalizers as,

$$\mathbf{F}_k^H = \frac{1}{\sqrt{t_k}}\mathbf{V}_k, \ \forall k \in [1, K], \tag{8}$$

where $\mathbf{V}_k$ is the L_k normalized left singular vectors of $\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k}$ associated with its L_k largest singular values and $t_k \geq c_k/\varepsilon_k$. Therefore, it is easy to prove that $\{\mathrm{Tr}(\mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k^H) \leq \varepsilon_k\}_{k=1}^K$. Furthermore, since $\{\mathrm{rank}(\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k}) \geq L_k\}_{k=1}^K$, none of the row vectors of $\mathbf{F}_k$ lie in the left null space of $\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k}$.

By using the equalizers in (8), we have an $L \times N$ matrix,

$$\begin{bmatrix} \mathbf{F}_1(\mathbf{H}_{r,1}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,1}) \\ \cdots \\ \mathbf{F}_K(\mathbf{H}_{r,K}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,K}) \end{bmatrix} = \oplus_{k=1}^K \mathbf{F}_k \cdot \mathbf{H}_0 = \mathbf{H}_1. \tag{9}$$

Since the row vectors in equalizers (8) are orthogonal vectors, the block diagonal structure in $\oplus_{k=1}^K \mathbf{F}_k$ makes its L row vectors are also orthogonal vectors. Furthermore, it can be extended from the upper paragraph that none of the row vectors of $\oplus_{k=1}^K \mathbf{F}_k$ lies in the null space of $\mathbf{H}_0^T$. Together with the condition $\mathrm{rank}(\mathbf{H}_0) \geq L$, the L row vectors of $\oplus_{k=1}^K \mathbf{F}_k \cdot \mathbf{H}_0$ are linearly independent, i.e., $\mathrm{rank}(\mathbf{H}_1) = L$. Therefore, letting $\mathbf{G} = \mathbf{H}_1^H(\mathbf{H}_1\mathbf{H}_1^H)^{-1}$, we have

$$\begin{bmatrix} \mathbf{F}_1(\mathbf{H}_{r,1}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,1})\mathbf{G} \\ \cdots \\ \mathbf{F}_K(\mathbf{H}_{r,K}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,K})\mathbf{G} \end{bmatrix} = \mathbf{H}_1\mathbf{G} = \mathbf{I}_L = \oplus_{k=1}^K \mathbf{I}_{L_k}, \tag{10}$$

which makes $\{\|\mathbf{F}_k(\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k})\mathbf{G} - \mathbf{D}_k\|_F^2 = 0\}_{k=1}^K$. Therefore, there exists a joint design $(\mathbf{G}, \boldsymbol{\theta}, \{\mathbf{F}_k\}_{k=1}^K)$ to make problem (6) feasible. ■

By using the simple singular value decomposition (SVD), the problem (6) is infeasible if the rank conditions in Proposition 1 are not satisfied. In this case, the optimization effort for problem (6) is saved. Since $\mathbf{\Theta} = \mathrm{Diag}(\boldsymbol{\theta})$ is a full rank diagonal matrix for any $|\boldsymbol{\theta}(n)| = 1$, the impact of $\mathbf{\Theta}$ on the rank of $\mathbf{H}_0$ is small. The rank of $\mathbf{H}_0$ depends heavily on the correlations in and between $\mathbf{H}$, $\{\mathbf{H}_{r,k}\}_{k=1}^K$, $\{\mathbf{H}_{d,k}\}_{k=1}^K$. Therefore, $\mathbf{\Theta} = \mathbf{I}_{N_R}$ is a good choice to check Proposition 1, and an example is illustrated in the simulation section.

Remark 1: The channel $\mathbf{H}_{d,k}$ can be estimated by switching off the IRS and using the conventional channel estimation method. With low-power receive RF chain in IRS [14], the channels $\mathbf{H}$ and $\mathbf{H}_{r,k}$ can be estimated at IRS by using the conventional channel estimation method and the channel reciprocity property in the time-division duplexing (TDD) system. To consider the channel estimation errors, the robust QoS constrained beamforming problem is quite different from the conventional robust designs with Gaussian perturbation [15]–[17]. Since the distribution of the product of the errors in $\mathbf{H}$ and $\mathbf{H}_{r,k}$ is difficult to obtain, the robust QoS constrained IRS design is a challenge task [12].

IV. Iterative IRS and Transceiver Design

Although problem (6) is a difficult nonconvex problem on $(\mathbf{G}, \boldsymbol{\theta}, \{\mathbf{F}_k\}_{k=1}^K)$, the block coordinate descent methodology is used to find $\mathbf{G}, \boldsymbol{\theta}, \{\mathbf{F}_k\}_{k=1}^K$ sequentially.

With fixed IRS processing matrix $\mathbf{\Theta}$ and receivers $\{\mathbf{F}_k\}_{k=1}^K$, the optimal transmitter $\mathbf{G}$ is obtained through

$$\begin{aligned} &\min_{\mathbf{G}} \ \mathrm{Tr}(\mathbf{G}\mathbf{G}^H) \\ &\text{s.t. } \|\mathbf{F}_k(\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k})\mathbf{G} - \mathbf{D}_k\|_F^2 + \mathrm{Tr}(\mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k^H) \\ &\quad \leq \varepsilon_k, \ \forall k \in [1, K], \end{aligned} \tag{11}$$

which is a convex QCQP problem.

With fixed IRS processing matrix $\mathbf{\Theta}$ and transmitter $\mathbf{G}$, the optimal receiver $\mathbf{F}_k$ is obtained through minimizing the guaranteed MSE,

$$\min_{\mathbf{F}_k} \|\mathbf{F}_k(\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H}+\mathbf{H}_{d,k})\mathbf{G}-\mathbf{D}_k\|_F^2+\mathrm{Tr}(\mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k^H), \quad (12)$$

which is a convex problem, its optimal solution occurs at the unique stationary point $\mathbf{F}_k=\mathbf{D}_k\mathbf{H}_{\mathbf{G},k}^H(\mathbf{H}_{\mathbf{G},k}\mathbf{H}_{\mathbf{G},k}^H+\mathbf{\Sigma}_k)^{-1}$ with $\mathbf{H}_{\mathbf{G},k}=(\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H}+\mathbf{H}_{d,k})\mathbf{G}$.

With fixed transmitter $\mathbf{G}$ and receivers $\{\mathbf{F}_k\}_{k=1}^K$, the optimization problem to get the optimal IRS is obtained by minimizing the total MSE of all K users under individual MSE constraints,

$$\begin{aligned}&\min_{\boldsymbol{\theta}} \sum_{k=1}^K \|\mathbf{F}_k(\mathbf{H}_{r,k}\mathrm{Diag}(\boldsymbol{\theta})\mathbf{H}+\mathbf{H}_{d,k})\mathbf{G}-\mathbf{D}_k\|_F^2\\ &\text{s.t. } \|\mathbf{F}_k(\mathbf{H}_{r,k}\mathrm{Diag}(\boldsymbol{\theta})\mathbf{H}+\mathbf{H}_{d,k})\mathbf{G}-\mathbf{D}_k\|_F^2+\mathrm{Tr}(\mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k^H)\\ &\quad\le\varepsilon_k,\ \forall k\in[1,K],\\ &\quad|\boldsymbol{\theta}(n)|=1,\ \forall n\in[1,N_R],\end{aligned} \quad (13)$$

which is a difficult nonconvex problem owing to the unit-modulus constraints. Note that the constant term $\mathrm{Tr}(\mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k^H)$ is eliminated from the cost function of (13).

A. IRS Design Based on SDR

In order to tackle the nonconvex constraint $|\boldsymbol{\theta}(n)|=1$, the conventional SDR approach needs the following reformulations. Denoting the i-th column vector of $\mathbf{H}$ as $\mathbf{H}[:,i]$, we have the algebraic transformations for the objective function of problem (13),

$$\begin{aligned}&\|\mathbf{F}_k\mathbf{H}_{r,k}\mathrm{Diag}(\boldsymbol{\theta})\overbrace{\mathbf{H}\mathbf{G}}^{\mathbf{H}_G}+\mathbf{F}_k\mathbf{H}_{d,k}\mathbf{G}-\mathbf{D}_k\|_F^2 && (14)\\ &=\|\mathbf{F}_k\mathbf{H}_{r,k}\left[\mathrm{Diag}(\mathbf{H}_G[:,1])\boldsymbol{\theta},\ldots,\mathrm{Diag}(\mathbf{H}_G[:,L])\boldsymbol{\theta}\right]\\ &\quad+\mathbf{F}_k\mathbf{H}_{d,k}\mathbf{G}-\mathbf{D}_k\|_F^2 && (15)\\ &=\|\left[\mathbf{F}_k\mathbf{H}_{r,k}\mathrm{Diag}(\mathbf{H}_G[:,1])\boldsymbol{\theta},\ldots,\mathbf{F}_k\mathbf{H}_{r,k}\mathrm{Diag}(\mathbf{H}_G[:,L])\boldsymbol{\theta}\right]\\ &\quad+\mathbf{F}_k\mathbf{H}_{d,k}\mathbf{G}-\mathbf{D}_k\|_F^2 && (16)\\ &=\|\mathrm{vec}\Big(\left[\mathbf{F}_k\mathbf{H}_{r,k}\mathrm{Diag}(\mathbf{H}_G[:,1])\boldsymbol{\theta},\ldots,\mathbf{F}_k\mathbf{H}_{r,k}\mathrm{Diag}(\mathbf{H}_G[:,L])\boldsymbol{\theta}\right]\\ &\quad+\mathbf{F}_k\mathbf{H}_{d,k}\mathbf{G}-\mathbf{D}_k\Big)\|_2^2 && (17)\\ &=\|\underbrace{\begin{bmatrix}\mathbf{F}_k\mathbf{H}_{r,k}\mathrm{Diag}(\mathbf{H}_G[:,1])\\ \cdots\\ \mathbf{F}_k\mathbf{H}_{r,k}\mathrm{Diag}(\mathbf{H}_G[:,L])\end{bmatrix}}_{\mathbf{A}_k}\boldsymbol{\theta}+\underbrace{\mathrm{vec}(\mathbf{F}_k\mathbf{H}_{d,k}\mathbf{G}-\mathbf{D}_k)}_{\mathbf{b}_k}\|_2^2 && (18)\\ &=\|\mathbf{A}_k\boldsymbol{\theta}+\mathbf{b}_k\|_2^2=\|[\mathbf{A}_k,\mathbf{b}_k]\underbrace{[\boldsymbol{\theta}^T,1]^T}_{\bar{\boldsymbol{\theta}}}\|_2^2 && (19)\\ &=\mathrm{Tr}([\mathbf{A}_k,\mathbf{b}_k]^H[\mathbf{A}_k,\mathbf{b}_k]\bar{\boldsymbol{\theta}}\bar{\boldsymbol{\theta}}^H) && (20)\end{aligned}$$

After letting $\mathbf{Q}_{\bar{\boldsymbol{\theta}}}\triangleq\bar{\boldsymbol{\theta}}\bar{\boldsymbol{\theta}}^H$, problem (13) is equivalent to

$$\begin{aligned}&\min_{\mathbf{Q}_{\bar{\boldsymbol{\theta}}}\succeq 0}\sum_{k=1}^K\mathrm{Tr}([\mathbf{A}_k,\mathbf{b}_k]^H[\mathbf{A}_k,\mathbf{b}_k]\mathbf{Q}_{\bar{\boldsymbol{\theta}}})\\ &\text{s.t. } \mathrm{Tr}([\mathbf{A}_k,\mathbf{b}_k]^H[\mathbf{A}_k,\mathbf{b}_k]\mathbf{Q}_{\bar{\boldsymbol{\theta}}})+\mathrm{Tr}(\mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k^H)\le\varepsilon_k,\forall k\in[1,K]\\ &\quad\mathrm{diag}(\mathbf{Q}_{\bar{\boldsymbol{\theta}}})=\mathbf{1},\mathrm{rank}(\mathbf{Q}_{\bar{\boldsymbol{\theta}}})=1,\end{aligned} \quad (21)$$

which is a nonconvex problem owing to the rank-one constraint. By dropping the rank-one constraint, the problem (21) becomes the convex semidefinite programming (SDP) problem

$$\begin{aligned}&\min_{\mathbf{Q}_{\bar{\boldsymbol{\theta}}}\succeq 0}\sum_{k=1}^K\mathrm{Tr}([\mathbf{A}_k,\mathbf{b}_k]^H[\mathbf{A}_k,\mathbf{b}_k]\mathbf{Q}_{\bar{\boldsymbol{\theta}}})\\ &\text{s.t. } \mathrm{Tr}([\mathbf{A}_k,\mathbf{b}_k]^H[\mathbf{A}_k,\mathbf{b}_k]\mathbf{Q}_{\bar{\boldsymbol{\theta}}})+\mathrm{Tr}(\mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k^H)\le\varepsilon_k,\forall k\in[1,K]\\ &\quad\mathrm{diag}(\mathbf{Q}_{\bar{\boldsymbol{\theta}}})=\mathbf{1}.\end{aligned} \quad (22)$$

If the rank of the optimal solution in the convex problem (22) is one, the optimal solution of the nonconvex problem (21) is also obtained. Although the SDR approach in (22) is not guaranteed to get a rank-one solution, the well-known Gaussian randomization method in [18] provides suboptimal solution for (13) and (21). First, the random realizations $\{\mathbf{r}_i\}_{i=1}^R$ are generated from $\mathcal{CN}(\mathbf{0},\mathbf{I}_{N_R+1})$. Under eigendecomposition $\mathbf{Q}_{\bar{\boldsymbol{\theta}}}=\mathbf{V}_q\mathbf{\Sigma}_q\mathbf{V}_q^H$ for the optimal solution $\mathbf{Q}_{\bar{\boldsymbol{\theta}}}$ in (22), $\{\mathbf{V}_q\mathbf{\Sigma}_q^{\frac{1}{2}}\mathbf{r}_i\}_{i=1}^R$ are constructed. Second, we can construct the candidates $\{\mathbf{Q}_{\bar{\boldsymbol{\theta}}}^i=\bar{\boldsymbol{\theta}}_i\bar{\boldsymbol{\theta}}_i^H\}_{i=1}^R$, where $\bar{\boldsymbol{\theta}}_i$ is the elements in vector $\mathbf{V}_q\mathbf{\Sigma}_q^{\frac{1}{2}}\mathbf{r}_i$ divided by their modulus. Third, the feasible $\mathbf{Q}_{\bar{\boldsymbol{\theta}}}^i$ with minimal objective value in problem (22) is selected, and the corresponding $\bar{\boldsymbol{\theta}}_i$ is denoted as $\bar{\boldsymbol{\theta}}^\star$. Finally, the solution for (13) is obtained as $\bar{\boldsymbol{\theta}}^\star(1{:}N_R)/\bar{\boldsymbol{\theta}}^\star(1+N_R)$, where $\bar{\boldsymbol{\theta}}^\star(1:N_R)$ is the first N_R elements of $\bar{\boldsymbol{\theta}}^\star$.

B. IRS Design Based on Relax-and-Retract

For the difficult nonconvex constraint $|\boldsymbol{\theta}(n)|=1$ in IRS design problem (13), instead of the conventional SDR approach, we propose to relax the constraint into a convex constraint $|\boldsymbol{\theta}(n)|\le 1$ and then retract the solution to the unit circle.

1) Relax: After relaxing the nonconvex constraint $|\boldsymbol{\theta}(n)|=1$ into the convex constraint $|\boldsymbol{\theta}(n)|\le 1$, the IRS design problem (13) is relaxed as,

$$\begin{aligned}&\min_{\boldsymbol{\theta}} \sum_{k=1}^K \|\mathbf{F}_k(\mathbf{H}_{r,k}\mathrm{Diag}(\boldsymbol{\theta})\mathbf{H}+\mathbf{H}_{d,k})\mathbf{G}-\mathbf{D}_k\|_F^2\\ &\text{s.t. } \|\mathbf{F}_k(\mathbf{H}_{r,k}\mathrm{Diag}(\boldsymbol{\theta})\mathbf{H}+\mathbf{H}_{d,k})\mathbf{G}-\mathbf{D}_k\|_F^2+\mathrm{Tr}(\mathbf{F}_k\mathbf{\Sigma}_k\mathbf{F}_k^H)\\ &\quad\le\varepsilon_k,\ \forall k\in[1,K]\\ &\quad|\boldsymbol{\theta}(n)|\le 1,\ \forall n\in[1,N_R],\end{aligned} \quad (23)$$

which is a convex QCQP problem. Denote the optimal solution of problem (23) as $\boldsymbol{\theta}_i$.

2) Retract: Since the optimal solution $\boldsymbol{\theta}_i$ in (23) may not occurs at the unit circle, direct projection to the unit circle is a way to satisfy the unit-modulus constraint, i.e.,

$$\boldsymbol{\theta}_i^p=[\boldsymbol{\theta}_i(1)/|\boldsymbol{\theta}_i(1)|,\ldots,\boldsymbol{\theta}_i(N_R)/|\boldsymbol{\theta}_i(N_R)|]^T. \quad (24)$$

However, the projection solution $\boldsymbol{\theta}_i^p$ may not satisfy the MSE constraint. From the projection point $\boldsymbol{\theta}_i^p$, we can construct tangent planes on the unit circles $|\boldsymbol{\theta}(n)|^2=1$ as,

$$\mathrm{Re}\left(\boldsymbol{\theta}_i^p(n)\boldsymbol{\theta}^H(n)\right)=1,\ \forall n\in[1,N_R]. \quad (25)$$

Next, in order to cover both $\boldsymbol{\theta}_i$ and its projection $\boldsymbol{\theta}_i^p$, we can shift the tangent planes (25) and construct half-spaces

$$\mathrm{Re}\left(\boldsymbol{\theta}_i^p(n)\boldsymbol{\theta}^H(n)\right)\ge 1-\boldsymbol{\Delta}(n),\ \boldsymbol{\Delta}(n)\ge 0,\ \forall n\in[1,N_R], \quad (26)$$

where the shift terms $\{\boldsymbol{\Delta}(n)\}_{n=1}^{N_R}$ represent the distances between the tangent planes (25) and their corresponding boundaries of the half-spaces (26). By pushing the boundaries of half-spaces (26) to the tangent

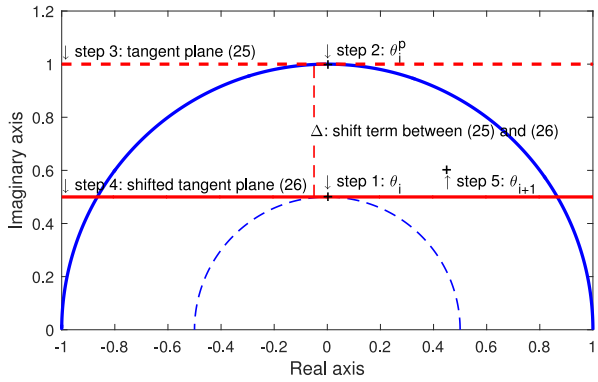

Fig. 2. Illustration of the retract operation. The region above the solid red line and below the solid blue curve contains the subset of the feasible set in problem (27). Then, any feasible solution in this region has larger modulus than $\boldsymbol{\theta}_i$, i.e., $\|\boldsymbol{\theta}_{i+1}\|_2^2 \geq \|\boldsymbol{\theta}_i\|_2^2$.

Algorithm 1: IRS Retract Operation.

1: **initialization:** Let $\boldsymbol{\theta}_i$ be the optimal solution of problem (23).
2: **repeat**
3: Find the projected solution $\boldsymbol{\theta}_i^p$, solve the convex QCQP problem (27), denote the optimal solution as $\boldsymbol{\theta}_{i+1}$. Update $i = i + 1$.
4: **until** $\|\boldsymbol{\theta}_i\|_2^2 - \|\boldsymbol{\theta}_{i-1}\|_2^2 \leq \epsilon_1$

planes (25), we can find better feasible solution with larger modulus, i.e.,

$$\begin{aligned} &\min_{\boldsymbol{\theta},\boldsymbol{\Delta}} \ \|\boldsymbol{\Delta}\|_2^2 \\ &\text{s.t. } \|\mathbf{F}_k(\mathbf{H}_{r,k}\text{Diag}(\boldsymbol{\theta})\mathbf{H} + \mathbf{H}_{d,k})\mathbf{G} - \mathbf{D}_k\|_F^2 + \text{Tr}(\mathbf{F}_k\boldsymbol{\Sigma}_k\mathbf{F}_k^H) \\ &\quad \leq \varepsilon_k, \ \forall k \in [1, K] \\ &\quad |\boldsymbol{\theta}(n)| \leq 1, \ \forall n \in [1, N_R], \\ &\quad \text{Re}\left(\boldsymbol{\theta}_i^p(n)\boldsymbol{\theta}^H(n)\right) \geq 1 - \boldsymbol{\Delta}(n), \ \boldsymbol{\Delta}(n) \geq 0, \ \forall n \in [1, N_R], \end{aligned} \tag{27}$$

Since $\boldsymbol{\theta}_i$ is a feasible solution of problem (27),[1] it is always feasible. Denote the optimal solution of problem (27) as $\boldsymbol{\theta}_{i+1}$, owing to the design criterion of problem (27), we have

$$\|\boldsymbol{\theta}_i\|_2^2 \leq \|\boldsymbol{\theta}_{i+1}\|_2^2. \tag{28}$$

Besides the optimization criterion in problem (27), the conclusion (28) can also be illustrated geometrically in Fig. 2. Based on the optimal solution $\boldsymbol{\theta}_{i+1}$, we can construct the projected solution $\boldsymbol{\theta}_{i+1}^p$ and similar half-spaces in (26), and find larger modulus solution again. Therefore, the IRS retract operation is proposed in Algorithm 1.

Since the optimal solution $\boldsymbol{\theta}_{i+1}$ of problem (27) satisfies $|\boldsymbol{\theta}_{i+1}(n)| \leq 1, \forall n \in [1, N_R]$, we have $\|\boldsymbol{\theta}_{i+1}\|_2^2 \leq N_R$. Together with the monotonic increasing property in (28), the Algorithm 1 is guaranteed to converge.

C. Iterative Algorithm and Complexity Analysis

The algorithm to solve the joint IRS and MU-MIMO transceiver design problem (6) is described in Algorithm 2. With a feasible transmitter from problem (11), the i-th transmit power is $\|\mathbf{G}_i\|_2^2$, the sequential

[1]Although the constraints $\{\boldsymbol{\Delta}(n) \geq 0\}_{n=1}^{N_R}$ are necessary for the geometric illustration, they are redundant for the optimization. Therefore, the positive constraints on the shift terms can be removed from the optimization (27).

Algorithm 2: Joint IRS and MU-MIMO Transceiver Design.

1: **initialization:** initialize $\boldsymbol{\theta}_0 = \mathbf{1}$, $\{\mathbf{F}_{k,0}\}_{k=1}^K$ is initialized as (8) and set $i = 0$.
2: **repeat**
3: With $\boldsymbol{\theta} = \boldsymbol{\theta}_i$, $\{\mathbf{F}_k = \mathbf{F}_{k,i}\}_{k=1}^K$, solve the convex problem (11), denote the optimal solution as $\mathbf{G}_{i+1}$.
4: With $\boldsymbol{\theta} = \boldsymbol{\theta}_i$, $\mathbf{G} = \mathbf{G}_{i+1}$, solve the convex problem (12) for all $k \in [1, K]$, denote the optimal solutions as $\{\mathbf{F}_{k,i+1}\}_{k=1}^K$.
5: Let $\mathbf{G} = \mathbf{G}_{i+1}$, $\{\mathbf{F}_k = \mathbf{F}_{k,i+1}\}_{k=1}^K$. For the SDR approach, solve the problem (22). For the relax-and-retract approach, solve the convex QCQP problem (23). Denote the solution as $\boldsymbol{\theta}_{i+1}$, and set $i = i + 1$.
6: **until** $\|\mathbf{G}_{i-1}\|_F^2 - \|\mathbf{G}_i\|_F^2 \leq \epsilon_2$ or $i \geq i_{\max}$
7: For the relax-and-retract approach, use Algorithm 1.

optimal MSE minimization in (12) makes its objective value smaller than the MSE target ε_k, the total MSE minimization in (13) makes the resulting MSEs smaller than the MSE targets $\{\varepsilon_k\}_{k=1}^K$. The MSE minimizations in (12) and (13) make the successive transmitter design have more space to reduce the transmission power, i.e., $\|\mathbf{G}_{i+1}\|_2^2 \leq \|\mathbf{G}_i\|_2^2$. Since the transmit power is monotonically nonincreasing and the transmit power is bounded below from zero, the iterative algorithm in Algorithm 2 converges.

Note that the IRS retract operation in Algorithm 1 is only added at the end of Algorithm 2, instead of adding Algorithm 1 in every loop of Algorithm 2. This is due to the fact that the IRS retract operation increases the modulus and satisfies the MSE requirements, while does not change the transceiver and its transmit power. Since the IRS retract operation does not change the transmit power and the feasible set of relaxed subproblem (23) is larger than that of problem (21), the transmit power of the relax-and-retract approach is less than or equal to that of the SDR approach.

The computational complexities of the relax-and-retract approach with problems (23) and (27) and the SDR approach with problem (22) are analyzed as follows. By using the interior point method, the computational complexities of the convex QCQP problems (23) and (27) are $\mathcal{O}((K + N_R)^{\frac{1}{2}}(K + 2N_R)N_R^2)$ [19, p.145]. Therefore, the computational complexity of the relax-and-retract algorithm is $\mathcal{O}((1 + i_1/i_2)N_R^{3.5})$, where i_1 is the iteration number of Algorithm 1 and i_2 is the iteration number of relax operation in Algorithm 2. Since the computational complexity of the SDR problem (22) is $\mathcal{O}((N_R + 1)^{6.5})$ [20], the computational complexity of the relax-and-retract approach is much lower than that of the SDR approach.

V. Simulation Results and Discussions

In the simulations, the number of transmit antennas is $N = 4$, the number of users is $K = 2$, the number of antennas is $M_k = 2$ for all users, the number of data streams is $L_k = 1$ for all users, and the number of reflecting element is $N_R = 40$ unless noted otherwise. The MSE requirements are fixed as $\{\varepsilon_k = 0.1\}_{k=1}^K$, and the noise covariance matrices are $\{\boldsymbol{\Sigma}_k = 0.01\mathbf{I}_{M_k}\}_{k=1}^K$. The elements in channels $\mathbf{H}$, $\{\mathbf{H}_{r,k}, \mathbf{H}_{d,k}\}_{k=1}^K$ are generated from $\mathcal{CN}(0, 1)$ divided by the square root of their element numbers. Since the objective of this paper is to show the computational advantage of the proposed method, and the computational complexity is independent of the pathloss model, it is not considered in the simulations. All the simulation results are averaged over 100 random channel realizations. The Gaussian randomization number is $R = 10^3$. The termination threshold for Algorithm 1 is

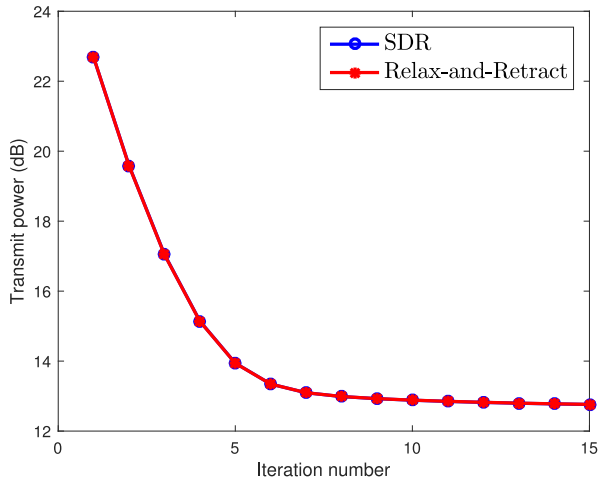

Fig. 3. Transmit powers of SDR and relax-and-retract methods.

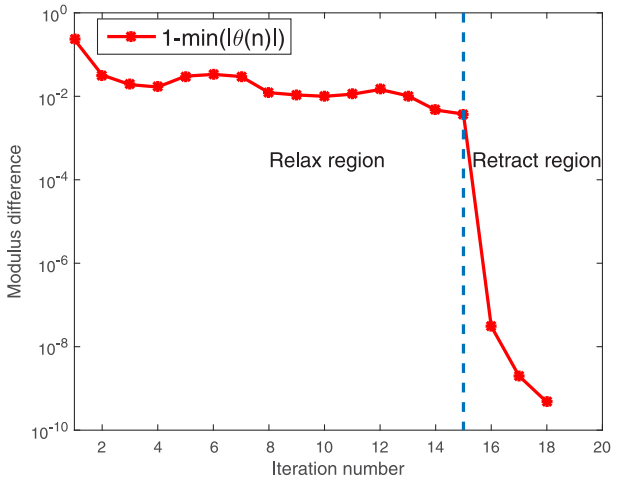

Fig. 4. Check the unit-modulus constraints of the relax-and-retract method.

$\epsilon_1 = 10^{-6}$, while for Algorithm 2 is $\epsilon_2 = 0$ and $i_{\max} = 15$, which is designed to make a clear distinction between the relax region and the retract region at the 15-th iteration. The optimization problems are solved on a laptop PC with 3.6 GHz CPU and 16 GB RAM, using the parser CVX and solver SDPT3 [21].

From the simulation setting, all users' channels $\{\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k}\}_{k=1}^{K}$ are 2×4 random matrices, therefore $\{\Pr(\mathbf{H}_{\mathrm{r,k}}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{\mathrm{d,k}} = \mathbf{0}) = 0\}_{\mathrm{k=1}}^{\mathrm{K}}$ and the probability of $\{\mathrm{rank}(\mathbf{H}_{r,k}\mathbf{\Theta}\mathbf{H} + \mathbf{H}_{d,k}) \geq 1\}_{k=1}^{K}$ is one. Furthermore, for the uncorrelated random 4×4 matrix $\mathbf{H}_0$, the probability of $\mathrm{rank}(\mathbf{H}_0) \geq 2$ is also almost one. Therefore, according to Proposition 1, all the simulations in this setting are feasible.

The transmit powers $(10\log_{10}\frac{\|\mathbf{G}\|_F^2}{\mathbf{\Sigma}_k(1,1)})$ of the relax-and-retract approach and the SDR approach are compared in Fig. 3. It can be seen from Fig. 3 that the transmit powers of the relax-and-retract approach are almost the same as that of the SDR approach, which validates the transmit power analysis in Section IV-C.

The unit-modulus constraint of the relax-and-retract approach is checked in Fig. 4. In the retract region of Fig. 4, $\min\{|\theta(n)|\}_{n=1}^{N_R} > 1 - 10^{-9}$. Together with the relaxed constraints $\{|\theta(n)| \leq 1\}_{n=1}^{N_R}$, we draw the conclusion that all IRS's modulus of the relax-and-retract approach are one, up to a numeric error 10^{-9}.

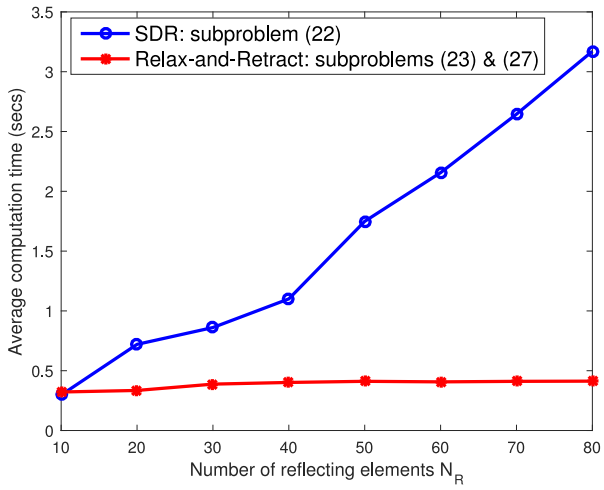

Fig. 5. Computational times of SDR and relax-and-retract methods.

The computational times of the relax-and-retract approach and the SDR approach are compared in Fig. 5. Since the retract operation in Algorithm 1 is implemented at the end of Algorithm 2, the computational time of the relax-and-retract method is the sum of the computational time of the relax operation in subproblem (23) and that of the retract operation in Algorithm 1 divided by the iteration number in the relax region. It can be seen that the computational time of the relax-and-retract approach is much smaller than that of the SDR approach, which validates the complexity analysis in Section IV-C. Therefore, Figs. 3 to 5 reveal that, compared to the conventional SDR approach, the proposed relax-and-retract approach achieves the same transmit power and the same unit-modulus implementation but with a much lower computational complexity.

VI. Conclusion

In this study, we have proposed a novel relax-and-retract approach to tackle the difficult nonconvex unit-modulus constraint in IRS design. The proposed relax-and-retract approach enables us to get convex QCQP subproblems, which has a much lower computational complexity than the SDR approach. Simulation results have shown that the proposed relax-and-retract approach has excellent performance in terms of computational time, while the transmit power and the unit-modulus hardware implementation are the same as those of the conventional SDR approach.

References

[1] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

[2] E. Björnson, Ö. Özdogan, E. G. Larsson, "Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 244–248, Feb. 2020.

[3] 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, Aug. 2019.

[4] H. Zhang, B. Di, L. Song, and Z. Han, "Reconfigurable intelligent surfaces assisted communications with limited phase shifts: How many phase shifts are enough?," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4498–4502, Apr. 2020.

[5] B. Di, H. Zhang, L. Li, L. Song, Y. Li, and Z. Han, "Practical hybrid beamforming with finite-resolution phase shifters for reconfigurable intelligent surface based multi-user communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4565–4570, Apr. 2020.

[6] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 518–522, Apr. 2020.
[7] C. Pan *et al.*, "Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1719–1734, Aug. 2020.
[8] 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, Oct. 2019.
[9] Z. Chu, W. Hao, P. Xiao, and J. Shi, "Intelligent reflecting surface aided multi-antenna secure transmission," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 108–112, Jan. 2020.
[10] X. Guan, Q. Wu, and R. Zhang, "Intelligent reflecting surface assisted secrecy communication: Is artificial noise helpful or not?," *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 778–782, Jun. 2020.
[11] C. Pan *et al.*, "Multicell mimo communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.
[12] M. M. Zhao, Q. Wu, M. J. Zhao, and R. Zhang, "Exploiting amplitude control in intelligent reflecting surface aided wireless communication with imperfect CSI," 2020, *arXiv:2005.07002*.
[13] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1179–1197, May 2004.
[14] 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, Jan. 2020.
[15] N. Vucic and H. Boche, "A tractable method for chance-constrained power control in downlink multiuser MISO systems with channel uncertainty," *IEEE Signal Process. Lett.*, vol. 16, no. 5, pp. 346–349, May 2009.
[16] K. Wang, A. M. So, T. Chang, W. Ma, and C. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.
[17] X. He and Y.-C. Wu, "Set squeezing procedure for quadratically perturbed chance-constrained programming," *IEEE Trans. Signal Process.*, vol. 69, pp. 682–694, 2021.
[18] M. C. So, J. Zhang, and Y. Ye, "On approximating complex quadratic optimization problems via semidefinite programming relaxations," *Math. Program.*, vol. 110, no. 1, pp. 93–110, 2007.
[19] A. Nemirovski, "Interior point polynomial time methods in convex programming," Lecture Notes, 1996.
[20] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1426–1438, Aug. 2006.
[21] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, version 2.1," Mar. 2014. [Online]. Avilable: http://cvxr.com/cvx