# Improper Gaussian Signaling for the $K$-user MIMO Interference Channels with Hardware Impairments

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

*Abstract*—This paper investigates the performance of improper Gaussian signaling (IGS) for the $K$-user multiple-input, multiple-output (MIMO) interference channel (IC) with hardware impairments (HWI). HWI may arise due to imperfections in the devices like I/Q imbalance, phase noise, etc. With I/Q imbalance, the received signal is a widely linear transformation of the transmitted signal and noise. Thus, the effective noise at the receivers becomes improper, which means that its real and imaginary parts are correlated and/or have unequal powers.

IGS can improve system performance with improper noise and/or improper interference. In this paper, we study the benefits of IGS for this scenario in terms of two performance metrics: achievable rate and energy efficiency (EE). We consider the rate region, the sum-rate, the EE region and the global EE optimization problems to fully evaluate the IGS performance. To solve these non-convex problems, we employ an optimization framework based on majorization-minimization algorithms, which allow us to obtain a stationary point of any optimization problem in which either the objective function and/or constraints are linear functions of rates. Our numerical results show that IGS can significantly improve the performance of the $K$-user MIMO IC with HWI and I/Q imbalance, where its benefits increase with the number of users, $K$, and the imbalance level, and decrease with the number of antennas.

*Index Terms*—Achievable rate region, convex/concave procedure, energy efficiency, generalized Dinkelbach algorithm, hardware impairments, improper Gaussian signaling, interference channel, majorization-minimization, MIMO systems.

## I. INTRODUCTION

Wireless communication devices are never completely ideal in practice, which can significantly degrade the system performance especially when the hardware non-idealities are not adequately modeled and accounted for the system design. In general, hardware impairments (HWI) may occur due to imperfections such as quantization noise, phase noise, amplifier nonlinearities, and I/Q imbalance [1]–[19]. In addition to hardware non-idealities, communication systems may suffer from strong interference because modern wireless communication systems are mostly interference-limited. Thus, interference-management techniques will play a key role in 5G and future generations of wireless communication systems [20]. In the last decade, the use of improper

*Mohammad Soleymani and Peter J. Schreier are with the Signal and System Theory Group, Universität Paderborn, 33098 Paderborn, Germany, http://sst.upb.de (emails: {mohammad.soleymani,peter.schreier}@sst.upb.de).

†Ignacio Santamaria is with the Department of Communications Engineering, University of Cantabria, 39005 Santander, Spain (email: i.santamaria@unican.es).

Gaussian signaling (IGS) has been proposed and extensively studied as an interference-management technique [3], [21]–[39]. The real and imaginary parts of complex improper signals are correlated and/or have unequal powers, for a full treatment of improper signals the reader is referred to [40]–[42].

### A. Related work

The impact of hardware imperfections has been studied for various wireless communications scenarios in [1]–[19]. For instance, [1]–[5] considered different interference-limited networks with single-antenna transceivers subject to additive hardware distortions (AHWD). When there is AHWD, the noise distortion power is a linear function of the signal power at the corresponding antenna [1]–[5]. In [1], the authors investigated the effect of AHWD on the performance of a dual-hop relay with both amplify-and-forward and decode-and-forward protocols, and derived closed-form expressions for the outage probabilities, as well as an upper bound for the ergodic capacity. The outage probability for a device-to-device millimeter wave communication system with complex proper AHWD was derived in [2]. However, AHWD may be, in general, improper due to I/Q imbalance[1] [3]–[5], [43]–[47]. The authors in [3] considered a relay channel with improper AHWD and maximized the achievable rate of the system by optimizing over complementary variances. In [4], the authors considered a full-duplex multihop relay channel with improper AHWD. The work in [5] considered the 2-user single-input, single-output (SISO) IC with improper AHWD and proposed two suboptimal IGS schemes to obtain the achievable rate region of the 2-user SISO IC.

Hardware non-idealities can be even more critical in multiple-antenna systems [6]–[17]. In [15], the authors studied the capacity limit and multiplexing gain of multiple-input, multiple-output (MIMO) point-to-point systems with AHWD at both transmitter and receiver sides. The papers [6]–[8] studied secure communications for massive MIMO systems with AHWD in different scenarios. The paper [9] investigated the impact of AHWD on the performance of cellular communication systems in which the base station employs a massive number of antennas. The papers [10]–[12] studied the performance of massive MIMO systems with AHWD in fading channels in different scenarios. In

---

[1] In this paper, AHWD noise refers to the model in [1], [5], [6], [8]–[10], [12], [15] in which the power of the AHWD noise is a linear function of the power of the received signal. On the other hand, HWI with I/Q imbalance refers to the model in [43], [44], where the received signal is a function of the widely linear transform of the transmitted signal and noise, and the variance of the noise is fixed and independent of signal powers.

[13], the authors investigated the system performance of a two-way massive MIMO relay channel with AHWD. In [14], the authors considered beamforming designs for a dual-hop massive MIMO amplify-and-forward relay channel in the presence of AHWD and analyzed the outage probabilities for the system.

In addition to AHWD, there might be other sources of hardware imperfections like I/Q imbalance. When I/Q imbalance occurs, the received signal can be modeled through a widely linear transformation of the transmitted signal and the aggregated noise. Hence, the received signal can be improper even if the transmitted signal and additive noise are both proper [43]. It has been shown that IGS can improve the system performance in the presence of improper noise or interference-plus-noise [3]–[5], [43]–[47]. For example, it is shown in [46] that IGS is the optimal signaling for a point-to-point single-input, multiple-output (SIMO) system with asymmetric or improper AHWD.

Improper signaling can be also used as an interference-management technique in modern wireless communications systems [3], [21]–[39]. It was shown that IGS can improve several performance metrics of different interference-limited systems. First, interference-alignment techniques transmitting IGS can increase the degrees of freedom (DoF) of different ICs, as proved in [21]–[23]. Second, IGS can provide significant gains in terms of achievable rate and/or power/energy-efficiency perspectives when treating interference as noise (TIN) is the decoding strategy [24]–[39]. IGS was considered as an interference-management technique for the first time in [21], where it was shown that IGS can increase the DoF of the 3-user IC. The papers [24]–[27] showed the superiority of IGS in the 2-user SISO IC in terms of achievable rate. A robust IGS design for the 2-user IC with imperfect channel state information (CSI) was proposed in [26]. The ergodic rate of IGS and proper Gaussian signaling (PGS) schemes in the 2-user IC with statistical CSI was studied in [27]. The works [28]–[31] investigated performance improvements by IGS in the Z-IC. In [33], the authors showed that IGS can decrease the outage probability of the secondary user (SU) in an overlay cognitive radio (CR) for a given rate target. The work [32] showed that IGS can increase the achievable rate of the SU in an underlay CR (UCR) if the power gain of the interference link is greater than a threshold. Energy-efficient designs for IGS were proposed in [34] for UCR and in [35] for the $K$-user SISO IC. In [38], the authors showed that IGS can increase the minimum achievable rate of the users in the MIMO broadcast channel. The paper [39] investigated the performance improvements by IGS in non-orthogonal multiple access systems.

The papers [21], [22], [24]–[39] studied the performance of IGS with ideal devices, which is not a realistic scenario. The papers [3]–[5], [45]–[47] consider the performance of IGS with AHWD as indicated before. However, to the best of our knowledge, there is no work on IGS in multiple-antenna interference-limited systems in the presence of HWI with I/Q imbalance.

The performance of IGS in the 2-user and/or $K$-user SISO ICs has been vastly studied, e.g., in [21], [24]–[28], [35]. However, the performance of IGS in the $K$-user MIMO IC still requires further investigation. In SISO systems, it is known that the benefits of IGS are greatly reduced when the

number of resources, e.g., time or frequency channel uses, increases. For instance, in [48], we showed that IGS does not provide a significant gain in orthogonal frequency division multiplexing (OFDM) UCR systems when the number of subcarriers grows. The same behavior is observed in [49], where PGS is proved to be optimal in the 2-user IC if coded time-sharing is allowed in which the average power consumption is constrained instead of the *instantaneous power*. Hence, it seems that increasing the number of temporal or frequency dimensions provides a more flexible power allocation for PGS, which might lead to minor improvements by IGS. In MIMO systems, the number of resources increases by allowing more antennas at the transceivers. Thus, the following questions arise: how does IGS perform in the $K$-user MIMO ICs? Is IGS still beneficial when the number of spatial dimensions (antennas) increases? In this paper, we answer these questions and analyze the performance of IGS by considering different rate and energy-efficiency metrics and solving various optimization problems.

*B. Contribution*

This paper investigates the performance of IGS in the $K$-user MIMO IC. To the best of our knowledge, this is the first work to study IGS in the $K$-user MIMO IC with HWI including I/Q imbalance. We employ the HWI model in [43] and assume non-ideal transceivers, which generate an additive proper white Gaussian noise. Moreover, we assume that the upconversion (at the transmitter side) and/or the downconversion (at the receiver side) chains suffer from I/Q imbalance, which makes the received signal a function of the widely linear transform of the transmitted signal and the aggregated noise. Thus, the aggregated noise becomes improper, which also motivates us to consider improper signals and improper signaling.

It is known that by making signals improper, we introduce structure into them by correlating their real and imaginary parts [40]–[42], which can bring benefits to the performance of interference-limited systems. Indeed, on the one hand, interference can be mitigated or suppressed more effectively at the receiver side when it has more structure. On the other hand, the differential entropy of a Gaussian signal decreases if the signal is more structured (e.g., non-circular). Improper signaling schemes can improve the overall system performance when the gain we get by receiving an improper interference at the non-intended receivers overcomes the rate loss caused by transmitting improper signals. Additionally, IGS provides more optimization parameters and hence brings more flexibility than PGS schemes to the design of interference-limited wireless communication systems. This feature can be exploited to improve the system performance. Note that IGS includes PGS as a special case, where the complementary variances are zero. Hence, PGS never outperforms the optimal IGS scheme. In this paper, we investigate whether this flexibility in design leads to performance improvement for the $K$-user MIMO IC with HWI.

Throughout this paper, we consider two main performance metrics: the achievable rate and the energy efficiency (EE). The EE of a user is defined as the ratio of its achievable rate to its total power consumption [50]. There are only a few works that study the energy-efficiency of IGS in SISO networks [34], [35], but the question of how these results

translate to the $K$-user MIMO IC remains unanswered. In this paper, we provide an answer and show that IGS can be beneficial in MIMO systems in terms of achievable rate and EE as well. Interestingly, IGS provides more relative gain in terms of achievable rate than in terms of energy efficiency. This is in agreement with our previous findings on energy-efficient IGS schemes for SISO systems [34], [35]. For instance, in [34], we derived the necessary and sufficient conditions for optimality of IGS in SISO UCR from an EE point of view and showed that these conditions are more restrictive than those obtained when the achievable rate is used as performance metric instead. Moreover, although there are some benefits for IGS in terms of global EE (defined as the ratio of the total achievable rate of the network to the total power consumption of the network) for the $K$-user MIMO IC, as we will show, these benefits may not be significant. In other words, our numerical results suggest that IGS does not provide a significant gain in terms of global EE for the $K$-user MIMO IC.

In order to analyze the performance of IGS, we consider different optimization problems such as achievable rate region, maximum sum-rate, energy-efficiency region and global energy efficiency. To solve these non-convex problems, we first formulate a general optimization problem that encompasses all performance metrics under study and then, employ a majorization-minimization framework to obtain solutions for each problem. The main idea of this framework is based on the structure of the achievable rate or energy-efficiency functions in interference-limited systems when interference is treated as noise. Specifically, the achievable rate with TIN is a difference of two concave/convex functions. We exploit this feature and employ a majorization-minimization (MM) approach to derive a stationary point[2] of every optimization problem in interference-limited systems with TIN in which the objective function and/or the constraints are linear functions of the rates.

Our numerical results show that IGS can improve the performance of the $K$-user MIMO IC with HWI. Additionally, the results show an interesting behavior of IGS as an interference-management technique. We observe that the IGS benefits decrease with the number of antennas either at the transmitter or receiver sides for a fixed number of users. This is due to the fact that interference can be managed more effectively by PGS when there are more available spatial dimensions, and consequently, IGS provides less gain as an interference-management technique in MIMO systems. Note however that with I/Q imbalance, IGS always performs better than PGS even if its benefits can be minor when the number of antennas increases. We also observe that, for a fixed number of antennas, the benefits of IGS increase when the number of users grows. The reason is that the interference level increases with $K$, and the higher the level of interference, the better the performance for IGS. Additionally, our results show that the benefits of IGS increase with the imbalance level. The more improper the noise is, the more benefits can be achieved by IGS.

The main contributions of this paper can be summarized as follows:

- We propose HWI-aware IGS schemes for the $K$-user MIMO IC. We study two general performance metrics,

---

[2]In a constrained optimization problem, a stationary point of satisfies the corresponding Karush-Kuhn-Tucker (KKT) conditions [51].

**TABLE I:** List of frequently used abbreviations.

| | |
|---|---|
| CCP | Convex-Concave Procedure |
| DCP | Difference of Convex Programming |
| EE | Energy Efficiency |
| GDA | Generalized Dinkelbach Algorithm |
| HWI | Hardware Impairment |
| IC | Interference Channel |
| IGS | Improper Gaussian Signaling |
| MIMO | Multiple-Input Multiple-Output |
| MM | Majorization Minimization |
| OFDM | Orthogonal Frequency Division Multiplexing |
| PGS | Proper Gaussian Signaling |
| QoS | Quality of Service |
| SNR | Signal to Noise Ratio |
| TIN | Treating Interference as Noise |
| UCR | Underlay Cognitive Radio |

i.e., the achievable rate and EE, and solve four different optimization problems. We derive a stationary point of the rate region, sum-rate maximization, EE region and global EE maximization problems.
- To solve these optimization problems, we cast them as a general optimization problem in which the objective function and/or the constraints are linear functions of the rates. We then apply a unified framework to obtain a stationary point of these optimization problems by majorization-minimization algorithms.
- Our results show that IGS can improve the performance of the $K$-user MIMO IC with HWI in terms of achievable rate and EE. We show that IGS provides more benefits in terms of achievable rate than in terms of energy efficiency.
- Our numerical simulations suggest that the benefit of IGS schemes increases with $K$ and with the level of impairment for a fixed number of antennas. However, IGS provides minor gains with respect to PGS when the number of antennas grows for a fixed number of users.

### C. Paper organization and notations

This paper is organized as follows. In Section II, we present some background on improper random signals, as well as the system model considered in this work. We define the optimization framework for the MIMO IC with HWI at the transceivers in Section III. We solve the corresponding optimization problems in Section IV. Finally, Section V provides some numerical examples along with an extensive discussion of them.

*Notations and abbreviations*: In this paper, matrices are denoted by bold-faced upper case letters, bold-face lower case letters denote column vectors, and scalars are denoted by light-face lower case letters. Furthermore, $\mathrm{Tr}(\mathbf{X})$ and $\det(\mathbf{X})$ denote, respectively, the trace and determinant of matrix $\mathbf{X}$. The notations $\Re\{\cdot\}$ and $\Im\{\cdot\}$ return, respectively, the real and imaginary part of $\{\cdot\}$ and can be applied to scalars, vectors and matrices. Additionally, $(\cdot)^H$, $(\cdot)^T$ and $(\cdot)^*$ denote, respectively, Hermitian, transpose and conjugate operations. We represent the $N \times N$ identity matrix by $\mathbf{I}_N$. Moreover, $\mathcal{CN}(0,1)$ denotes a proper complex Gaussian distribution with zero mean and unit variance, $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$ denotes a proper complex Gaussian vector with zero mean and covariance $\mathbf{R}$. Finally, we provide a list of the most frequently used abbreviations in Table I.
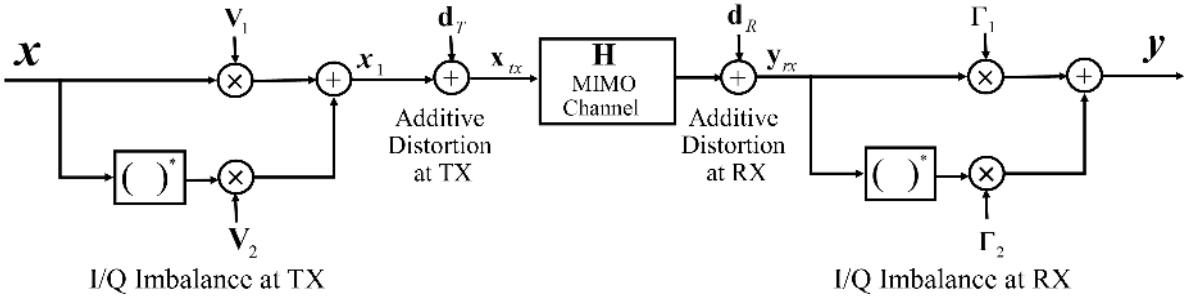
**Fig. 1:** The transceiver model of a point-to-point communications link with imperfect devices.

## II. PRELIMINARIES AND SYSTEM MODEL

We provide some preliminaries on the real-decomposition method in Section II-A, and on improper signaling in Section II-B. We then describe the HWI model in Section II-C. We finally present the considered scenario in Section II-D.

### A. Real decomposition of a complex system

Consider the following point-to-point MIMO communication system

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

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

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

Assume $\mathbf{n}$ is a random vector with Gaussian distribution as $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_n)$. The achievable rate of the system is [52]

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

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

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

### B. Preliminaries of IGS

A zero-mean complex Gaussian random variable $x$ with variance $p_t = \mathbb{E}\{|t|^2\}$ is called proper if $\mathbb{E}\{t^2\} = 0$; otherwise, it is improper [40], [41]. When the variable $t$ is improper, its real and imaginary parts are not independent and identically distributed (i.i.d). We can extend the definition of improper scalar variables to vectors. A zero-mean complex Gaussian random vector $\mathbf{t} \in \mathbb{C}^{N \times 1}$ with covariance matrix $\mathbf{P} = \mathbb{E}\{\mathbf{t}\mathbf{t}^H\}$ is called proper if $\mathbb{E}\{\mathbf{t}\mathbf{t}^T\} = \mathbf{0}$; otherwise, it is improper [40], [41].

To deal with improper signals, there are generally two approaches: augmented covariance matrix [41] and real decomposition method [53]. In the augmented-covariance-matrix approach, complex-domain signals are considered, and the optimized variables are covariance and complementary covariance matrices. However, in the real-decomposition method, every variable is written in the real domain, and the optimization variable is the covariance matrix of the real decomposition of the signals. The main differences of these two approaches are in the structure of the optimization variables as well as in the corresponding optimization

problems. That is, a complementary covariance matrix has to follow a specific structure for improper signals, while the real covariance matrices are required to be only positive semi-definite. On the one hand, the use of the augmented-covariance-matrix approach can provide insights in some problems. For example, we might be able to derive some conditions for optimality of proper or improper signaling by considering complementary variances. On the other hand, depending on the scenario, the optimization over the real domain might be simpler. For instance, in MIMO systems, the achievable rate by IGS is a complicated function of the covariance and complementary covariance matrices (please refer to [25, Eq. (10)] or [43, Section III]). Then, using the standard complex formulation, it is not possible to express the rates as a concave/convex function or a difference of two concave functions in the optimization parameters, which makes the analysis intractable. This is in contrast to the real-decomposition method by means of which the rates can be written as a difference of two concave functions, as will be shown in Section II-D. Therefore, the real-decomposition method is used to simplify the optimization problems throughout this paper.

It is worth emphasizing that, in the real decomposition model, an improper random vector can have any arbitrary symmetric and positive semi-definite covariance matrix. However, a proper Gaussian signal has a covariance matrix patterned as [40]

$$\mathbf{P} = \mathbb{E}\left\{ \begin{bmatrix} \Re\{\mathbf{t}\}^T & \Im\{\mathbf{t}\}^T \end{bmatrix}^T \begin{bmatrix} \Re\{\mathbf{t}\}^T & \Im\{\mathbf{t}\}^T \end{bmatrix} \right\}$$
$$= \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{bmatrix}, \tag{5}$$

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.

### C. HWI model for MIMO systems

In this subsection, we present the HWI model for a MIMO system with $N_T$ transmitter antennas and $N_R$ receiver antennas (see Fig. 1). We employ the non-ideal-hardware model in [43] and assume that the transceivers suffer from I/Q imbalance and generate additive distortion noise. For the sake of completeness, we briefly present the model in [43] in this subsection.

The I/Q imbalance at the transmitter side is modeled as a widely linear transformation of the transmit signal $\mathbf{x} \in \mathbb{C}^{N_T \times 1}$ as

$$\mathbf{x}_1 = \mathbf{V}_1 \mathbf{x} + \mathbf{V}_2 \mathbf{x}^*, \tag{6}$$

where the matrices $\mathbf{V}_1 \in \mathbb{C}^{N_T \times N_T}$ and $\mathbf{V}_2 \in \mathbb{C}^{N_T \times N_T}$ capture the amplitude and rotational imbalance and can be expressed as [43]

$$\mathbf{V}_1 = \frac{\mathbf{I}_{N_T} + \mathbf{A}_T e^{j\boldsymbol{\theta}_T}}{2}, \tag{7}$$

$$\mathbf{V}_2 = \mathbf{I}_{N_T} - \mathbf{V}_1^* = \frac{\mathbf{I}_{N_T} - \mathbf{A}_T e^{-j\boldsymbol{\theta}_T}}{2}. \tag{8}$$

Moreover, the matrices $\mathbf{A}_T$ and $\boldsymbol{\theta}_T$ are diagonal and, respectively, reflect the amplitude and phase errors of each branch at the transmitter side [43]. There are different methods to estimate the parameters of I/Q imbalance [54]–[57]. There is no I/Q imbalance if $\mathbf{A}_T = \mathbf{I}$ and $\boldsymbol{\theta}_T = \mathbf{0}$ or equivalently, $\mathbf{V}_1 = \mathbf{I}$ and $\mathbf{V}_2 = \mathbf{0}$.

We also assume that the transmitter is not perfect and may generate an additive proper Gaussian noise in addition to the I/Q imbalance with probability distribution $\mathbf{d}_T \in \mathbb{C}^{N_T \times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_T)$ [43]. Hence, the transmitted signal is

$$\mathbf{x}_{tx} = \mathbf{x}_1 + \mathbf{d}_T. \tag{9}$$

The transmitted signal is delivered to the receiver over a MIMO channel with additive white Gaussian noise. Hence, the received signal is

$$\mathbf{y}_{rx} = \mathbf{H}\mathbf{x}_{tx} + \mathbf{d}_R, \tag{10}$$

where the vector $\mathbf{d}_R \in \mathbb{C}^{N_R \times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_R)$ accounts for the aggregate effect of the additive white Gaussian noise of the channel and the additive distortion of the receive devices. The receiver can suffer from an I/Q imbalance similar to the transmitter. Thus, the received signal after I/Q imbalance is

$$\mathbf{y} = \boldsymbol{\Gamma}_1 \mathbf{y}_{rx} + \boldsymbol{\Gamma}_2 \mathbf{y}_{rx}^*, \tag{11}$$

where the matrices $\boldsymbol{\Gamma}_1 \in \mathbb{C}^{N_R \times N_R}$ and $\boldsymbol{\Gamma}_2 \in \mathbb{C}^{N_R \times N_R}$ are, respectively, given by

$$\boldsymbol{\Gamma}_1 = \frac{\mathbf{I}_{N_R} + \mathbf{A}_R e^{j\boldsymbol{\theta}_R}}{2} \tag{12}$$

$$\boldsymbol{\Gamma}_2 = \mathbf{I}_{N_R} - \boldsymbol{\Gamma}_1^* = \frac{\mathbf{I}_{N_R} - \mathbf{A}_R e^{-j\boldsymbol{\theta}_R}}{2}. \tag{13}$$

Similar to $\mathbf{A}_T$ and $\boldsymbol{\theta}_T$, the matrices $\mathbf{A}_R$ and $\boldsymbol{\theta}_R$ are diagonal and, respectively, reflect the amplitude and phase errors of each branch at the receiver side [43]. The following lemmas present the aggregate effect of the impairments at the transmitter and receiver sides.

**Lemma 1** ([43]). *The transceiver of a MIMO system with HWI can be modeled as*

$$\mathbf{y} = \bar{\mathbf{H}}_1 \mathbf{x} + \bar{\mathbf{H}}_2 \mathbf{x}^* + \mathbf{z}, \tag{14}$$

*where*

$$\bar{\mathbf{H}}_1 = \boldsymbol{\Gamma}_1 \mathbf{H} \mathbf{V}_1 + \boldsymbol{\Gamma}_2 \mathbf{H}^* \mathbf{V}_2^* \in \mathbb{C}^{N_R \times N_T}, \tag{15}$$

$$\bar{\mathbf{H}}_2 = \boldsymbol{\Gamma}_1 \mathbf{H} \mathbf{V}_2 + \boldsymbol{\Gamma}_2 \mathbf{H}^* \mathbf{V}_1^* \in \mathbb{C}^{N_R \times N_T}, \tag{16}$$

$$\mathbf{z} = \boldsymbol{\Gamma}_1 (\mathbf{H}\mathbf{d}_T + \mathbf{d}_R) + \boldsymbol{\Gamma}_2 (\mathbf{H}\mathbf{d}_T + \mathbf{d}_R)^* \in \mathbb{C}^{N_R \times 1}. \tag{17}$$

*Proof.* Please refer to [43, Lemma 1]. $\square$

**Lemma 2.** *The real decomposition of the MIMO system with HWI in Lemma 1 is*

$$\underline{\mathbf{y}} = \tilde{\mathbf{H}}\underline{\mathbf{x}} + \underline{\mathbf{z}}, \tag{18}$$

*where* $\underline{\mathbf{y}} = \begin{bmatrix} \Re\{\mathbf{y}\}^T & \Im\{\mathbf{y}\}^T \end{bmatrix}^T$, $\underline{\mathbf{x}} = \begin{bmatrix} \Re\{\mathbf{x}\}^T & \Im\{\mathbf{x}\}^T \end{bmatrix}^T$, *and* $\underline{\mathbf{z}} = \begin{bmatrix} \Re\{\mathbf{z}\}^T & \Im\{\mathbf{z}\}^T \end{bmatrix}^T$ *are,*
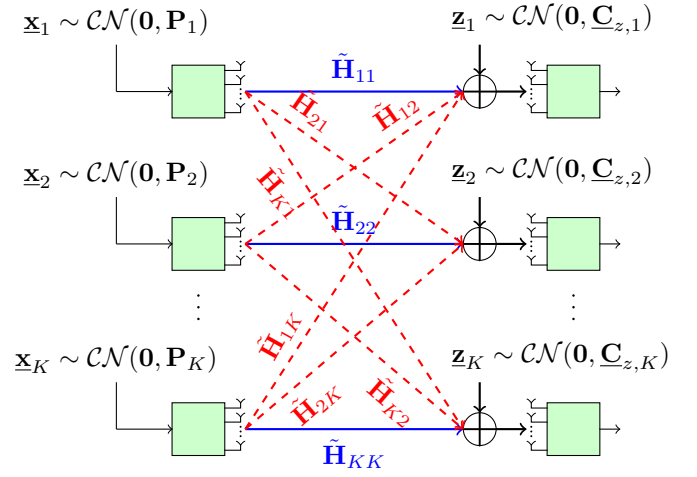


**Fig. 2:** The equivalent real-decomposition channel model for the $K$-user MIMO IC.

*respectively, the real decomposition of* $\mathbf{y}$, $\mathbf{x}$, *and* $\mathbf{z}$ *in* (14). *Moreover,* $\tilde{\mathbf{H}}$ *is*

$$\tilde{\mathbf{H}} = \begin{bmatrix} \Re\{\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_2\} & -\Im\{\bar{\mathbf{H}}_1 - \bar{\mathbf{H}}_2\} \\ \Im\{\bar{\mathbf{H}}_1 + \bar{\mathbf{H}}_2\} & \Re\{\bar{\mathbf{H}}_1 - \bar{\mathbf{H}}_2\} \end{bmatrix}. \tag{19}$$

*The statistics of the vector* $\underline{\mathbf{z}} \in \mathbb{R}^{2N_R \times 1}$ *are* $\mathbb{E}\{\underline{\mathbf{z}}\} = \mathbf{0}$, *and*

$$\mathbb{E}\{\underline{\mathbf{z}}\,\underline{\mathbf{z}}^T\} = \underline{\mathbf{C}}_z = \underline{\boldsymbol{\Gamma}}\mathbf{C}_d \underline{\boldsymbol{\Gamma}}^T, \tag{20}$$

*where* $\mathbf{C}_d = \underline{\mathbf{H}}\,\underline{\mathbf{C}}_T \underline{\mathbf{H}}^T + \underline{\mathbf{C}}_R$, *and*

$$\underline{\boldsymbol{\Gamma}} \triangleq \begin{bmatrix} \Re\{\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2\} & -\Im\{\boldsymbol{\Gamma}_1 - \boldsymbol{\Gamma}_2\} \\ \Im\{\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2\} & \Re\{\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2\} \end{bmatrix}. \tag{21}$$

*Additionally,* $\underline{\mathbf{H}}$, $\underline{\mathbf{C}}_T$, *and* $\underline{\mathbf{C}}_R$ *are, respectively, the real decomposition of* $\mathbf{H}$, $\mathbf{C}_T$, *and* $\mathbf{C}_R$. *For example, if* $\mathbf{C}_T = \sigma^2 \mathbf{I}_{N_T}$, *then* $\underline{\mathbf{C}}_T = \frac{1}{2}\sigma^2 \mathbf{I}_{2N_T}$.

*Proof.* We can easily construct the real decomposition model in (18) from the complex model in (14). Now we would like to derive the statistics of $\underline{\mathbf{z}}$ in (18). To this end, we first write the real decomposition of $\mathbf{z}$ in (17) as

$$\begin{bmatrix} \Re\{\mathbf{z}\} \\ \Im\{\mathbf{z}\} \end{bmatrix} = \begin{bmatrix} \Re\{\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2\} & -\Im\{\boldsymbol{\Gamma}_1 - \boldsymbol{\Gamma}_2\} \\ \Im\{\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2\} & \Re\{\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2\} \end{bmatrix}$$
$$\times \begin{bmatrix} \Re\{\mathbf{H}\mathbf{d}_T + \mathbf{d}_R\} \\ \Im\{\mathbf{H}\mathbf{d}_T + \mathbf{d}_R\} \end{bmatrix}, \tag{22}$$

which can be represented as $\underline{\mathbf{z}} = \underline{\boldsymbol{\Gamma}}\,(\underline{\mathbf{H}}\,\underline{\mathbf{d}}_T + \underline{\mathbf{d}}_R)$, where $\underline{\mathbf{d}}_T$ and $\underline{\mathbf{d}}_R$ are, respectively, the real decomposition of $\mathbf{d}_T$ and $\mathbf{d}_R$. The average of $\underline{\mathbf{z}}$ is simply $\mathbf{0}$, since $\mathbf{d}_R$ and $\mathbf{d}_T$ are zero-mean random vectors. Furthermore, the covariance matrix of $\underline{\mathbf{z}}$ can be derived as (20). $\square$

*D. System model*

We consider a $K$-user MIMO IC with imperfect transceivers, as shown in Fig. 2. Without loss of generality, we assume that the transceivers have the same number of antennas and produce a noise with the same statistics to simplify the notation and the expressions. Obviously, it is very straightforward to extend this model to the most general case with asymmetric devices. According to Lemma 2, the real decomposition of the received signal at the receiver of user $k$ is

$$\underline{\mathbf{y}}_k = \sum_{i=1}^{K} \tilde{\mathbf{H}}_{ki}\underline{\mathbf{x}}_i + \underline{\mathbf{z}}_k \tag{23}$$

where $\underline{\mathbf{x}}_i$ is the real decomposition of the transmitted signal of user $i$, and

$$\tilde{\mathbf{H}}_{ki} = \begin{bmatrix} \Re\{\bar{\mathbf{H}}_{1,ki} + \bar{\mathbf{H}}_{2,ki}\} & -\Im\{\bar{\mathbf{H}}_{1,ki} - \bar{\mathbf{H}}_{2,ki}\} \\ \Im\{\bar{\mathbf{H}}_{1,ki} + \bar{\mathbf{H}}_{2,ki}\} & \Re\{\bar{\mathbf{H}}_{1,ki} - \bar{\mathbf{H}}_{2,ki}\} \end{bmatrix},$$ (24)

where $\bar{\mathbf{H}}_{1,ki}$ and $\bar{\mathbf{H}}_{2,ki}$ can be derived, respectively, by replacing $\mathbf{H}_{ki}$ in (15) and (16). Note that $\mathbf{H}_{ki}$ is the channel matrix for the link between transmitter $i$ and receiver $k$. Moreover, $\underline{\mathbf{z}}_k$ is the real decomposition of the noise vector $\mathbf{z}_k$, which is given by

$$\mathbf{z}_k = \mathbf{\Gamma}_1 \left( \sum_{i=1}^{K} \mathbf{H}_{ki} \mathbf{d}_{T,i} + \mathbf{d}_{R,k} \right) + \mathbf{\Gamma}_2 \left( \sum_{i=1}^{K} \mathbf{H}_{ki} \mathbf{d}_{T,i} + \mathbf{d}_{R,k} \right)^*.$$ (25)

According to Lemma 2, the covariance matrix of $\underline{\mathbf{z}}_k$ is

$$\underline{\mathbf{C}}_{z,k} = \underline{\mathbf{\Gamma}} \left( \sum_{i=1}^{K} \underline{\mathbf{H}}_{ki} \underline{\mathbf{C}}_T \underline{\mathbf{H}}_{ki}^T + \underline{\mathbf{C}}_R \right) \underline{\mathbf{\Gamma}}^T,$$ (26)

where $\underline{\mathbf{H}}_{ki}$ is the real decomposition of $\mathbf{H}_{ki}$, and $\underline{\mathbf{\Gamma}}$ is given by (21). Treating interference as noise, we can derive the rate of user $k \in \{1, 2, ..., K\}$ as [40], [52], [58]

$$R_k = \underbrace{\frac{1}{2} \log_2 \det \left( \underline{\mathbf{C}}_{z,k} + \sum_{i=1}^{K} \tilde{\mathbf{H}}_{ki} \mathbf{P}_i \tilde{\mathbf{H}}_{ki}^T \right)}_{\triangleq r_{k,1}}$$
$$- \underbrace{\frac{1}{2} \log_2 \det \left( \underline{\mathbf{C}}_{z,k} + \sum_{i=1, i\neq k}^{K} \tilde{\mathbf{H}}_{ki} \mathbf{P}_i \tilde{\mathbf{H}}_{ki}^T \right)}_{\triangleq r_{k,2}}.$$ (27)

As can be observed through (27), the rate of user $k$ is a difference of two concave functions, i.e., $R_k = r_{k,1} - r_{k,2}$, where $r_{k,1}$ and $r_{k,2}$ are concave. This feature allows us to employ MM and convex/concave procedure (CCP) for optimization problems in which the objective and/or constraints are linear functions of the rates as will be shown in Section III and Section IV.

## III. OPTIMIZATION FRAMEWORK FOR MIMO SYSTEMS BASED ON MM

In this section, we present a framework based on MM to solve a family of optimization problems in which either the objective function and/or the constraints are linear functions of the rates. In this approach, we exploit the fact that the rate is a difference of two concave functions and solve the corresponding optimization problem iteratively. To this end, we apply the CCP to the rates and approximate the convex part of the rates, $-r_{k,2}$, by a linear function through a first-order Taylor expansion.

This framework can be applied to both IGS and PGS schemes. The only difference of IGS and PGS schemes in the framework is the feasibility set of the covariance matrices. As indicated in Section II-B, an improper Gaussian random variable can have an arbitrary symmetric and positive semi-definite covariance matrix. Thus, the feasibility set of the covariance matrices of users $\{\mathbf{P}_k\}_{k=1}^{K}$ for IGS is

$$\mathcal{P}_{IGS} = \left\{ \{\mathbf{P}_k\}_{k=1}^{K} : \mathrm{Tr}(\mathbf{P}_k) \leq P_k, \ \mathbf{P}_k \succcurlyeq \mathbf{0}, \forall k \right\},$$ (28)

where $P_k$ is the power budget of user $k$. It is in contrast with a proper Gaussian signal, which has a covariance matrix with

the specific structure in (5). In this case, the feasibility set is

$$\mathcal{P}_{PGS} = \left\{ \{\mathbf{P}_k\}_{k=1}^{K} : \mathrm{Tr}(\mathbf{P}_k) \leq P_k, \mathbf{P}_k = \mathbf{P}_{x_k}, \mathbf{P}_k \succcurlyeq \mathbf{0}, \forall k \right\},$$ (29)

where $\mathbf{P}_{x_k}$ has the structure in (5). In order to include both IGS and PGS schemes in the derivations to follow, we denote the feasibility set of the covariance matrices as $\mathcal{P}$ hereafter.

Consider the following optimization problem

$$\max_{\{\mathbf{P}_k\}_{k=1}^{K} \in \mathcal{P}} \quad f_0 \left( \{\mathbf{P}_k\}_{k=1}^{K} \right)$$ (30a)

$$\text{s.t.} \quad f_i \left( \{\mathbf{P}_k\}_{k=1}^{K} \right) \geq 0, \quad i = 1, 2, ..., I.$$ (30b)

If $f_0(\cdot)$ and $f_i(\cdot)$ for $i = 1, 2, ..., I$ are concave, the optimization problem (30) is known to be convex[3] and can be solved in polynomial time. If $f_0(\cdot)$ and $f_i(\cdot)$ for $i = 1, 2, ..., I$ are neither concave nor pseudo-concave, it is not straightforward to derive the global optimal solution of (30) in polynomial time [59]–[62]. A way to solve non-convex optimization problems is to employ iterative optimization algorithms such as MM. The MM algorithm consists of two steps at each iteration: majorization and minimization. In the majorization step, the functions $f_0(\cdot)$ and $f_i(\cdot)$ for $i = 1, 2, ..., I$ are approximated by surrogate functions. Then, the corresponding surrogate problem is solved in the minimization step. In the following lemma, we present convergence conditions of MM iterative algorithms.

**Lemma 3** ([59]). *Let us define $\tilde{f}_i^{(l)}(\cdot)$ for $l \in \mathbb{N}$ as surrogate functions of $f_i(\cdot)$ for $i = 0, 1, 2, ..., I$ such that the following conditions are fulfilled:*

- $\tilde{f}_i^{(l)} \left( \{\mathbf{P}_k^{(l)}\}_{k=1}^{K} \right) = f_i \left( \{\mathbf{P}_k^{(l)}\}_{k=1}^{K} \right)$ *for* $i = 0, 1, 2, \cdots, I$.
- $\frac{\partial \tilde{f}_i^{(l)}\left(\{\mathbf{P}_k^{(l)}\}_{k=1}^{K}\right)}{\partial \mathbf{P}_k} = \frac{\partial f_i\left(\{\mathbf{P}_k^{(l)}\}_{k=1}^{K}\right)}{\partial \mathbf{P}_k}$ *for* $i = 0, 1, 2, \cdots, I$ *and* $k = 1, 2, \cdots, K$.
- $\tilde{f}_i^{(l)}(\cdot) \leq f_i(\cdot)$ *for* $i = 0, 1, 2, \cdots, I$ *for all feasible* $\{\mathbf{P}_k\}_{k=1}^{K}$,

*where $\{\mathbf{P}_k^{(l)}\}_{k=1}^{K}$ is the initial point at the l-th iteration of the MM algorithm, which is obtained by solving*

$$\max_{\{\mathbf{P}_k\}_{k=1}^{K} \in \mathcal{P}} \quad \tilde{f}_0^{(l-1)} \left( \{\mathbf{P}_k\}_{k=1}^{K} \right)$$ (31a)

$$\text{s.t.} \quad \tilde{f}_i^{(l-1)} \left( \{\mathbf{P}_k\}_{k=1}^{K} \right) \geq 0, \quad \forall i.$$ (31b)

*Then, the sequence of $\{\mathbf{P}_k^{(l)}\}_{k=1}^{K}$ converges to a stationary point of* (30).

**Remark 1.** *The surrogate optimization problem* (31) *is not necessarily convex; however, we can obtain the global optimal solution of* (31) *much more easily than* (30).

Note that finding surrogate functions depends on the structure of the objective and constraint functions. In general, there might be different approaches to obtain a surrogate function (see, e.g., [62]). As indicated in Section II-D, the rate of each user is a difference of two concave functions, which allows us to apply CCP to obtain a suitable surrogate function. That is, we approximate the convex part of the rate expressions in (27) by its first-order Taylor series expansion,

---

[3]In [59], it is defined as a concave optimization problem. However, we call it convex since such an optimization problem is also widely known as a convex optimization problem [60].

which is a linear function. By MM and CCP, we are able to obtain a stationary point of optimization problems in which either the objective or constraint functions are linear functions of the rates of the users, as will be discussed in Section IV. In the following lemmas, we present the surrogate functions for the rates.

**Lemma 4.** *Using CCP, we can obtain an affine upper bound for* $\log \det(\mathbf{Q})$ *as*

$$\log_2 \det(\mathbf{Q}) \leq \log_2 \det\left(\mathbf{Q}^{(l)}\right) + \frac{1}{\ln 2} \mathrm{Tr}((\mathbf{Q}^{(l)})^{-1}(\mathbf{Q}-\mathbf{Q}^{(l)})), \tag{32}$$

*where* $\mathbf{Q}^{(l)}$ *is any feasible fixed point.*

*Proof.* A concave function can be majorized by an affine function if the two functions have the same value and the same derivative in a point [62]. The logarithmic function is concave. Furthermore, the left-hand and the right-hand sides of (32) hold these conditions at $\mathbf{Q} = \mathbf{Q}^{(l)}$. Thus, the upper-bound in (32) holds for all feasible $\mathbf{Q}$. Note that the derivative of $\log \det(\mathbf{Q})$ with respect to $\mathbf{Q}$ is $\mathbf{Q}^{-1}$. $\qquad\square$

**Lemma 5.** *A concave approximation of the rates in* (27) *can be obtained by CCP as*

$$R_k \geq \tilde{R}_k^{(l)} = r_{k,1} - r_{k,2}\left(\{\mathbf{P}_i^{(l)}\}_{i=1}^K\right)$$
$$- \mathrm{Tr}\left(\sum_{i=1,i\neq k}^K \frac{\partial r_{k,2}\left(\{\mathbf{P}_i^{(l)}\}_{i=1}^K\right)^T}{\partial \mathbf{P}_i}(\mathbf{P}_i - \mathbf{P}_i^{(l)})\right) \tag{33}$$

*where* $r_{k,1}$ *and* $r_{k,2}$ *are, respectively, the concave and convex parts of* $R_k$ *in* (27). *Moreover,* $\frac{\partial r_{k,2}\left(\{\mathbf{P}_i^{(l)}\}_{i=1}^K\right)}{\partial \mathbf{P}_i}$ *is the derivative of* $r_{k,2}$ *with respect to* $\mathbf{P}_i$ *at the previous iteration as*

$$\frac{\partial r_{k,2}\left(\{\mathbf{P}_i^{(l)}\}_{i=1}^K\right)}{\partial \mathbf{P}_i} =$$
$$\frac{1}{\ln 2}\tilde{\mathbf{H}}_{ki}^T\left(\underline{\mathbf{C}}_{z,k} + \sum_{i=1,i\neq k}^K \tilde{\mathbf{H}}_{ki}\mathbf{P}_i^{(l)}\tilde{\mathbf{H}}_{ki}^T\right)^{-1}\tilde{\mathbf{H}}_{ki}. \tag{34}$$

*Note that* $r_{k,2}\left(\{\mathbf{P}_i^{(l)}\}_{i=1}^K\right)$ *is constant and is given by* $r_{k,2}$ *at the previous step. Additionally,* $R_k$ *and* $\tilde{R}_k^{(l)}$ *fulfill the conditions in Lemma 3.*

## IV. OPTIMIZATION PROBLEMS

In this section, we obtain a stationary point of the rate region, sum-rate maximization, EE region and global EE maximization problems by the framework described in Section III.

### A. Achievable rate region

The rate region for the $K$-user MIMO IC with HWI can be derived by the rate-profile technique as [25]

$$\max_{R,\{\mathbf{P}_k\}_{k=1}^K \in \mathcal{P}} R \qquad \text{s.t.} \qquad R_k \geq \alpha_k R, \quad \forall k, \tag{35}$$

where $\alpha_k \geq 0$ for $k = 1, 2, \cdots, K$ are given constants, and $\sum_{k=1}^K \alpha_k = 1$. The boundary of the achievable rate region can be derived by solving (35) for different values of the $\alpha_k$s. The optimization problem (35) is not convex;

however, we can obtain its stationary point by the framework proposed in Section III. That is, we solve (35) iteratively, and in each iteration, we employ the surrogate function in Lemma 5 for the rates. Since the corresponding surrogate optimization problem is convex, we can efficiently derive the global optimal solution of each surrogate optimization problem and consequently, obtain a stationary point of (35).

### B. Maximizing sum-rate

The sum-rate of the $K$-user MIMO IC with HWI can be obtained by solving

$$\max_{\{\mathbf{P}_k\}_{k=1}^K \in \mathcal{P}} \sum_{k=1}^K R_k \tag{36a}$$
$$\text{s.t.} \qquad R_k \geq R_{\mathrm{th},k}, \qquad \forall k, \tag{36b}$$

where (36b) is the quality of service (QoS) constraint, and $R_{\mathrm{th},k}$ is a given threshold for the rate of user $k$. Note that the $R_{\mathrm{th},k}$s have to be set to make (36) feasible. Similar to (35), we can solve (36) by the framework in Section III and obtain its stationary point. Note that each surrogate optimization problem is convex, which can be solved efficiently.

### C. Energy-efficiency region

Now we consider the EE of the $K$-user MIMO IC with HWI. The EE of user $k$ is defined as the ratio of its achievable rate to its power consumption [50]

$$E_k = \frac{R_k}{\eta_k \mathrm{Tr}(\mathbf{P}_k) + P_{c,k}}, \qquad \text{(bits/Joule)} \tag{37}$$

where $\eta_k^{-1}$, and $P_{c,k}$ are, respectively, the power transmission efficiency of user $k$, and the constant power consumption of the $k$-th transceiver. The EE function is a linear function of the rates, which allows us to apply the framework in Section III to optimize the EE. EE function has a fractional structure, which makes its optimization more difficult than the rate analysis, as will be discussed in the following.

The EE region of the $K$-user MIMO IC with HWI can be derived by solving [50]

$$\max_{E,\{\mathbf{P}_k\}_{k=1}^K \in \mathcal{P}} E \tag{38a}$$
$$\text{s.t.} \qquad E_k = \frac{R_k}{\eta_k \mathrm{Tr}(\mathbf{P}_k) + P_{c,k}} \geq \alpha_k E, \quad \forall k, \tag{38b}$$
$$R_k \geq R_{\mathrm{th},k}, \qquad \forall k, \tag{38c}$$

where the constraint (38c) is the QoS constraint, similar to (36b), and $R_{\mathrm{th},k}$ has to be chosen such that the feasible set of parameters is not empty. Similar to (35), the boundary of the EE region can be derived by solving (38) for all possible $\alpha_k$s. Since $E_k$ is a linear function of $R_k$, we can apply the framework in Section III to derive a stationary point of (38). The surrogate optimization problem at the $l$-th iteration is

$$\max_{E,\{\mathbf{P}_k\}_{k=1}^K \in \mathcal{P}} E \tag{39a}$$
$$\text{s.t.} \qquad \tilde{E}_k^{(l)} = \frac{\tilde{R}_k^{(l)}}{\eta_k \mathrm{Tr}(\mathbf{P}_k) + P_{c,k}} \geq \alpha_k E, \quad \forall k, \tag{39b}$$
$$\tilde{R}_k^{(l)} \geq R_{\mathrm{th},k}, \qquad \forall k. \tag{39c}$$

Note that we can rewrite (39) as a maximin fractional optimization problem by removing $E$ as

$$\max_{\{\mathbf{P}_k\}_{k=1}^K \in \mathcal{P}} \quad \min_{1 \le k \le K} \left\{ \frac{\tilde{E}_k^{(l)}}{\alpha_k} \right\} \quad \text{s.t.} \quad \tilde{R}_k^{(l)} \ge R_{\text{th},k}, \quad \forall k. \tag{40}$$

The optimization problem (39) (or equivalently (40)) is not convex; however, its global optimum can be derived by employing the generalized Dinkelbach algorithm (GDA). The GDA is a powerful tool to solve maximin fractional optimization problems and is presented in the following Lemma.

**Lemma 6.** *Consider the following fractional optimization problem*

$$\max_{\{\mathbf{X}\} \in \mathcal{X}} \quad \min_{1 \le k \le K} \left\{ \frac{v_k(\mathbf{X})}{u_k(\mathbf{X})} \right\}, \tag{41}$$

*where $v_k(\cdot)$ is a concave function in $\mathbf{X}$, $u_k(\cdot)$ is a convex function in $\mathbf{X}$, and $\mathcal{X}$ is a compact set. The global optimal solution of (41) can be derived, iteratively, by the GDA, i.e., by solving*

$$\max_{t, \{\mathbf{X}\} \in \mathcal{X}} t \quad \text{s.t.} \quad v_k(\mathbf{X}) - \mu^{(m)} u_k(\mathbf{X}) \ge t, \quad \forall k, \tag{42}$$

*where $\mu^{(m)}$ is*

$$\mu^{(m)} = \min_{1 \le k \le K} \left\{ \frac{v_k(\mathbf{X}^{(m-1)})}{u_k(\mathbf{X}^{(m-1)})} \right\}, \tag{43}$$

*where $\mathbf{X}^{(m-1)}$ is the solution of (42) at the $(m-1)$th iteration. The GDA converges to the global optimum of (41) linearly.*

*Proof.* Please refer to [5], [50], [63]. $\quad\square$

**Remark 2.** *In order to apply the GDA, it is not necessary that $v_k$ and $-u_k$ are concave in $\mathbf{X}$. However, the GDA is ensured to obtain the global optimum if $v_k$ and $-u_k$ are concave in $\mathbf{X}$.*

Applying the GDA to (39), we have

$$\max_{E, \{\mathbf{P}_k\}_{k=1}^K \in \mathcal{P}} E \tag{44a}$$

$$\text{s.t.} \quad \tilde{R}_k^{(l)} - \mu^{(m)} \left( \eta_k \text{Tr}(\mathbf{P}_k) + P_{c,k} \right) \ge \alpha_k E, \quad \forall k, \tag{44b}$$

$$\tilde{R}_k^{(l)} \ge R_{\text{th},k}, \quad \forall k, \tag{44c}$$

where

$$\mu^{(m)} = \min_{1 \le k \le K} \left\{ \frac{E_k \left( \{\mathbf{P}_i^{(l,m-1)}\}_{i=1}^K \right)}{\alpha_k} \right\}, \tag{45}$$

where $\{\mathbf{P}_i^{(l,m-1)}\}_{i=1}^K$ is the solution of (44) at the $(m-1)$th iteration. Note that the GDA converges to the global optimum of (39) linearly, and the whole algorithm converges to a stationary point of (38).

### D. Global energy-efficiency

In this subsection, we consider the global EE of the $K$-user MIMO IC with HWI, which can be cast as [50]

$$\max_{\{\mathbf{P}_k\}_{k=1}^K \in \mathcal{P}} G = \frac{\sum_{k=1}^K R_k}{\sum_{k=1}^K \left( \eta_k \text{Tr}(\mathbf{P}_k) + P_{c,k} \right)} \tag{46a}$$

$$\text{s.t.} \quad R_k \ge R_{\text{th},k}, \quad \forall k. \tag{46b}$$

Similar to (38), since the EE is a linear function of the rates, we can apply the framework in Section III to obtain a stationary point of (46). Thus, the surrogate optimization problem at the $l$-th iteration is

$$\max_{\{\mathbf{P}_k\}_{k=1}^K \in \mathcal{P}} \tilde{G} = \frac{\sum_{k=1}^K \tilde{R}_k}{\sum_{k=1}^K \left( \eta_k \text{Tr}(\mathbf{P}_k) + P_{c,k} \right)} \tag{47a}$$

$$\text{s.t.} \quad \tilde{R}_k \ge R_{\text{th},k}, \quad \forall k. \tag{47b}$$

Similar to (39), the optimization problem is not convex; however, its global optimal solution can be derived by the Dinkelbach algorithm. That is, we obtain $\mathbf{P}_i^{(l,m)}$ by solving

$$\max_{\{\mathbf{P}_k\}_{k=1}^K \in \mathcal{P}} \sum_{k=1}^K \tilde{R}_k - \mu^{(m)} \sum_{k=1}^K \left( \eta_k \text{Tr}(\mathbf{P}_k) + P_{c,k} \right) \tag{48a}$$

$$\text{s.t.} \quad \tilde{R}_k \ge R_{\text{th},k}, \quad \forall k, \tag{48b}$$

where $\mu^{(m)} = \tilde{G}\left( \{\mathbf{P}_i^{(l,m-1)}\}_{i=1}^K \right)$, in which $\{\mathbf{P}_i^{(l,m-1)}\}_{i=1}^K$ is the solution of (48) at the $(m-1)$th iteration. The global optimal solution of (47a) can be achieved by iteratively solving (48) and updating $\mu^{(m)}$ until a convergence metric is met. Moreover, as indicated, the whole algorithm converges to a stationary point of (46).

## V. NUMERICAL EXAMPLES

In this section, we provide some numerical examples. We employ Monte Carlo simulations and average the results over 100 independent channel realizations. In each channel realization, the channel entries are drawn from a zero-mean complex proper Gaussian distribution with unit variance, i.e., $\mathcal{CN}(0,1)$. For all simulations, the maximum number of the iterations of the MM algorithm is set to 40. We also consider $\mathbf{C}_T = \sigma_T^2 \mathbf{I}_{N_T}$ and $\mathbf{C}_R = \sigma_R^2 \mathbf{I}_{N_R}$ [43], or equivalently $\underline{\mathbf{C}}_T = \frac{1}{2}\sigma_T^2 \mathbf{I}_{2N_T}$ and $\underline{\mathbf{C}}_R = \frac{1}{2}\sigma_R^2 \mathbf{I}_{2N_R}$. In all simulations, we assume $\sigma_T^2 = 0.2$ and $\sigma_R^2 = 1$. We assume that the I/Q imbalance by each antenna is the same. In other words, the matrices $\mathbf{A}_T = a_T \mathbf{I}_{N_T}$, $\boldsymbol{\theta}_T = \phi_T \mathbf{I}_{N_T}$, $\mathbf{A}_R = a_R \mathbf{I}_{N_R}$, and $\boldsymbol{\theta}_R = \phi_R \mathbf{I}_{N_R}$ are scaled identity matrices. We consider $\phi_T = \phi_R = 5$ degrees in all simulations. We also define the signal-to-noise ratio (SNR) as the ratio of the power budget to $\sigma^2$, i.e., SNR$= \frac{P}{\sigma^2}$.

To the best of our knowledge, there are no other IGS algorithms in the literature that optimize EE or rate functions in the K-user MIMO IC. Therefore, we compare our proposed algorithms for PGS and IGS with the PGS algorithm for ideal devices. The considered schemes in this section are as follows:

- **IGS**: The IGS scheme.
- **PGS**: The PGS scheme.
- **I-PGS**: The PGS scheme for $K$-user IC without considering the I/Q imbalance in the design.

Note that the performance of MM algorithms depend on the initial point. In PGS and I-PGS algorithms, we start with a uniform power allocation $\mathbf{P}_k = \frac{P}{2N_T}\mathbf{I}_{2N_T}$ for $k = 1, \cdots, K$ for optimization problems (35) and (36); and $\mathbf{P}_k = \frac{0.3P}{2N_T}\mathbf{I}_{2N_T}$ for $k = 1, \cdots, K$ for optimization problems (38) and (46). On the other hand, the IGS algorithm takes the solution of the PGS algorithm as an initial point.
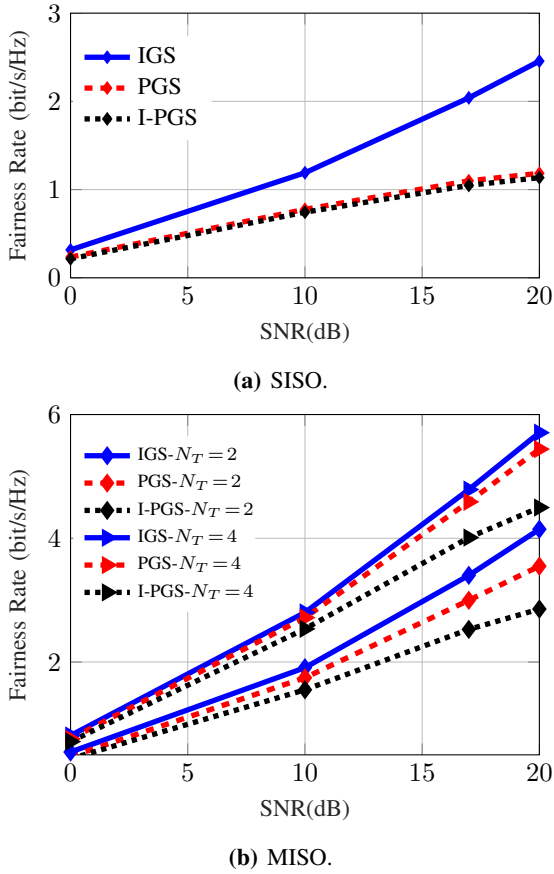
**(a)** SISO.



**(b)** MISO.

**Fig. 3:** Average fairness rate versus SNR for the 2-user SISO and MISO IC.



**(a)** $K = 2$ and $K = 4$.



**(b)** $K = 6$.

**Fig. 4:** Average fairness rate versus SNR for the $K$-user $2 \times 2$ MIMO IC.

## A. Achievable rate region

In this subsection, we consider a specific point of the rate region, which maximizes the minimum rate of users. The minimum rate of the $K$-user MIMO IC is maximized for $\alpha_k = \frac{1}{K}$. This point is also referred as the maximin point or the fairness point. Hereafter, we call the maximin rate the fairness rate. We show the fairness rate of the 2-user SISO and MISO IC for $a_T = 0.6$ and different number of antennas at the transmitter side in Fig. 3. As can be observed, there is a huge performance improvement by IGS in the 2-user SISO IC, especially at high SNR. However, the benefits of employing IGS become less substantial when the number of antennas increases. This is due to the fact that, by increasing the number of spatial resources (e.g. antennas) for a fixed number of users, the interference of PGS schemes can be managed more easily, and hence, IGS as an interference-management tool does not provide an additional significant gain. As indicated, this is in line with the results in [48], in which it was shown that IGS might not provide significant benefits in OFDM UCR systems due to the existence of multiple parallel channels over which interference can be managed efficiently without resorting to IGS. Moreover, in [49], it was shown that the IGS does not provide any benefit in comparison to PGS with time sharing when the average power consumption is constrained instead of the instantaneous power, which allows a more flexible power allocation. To sum up, the benefits of IGS decrease or even vanish when increasing the number of resources either by increasing the number of antennas or number of time slots, by time sharing, and/or the number of parallel
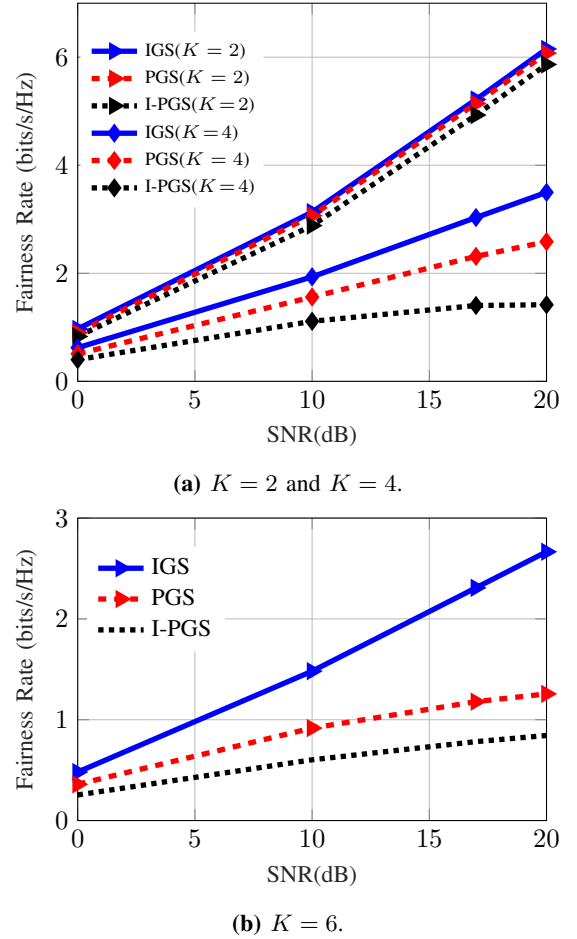
channels by OFDM. Nevertheless, it is worth mentioning that, even with a large number of antennas, IGS and HWI-aware PGS outperform PGS, which is designed for ideal devices. Furthermore, IGS always performs not worse than PGS since IGS includes PGS as a particular case. With I/Q imbalance, IGS performs better than PGS even in the MISO case with $N_T = 4$. However, the benefits of IGS as an interference-management technique are not significant in this case.

Figure 4 shows the fairness rate of the $K$-user $2 \times 2$ MIMO IC versus the SNR for $a_T = 0.6$. As can be observed, the benefit of IGS is minor when $K = 2$. However, by increasing the number of users, the performance improvement of IGS increases. The reason is that, by increasing the number of users, the interference level increases as well, which results in more performance improvements by transmitting improper signals. Moreover, IGS performs much better at high SNR for $K = 6$, similar to the 2-user SISO IC as depicted in Fig. 3a.

Figure 5 shows the fairness rate versus the level of the I/Q imbalance, $1 - a_T$, for SNR= 0 dB and $N_T = N_R = 2$. As can be observed, the IGS design is less affected by the HWI level for different $K$. When $K = 2$, the IGS and PGS schemes perform very similarly in low HWI level. However, the performance of the PGS scheme drastically decreases with the HWI level, while the fairness rate of the IGS scheme decreases only slightly. When $K = 4$ and $K = 6$, the same trend is observed, but the relative performance of the IGS scheme over the PGS scheme increases with $K$. Moreover,
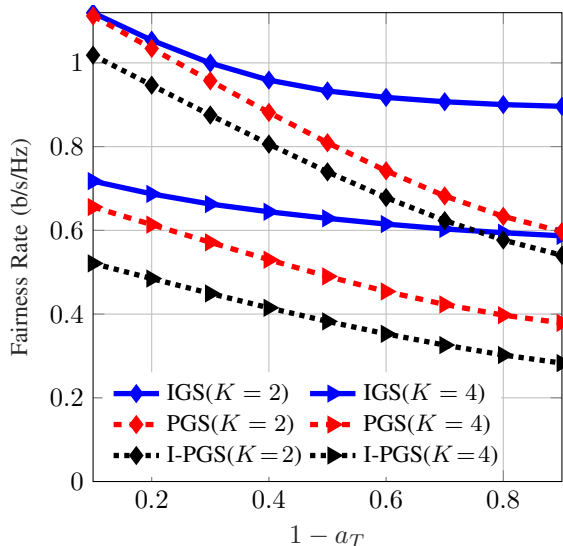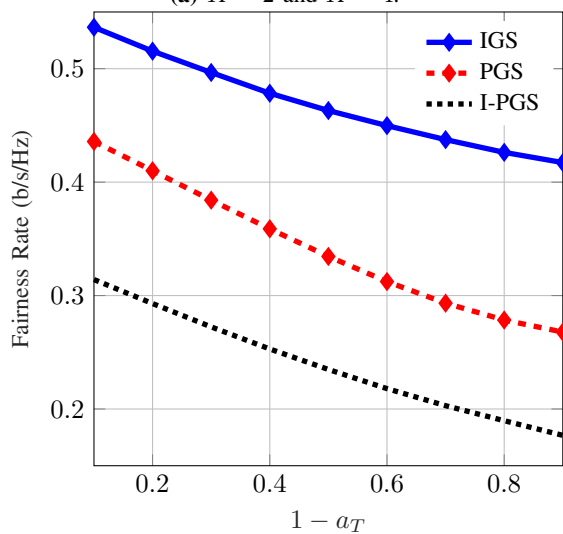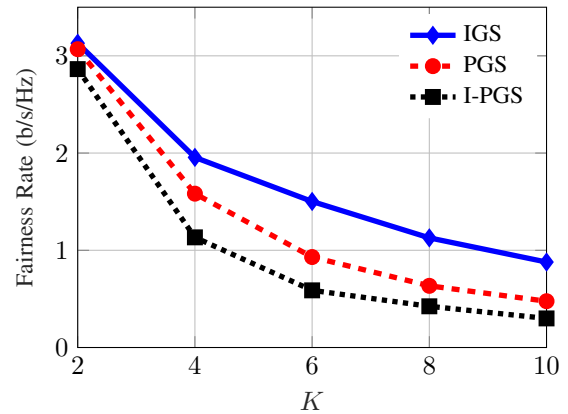
**(a)** $K = 2$ and $K = 4$.



**(b)** $K = 6$.

**Fig. 5:** Average fairness rate versus the I/Q imbalance level for the $K$-user $2 \times 2$ MIMO IC with SNR$= 0$ dB.



**(a)** Achievable rate



**(b)** Relative improvement of IGS over PGS

**Fig. 6:** Average fairness rate versus the number of users for the $K$-user $2 \times 2$ MIMO IC with SNR$= 10$ dB.



**(a)** Achievable sum-rate



**(b)** Relative improvement of IGS over PGS

**Fig. 7:** Average achievable sum-rate and relative performance of IGS versus the number of users for the $K$-user $2 \times 2$ MIMO IC with SNR$= 10$ dB.

for a given $K$, the benefits of IGS increase with the level of the I/Q imbalance, as expected.

Figure 6 considers the effect of the number of users on the fairness rate as well as on the performance of IGS for SNR$=$ 10 dB, $a_T = 0.6$ and $N_T = N_R = 2$. As can be observed, the fairness rates decreases when $K$ increases. Additionally, the relative performance improvement by IGS increases with the number of users. For $K = 10$, the relative improvement of IGS over PGS is more than $80\%$. The reason is that more users provoke more interference, which results in turn in more improvements by IGS, as indicated before.

### B. Achievable sum-rate

In Fig. 7, we show the effect of the number of users on the achievable sum-rate of the $K$-user $2 \times 2$ MIMO IC for SNR$= 10$ dB, and $a_T = 0.6$. In this figure, we set the threshold in (36) to $R_{\text{th},k} = 0$. We can observe that the sum-rate and also the relative performance improvement of IGS over PGS are increasing with respect to $K$. Since we maximize the sum-rate without considering a QoS constraint, the rate of some users with weak direct links might be even
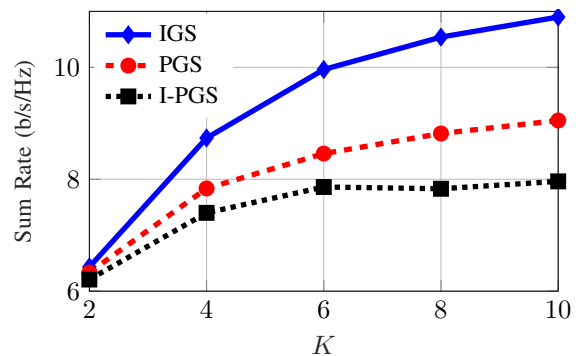
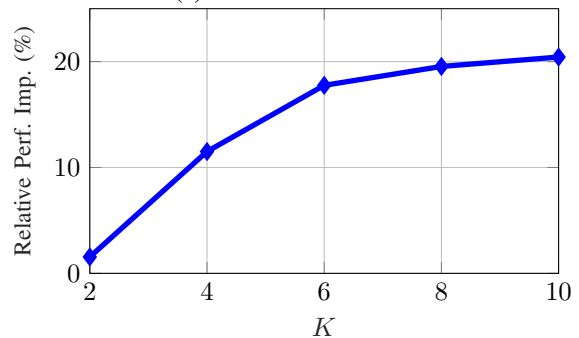0, which causes less interference. As a result, the relative performance improvement of IGS is less significant than the improvements for the fairness rate observed in Fig. 6.
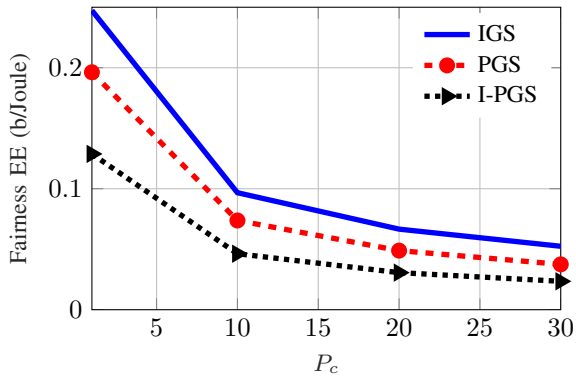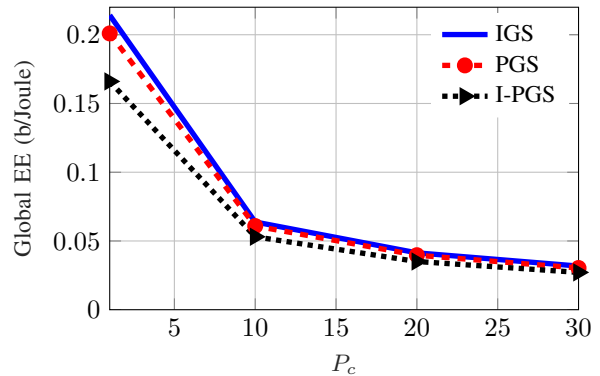
**Fig. 8:** Average fairness EE of the 6-user $2 \times 2$ MIMO IC versus $P_c$.



**Fig. 9:** Relative performance of IGS with respect to PGS and I-PGS versus $P_c$ for the 6-user $2 \times 2$ MIMO IC.



**Fig. 10:** Average global EE of the 6-user $2 \times 2$ MIMO IC versus $P_c$.

### C. Energy efficiency region

In this subsection, we consider the EE region in (38). In general, IGS provides less EE benefits than rate benefits. For example, in [34], it was shown that the conditions for the optimality of IGS over PGS in an UCR network are more stringent for EE than for rate. In other words, it might happen that IGS must be used for the SU rate to be maximized, while PGS must be used for the EE to be maximized.

In Fig. 8, we show the fairness EE of the 6-user $2 \times 2$ MIMO IC versus $P_c$ for SNR= 10 dB, and $a_T = 0.6$. As can be observed, the fairness EE decreases with $P_c$. Moreover, the proposed IGS scheme outperforms the PGS scheme as well as I-PGS. Figure 9 shows the relative performance improvement by the IGS scheme with respect to the PGS and I-PGS schemes for the results in Fig. 8. As can be observed in these figures, the fairness EE decreases with $P_c$; however, the benefits of employing IGS is increasing in $P_c$. The reason is that when $P_c$ is very large, the EE-region-optimization problem becomes equivalent to the achievable-rate-region problem, and as indicated, IGS can provide more gain in achievable-rate optimizations.

### D. Global Energy efficiency

Figure 10 shows the global EE of the 6-user $2 \times 2$ MIMO IC versus $P_c$ for SNR= 10 dB, and $a_T = 0.6$. In this figure, we assume $R_{\text{th},k} = 0$ in (46). As can be observed, IGS provides minor benefits in terms of the global EE. Since the QoS constraint is not considered, it might happen that some users are switched off, thus reducing the total level of interference. Moreover, the lower the interference level, the less need for advanced interference-management techniques. Thus, we can expect that the benefits of employing IGS decrease in global EE with respect to per-user EE. Note that IGS still performs slightly better than PGS, which can be due to the I/Q imbalance.

## VI. Conclusion

This paper studied the performance of IGS of a $K$-user MIMO IC with HWI including I/Q imbalance at the transceivers. In the presence of I/Q imbalance, the received signal is a function of the widely linear transform of the transmitted signal and the aggregated noise. Hence, the effective noise is modeled as improper at the receiver side, which motivated us to consider the use of IGS. Considering achievable rates and EE as performance metrics, we proposed HWI-aware IGS schemes for the $K$-user MIMO IC. We employed an optimization framework, which can obtain a stationary point of any optimization problem for interference-limited systems with TIN in which the objective function and/or constraints are linear functions of the achievable rate. In this paper, we derived a stationary point of the achievable rate-region, sum-rate maximization, EE region and global EE maximization problems. Our numerical results showed that IGS can improve the performance of the $K$-user MIMO IC with HWI from both achievable rate and EE points of view. We observed that the benefits of IGS as an interference-management technique increase with the number of users and decrease with the number of antennas. This is due to the fact that higher interference levels result in an increased need for interference management and consequently, more improvements by IGS. We also observed that the benefit of employing IGS increases with impairment level.

As future work, it may be interesting to find out how close the solution of this algorithm is to the corresponding global optimum solutions. Additionally, our scheme is a centralized approach, which might not be applicable in some practical scenarios. Hence, distributed algorithms should also be developed.

## REFERENCES

[1] E. Bjornson, M. Matthaiou, and M. Debbah, "A new look at dual-hop relaying: Performance limits with hardware impairments," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4512–4525, 2013.

[2] L. Tlebaldiyeva, B. Maham, and T. A. Tsiftsis, "Device-to-device mmWave communication in the presence of interference and hardware distortion noises," *IEEE Commun. Lett.*, 2019.

[3] S. Javed, O. Amin, and M.-S. Alouini, "Full-duplex relaying under I/Q imbalance using improper Gaussian signaling," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Processing (ICASSP)*, 2017, pp. 6538–6542.

[4] 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.

[5] 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.

[6] 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.

[7] X. Zhang, D. Guo, K. An, and B. Zhang, "Secure communications over cell-free massive MIMO networks with hardware impairments," *IEEE Systems Journal*, 2019.

[8] 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, 2017.

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

[10] J. Zhang, L. Dai, X. Zhang, E. Björnson, and Z. Wang, "Achievable rate of Rician large-scale MIMO channels with transceiver hardware impairments," *IEEE Trans. Veh. Technol*, vol. 65, no. 10, pp. 8800–8806, 2016.

[11] Q. Zhang, T. Q. Quek, and S. Jin, "Scaling analysis for massive MIMO systems with hardware impairments in Rician fading," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4536–4549, 2018.

[12] X. Xia, D. Zhang, K. Xu, W. Ma, and Y. Xu, "Hardware impairments aware transceiver for full-duplex massive MIMO relaying," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6565–6580, 2015.

[13] J. Feng, S. Ma, S. Aïssa, and M. Xia, "Two-way massive MIMO relaying systems with non-ideal transceivers: Joint power and hardware scaling," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8273–8289, 2019.

[14] S. Cheng, R. Wang, J. Wu, W. Zhang, and Z. Fang, "Performance analysis and beamforming designs of MIMO AF relaying with hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6229–6243, 2018.

[15] 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, 2013.

[16] J. Zhang, Y. Wei, E. Björnson, Y. Han, and S. Jin, "Performance analysis and power control of cell-free massive MIMO systems with hardware impairments," *IEEE Access*, vol. 6, pp. 55 302–55 314, 2018.

[17] T. Younas, J. Li, and J. Arshad, "On bandwidth efficiency analysis for LS-MIMO with hardware impairments," *IEEE Access*, vol. 5, pp. 5994–6001, 2017.

[18] G. Mohamed, "Massive relaying communication system under I/Q imbalance and hardware manufacturing problems," *IEEE Systems Journal*, vol. 13, no. 4, pp. 3889–3896, 2019.

[19] L. Tlebaldiyeva, T. A. Tsiftsis, and B. Maham, "Performance analysis of improved energy detector with hardware impairments for accurate spectrum sensing," *IEEE Access*, vol. 7, pp. 13 927–13 938, 2019.

[20] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[21] 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.

[22] L. Yang and W. Zhang, "Interference alignment with asymmetric complex signaling on MIMO X channels," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3560–3570, 2014.

[23] C. Lameiro and I. Santamaría, "Degrees-of-freedom for the 4-user SISO interference channel with improper signaling," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2013, pp. 3053–3057.

[24] Z. K. Ho and E. Jorswieck, "Improper Gaussian signaling on the two-user SISO interference channel," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3194–3203, 2012.

[25] 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.

[26] 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.

[27] 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.

[28] H. D. Nguyen, R. Zhang, and S. Sun, "Improper signaling for symbol error rate minimization in $K$-user interference channel," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 857–869, 2015.

[29] 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.

[30] E. Kurniawan and S. Sun, "Improper Gaussian signaling scheme for the Z-interference channel," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3912–3923, 2015.

[31] C. Lameiro, I. Santamaría, and P. J. Schreier, "Rate region boundary of the SISO Z-interference channel with improper signaling," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1022–1034, 2017.

[32] ——, "Benefits of improper signaling for underlay cognitive radio," *IEEE Wireless Commun. Lett.*, vol. 4, no. 1, pp. 22–25, 2015.

[33] 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.

[34] 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.

[35] M. Soleymani, C. Lameiro, I. Santamaria, and P. J. Schreier, "Energy-efficient improper signaling for $K$-User interference channels," in *Proc. IEEE European Signal Process. Conference (EUSIPCO)*, 2019.

[36] S. Lagen, A. Agustin, and J. Vidal, "Coexisting linear and widely linear transceivers in the MIMO interference channel," *IEEE Trans. Signal Process.*, vol. 64, no. 3, pp. 652–664, 2016.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5101–5125, 2011.

[42] T. Adali and P. J. Schreier, "Optimization and estimation of complex-valued signals: Theory and applications in filtering and blind source separation," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 112–128, 2014.

[43] 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.

[44] A.-A. A. Boulogeorgos, 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.

[45] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Asymmetric modulation for hardware impaired systems-error probability analysis and receiver design," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1723–1738, 2019.

[46] ——, "Asymmetric hardware distortions in receive diversity systems: Outage performance analysis," *IEEE Access*, vol. 5, pp. 4492–4504, 2017.

[47] ——, "Impact of improper Gaussian signaling on hardware impaired systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2017, pp. 1–6.

[48] 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.

[49] 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.

[50] 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.

[51] G. R. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Proc. Adv. Neural Inf. Process. Syst*, 2009, pp. 1759–1767.

[52] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.

[53] C. Hellings and W. Utschick, "Block-skew-circulant matrices in complex-valued signal processing," *IEEE Trans. Signal Process.*, vol. 63, no. 8, pp. 2093–2107, 2015.

[54] Y.-H. Chung and S.-M. Phoong, "Joint estimation of I/Q imbalance, CFO and channel response for MIMO OFDM systems," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1485–1492, 2010.

[55] G.-T. Gil, I.-H. Sohn, J.-K. Park, and Y. H. Lee, "Joint ML estimation of carrier frequency, channel, I/Q mismatch, and DC offset in communication receivers," *IEEE Trans. Veh. Technol.*, vol. 54, no. 1, pp. 338–349, 2005.

[56] K.-Y. Sung and C.-c. Chao, "Estimation and compensation of I/Q imbalance in OFDM direct-conversion receivers," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 3, pp. 438–453, 2009.

[57] X. Cai, Y.-C. Wu, H. Lin, and K. Yamashita, "Estimation and compensation of CFO and I/Q imbalance in OFDM systems under timing ambiguity," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1200–1205, 2011.

[58] L. You, J. Xiong, A. Zappone, W. Wang, and X. Gao, "Spectral efficiency and energy efficiency tradeoff in massive MIMO downlink transmission with statistical CSIT," *IEEE Trans. on Signal Process.*, vol. 68, pp. 2645–2659, 2020.

[59] A. Aubry, A. De Maio, A. Zappone, M. Razaviyayn, and Z.-Q. Luo, "A new sequential optimization procedure and its applications to resource allocation for wireless systems," *IEEE Trans. Signal Process.*, vol. 66, no. 24, pp. 6518–6533, 2018.

[60] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[61] Y. Yang and M. Pesavento, "A unified successive pseudoconvex approximation framework," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3313–3328, 2017.

[62] 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.

[63] J.-P. Crouzeix and J. A. Ferland, "Algorithms for generalized fractional programming," *Math. Programm.*, vol. 52, no. 1-3, pp. 191–207, May 1991.

**Ignacio Santamaria (M'96, SM'05)** received the Telecommunication Engineer degree and the Ph.D. degree in electrical engineering from the Universidad Politécnica 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 Full Professor since 2007. 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 at the Colorado State University. Prof. Santamaria was general co-chair of the 2012 IEEE Workshop on Machine Learning for Signal Processing (MLSP 2012). From 2009 to 2014, he was a member of the IEEE Machine Learning for Signal Processing Technical Committee. 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 EEEfCOM 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.

**Mohammad Soleymani** received the B.Sc. degree in electrical engineering from the Amirkabir University of Technology (Tehran Polytechnic) and the M.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran. He is pursuing the Ph.D. degree with the Signal and System Theory Group, University of Paderborn, Germany. He was a Visiting Researcher at University of Cantabria, Spain. His research interests include multiuser communications, wireless networking, convex optimization and statistical signal processing.