Improper Signaling for SISO Two-user Interference Channels with Asymmetric Hardware Impairment

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

Abstract

Hardware non-idealities are among the main performance restrictions for upcoming wireless communication systems. Asymmetric hardware impairments (HWI) happen when the impairments of the I/Q branches are correlated or imbalanced. In this paper, we investigate the benefits of improper Gaussian signaling (IGS) in a two-user interference channel (IC) with asymmetric HWI when interference is treated as noise. The real and imaginary parts of improper or asymmetric Gaussian signals are correlated and/or have unequal powers. IGS has been recently shown to improve the performance of different interference-limited systems. We propose two iterative algorithms to optimize the parameters of the improper transmit signals. We first rewrite the rate region as an energy region and employ fractional programming and the generalized Dinkelbach algorithm to find a suboptimal solution for the achievable rate pairs. Then, we propose a simplified algorithm based on a separate optimization of the powers and complementary variances of the users, which exhibits lower computational complexity. We show that IGS can improve the performance of the two-user IC with HWI. Our proposed algorithms outperform proper Gaussian signaling as well as competing IGS algorithms in the literature that do not consider asymmetric HWI.

*Mohammad Soleymani, Christian Lameiro and Peter J. Schreier are with the Signal and System Theory Group, Universität Paderborn, Germany, [http://sst.upb.de](http://sst.upb.de) (emails: {mohammad.soleymani,christian.lameiro,peter.schreier}@sst.upb.de).
†Ignacio Santamaria is with the Department of Communications Engineering, University of Cantabria (email: i.santamaria@unican.es).
Index Terms

Achievable rate region, asymmetric hardware impairments, difference of convex programming, generalized Dinkelbach algorithm, improper Gaussian signaling, interference channel.

I. Introduction

One of the targets of 5G is reaching a data rate more than 1000 times greater than the data rate of current cellular systems [1]. However, reaching this goal entails many challenges. Among them is to overcome the non-idealities, i.e., hardware impairments (HWI), of devices which can result in a substantial performance degradation [2]–[4]. HWI are due to various imperfections in transceivers, including I/Q imbalance, non-linear power amplifiers, imperfect and/or low resolution analog-to-digital and digital-to-analog converters, frequency/phase offset and so on [3]–[7]. Another main challenge for data-rate enhancement is interference from other users, and hence interference management techniques play a key role in 5G [1]. Recently, it has been shown that improper Gaussian signaling (IGS) can improve the performance of various interference-limited systems [8]–[19]. In IGS, the real and imaginary parts of the signal are correlated and/or have unequal powers [20], [21]. While proper Gaussian signaling (PGS) achieves channel capacity for point-to-point communications in the presence of proper Gaussian noise [22], this is not the case under asymmetric HWI [4], [8], [23].

The effect of HWI is studied in [5]–[7] for various scenarios. In [5], the authors considered a full-duplex massive multiple-input multiple-output (MIMO) relay with HWI and proposed a scheme to mitigate the distortion by exploiting statistical knowledge of channels. The paper [6] analyzed the achievable rate of massive MIMO systems with Rician channels and HWI. In [7], the secrecy performance of downlink massive MIMO systems was considered with HWI and a passive multiple-antenna eavesdropper. In the aforementioned papers, symmetric HWI are considered. Nevertheless, HWI can, in general, provoke asymmetric or improper distortion in both the transmitted and received signal [4], [8], [23], [24]. The paper [4] considered IGS in a single-input, multiple-output (SIMO) system with asymmetric HWI and showed that IGS improved the
performance of the system. In [8], the authors investigated the effect of IGS in a relay network with asymmetric HWI. They maximized the achievable rate of the relay network by optimizing the complementary variance of the transmitted signal in the source and relay nodes.

In [9]–[19], the authors investigated the benefits of IGS in different interference-limited systems with ideal devices. In [9], IGS was considered as an interference management tool for the first time in the literature, where the authors considered a three-user interference channel (IC) and showed that IGS can improve the degrees-of-freedom (DoF) in the IC. The paper [10] showed that IGS can increase the DoF of MIMO X channels. In [11]–[18], the authors studied the performance of IGS when Treating Interference as Noise (TIN) was the strategy used for decoding. The paper [11] showed that IGS can improve the performance of the two-user interference channel. In [12], the authors used IGS to optimize the rate of the $K$-user MIMO interference channel. Moreover, they derived the rate region of the two-user IC with TIN by solving a semidefinite programming (SDP) problem. They showed that IGS can enlarge the rate region and improve the performance of the system. The paper [13] showed that IGS can reduce the symbol error rate of the $K$-user IC. In [14]–[16], benefits of IGS were studied in different Z-IC scenarios. In [17], [18], the authors showed that IGS improves the performance of underlay and overlay cognitive radio systems, respectively. Finally, the paper [19] showed that IGS can improve the performance of full-duplex relaying systems with fading channels.

In this paper, we study the performance of IGS in a two-user IC with asymmetric HWI with TIN. We assume that the transceivers of both users produce asymmetric HWI noise, and model the HWI as an additive improper Gaussian noise, similar to [4], [8], [23], [24]. We devise two iterative algorithms for the two-user IC to operate as closely as possible to the Pareto-optimal region. Our first proposed algorithm is based on fractional programming (FP) and the well-known generalized Dinkelbach algorithm. In order to apply FP to the rate-region problem, we rewrite the problem as an energy region, to which we apply the generalized Dinkelbach algorithm. The generalized Dinkelbach algorithm is a powerful tool to solve multiple ratio maximin problems [25], [26]. In Dinkelbach-based algorithms, an iterative optimization problem is proposed, in which the frac-
tional functions are replaced by surrogate functions at each iteration. The generalized Dinkelbach algorithm permits solving fractional programming efficiently and results in the global optimal solution of the original optimization problem if the optimization problem at each iteration is perfectly solved [25]–[28]. In our proposed algorithm, each iteration of the generalized Dinkelbach algorithm results in a non-convex optimization problem. To efficiently solve these problems, we employ difference of convex programming (DCP), which is a special case of sequential convex programming (SCP).

We also propose a simplified algorithm that is computationally less expensive than our proposed algorithm with FP. This simplified algorithm is based on a separate optimization of powers and complementary variances of users. That is, we first optimize the transmission power of users by employing the well-known bisection method, which transforms the original problem into a sequence of feasibility problems, and derive a closed-form solution for the feasibility problem. Then, we employ DCP to optimize the complementary variances of the users. This simplified decoupled algorithm is much faster than the algorithm based on FP, which jointly optimizes the variances and complementary variances of the users. Our results show that IGS enlarges the achievable rate of the two-user IC in the presence of asymmetric HWI, and that there is a significant performance improvement by IGS for highly asymmetric HWI noise. Moreover, both of our proposed algorithms outperform PGS and other existing IGS algorithms.

The rest of this paper is organized as follows. Section II describes the scenario and formulates the achievable-rate-region problem. In Section III, we propose our algorithm based on FP, and in Section IV, we develop a simplified version of this algorithm. Finally, Section V presents some numerical results.

II. System Model

A. Preliminaries of IGS

Let us consider a zero-mean complex Gaussian random variable $x$ with variance $p_x = \mathbb{E}\{|x|^2\}$ and complementary variance $q_x = \mathbb{E}\{x^2\}$ [20], [21]. Note that the complementary variance is
complex and $|q_x| \leq p_x$. We denote the probability distribution of $x$ by $x \sim \mathcal{CN}(m_x, p_x, q_x)$, where $m_x = 0$ is the mean of $x$. We define the complex correlation coefficient of $x$ as $\tilde{\kappa}_x = \frac{q_x}{p_x}$, where $|\tilde{\kappa}_x| \in [0, 1]$ is the so-called circularity coefficient. If $\tilde{\kappa}_x = 0$, $x$ is proper; otherwise, it is improper \cite{20, 21}. We call $x$ maximally improper if $|\tilde{\kappa}_x| = 1$.

B. Problem Statement

We consider a two-user IC with asymmetric HWI at the transmitters and receivers of both users, as depicted in Fig. 1. We assume that users are allowed to employ IGS and treat the interference as noise. We model the aggregated HWI at the transceiver of a communication link with an improper Gaussian additive noise as \cite{4, 8, 23, 24}

$$y = \sqrt{P}h(x + \eta) + n,$$

where $y$, $x$, $P$, $h$, $\eta$, and $n$ are the received signal, transmitted symbol, transmission power, channel coefficient, HWI noise and additive complex proper Gaussian noise, respectively. The HWI noise is modeled as an improper complex Gaussian random variable with probability distribution $\eta \sim \mathcal{CN}(0, \sigma^2_{\eta}, \tilde{\sigma}^2_{\eta})$, where $\sigma^2_{\eta} = \sigma^2_{\eta_{\text{TX}}} + \sigma^2_{\eta_{\text{RX}}}$ and $\tilde{\sigma}^2_{\eta} = \tilde{\sigma}^2_{\eta_{\text{TX}}} + \tilde{\sigma}^2_{\eta_{\text{RX}}}$ are variance and complementary variance of $\eta$, respectively, which are each composed of contributions at the transmitter side (denoted TX) and the receiver side (denoted RX). Note that the variances and complementary variances of HWI noise are not only a function of device parameters, but also a linear function of the transmission power and channel gain, meaning that higher transmission power results in higher

\footnote{It is worth mentioning that our algorithms can easily be extended to the $K$-user IC. However, we consider only the 2-user IC for the ease of illustration.}
HWI noise. Moreover, even if the channel noise is proper, the aggregated distortion is improper due to the asymmetric HWI.

Using this model, the received signals at receivers of user 1 and 2 are

\[ y_1 = \sqrt{p_1} h_{11}(x_1 + n_1) + \sqrt{p_2} h_{21}(x_2 + n_2) + n_1, \]  
\[ y_2 = \sqrt{p_1} h_{12}(x_1 + n_1) + \sqrt{p_2} h_{22}(x_2 + n_2) + n_2, \]  

respectively, where \( x_i, h_{ij}, n_i, \) and \( \eta_{ij} \) for \( i, j \in \{1, 2\} \) are the transmit signal of user \( i \), channel coefficient of the link between transmitter \( i \) and receiver \( j \), independent zero-mean proper complex Gaussian noise with variance \( \sigma^2 \), and aggregated HWI noise of the link between transmitter \( i \) and receiver \( j \), respectively. Without loss of generality, we assume that the variance of the transmitted symbol of user \( k \), for \( k \in \{1, 2\} \), is equal to 1, i.e., \( \sigma^2_{n_k} = 1 \). Since the transmitted signals \( x_1 \) and \( x_2 \) are improper complex Gaussian, the achievable rates of users 1 and 2 are \([8], [12], [16]\)

\[ R_1 = \frac{1}{2} \log_2 \left( \frac{\left( \sigma^2 + \sum_{j=1}^{2} p_j |h_{j1}|^2 (1 + \sigma^2_{\eta_{j1}}) \right)^2 - \left| \sum_{j=1}^{2} (q_j + p_j \bar{\sigma}^2_{\eta_{j1}}) h_{j1} \right|^2}{\left( \sigma^2 + p_2 |h_{21}|^2 (1 + \sigma^2_{\eta_{21}}) + p_1 |h_{11}|^2 \sigma^2_{\eta_{11}} \right)^2 - \left| p_1 \bar{\sigma}^2_{\eta_{11}} h_{11} + (q_2 + p_2 \bar{\sigma}^2_{\eta_{21}}) h_{21} \right|^2} \), \]  
\[ R_2 = \frac{1}{2} \log_2 \left( \frac{\left( \sigma^2 + \sum_{j=1}^{2} p_j |h_{j2}|^2 (1 + \sigma^2_{\eta_{j2}}) \right)^2 - \left| \sum_{j=1}^{2} (q_j + p_j \bar{\sigma}^2_{\eta_{j2}}) h_{j2} \right|^2}{\left( \sigma^2 + p_1 |h_{12}|^2 (1 + \sigma^2_{\eta_{12}}) + p_2 |h_{22}|^2 \sigma^2_{\eta_{22}} \right)^2 - \left| (q_1 + p_1 \bar{\sigma}^2_{\eta_{12}}) h_{12} + p_2 \bar{\sigma}^2_{\eta_{22}} h_{22} \right|^2} \), \]

where \( p_i, q_i, \sigma^2_{\eta_{ij}}, \) and \( \bar{\sigma}^2_{\eta_{ij}} \) for \( i, j \in \{1, 2\} \) are, respectively, the transmission power of user \( i \), complementary variance of the transmitted signal of user \( i \), aggregated variance and complementary variance of the HWI noise in the link between user \( i \) and user \( j \). The rate of user \( k \in \{1, 2\} \) can be written using vector notation as

\[ R_k = \frac{1}{2} \log_2 \left( \frac{(\sigma^2 + \mathbf{a}_k^T \mathbf{p})^2 - |\mathbf{f}_k^H \mathbf{q} + \bar{\mathbf{f}}_k^H \mathbf{p}|^2}{(\sigma^2 + \mathbf{b}_k^T \mathbf{p})^2 - |\mathbf{g}_k^H \mathbf{q} + \bar{\mathbf{g}}_k^H \mathbf{p}|^2} \right), \]  

\( \mathbf{a}_k \) and \( \mathbf{b}_k \) are the link coefficients.
where

\[ a_k = \left[ |h_{1k}|^2 (1 + \sigma^2_{\eta_{1k}}) \quad |h_{2k}|^2 (1 + \sigma^2_{\eta_{2k}}) \right]^T, \quad f_k = \left[ h_{1k}^2 \quad h_{2k}^2 \right]^H, \]

\[ \tilde{f}_k = \left[ h_{1k}^2 \tilde{\sigma}^2_{\eta_{1k}} \quad h_{2k}^2 \tilde{\sigma}^2_{\eta_{2k}} \right]^H, \]

\[ b_1 = \left[ |h_{11}|^2 \sigma^2_{\eta_{11}} \quad |h_{21}|^2 (1 + \sigma^2_{\eta_{21}}) \right]^T, \quad g_1 = \left[ 0 \quad h_{21}^2 \right]^H, \]

\[ b_2 = \left[ |h_{12}|^2 (1 + \sigma^2_{\eta_{12}}) \quad |h_{22}|^2 \sigma^2_{\eta_{22}} \right]^T, \quad g_2 = \left[ h_{12}^2 \quad 0 \right]^H, \]

\[ q = \left[ q_1 \quad q_2 \right]^T, \quad p = \left[ p_1 \quad p_2 \right]^T. \]

We also define \( \Omega = \{ p_k, q_k : 0 \leq p_k \leq P_k, |q_k| \leq p_k, k = 1, 2 \} \) as the feasible set of the design parameters, where \( P_k \) is the power budget of user \( k \).

In this paper, we aim at obtaining the boundary of the achievable rate region for the described two-user IC. To this end, we employ the following definition of the Pareto boundary for the achievable rate region.

**Definition 1** ([16], [29]). The rate pair \((R_1, R_2)\) is called Pareto-optimal if \((R_1', R_2)\) and \((R_1, R_2')\), with \( R_1' > R_1 \) and \( R_2' > R_2 \), are not achievable.

The rate region is the union of all these achievable rate tuples, i.e., \( \mathcal{R} = \bigcup_{\{p,q\} \in \Omega} (R_1, R_2) \), and its boundary can be derived by the rate profile technique as in the optimization problem [12]

\[
\begin{align*}
\text{maximize} & \quad R \\
\text{s.t.} & \quad R_k \geq \alpha_k R, \quad k = 1, 2, \\
& \quad 0 \leq p_k \leq P_k, \quad k = 1, 2, \\
& \quad |q_k| \leq p_k, \quad k = 1, 2,
\end{align*}
\]

where \( \alpha_1, \alpha_2 \geq 0 \) are fixed and \( \alpha_1 + \alpha_2 = 1 \). We can obtain the boundary of the rate region by solving (10) for different rate-profile parameters, i.e., \( \alpha_1 \) and \( \alpha_2 \). The optimization problem (10)
can also be written as a minimum weighted rate maximization by eliminating variable $R_k$ as

\[
\begin{align*}
\text{maximize}_{p,q} & \quad \min_{k=1,2} \frac{R_k}{\alpha_k} \\
\text{s.t.} & \quad (10c), (10d).
\end{align*}
\] (11a)
\[(11b)

III. BOUNDARY OF THE RATE REGION BY FRACTIONAL PROGRAMMING

In this section, we propose an iterative algorithm based on fractional programming and the well-known generalized Dinkelbach algorithm to obtain a suboptimal solution, operating as closely as possible to the Pareto boundary of the rate region. To this end, we first rewrite the rate profile problem as an energy profile in Section III.A, which allows us to apply FP techniques. Then, we derive the energy region in Section III.B. The solution of the energy region yields the achievable rate region as will be discussed below.

A. Energy profile

In this subsection, we rewrite (10) and (11) such that they are more suitable to be solved with fractional programming. To this end, we employ the energy profile technique in [30], [31] to write an optimization problem that results in the solution of (10). We define the energy profile as

\[
\begin{align*}
\text{maximize}_{E,p,q} & \quad E \\
\text{s.t.} & \quad E_k(p, q) \geq 1 + \alpha_k E, \quad k = 1, 2, \\
& \quad 0 \leq p_k \leq P_k, \quad k = 1, 2, \\
& \quad |q_k| \leq p_k, \quad k = 1, 2,
\end{align*}
\] (12a)
\[(12b)
\[(12c)
\[(12d)

where $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$ are constants, $\alpha_1 + \alpha_2 = 1$, and

\[
E_k(p, q) \triangleq \frac{(\sigma^2 + a_k^T p)^2 - |f_k^H q + \tilde{f}_k^H p|^2}{\left(\sigma^2 + b_k^T p\right)^2 - |g_k^H q + \tilde{f}_k^H p|^2},
\] (13)
We can derive the boundary of the energy region by varying $\alpha_1 \in [0, 1]$. Note that $E_k(p, q) \geq 1$ for $k = 1, 2$ since the rate of users is non-negative. In the following lemma, we show that this technique results in the boundary of the rate region in (10).

**Lemma 1** ([30], [31]). Every point in the boundary of the rate region can be derived by the energy profile technique.

*Proof:* Assume there exists a pair $(R_1, R_2)$ on the boundary of the achievable rate region that is not on the boundary of the energy region. In other words, the pair $(E_1 = 2^{R_1}, E_2 = 2^{R_2})$, which is a feasible energy pair, is not on the boundary of the energy region, and hence there exist $E'_1$ and/or $E'_2$ such that the pairs $(E'_1 > E_1, E_2)$ and/or $(E_1, E'_2 > E_2)$ are feasible. Since the logarithm functions are monotonically increasing, the rate pairs $(0.5 \log_2(E'_1) > R_1, R_2)$ or $(R_1, 0.5 \log_2(E'_2) > R_2)$ are achievable, which implies that $(R_1, R_2)$ is not on the boundary of the rate region.

Note that, similar to (10), we can rewrite (12) as the following maximin optimization problem by removing the variable $E$

\[
\text{maximize } \min_{p, q} \left\{ \frac{E_k(p, q) - 1}{\alpha_k} \right\} \quad (14a)
\]

\[0 \leq p_k \leq P_k, \quad k = 1, 2, \quad (14b)\]

\[|q_k| \leq p_k, \quad k = 1, 2. \quad (14c)\]

**B. Dinkelbach’s algorithm**

In this subsection, we solve the energy profile problem in (12) by FP and the generalized Dinkelbach algorithm [26]–[28]. Dinkelbach’s algorithm is a powerful tool that solves FP problems, which was proposed to handle single-ratio functions. The generalized Dinkelbach algorithm is a modified Dinkelbach algorithm to solve maximin multiple-ratio problems [25]. The generalized
Dinkelbach algorithm is an iterative approach, in which the fractional functions are approximated by surrogate functions at each iteration. In the following lemma, we present some conditions that are used in the generalized Dinkelbach algorithm.

Lemma 2 ([25], [26]). Consider the fractional functions \( u_i(x) \) and \( v_i(x) \) are continuous in \( x \), \( v_i(x) \) is strictly positive in \( x \), and \( x \) is a vector with dimension \( n \) that belongs to a compact set \( \mathcal{X} \). Let us define

\[
V(\mu) = \max_x \min_i (u_i(x) - \mu v_i(x)),
\]

\[
\bar{\mu} = \max_x \min_i \left( \frac{u_i(x)}{v_i(x)} \right),
\]

where \( V(\mu) \), \( \bar{\mu} \), and \( \mu \) are real and scalar, and have the following properties.

1) \( V(\mu) \) is continuous and strictly decreasing in \( \mu \).

2) The optimization problems (15) and (16) always have optimal solutions.

3) \( \bar{\mu} \) is finite and \( V(\bar{\mu}) = 0 \).

4) \( V(\mu) \) has a unique root, and \( V(\mu) = 0 \) implies \( \mu = \bar{\mu} \).

The generalized Dinkelbach algorithm employs the surrogate function \( V(\mu) \) in (15) and tries to iteratively find the unique root of \( V(\mu) \), i.e., \( \bar{\mu} \). The algorithm starts with an initial point, e.g., \( \mu^{(0)} = 0 \), then it updates \( \mu \) to obtain \( \bar{\mu} \). Assuming \( u_i(x) \geq 0 \), which is the case we consider in this paper, \( V(0) = \max_x \min_i (u_i(x)) > 0 \). Since \( V(\mu) \) is continuous and strictly decreasing in \( \mu \), \( \mu \) is chosen monotonically increasing at each iteration \( (\mu^{(l)}) > (\mu^{(l-1)}) \) until \( V(\mu) \) approaches 0. At the \( l \)th iteration, \( \mu^{(l)} \) is

\[
\mu^{(l)} = \min \left( \frac{u_1(x^{(l-1)})}{v_1(x^{(l-1)})}, \frac{u_2(x^{(l-1)})}{v_2(x^{(l-1)})} \right) > 0,
\]

where \( x^{(l-1)} \) is

\[
x^{(l-1)} = \arg \max_x \min_i \left( u_i(x) - \mu^{(l-1)} v_i(x) \right).
\]

The generalized Dinkelbach algorithm updates \( \mu^{(l)} \) and \( x^{(l-1)} \) based on (17) and (18), respectively,
until a convergence metric is met, e.g., $V(\mu^{(l)}) < \epsilon$, where $\epsilon > 0$. This algorithm converges linearly to the optimal solution $[25]$.

Note that in order to apply the generalized Dinkelbach algorithm, it is not required that $u_i(x)$ and $v_i(x)$ fulfill any other condition (except those in the lemma), which makes this algorithm a powerful tool to solve different types of fractional problems. If $u_i(x)$ and $v_i(x)$ are concave and convex functions, respectively, the optimization problem at each iteration is convex and can easily be solved. However, in the general case, it might be difficult to efficiently solve the optimization problem at each iteration.

We can apply the generalized Dinkelbach algorithm to derive the boundary of the energy region since the optimization problem can be written as a maximin weighted problem as indicated in [14]. In order to solve the energy profile problem in (12) by fractional programming, we introduce the following functions, which are the corresponding surrogate functions of $\frac{E_{k-1}}{\alpha_k}$ for $k = 1, 2$:

$$E_k(p, q, \mu) \triangleq (\sigma^2 + a_k^T p)^2 - |f_k^H q + \tilde{f}_k^H p|^2 - (\mu \alpha_k + 1) \left( (\sigma^2 + b_k^T p)^2 - |g_k^H q + \tilde{g}_k^H p|^2 \right),$$ (19)

where $\mu \in \mathbb{R}$. By substituting (19) in (12), the optimization problem at each iteration is

$$\text{maximize} \quad E' \quad \text{s.t.} \quad \tilde{E}_k(p, q, \mu) \geq E', \quad k = 1, 2, \quad 0 \leq p_k \leq P_k, \quad k = 1, 2, \quad |q_k| \leq p_k, \quad k = 1, 2.$$ (20a)

Unfortunately, (20) is not a convex optimization problem since $\tilde{E}_k(p, q, \mu)$ for $k = 1, 2$ is not concave in the optimization variables [32]. Further below, we describe how to solve (20), which is not a convex problem. At the $l$th iteration, we fix $\mu$ as

$$\mu = \min \left( \frac{E_1(p^{(l-1)}, q^{(l-1)}) - 1}{\alpha_1}, \frac{E_2(p^{(l-1)}, q^{(l-1)}) - 1}{\alpha_2} \right)$$ (21)
Algorithm I Generalized Dinkelbach algorithm.

Initialization
Set $\epsilon$, $L$, $\mu = 0$, $l = 0$
Compute $\tilde{E}_k(p, q, \mu)$ for $k = 1, 2$ by (19)

While $\min_{k=1,2} \{\tilde{E}_k(p, q, \mu)\} \geq \epsilon$ and $l \leq L$ do
    $l = l + 1$
    Obtain $p^{(l)}$ and $q^{(l)}$ by solving (20), i.e., run Algorithm II
    If $\min_{k=1,2} \{\tilde{E}_k(p, q, \mu)\} < \epsilon$
        $p^* = p^{(l)}$ and $q^* = q^{(l)}$
    Else
        Update $\mu$ by (21)
    End(If)
End (While)
Return $p^*$ and $q^*$.

and solve (20) for the given $\mu$, which results in $p^{(l)}$ and $q^{(l)}$. Then, we update $\mu$ and repeat the procedure until a convergence metric is met. We summarize this procedure in Algorithm I.

In order to solve (20), we employ the difference of convex programming (DCP), which is a special case of sequential convex programming (SCP). DCP can be applied to non-convex problems where the objective function and/or the constraints are the difference of two convex, or equivalently concave, functions [33]–[37]. The optimization problem in (20) can be efficiently solved by DCP since $\tilde{E}_k(p, q, \mu)$ is the difference of two convex/concave functions:

$$\tilde{E}_k(p, q, \mu) = -|f_k^H q + \tilde{f}_k^H p|^2 - (\mu \alpha_k + 1)(\sigma^2 + b_k^T p)^2$$

$$+ (\mu \alpha_k + 1)|g_k^H q + \tilde{g}_k^H p|^2 + (\sigma^2 + a_k^T p)^2.$$  \hspace{1cm} (22)

The convex part in (22) makes (20) non-convex and has to be approximated by a concave function [32]. This procedure is also known as convex-concave procedure (CCP), in which the convex part is approximated as an affine function by the first-order approximation of the Taylor expansion. Note that we take the first-order term and employ an affine approximation since an affine function is the nearest concave approximation to a convex function. The first-order approximation of a real
function $u(x)$ around the point $x_0$ is obtained through its Taylor expansion as \[20\], \[38\]

$$u(x) \approx u(x_0) + 2\Re \left[ \left. \frac{\partial u(x)}{\partial x} \right|_{x=x_0} \right]^T (x - x_0), \quad (23)$$

where $x$ is a complex vector, and $\Re \{ x \}$ takes the real part of $x$. In order to apply the CCP to \(22\), we have to differentiate the convex part in \(22\) with respect to $p$ and $q$. The derivative of $(\sigma^2 + a_k^T p)^2$ with respect to $p$ is straightforward since it is a real function on a real domain and consequently, analytic in $p$. The derivative of $(\sigma^2 + a_k^T p)^2$ with respect to $p$ is

$$\frac{\partial (\sigma^2 + a_k^T p)^2}{\partial p} = 2a_k(\sigma^2 + a_k^T p), \quad (24)$$

and the resulting first-order approximation around the power vector in the $m$th iteration, $p^{(m)}$, is given by

$$(\sigma^2 + a_k^T p)^2 \simeq (\sigma^2 + a_k^T p^{(m)})^2 + 2(\sigma^2 + a_k^T p^{(m)})a_k^T (p - p^{(m)}). \quad (25)$$

The derivative of $|g_k^H q + \tilde{f}_k^H p|^2$ with respect to $p$ is also straightforward since it is analytic in $p$:

$$\frac{\partial |g_k^H q + \tilde{f}_k^H p|^2}{\partial p} = 2\Re \left[ \tilde{f}_k (g_k^H q + \tilde{f}_k^H p) \right]. \quad (26)$$

The term $|g_k^H q + \tilde{f}_k^H p|^2$, on the other hand, is not analytic in $q$ since it is a real-valued function while $q$ is a complex vector \[20\], \[38\]. Thus, we have to employ Wirtinger calculus to obtain the derivative of $|g_k^H q + \tilde{f}_k^H p|^2$ with respect to $q$. By Wirtinger calculus, we treat $q$ and $q^*$ as two independent complex variables \[20\], \[38\]. Thus, we take the derivative of $|g_k^H q + \tilde{f}_k^H p|^2$ with respect to $q$ while treating $q^*$ as a constant, which results in

$$\frac{\partial |g_k^H q + \tilde{f}_k^H p|^2}{\partial q} = g_k^* (g_k^H q + \tilde{f}_k^H p)^*. \quad (27)$$
Now by (23), we can approximate \(|g_k^H q + \tilde{f}_k^H p|^2\) as an affine function as

\[
|g_k^H q + \tilde{f}_k^H p|^2 \simeq |g_k^H q^{(m)} + \tilde{f}_k^H p^{(m)}|^2 + 2\Re \left( \tilde{f}_k^H (g_k^H q^{(m)} + \tilde{f}_k^H p^{(m)})^* (p - p^{(m)}) \right) \\
+ 2\Re \left( (g_k^H q^{(m)} + \tilde{f}_k^H p^{(m)})^* g_k^H (q - q^{(m)}) \right).
\]

(28)

By substituting (25) and (28) in (22), we have

\[
\tilde{E}_k(p, q, \mu) \approx \hat{E}_k(p, q, \mu) = -|f_k^H q + \tilde{f}_k^H p|^2 - (\mu \alpha_k + 1)(\sigma^2 + b_k^T p)^2 \\
+ (\sigma^2 + a_k^T p^{(m)})^2 + 2(\sigma^2 + a_k^T p^{(m)})a_k^T (p - p^{(m)}) \\
+ (\mu \alpha_k + 1)|g_k^H q^{(m)} + \tilde{f}_k^H p^{(m)}|^2 + 2(\mu \alpha_k + 1)\Re \left[ \tilde{f}_k^H (g_k^H q^{(m)} + \tilde{f}_k^H p^{(m)})^* (p - p^{(m)}) \right] \\
+ 2(\mu \alpha_k + 1)\Re \left[ (g_k^H q^{(m)} + \tilde{f}_k^H p^{(m)})^* g_k^H (q - q^{(m)}) \right].
\]

(29)

Now, we are able to solve (20) by DCP, which is also an iterative algorithm. In the \((m + 1)\)th iteration, \(\tilde{E}_k(p, q, \mu)\) in (22) is approximated by \(\hat{E}_k(p, q, \mu)\) in (29) using the solution \(p^{(m)}\) and \(q^{(m)}\) from the previous iteration. The resulting convex optimization problem, which is an approximation of (20) in the \((m + 1)\)th iteration, is

\[
\begin{align*}
\text{maximize} & \quad E' \\
\text{s.t.} & \quad \hat{E}_k(p, q, \mu) \geq E', \quad k = 1, 2, \\
& \quad (20c), (20d).
\end{align*}
\]

(30)

We obtain \(p^{(m+1)}\) and \(q^{(m+1)}\) as the solution of (30) and repeat the procedure until a convergence criterion is met. This procedure is summarized in Algorithm II. Note that, since \(\tilde{E}_k(p, q, \mu) \geq \hat{E}_k(p, q, \mu)\) for all \(p\) and \(q\) at each iteration, no trust region is required [33], [34]. The DCP algorithm converges to a stationary point\(^2\) of the original non-convex problem [33]–[37].

\(^2\)A stationary point of a constrained optimization problem satisfies the corresponding Karush-Kuhn-Tucker (KKT) conditions [36].
Algorithm II Proposed DCP.

Initialization
Set $\epsilon$, $M$, $p^{(0)} = 0$, $q^{(0)} = 0$, $m = 1$, convergence=0

While convergence=0 and $m \leq M$ do
     Construct $\hat{E}_k$ for $k = 1, 2$ by (29)
     Obtain $p^{(m+1)}$ and $q^{(m+1)}$ by solving (30)
     If $\|p^{(m)} - p^{(m+1)}\|/\|p^{(m)}\| < \epsilon$ and $\|q^{(m)} - q^{(m+1)}\|/\|q^{(m)}\| < \epsilon$
           convergence=1
           $p^* = p^{(m+1)}$ and $q^* = q^{(m+1)}$
     End(If)
     $m = m + 1$
End (While)
Return $p^*$ and $q^*$.

To sum up, the proposed algorithm works as follows. We rewrite the rate profile as an energy profile and solve it using the generalized Dinkelbach algorithm. In each iteration of the generalized Dinkelbach algorithm, we have to solve a non-convex optimization problem. We solve this problem by DCP, which is also an iterative solver. In other words, at every iteration of the generalized Dinkelbach algorithm, we perform another iterative algorithm to solve the inner problem of the generalized Dinkelbach algorithm, i.e., problem (20).

IV. SIMPLIFIED ALGORITHM

In this section, we propose a simplified version of the algorithm from Section III which exhibits a lower computational complexity. In the simplified algorithm, we first optimize the transmission power $p$ for PGS, i.e., for $q = 0$. This problem is addressed in Section IV-A Then, in Section IV-B, we optimize the complementary variances for the resulting transmit power $p$ such that the rates of all users is simultaneously increased.
A. Transmission power design

In this subsection, we optimize the transmission power vector $p$ for PGS, i.e., when $q = 0$. In this case, deriving the boundary of the energy region can be cast as the optimization problem

$$\begin{align*}
\max_{E, p} & \quad E \\
s.t. & \quad \frac{(\sigma^2 + a_i^T p)^2 - |\tilde{f}_i^H p|^2}{(\sigma^2 + b_i^T p)^2 - |\tilde{f}_i^H p|^2} \geq 1 + \alpha_k E, \quad k = 1, 2, \quad (31a) \\
& \quad 0 \leq p_k \leq P_k, \quad k = 1, 2, \quad (31b)
\end{align*}$$

for $\alpha_1, \alpha_2 \geq 0$ and $\alpha_1 + \alpha_2 = 1$. Unfortunately, the optimization problem in (31) is not convex due to (31b). In the following lemma, we derive a lower bound for (31b), which allows us to simplify (31) and derive a low-complexity algorithm.

**Lemma 3.** A lower bound for (31b) is

$$\frac{(\sigma^2 + a_i^T p)^2 - |\tilde{f}_i^H p|^2}{(\sigma^2 + b_i^T p)^2 - |\tilde{f}_i^H p|^2} \geq \frac{(\sigma^2 + a_i^T p)^2}{(\sigma^2 + b_i^T p)^2}, \quad (32)$$

where the equality in (32) holds if and only if the HWI noise is proper, i.e., $\tilde{f}_i = 0$.

**Proof:** It is easy to verify that $0 \leq |\tilde{f}_i^H p|^2 < (\sigma^2 + b_i^T p)^2 < (\sigma^2 + a_i^T p)^2$. Let us define

$$f(t) = \frac{\beta_1 - t}{\beta_2 - t}, \quad (33)$$

where $0 \leq t < \beta_2 < \beta_1$. The lower bound in (32) is then satisfied if $f(t)$ is increasing in $t$. This function is strictly increasing in $t \in [0, \beta_2)$ since

$$\frac{\partial f(t)}{\partial t} = \frac{\beta_1 - \beta_2}{(\beta_2 - t)^2} > 0, \quad (34)$$

Thus, we have

$$\frac{\beta_1 - t}{\beta_2 - t} \geq \frac{\beta_1}{\beta_2}, \quad (35)$$
with equality if and only if $t = 0$.

For each point characterized by $\alpha_1$ and $\alpha_2$, we solve (31) for the lower bound in (32) as the optimization problem

$$\begin{align*}
\text{maximize} & \quad E \\
\text{s.t.} & \quad \frac{\sigma^2 + a_k^T p}{\sigma^2 + b_k^T p} \geq \sqrt{1 + \alpha_k E}, \quad k = 1, 2, \\
& \quad 0 \leq p_k \leq P_k, \quad k = 1, 2. \tag{36b}
\end{align*}\tag{36}$$

Note that the region achieved by (31) includes the region achieved by (36). If the HWI noise is proper, (36) is equivalent to (31). The optimization problem in (36) can be solved by a sequence of feasibility problems. That is, we fix $E$ as $E'$ and consider the feasibility problem

$$\begin{align*}
\text{find} & \quad p \in \mathbb{R}^2, \\
\text{s.t.} & \quad (a_k^T - \sqrt{1 + \alpha_k E' b_k^T}) p \geq (\sqrt{1 + \alpha_k E'} - 1) \sigma^2, \quad k = 1, 2, \tag{37b} \\
& \quad 0 \leq p_k \leq P_k, \quad k = 1, 2. \tag{37c}
\end{align*}\tag{37}$$

If (37) is feasible for a given $E'$, the optimal solution of (36) is greater than or equal to $E'$, i.e., $E^* \geq E'$. Otherwise, $E^* < E'$. To this end, we employ the well-known bisection method over $E'$ solving (37) at each iteration, which yields, upon convergence, the optimal solution of (36) [32]. Constraints (37b) and (37c) are linear in $p$, which permits deriving a closed-form expression for a feasible point, as presented in the following theorem.

**Theorem 1.** The optimization problem in (37) is feasible for a given $E'$ if and only if $0 \leq p_k' \leq P_k$.
for $k = 1, 2$, where

\[
\begin{bmatrix}
p_1' \\
p_2'
\end{bmatrix} = A^{-1} \begin{bmatrix}
(\sqrt{1 + \alpha_1 E'} - 1)\sigma^2 \\
(\sqrt{1 + \alpha_2 E'} - 1)\sigma^2
\end{bmatrix},
\]
(38)

\[
A = \begin{bmatrix}
a_1^T - \sqrt{1 + \alpha_1 E'} b_1^T \\
a_2^T - \sqrt{1 + \alpha_2 E'} b_2^T
\end{bmatrix}
\]

\[
= \begin{bmatrix}
|h_{11}|^2 \left(1 - \sigma_{\eta_{11}}^2 (\sqrt{1 + \alpha_1 E'} - 1)\right) - |h_{21}|^2 \left(1 + \sigma_{\eta_{21}}^2 (\sqrt{1 + \alpha_1 E'} - 1)\right) \\
|h_{12}|^2 \left(1 + \sigma_{\eta_{12}}^2 (\sqrt{1 + \alpha_2 E'} - 1)\right) - |h_{22}|^2 \left(1 - \sigma_{\eta_{22}}^2 (\sqrt{1 + \alpha_2 E'} - 1)\right)
\end{bmatrix}.
\]
(39)

Proof: Please refer to Appendix A.

B. Complementary variance design

In this subsection, we optimize the complementary variances $q$ for a given $p^*$, which has been obtained by solving (36). We obtain $q$ such that the rates of both users exceed the rates achieved by PGS, which are the rates that are achievable with $q = 0$ and the power vector $p^*$ obtained by solving (36). In other words, we want to solve the following optimization problem

\[
\begin{align*}
\text{maximize} & \quad t, q \\
\text{s.t.} & \quad \frac{(\sigma^2 + a_k^T p^*)^2 - |f_k^H q + \tilde{f}_k^H p^*|^2}{(\sigma^2 + b_k^T p^*)^2 - |g_k^H q + \tilde{f}_k^H p^*|^2} \geq E_{p,k} + \alpha_k t, \quad k = 1, 2, \\
& \quad |q_k| \leq p_k^*, \quad k = 1, 2,
\end{align*}
\]
(40)

where $p_k^*$ is the $k$th element of $p^*$. Moreover, $E_{p,k}$ is fixed and given by

\[
E_{p,k} = \frac{(\sigma^2 + a_k^T p^*)^2 - |f_k^H p^*|^2}{(\sigma^2 + b_k^T p^*)^2 - |f_k^H p^*|^2}.
\]
(41)

Unfortunately, (40) is not convex due to (40b). Hence, in order to efficiently solve (40), we first rewrite (40b) as

\[
\frac{(\sigma^2 + a_k^T p^*)^2 - |f_k^H q + \tilde{f}_k^H p^*|^2 - \alpha_k t_k}{(\sigma^2 + b_k^T p^*)^2 - |g_k^H q + \tilde{f}_k^H p^*|^2} \geq E_{p,k},
\]
(42)
where \( t_k = t \left[ (\sigma^2 + b^T_k p^*)^2 - |g^H_k q + \tilde{f}^H_k p^*|^2 \right] \). We then relax the relation between \( t_1 \) and \( t_2 \) and approximate (40) as

\[
\text{maximize}_{t_1, t_2, q} \quad \min(t_1, t_2)
\]

\[
\text{s.t.} \quad \frac{(\sigma^2 + a^T_k p^*)^2 - |f^H_k q + \tilde{f}^H_k p^*|^2}{(\sigma^2 + b^T_k p^*)^2 - |g^H_k q + \tilde{f}^H_k p^*|^2} \geq E_{p,k}, \quad k = 1, 2,
\]

\[
|q_k| \leq p_k, \quad k = 1, 2. \tag{43c}
\]

If \( \min(t_1, t_2) > 0 \), the rates of both users are simultaneously increased by employing IGS. Otherwise, we set \( q = 0 \) and employ PGS. Note that the constraint (43b) can be rewritten as

\[
E_{p,k} |g^H_k q + \tilde{f}^H_k p^*|^2 - |f^H_k q + \tilde{f}^H_k p^*|^2 + (\sigma^2 + a^T_k p^*)^2 - E_{p,k} (\sigma^2 + b^T_k p^*)^2 \geq \alpha_k t_k, \tag{44}
\]

which is a difference of two convex functions. Thus, (43) is not a convex optimization problem, but it can be efficiently solved by difference of convex programming and a convex-concave procedure similar to (20) [33]–[37]. Hence, we employ DCP and solve (43) iteratively. At each iteration, we approximate the left-hand side of (44) by a concave function. To this end, we employ the first-order Taylor expansion and approximate the convex part of (44) around the point \( q^{(l)} \) by an affine function as

\[
|g^H_k q + \tilde{f}^H_k p^*|^2 \approx |g^H_k q^{(l)} + \tilde{f}^H_k p^*|^2 + 2 \Re \left( (g^H_k q^{(l)} + \tilde{f}^H_k p^*)^* g^H_k (q - q^{(l)}) \right), \tag{45}
\]

where \( q^{(l)} \) contains the complementary variances of the users in the \( l \)th iteration. It is worth mentioning that \( |g^H_k q + \tilde{f}^H_k p^*|^2 \) is always greater than or equal to the right-hand side of (45), and consequently, no trust region is required in DCP as pointed out before [33]–[35]. Finally, in the \( l \)th each iteration, (44) can be approximated by

\[
- |f^H_k q + \tilde{f}^H_k p^*|^2 + 2E_{p,k} \Re \left( (g^H_k q^{(l)} + \tilde{f}^H_k p^*)^* g^H_k (q - q^{(l)}) \right) + E_{p,k} |g^H_k q^{(l)} + \tilde{f}^H_k p^*|^2
\]

\[
+ (\sigma^2 + a^T_k p^*)^2 - E_{p,k} (\sigma^2 + b^T_k p^*)^2 \geq \alpha_k t_k. \tag{46}
\]
Finally, the convex optimization problem in the $l$th iteration is

$$\begin{align*}
\text{maximize} \quad & t_1, t_2, q \\
\text{subject to} \quad & (46), \ (43c), \ (47b)
\end{align*}$$

This problem can be easily solved by standard numerical tools [32]. Moreover, the proposed DCP algorithm converges to a stationary point of (43) [33]–[37].

The proposed simplified algorithm can be summarized as follows. The joint optimization problem for $p$ and $q$ is decoupled into two separate optimization problems. We derive the transmission powers by employing the well-known bisection method, which results, in each iteration, in a feasibility problem that has a closed-form solution. Then, we employ the DCP algorithm to derive the complementary variances for the given transmission powers.

V. NUMERICAL RESULTS

In this section, we present some numerical results to illustrate our findings. For all examples, we set $\sigma^2 = 1$, $P_1 = P_2 = P$, $\epsilon = 10^{-4}$, and $L = M = 20$, where $\epsilon$, $L$, and $M$ are, respectively, the threshold for convergence, and the maximum number of iterations for Algorithms I and II. Moreover, the maximum number of iterations for the algorithm in Section IV-B is 40. We also define the signal-to-noise ratio (SNR) as the ratio of the power budget to $\sigma^2$, i.e., $\text{SNR} = \frac{P}{\sigma^2}$. We compare our proposed algorithms with PGS and the joint variance and complementary variance optimization algorithm in [12] for IGS, which is designed for ideal devices. To the best of our knowledge, there exists no PGS algorithm for asymmetric HWI in the literature. Because of that, we optimize the PGS scheme by using the first step of our simplified algorithm (see Section IV-A).

In the figures, we use the following labels:

- **S-IGS**: our proposed simplified design in Section IV,
- **FP-IGS**: our proposed design with FP in Section III,
- **PGS**: the proposed PGS design in Section IV-A,
- **I-IGS**: the joint variance and complementary variance IGS design in [12] for ideal devices.
A. Ideal devices

In this subsection, we compare the performance of our proposed algorithms with the IGS algorithm in [12] when there is no HWI. The IGS algorithm in [12] is an iterative algorithm, based on a bisection method over the minimum weighted rates of users, and is proposed for ideal devices. The algorithm employs semidefinite relaxation (SDR) programming in order to solve the corresponding feasibility problem at each iteration of the bisection method. The main difference of our proposed algorithm based on FP and the algorithm in [12], for ideal devices, is the approach to solve the corresponding optimization problems at each iteration, where we employ DCP instead of SDR. Note also that our proposed algorithms are more general since they consider asymmetric HWI, while the algorithm in [12] can only be applied for ideal devices.

In Fig. 2, we show the average symmetric rate, i.e., the minimum rate allocated to the users, which is the fairness point of the rate region boundary and obtained by $\alpha_1 = \alpha_2 = 0.5$. We average the results over 100 channel realizations, where each channel realization is taken from a complex proper Gaussian distribution with variance 1, i.e., $\mathcal{CN}(0, 1, 0)$. As can be observed, our proposed algorithm based on FP outperforms the proposed algorithm in [12], especially at high SNR. Our simplified algorithm performs similarly to the proposed algorithm in [12] for low SNR. However, the algorithm in [12] performs better than the simplified algorithm in the moderate SNR regime.
The reason is that the benefit of employing IGS increases with SNR. Thus, the performance differences of the IGS algorithms are clearer at higher SNR. Although the symmetric rate of the users for our simplified algorithm is not better than the algorithm in [12] at moderate SNR, our simplified algorithm is much faster than the other algorithms, as we will show in Section V.C.

In Fig. 3, we also provide rate region examples for ideal devices and the channel realization

\[ H_1 = \begin{bmatrix} 1.4070e^{10.2721} & 0.9288e^{11.8320} \\ 0.9288e^{11.8320} & 1.7367e^{11.1136} \end{bmatrix}, \] (48)

where \([H_1]_{ij} = h_{ij}\) for \(i, j \in \{1, 2\}\). In this channel realization, the direct links are stronger than the interference links, which can be considered as moderate interference. As can be observed, IGS can enlarge the achievable rate region for this channel realization and \(P = 10\). Since the benefits of IGS are minor for low SNR, IGS does not provide any gain for \(P = 1\). This is also in line with the averaging results in Fig. 2 where IGS has minor benefits at low SNR, while it improves the performance of the system significantly at moderate SNR. For this channel realization, our proposed algorithms and the algorithm in [12] perform very closely to each other.
B. Non-ideal devices

In this subsection, we consider the effect of HWI on the performance of the two-user IC. Throughout this subsection, we consider the same statistics for HWI in all devices. In Fig. 4, we show the rate region for $H_1$ and $P = 1$ under maximally improper HWI noise. As shown in Fig. 3, IGS brings negligible gains when the transceivers are ideal, but, as observed in Fig. 4, IGS can significantly enlarge the rate region if there is asymmetric HWI. Note that even in point-to-point communications, PGS is in general suboptimal for asymmetric HWI, as it is shown in Fig. 4 for either $R_1 = 0$ or $R_2 = 0$.

In Fig. 5, we show the achievable rate region for $\hat{\sigma}^2_\eta = 0$, $P = 10$ (SNR = 10 dB), and channel realization

$$H_2 = \begin{bmatrix}
0.3764 e^{i1.4381} & 0.4029 e^{i0.9486} \\
1.8542 e^{i2.8153} & 0.6277 e^{i2.3697}
\end{bmatrix}. \tag{49}$$

In this channel realization, one of the interference links is much stronger than the direct links, which can be considered as high interference. We take $\hat{\sigma}^2_\eta = 0$, i.e., symmetric (proper) HWI. We can observe that IGS enlarges the rate region even for proper HWI with high noise variance, i.e., $\sigma^2_\eta = 0.5$ and $\sigma^2_\eta = 1$.

$^3$Maximally improper HWI happens when the in-phase and quadrature-phase noises are completely correlated [24].
In the following, we provide some averaged results for different parameters to illustrate different aspects of employing IGS. Similar to Fig. 2, we average the results over 100 channel realizations, where each channel realization is taken from a complex proper Gaussian distribution with variance 1, i.e., $CN(0, 1, 0)$.

In Fig. 6, we consider the effect of the variance of the HWI noise on the average symmetric rate of users ($\alpha_1 = \alpha_2 = 0.5$) for $P = 20$. In this figure, we consider proper ($\tilde{\sigma}^2_{\eta} = 0$) and maximally improper ($\tilde{\sigma}^2_{\eta} = \sigma^2_{\eta}$) HWI noise. We observe that our proposed algorithm with FP outperforms the other algorithms for maximally improper HWI noise. Moreover, our proposed IGS algorithms
always perform better than PGS for proper HWI noise. Furthermore, our simplified algorithm outperforms the IGS algorithm in [12] in the presence of HWI, exhibiting also a low computational complexity as we will discuss in Section V-C. However, the performance improvement by our algorithms is minor for proper HWI with high noise variance, where our algorithms only provide 5% improvement over PGS when $\sigma^2 = 1$.

Figure 7 shows the effect of the circularity coefficient of the HWI noise on the symmetric rate for $P = 20$. As can be observed, the benefits of employing IGS increase with the circularity coefficient of the HWI noise, and there is a considerable performance improvement by IGS in maximally improper HWI noise. Our proposed IGS design with FP outperforms the other algorithms, especially in highly asymmetric HWI noise. When the variance of the HWI noise is small, the gain of employing IGS is larger. The other interesting result in this figure is that our simplified algorithm performs very similarly to our proposed algorithm based on FP for proper HWI. Since the simplified algorithm is much faster, it can be employed for proper HWI noise when the variance of the HWI noise is high, i.e., $\sigma^2 \geq 0.5$. However, our proposed algorithm based on FP outperforms the other algorithm in low-power HWI noise and/or highly asymmetric HWI noise. Note that, since the IGS algorithm in [12] is proposed for ideal devices and does not consider HWI, it performs worse than the proposed PGS, which considers symmetric HWI, from
In Fig. 8, we consider the effect of the power budget on the symmetric rate of users. There is an almost constant performance gap between our proposed algorithms and the other algorithms. Similar to the other figures, our proposed IGS with FP outperforms our simplified algorithm.

### C. Algorithms complexity

In this subsection, we evaluate the complexity of the algorithms by the average computation time per channel realization on an Intel(R) Core(TM) i7-7560U with 2.4 GHz CPU. In Table I, we provide the average runtime per channel realization in seconds to obtain the results depicted in Fig. 2 for ideal devices. As can be observed, our simplified algorithm is significantly faster than the other algorithms. The computational time of our proposed algorithm with FP increases with the power budget because the algorithm needs more iterations when increasing the power budget. However, it converges faster than the previously proposed algorithm in [12] for low SNR.

---

**TABLE I:** Average runtime in seconds per channel realization to obtain the results depicted in Fig. 2.

| Power budget | 0 dB | 3 dB | 8 dB | 13 dB | 18 dB |
|--------------|------|------|------|-------|-------|
| S-IGS        | 2.58 | 2.66 | 2.55 | 2.75  | 3.20  |
| FP-IGS       | 5.59 | 5.82 | 6.12 | 7.75  | 8.66  |
| I-IGS        | 5.67 | 6.05 | 5.88 | 6.35  | 6.79  |
TABLE II: Average runtime in seconds per channel realization to obtain the results of Fig. 8a.

| Power budget | 3 dB | 8 dB | 13 dB | 18 dB |
|--------------|------|------|-------|-------|
| S-IGS        | 6.19 | 5.82 | 6.30  | 6.42  |
| FP-IGS       | 11.28| 23.86| 28.77 | 38.16 |
| I-IGS        | 6.49 | 8.34 | 7.94  | 8.83  |

In Table II, we present the average runtime per channel realization in seconds for deriving the results in Fig. 8a. As can be observed, our simplified algorithm is significantly faster than our algorithm based on FP. Moreover, the performance of our simplified algorithm is only slightly worse than the algorithm based on FP.

VI. CONCLUSION

In this paper, we considered a two-user IC with asymmetric HWI at the transceivers. Treating interference as noise, we addressed the problem of obtaining the achievable rate region for IGS and proposed two suboptimal algorithms. The first approach is based on FP and the generalized Dinkelbach algorithm. Then, we proposed a simplified algorithm that has lower computational complexity. This simplified algorithm is based on the separate optimization of the powers and complementary variances. We showed that the proposed approaches outperform PGS and existing IGS algorithms, especially as the HWI becomes more asymmetric. Furthermore, we showed that the proposed simplified algorithm exhibits significantly lower computational complexity with only small performance degradation.

ACKNOWLEDGMENTS

The work of M. Soleymani, C. Lameiro and P. J. Schreier was supported by the German Research Foundation (DFG) under grants LA 4107/1-1 and SCHR 1384/8-1. The work of I. Santamaria was supported by MINECO of Spain and AEI/FEDER funds of the E.U., under grant TEC2016-75067-C4-4-R (CARMEN).
APPENDIX A

PROOF OF THEOREM 1

A given \( E' \) is feasible if and only if there exists at least a pair \((p_1, p_2)\) that satisfies all the constraints in (37). Let us first consider the two linear constraints in (37b), which can be written as

\[
|h_{11}|^2 \left( 1 - \sigma^2_{n1} (\sqrt{1 + \alpha_1 E' - 1}) \right) p_1 - |h_{21}|^2 (1 + \sigma^2_{n21}) (\sqrt{1 + \alpha_1 E' - 1}) p_2 \\
\geq (\sqrt{1 + \alpha_1 E' - 1}) \sigma^2,
\]

\[
|h_{12}|^2 (1 + \sigma^2_{n2}) (\sqrt{1 + \alpha_2 E' - 1}) p_1 + |h_{22}|^2 (1 - \sigma^2_{n22}) (\sqrt{1 + \alpha_2 E' - 1}) p_2 \\
\geq (\sqrt{1 + \alpha_2 E' - 1}) \sigma^2.
\]

We can construct \( A \) in (39) by the coefficients of \( p_1 \) and \( p_2 \) in (50) and (51) as

\[
A = \begin{bmatrix}
|h_{11}|^2 \left( 1 - \sigma^2_{n1} (\sqrt{1 + \alpha_1 E' - 1}) \right) & -|h_{21}|^2 (1 + \sigma^2_{n21}) (\sqrt{1 + \alpha_1 E' - 1}) \\
|h_{12}|^2 (1 + \sigma^2_{n2}) (\sqrt{1 + \alpha_2 E' - 1}) & |h_{22}|^2 (1 - \sigma^2_{n22}) (\sqrt{1 + \alpha_2 E' - 1})
\end{bmatrix}.
\]

Since (50) and (51) must hold with equality as we will show later, we first consider equality in (50) and (51). Assuming equality in (50) and (51), these two constraints can be expressed as

\[
Ap = y,
\]

where \( y = \begin{bmatrix} (\sqrt{1 + \alpha_1 E'} - 1) \sigma^2 & (\sqrt{1 + \alpha_2 E'} - 1) \sigma^2 \end{bmatrix}^T \). The problem is then feasible if the solution to (52), which is given by (38), is non-negative and does not exceed the power budget as indicated in Theorem 1. It is worth mentioning that the non-diagonal elements of \( A \) in (39) are non-positive since \( \sqrt{1 + \alpha_1 E'} \geq 1 \) and \( \sqrt{1 + \alpha_2 E'} \geq 1 \). Thus, if the diagonal elements of \( A \) are not positive, there is no positive power pair that satisfies (50) and (51) simultaneously, and (52) gives negative solutions. Thus, in the following, we assume without loss of generality that \( A \) has strictly positive diagonal elements and strictly negative non-diagonal elements.

Now we show that it is sufficient to consider the equality for feasibility of \( E' \). We can rewrite
Fig. 9: The constraints of (37) in the power plane.

(50) and (51) as

\[
\begin{align*}
[A]_{11}p_1 & \geq -[A]_{12}p_2 + y_1, \\
[A]_{22}p_2 & \geq -[A]_{21}p_1 + y_2,
\end{align*}
\]

where \([A]_{ij}\) and \(y_i\) for \(i, j \in \{1, 2\}\) are the \(ij\)th element of \(A\), and the \(i\)th element of \(y\), respectively.

If we decouple the inequalities, we end up with

\[
\begin{align*}
det(A)p_1 & \geq [A]_{22}y_1 - [A]_{12}y_2, \\
det(A)p_2 & \geq -[A]_{21}y_1 + [A]_{11}y_2.
\end{align*}
\]

The right-hand sides (RHS) in (55) and (56) are positive for a feasible \(E'\) as mentioned before. Note that if \(det(A) < 0\), there are no positive power pairs that satisfy (50) and (51) for the given structure of \(A\) in (39), and (52) gives negative solutions. Thus, we consider \(det(A) > 0\), which
yields

\[
\begin{align*}
    p_1 & \geq p_1' = \frac{[A]_{22}y_1 - [A]_{12}y_2}{\det(A)}, \\
    p_2 & \geq p_2' = -\frac{[A]_{21}y_1 + [A]_{11}y_2}{\det(A)},
\end{align*}
\]

(57) \quad (58)

where \( p_1' \) and \( p_2' \) are the intersecting point given in (58). Hence, the intersecting point provides the minimum positive power pairs that satisfy (50) and (51). Figure 9 shows the constraints of the feasibility problem (37) for the case that \( \det(A) > 0 \) and the diagonal elements of \( A \) are strictly positive. It can be easily verified that in the blue region, both constraints in (53) and (56) are simultaneously satisfied. Hence, the minimum power pair \( (p_1, p_2) \) that satisfies both of the constraints in (53) and (56) can be obtained by the intersecting point.

REFERENCES

[1] 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.

[2] 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.

[3] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, “An overview of signal processing techniques for millimeter wave MIMO systems,” IEEE J. Sel. Topics Signal Process., vol. 10, no. 3, pp. 436–453, 2016.

[4] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, “Asymmetric hardware distortions in receive diversity systems: Outage performance analysis,” IEEE Access, vol. 5, pp. 4492–4504, 2017.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] V. R. Cadambe, S. A. Jafar, and C. Wang, “Interference alignment with asymmetric complex signaling–Settling the Host-Madsen-Nosratinia conjecture,” IEEE Trans. Inf. Theory, vol. 56, no. 9, pp. 4552–4565, 2010.
[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

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

[18] 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.

[19] 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.

[20] 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.

[21] 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.

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

[23] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, “Impact of improper Gaussian signaling on hardware impaired systems,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2017, pp. 1–6.

[24] S. Javed, O. Amin, B. Shihada, and M.-S. Alouini, “Improper Gaussian signaling for hardware impaired multihop full-duplex relaying systems,” *IEEE Trans. Commun.*, 2018.

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

[26] 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.

[27] W. Dinkelbach, “On nonlinear fractional programming,” *Management Science*, vol. 13, no. 7, pp. 492–498, 1967.
[28] K. Shen and W. Yu, “Fractional programming for communication systems–Part I: Power control and beamforming,” *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, 2018.

[29] E. A. Jorswieck, E. G. Larsson, and D. Danev, “Complete characterization of the Pareto boundary for the MISO interference channel,” *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5292–5296, 2008.

[30] C. Lameiro, J. Vía, and I. Santamaría, “Amplify-and-forward strategies in the two-way relay channel with analog Tx-Rx beamforming,” *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 642–654, 2013.

[31] I. Santamaria, J. Vía, A. Nazábal, and C. Lameiro, “Capacity region of the multiantenna Gaussian broadcast channel with analog Tx-Rx beamforming,” in *Proc. 5th Int. ICST CHINACOM*, Aug. 2010, pp. 1–6.

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

[33] J. Duchi, “Sequential convex programming,” Lecture Notes, EE364b, Stanford University, 2018.

[34] T. Lipp and S. Boyd, “Variations and extension of the convex–concave procedure,” *Optim. Eng.*, vol. 17, no. 2, pp. 263–287, 2016.

[35] A. L. Yuille and A. Rangarajan, “The concave-convex procedure,” *Neural computation*, vol. 15, no. 4, pp. 915–936, 2003.

[36] 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.

[37] 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.

[38] D. Brandwood, “A complex gradient operator and its application in adaptive array theory,” in *IEE Proc. F Commun., Radar Signal Process.*, vol. 130, no. 1. IET, 1983, pp. 11–16.