Beamforming Optimization for IRS-Aided Communications with Transceiver Hardware Impairments

Hong Shen, *Member, IEEE,* Wei Xu, *Senior Member, IEEE,* Shulei Gong,
Chunming Zhao, *Member, IEEE,* and Derrick Wing Kwan Ng, *Senior Member, IEEE*

Abstract—In this paper, we focus on intelligent reflecting surface (IRS) assisted multi-antenna communications with transceiver hardware impairments encountered in practice. In particular, we aim to maximize the received signal-to-noise ratio (SNR) taking into account the impact of hardware impairments, where the source transmit beamforming and the IRS reflect beamforming are jointly designed under the proposed optimization framework. To circumvent the non-convexity of the formulated design problem, we first derive a closed-form optimal solution to the source transmit beamforming. Then, for the optimization of IRS reflect beamforming, we obtain an upper bound to the optimal objective value via solving a single convex problem. A low-complexity minorization-maximization (MM) algorithm was developed to approach the upper bound. Simulation results demonstrate that the proposed beamforming design is more robust to the hardware impairments than that of the conventional SNR maximized scheme. Moreover, compared to the scenario without deploying an IRS, the performance gain brought by incorporating the hardware impairments is more evident for the IRS-aided communications.

Index Terms—Intelligent reflecting surface (IRS), hardware impairment, beamforming optimization, minorization-maximization (MM).

I. INTRODUCTION

With the rapid development of electromagnetic materials, intelligent reflecting surface (IRS) has become a cost-effective promising solution for future wireless communications [1]–[7]. By deploying a number of passive reflecting elements with controllable phase shifts, IRS has the capability of improving communication quality via enhancing received signal power and harnessing interference power at desired receivers, which paves the way to realizing smart radio environments

[4], [7]. Compared to the well-known massive multiple-input multiple-output (MIMO) technique [8], [9], IRS can provide excellent system performance without requiring power-hungry components such as power amplifiers[1]. As a result, resource allocation design for IRS-assisted systems has drawn significant research interest recently. For example, the transmission optimization for various IRS-assisted communication systems has been studied in a number of recent works [11]–[24]. Specifically, the received signal-to-noise ratio (SNR) maximization [11], [12] and quality-of-service (QoS) constrained transmit power minimization [13], [14] problems were investigated for an IRS-aided single-user multiple-input single-output (MISO) system. Regarding a multi-user MISO system assisted by an IRS, different design objectives, e.g., transmit power minimization [13], [14], achievable rate maximization [15], [16], and energy efficiency maximization [17], [18], have been adopted for optimizing the beamforming scheme. In addition to single-user and multi-user systems, beamforming optimization for IRS-assisted multi-antenna wiretap channels was investigated in [19]–[23], where the design goal is to maximize the system achievable secrecy rate. Alternatively, focusing on an IRS-aided multi-antenna simultaneous wireless information and power transfer (SWIPT) system, the authors of [24] performed beamforming optimization to maximize the weighted sum power received by energy harvesting receivers while fulfilling the QoS constraints for information decoding receivers. As shown in all the above works, the presence of an IRS complicates the optimal beamforming design due to the fact that the phase shifts of an IRS are not only coupled with the transmit beamforming, but they should also satisfy unit-modulus constraints or belong to a discrete set [12], [14], [16], [18].

For practical communication systems, there exist inevitable impairments in transceiver hardware such as amplifiers, oscillators, digital-to-analog converters (DACs), and analog-to-digital converters (ADCs) [25], [26]. Note that different from the traditional noise at the receiver side, hardware impairments lead to distortions on both transmitted and received signals. More importantly, the power of distortions scales with the power of the useful signal at the wireless transceivers. Hence, several efforts have been made to analyze the impact of hardware impairments on various MIMO systems [27]–[29].

---

[1]IRS is also different from the large intelligent surface (LIS) recently proposed in [10] due to the fact that IRS is a passive device while LIS is an active one.

As a remedy, the authors of [30]–[33] conducted transceiver optimization for MIMO systems such as multicell MIMO and MIMO relaying, in order to relieve the performance deterioration caused by the hardware impairments. Concerning the IRS-aided multi-antenna communications, the existence of transceiver hardware impairments, if not carefully handled, will affect both the source transmit beamforming and the IRS reflect beamforming designs, thus leading to severer performance degradation than that in conventional systems without exploiting an IRS. Hence, it is necessary to take into account the hardware impairments when performing beamforming optimization for IRS-aided multi-antenna systems, which, to the best of our knowledge, has not been studied in literature yet.

In this paper, motivated by the aforementioned observations, we study the beamforming optimization for an IRS-aided multi-antenna communication system by incorporating signal distortions caused by hardware impairments. Both source transmit beamforming and IRS reflect beamforming are optimized to maximize the received SNR at the destination. Compared with a relevant work [11] under the assumption of perfect hardware, the considered problem in this paper is more sophisticated due to the presence of hardware impairments and cannot be directly solved with the methods developed in [11]. On the other hand, existing robust designs, e.g., [30]–[33], for conventional multi-antenna systems are not applicable to IRS-assisted systems since the IRS phase shifts also need to be optimized, which is non-trivial even for the perfect hardware case. Our main contributions are summarized as follows:

- We determine the optimal source transmit beamforming, which has a closed form and is in fact a weighted version of matched filter (MF) beamforming.
- We acquire an upper bound to the optimal objective value of the IRS reflect beamforming optimization, which can be efficiently calculated by solving one convex problem.
- We further employ a low-complexity minorization-maximization (MM) technique to handle the IRS reflect beamforming optimization. The method only requires calculating a closed-form expression iteratively while almost achieving the upper bound.

Notations: Vectors and matrices are represented by boldface lower-case and boldface upper-case letters, respectively. $|\cdot|$ and $\|\cdot\|$ stand for the absolute value of the input scalar and the Euclidean norm of the input vector, respectively. $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote the conjugate, the transpose, and the Hermitian transpose operations, respectively. $\arg(\cdot)$ and $\Re(\cdot)$ return the phase and the real part of the input complex number, respectively. $\text{diag}\{\cdot\}$ represents a diagonal matrix with diagonal elements being the entries of the input vector and $\widetilde{\text{diag}}\{\cdot\}$ denotes a diagonal matrix which has the same diagonal elements as the input matrix. $(\cdot)^{-1}$, $\text{tr}(\cdot)$, and $\lambda_{\max}(\cdot)$ denote the inverse, the trace, and the maximum eigenvalue of the input matrix, respectively. $(\cdot)_{i,j}$ denotes the (i, j)-th entry of the input matrix. $\mathbf{a}(1 : N)$ denotes the vector whose entries are the first N entries of vector $\mathbf{a}$. $a \sim \mathcal{CN}(\mu, \sigma^2)$ means that scalar a is a circularly symmetric complex Gaussian random variable with mean μ and variance σ^2. $\mathbf{a} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ means
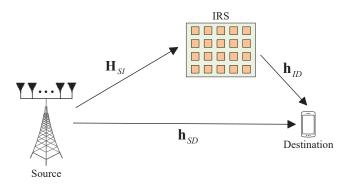

Fig. 1. A wireless communication system assisted by an IRS.

that vector $\mathbf{a}$ follows circularly symmetric complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. $\mathbf{A} \succeq \mathbf{0}$ denotes that matrix $\mathbf{A}$ is positive semidefinite. $\mathbb{E}\{\cdot\}$ is the expectation operation. $\mathbb{C}^{m\times n}$ stands for the set of $m \times n$ complex matrices. $\mathbf{e}_m$ is the unit vector with the m-th element being 1. $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. $O(\cdot)$ represents the big-O notation.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

We consider an IRS-aided communication system consisting of one source, one IRS, and one destination, as depicted in Fig. 1. The source is equipped with N_S antennas, the IRS has N_I passive reflecting elements, and the destination is a single-antenna device.

Different from prior works on IRS-aided communications, e.g., [11]–[24], we incorporate practical transceiver hardware impairments in the considered system. Specifically, the source transmit signal is expressed by

$$\mathbf{t} = \mathbf{w}x + \mathbf{z}_S, \tag{1}$$

where $\mathbf{w} \in \mathbb{C}^{N_S \times 1}$ is the source transmit beamforming, x is the transmit symbol of the source with unit power, and $\mathbf{z}_S$ represents the transmit distortion noise independent of x, which models the joint effects of the power amplifier noise, the non-linearities in DAC and power amplifier, and the oscillator phase noise [25]. In particular, the entries of $\mathbf{z}_S$ are independent zero-mean Gaussian random variables and the variance of the i-th entry is proportional to the power of the intended transmit signal at the i-th antenna, i.e., $\mathbf{z}_S \sim \mathcal{CN}(\mathbf{0}, \kappa_S \widetilde{\text{diag}}\{\mathbf{w}\mathbf{w}^H\})$, where $\kappa_S \in (0, 1)$ denotes the normalized variance of the transmit distortion noise.

The received signal at the destination is expressed by

$$y = \underbrace{\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{H}_{SI} \mathbf{t}}_{\text{reflected link}} + \underbrace{\mathbf{h}_{SD}^H \mathbf{t}}_{\text{direct link}} + n + z_D$$

$$\triangleq \tilde{y} + z_D, \tag{2}$$

where $\mathbf{h}_{ID} \in \mathbb{C}^{N_I \times 1}$ is the IRS-to-destination channel, $\boldsymbol{\Theta} \triangleq \text{diag}\{[e^{j\phi_1}, \cdots, e^{j\phi_i}, \cdots, e^{j\phi_{N_I}}]\}$ is the reflect beamforming matrix of the IRS with ϕ_i denoting the phase shift incurred by the i-th reflecting element, $\mathbf{H}_{SI} \in \mathbb{C}^{N_I \times N_S}$ is

$$\text{SNR}(\mathbf{w}, \boldsymbol{\Theta}) = \frac{|(\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\mathbf{w}|^2}{(\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H)(\kappa_D \mathbf{w}\mathbf{w}^H + (1+\kappa_D)\kappa_S \widetilde{\text{diag}\{\mathbf{w}\mathbf{w}^H\}})(\mathbf{H}_{SI}^H \boldsymbol{\Theta}^H \mathbf{h}_{ID} + \mathbf{h}_{SD}) + (1+\kappa_D)\sigma_n^2} \tag{3}$$

$$\begin{aligned} \underset{\mathbf{w}}{\text{maximize}} \quad & \frac{|(\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\mathbf{w}|^2}{(\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H)(\kappa_D \mathbf{w}\mathbf{w}^H + (1+\kappa_D)\kappa_S \widetilde{\text{diag}\{\mathbf{w}\mathbf{w}^H\}})(\mathbf{H}_{SI}^H \boldsymbol{\Theta}^H \mathbf{h}_{ID} + \mathbf{h}_{SD}) + (1+\kappa_D)\sigma_n^2} \\ \text{subject to} \quad & \|\mathbf{w}\|^2 \leq \tilde{P} \end{aligned} \tag{5}$$

$$\begin{aligned} \underset{\mathbf{w}}{\text{maximize}} \quad & \frac{|(\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\mathbf{w}|^2}{\mathbf{w}^H(\kappa_D \boldsymbol{\Upsilon}(\boldsymbol{\Theta}) + (1+\kappa_D)\kappa_S \widetilde{\text{diag}\{\boldsymbol{\Upsilon}(\boldsymbol{\Theta})\}})\mathbf{w} + (1+\kappa_D)\sigma_n^2} \\ \text{subject to} \quad & \|\mathbf{w}\|^2 \leq \tilde{P} \end{aligned} \tag{6}$$

the source-to-IRS channel, $\mathbf{h}_{SD} \in \mathbb{C}^{N_S \times 1}$ is the source-to-destination channel, and $n \sim \mathcal{CN}(0, \sigma_n^2)$ is the additive white Gaussian noise (AWGN) at the destination. Variable z_D stands for the receive distortion noise independent of $\tilde{y}$, which models the joint effects of the automatic gain control (AGC) noise, the non-linearities in ADC and AGC, and the oscillator phase noise [26]. Note that z_D follows zero-mean Gaussian distribution and its variance is proportional to the power of the undistorted received signal, i.e., $z_D \sim \mathcal{CN}(0, \kappa_D \mathbb{E}\{|\tilde{y}|^2\})$, where $\kappa_D \in (0,1)$ denotes the normalized variance of the receive distortion noise.

Based on (1) and (2), we obtain the received SNR at the destination which is given by (3). Since the source transmit beamformer $\mathbf{w}$ appears in both the numerator and the denominator of the above SNR expression, it is expected that the received SNR of the considered system in the presence of transceiver hardware impairments will become saturated when the transmit power is exceedingly large.

We note that the reflecting element of the IRS is subject to the phase errors [34], [35] caused by quantized phase shifts in practice. This is because the IRS usually has a large number of reflecting elements and the implementation cost per reflecting element should be as low as possible to reduce the overall cost. In fact, the reflecting element of the IRS typically has a small number of discrete phase shifts, e.g., 2 or 4 [36]–[38], which results in non-negligible phase errors. To capture the impacts of IRS phase errors in the current framework, we can simply replace the ideal continuous phase shift constraints by discrete constraints as in [12], [16], [18], where the discrete phase shift set is given by $\mathcal{F}_d \triangleq \left\{ \phi_i \mid \phi_i \in \left\{ 0, \frac{2\pi}{L}, \cdots, \frac{2\pi(L-1)}{L} \right\} \right\}$ with $L = 2^B$ being the number of phase shift levels and B denoting the number of bits used to represent the L phase shift levels.

B. Problem Formulation

In this paper, we aim to maximize the received SNR in (3) by jointly optimizing the source transmit beamformer $\mathbf{w}$ and the IRS reflect beamforming matrix $\boldsymbol{\Theta}$, where we also impose a power constraint on $\mathbf{w}$ and unit-modulus constraints

on the diagonal elements of $\boldsymbol{\Theta}$. More specifically, according to (1) and the distribution of $\mathbf{z}_S$, we can readily derive the total transmit power of the source node as $(1 + \kappa_S)\|\mathbf{w}\|^2$. Moreover, both continuous and discrete phase shift constraints will be considered in this work, where the result of the continuous phase shift case serves as a performance benchmark. Therefore, we formulate the design problem as

$$\begin{aligned} \underset{\mathbf{w}, \boldsymbol{\Theta}}{\text{maximize}} \quad & \text{SNR}(\mathbf{w}, \boldsymbol{\Theta}) \\ \text{subject to} \quad & (1+\kappa_S)\|\mathbf{w}\|^2 \leq P, \\ & \theta_i \in \mathcal{F}_1 \text{ or } \mathcal{F}_2, \ i = 1, \cdots, N_I, \end{aligned} \tag{4}$$

where P denotes the maximal source transmit power and θ_i is the i-th diagonal element of $\boldsymbol{\Theta}$. $\mathcal{F}_1$ and $\mathcal{F}_2$ are defined by $\mathcal{F}_1 \triangleq \{\theta_i \mid |\theta_i| = 1\}$ and $\mathcal{F}_2 \triangleq \left\{ \theta_i \mid \theta_i = e^{j\phi_i}, \phi_i \in \left\{ 0, \frac{2\pi}{L}, \cdots, \frac{2\pi(L-1)}{L} \right\} \right\}$, respectively.

The above problem has a complicated objective function and nonconvex unit-modulus constraints, whose optimal solution cannot be readily determined. In particular, compared to the SNR maximization problem in [11], the presence of hardware impairments makes the beamforming optimization more challenging and thus existing methods designed for the case of perfect hardware are not applicable to the problem at hand.

III. CLOSED-FORM OPTIMAL SOLUTION TO SOURCE TRANSMIT BEAMFORMING

In this section, we determine a closed-form optimal solution to $\mathbf{w}$ for an arbitrary $\boldsymbol{\Theta}$. Based on problem (4), the optimization of $\mathbf{w}$ with a fixed $\boldsymbol{\Theta}$ is expressed by (5), where $\tilde{P} \triangleq \frac{P}{1+\kappa_S}$. To find the optimal solution to this problem in a closed form, we need to first transform the denominator of the objective function into a tractable form. In fact, by invoking the equality $\mathbf{a}^H \widetilde{\text{diag}\{\mathbf{b}\mathbf{b}^H\}}\mathbf{a} = \mathbf{b}^H \widetilde{\text{diag}\{\mathbf{a}\mathbf{a}^H\}}\mathbf{b}$, we transform problem (5) into an equivalent form in (6), where $\boldsymbol{\Upsilon}(\boldsymbol{\Theta}) \triangleq (\mathbf{H}_{SI}^H \boldsymbol{\Theta}^H \mathbf{h}_{ID} + \mathbf{h}_{SD})(\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H)$. By dividing the numerator and the denominator of the above objective function by $\|\mathbf{w}\|^2$, we can readily confirm that

$$\mathbf{w}^{\star} = \sqrt{\tilde{P}} \frac{\left(\kappa_D \mathbf{\Upsilon}(\mathbf{\Theta}) + (1+\kappa_D)\kappa_S \widetilde{\text{diag}}\{\mathbf{\Upsilon}(\mathbf{\Theta})\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S}\right)^{-1}(\mathbf{H}_{SI}^H \mathbf{\Theta}^H \mathbf{h}_{ID} + \mathbf{h}_{SD})}{\left\|\left(\kappa_D \mathbf{\Upsilon}(\mathbf{\Theta}) + (1+\kappa_D)\kappa_S \widetilde{\text{diag}}\{\mathbf{\Upsilon}(\mathbf{\Theta})\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S}\right)^{-1}(\mathbf{H}_{SI}^H \mathbf{\Theta}^H \mathbf{h}_{ID} + \mathbf{h}_{SD})\right\|} \tag{7}$$

$$
\begin{aligned}
\underset{\mathbf{\Theta}}{\text{maximize}} \quad & \tilde{P}(\mathbf{h}_{ID}^H \mathbf{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\left(\kappa_D \mathbf{\Upsilon}(\mathbf{\Theta}) + (1+\kappa_D)\kappa_S \widetilde{\text{diag}}\{\mathbf{\Upsilon}(\mathbf{\Theta})\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S}\right)^{-1}(\mathbf{H}_{SI}^H \mathbf{\Theta}^H \mathbf{h}_{ID} + \mathbf{h}_{SD}) \\
\text{subject to} \quad & |\theta_i| = 1, \ i = 1, \cdots, N_I
\end{aligned} \tag{9}
$$

it is monotonically increasing with $\|\mathbf{w}\|$. Hence, the power consumption constraint must be active at optimality. Moreover, by utilizing the generalized Rayleigh quotient [39], we obtain the optimal solution to $\mathbf{w}$ in (7).

Note that a matrix inversion is required by (7), which can be computationally inefficient particularly when N_S is large. To deal with this issue, we show that $\mathbf{w}^{\star}$ has an equivalent form that only involves the inversion of a diagonal matrix in the following proposition.

Proposition 1: The optimal solution to problem (6) can be expressed by

$$\mathbf{w}^{\star} = \sqrt{\tilde{P}} \frac{\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H \mathbf{\Theta}^H \mathbf{h}_{ID} + \mathbf{h}_{SD})}{\left\|\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H \mathbf{\Theta}^H \mathbf{h}_{ID} + \mathbf{h}_{SD})\right\|}, \tag{8}$$

where $\tilde{\mathbf{\Upsilon}}(\mathbf{\Theta}) \triangleq (1+\kappa_D)\kappa_S \widetilde{\text{diag}}\{\mathbf{\Upsilon}(\mathbf{\Theta})\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S}$.

Proof: See Appendix A. ∎

Compared to (7), we only need to calculate the inversion of a diagonal matrix, thus requiring much lower computational complexity. Moreover, it can be observed from (8) that $\mathbf{w}^{\star}$ is actually a product of a diagonal weighted matrix and the MF beamforming, where the factor of hardware impairments is included in the weighted matrix. When there exist severe hardware impairments, i.e., κ_S or κ_D is large, the weights become smaller such that the performance degradation caused by hardware impairments is alleviated. On the other hand, for the ideal scenario without transceiver hardware impairments, i.e., $\kappa_S = \kappa_D = 0$, it is clear that (8) reduces to the conventional MF beamforming.

IV. OPTIMIZATION OF IRS REFLECT BEAMFORMING

Now, let us first consider the IRS reflect beamforming optimization under the continuous phase shift constraints, i.e., $|\theta_i| = 1$, $i = 1, \cdots, N_I$. Recall that the optimal transmit beamforming was obtained in a closed form given any reflect beamforming in the previous section. By plugging the optimal $\mathbf{w}^{\star}$ in (7) into the original problem in (4), we transform problem (4) into (9) without loss of optimality. Denote the above objective function by $\psi(\mathbf{\Theta})$. Then, by left multiplying $\mathbf{h}_{ID}^H \mathbf{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H$ on both sides of (25) in Appendix A, we have

$$\psi(\mathbf{\Theta}) = \frac{\tilde{P}\tilde{\psi}(\mathbf{\Theta})}{\kappa_D \tilde{\psi}(\mathbf{\Theta}) + 1}, \tag{10}$$

where $\tilde{\psi}(\mathbf{\Theta}) \triangleq (\mathbf{h}_{ID}^H \mathbf{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H \mathbf{\Theta}^H \mathbf{h}_{ID} + \mathbf{h}_{SD})$. Hence, we are able to convert problem (9) to the simplified form as follows:

$$
\begin{aligned}
\underset{\mathbf{\Theta}}{\text{maximize}} \quad & (\mathbf{h}_{ID}^H \mathbf{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H \mathbf{\Theta}^H \mathbf{h}_{ID} + \mathbf{h}_{SD}) \\
\text{subject to} \quad & |\theta_i| = 1, \ i = 1, \cdots, N_I.
\end{aligned} \tag{11}
$$

Different from problem (5) with respect to $\mathbf{w}$, it is quite difficult to optimally solve this problem using the generalized Rayleigh quotient due to the more complicated objective function and nonconvex constraints. In the following, we first obtain an upper bound to the optimal objective value of problem (11). Then, we propose an MM-based algorithm to approach this upper bound with low computational complexity. Finally, based on the obtained continuous solution, we provide a method to achieve an efficient solution for the discrete phase shift case.

A. Upper Bound to Optimal Objective Value

Although it is challenging to achieve the globally optimal solution to problem (11) in polynomial time due to its nonconvexity, we are able to acquire an upper bound to the optimal objective value via solving a convex problem as given in the subsequent theorem.

Theorem 1: The optimal objective value of problem (11) is upper bounded by that of the following convex problem:

$$
\begin{aligned}
\underset{\tilde{\mathbf{\Theta}} \succeq \mathbf{0}}{\text{maximize}} \quad & \sum_{m=1}^{N_S} \frac{(\mathbf{\Psi}\tilde{\mathbf{\Theta}}\mathbf{\Psi}^H)_{m,m}}{(1+\kappa_D)\kappa_S(\mathbf{\Psi}\tilde{\mathbf{\Theta}}\mathbf{\Psi}^H)_{m,m} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}} \\
\text{subject to} \quad & (\tilde{\mathbf{\Theta}})_{i,i} = 1, \ i = 1, \cdots, N_I + 1,
\end{aligned} \tag{12}
$$

where $\mathbf{\Psi} \triangleq [\mathbf{H}_{SI}^H \text{diag}\{\mathbf{h}_{ID}\} \ \mathbf{h}_{SD}]$ and $\tilde{\mathbf{\Theta}}$ is a positive semidefinite matrix of size $N_I + 1$.

Proof: See Appendix B. ∎

According to the above conclusion, we can determine an upper bound to the optimal objective value of problem (11) by solving problem (12) with the off-the-shelf convex optimization tools such as CVX [40]. This upper bound serves as a benchmark to evaluate the performance of our proposed algorithm.

$$\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H \left((1+\kappa_D)\kappa_S \widetilde{\mathrm{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S} \right)^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}$$

$$\geq 2\Re\{\tilde{\boldsymbol{\theta}}_0^H (\boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1}\boldsymbol{\Psi} - (1+\kappa_D)\kappa_S(\boldsymbol{\Omega} - \lambda_{\max}(\boldsymbol{\Omega})\mathbf{I}_{N_I+1}))\tilde{\boldsymbol{\theta}}\} - 2(1+\kappa_D)\kappa_S(N_I+1)\lambda_{\max}(\boldsymbol{\Omega})$$

$$+ \tilde{\boldsymbol{\theta}}_0^H (2(1+\kappa_D)\kappa_S\boldsymbol{\Omega} - \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1}\boldsymbol{\Psi})\tilde{\boldsymbol{\theta}}_0 \qquad (13)$$

$$\underset{\tilde{\boldsymbol{\theta}}}{\text{maximize}} \quad \Re\left\{ \left(\tilde{\boldsymbol{\theta}}^{(k-1)}\right)^H (\boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1}\boldsymbol{\Psi} - (1+\kappa_D)\kappa_S(\boldsymbol{\Omega} - \lambda_{\max}(\boldsymbol{\Omega})\mathbf{I}_{N_I+1}))\tilde{\boldsymbol{\theta}} \right\}$$

$$\text{subject to} \quad |\tilde{\theta}_i| = 1, \ i = 1, \cdots, N_I + 1 \qquad (14)$$

Algorithm 1 MM Algorithm for Problem (11)

1: *Initialization:* Initialize $\tilde{\boldsymbol{\theta}}^{(0)}$, iteration index $k = 0$, and convergence accuracy $\epsilon \to 0$.
2: **repeat**
3:   $k = k + 1$.
4:   Calculate $\tilde{\boldsymbol{\theta}}^\star$ using (15).
5:   $\tilde{\boldsymbol{\theta}}^{(k)} = \tilde{\boldsymbol{\theta}}^\star$.
6: **until** convergence.
7: Calculate $\boldsymbol{\Theta}^\star = \mathrm{diag}\{(\tilde{\boldsymbol{\theta}}^{(k)}(1:N_I)/\tilde{\theta}_{N_I+1}^{(k)})^*\}$.

B. MM Algorithm for Solving Problem (11)

We consider solving problem (11), or, equivalently, problem (28) in Appendix B, via the popular MM algorithm [41], which achieves almost the same objective value as the upper bound obtained in Section IV-A with low computational cost.

According to the philosophy of the MM technique, we first construct a surrogate function that serves as a lower bound of the objective function of problem (28), which is shown in the following proposition.

Proposition 2: For problem (28) in Appendix B, a lower bound of the objective function used for the MM algorithm is given by (13), where $\boldsymbol{\Xi}_0 \triangleq (1+\kappa_D)\kappa_S\widetilde{\mathrm{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H\boldsymbol{\Psi}^H\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S}$ with $\tilde{\boldsymbol{\theta}}_0$ fulfilling the unit-modulus constraints and $\boldsymbol{\Omega} \triangleq \boldsymbol{\Psi}^H\widetilde{\mathrm{diag}}\{\boldsymbol{\Xi}_0^{-1}\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H\boldsymbol{\Psi}^H\boldsymbol{\Xi}_0^{-1}\}\boldsymbol{\Psi}$.

Proof: See Appendix C. ∎

Then, the MM algorithm is adopted to iteratively tighten the lower bound of the objective function. Specifically, in the k-th iteration of the MM algorithm, we set $\tilde{\boldsymbol{\theta}}_0 = \tilde{\boldsymbol{\theta}}^{(k-1)}$ with $\tilde{\boldsymbol{\theta}}^{(k-1)}$ denoting the optimal solution to $\tilde{\boldsymbol{\theta}}$ obtained in the $(k-1)$-th iteration, and consider maximizing the lower bound in (13). Accordingly, the corresponding problem to be solved in the k-th iteration is given by (14), where we remove the constant terms in (13). Define $\boldsymbol{\alpha} \triangleq (\boldsymbol{\Psi}^H\boldsymbol{\Xi}_0^{-1}\boldsymbol{\Psi} - (1+\kappa_D)\kappa_S(\boldsymbol{\Omega}-\lambda_{\max}(\boldsymbol{\Omega})\mathbf{I}_{N_I+1}))\tilde{\boldsymbol{\theta}}^{(k-1)}$. Then, we can readily obtain the optimal solution to problem (14) by

$$\tilde{\theta}_i^\star = e^{j\arg(\alpha_i)}, \ i = 1, \cdots, N_I + 1, \qquad (15)$$

where α_i is the i-th entry of $\boldsymbol{\alpha}$.

The proposed MM algorithm for solving problem (11) is summarized in Algorithm 1, where the last step follows from the definition of $\tilde{\boldsymbol{\theta}}$ in Appendix B. Regarding the convergence property of Algorithm 1, we can prove the following conclusion.

Theorem 2: Algorithm 1 converges to a stationary point of problem (11) when convergence is reached.

Proof: See Appendix D. ∎

After obtaining $\boldsymbol{\Theta}^\star$, i.e., the solution to IRS reflect beamforming, via Algorithm 1, we substitute it into (8) to calculate $\mathbf{w}^\star$, i.e., the solution to source transmit beamforming. We would like to emphasize that the proposed solution framework does not optimize the source transmit beamforming and the IRS reflect beamforming in an alternating manner as the one commonly applied in the literature, e.g., [19]–[23]. Moreover, based on Theorem 2 and the equivalence between problem (11) and the original beamforming optimization problem in (4), we arrive at the conclusion that $(\mathbf{w}^\star, \boldsymbol{\Theta}^\star)$ is a stationary point of problem (4).

Based on the continuous solution to the IRS phase shifts obtained by Algorithm 1, we provide a simple but effective method to acquire discrete phase shifts. Specifically, we project the IRS phase shifts $\phi_i^\star$, $i = 1, \cdots, N_I$, calculated by Algorithm 1 onto the discrete set $\mathcal{F}_d$ defined in Section II-A, i.e., $\hat{\phi}_i = \arg\min_{\phi\in\mathcal{F}_d} |\phi_i^\star - \phi|$, $i = 1, \cdots, N_I$, where $\hat{\phi}_i$ is the quantized phase shift of the i-th reflecting element. As will be shown in Section V, we can achieve a performance approaching that of the continuous phase shift case by setting only $B = 2$.

Remark 1: We find empirically that the convergence of Algorithm 1 can be further accelerated by applying the acceleration scheme based on SQUAREM [42, Section V-B]. In fact, from the comparison results in Table III of Section V, using the acceleration method significantly reduces the required average number of iterations.

Remark 2: For the proposed design, the optimized solutions to the source transmit beamforming and the IRS reflect beamforming can be calculated at the source with acceptable computational complexity (see Section V-C for details). Also, the calculated IRS discrete phase shifts (represented by bits) are conveyed to the IRS via a dedicated wireless control link as in prior works [13], [14]. On the other hand, the beamforming

$$\underset{\boldsymbol{\Theta}}{\text{maximize}} \quad \mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{H}_{SI} \left((1+\kappa_D)\kappa_S \widetilde{\text{diag}} \{\mathbf{H}_{SI}^H \boldsymbol{\Theta}^H \mathbf{h}_{ID} \mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{H}_{SI}\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}} \mathbf{I}_{N_S} \right)^{-1} \mathbf{H}_{SI}^H \boldsymbol{\Theta}^H \mathbf{h}_{ID}$$

$$\text{subject to} \quad |\theta_i| = 1, \; i = 1, \cdots, N_I \tag{16}$$

$$|\eta|^2 |\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{a}_I|^2 \mathbf{a}_S^H \left((1+\kappa_D)\kappa_S \widetilde{\text{diag}} \{|\eta|^2 |\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{a}_I|^2 \mathbf{a}_S \mathbf{a}_S^H\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}} \mathbf{I}_{N_S} \right)^{-1} \mathbf{a}_S$$

$$= \frac{N_S |\eta|^2 |\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{a}_I|^2}{(1+\kappa_D)\kappa_S |\eta|^2 |\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{a}_I|^2 + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}} \tag{17}$$

calculation at the source requires the normalized distortion noise variances κ_S and κ_D. The two constants can be measured at the source and the destination in an offline manner, respectively, and κ_D can be fed back to the source for beamforming calculation. In addition, channel state information (CSI) is also needed for beamforming calculation. Fortunately, we find that it suffices to acquire $\boldsymbol{\Psi} \triangleq [\mathbf{H}_{SI}^H \text{diag}\{\mathbf{h}_{ID}\} \; \mathbf{h}_{SD}]$ for the calculation of the proposed beamforming solution. In fact, by properly setting the IRS phase shifts during the channel training, the source can directly estimate $\boldsymbol{\Psi}$ with some conventional channel estimation methods such as least-square (LS) channel estimation [43], [44].

Remark 3: The proposed beamforming design is also applicable to a time division multiple access (TDMA)-based multi-user system. For this scenario, we only need to apply the proposed method to each user in each independent time slot. Nevertheless, we recognize that the extension to other multi-user systems, e.g., spatial division multiple access (SDMA)-based multi-user system, is non-trivial. In fact, the beamforming optimization method developed in our work may no longer be optimal due to the more complicated objective function. The study of the beamforming optimization for the SDMA-based multi-user system with hardware impairments is left for future works.

C. Special Case Without Direct Link

In practice, the direct link between the source and the destination is negligible if there exists an exceedingly large path loss between them. By neglecting the direct link, i.e., $\mathbf{h}_{SD}^H = \mathbf{0}$, the IRS reflect beamforming optimization problem in (11) becomes (16). In order to obtain a simple and insightful solution to the above problem, we further assume that the source-to-IRS channel $\mathbf{H}_{SI}$ only has the line-of-sight (LOS) component, i.e., $\mathbf{H}_{SI} = \eta \mathbf{a}_I \mathbf{a}_S^H$, where η is the complex channel gain, and $\mathbf{a}_I$ and $\mathbf{a}_S$, with normalized entries, are the array response vectors of the IRS and the source, respectively. Then, the objective function of problem (16) is rewritten by (17) since $\|\mathbf{a}_S\|^2 = N_S$ and $\widetilde{\text{diag}}\{\mathbf{a}_S \mathbf{a}_S^H\} = \mathbf{I}_{N_S}$. Therefore, it follows that problem (16) can be simplified to

$$\underset{\boldsymbol{\Theta}}{\text{maximize}} \quad |\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{a}_I|$$

$$\text{subject to} \quad |\theta_i| = 1, \; i = 1, \cdots, N_I. \tag{18}$$

According to [11, Section IV], the optimal solution to the above problem is given by

$$\theta_i^\star = e^{-j(\arg(h_{ID,i}^*) + \arg(a_{I,i}))}, \; i = 1, \cdots, N_I, \tag{19}$$

where $h_{ID,i}$ is the i-th entry of $\mathbf{h}_{ID}$ and $a_{I,i}$ is the i-th entry of $\mathbf{a}_I$. Moreover, by substituting (19) into the optimal source transmit beamforming in (8), we can readily obtain

$$\mathbf{w}^\star = \sqrt{\frac{\tilde{P}}{N_S}} \mathbf{a}_S. \tag{20}$$

Therefore, when the direct link is negligible and the source-to-IRS channel only contains the LOS component, the optimal source transmit beamforming should be aligned with the source array response vector while the optimal IRS phase shifts should match the sum of the phase shifts of the IRS-to-destination channel and the IRS array response vector.

By substituting (19) and (20) into the expression of the received SNR in (3), we have

$$\text{SNR} = \frac{\tilde{P} N_S |\eta|^2 \|\mathbf{h}_{ID}\|_1^2}{\tilde{P}(\kappa_D N_S + (1+\kappa_D)\kappa_S)|\eta|^2 \|\mathbf{h}_{ID}\|_1^2 + (1+\kappa_D)\sigma_n^2}. \tag{21}$$

To gain further insights, we assume that $\mathbf{h}_{ID} \sim \mathcal{CN}(\mathbf{0}, \sigma_{ID}^2 \mathbf{I}_{N_I})$, where σ_{ID}^2 denotes the variance of each entry of $\mathbf{h}_{ID}$. Then, $|h_{ID,i}|$ follows Rayleigh distribution with mean $\sqrt{\pi}\sigma_{ID}/2$. Accordingly, when $N_I \to \infty$, $\frac{\|\mathbf{h}_{ID}\|_1}{N_I} \to \sqrt{\pi}\sigma_{ID}/2$ and thus we obtain the asymptotic SNR as

$$\text{SNR} \to \frac{\tilde{P} N_S |\eta|^2 \pi N_I^2 \sigma_{ID}^2}{\tilde{P}(\kappa_D N_S + (1+\kappa_D)\kappa_S)|\eta|^2 \pi N_I^2 \sigma_{ID}^2 + 4(1+\kappa_D)\sigma_n^2}. \tag{22}$$

It can be observed from (22) that when the destination noise variance σ_n^2 is non-negligible, the received SNR is monotonically increasing with respect to N_I and σ_{ID}^2. Hence, we can compensate the performance degradation caused by hardware impairments by increasing the size of the IRS and/or reducing the distance between the IRS and the destination. On the other hand, when the destination noise is negligible, the received

TABLE I
SIMULATION PARAMETERS

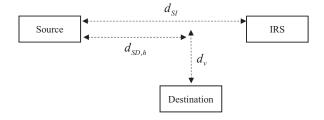| Parameter Name	Parameter Value
Number of Source Transmit Antennas N_S	4
Number of Reflecting Elements N_I	50
Maximum Source Transmit Power P	12 dBW
Normalized Distortion Power $\kappa = \kappa_S = \kappa_D$	0.07
Destination Noise Power σ_n^2	−85 dBW
Reference Distance d_0	1 m
PL at Reference Distance PL_0	−30 dB
PLE of Source-to-IRS Link γ_{SI}	2.5
PLE of IRS-to-Destination Link γ_{ID}	2.5
PLE of Source-to-Destination Link γ_{SD}	3.5
Distance Between Source and IRS d_{SI}	50 m
Vertical Distance Between Destination and Source-to-IRS Link d_v	2 m
Horizontal Distance Between Source and Destination $d_{SD,h}$	49 m

Fig. 2. Simulation setup and the definitions of d_{SI}, d_v, and $d_{SD,h}$ [11].

SNR is less sensitive to the size of the IRS or the distance between the IRS and the destination in the large limit of N_I.

V. SIMULATION RESULTS AND COMPLEXITY ANALYSIS

A. Simulation Setup

In this section, we perform simulations to compare the performance of the following beamforming designs:

- Proposed robust design incorporating hardware impairments for IRS-aided systems ("Robust w/ IRS");
- Nonrobust design neglecting hardware impairments for IRS-aided systems ("Nonrobust w/ IRS");
- Robust design incorporating hardware impairments for conventional systems without IRS ("Robust w/o IRS");
- Nonrobust design neglecting hardware impairments for conventional systems without IRS ("Nonrobust w/o IRS").

For the reflect beamforming optimization of "Robust w/ IRS", we generate an initial point $\tilde{\boldsymbol{\theta}}^{(0)}$ by normalizing each entry of a random $(N_I + 1) \times 1$ vector. "Nonrobust w/ IRS" is obtained by setting $\kappa_S = \kappa_D = 0$ for "Robust w/ IRS". "Robust w/o IRS" is obtained by setting $\boldsymbol{\Theta} = \boldsymbol{0}$ in (8) and "Nonrobust w/o IRS" is achieved by applying MF beamforming at the source pointing towards the desired receiver.

We investigate both received SNR and symbol error rate (SER) performance for the above beamforming designs. For the SNR simulation, we obtain the results by averaging over 500 Rayleigh-faded channels, where we also adopt a theoretical SNR upper bound as a benchmark, which is

calculated by solving problem (12) and substituting the optimal objective value into (10). For the SER simulation, we generate 500 Rayleigh-faded channel realizations and 2,000 random quadrature phase shift keying (QPSK) symbols for each channel realization. Moreover, QPSK demodulation at the destination is performed based on the equalized signal, i.e., $\frac{y}{(\mathbf{h}_{ID}^H \boldsymbol{\Theta} \mathbf{H}_{SI} + \mathbf{h}_{SD}^H) \mathbf{w}}$.

During the simulation, the path loss (PL) of all the channels is modeled by

$$PL = \left(PL_0 - 10\gamma \log_{10} \left(\frac{d}{d_0} \right) \right) \text{ dB}, \quad (23)$$

where PL_0 denotes the path loss at the reference distance d_0, γ represents the path loss exponent (PLE), and d stands for the distance between the transmitter and the receiver. The simulation parameters are set according to Table I unless otherwise specified, where the definitions of d_{SI}, d_v, and $d_{SD,h}$ are illustrated in Fig. 2.

B. Simulation Results

Fig. 3(a) and Fig. 3(b) show the SNR and the SER of various beamforming designs versus N_I, respectively. We can find from both figures that when N_I increases, the performance of both robust and nonrobust designs for IRS-aided systems is improved due to the increasing beamforming gains brought by the IRS. Also, the proposed "Robust w/ IRS" outperforms the conventional "Nonrobust w/ IRS" by evident gains under both continuous and discrete phase shift cases. In particular, the gain is enlarged with the increase of N_I. In contrast, the performance advantage of "Robust w/o IRS" compared with "Nonrobust w/o IRS" is minor. This is due to the fact that both the IRS reflect beamforming and the source transmit beamforming designs are affected by the transceiver hardware impairments for the IRS-assisted communications. It is also worthwhile pointing out that "Robust w/ IRS" achieves almost the same received SNR as the theoretical upper bound in Fig. 3(a), which validates the effectiveness of the proposed design. Moreover, compared to the ideal continuous phase shift case, the discrete phase shifts can lead to performance
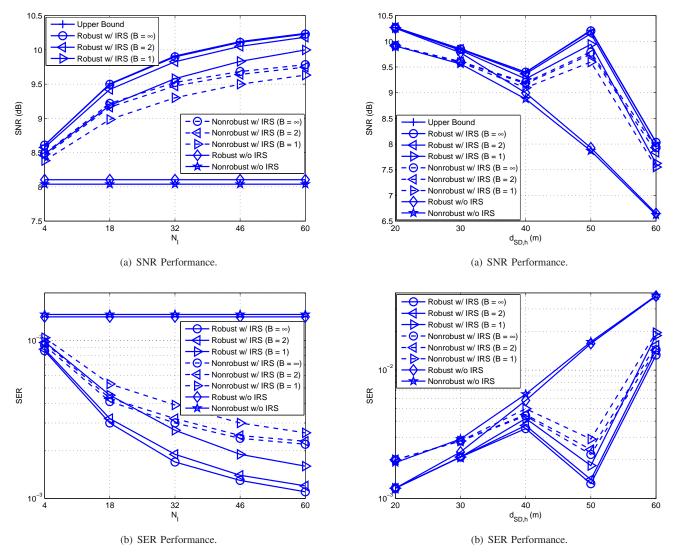
(a) SNR Performance.

(b) SER Performance.

Fig. 3. SNR and SER versus the number of reflecting elements N_I for different designs.

(a) SNR Performance.

(b) SER Performance.

Fig. 4. SNR and SER versus the horizontal distance between source and destination $d_{SD,h}$ for different designs.

degradation especially for large N_I. However, the performance loss becomes minor when we adopt 2 bits to represent the IRS phase shifts. This is because the obtained continuous solution is near optimal and it is possible to find an effective discrete solution by mapping the obtained continuous solution to the discrete phase set. Note that similar observations have been made in prior works such as [16].

We compare the SNR and the SER of various schemes versus $d_{SD,h}$ in Fig. 4(a) and Fig. 4(b), respectively. As can be found from both figures, the performance of the conventional system without IRS degrades rapidly with the increase of $d_{SD,h}$ due to the increase of the associated path loss. In contrast, the performance of the IRS-assisted transmission first degrades, then improves, and finally degrades again when $d_{SD,h}$ becomes larger. It can be seen that the best performance is achieved when the destination is closest to the IRS, i.e., $d_{SD,h} = d_{SI} = 50$ m. This is because the introduction of an IRS helps to enhance the received signal strength via signal reflection when the destination gets close to the IRS,

thus improving the system performance. Note that similar phenomenon has also been observed in [11]. We can also find from both figures that the proposed "Robust w/ IRS" outperforms "Nonrobust w/ IRS" under different values of $d_{SD,h}$, which again validates the necessity of incorporating the hardware impairments into the beamforming optimization.

The SNR and the SER versus P are shown in Fig. 5(a) and Fig. 5(b), respectively. We can find that the performance of all the methods is enhanced as P increases. Similarly to previous observations, the performance improvement brought by the schemes taking into account the hardware impairments is significant, particularly for the IRS-assisted transmission. Moreover, there exist error floors for the IRS-aided communications when P is sufficiently large in Fig. 5(b), which is due to the fact that the power of the distortions caused by hardware impairments also grows with the increase of P, as predicted in (3). Nevertheless, the proposed optimization framework can effectively lower the error floors for a given transmit power P compared with the nonrobust scheme.
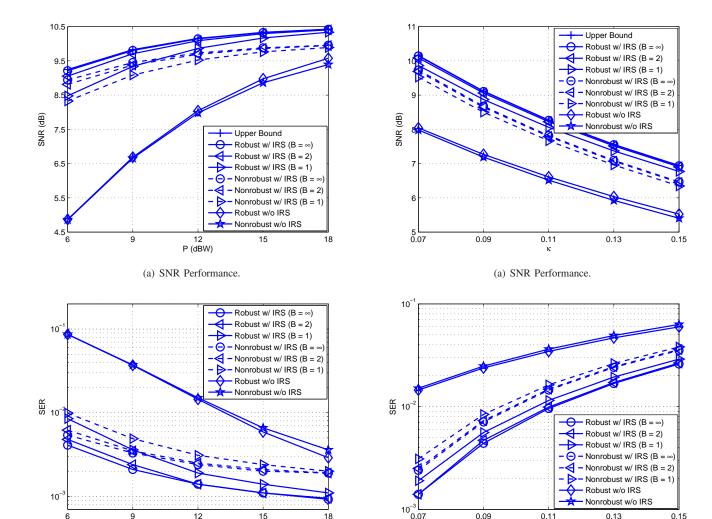
(a) SNR Performance.

(a) SNR Performance.

(b) SER Performance.

(b) SER Performance.

Fig. 5. SNR and SER versus source transmit power P for different designs.

Fig. 6. SNR and SER versus normalized distortion power κ for different designs.

We show the SNR and the SER comparisons versus κ in Fig. 6(a) and Fig. 6(b), respectively. It can be seen that the performance of various designs deteriorates with the increase of κ since the negative impact of hardware impairments becomes more prominent. It can also been found that large performance gains can be achieved due to the use of the IRS for relatively small κ. Yet, the gains become smaller for large κ since the hardware impairments severely degrade the performance of both IRS-aided and conventional communications. In addition, as expected, by taking into account the hardware impairments, the proposed "Robust w/ IRS" outperforms "Nonrobust w/ IRS" under different κ's.

C. Complexity Analysis

Now we compare the computational complexity of all the above considered beamforming schemes as shown in Table II, N_{iter} is the number of iterations required by "Robust w/ IRS", and N'_{iter} is the number of iterations needed by "Nonrobust w/ IRS". It can be found that all the methods have polynomial-time computational complexity. The major

computation of "Robust w/ IRS" in each iteration lies in the calculation of the vector $(\mathbf{\Psi}^H \mathbf{\Xi}_0^{-1} \mathbf{\Psi} - (1 + \kappa_D)\kappa_S (\mathbf{\Omega} - \lambda_{\max}(\mathbf{\Omega}) \mathbf{I}_{N_I+1})) \tilde{\boldsymbol{\theta}}^{(k-1)}$ whose complexity is $\mathcal{O}(N_S N_I + N_I^2)$. Note that $\lambda_{\max}(\mathbf{\Omega})$ is computed via the efficient power iteration method [45]. "Nonrobust w/ IRS" mainly needs to calculate the vector $\mathbf{\Psi}^H \mathbf{\Psi} \boldsymbol{\theta}^{(k-1)}$ in each iteration, whose complexity is $\mathcal{O}(N_S N_I)$.

We compare the average number of iterations required by "Robust w/ IRS" and "Nonrobust w/ IRS" in Table III for different values of N_I. It can be seen that the algorithm convergence becomes slower when N_I increases. Moreover, the use of the acceleration scheme [42, Section V-B] results in faster convergence especially for "Robust w/ IRS". "Robust w/ IRS" needs more iterations to converge than "Nonrobust w/ IRS" while the gap becomes smaller after applying the acceleration method.

Based on the above simulation results and complexity analysis, we can conclude that the proposed robust design incorporating hardware impairments can strike a good balance

$$\left(\kappa_D\mathbf{\Upsilon}(\mathbf{\Theta}) + (1+\kappa_D)\kappa_S\widetilde{\text{diag}}\{\mathbf{\Upsilon}(\mathbf{\Theta})\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S}\right)^{-1}$$

$$=\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta}) - \frac{\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD})(\mathbf{h}_{ID}^H\mathbf{\Theta}\mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})}{(\mathbf{h}_{ID}^H\mathbf{\Theta}\mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD}) + (1/\kappa_D)} \tag{24}$$

$$\left(\kappa_D\mathbf{\Upsilon}(\mathbf{\Theta}) + \tilde{\mathbf{\Upsilon}}(\mathbf{\Theta})\right)^{-1}(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD})$$

$$=\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD}) - \frac{\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD})(\mathbf{h}_{ID}^H\mathbf{\Theta}\mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD})}{(\mathbf{h}_{ID}^H\mathbf{\Theta}\mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD}) + (1/\kappa_D)}$$

$$=\left(1 - \frac{(\mathbf{h}_{ID}^H\mathbf{\Theta}\mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD})}{(\mathbf{h}_{ID}^H\mathbf{\Theta}\mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD}) + (1/\kappa_D)}\right)\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD})$$

$$=\frac{1}{\kappa_D(\mathbf{h}_{ID}^H\mathbf{\Theta}\mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD}) + 1}\tilde{\mathbf{\Upsilon}}^{-1}(\mathbf{\Theta})(\mathbf{H}_{SI}^H\mathbf{\Theta}^H\mathbf{h}_{ID} + \mathbf{h}_{SD}) \tag{25}$$

$$\underset{\boldsymbol{\theta}}{\text{maximize}} \quad (\boldsymbol{\theta}^H\text{diag}\{\mathbf{h}_{ID}^H\}\mathbf{H}_{SI} + \mathbf{h}_{SD}^H)\left((1+\kappa_D)\kappa_S\widetilde{\text{diag}}\{\mathbf{\Upsilon}(\boldsymbol{\theta})\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S}\right)^{-1}(\mathbf{H}_{SI}^H\text{diag}\{\mathbf{h}_{ID}\}\boldsymbol{\theta} + \mathbf{h}_{SD})$$

$$\text{subject to} \quad |\theta_i| = 1, \ i = 1, \cdots, N_I \tag{26}$$

$$\underset{\boldsymbol{\theta},t}{\text{maximize}} \quad (\boldsymbol{\theta}^H\text{diag}\{\mathbf{h}_{ID}^H\}\mathbf{H}_{SI} + t^*\mathbf{h}_{SD}^H)\left((1+\kappa_D)\kappa_S\widetilde{\text{diag}}\{\mathbf{\Upsilon}(\boldsymbol{\theta},t)\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S}\right)^{-1}(\mathbf{H}_{SI}^H\text{diag}\{\mathbf{h}_{ID}\}\boldsymbol{\theta} + t\mathbf{h}_{SD})$$

$$\text{subject to} \quad |\theta_i| = 1, \ i = 1, \cdots, N_I,$$

$$\quad\quad\quad\quad\quad |t| = 1 \tag{27}$$

---

of the tradeoff between performance and computational complexity.

VI. CONCLUSIONS

We investigated the beamforming optimization for IRS-aided multi-antenna communications in the presence of transceiver hardware impairments. By incorporating signal distortions caused by the hardware impairments, we jointly optimized the source transmit beamforming and IRS reflect beamforming to maximize the received SNR at the destination. The optimal solution to the source transmit beamforming was obtained in a closed form, which turns out to be weighted MF beamforming. An MM-based algorithm was developed to optimize the IRS reflect beamforming, which approaches the theoretical upper bound with low computational complexity. The performance advantage of the proposed beamforming design was also validated via extensive simulations and the importance of considering hardware impairments in the beamforming design was unveiled.

APPENDIX A
PROOF OF PROPOSITION 1

We first apply the matrix inversion lemma [46] and rewrite the matrix inversion term in (7) by (24), where $\tilde{\mathbf{\Upsilon}}(\mathbf{\Theta}) \triangleq (1+\kappa_D)\kappa_S\widetilde{\text{diag}}\{\mathbf{\Upsilon}(\mathbf{\Theta})\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S}$. Then, it follows that (25) holds. Substituting (25) into (7), we obtain (8).

APPENDIX B
PROOF OF THEOREM 1

We first rewrite problem (11) by (26), where $\boldsymbol{\theta} \triangleq [\theta_1^*, \cdots, \theta_{N_I}^*]^T$ and $\mathbf{\Upsilon}(\boldsymbol{\theta}) \triangleq (\mathbf{H}_{SI}^H\text{diag}\{\mathbf{h}_{ID}\}\boldsymbol{\theta} + \mathbf{h}_{SD})(\boldsymbol{\theta}^H\text{diag}\{\mathbf{h}_{ID}^H\}\mathbf{H}_{SI} + \mathbf{h}_{SD}^H)$. Then, we introduce a complex slack variable t satisfying $|t| = 1$ and recast the above problem as (27), where $\mathbf{\Upsilon}(\boldsymbol{\theta},t) \triangleq (\mathbf{H}_{SI}^H\text{diag}\{\mathbf{h}_{ID}\}\boldsymbol{\theta} + t\mathbf{h}_{SD})(\boldsymbol{\theta}^H\text{diag}\{\mathbf{h}_{ID}^H\}\mathbf{H}_{SI} + t^*\mathbf{h}_{SD}^H)$. Since $|t| = 1$, it can be readily confirmed that, if $(\boldsymbol{\theta}^\star, t^\star)$ is optimal to problem (27), $\boldsymbol{\theta}^\star/t^\star$ will also be optimal to problem (26) and vice versa. Therefore, problems (26) and (27) are equivalent.

Define $\tilde{\boldsymbol{\theta}} \triangleq [\boldsymbol{\theta}^T \ t]^T$ and $\mathbf{\Psi} \triangleq [\mathbf{H}_{SI}^H\text{diag}\{\mathbf{h}_{ID}\} \ \mathbf{h}_{SD}]$. Then, we rewrite problem (27) by (28) on the top of next page, where $\tilde{\theta}_i$ is the i-th entry of $\tilde{\boldsymbol{\theta}}$. After some matrix manipulations,

$$\text{maximize}_{\tilde{\boldsymbol{\theta}}} \quad \tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H \left((1+\kappa_D)\kappa_S \widetilde{\text{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S} \right)^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}$$
$$\text{subject to} \quad |\tilde{\theta}_i| = 1, \ i = 1,\cdots, N_I + 1 \tag{28}$$

$$[(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}})_1^*,\cdots,(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}})_{N_S}^*]\text{diag}\left\{ \left[\left((1+\kappa_D)\kappa_S(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H)_{1,1} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}} \right)^{-1}, \cdots, \right. \right.$$
$$\left. \left. \left((1+\kappa_D)\kappa_S(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H)_{N_S,N_S} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}} \right)^{-1} \right] \right\} [(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}})_1, \cdots, (\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}})_{N_S}]^T$$
$$= \sum_{m=1}^{N_S} \frac{|(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}})_m|^2}{(1+\kappa_D)\kappa_S(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H)_{m,m} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}} \tag{29}$$

TABLE II
BEAMFORMING DESIGNS COMPLEXITY

Beamforming Design Method	Computational Complexity
Robust w/ IRS	$\mathcal{O}(N_{\text{iter}}(N_S N_I + N_I^2))$
Nonrobust w/ IRS	$\mathcal{O}(N_{\text{iter}}' N_S N_I)$
Robust w/o IRS	$\mathcal{O}(N_S)$
Nonrobust w/o IRS	$\mathcal{O}(N_S)$

TABLE III
COMPARISON OF AVERAGE NUMBER OF ITERATIONS (CONVERGENCE ACCURACY $\epsilon = 10^{-5}$)

Beamforming Design Method	Average Number of Iterations				
	$N_I = 4$	$N_I = 18$	$N_I = 32$	$N_I = 46$	$N_I = 60$
Robust w/o Acceleration	161.69	280.61	411.24	492.84	518.17
Robust w/ Acceleration	8.31	19.92	27.1	33.44	36.57
Nonrobust w/o Acceleration	5.13	9.82	14.63	15.65	17.83
Nonrobust w/ Acceleration	2.97	4.31	5.73	6.2	6.52

we express the objective function of problem (28) by (29) on the top of next page. Since $|(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}})_m|^2 = (\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H)_{m,m}$, problem (28) now becomes

$$\text{maximize}_{\tilde{\boldsymbol{\theta}}} \quad \sum_{m=1}^{N_S} \frac{(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H)_{m,m}}{(1+\kappa_D)\kappa_S(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H)_{m,m} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}}$$
$$\text{subject to} \quad |\tilde{\theta}_i| = 1, \ i = 1,\cdots, N_I + 1. \tag{30}$$

By defining $\tilde{\boldsymbol{\Theta}} \triangleq \tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H$ and relaxing the rank-one constraint with respect to $\tilde{\boldsymbol{\Theta}}$, we obtain problem (12), i.e.,

$$\text{maximize}_{\tilde{\boldsymbol{\Theta}} \succeq \mathbf{0}} \quad \sum_{m=1}^{N_S} \frac{(\boldsymbol{\Psi}\tilde{\boldsymbol{\Theta}}\boldsymbol{\Psi}^H)_{m,m}}{(1+\kappa_D)\kappa_S(\boldsymbol{\Psi}\tilde{\boldsymbol{\Theta}}\boldsymbol{\Psi}^H)_{m,m} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}}$$
$$\text{subject to} \quad (\tilde{\boldsymbol{\Theta}})_{i,i} = 1, \ i = 1,\cdots, N_I + 1. \tag{31}$$

Clearly, the optimal objective value of this problem must be an upper bound to that of problem (11) due to the relaxation of the rank-one constraint.

To prove that the above problem is convex, we only need to show that the objective function is concave with respect to $\tilde{\boldsymbol{\Theta}}$. We express the objective function by

$$\sum_{m=1}^{N_S} \frac{(\boldsymbol{\Psi}\tilde{\boldsymbol{\Theta}}\boldsymbol{\Psi}^H)_{m,m}}{(1+\kappa_D)\kappa_S(\boldsymbol{\Psi}\tilde{\boldsymbol{\Theta}}\boldsymbol{\Psi}^H)_{m,m} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}} = \frac{1}{(1+\kappa_D)\kappa_S}$$
$$\times \sum_{m=1}^{N_S} \left(1 - \frac{\frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}}{(1+\kappa_D)\kappa_S(\boldsymbol{\Psi}\tilde{\boldsymbol{\Theta}}\boldsymbol{\Psi}^H)_{m,m} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}} \right). \tag{32}$$

Define $q_m \triangleq (\boldsymbol{\Psi}\tilde{\boldsymbol{\Theta}}\boldsymbol{\Psi}^H)_{m,m}$ and $u(q_m) \triangleq -\frac{1}{(1+\kappa_D)\kappa_S q_m + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}}$. Then, it can be readily verified that $u(q_m)$ is concave. Moreover, q_m is an affine function of $\tilde{\boldsymbol{\Theta}}$. Therefore, the objective function is also concave according to [47, Section 3.2.2].

APPENDIX C
PROOF OF PROPOSITION 2

Define $\boldsymbol{\Xi} \triangleq (1+\kappa_D)\kappa_S \widetilde{\text{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H\} + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\mathbf{I}_{N_S}$. Since the function $f(\mathbf{x}, \mathbf{Y}) = \mathbf{x}^H \mathbf{Y}^{-1}\mathbf{x}$ is jointly convex with

$$\hat{f}(\tilde{\boldsymbol{\theta}},\boldsymbol{\Xi},\tilde{\boldsymbol{\theta}}_0,\boldsymbol{\Xi}_0) \overset{(a)}{=} \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0 + 2\Re\{\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}(\tilde{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}_0)\} + \text{tr}(-\boldsymbol{\Xi}_0^{-1}(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H)^T \boldsymbol{\Xi}_0^{-1}$$
$$\times ((1+\kappa_D)\kappa_S \widetilde{\text{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H\} - (1+\kappa_D)\kappa_S \widetilde{\text{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H\}))$$
$$= 2\Re\{\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\} - (1+\kappa_D)\kappa_S \text{tr}(\boldsymbol{\Xi}_0^{-1}(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H)^T \boldsymbol{\Xi}_0^{-1} \widetilde{\text{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H\})$$
$$- \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0 + (1+\kappa_D)\kappa_S \text{tr}(\boldsymbol{\Xi}_0^{-1}(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H)^T \boldsymbol{\Xi}_0^{-1} \widetilde{\text{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H\})$$
$$\overset{(b)}{=} 2\Re\{\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\} - (1+\kappa_D)\kappa_S \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \widetilde{\text{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H\} \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0$$
$$- \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0 + (1+\kappa_D)\kappa_S \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \widetilde{\text{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H\} \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0$$
$$\overset{(c)}{=} 2\Re\{\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\} - (1+\kappa_D)\kappa_S \tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H \widetilde{\text{diag}}\{\boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1}\} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}$$
$$- \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0 + (1+\kappa_D)\kappa_S \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \widetilde{\text{diag}}\{\boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1}\} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0$$
$$\triangleq \hat{f}(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\theta}}_0, \boldsymbol{\Xi}_0) \tag{34}$$

$$\bar{f}(\tilde{\boldsymbol{\theta}}_0; \tilde{\boldsymbol{\theta}}_0) = 2\Re\{\tilde{\boldsymbol{\theta}}_0^H(\boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi} - (1+\kappa_D)\kappa_S(\boldsymbol{\Omega} - \lambda_{\max}(\boldsymbol{\Omega})\mathbf{I}_{N_I+1}))\tilde{\boldsymbol{\theta}}_0\} - 2(1+\kappa_D)\kappa_S(N_I+1)\lambda_{\max}(\boldsymbol{\Omega})$$
$$+ \tilde{\boldsymbol{\theta}}_0^H(2(1+\kappa_D)\kappa_S\boldsymbol{\Omega} - \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi})\tilde{\boldsymbol{\theta}}_0$$
$$= \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0 + 2(1+\kappa_D)\kappa_S \|\tilde{\boldsymbol{\theta}}_0\|^2 \lambda_{\max}(\boldsymbol{\Omega}) - 2(1+\kappa_D)\kappa_S(N_I+1)\lambda_{\max}(\boldsymbol{\Omega})$$
$$\overset{(a)}{=} \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0$$
$$\overset{(b)}{=} f(\tilde{\boldsymbol{\theta}}_0) \tag{36}$$

$(\mathbf{x}, \mathbf{Y})$ for $\mathbf{Y} \succ \mathbf{0}$ [47, Section 3.1.7] and $\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}$ is an affine transformation of $\tilde{\boldsymbol{\theta}}$, $f(\tilde{\boldsymbol{\theta}}, \boldsymbol{\Xi}) = \tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}$ is also jointly convex with $(\tilde{\boldsymbol{\theta}}, \boldsymbol{\Xi})$ [47, Section 3.2.2]. Recall that a convex function is lower bounded by its first-order Taylor expansion. Hence, by applying the complex-valued derivatives in [48, Section 4.2], we have

$$f(\tilde{\boldsymbol{\theta}}, \boldsymbol{\Xi}) \geq \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0 + 2\Re\{\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}(\tilde{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}_0)\}$$
$$+ \text{tr}(-\boldsymbol{\Xi}_0^{-1}(\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H)^T \boldsymbol{\Xi}_0^{-1}(\boldsymbol{\Xi} - \boldsymbol{\Xi}_0))$$
$$\triangleq \hat{f}(\tilde{\boldsymbol{\theta}}, \boldsymbol{\Xi}, \tilde{\boldsymbol{\theta}}_0, \boldsymbol{\Xi}_0), \tag{33}$$

where $\boldsymbol{\Xi}_0 \triangleq (1+\kappa_D)\kappa_S \widetilde{\text{diag}}\{\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H\} + \frac{(1+\kappa_D)\sigma_n^2}{\bar{P}}\mathbf{I}_{N_S}$ and $\tilde{\boldsymbol{\theta}}_0$ is a feasible point. Furthermore, we express $\hat{f}(\tilde{\boldsymbol{\theta}}, \boldsymbol{\Xi}, \tilde{\boldsymbol{\theta}}_0, \boldsymbol{\Xi}_0)$ by (34), where (a) follows from the definitions of $\boldsymbol{\Xi}$ and $\boldsymbol{\Xi}_0$, (b) holds due to $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$, and (c) follows from the equation $\mathbf{a}^H \widetilde{\text{diag}}\{\mathbf{bb}^H\}\mathbf{a} = \mathbf{b}^H \widetilde{\text{diag}}\{\mathbf{aa}^H\}\mathbf{b}$.

According to the inequality given by [41, Eq. (26)], the quadratic term in $\hat{f}(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\theta}}_0, \boldsymbol{\Xi}_0)$ is upper bounded by

$$\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H \widetilde{\text{diag}}\{\boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1}\} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}$$
$$\overset{(a)}{\leq} \lambda_{\max}(\boldsymbol{\Omega})\|\tilde{\boldsymbol{\theta}}\|^2 + 2\Re\{\tilde{\boldsymbol{\theta}}_0^H(\boldsymbol{\Omega} - \lambda_{\max}(\boldsymbol{\Omega})\mathbf{I}_{N_I+1})\tilde{\boldsymbol{\theta}}\}$$
$$+ \tilde{\boldsymbol{\theta}}_0^H(\lambda_{\max}(\boldsymbol{\Omega})\mathbf{I}_{N_I+1} - \boldsymbol{\Omega})\tilde{\boldsymbol{\theta}}_0$$
$$\overset{(b)}{=} 2\Re\{\tilde{\boldsymbol{\theta}}_0^H(\boldsymbol{\Omega} - \lambda_{\max}(\boldsymbol{\Omega})\mathbf{I}_{N_I+1})\tilde{\boldsymbol{\theta}}\} + 2(N_I+1)\lambda_{\max}(\boldsymbol{\Omega})$$
$$- \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Omega}\tilde{\boldsymbol{\theta}}_0, \tag{35}$$

where $\boldsymbol{\Omega} \triangleq \boldsymbol{\Psi}^H \widetilde{\text{diag}}\{\boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1}\} \boldsymbol{\Psi}$ in (a) and (b) holds because $|\tilde{\theta}_i| = |(\tilde{\theta}_0)_i| = 1$, $i = 1, \cdots, N_I + 1$. By substituting the above inequality into (34), we obtain (13).

APPENDIX D
PROOF OF THEOREM 2

Since problems (28) and (11) are equivalent, we focus on the convergence proof for problem (28). Define the objective function of problem (28) and its lower bound in Proposition 2 by $f(\tilde{\boldsymbol{\theta}})$ and $\bar{f}(\tilde{\boldsymbol{\theta}}; \tilde{\boldsymbol{\theta}}_0)$, respectively. Denote the constraint set of problem (28) by $\mathcal{S}$. Then, in order to prove Theorem 2, we need to verify the following four conditions according to [49, Section III]:

1) $f(\tilde{\boldsymbol{\theta}}) \geq \bar{f}(\tilde{\boldsymbol{\theta}}; \tilde{\boldsymbol{\theta}}_0), \forall \tilde{\boldsymbol{\theta}} \in \mathcal{S}$;
2) $f(\tilde{\boldsymbol{\theta}}_0) = \bar{f}(\tilde{\boldsymbol{\theta}}_0; \tilde{\boldsymbol{\theta}}_0)$;
3) $\nabla f(\tilde{\boldsymbol{\theta}}_0) = \nabla \bar{f}(\tilde{\boldsymbol{\theta}}_0; \tilde{\boldsymbol{\theta}}_0)$;
4) $\bar{f}(\tilde{\boldsymbol{\theta}}; \tilde{\boldsymbol{\theta}}_0)$ is continuous in both $\tilde{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}_0$.

Note that the convergence of the MM algorithm is guaranteed by the first two conditions while conditions 3) and 4) guarantee convergence to a stationary point.

Clearly, conditions 1) and 4) hold due to Proposition 2. Moreover, from (13), we have (36), where (a) holds since $\tilde{\boldsymbol{\theta}}_0 \in \mathcal{S}$, i.e., $|(\tilde{\theta}_0)_i| = 1$, $i = 1, \cdots, N_I + 1$, and (b) is obtained based on the definition of $\boldsymbol{\Xi}_0$ in Proposition 2. Therefore, condition 2) is also verified.

Concerning condition 3), we first rewrite $f(\tilde{\boldsymbol{\theta}})$ in (29) by

$$f(\tilde{\boldsymbol{\theta}}) = \sum_{m=1}^{N_S} \frac{\|\mathbf{e}_m^T \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\|^2}{(1+\kappa_D)\kappa_S \|\mathbf{e}_m^T \boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\|^2 + \frac{(1+\kappa_D)\sigma_n^2}{\bar{P}}}. \tag{37}$$

$$
\begin{aligned}
\nabla \bar{f}(\tilde{\boldsymbol{\theta}}; \tilde{\boldsymbol{\theta}}_0) =& 2\tilde{\boldsymbol{\theta}}_0^H (\boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi} - (1+\kappa_D)\kappa_S(\boldsymbol{\Omega} - \lambda_{\max}(\boldsymbol{\Omega})\mathbf{I}_{N_I+1})) - 2(1+\kappa_D)\kappa_S\lambda_{\max}(\boldsymbol{\Omega})\tilde{\boldsymbol{\theta}}^H \\
=& 2\tilde{\boldsymbol{\theta}}_0^H (\boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi} - (1+\kappa_D)\kappa_S\boldsymbol{\Omega}) + 2(1+\kappa_D)\kappa_S\lambda_{\max}(\boldsymbol{\Omega})(\tilde{\boldsymbol{\theta}}_0^H - \tilde{\boldsymbol{\theta}}^H) \\
\stackrel{(a)}{=}& 2\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H (\boldsymbol{\Xi}_0^{-1} - (1+\kappa_D)\kappa_S\widetilde{\mathrm{diag}}\{\boldsymbol{\Xi}_0^{-1}\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^H\boldsymbol{\Psi}^H\boldsymbol{\Xi}_0^{-1}\})\boldsymbol{\Psi} + 2(1+\kappa_D)\kappa_S\lambda_{\max}(\boldsymbol{\Omega})(\tilde{\boldsymbol{\theta}}_0^H - \tilde{\boldsymbol{\theta}}^H) \\
\stackrel{(b)}{=}& 2 \sum_{m=1}^{N_S} \tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \mathbf{e}_m \mathbf{e}_m^T \boldsymbol{\Psi} \left(\frac{1}{(1+\kappa_D)\kappa_S\|\mathbf{e}_m^T\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\|^2 + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}} - \frac{(1+\kappa_D)\kappa_S\|\mathbf{e}_m^T\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\|^2}{\left((1+\kappa_D)\kappa_S\|\mathbf{e}_m^T\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\|^2 + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\right)^2} \right) \\
& + 2(1+\kappa_D)\kappa_S\lambda_{\max}(\boldsymbol{\Omega})(\tilde{\boldsymbol{\theta}}_0^H - \tilde{\boldsymbol{\theta}}^H) \\
=& \frac{2(1+\kappa_D)\sigma_n^2}{\tilde{P}} \sum_{m=1}^{N_S} \frac{\tilde{\boldsymbol{\theta}}_0^H \boldsymbol{\Psi}^H \mathbf{e}_m \mathbf{e}_m^T \boldsymbol{\Psi}}{\left((1+\kappa_D)\kappa_S\|\mathbf{e}_m^T\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}_0\|^2 + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\right)^2} + 2(1+\kappa_D)\kappa_S\lambda_{\max}(\boldsymbol{\Omega})(\tilde{\boldsymbol{\theta}}_0^H - \tilde{\boldsymbol{\theta}}^H) \quad (40)
\end{aligned}
$$

Then, we can readily calculate $\nabla f(\tilde{\boldsymbol{\theta}})$ by

$$
\nabla f(\tilde{\boldsymbol{\theta}}) = \frac{2(1+\kappa_D)\sigma_n^2}{\tilde{P}}
$$
$$
\times \sum_{m=1}^{N_S} \frac{\tilde{\boldsymbol{\theta}}^H \boldsymbol{\Psi}^H \mathbf{e}_m \mathbf{e}_m^T \boldsymbol{\Psi}}{\left((1+\kappa_D)\kappa_S\|\mathbf{e}_m^T\boldsymbol{\Psi}\tilde{\boldsymbol{\theta}}\|^2 + \frac{(1+\kappa_D)\sigma_n^2}{\tilde{P}}\right)^2}. \quad (38)
$$

On the other hand, since $\|\tilde{\boldsymbol{\theta}}\|^2 = N_I + 1$, we express $\bar{f}(\tilde{\boldsymbol{\theta}}; \tilde{\boldsymbol{\theta}}_0)$ by

$$
\bar{f}(\tilde{\boldsymbol{\theta}}; \tilde{\boldsymbol{\theta}}_0) =
$$
$$
2\Re\{\tilde{\boldsymbol{\theta}}_0^H (\boldsymbol{\Psi}^H \boldsymbol{\Xi}_0^{-1} \boldsymbol{\Psi} - (1+\kappa_D)\kappa_S(\boldsymbol{\Omega} - \lambda_{\max}(\boldsymbol{\Omega})\mathbf{I}_{N_I+1}))\tilde{\boldsymbol{\theta}}\}
$$
$$
- (1+\kappa_D)\kappa_S\|\tilde{\boldsymbol{\theta}}\|^2\lambda_{\max}(\boldsymbol{\Omega}) - (1+\kappa_D)\kappa_S(N_I+1)\lambda_{\max}(\boldsymbol{\Omega})
$$
$$
+ \tilde{\boldsymbol{\theta}}_0^H (2(1+\kappa_D)\kappa_S\boldsymbol{\Omega} - \boldsymbol{\Psi}^H\boldsymbol{\Xi}_0^{-1}\boldsymbol{\Psi})\tilde{\boldsymbol{\theta}}_0. \quad (39)
$$

Accordingly, $\nabla \bar{f}(\tilde{\boldsymbol{\theta}}; \tilde{\boldsymbol{\theta}}_0)$ is obtained by (40), where (a) and (b) are achieved based on the definitions of $\boldsymbol{\Omega}$ and $\boldsymbol{\Xi}_0$, respectively. By setting $\tilde{\boldsymbol{\theta}} = \tilde{\boldsymbol{\theta}}_0$ for $\nabla f(\tilde{\boldsymbol{\theta}})$ in (38) and $\nabla \bar{f}(\tilde{\boldsymbol{\theta}}; \tilde{\boldsymbol{\theta}}_0)$ in (40), we eventually arrive at condition 3). Therefore, the proof is completed.

REFERENCES

[1] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, Sep. 2018.

[2] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless networks," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.

[3] E. Basar, M. D. Renzo, J. de Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753–116773, 2019.

[4] M. Di Renzo, *et al.*, "Smart radio environments empowered by AI reconfigurable meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, 2019, Art. no. 129.

[5] J. Zhang, *et al.*, "Prospective multiple antenna technologies for beyond 5G," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1637–1660, Aug. 2020.

[6] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," 2020, [Online] Available: https://arxiv.org/abs/2002.03491.

[7] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and road ahead", *accepted by IEEE J. Sel. Areas Commun.*, 2020.

[8] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[9] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[10] S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of data transmission with large intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2746–2758, May 2018.

[11] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

[12] Q. Wu and R. Zhang, "Beamforming optimization for intelligent reflecting surface with discrete phase shifts," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Brighton, UK, May 2019, pp. 7830–7833.

[13] 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.

[14] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, Mar. 2020.

[15] C. Huang, A. Zappone, M. Debbah, and C. Yuen, "Achievable rate maximization by passive intelligent mirrors," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Calgary, Canada, Apr. 2018, pp. 1–5.

[16] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3064–3076, May 2020.

[17] 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, Aug. 2019.

[18] C. Huang, G. C. Alexandropoulos, A. Zappone, M. Debbah, and C. Yuen, "Energy efficient multi-user MISO communication using low resolution large intelligent surfaces," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

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

[20] X. Yu, D. Xu, and R. Schober, "Enabling secure wireless communications via intelligent reflecting surfaces," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

[21] 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.

[22] 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, p. 778–782, Jun. 2020.

[23] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," *accepted by IEEE J. Sel. Areas Commun.*, 2020.

[24] Q. Wu and R. Zhang, "Weighted sum power maximization for intelligent reflecting surface aided SWIPT," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 586–590, May 2020.

[25] H. Suzuki, T. V. A. Tran, I. B. Collings, G. Daniels, and M. Hedley, "Transmitter noise effect on the performance of a MIMO-OFDM hardware implementation achieving improved coverage," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 6, pp. 867–876, Aug. 2008.

[26] W. Namgoong, "Modeling and analysis of nonlinearities and mismatches in AC-coupled direct-conversion receiver," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 163–173, Jan. 2005.

[27] E. Bjornson, P. Zetterberg, M. Bengtsson, and B. Ottersten, "Capacity limits and multiplexing gains of MIMO channels with transceiver impairments," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 91–94, Jan. 2013.

[28] E. Bjornson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inform. Theory*, vol. 60, no. 11, pp. 7112–7139, Nov. 2014.

[29] J. Zhu, D. W. K. Ng, N. Wang, R. Schober, and V. K. Bhargava, "Analysis and design of secure massive MIMO systems in the presence of hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 2001–2016, Mar. 2017.

[30] E. Bjornson, P. Zetterberg, and M. Bengtsson, "Optimal coordinated beamforming in the multicell downlink with transceiver impairments," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Anaheim, CA, USA, Dec. 2012, pp. 4775–4780.

[31] J. You, E. Liu, R. Wang, and W. Su, "Joint source and relay precoding design for MIMO two-way relay systems with transceiver impairments," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 572–575, Mar. 2017.

[32] H. Shen, C. Liu, W. Xu, and C. Zhao, "Optimized full-duplex MIMO DF relaying with limited dynamic range," *IEEE Access*, vol. 5, pp. 20726–20735, 2017.

[33] O. Taghizadeh, A. C. Cirik, and R. Mathar, "Hardware impairments aware transceiver design for full-duplex amplify-and-forward MIMO relaying," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1644–1659, Mar. 2018.

[34] L. C. Godara, "The effect of phase-shifter errors on the performance of an antenna-array beamformer," *IEEE J. Oceanic Eng.*, vol. 10, no. 3, pp. 278–284, Jul. 1985.

[35] B. Peng, S. Priebe, and T. Kurner, "Effects of phase shift errors on the antenna directivity of phased arrays in indoor terahertz communications," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Barcelona, Spain, Aug. 2014, pp. 355–359.

[36] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, "Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Honolulu, USA, Apr. 2018, pp. 270–278.

[37] L. Dai *et al.*, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45913–45923, 2020.

[38] H. Xu, S. Xu, F. Yang, and M. Li, "Design and experiment of a dual-band 1-bit reconfigurable reflectarray antenna with independent large-angle beam scanning capability," *accepted by IEEE Antennas Wireless Propag. Lett.*, 2020.

[39] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[40] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," Dec. 2018, software available at http://cvxr.com/cvx/.

[41] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

[42] J. Song, P. Babu, and D. P. Palomar, "Sequence design to minimize the weighted integrated and peak sidelobe levels," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2051–2064, Apr. 2016.

[43] T. L. Jensen and E. De Carvalho, "An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Barcelona, Spain, May 2020, pp. 5000–5004.

[44] C. You, B. Zheng, and R. Zhang, "Intelligent reflecting surface with discrete phase shifts: Channel estimation and passive beamforming," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, Jun. 2020, pp. 1–6.

[45] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. Oxford, U.K.: Clarendon Press, 1965.

[46] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York, NY, USA: Cambridge Univ. Press, 1985.

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

[48] A. Hjoungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[49] L. Zhao, J. Song, and D. P. Palomar, "A unified framework for low autocorrelation sequence design via majorization-minimization," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 438–453, Jan. 2017.