

Improper Signaling for Multicell MIMO RIS-assisted Broadcast Channels with I/Q Imbalance

Mohammad Soleymani*, Ignacio Santamaria† *Senior Member, IEEE* and Peter J. Schreier*, *Senior Member, IEEE*

Abstract—This paper studies the performance of improper Gaussian signaling (IGS) in a multicell broadcast multiple-input, multiple-output (MIMO) reconfigurable-intelligent-surface (RIS)-assisted channel. The transceivers are non-ideal devices with I/Q imbalance, which further motivates the use of IGS. We propose IGS schemes to improve the spectral and energy efficiency (EE) of the system by solving different optimization problems such as minimum-weighted-rate, weighted-sum-rate, minimum-weighted-EE and global-EE maximization. Two different RIS implementations are considered: in the first one, the phase and amplitude of the RIS components can be optimized independently, which provides an upper-bound for RIS performance. In the second implementation, the amplitude of each RIS component is fixed, and only its phase can be optimized, which is referred to as reflecting surfaces. We show that RIS can significantly increase the spectral and energy efficiency of the system, while the reflecting surfaces perform very close to the upper-bound performance of RISs. Moreover, distributed RIS implementations that use spatially separated RIS with fewer components, can outperform centralized implementations consisting of a single RIS with co-located reflecting elements. Additionally, our results indicate that IGS can provide considerable gains from both spectral and energy-efficiency perspectives, and the benefits of IGS can be even higher in RIS-assisted systems.

Index Terms—Energy efficiency, fairness rate, improper Gaussian signaling, I/Q imbalance, majorization minimization, MIMO broadcast channels, reflecting intelligent surface, sum-rate maximization.

I. INTRODUCTION

Reconfigurable intelligent surfaces (RISs) contain a large number of low-cost controllable elements, which have sizes and inter-distances much smaller than the wavelength, and are able to dynamically modulate wireless channels [1]–[3]. RISs are currently receiving a lot of attention as a promising technology to improve the spectral and energy efficiency of beyond 5G (B5G)/6G wireless communication systems [4]–[12]. Due to ever-increasing demand for data rate and bandwidth shortage, modern wireless communication systems are mainly restricted by interference and thus, it is interesting to investigate the performance of RISs in interference-limited systems under realistic assumptions.

A. Related work

RIS is a new promising technology, which is expected to play a key role in next generations of wireless communi-

cation systems. The performance of RISs has been studied in various systems [1]–[12]. For instance, the authors in [4] considered the downlink of a single-cell system with a multiple-antenna base station (BS) and multiple single-antenna users in the presence of RIS. They assumed that there is no direct link between the BS and users, and proposed schemes to improve the sum rate and global EE of the system. The authors in [5] considered a multi-user single-cell multiple-input, single-output (MISO) RIS-assisted system and proposed beamforming schemes to maximize the signal-to-interference-plus-noise ratio (SINR) of users. They showed that RIS can substantially improve the system performance if the RIS components are optimized properly. The paper [6] proposed an algorithm to maximize the minimum SINR of a single-cell MISO RIS-assisted system. In [10], the authors studied the uplink of a single-user single-input, single-output (SISO) orthogonal frequency division multiplexing (OFDM) RIS-assisted with frequency-selective channels and proposed a power allocation scheme as well as a channel estimation algorithm to increase the user rates. Algorithms to improve the spectral efficiency of different MIMO RIS-assisted systems have also been proposed in [8], [9]. Finally, a power allocation algorithm to maximize the sum rate of a multi-cell SISO RIS-assisted system with non-orthogonal multiple-access (NOMA) was proposed in [12].

The aforementioned works considered ideal devices. Unfortunately, devices in practice suffer from hardware impairments (HWI), which is one of the most limiting factors in wireless communication systems [13]–[17]. HWI may drastically degrade the system performance especially when these imperfections are overlooked in the design [18]. A particularly important source of distortion in non-ideal devices is I/Q imbalance (IQI), which is modeled as a widely linear transformation of the input signal [16]–[18]. Hence, IQI makes the output signal improper and motivates us to employ improper signaling to compensate this imperfection. In a complex improper variable, the real and imaginary parts are correlated and/or have unequal powers [19].

Another main performance limitation in modern wireless communication systems is interference from other users sharing the same communication channel. It has been shown that improper Gaussian signaling (IGS) is able to improve the spectral and energy-efficiency of various interference-limited systems such as cognitive radio systems [20]–[23], multi-user interference channels (IC) [14], [18], [24]–[31], NOMA systems [32], [33], broadcast channels [7], [34], device-to-device communications [35], and energy harvesting systems [36], [37], to mention a few. Employing IGS,

*Mohammad Soleymani and Peter J. Schreier are with the Signal and System Theory Group, Universität Paderborn, Germany, <http://sst.upb.de> (emails: {mohammad.soleymani,peter.schreier}@sst.upb.de).

†Ignacio Santamaria is with the Departamento de Ingeniería de Comunicaciones, Universidad de Cantabria (email: i.santamaria@unican.es).

The work of I. Santamaria has been partly supported by the project ADELE PID2019-104958RB-C43, funded by MCIN/AEI/10.13039/501100011033.

on the one hand, decreases the undesired consequences of interference and, on the other hand, reduces the entropy of the desired signal at the receiver side. Thus, we should design the parameters of the transmit signals such that the overall performance is improved. Employing IGS has also another advantage, which is to provide more parameters to optimize and consequently, more flexibility in the design. In a zero-mean proper Gaussian signal, the real and imaginary parts are independent and identically distributed (i.i.d). However, when IGS is used the real and imaginary parts can be correlated and/or have unequal powers. In other words, the powers of the real and imaginary parts can be considered as two different optimization variables. This feature has been used for the first time in [27], where the authors showed that by using IGS the degrees of freedom (DoF) of the 3-user SISO ICs are maximized. This idea was later used in other works such as [18], [26] to propose spectral and/or energy-efficient IGS schemes. Note that proper Gaussian signaling (PGS) schemes can never outperform the optimal IGS scheme since IGS includes PGS as a special case.

B. Motivation

In RIS-assisted systems, it is expected that interference is more easily managed since RISs can shape the channels, making for example the interference links weaker. In other words, RISs give us some freedom in designing the channels, which can help us to mitigate undesired interference effects. Therefore, it would be reasonable to think that in a RIS-assisted system, the benefits provided by IGS would be minor (if there are any). The reason is that the benefits of IGS have been shown to decrease (or even vanish) as the number of resources either in time (by allowing time sharing [38]), frequency (by employing OFDM [39]) or spatial (by employing MIMO [18]) increases. This happens because users experience less interference when the number of resources increases for a fixed number of users. Thus, the following question arises: *Can RISs manage interference in a multicell MIMO network well enough so that there is no need to employ IGS?* In this paper, we answer this question in the negative, and show that IGS can still provide significant gains even in the presence of RIS. Indeed, although the use of RIS improves the system performance, the main performance bottleneck is the number of transmit/receive antennas. In other words, RIS improves the performance of both IGS and PGS schemes, but the relative performance may remain unchanged with/without RIS. Interestingly, RIS may even increase the benefits of IGS in some cases, as will be shown.

C. Contribution

In this paper, we consider a multi-cell broadcast MIMO RIS-assisted system possibly with IQI at the transmit and receive sides. To the best of our knowledge, it is the first work that considers the performance of IGS in MIMO RIS-assisted systems. Additionally, it is the first work that studies EE metrics for IGS in RIS-assisted systems. To evaluate the performance of RIS, we consider two possible implementations regarding the reflecting coefficients. In the first implementation, we assume that the amplitude and the phase of each reflecting component can be independently optimized [10]. Although this may not be a realistic assumption, it can

give us an upper bound for the performance of RISs [1]. Second, we consider a more realistic RIS implementation in which the amplitudes of their reflecting elements are fixed, and only the phases can be optimized, similar to [1], [2], [5]–[9]. In both scenarios, we consider treating interference as noise (TIN) as the decoding strategy and propose IGS schemes to maximize different utility functions such as minimum weighted rate, weighted sum rate, global EE and minimum weighted EE. Note that TIN is a simple, but practical, decoding scheme, which is shown to be optimal in terms of generalized DoF (GDoF) when the desired signals are strong enough at receivers [40].

In order to develop IGS schemes, we have to tackle complicated non-convex optimization problems. To this end, we employ a generalized framework based on majorization-minimization (MM) and alternating optimization, which conducts a separate optimization of the transmit covariance matrices and the reflecting coefficients. This framework belongs to the family of MM algorithms and converges to a stationary point of the considered problems when the amplitude and phase of each RIS component can be optimized separately. Additionally, the framework can be applied to every optimization problem in which the objective and/or constraints are linear functions of achievable rates. In the numerical results, we show that RIS may provide negligible gains if the reflecting coefficients are not properly optimized. This suggests the importance of the generalized framework to fully exploit the RIS benefits. We also show that the reflecting surfaces, in which only the reflecting elements' phases can be optimized, perform closely to the upper-bound attainable when the amplitudes and phases can both be optimized.

In this work, we compare a centralized implementation of RIS with co-located reflecting elements versus a distributed implementation consisting of smaller RISs deployed in a larger area. We show that a distributed implementation can improve the system performance. In particular, our results show that the position of RIS components can play a key role in the system performance. The reason is that RIS may suffer from more severe path loss than a direct link since the RIS path loss is a function of the product of the distance between RIS and BS and the distance between RIS and users, rather than a sum distance [1]. Thus, RIS should be located as close as possible to users in order to provide more benefits. This result indicates that it is advisable to use a distributed RIS implementation, especially when users are not clustered or located close to each other. Additionally, we show that the position of RIS as well as the system topology can play major roles in the performance of RIS. For instance, RIS can significantly improve the performance of cell-edge users but may not provide a considerable gain for the users located close to the BSs.

The main contributions of the paper are as follows:

- We provide a general framework to solve a variety of optimization problems in RIS-assisted MIMO systems possibly transmitting improper signals. This framework is also able to account for hardware impairments such IQI.
- We propose IGS schemes to improve the spectral and energy efficiency of a multicell broadcast MIMO RIS-assisted system with IQI showing that IGS can provide considerable system gains. Our results show that IGS

can provide more gains with RIS than without RIS, specially in SISO systems.

- RIS can significantly improve the system performance, while the performance of a reflecting surface, in which only the phases can be optimized, is very close to the upper-bound performance for RIS.
- We show that a distributed RIS implementation outperforms a centralized RIS with co-located elements. Moreover, we show that the system topology and RIS position play key roles in performance.

D. Paper outline

This paper is organized as follows. In Section II, we provide some preliminaries on improper signaling and present the IQI model. We present the system model in Section III. Section IV and Section V propose algorithms to improve the spectral and energy efficiency of the considered system, respectively. Section VI presents some numerical results, and finally, Section VII concludes the paper.

II. PRELIMINARIES

In this section, we provide some preliminaries on improper signaling as well as on modeling IQI. In Section II-A, we provide the definition of improper signals and briefly describe the real decomposition method to analyze improper signals. In Section II-B, we describe the considered IQI model, which is based on the model in [16], [18].

A. Improper signaling

A zero-mean complex Gaussian scalar variable x is called proper if its complementary variance is zero $\mathbb{E}\{x^2\} = 0$, and it is called improper otherwise [19]. The concept of impropriety can be extended to random vectors. A zero-mean complex Gaussian random vector \mathbf{x} is proper if its complementary variance $\mathbb{E}\{\mathbf{x}\mathbf{x}^T\} = \mathbf{0}$. Otherwise, it is called improper [19].

There are different analytical tools to deal with impropriety. In this paper, we employ the real decomposition method since it simplifies the analysis in multiple-antenna systems, as also discussed in [18]. In the real decomposition model, every variable is written in the real domain. For instance, consider a typical point-to-point MIMO link as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

where $\mathbf{y} \in \mathbb{C}^{N_R \times 1}$ is the received signal, $\mathbf{x} \in \mathbb{C}^{N_T \times 1}$ is the transmitted signal, $\mathbf{n} \in \mathbb{C}^{N_R \times 1}$ is the additive proper Gaussian noise, and $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ is the channel matrix. The real decomposition model for the link is

$$\begin{bmatrix} \Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\} \end{bmatrix} = \begin{bmatrix} \Re\{\mathbf{H}\} & -\Im\{\mathbf{H}\} \\ \Im\{\mathbf{H}\} & \Re\{\mathbf{H}\} \end{bmatrix} \begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix} + \begin{bmatrix} \Re\{\mathbf{n}\} \\ \Im\{\mathbf{n}\} \end{bmatrix}. \quad (2)$$

Moreover, the achievable rate of the system for proper and/or improper \mathbf{x} and \mathbf{n} is [18], [41]

$$R = \frac{1}{2} \log_2 \det(\underline{\mathbf{C}}_n + \underline{\mathbf{H}}\underline{\mathbf{P}}\underline{\mathbf{H}}^T) - \frac{1}{2} \log_2 \det(\underline{\mathbf{C}}_n), \quad (3)$$

where $\underline{\mathbf{C}}_n$ is the covariance matrix of $\begin{bmatrix} \Re\{\mathbf{n}\} \\ \Im\{\mathbf{n}\} \end{bmatrix}^T$, $\underline{\mathbf{P}}$ is the covariance matrix of $\begin{bmatrix} \Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\} \end{bmatrix}^T$, and $\underline{\mathbf{H}}$ is

$$\underline{\mathbf{H}} = \begin{bmatrix} \Re\{\mathbf{H}\} & -\Im\{\mathbf{H}\} \\ \Im\{\mathbf{H}\} & \Re\{\mathbf{H}\} \end{bmatrix}. \quad (4)$$

Note that we can use similar analytical tools for improper and proper signaling when using the real decomposition method. The main difference of improper and proper signals with the real decomposition method lies in the structure of the covariance matrices and consequently in the feasibility set of their parameters. The covariance matrix of an improper Gaussian random vector in the real decomposition model can be any arbitrary symmetric and positive semi-definite covariance matrix [19]. However, the covariance matrix of a proper Gaussian signal \mathbf{x} has the following structure [18], [19]

$$\begin{aligned} \mathbf{P} &= \mathbb{E} \left\{ \begin{bmatrix} \Re\{\mathbf{x}\}^T & \Im\{\mathbf{x}\}^T \end{bmatrix}^T \begin{bmatrix} \Re\{\mathbf{x}\}^T & \Im\{\mathbf{x}\}^T \end{bmatrix} \right\} \\ &= \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{bmatrix}, \end{aligned} \quad (5)$$

where $\mathbf{A} \in \mathbb{R}^{N \times N}$ is symmetric and positive semi-definite, and $\mathbf{B} \in \mathbb{R}^{N \times N}$ is skew-symmetric, i.e., $\mathbf{B} = -\mathbf{B}^T$, which implies that its diagonal elements are zero.

B. IQI model

When there is IQI at either the transmit or the receive side, the received signal can be represented as a widely linear transformation of the transmitted signal. The widely linear transformation is actually a linear transformation of the signal and its conjugate [19]. In this paper, we employ the IQI model in [16], [18]. For the sake of completeness, we briefly describe the IQI model in the following and refer the reader to [16], [18] for more details.

Consider a MIMO system with N_t transmit antennas and N_r receive antennas with IQI at the transceivers. The received signal can be written as

$$\mathbf{y} = \mathbf{\Gamma}_{r,1}(\mathbf{H}\mathbf{x}_{tx} + \mathbf{r}) + \mathbf{\Gamma}_{r,2}(\mathbf{H}\mathbf{x}_{tx} + \mathbf{r})^*, \quad (6)$$

where \mathbf{H} is the channel matrix, \mathbf{r} is the additive proper Gaussian noise, and \mathbf{x}_{tx} is the transmitted signal. Moreover, $\mathbf{\Gamma}_{r,1} \in \mathbb{C}^{N_r \times N_r}$ and $\mathbf{\Gamma}_{r,2} \in \mathbb{C}^{N_r \times N_r}$ capture the amplitude and rotational imbalance and are given by [16]

$$\mathbf{\Gamma}_{r,1} = \frac{\mathbf{I} + \mathbf{A}_r e^{j\phi_r}}{2}, \quad \mathbf{\Gamma}_{r,2} = \mathbf{I} - \mathbf{\Gamma}_{r,1}^* = \frac{\mathbf{I} - \mathbf{A}_r e^{-j\phi_r}}{2}, \quad (7)$$

where \mathbf{I} is the identity matrix, and the matrices \mathbf{A}_r and ϕ_r are diagonal and, respectively, reflect the amplitude and phase errors of each branch at the receiver side [16]. Additionally, the transmit signal \mathbf{x}_{tx} is also a widely linear transformation of the input signal \mathbf{x} as

$$\mathbf{x}_{tx} = \mathbf{\Gamma}_{t,1}\mathbf{x} + \mathbf{\Gamma}_{t,2}\mathbf{x}^*, \quad (8)$$

where the matrices $\mathbf{\Gamma}_{t,1} \in \mathbb{C}^{N_t \times N_t}$ and $\mathbf{\Gamma}_{t,2} \in \mathbb{C}^{N_t \times N_t}$ are [16]

$$\mathbf{\Gamma}_{t,1} = \frac{\mathbf{I} + \mathbf{A}_t e^{j\phi_t}}{2}, \quad \mathbf{\Gamma}_{t,2} = \mathbf{I} - \mathbf{\Gamma}_{t,1}^* = \frac{\mathbf{I} - \mathbf{A}_t e^{-j\phi_t}}{2}, \quad (9)$$

where the matrices \mathbf{A}_t and ϕ_t are diagonal and, respectively, reflect the amplitude and phase errors of each branch at the transmitter side [16]. There is no I/Q imbalance at transmitter (receiver) if $\mathbf{\Gamma}_{t,1} = \mathbf{I}$ ($\mathbf{\Gamma}_{r,1} = \mathbf{I}$) and consequently $\mathbf{\Gamma}_{t,2} = \mathbf{0}$ ($\mathbf{\Gamma}_{r,2} = \mathbf{0}$).

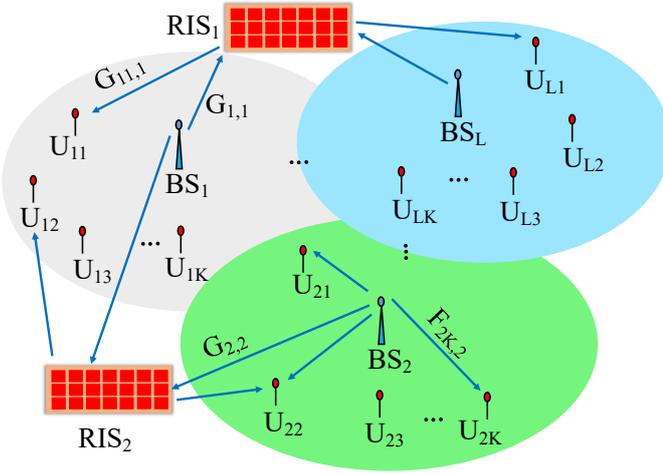


Fig. 1: The system model for a multicell broadcast system with RIS.

Lemma 1 ([18]). *The real decomposition of the aforementioned MIMO system with IQI is*

$$\underline{\mathbf{y}} = \underline{\mathbf{H}}\underline{\mathbf{x}} + \underline{\mathbf{n}}, \quad (10)$$

where $\underline{\mathbf{y}} = [\Re\{\mathbf{y}\}^T \Im\{\mathbf{y}\}^T]^T$, $\underline{\mathbf{x}} = [\Re\{\mathbf{x}\}^T \Im\{\mathbf{x}\}^T]^T$, and $\underline{\mathbf{n}} = [\Re\{\mathbf{n}\}^T \Im\{\mathbf{n}\}^T]^T$ are, respectively, the real decomposition of \mathbf{y} , \mathbf{x} , and $\mathbf{n} = \Gamma_{r,1}\mathbf{r} + \Gamma_{r,2}\mathbf{r}^*$. Moreover, $\underline{\mathbf{H}}$ is

$$\underline{\mathbf{H}} = \begin{bmatrix} \Re\{\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_2\} & -\Im\{\bar{\mathbf{H}}_1 - \bar{\mathbf{H}}_2\} \\ \Im\{\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_2\} & \Re\{\bar{\mathbf{H}}_1 - \bar{\mathbf{H}}_2\} \end{bmatrix}, \quad (11)$$

where

$$\bar{\mathbf{H}}_1 = \Gamma_{r,1}\mathbf{H}\Gamma_{t,1} + \Gamma_{r,2}\mathbf{H}^*\Gamma_{t,2}^* \in \mathbb{C}^{N_R \times N_T}, \quad (12a)$$

$$\bar{\mathbf{H}}_2 = \Gamma_{r,1}\mathbf{H}\Gamma_{t,2} + \Gamma_{r,2}\mathbf{H}^*\Gamma_{t,1}^* \in \mathbb{C}^{N_R \times N_T}. \quad (12b)$$

The statistics of the vector $\underline{\mathbf{n}} \in \mathbb{R}^{2N_r \times 1}$ are $\mathbb{E}\{\underline{\mathbf{n}}\} = \mathbf{0}$, and $\mathbb{E}\{\underline{\mathbf{n}}\underline{\mathbf{n}}^T\} = \underline{\mathbf{C}}_n = \underline{\Gamma}\underline{\mathbf{C}}_r\underline{\Gamma}^T$, where

$$\underline{\Gamma} \triangleq \begin{bmatrix} \Re\{\Gamma_{r,1} + \Gamma_{r,2}\} & -\Im\{\Gamma_{r,1} - \Gamma_{r,2}\} \\ \Im\{\Gamma_{r,1} + \Gamma_{r,2}\} & \Re\{\Gamma_{r,1} - \Gamma_{r,2}\} \end{bmatrix}, \quad (13)$$

and $\underline{\mathbf{C}}_r$ is the real decomposition of \mathbf{C}_r . For example, if $\mathbf{C}_r = \sigma^2\mathbf{I}_{N_t}$, then $\underline{\mathbf{C}}_r = \frac{1}{2}\sigma^2\mathbf{I}_{2N_t}$.

Proof. Please refer to [18, Lemma 2]. \square

III. SYSTEM MODEL

We consider the downlink of a cellular system with L base stations (BS) as show in Fig. 1. To simplify the notations, we assume that each BS has N_{BS} antennas and serves K users with N_U antennas without loss of generality. We also assume that the devices at the BSs and users are imperfect and suffer from IQI according to the model described in Sec. II-B. Obviously, it is straightforward to extend the results to a more general case in which the number of users in each cell and/or the number of transmit/receive antennas and/or the IQI parameters at the transmit/receive transceivers are different. We further assume that there are M RISs with N_{RIS} antennas each that help the BSs to serve the users. Note that the considered scenario is a standard multicell BS, which is also considered in, e.g., [8], [42]. For the ease of readers, we provide the most frequently used notations of this paper in Table I.

A. Channel model

There are two types of links between a BS and a user: a direct link and a link through the RISs. Hence, the channel between BS l and k th associated user to BS j , i.e., \mathbf{u}_{jk} , for $1 \leq l, j \leq L$ and $1 \leq k \leq K$ is [8], [9]

$$\mathbf{H}_{jk,l}(\{\Theta\}) = \underbrace{\sum_{m=1}^M \mathbf{G}_{jk,m}\Theta_m\mathbf{G}_{m,l}}_{\text{Link through RIS}} + \underbrace{\mathbf{F}_{jk,l}}_{\text{Direct link}} \in \mathbb{C}^{N_U \times N_{BS}}, \quad (14)$$

where $\mathbf{F}_{jk,l} \in \mathbb{C}^{N_U \times N_{BS}}$ is the channel matrix between the BS l and \mathbf{u}_{jk} , $\mathbf{G}_{jk,m} \in \mathbb{C}^{N_U \times N_{RIS}}$ is the channel matrix between the m th RIS and \mathbf{u}_{jk} , $\mathbf{G}_{m,l} \in \mathbb{C}^{N_{RIS} \times N_{BS}}$ is the channel matrix between the BS l and the m th RIS. Additionally, $\{\Theta\}$ is the set of $\{\Theta_m\}_{m=1}^M$, where $\Theta_m \in \mathbb{C}^{N_{RIS} \times N_{RIS}}$ is the matrix of the reflecting coefficient for the m th RIS

$$\Theta_m = \text{diag}(\theta_{m_1}, \theta_{m_2}, \dots, \theta_{m_{N_{RIS}}}), \quad (15)$$

in which θ_{m_n} for all m, n are complex-valued optimization parameters. Ideally, the amplitude and the phase of each reflecting coefficient can be treated as independent optimization variables [1], [10], [11]. In this case, the feasibility set of the reflecting coefficients is [1, Eq. (11)]

$$\mathcal{T}_I = \{\theta_{m_n} : |\theta_{m_n}|^2 \leq 1 \ \forall m, n\}. \quad (16)$$

Unfortunately, this assumption does not hold in practice [1], [2]; however, it can give us the theoretical performance limit of the RIS-assisted systems [1]. Thus, in this paper, we consider the feasibility set \mathcal{T}_I as an upper bound on performance. In addition, we consider a more practical assumption regarding the reflecting coefficients in which the amplitude of the coefficients is fixed, and we can only control the phase or equivalently the reflecting angles of each RIS component, similar to the model in [1], [2], [5]–[9]. This model leads to the following feasibility set

$$\mathcal{T}_R = \{\theta_{m_n} : |\theta_{m_n}| = 1 \ \forall m, n\}. \quad (17)$$

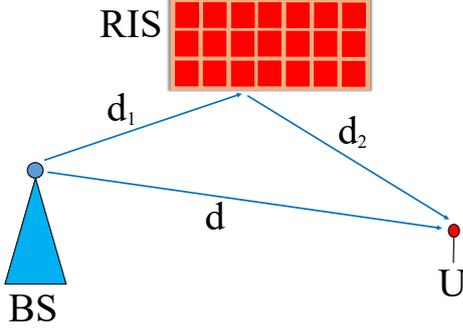
This case is also referred to as intelligent reflecting surface (IRS) to emphasize that there is only a passive phase-shifting beamforming at the intelligent surfaces [2]. It is worth emphasizing that this assumption may accurately hold if an RIS is made of discrete tiny antenna elements, which are sufficiently distant from each other [2]. Since we consider both feasibility sets \mathcal{T}_I and \mathcal{T}_R throughout this paper; hereafter, we use \mathcal{T} for the feasibility to simplify the notations when we do not refer to a specific feasibility set.

The channel matrices $\mathbf{F}_{jk,l}$, $\mathbf{G}_{jk,m}$ and $\mathbf{G}_{m,j}$ for all j, k, l, m in (14) are not controllable and, in general, depend on the path-loss, large-scale shadowing as well as small-scale fading [1]. Depending on the presence of line-of-sight (LoS) link, the small-scale fading can be modeled as Rayleigh (for non-LoS link) [4] or Rician (for LoS link) [8], [9] fading. The main difference of the link through RISs and the direct link is in the path-loss effect, where the RIS link follows the product-distance path-loss model rather than the sum-distance one [1]. That is, the channel gain is proportional to the product of the distance between the BS and RIS, d_1 , as well as distance between the RIS and user, d_2 , as

$$\beta_{RIS} \propto \frac{1}{d_1^{\alpha_1} d_2^{\alpha_2}}, \quad (18)$$

TABLE I: List of frequently used notations.

L/M	Total number of BSs/RISs
K	Total number of associated users to each BS
$N_{BS}/N_U/N_{RIS}$	Number of transmit antennas at each BS/user/RIS
\mathbf{u}_{lk}	k -th associated user to BS l
$\mathbf{x}_{lk} \in \mathbb{C}^{N_{BS} \times 1}$	Transmit signal of BS l intended for \mathbf{u}_{lk}
$\mathbf{P}_{lk} \in \mathbb{R}^{2N_{BS} \times 2N_{BS}}$	Covariance matrix of \mathbf{x}_{lk}
$\mathbf{H}_{lk,l} \in \mathbb{C}^{N_U \times N_{BS}}$	Equivalent channel between BS l and \mathbf{u}_{lk}
$\mathbf{F}_{jk,l} \in \mathbb{C}^{N_U \times N_{BS}}$	Channel matrix between the BS l and \mathbf{u}_{jk}
$\mathbf{G}_{jk,m} \in \mathbb{C}^{N_U \times N_{RIS}}$	Channel matrix between the m th RIS and \mathbf{u}_{jk}
$\mathbf{G}_{m,l} \in \mathbb{C}^{N_{RIS} \times N_{BS}}$	Channel matrix between the BS l and the m th RIS
$\mathbf{\Theta}_m \in \mathbb{C}^{N_{RIS} \times N_{RIS}}$	Matrix of the reflecting coefficient for the m th RIS
$\mathbf{D}_{lk}(\cdot) \in \mathbb{R}^{2N_U \times 2N_U}$	Interference-plus-noise covariance matrix at \mathbf{u}_{lk}
$\mathbf{n}_{lk} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$	Additive white Gaussian noise at \mathbf{u}_{lk}
E_{lk}/r_{lk}	Energy efficiency/rate of \mathbf{u}_{lk}
p_l	Power budget of BS l

**Fig. 2:** A typical RIS-assisted system.

where α_1 and α_2 are, respectively, the path-loss component for the BS-RIS and RIS-U links (see Fig. 2). The link through RIS can be much weaker than the direct link if $d_1^{\alpha_1} d_2^{\alpha_2} \gg d^\alpha$, where d and α are, respectively, the distance and path-loss component of the direct link. This implies that the RIS position can play a major role on the performance of RIS-assisted systems. When there is a LoS channel between the RISs and BSs/users the value of α_1 and α_2 are small, which makes the link through RISs stronger. Additionally, the RISs should be positioned relatively close to the users or the BS such that $d_1^{\alpha_1} d_2^{\alpha_2}$ is minimized. We will discuss the effect of the position of RIS with more details in the numerical results section.

B. Signal model

We represent the transmit signal of BS l by

$$\mathbf{x}_l = \sum_{k=1}^K \mathbf{x}_{lk} \in \mathbb{C}^{N_{BS} \times 1}, \quad (19)$$

where $\mathbf{x}_{lk} \in \mathbb{C}^{N_{BS} \times 1}$ is the transmit signal of BS l intended for its k th associated user, i.e., \mathbf{u}_{lk} , and \mathbf{x}_{lk} s are uncorrelated, i.e., $\mathbb{E}\{\mathbf{x}_{lk} \mathbf{x}_{lj}^H\} = \mathbf{0}$ for $k \neq j$. We assume that \mathbf{x}_{lk} s for all l, k are zero-mean Gaussian and can be improper. To model the impropriety, we employ the real decomposition method and represent the covariance matrix of $\mathbf{x}_{lk} = [\Re\{\mathbf{x}_{lk}\}^T \Im\{\mathbf{x}_{lk}\}^T]^T$ by $\mathbf{P}_{lk} = \mathbb{E}\{\mathbf{x}_{lk} \mathbf{x}_{lk}^T\} \in \mathbb{R}^{2N_{BS} \times 2N_{BS}}$. We define the feasibility set of the covariance matrices for IGS as

$$\mathcal{P}_{IGS} = \left\{ \{\mathbf{P}_{lk}\}_{\forall l,k} : \sum_{k=1}^K \text{Tr}(\mathbf{P}_{lk}) \leq p_l, \mathbf{P}_{lk} \succcurlyeq \mathbf{0}, \forall l, k \right\}, \quad (20)$$

where p_l is the power budget of BS l , and for PGS, as

$$\mathcal{P}_{PGS} = \left\{ \{\mathbf{P}_{lk}\}_{\forall l,k} : \sum_{k=1}^K \text{Tr}(\mathbf{P}_{lk}) \leq p_l, \mathbf{P}_{lk} = \mathbf{P}_t, \mathbf{P}_{lk} \succcurlyeq \mathbf{0}, \forall l, k \right\}, \quad (21)$$

where \mathbf{P}_t fulfills the structure in (5). Since the algorithms proposed in this paper can be applied to both improper and proper signaling schemes, the feasibility set for the covariance matrices is denoted as \mathcal{P} to simplify the notations.

Considering the channel link in (14) and the IQI model in Section II-B, the real decomposition model for the received signal at \mathbf{u}_{lk} is [8], [9]

$$\begin{aligned} \mathbf{y}_{lk} &= \sum_{i=1}^L \mathbf{H}_{lk,i}(\{\mathbf{\Theta}\}) \sum_{j=1}^K \mathbf{x}_{ij} + \mathbf{n}_{lk} \quad (22a) \\ &= \underbrace{\mathbf{H}_{lk,l}(\{\mathbf{\Theta}\}) \mathbf{x}_{lk}}_{\text{Desired signal}} + \underbrace{\mathbf{H}_{lk,l}(\{\mathbf{\Theta}\}) \sum_{j=1, j \neq k}^K \mathbf{x}_{lj}}_{\text{Intra-cell interference}} \\ &\quad + \underbrace{\sum_{i=1, i \neq l}^L \mathbf{H}_{lk,i}(\{\mathbf{\Theta}\}) \sum_{j=1}^K \mathbf{x}_{ij}}_{\text{Inter-cell interference}} + \underbrace{\mathbf{n}_{lk}}_{\text{noise}}, \quad (22b) \end{aligned}$$

where \mathbf{y}_{lk} , \mathbf{x}_{ij} , $\mathbf{H}_{lk,i}(\cdot) \in \mathbb{R}^{2N_U \times 2N_{BS}}$, and $\mathbf{n}_{lk} \in \mathbb{R}^{2N_U \times 1}$ are given by Lemma 1. We consider an additive white Gaussian noise at users with the covariance $\mathbf{C}_n = \frac{1}{2} \sigma^2 \mathbf{I}$. In this paper, due to its simplicity and usefulness we assume that each user treats the interference (both inter-cell and intra-cell) as noise, as in other works such as [8], [42].

C. Rate and energy-efficiency expressions

Treating interference as noise, the achievable rate of \mathbf{u}_{lk} is (23) and (24) on the top of the next page, where $\{\mathbf{P}\}$ is the set including all possible covariance matrices, i.e., $\{\mathbf{P}_{ij}\}_{\forall i,j}$, and $\mathbf{D}_{lk}(\cdot)$ is the interference-plus-noise covariance matrix at \mathbf{u}_{lk}

$$\begin{aligned} \mathbf{D}_{lk}(\cdot) &= \sum_{i=1, i \neq l}^L \sum_{j=1}^K \mathbf{H}_{lk,i}(\{\mathbf{\Theta}\}) \mathbf{P}_{ij} \mathbf{H}_{lk,i}^T(\{\mathbf{\Theta}\}) \\ &\quad + \sum_{j=1, j \neq k}^K \mathbf{H}_{lk,l}(\{\mathbf{\Theta}\}) \mathbf{P}_{lj} \mathbf{H}_{lk,l}^T(\{\mathbf{\Theta}\}) + \frac{\sigma^2}{2} \mathbf{I}. \quad (25) \end{aligned}$$

$$r_{lk} = \frac{1}{2} \log_2 \left| \mathbf{I} + \underline{\mathbf{H}}_{lk,l}(\{\boldsymbol{\Theta}\}) \mathbf{P}_{lk} \underline{\mathbf{H}}_{lk,l}^T(\{\boldsymbol{\Theta}\}) \mathbf{D}_{lk}^{-1}(\{\mathbf{P}\}, \{\boldsymbol{\Theta}\}) \right| \quad (23)$$

$$= \frac{1}{2} \log_2 \left| \underbrace{\underline{\mathbf{H}}_{lk,l}(\{\boldsymbol{\Theta}\}) \mathbf{P}_{lk} \underline{\mathbf{H}}_{lk,l}^T(\{\boldsymbol{\Theta}\})}_{r_{lk,1}} + \mathbf{D}_{lk}(\{\mathbf{P}\}, \{\boldsymbol{\Theta}\}) \right| - \frac{1}{2} \log_2 \left| \underbrace{\mathbf{D}_{lk}(\{\mathbf{P}\}, \{\boldsymbol{\Theta}\})}_{r_{lk,2}} \right|, \quad (24)$$

Note that the rates depend on all optimization parameters, i.e., $\{\mathbf{P}_{ij}\}_{\forall i,j}, \{\boldsymbol{\Theta}\}_{m=1}^M$. To simplify the notations, we drop this dependency and represent the rate of u_{lk} by r_{lk} .

The global EE (GEE) for the system is defined as the ratio between the sum rate and total power consumption as [43]

$$GEE = \frac{\sum_{l=1}^L \sum_{k=1}^K r_{lk}}{\sum_{l=1}^L \sum_{k=1}^K (P_c + \eta \text{Tr}(\mathbf{P}_{lk}))}, \quad (26)$$

where η^{-1} is the power transmission efficiency of each BS, and P_c is the constant power consumption for transmitting data to a user, which can be obtained as

$$P_c = \frac{MP_{RIS}}{KL} + \frac{P_{BS}}{K} + P_{UE}, \quad (27)$$

where P_{RIS} is the power consumption by an RIS, P_{BS}/P_{UE} is the constant power consumption by a BS/user. Note that, for notational simplicity, we assume the same power transmission efficiency and constant power consumption for all BS/RIS/users without loss of generality. Obviously, it can be easily extended to a more general scenario with asymmetric BS/RIS/users. The GEE is a metric to measure the overall system EE and does not consider the EE of each individual user. A metric that provides fairness among users maximizes the minimum weighted EE of users [44]. The EE of a user can be defined as the ratio between its achievable rate and the power intended for the data transmission to the user, i.e., [43], [44]

$$EE_{lk} = \frac{r_{lk}}{P_c + \eta \text{Tr}(\mathbf{P}_{lk})}. \quad (28)$$

IV. SPECTRAL EFFICIENCY OPTIMIZATION

In this paper, we consider different utility functions such as minimum weighted rate, weighted sum rate, global EE, and minimum weighted EE. In all these problems, the objective function and/or constraints are linear functions of the achievable rates. We propose a general optimization framework to solve all these problems using similar optimization tools. It would be possible to study specific solutions for any of the problems considered in this work using suboptimal or heuristic techniques, perhaps with a lower computational cost. However, in this paper we have preferred to consider a general methodology for the solution of the different problems, thus emphasizing the common aspects of the different cost functions, as well as providing a more comprehensive overview on the performance of IGS in RIS-assisted systems.

In this section, we consider the spectral efficiency metrics, i.e., the minimum-weighted and weighted-sum rate maximization. Note that the minimum [weighted] rate can be considered as a metric of fairness among the users since it is very often the case that all users have the same rate when the minimum rate is maximized [45]. Clearly, there are other fairness metrics such as the geometric mean of rates [46]. As shown in [46, Eq. (10)], the maximization of the

geometric mean of rates can be solved through a sequence of weighted-sum-rate maximization problems. Hence, the algorithms proposed in this section could be also applied to the maximization of the geometric mean of rates. However, due to the space limitation, we only consider the minimum-weighted and weighted-sum rate maximization problems.

The minimum-weighted-rate maximization (MWRM) problem can be written as

$$\max_{r, \{\boldsymbol{\Theta}\} \in \mathcal{T}, \{\mathbf{P}\} \in \mathcal{P}} r, \quad \text{s.t.} \quad \lambda_{lk} r_{lk} \geq r, \quad \forall l, k, \quad (29)$$

where $\lambda_{lk} \geq 0$ is the corresponding weight representing the priority of users. Note that the achievable rate region problem can be written as an MWRM problem by employing the rate profile technique [47]. In this case, the whole achievable rate region can be characterized by solving (29) for all possible $\sum_{\forall l} \sum_{\forall k} \lambda_{lk}^{-1} = 1$. The weighted-sum-rate maximization (WSRM) problem can be also written as

$$\max_{\{\boldsymbol{\Theta}\} \in \mathcal{T}, \{\mathbf{P}\} \in \mathcal{P}} \sum_{l=1}^L \sum_{k=1}^K \lambda_{lk} r_{lk}, \quad \text{s.t.} \quad r_{lk} \geq \bar{r}_{lk}, \quad \forall l, k, \quad (30)$$

where $r_{lk} \geq \bar{r}_{lk}$ is a quality-of-service (QoS) constraint, and \bar{r}_{lk} is the minimum required data rate for u_{lk} . Note that \bar{r}_{lk} has to be chosen such that (30) is feasible. In this paper, we propose a centralized approach to solve (30) (or (29)). This means that a central processing unit (CPU) with sufficient computational resources solves the problem and then sends, through a dedicated channel, the optimal transmission parameters as well as the optimal reflecting coefficients to the BSs and RISs, respectively. Note that all parameters are computed simultaneously, and the problem is solved only once.

Unfortunately, the MWRM and WSRM problems are not convex due to the structure of the rates as well as the feasibility sets \mathcal{T}_R . Furthermore, the joint optimization of $\{\boldsymbol{\Theta}\}, \{\mathbf{P}\}$ requires an excessive amount of computations, which might not be practical. Thus, most of the recent works on RIS employ a disjoint alternating optimization [8]–[10]. In [7], in addition to the alternating approach, a joint optimization was proposed. However, the joint optimization involved high computational complexity, and the authors applied the joint algorithm as refinement procedure after the disjoint optimization to reduce the number of computations. Hence, in this paper, we propose a disjoint alternation optimization approach with affordable computational cost.

In the alternating approach, we first solve (30) (and/or (29)) for a given $\{\boldsymbol{\Theta}^{(t-1)}\}$ and obtain $\{\mathbf{P}^{(t)}\}$. We then solve (30) over $\{\boldsymbol{\Theta}\}$ for the given $\{\mathbf{P}^{(t)}\}$ to obtain $\{\boldsymbol{\Theta}^{(t)}\}$. The convergence of this algorithm is ensured since this approach generates a non-decreasing sequence for the objective function. Unfortunately, even if we fix $\{\boldsymbol{\Theta}^{(t-1)}\}$ (or $\{\mathbf{P}^{(t)}\}$), the corresponding optimization problems are not convex, so we propose to find a suboptimal solution by employing MM.

In the following, we first consider the WSRM problem and then, modify the algorithm to solve the MWRM problem.

A. Optimization of covariance matrices

We fix $\{\Theta^{(t-1)}\}$, which simplifies the optimization problem as

$$\max_{\{\mathbf{P}\} \in \mathcal{P}} \sum_{l=1}^L \sum_{k=1}^K \lambda_{lk} r_{lk}, \quad \text{s.t.} \quad r_{lk} \geq \bar{r}_{lk}, \quad \forall l, k, \quad (31)$$

This problem is not convex since the rates are not concave in $\{\mathbf{P}\}$. Thus, we apply MM to obtain a stationary point of (31). To this end, we first find a lower-bound concave function for the user rate. As can be observed through (24), the rate of each user can be written as a difference of two concave/convex functions. This feature of the rate expressions allows us to apply convex-concave procedure (CCP) to derive a lower-bound for the rates [48]. That is, we approximate the convex part of the rate ($-r_{lk,2}$) by an affine (linear) function by employing the first-order Taylor expansion expansion given by (32) on the top of the next page, where $r_{lk,2}^{(t-1)} = r_{lk,2}(\{\Theta^{(t-1)}\}, \{\mathbf{P}^{(t-1)}\})$ and $\frac{\partial r_{lk,2}(\{\Theta^{(t-1)}\}, \{\mathbf{P}^{(t-1)}\})}{\partial \mathbf{P}_{ij}}$ is the derivative of $r_{lk,2}$ with respect to \mathbf{P}_{ij} at $\{\Theta^{(t-1)}\}, \{\mathbf{P}^{(t-1)}\}$. The derivative is

$$\frac{\partial r_{lk,2}(\{\Theta\}, \{\mathbf{P}\})}{\partial \mathbf{P}_{ij}} = \frac{1}{2 \ln 2} \mathbf{H}_{ij,i}^T \mathbf{D}_{lk}(\{\mathbf{P}\}, \{\Theta\})^{-1} \mathbf{H}_{ij,i}. \quad (33)$$

Note that we employ the first-order Taylor expansion to approximate the convex part of the rates since an affine function is the closest concave lower bound for a convex function. Replacing the lower bound $\hat{r}_{lk}^{(t-1)}$ in (31), the corresponding surrogate optimization problem is convex and can be efficiently solved by the existing numerical solvers. It is worth emphasizing that the error in the lower-bound approximation in (32) does not affect the optimality convergence of our algorithm since it falls into MM algorithms [48].

B. Optimization of the reflecting-coefficient matrix

In this subsection, we tackle the optimization of the reflecting coefficients for the two feasibility sets. To this end, we first consider feasibility set \mathcal{T}_I since it is a convex set. Note that $|\theta_{m_n}|^2$ is a convex function, and consequently, the constraint $|\theta_{m_n}|^2 \leq 1$ is a convex constraint. In this case, the WSRM problem for the given $\{\mathbf{P}^{(t)}\}$ is

$$\max_{\{\Theta\} \in \mathcal{T}_I} \sum_{l=1}^L \sum_{k=1}^K \lambda_{lk} r_{lk}(\{\Theta\}, \{\mathbf{P}^{(t)}\}), \quad (34a)$$

$$\text{s.t.} \quad r_{lk}(\{\Theta\}, \{\mathbf{P}_{lk}^{(t)}\}_{\forall l,k}) \geq \bar{r}_{lk}, \quad \forall l, k. \quad (34b)$$

This optimization problem is non-convex because of the rate functions. To solve (34), we first employ MM by obtaining suitable surrogate functions for the rates. To this end, we use the inequality in the following lemma.

Lemma 2 ([7], [36]). *Consider arbitrary matrices \mathbf{V} and $\bar{\mathbf{V}}$, and positive definite matrices \mathbf{Y} and $\bar{\mathbf{Y}}$, where all these matrices are $N \times N$. Then the following inequality holds:*

$$\begin{aligned} \ln |\mathbf{I} + \mathbf{V}\mathbf{V}^H \mathbf{Y}^{-1}| &\geq \ln |\mathbf{I} + \bar{\mathbf{V}}\bar{\mathbf{V}}^H \bar{\mathbf{Y}}^{-1}| \\ &\quad - \text{Tr}(\bar{\mathbf{V}}\bar{\mathbf{V}}^H \bar{\mathbf{Y}}^{-1}) + 2\Re \{ \text{Tr}(\bar{\mathbf{V}}^H \bar{\mathbf{Y}}^{-1} \mathbf{V}) \} \\ &\quad - \text{Tr}((\bar{\mathbf{Y}}^{-1} - (\bar{\mathbf{V}}\bar{\mathbf{V}}^H + \bar{\mathbf{Y}})^{-1})^H (\mathbf{V}\mathbf{V}^H + \mathbf{Y})) \end{aligned} \quad (35)$$

Theorem 1. *A concave lower-bound function $\hat{r}_{lk}^{(t-1)}(\{\mathbf{P}^{(t)}\}, \{\Theta\})$ for the rate of users $r_{lk}(\{\mathbf{P}^{(t)}\}, \{\Theta\})$ can be found as in (36), shown on the top of the next page, where $r_{lk}^{(t-1)} = r_{lk}(\{\mathbf{P}^{(t)}\}, \{\Theta^{(t-1)}\})$ is the rate of u_{lk} at the beginning of the step, and*

$$\mathbf{V}_{lk,l} = \mathbf{H}_{lk,l}(\{\Theta\}) \mathbf{P}_{lk}^{(t)1/2}, \quad (37)$$

$$\bar{\mathbf{V}}_{lk,l} = \mathbf{H}_{lk,l}(\{\Theta^{(t-1)}\}) \mathbf{P}_{lk}^{(t)1/2}, \quad (38)$$

$$\mathbf{Y}_{lk,l} = \mathbf{D}_{lk}(\{\mathbf{P}^{(t)}\}, \{\Theta\}), \quad (39)$$

$$\bar{\mathbf{Y}}_{lk,l} = \mathbf{D}_{lk}(\{\mathbf{P}^{(t)}\}, \{\Theta^{(t-1)}\}). \quad (40)$$

Proof. We can easily obtain (36) by considering the rate function in (23) and employing the inequality in Lemma 2. \square

Note that we can easily compute $\mathbf{P}_{lk}^{(t)1/2}$ since $\mathbf{P}_{lk}^{(t)}$ is a positive semi-definite matrix. Considering \mathcal{T}_I , $\{\Theta^{(t)}\}$ can be obtained by solving

$$\max_{\{\Theta\}} \sum_{l=1}^L \sum_{k=1}^K \lambda_{lk} \hat{r}_{lk}^{(t-1)}(\{\Theta\}, \{\mathbf{P}^{(t)}\}), \quad (41a)$$

$$\text{s.t.} \quad \hat{r}_{lk}^{(t-1)}(\{\Theta\}, \{\mathbf{P}^{(t)}\}) \geq \bar{r}_{lk}, \quad \forall l, k, \quad (41b)$$

$$|\theta_{m_n}|^2 \leq 1 \quad \forall m, n, \quad (41c)$$

where $\hat{r}_{lk}^{(t-1)}(\cdot)$ is given by Theorem 1. The optimization problem (41) is convex and can be efficiently solved. The whole algorithm for the feasibility set \mathcal{T}_I falls into MM and converges to a stationary point of the original problem.

We now consider the feasibility set \mathcal{T}_R in which $|\theta_{m_n}| = 1$. We rewriting $|\theta_{m_n}| = 1$ as the two following constraints: (41c) and

$$|\theta_{m_n}|^2 \geq 1. \quad (42)$$

As can be observed, the only difference between \mathcal{T}_I and \mathcal{T}_R is in the non-convex constraint (42). To deal with (42), we employ the convex-concave procedure (CCP) and approximate $|\theta_{m_n}|^2$ with a linear lower-bound function:

$$|\theta_{m_n}|^2 \geq 2\Re \left\{ \theta_{m_n}^{(t-1)*} \theta_{m_n} \right\} - |\theta_{m_n}^{(t-1)}|^2 \geq 1 \quad \forall m, n, \quad (43)$$

where $\theta_{m_n}^{(t-1)}$ is the value of θ_{m_n} at the previous step. To speed up the convergence of the algorithm, we relax the constraint (43) as

$$2\Re \left\{ \theta_{m_n}^{(t-1)*} \theta_{m_n} \right\} - |\theta_{m_n}^{(t-1)}|^2 \geq 1 - \epsilon \quad \forall m, n, \quad (44)$$

where ϵ is a small value. Hence, the surrogate optimization problem for \mathcal{T}_R is

$$\max_{\{\Theta\}} \sum_{l=1}^L \sum_{k=1}^K \lambda_{lk} \hat{r}_{lk}^{(t-1)}(\{\Theta\}, \{\mathbf{P}^{(t)}\}), \quad (45a)$$

$$\text{s.t.} \quad (41b), (41c), (44). \quad (45b)$$

This optimization problem is convex, and therefore can be efficiently solved. Let us call the solution of (45a) as $\{\Theta^{(*)}\}$. Since we relax (43) by introducing ϵ , $\{\Theta^{(*)}\}$ might not satisfy $|\theta_{m_n}| = 1$. Thus, we normalize $\{\Theta^{(*)}\}$ as $\{\hat{\Theta}^{(*)}\}$ to ensure that the new $\{\Theta\}$ satisfies $|\theta_{m_n}| = 1$ for all m, n . It may happen that the normalized $\{\hat{\Theta}^{(*)}\}$ does not satisfy

$$r_{lk} \geq \tilde{r}_{lk}^{(t-1)} = r_{lk,1} - r_{lk,2}^{(t-1)} - \sum_{\forall i,j} \text{Tr} \left(\frac{\partial r_{lk,2}(\{\boldsymbol{\Theta}^{(t-1)}\}, \{\mathbf{P}^{(t-1)}\})}{\partial \mathbf{P}_{ij}} (\mathbf{P}_{ij} - \mathbf{P}_{ij}^{(t-1)}) \right), \quad (32)$$

$$r_{lk}(\cdot) \geq \hat{r}_{lk}^{(t-1)}(\cdot) = r_{lk}^{(t-1)} - \frac{1}{2 \ln 2} \text{Tr} \left(\bar{\mathbf{V}}_{lk,l} \bar{\mathbf{V}}_{lk,l}^H \bar{\mathbf{Y}}_{lk,l}^{-1} \right) + \frac{1}{\ln 2} \Re \left\{ \text{Tr} \left(\bar{\mathbf{V}}_{lk,l}^H \bar{\mathbf{Y}}_{lk,l}^{-1} \mathbf{V}_{lk,l} \right) \right\} \\ - \frac{1}{2 \ln 2} \text{Tr} \left((\bar{\mathbf{Y}}_{lk,l}^{-1} - (\bar{\mathbf{V}}_{lk,l} \bar{\mathbf{V}}_{lk,l}^H + \bar{\mathbf{Y}}_{lk,l})^{-1})^H (\mathbf{V}_{lk,l} \mathbf{V}_{lk,l}^H + \mathbf{Y}_{lk,l}) \right) \quad (36)$$

Algorithm I Proposed IGS algorithm for WMRM with \mathcal{T}_I .

Initialization

Set $t = 1$, $\{\mathbf{P}\} = \{\mathbf{P}^{(t-1)}\}$, and $\{\boldsymbol{\Theta}\} = \{\boldsymbol{\Theta}^{(t-1)}\}$
While $\left(\min_{\forall l,k} r_{lk}^{(t)} - \min_{\forall l,k} r_{lk}^{(t-1)} \right) / \min_{\forall l,k} r_{lk}^{(t-1)} \geq \epsilon$
Optimizing over $\{\mathbf{P}^{(t-1)}\}$ **by fixing** $\{\boldsymbol{\Theta}\} = \{\boldsymbol{\Theta}^{(t-1)}\}$

Derive $\tilde{r}_{lk}^{(t-1)}(\{\mathbf{P}\}, \{\boldsymbol{\Theta}^{(t-1)}\})$ in (32)

Obtain $\{\mathbf{P}^{(t)}\}$ by solving (47)

Optimizing over $\{\boldsymbol{\Theta}\}$ **by fixing** $\{\mathbf{P}\} = \{\mathbf{P}^{(t-1)}\}$

Derive $\hat{r}_{lk}^{(t-1)}(\{\mathbf{P}^{(t)}\}, \{\boldsymbol{\Theta}\})$ in Theorem 1

Obtain $\{\boldsymbol{\Theta}^{(t)}\}$ by solving (48)

 $t = t + 1$
End (While)
Return $\{\mathbf{P}^*\}$ and $\{\boldsymbol{\Theta}^*\}$.

$r_{lk}(\{\mathbf{P}^{(t)}\}, \{\hat{\boldsymbol{\Theta}}^{(*)}\}) \geq r_{lk}^{(t-1)}$. To address this issue, we choose $\{\boldsymbol{\Theta}^{(t)}\}$ as

$$\{\boldsymbol{\Theta}^{(t)}\} = \begin{cases} \{\hat{\boldsymbol{\Theta}}^{(*)}\} & \text{if } r_{lk}(\{\mathbf{P}^{(t)}\}, \{\hat{\boldsymbol{\Theta}}^{(*)}\}) \geq r_{lk}^{(t-1)} \\ \{\boldsymbol{\Theta}^{(t-1)}\} & \text{Otherwise.} \end{cases} \quad (46)$$

C. Minimum-weighted-rate maximization

The approach for the WSRM can easily be applied to the MWRM problem since the WSRM and MWRM problems have a very similar structure. Indeed, we only need to replace the objective function with a new optimization variable r similar to (29). For the sake of completeness, we only mention the final surrogate optimization problems for each part. The covariance matrices $\{\mathbf{P}^{(t)}\}$ can be obtained by solving

$$\max_{r, \{\mathbf{P}\} \in \mathcal{P}} r, \quad \text{s.t.} \quad \lambda_{lk} \tilde{r}_{lk}^{(t-1)} \geq r, \quad \forall l, k, \quad (47)$$

where $\tilde{r}_{lk}^{(t-1)}$ is given by (32). Additionally, $\{\boldsymbol{\Theta}^{(t)}\}$ for \mathcal{T}_I is given by

$$\max_{r, \{\boldsymbol{\Theta}\}} r, \quad (48a)$$

$$\text{s.t.} \quad \lambda_{lk} \hat{r}_{lk}^{(t-1)}(\{\boldsymbol{\Theta}\}, \{\mathbf{P}^{(t)}\}) \geq \tilde{r}_{lk}, \quad (48b)$$

$$|\theta_{m_n}|^2 \leq 1 \quad \forall l, k, m, n, \quad (48c)$$

where $\hat{r}_{lk}^{(t-1)}(\cdot)$ is given by Theorem 1. For the ease of readers, the proposed IGS scheme for WMRM with the feasibility set \mathcal{T}_I is summarized in Algorithm I. Note that it is also straightforward to derive $\{\boldsymbol{\Theta}^{(t)}\}$ for \mathcal{T}_R as in the previous subsection.

V. ENERGY EFFICIENCY OPTIMIZATION

In this section, we consider the EE metrics for the system and solve the GEE maximization and minimum-weighted EE (MWEE) problems. The EE and GEE functions have

a fractional structure in $\{\mathbf{P}\}$. Thus, it is more complicated to optimize the EE metrics rather than spectral efficiency ones, especially when we optimize over $\{\mathbf{P}\}$. To solve these problems, we employ Dinkelbach-based algorithms as well as MM and alternating approach [14], [43], [49].

A. GEE maximization

The GEE maximization problem is

$$\max_{\{\boldsymbol{\Theta}\} \in \mathcal{T}, \{\mathbf{P}\} \in \mathcal{P}} GEE, \quad \text{s.t.} \quad r_{lk} \geq \tilde{r}_{lk}, \quad \forall l, k, \quad (49)$$

where $r_{lk} \geq \tilde{r}_{lk}$ is the QoS constraint similar to (30) and has to be chosen such that (49) is feasible. The optimization problem (49) is not convex, but we can obtain a suboptimal solution for it by an alternating optimization approach similar to the proposed algorithm in Section IV.

1) *Optimization of covariance matrices:* We first fix $\{\boldsymbol{\Theta}^{(t-1)}\}$ and maximize the GEE over $\{\mathbf{P}\}$. To this end, we employ MM and the Dinkelbach algorithm (DA) to solve the GEE maximization problem [18], [43]. That is, we first employ the concave lower-bound functions for the rates in (32), which leads to the following surrogate optimization problem

$$\max_{\{\mathbf{P}\} \in \mathcal{P}} \frac{\sum_{l=1}^L \sum_{k=1}^K \tilde{r}_{lk}^{(t-1)}}{\sum_{l=1}^L \sum_{k=1}^K (P_c + \eta \text{Tr}(\mathbf{P}_{lk}))}, \quad (50a)$$

$$\text{s.t.} \quad \tilde{r}_{lk}^{(t-1)} \geq \tilde{r}_{lk}, \quad \forall l, k, \quad (50b)$$

This optimization problem is not convex; however, we can obtain its global optimal solution by the Dinkelbach algorithm as indicated in Lemma 3.

Lemma 3. *The global optimal solution of (50) can be obtained by iteratively solving*

$$\max_{\{\mathbf{P}\} \in \mathcal{P}} \sum_{l=1}^L \sum_{k=1}^K \tilde{r}_{lk}^{(t-1)} - \mu^{(i)} \left(\sum_{l=1}^L \sum_{k=1}^K (P_c + \eta \text{Tr}(\mathbf{P}_{lk})) \right), \quad (51a)$$

$$\text{s.t.} \quad \tilde{r}_{lk}^{(t-1)} \geq \tilde{r}_{lk}, \quad \forall l, k, \quad (51b)$$

and updating $\mu^{(i)}$ as $\mu^{(i)} = \frac{\sum_{l=1}^L \sum_{k=1}^K \tilde{r}_{lk}(\mathbf{P}_{lk}^{(i-1)})}{\sum_{l=1}^L \sum_{k=1}^K (P_c + \eta \text{Tr}(\mathbf{P}_{lk}^{(i-1)}))}$, where $\mathbf{P}_{lk}^{(i-1)}$ is the solution of (51) at the $(i-1)$ th step.

2) *Optimization of the reflecting-coefficient matrix:* For a given $\{\mathbf{P}^{(t)}\}$, the GEE maximization problem has a structure similar to the WSRM and MWRM problems. Hence, we can solve it by the proposed algorithm in Section IV. That is, we first approximate the rates by the lower bound in Theorem

1. Considering the feasibility set \mathcal{T}_R , the GEE maximization is

$$\max_{\{\Theta\}} \frac{\sum_{l=1}^L \sum_{k=1}^K \lambda_{lk} \hat{r}_{lk}^{(t-1)}(\{\Theta\}, \{\mathbf{P}^{(t)}\})}{\sum_{l=1}^L \sum_{k=1}^K (P_c + \eta \text{Tr}(\mathbf{P}_{lk}^{(i-1)}))}, \quad (52a)$$

$$\text{s.t.} \quad (41c), (44), \quad (52b)$$

$$\hat{r}_{lk}^{(t-1)}(\{\Theta\}, \{\mathbf{P}^{(t)}\}) \geq \bar{r}_{lk}, \quad \forall l, k, \quad (52c)$$

which is a convex optimization problem and can be solved efficiently. Finally, $\{\Theta\}$ is updated according to (46). Note that, for feasibility set \mathcal{T}_I , $\{\Theta\}$ is updated by solving

$$\max_{\{\Theta\}} \frac{\sum_{l=1}^L \sum_{k=1}^K \lambda_{lk} \hat{r}_{lk}^{(t-1)}(\{\Theta\}, \{\mathbf{P}^{(t)}\})}{\sum_{l=1}^L \sum_{k=1}^K (P_c + \eta \text{Tr}(\mathbf{P}_{lk}^{(i-1)}))}, \quad (53a)$$

$$\text{s.t.} \quad (52c), (41c). \quad (53b)$$

B. Minimum weighted EE maximization

The MWEE maximization problem can be written as

$$\max_{\{\Theta_m\} \forall m \in \mathcal{T}, \{\mathbf{P}_{lk}\} \forall l, k \in \mathcal{P}} \min_{\forall l, k} \{\lambda_{lk} E E_{lk}\}, \quad (54a)$$

$$\text{s.t.} \quad r_{lk} \geq \bar{r}_{lk}, \quad \forall l, k. \quad (54b)$$

which is very similar to the maximization of the GEE. The main difference of these two problems is that MWEE is multiple-ratio fractional program (instead of single ratio), and we have to employ the generalized DA (GDA) for optimizing covariance matrices [14], [43], [49].

1) *Optimization of covariance matrices:* At this step, we fix $\{\Theta^{(t-1)}\}$ and solve the MWEE problem for $\{\mathbf{P}\}$ by employing the GDA and MM. That is, we first approximate the rates with the concave lower-bound in (32), which yields the following optimization problem

$$\max_{\{\mathbf{P}_{lk}\} \forall l, k \in \mathcal{P}} \min_{\forall l, k} \left\{ \lambda_{lk} \frac{\tilde{r}_{lk}^{(t-1)}}{P_c + \eta \text{Tr}(\mathbf{P}_{lk})} \right\}, \quad (55a)$$

$$\text{s.t.} \quad \tilde{r}_{lk}^{(t-1)} \geq \bar{r}_{lk}, \quad \forall l, k. \quad (55b)$$

Although this optimization problem is not convex, we can obtain its global optimal solution by employing the GDA as mentioned in the following lemma [18], [43].

Lemma 4. *The global optimal solution of (55) can be derived by iteratively solving*

$$\max_{\{\mathbf{P}_{lk}\} \forall l, k \in \mathcal{P}} \min_{\forall l, k} \left\{ \lambda_{lk} \tilde{r}_{lk}^{(t-1)} - \mu^{(i)} (P_c + \eta \text{Tr}(\mathbf{P}_{lk})) \right\}, \quad (56a)$$

$$\text{s.t.} \quad \tilde{r}_{lk}^{(t-1)} \geq \bar{r}_{lk}, \quad \forall l, k. \quad (56b)$$

and $\left\{ \lambda_{lk} \frac{\tilde{r}_{lk}^{(t-1)}(\{\Theta^{(t-1)}\}, \{\mathbf{P}^{(i-1)}\})}{P_c + \eta \text{Tr}(\mathbf{P}_{lk})} \right\}$, where $\mathbf{P}_{lk}^{(i-1)}$ is the solution of (56) at the $(i-1)$ -th step.

2) *Optimization of the reflecting-coefficient matrix:* For a given $\{\mathbf{P}^{(t)}\}$, the MWEE optimization problem can be solved similar to the GEE maximization problems. In other words, we can approximate the rates by the lower-bound in

Theorem 1, which results in the following convex optimization for the feasibility set \mathcal{T}_R

$$\max_{\{\Theta\}} \min_{\forall l, k} \left\{ \frac{\lambda_{lk} \hat{r}_{lk}^{(t-1)}(\{\Theta\}, \{\mathbf{P}^{(t)}\})}{P_c + \eta \text{Tr}(\mathbf{P}_{lk}^{(i-1)})} \right\}, \quad (57a)$$

$$\text{s.t.} \quad (41c), (44), \quad (57b)$$

$$\hat{r}_{lk}^{(t-1)}(\{\Theta\}, \{\mathbf{P}^{(t)}\}) \geq \bar{r}_{lk}, \quad \forall l, k, \quad (57c)$$

and $\{\Theta\}$ is updated according to (46). Additionally, for feasibility set \mathcal{T}_I , $\{\Theta\}$ is updated by solving the following convex problem

$$\max_{\{\Theta\}} \min_{\forall l, k} \left\{ \frac{\lambda_{lk} \hat{r}_{lk}^{(t-1)}(\{\Theta\}, \{\mathbf{P}^{(t)}\})}{P_c + \eta \text{Tr}(\mathbf{P}_{lk}^{(i-1)})} \right\}, \quad (58a)$$

$$\text{s.t.} \quad (57c), (41c). \quad (58b)$$

VI. NUMERICAL RESULTS

In this section, we provide some numerical results. We consider both large-scale and small-scale fading to accurately evaluate the performance of RISs. The large-scale path loss in dB is given by

$$PL = PL_0 + G_0 - 10\alpha \log_{10} \left(\frac{d}{d_0} \right), \quad (59)$$

where PL_0 is the path loss at the reference distance $d_0 = 1m$, d is the link distance, α is the path-loss exponent, and G_0 is the antenna gain at the transmitter side. Note that the channel attenuation coefficient is $\beta = 10^{PL/10}$. In the numerical results, we consider a two-cell MIMO broadcast channel with two users in each cell as depicted in Fig. 3. We assume that the BSs are located at $(0, 0, 25)$ and $(400, 0, 25)$, where $25m$ is the height of BSs. We further assume that the users with height $1.5m$ are uniformly located in a $20m \times 20m$ area centered at $(x, 0, 1.5)$ in cell 1 and $(400 - x, 0, 1.5)$ in cell 2. We also assume there is only one RIS with height $15m$, located at $(200, 0, 15)$ unless otherwise is explicitly mentioned. We choose the noise power density -174 dBm/Hz, channel bandwidth 1.5 MHz, $G_0 = 6$ dB for BSs and $G_0 = 5.5$ dB for RISs, $PL_0 = -30$ dB, $\alpha = 2.2$ for the links related to RISs, and $PL_0 = -35.9$ dB, $\alpha = 3.75$ for the direct links between the BSs and users. In other words, we consider a scenario that the direct links between the BSs and users is weaker than the links related to RISs, similar to, e.g., [8]. Some important parameters are summarized in Table II. We consider a Rayleigh fading as a small-scale fading for the link between the BSs and the users, which means that each elements of $\tilde{\mathbf{F}}_{jk,l}$ for all j, k, l is derived from a zero-mean complex proper Gaussian distribution with a unit variance, where $\mathbf{F}_{jk,l} = \beta_{jk,l} \tilde{\mathbf{F}}_{jk,l}$, and $\beta_{jk,l}$ is the channel attenuation coefficient (related to the large-scale path loss) for the direct link between u_{jk} and BS l . Moreover, we consider a Rician fading as a small-scale fading for the links related to the RISs. That is, we assume [8]

$$\tilde{\mathbf{G}} = \sqrt{\frac{\gamma}{1+\gamma}} \tilde{\mathbf{G}}^{\text{LoS}} + \sqrt{\frac{1}{1+\gamma}} \tilde{\mathbf{G}}^{\text{NLoS}}, \quad (60)$$

where $\gamma = 3$ is the Rician factor, $\tilde{\mathbf{G}}^{\text{LoS}}$ is the line-of-sight (LoS) component, and $\tilde{\mathbf{G}}^{\text{NLoS}}$ is the non-LoS component, which is assumed to follow a Rayleigh fading similar to $\tilde{\mathbf{F}}_{jk,l}$. We set $\gamma = 3$ similar to [8]. $\tilde{\mathbf{G}}^{\text{LoS}}$ is deterministic

TABLE II: List of Most Important Parameters.

Parameter	α_{BR}	α_{RU}	α_{BU}	$PL_{0,BR}$	$PL_{0,RU}$	$PL_{0,BU}$	G_{BS}	G_{RIS}	γ	Noise Power D.	Ch. Bw.
Value	2.2	2.2	3.75	30	30	35.9	6	5.5	3	-174 dBm/Hz	1.5 MHz

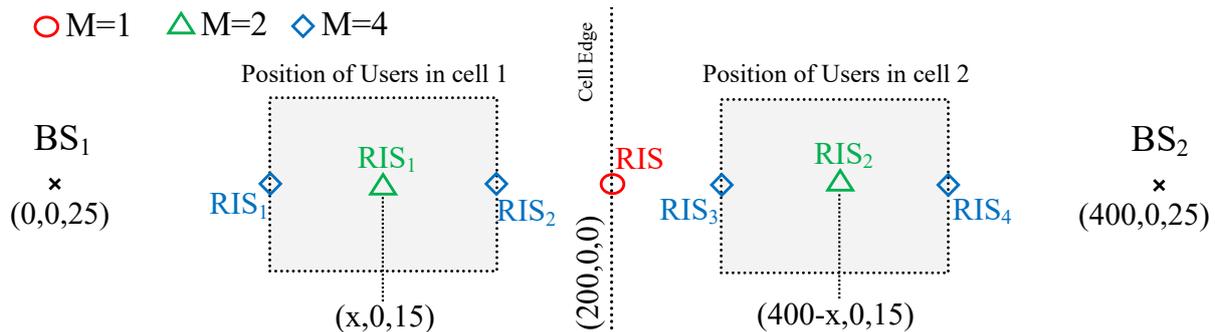


Fig. 3: System topology.

and can be derived as $\tilde{\mathbf{G}}^{\text{LoS}} = \mathbf{a}_{N_r}(\phi^A) \mathbf{a}_{N_t}^H(\phi^D)$, where $\phi^A \sim \text{Unif}[0, 2\pi]$ is angle of arrival, $\phi^D \sim \text{Unif}[0, 2\pi]$ is angle of departure, N_t/N_r is the number of transmit/receive antennas, and

$$\mathbf{a}_{N_r}(\cdot) = \left[1, e^{j\frac{2\pi d \sin(\phi^A)}{\lambda}}, \dots, e^{j\frac{2(N_r-1)\pi d \sin(\phi^A)}{\lambda}} \right], \quad (61)$$

$$\mathbf{a}_{N_t}(\cdot) = \left[1, e^{j\frac{2\pi d \sin(\phi^D)}{\lambda}}, \dots, e^{j\frac{2(N_t-1)\pi d \sin(\phi^D)}{\lambda}} \right], \quad (62)$$

where d/λ is chosen 1/2 for simplicity [8]. In the simulations, we consider an equal weight for all users, i.e., $\alpha_{lk} = 1$ for all l, k as well as an equal power budget P for all users. We average the results over 100 channel realizations. We consider the IQI parameters as in [18]. Due to space restriction, we skip the parameters and refer the readers to [18]. We will provide our simulation codes in our GitHub at <https://github.com/SSTGroup> after publishing the paper.

To the best of our knowledge, there is no other work that considers IGS with IQI in RIS-assisted MIMO systems. Hence, we compare our proposed algorithm with the proposed PGS scheme as well as a PGS scheme that does not consider IQI in the design. Moreover, we consider our algorithm with a system without RIS. To summarize, the considered algorithms in the simulations are as follows:

- **IGS-M**: The proposed IGS scheme for the feasibility set \mathcal{T}_I , which can be referred to as an upper-bound for the IGS performance.
- **IGS**: The proposed IGS scheme for the feasibility set \mathcal{T}_U .
- **IGS-N**: The IGS scheme for traditional multi-cell BCs (i.e., without RIS).
- **IGS-R**: The IGS scheme with a random reflecting coefficients.
- **PGS-M**: The proposed PGS scheme for the feasibility set \mathcal{T}_I , which can be referred to as an upper-bound for the performance of PGS schemes.
- **PGS**: The proposed PGS scheme for the feasibility set \mathcal{T}_U .
- **PGS-N**: The PGS scheme for traditional multi-cell BCs (i.e., without RIS).
- **PGS-R**: The PGS scheme with a random reflecting coefficients.
- **PGS-U**: The PGS scheme, which is unaware of IQI, with optimizing over reflecting coefficients for the feasibility set \mathcal{T}_U .

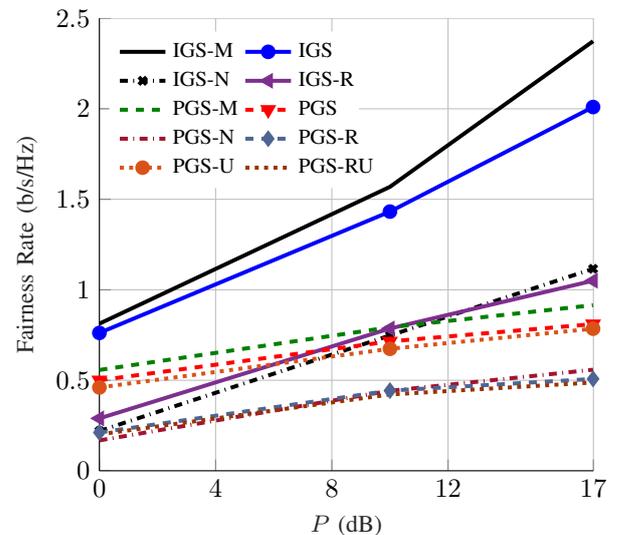
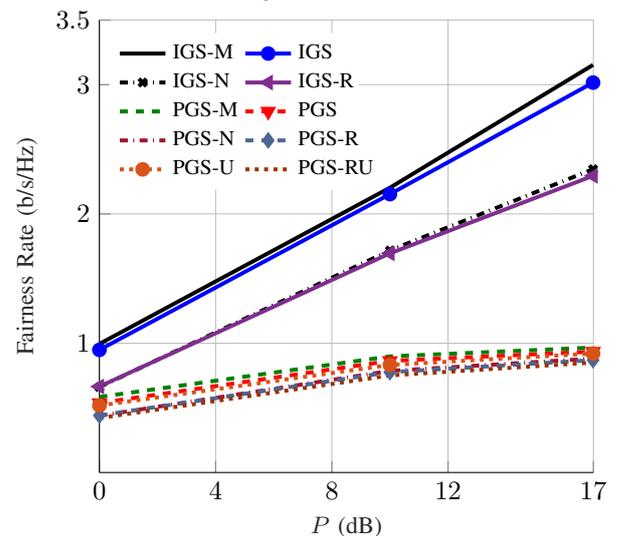
(a) Cell-edge users ($x = 180m$).(b) Middle position for users ($x = 150m$).

Fig. 4: The average fairness rate versus P for $N_{BS} = N_u = 1$, $N_{RIS} = 100$, $L = 2$, $K = 2$, $M = 1$ with different position of users.

- **PGS-RU**: The PGS scheme, which is unaware of IQI, with a random reflecting coefficients.

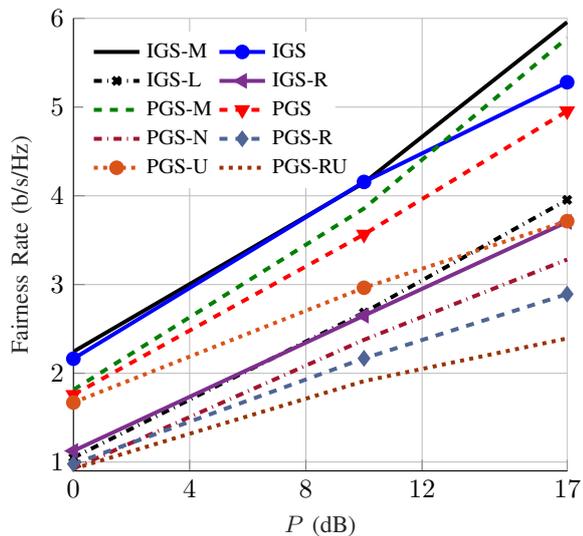
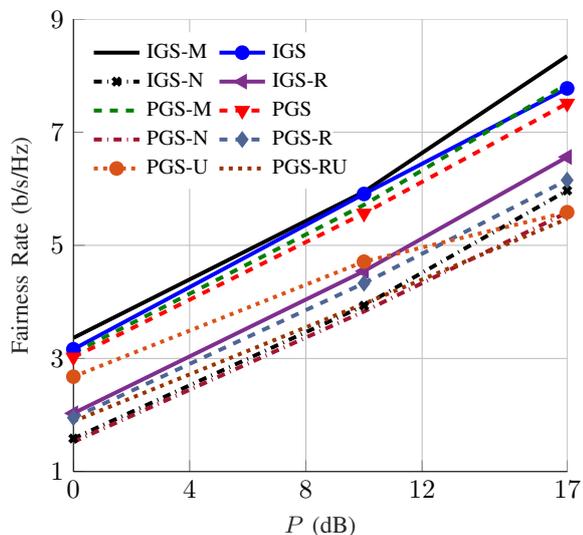
(a) $N_{BS} = N_u = 2$.(b) $N_{BS} = N_u = 3$.

Fig. 5: The average fairness rate versus P for $N_{RIS} = 100$, $L = 2$, $K = 2$, $M = 1$.

A. Fairness rate

In this subsection, we consider the WRM problem. The minimum rate of the users is also referred to as the fairness rate [18]. Figure 4 shows the average fairness rate of a two-cell system with two users in each cell for $N_{BS} = N_u = 1$, and $N_{RIS} = 100$ with different position of users. As can be observed, the IGS design can substantially outperform the PGS scheme. The reason is that the rate of PGS schemes in a broadcast channel with TIN is bounded due to the intracell interference, but IGS can efficiently manage the intracell interference. For instance, by employing maximally IGS, the BS can transmit the signal of each user in orthogonal dimensions and thus, manage the intracell interference more effectively [30].

In Fig. 4, we study how the position of users impact on the RIS performance. As can be observed, the benefits of RIS decrease when the distance between the RIS and users increases. It is expected that the benefits of RIS vanishes if the users are sufficiently close to the BS and far from RIS as discussed in Section III-A. Additionally, Fig. 4 shows that RIS may not increase the fairness rate if the reflecting

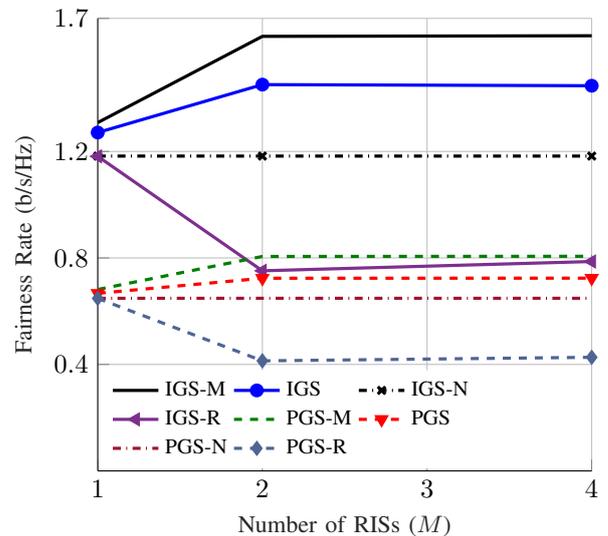
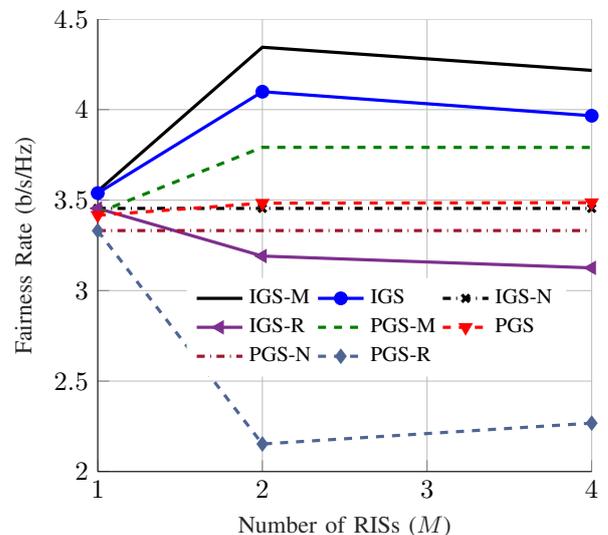
(a) $N_{BS} = N_u = 1$.(b) $N_{BS} = N_u = 2$.

Fig. 6: The average fairness rate versus the number of RISs (M) for $N_{RIS} = \frac{100}{M}$, $L = 2$, $K = 2$, $M = 1$.

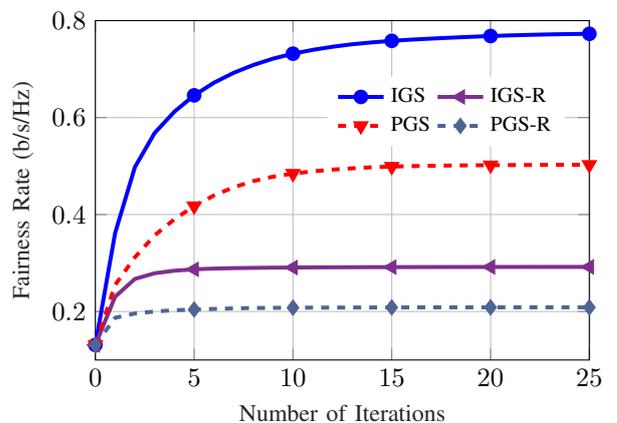


Fig. 7: The fairness rate versus number of iterations for $P=1W$, $N_{BS} = N_u = M = 1$, $N_{RIS} = 100$.

coefficients are not optimized. We can also observe that the performance of the reflecting surface is very close to the upper bound attainable when both the amplitudes and phases of the reflecting RIS elements are optimized. It means that

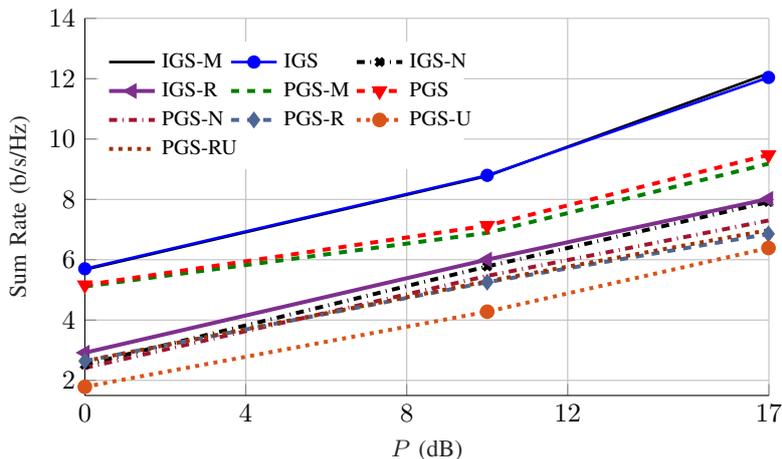


Fig. 8: The average sum rate versus P for $N_{BS} = N_u = 1$, $N_{RIS} = 100$, $L = 2$, $K = 2$, $M = 1$.

we do not lose a considerable gain when only the phases of the RIS elements are optimized.

Figure 5 shows the fairness rate of a two-cell broadcast channel with two-users in each cell for $N_{RIS} = 100$ and different number of antennas at the BSs and users. As can be observed, the benefits of employing IGS decrease with the number of antennas, which is in line with our previous study in [18] for the K -user IC. The reason is that the interference is more easily managed by PGS when the number of spatial transmit dimensions increases, and the improvement provided by IGS can be marginal when the number of resources is much greater than the number of users. We also observe that our IQI-aware algorithms can significantly outperform the PGS scheme, which is unaware of IQI. We again observe that RIS does not provide a considerable gain when the reflecting coefficients are not optimized. Additionally, the results suggest that the performance of reflecting surfaces is very close to the upper-bound performance specially at low SNRs.

In Fig. 6, we compare centralized and distributed implementations of RIS. We consider three different scenarios: a centralized RIS with $N_{RIS} = 100$, two distributed RISs with $N_{RIS} = 50$, and four distributed RISs with $N_{RIS} = 25$ as illustrated in Fig. 3. Indeed, the total number of RIS elements is fixed in these scenarios, and only the position of the elements varies. As can be observed, the average fairness rate increases when we employ distributed implementations. As indicated, the distance between RIS elements and users plays a key role in the performance of RIS. Thus, we should position the RISs as close as possible to the users, which explains the superiority of distributed implementations of RISs. Note that Fig. 6b suggests that there can be an optimal value for the number of co-located RISs, which may depend on the position of users. It means that the number and position of RISs can be considered as optimization variables, whose optimal values depend on the topology of the network as well as system parameters such as path-loss components and antenna gains.

In Fig. 7, we show the average fairness rate versus the number of iterations for a two-cell broadcast channel with two-users in each cell. As can be observed, the algorithms for RIS-assisted systems require more iterations to converge, which implies that these algorithms are slower. Note that one iteration of the proposed algorithm consists of solving two

convex optimization problems, and it is therefore computationally more costly than one iteration of the algorithms that do not optimize the reflecting coefficients. However, interestingly, we can observe that after *a single iteration* the algorithms with optimized reflecting coefficients provide a better minimum rate than the final value of the algorithms that do not optimize the reflecting coefficients. This improvement outweighs the complexity of our proposed algorithms.

B. Sum-rate maximization

Figure 8 shows the average sum rate of a two-cell broadcast channel with two-users in each cell for $N_{BS} = N_u = 1$, $N_{RIS} = 100$, and $M = 1$. In this figure, we assume that the minimum required rate of users can be 0, which means that some users may be switched off if they do not have good channel gains. Thus, the interference level can be lower, which in turn reduces the benefits of IGS as an interference-management technique. This is due to the fact that the higher the interference level is, the more benefits by IGS can be achieved as also showed in [18]. We can observe this issue in Fig. 8, where the proposed IGS scheme outperforms the PGS schemes, but the IGS benefits are lower than those observed in terms of the fairness rate. Moreover, we observe that there are only minor gains by employing RIS when the RIS components are chosen randomly and not optimized. On the contrary, RIS can provide a considerable gain when the RIS elements are optimized properly. Additionally, we observe that the reflecting surfaces perform very close to the upper-bound performance of RIS. Finally, we observe that PGS, which is unaware of IQI, performs even worse than the PGS scheme without RIS. This result shows the importance of implementing IQI-aware schemes.

C. Fairness EE

Figure 9 shows the fairness EE of a two-cell broadcast channel with two-users in each cell for $N_{RIS} = 100$ and different number of antennas at the BSs and users. As can be observed, the proposed IGS scheme outperforms PGS schemes. Moreover, there is a huge performance improvement by optimizing RIS components in both IGS and PGS schemes. Similar to WMRM and WSRM, RIS does not provide a considerable gain when the RIS components are chosen randomly. Additionally, the reflecting surface

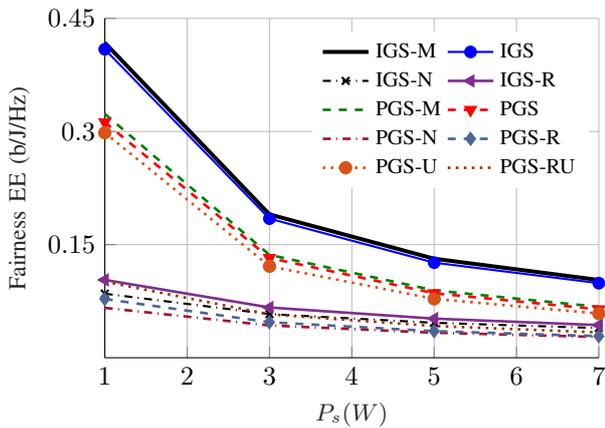
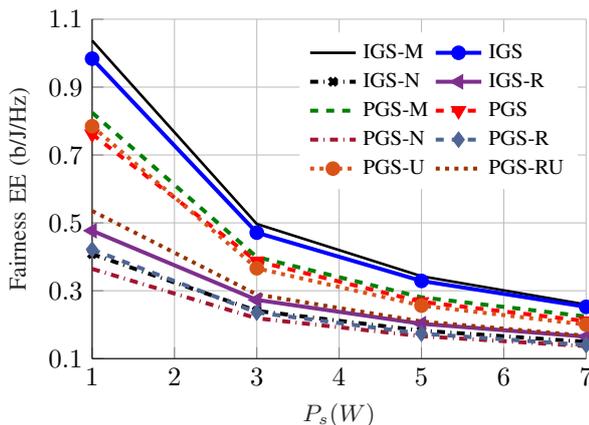
(a) $N_{BS} = N_u = 1$.(b) $N_{BS} = N_u = 2$.

Fig. 9: The average fairness EE versus P_c for $N_{RIS} = 100$, $L = 2$, $K = 2$, $M = 1$ with different N_{BS}, N_u .

performs very close to the upper-bound performance, which is given by jointly optimizing the amplitude and phase of RIS components. Finally, we observe that the IQI-aware scheme significantly outperforms the PGS scheme, which is unaware of IQI.

D. Global-EE maximization

Figure 10 shows the global EE of a two-cell broadcast channel with two users in each cell for $N_{RIS} = 100$ and different number of antennas at the BSs and users. As can be observed, we can get a considerable improvement by optimizing over the RIS components. However, the performance of RIS with random but fixed phases is very close to the case without RIS. Additionally, we observe that the IGS and PGS schemes perform very close to each other. In other words, IGS provides only minor improvements over the PGS scheme. This is in line with the results in [18] in which it was shown that the benefits of IGS from global EE perspective vanishes in the K -user MIMO ICs with HWI. Note that our IQI-aware schemes always outperform the PGS scheme, which is unaware of IQI.

E. Summary of numerical results

Our results show that IGS can substantially improve the spectral and energy-efficiency of RIS-assisted systems. It is precisely the combined use of RISs together with an interference-management technique such as IGS that makes

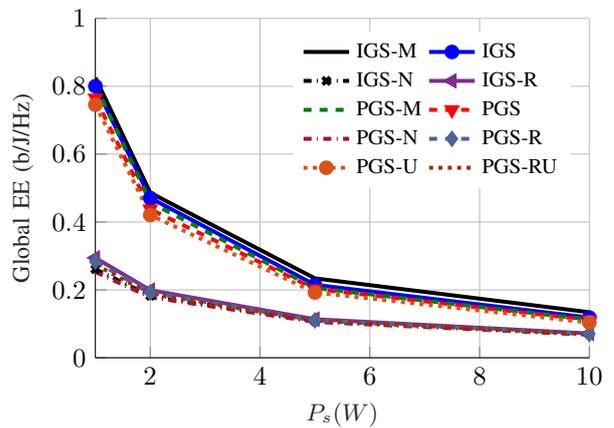
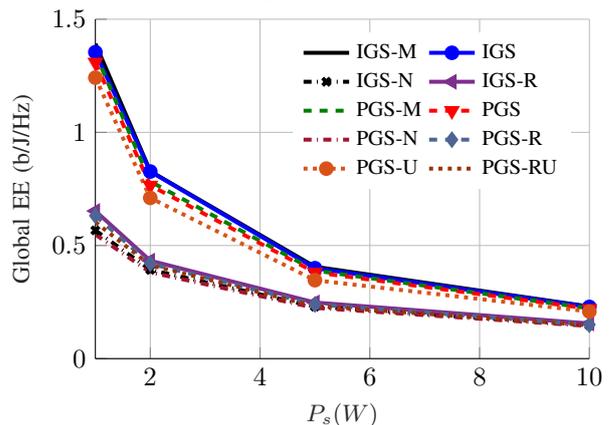
(a) $N_{BS} = N_u = 1$.(b) $N_{BS} = N_u = 2$.

Fig. 10: The average global EE versus P_c for $N_{RIS} = 100$, $L = 2$, $K = 2$, $M = 1$.

it possible to achieve satisfactory results in interference-limited multicell MIMO scenarios. Interestingly, our numerical results suggest that employing RIS substantially boosts the benefits of IGS especially in SISO systems. Finally, our results also show that a distributed implementation of RIS outperforms a centralized implementation, thus indicating that the location of the different RISs in the network should be optimized to maximize performance.

VII. CONCLUSION AND FUTURE WORK

In this paper, we proposed IGS schemes to improve the spectral and energy efficiency of a multicell RIS-assisted broadcast channel with IQI, showing the performance improvements that improper signaling schemes can bring to this scenario. Additionally, we proposed schemes to optimize the RIS elements. Our numerical results showed that RIS may provide only minor benefits if the RIS components are chosen randomly. We also showed that the benefits of RISs highly depend on their position. If the distance between the RIS and users is large, it may not offer considerable gains. Furthermore, we showed that distributed RIS implementations can bring more benefits than centralized implementations. Altogether, these results suggest that the number and position of RISs should be considered as optimization parameters. We also considered a RIS implementation for which the amplitude and phase of each RIS component can be independently optimized, which can be seen as an upper bound on the RIS performance. Our numerical results

suggest that reflecting surfaces, in which only the phases are optimized, can perform very close to the upper bound where both amplitudes and phases are optimized. Finally, we showed that neglecting IQI in the design can result in a huge performance degradation, while IQI-aware schemes significantly improve system performance.

As future lines of work, the asymptotic behavior of the spectral and energy efficiency metrics can be studied for BSs with a large number of antennas (massive MIMO). Furthermore, the performance of IGS in the presence of imperfect and/or statistical CSI is undoubtedly another aspect to be investigated. It can be also interesting to study the performance of IGS when a NOMA-based technique is employed in RIS-assisted systems. Finally, the development of distributed and/or low complexity versions of our algorithms is an interesting research avenue.

REFERENCES

- [1] Q. Wu *et al.*, "Intelligent reflecting surface aided wireless communications: A tutorial," *IEEE Trans. Commun.*, 2021.
- [2] M. Di Renzo *et al.*, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, 2020.
- [3] C. Huang, *et al.*, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 118–125, 2020.
- [4] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, 2019.
- [5] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, 2019.
- [6] Q.-U.-A. Nadeem, *et al.*, "Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7748–7764, 2020.
- [7] H. Yu *et al.*, "Joint design of reconfigurable intelligent surfaces and transmit beamforming under proper and improper Gaussian signaling," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2589–2603, 2020.
- [8] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, 2020.
- [9] L. Zhang, Y. Wang, W. Tao, Z. Jia, T. Song, and C. Pan, "Intelligent reflecting surface aided MIMO cognitive radio systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11 445–11 457, 2020.
- [10] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Trans. on Commun.*, vol. 68, no. 7, pp. 4522–4535, 2020.
- [11] M. A. ElMossallamy *et al.*, "Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 3, pp. 990–1002, 2020.
- [12] W. Ni, X. Liu, Y. Liu, H. Tian, and Y. Chen, "Resource allocation for multi-cell IRS-aided NOMA networks," *IEEE Trans. Wireless Commun.*, 2021.
- [13] E. Boshkovska, D. W. K. Ng, L. Dai, and R. Schober, "Power-efficient and secure WPCNs with hardware impairments and non-linear EH circuit," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2642–2657, 2018.
- [14] M. Soleymani, C. Lameiro, I. Santamaria, and P. J. Schreier, "Improper signaling for SISO two-user interference channels with additive asymmetric hardware distortion," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8624–8638, 2019.
- [15] S. Javed, O. Amin, B. Shihada, and M.-S. Alouini, "Improper Gaussian signaling for hardware impaired multihop full-duplex relaying systems," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 1858–1871, 2019.
- [16] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Multiple antenna systems with hardware impairments: New performance limits," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1593–1606, 2019.
- [17] A.-A. A. Boulgeorgos, N. D. Chatzidiamantis, and G. K. Karagiannidis, "Energy detection spectrum sensing under RF imperfections," *IEEE Trans. Wireless Commun.*, vol. 64, no. 7, pp. 2754–2766, 2016.
- [18] M. Soleymani, I. Santamaria, and P. J. Schreier, "Improper Gaussian signaling for the K -user MIMO interference channels with hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11 632–11 645, 2020.
- [19] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: the Theory of Improper and Noncircular Signals*. Cambridge University Press, 2010.
- [20] M. Gaafar, O. Amin, W. Abediseid, and M.-S. Alouini, "Underlay spectrum sharing techniques with in-band full-duplex systems using improper Gaussian signaling," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 235–249, 2017.
- [21] O. Amin, W. Abediseid, and M.-S. Alouini, "Overlay spectrum sharing using improper Gaussian signaling," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 1, pp. 50–62, 2017.
- [22] M. Soleymani, C. Lameiro, P. J. Schreier, and I. Santamaria, "Energy-efficient design for underlay cognitive radio using improper signaling," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Processing (ICASSP)*, 2019, pp. 4769–4773.
- [23] C. Lameiro, I. Santamaria, and P. J. Schreier, "Benefits of improper signaling for underlay cognitive radio," *IEEE Wireless Commun. Lett.*, vol. 4, no. 1, pp. 22–25, 2015.
- [24] —, "Rate region boundary of the SISO Z-interference channel with improper signaling," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1022–1034, 2017.
- [25] S. Lagen, A. Agustin, and J. Vidal, "On the superiority of improper Gaussian signaling in wireless interference MIMO scenarios," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3350–3368, 2016.
- [26] —, "Coexisting linear and widely linear transceivers in the MIMO interference channel," *IEEE Trans. Signal Process.*, vol. 64, no. 3, pp. 652–664, 2016.
- [27] V. R. Cadambe, S. A. Jafar, and C. Wang, "Interference alignment with asymmetric complex signaling—Settling the Høst-Madsen-Nosratinia conjecture," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4552–4565, 2010.
- [28] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Transmit optimization with improper Gaussian signaling for interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2899–2913, 2013.
- [29] M. Soleymani, C. Lameiro, I. Santamaria, and P. J. Schreier, "Robust improper signaling for two-user SISO interference channels," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 4709–4723, 2019.
- [30] M. Soleymani, I. Santamaria, C. Lameiro, and P. J. Schreier, "Ergodic rate for fading interference channels with proper and improper Gaussian signaling," *Entropy*, vol. 21, no. 10, p. 922, 2019.
- [31] M. Soleymani, C. Lameiro, I. Santamaria, and P. J. Schreier, "Energy-Efficient improper signaling for K -User interference channels," in *Proc. IEEE Eu. Signal Process. Conf. (EUSIPCO)*, 2019, pp. 1–5.
- [32] H. D. Tuan, A. A. Nasir, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Non-orthogonal multiple access with improper Gaussian signaling," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 496–507, 2019.
- [33] A. A. Nasir, H. D. Tuan, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Signal superposition in NOMA with proper and improper Gaussian signaling," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6537–6551, 2020.
- [34] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Improper Gaussian signaling for broadcast interference networks," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 808–812, 2019.
- [35] H. Nguyen, H. Tuan, D. Niyato, D. Kim, and H. Poor, "Improper Gaussian signaling for D2D communication coexisting MISO cellular networks," *IEEE Trans. Wireless Commun.*, 2021.
- [36] H. Yu, T. D. Hoang, A. A. Nasir, T. Q. Duong, and L. Hanzo, "Improper Gaussian signaling for computationally tractable energy and information beamforming," *IEEE Trans. Veh. Technol.*, 2020.
- [37] H. Yu, H. D. Tuan, T. Q. Duong, Y. Fang, and L. Hanzo, "Improper gaussian signaling for integrated data and energy networking," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3922–3934, 2020.
- [38] C. Hellings and W. Utschick, "Improper signaling versus time-sharing in the two-user Gaussian interference channel with TIN," *IEEE Trans. Inf. Theory*, vol. 66, no. 5, pp. 2988–2999, 2020.
- [39] M. Soleymani, C. Lameiro, P. J. Schreier, and I. Santamaria, "Improper signaling for OFDM underlay cognitive radio systems," in *2018 IEEE Statistical Signal Process. Workshop (SSP)*. IEEE, 2018, pp. 722–726.
- [40] C. Geng, N. Naderialzadeh, A. S. Avestimehr, and S. A. Jafar, "On the optimality of treating interference as noise," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1753–1767, 2015.
- [41] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.
- [42] Q. Shi *et al.*, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, 2011.
- [43] A. Zappone and E. Jorswieck, "Energy efficiency in wireless networks via fractional programming theory," *Found Trends in Commun. Inf. Theory*, vol. 11, no. 3-4, pp. 185–396, 2015.
- [44] S. Buzzi, I. Chih-Lin, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5G networks and

challenges ahead,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 697–709, 2016.

- [45] H. Park, S.-H. Park, J.-S. Kim, and I. Lee, “SINR balancing techniques in coordinated multi-cell downlink systems,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 626–635, 2013.
- [46] H. Yu, H. D. Tuan, E. Dutkiewicz, H. V. Poor, and L. Hanzo, “Maximizing the geometric mean of user-rates to improve rate-fairness: Proper vs. improper Gaussian signaling,” *IEEE Trans. Wireless Commun.*, 2021.
- [47] M. Soleymani, I. Santamaria, B. Maham, and P. J. Schreier, “Rate region of the K -user MIMO interference channel with imperfect transmitters,” in *Proc. IEEE Eu. Signal Process. Conf. (EUSIPCO)*, 2020, pp. 1–5.
- [48] 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, 2017.
- [49] J.-P. Crouzeix and J. A. Ferland, “Algorithms for generalized fractional programming,” *Math. Programm.*, vol. 52, no. 1-3, pp. 191–207, May 1991.



Mohammad Soleymani was born in Arak, Iran. He received the B.Sc. degree from Amirkabir University of Technology (Tehran Polytechnic), the M.Sc. degree from Sharif University of Technology, Tehran, Iran, and the Ph.D. degree from the University of Paderborn, Germany, all in electrical engineering. He is currently an assistant professor (Akademischer Rat) at the Signal and System Theory Group, University of Paderborn. He was a Visiting Researcher at the University of Cantabria, Spain. His research interests include multiuser

communications, wireless networking, convex optimization and statistical signal processing.



Ignacio Santamaria (M’96-SM’05) received the Telecommunication Engineer degree and the Ph.D. degree in electrical engineering from the Universidad Politecnica de Madrid (UPM), Spain, in 1991 and 1995, respectively. In 1992, he joined the Department of Communications Engineering, University of Cantabria, Spain, where he is currently Full Professor. He has co-authored more than 200 publications in refereed journals and international conference papers, and holds two patents. His current research interests include signal processing

algorithms and information-theoretic aspects of multiuser multiantenna wireless communication systems, multivariate statistical techniques and machine learning theories. He has been involved in numerous national and international research projects on these topics.

He has been a visiting researcher at the University of Florida (in 2000 and 2004), at the University of Texas at Austin (in 2009), and more recently at the Colorado State University. Prof. Santamaria was general Co-Chair of the 2012 IEEE Workshop on Machine Learning for Signal Processing (MLSP 2012). He is a member of the IEEE Signal Processing Theory and Methods Technical Committee, and Steering Committee member of the SP Data Science Initiative. In the past, he was a member of the IEEE Machine Learning for Signal Processing Technical Committee (2009-2014) and he served as Associate Editor and Senior Area Editor of the IEEE Transactions on Signal Processing (2011-2015). He was a co-recipient of the 2008 IEEE/COM Innovation Award, as well as coauthor of a paper that received the 2012 IEEE Signal Processing Society Young Author Best Paper Award.



Peter Schreier (S’03, M’04, SM’09) was born in Munich, Germany, in 1975. He received a Master of Science from the University of Notre Dame, IN, USA, in 1999, and a Ph.D. from the University of Colorado at Boulder, CO, USA, in 2003, both in electrical engineering. From 2004 until 2011, he was on the faculty of the School of Electrical Engineering and Computer Science at the University of Newcastle, NSW, Australia. Since 2011, he has been Chaired Professor of Signal and System Theory, and since Oct. 2018 he

serves as Dean of the Faculty of Electrical Engineering, Computer Science, and Mathematics of Paderborn University, Germany. He has spent sabbatical semesters at the University of Hawaii at Manoa, Honolulu, HI, and Colorado State University, Ft. Collins, CO. He is also a co-founder and the CEO of metamorphosis, which develops artificial-intelligence-based solutions for trauma surgery.

From 2008 until 2012, he was an Associate Editor of the IEEE Transactions on Signal Processing, from 2010 until 2014 a Senior Area Editor for the same Transactions, and from 2015 to 2018 an Associate Editor for the IEEE Signal Processing Letters. From 2009 until 2014, he was a member of the IEEE Technical Committee on Machine Learning for Signal Processing. He currently chairs the Steering Committee of the IEEE Signal Processing Society (SPS) Data Science Initiative and serves on the IEEE SPS Technical Committee on Signal Processing Theory and Methods, the Technical Directions Board of SPS, and the SPS Regional Committee for IEEE Region 8. He was the General Chair of the 2018 IEEE Statistical Signal Processing Workshop in Freiburg, Germany.