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

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 distortions (HWD) happen when the impairments of the I/Q branches are correlated or imbalanced, which in turn generate improper additive interference at the receiver side. When the interference is improper, as well as in other interference-limited scenarios, improper Gaussian signaling (IGS) has been shown to provide rate and/or power efficiency benefits. In this paper, we investigate the rate benefits of IGS in a two-user interference channel (IC) with additive asymmetric HWD when interference is treated as noise. We propose two iterative algorithms to optimize the parameters of the improper transmit signals. We first rewrite the rate region as an pseudo-signal-to-interference-plus-noise-ratio (PSINR) region and employ majorization minimization and fractional programming to find a suboptimal solution for the achievable user rates. 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 additive HWD. Our proposed algorithms outperform proper Gaussian signaling and competing IGS algorithms in the literature that do not consider asymmetric HWD.

Index Terms—Achievable rate region, asymmetric hardware distortions, 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 distortions (HWD), of devices which can result in a substantial performance degradation [2]–[4]. HWD 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]–[10]. Another main challenge for data-rate enhancement is to handle 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 [11]–[26]. In IGS schemes, the real and imaginary parts of the signal are correlated and/or have unequal powers [27], [28]. While proper Gaussian signaling (PGS) achieves channel capacity for point-to-point communications in the presence of proper Gaussian noise [29], this is not the case under improper Gaussian noise that arises as a result of asymmetric HWD [4], [11], [30], [31].

A. Related work

The effect of HWD is studied in [5]–[10] for various scenarios. In [5], the secrecy performance of downlink massive multiple-input multiple-output (MIMO) systems was considered with HWD and a passive multiple-antenna eavesdropper. The paper [6] analyzed the achievable rate of massive MIMO systems with Rician channels and HWD. In [7], the authors considered a full-duplex massive MIMO relay with HWD and proposed a scheme to mitigate the distortion by exploiting statistical knowledge of the channels. In [8], the authors studied a massive MIMO system with a new system model for HWD at the transceivers. The paper [10] studied the performance of dual-hop relaying with different protocols in the presence of HWD.

In the aforementioned papers, symmetric HWD are considered. Nevertheless, HWD can, in general, provoke asymmetric or improper distortion in both the transmitted and received signal [4], [11], [30]–[32]. The paper [4] considered IGS in a single-input, multiple-output (SIMO) system with additive asymmetric HWD and showed that IGS improved the performance of the system. In [11], the authors investigated the effect of IGS in a relay network with additive asymmetric HWD. They maximized the achievable rate of the relay network by optimizing the complementary variance of the transmitted signal in the source and relay nodes.

Improper signaling schemes have also been proposed to improve different performance metrics in interference-limited networks with ideal devices [12]–[23], [26]. In [12], 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 this scenario. The paper [13] showed that IGS can increase the DoF of MIMO X channels. In [14]–[23], [25], the authors studied the performance of IGS when Treating Interference as Noise (TIN) was the strategy used for decoding. The paper [14] showed that IGS can improve the performance of the two-user interference channel, while in [15] IGS was used to optimize the rate of the K-user MIMO interference channel. Moreover, the authors in [15] derived the rate region of

the two-user single-input, single-output (SISO) IC with TIN by solving a semidefinite programming (SDP) problem, showing that IGS can enlarge the rate region and improve the performance of the system. The paper [16] showed that IGS can reduce the symbol error rate of the K-user IC. In [17]–[19], benefits of IGS were studied in different Z-IC scenarios. In [20], [21], the authors showed that IGS improves the performance of underlay and overlay cognitive radio systems, respectively. Finally, [26] showed that IGS can improve the performance of full-duplex relaying systems with fading channels.

B. Contribution

In this paper, we study the performance of IGS in a two-user IC with additive asymmetric HWD with TIN. To the best of our knowledge, it is the first work addressing the SISO IC with asymmetric HWD. We assume that the transceivers of both users produce additive asymmetric HWD noise, and model the HWD as an additive improper Gaussian noise, similar to [4], [11], [30], [31]. We devise two iterative algorithms to derive suboptimal solutions for the achievable rate region of the two-user IC. To this end, we rewrite the rate region as a pseudo-signal-to-interference-plus-noise-ratio (PSINR) region and employ sequential optimization approaches to solve the resulting problems.

In our first proposed algorithm, we employ majorization minimization (MM) as well as fractional programming (FP) and the well-known generalized Dinkelbach algorithm. MM is an iterative algorithm and consists of two steps in every iteration: i) majorization, and ii) minimization [33]. In the majorization step, the objective function is approximated by a surrogate function. Then, the approximated problem is solved in the minimization step. In other words, MM solves a non-convex optimization problem by solving a sequence of surrogate optimization problems, which can be solved easier than the original problem [33]. In our algorithm, to solve each surrogate problem, we apply the generalized Dinkelbach algorithm, which is a powerful tool to solve multiple ratio maximin problems [34], [35]. In Dinkelbach-based algorithms, an iterative optimization is performed, in which the fractional 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, i.e., its global optimum is obtained [34]–[37].

In our second proposed algorithm, we employ a separate optimization of powers and complementary variances. We first optimize the powers transmitted by the 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. In order to obtain the complementary variances, we employ difference of convex programming (DCP), which is a special case of sequential convex programming (SCP) and falls into MM [33], [38]. In DCP, the objective function and/or constraints are difference of two convex/concave functions. DCP solves a non-convex problem by solving a sequence of convex optimization problems and converges to a stationary point[1] of the original problem [38].

---

[1]A stationary point of a constrained optimization problem satisfies the corresponding Karush-Kuhn-Tucker (KKT) conditions [38].
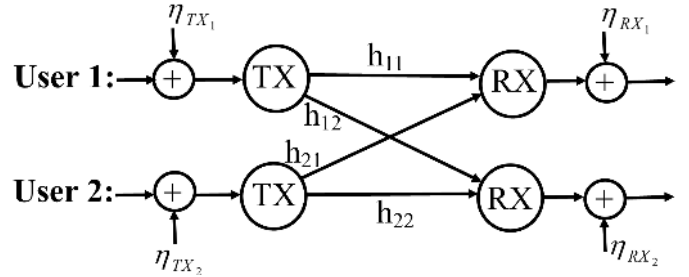

Fig. 1: The channel model for the SISO two-user IC.

The main contributions of this paper are as in the following:

- We first propose an iterative algorithm based on a sequential optimization method, in which we solve a sequence of fractional optimization problems [33], [39]. We derive the global optimal solution of each surrogate problem by FP and the generalized Dinkelbach algorithm. Our first proposed algorithm obtains a stationary point of the PSINR region.
- 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. We employ a bisection method to obtain the powers and derive a closed-form solution for powers in each iteration. Then, we employ DCP to find the complementary variances.
- Our results show that IGS enlarges the achievable rate of the two-user IC in the presence of additive asymmetric HWD, and that there is a significant performance improvement by IGS for highly asymmetric HWD noise. Moreover, both of our proposed algorithms outperform existing PGS and other existing IGS algorithms.

C. Paper outline

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 MM and 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 $q_x = \mathbb{E}\{x^2\}$ [27], [28]. 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 [27], [28]. We call x maximally improper if $|\tilde{\kappa}_x| = 1$.

B. Hardware distortion model

In this paper, we employ the distortion model in [4], [11], [30], [31] and model the aggregated effect of HWD on

the transceiver of a communication link with an improper Gaussian additive noise as

$$y = \sqrt{P}h(x + \eta) + n, \tag{1}$$

where y, x, P, h, η, and n are the received signal, transmitted symbol, transmission power, channel coefficient, aggregated HWD noise and additive complex proper Gaussian noise, respectively. The aggregated HWD noise is modeled as an improper complex Gaussian random variable with probability distribution $\eta \sim \mathcal{CN}(0, \sigma_\eta^2, \tilde{\sigma}_\eta^2)$, where $\sigma_\eta^2 = \sigma_{\eta_{\text{TX}}}^2 + \sigma_{\eta_{\text{RX}}}^2$ and $\tilde{\sigma}_\eta^2 = \tilde{\sigma}_{\eta_{\text{TX}}}^2 + \tilde{\sigma}_{\eta_{\text{RX}}}^2$ are the variance and complementary variance of η, respectively, both of which are composed of contributions at the transmitter side (denoted TX) and the receiver side (denoted RX). Please refer to [30, Lemma 1] for more details about this model.

It is worth mentioning that this model is an extension of the model in [5]–[10], where the HWD is modeled as additive proper Gaussian noise. However, as indicated in e.g., [4], [11], [30]–[32], [40]–[42], the aggregated HWD is, in general, improper due to I/Q imbalance. Note that the variances and complementary variances of HWD 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 HWD noise [4], [30]. Moreover, even if the channel noise is proper, the aggregated distortion is improper due to the asymmetric HWD.

C. Network scenario and signal model

We consider a two-user IC with additive asymmetric HWD at the transmitters and receivers of both users, as depicted in Fig. 1.[2] We assume that users are allowed to employ IGS and treat the interference as noise. Using the proposed HWD model, the received signals at receiver k is

$$y_k = \sqrt{p_1}h_{1k}(x_1 + \eta_{1k}) + \sqrt{p_2}h_{2k}(x_2 + \eta_{2k}) + n_k, \tag{2}$$

respectively, where x_k, h_{jk}, n_k, and η_{jk} for $j, k \in \{1, 2\}$ are the transmit signal of user k, channel between transmitter j and receiver k, independent zero-mean proper complex Gaussian noise with variance σ^2, and the aggregated HWD noise of the link between transmitter j and receiver k, respectively. Since the transmitted signals x_1 and x_2 are improper complex Gaussian, the achievable rate of user $k \in \{1, 2\}$ is [11], [15], [19] given by (3), shown at the top of the next page, where p_k, q_k, $\sigma_{\eta_{jk}}^2$, and $\tilde{\sigma}_{\eta_{jk}}^2$ for $i, j \in \{1, 2\}$ are, respectively, the transmission power of user k, the complementary variance of the transmitted signal of user k, the aggregated variance and the complementary variance of the HWD noise in the link between user j and user k. The rate of user $k \in \{1, 2\}$ can be written using vector notation as

$$R_k = \frac{1}{2} \log_2 \left(\frac{(\sigma^2 + \mathbf{a}_k^T \mathbf{p})^2 - |\mathbf{f}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}|^2}{(\sigma^2 + \mathbf{b}_k^T \mathbf{p})^2 - |\mathbf{g}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}|^2} \right), \tag{4}$$

where the corresponding parameters are defined in (5)-(9), shown at the top of the next page. 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

[2] 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.

user k. Note that since q_k for $k = 1, 2$ is the complementary variance of user k, its absolute value has to be not greater than the transmission power of user k, i.e., $|q_k| \leq p_k$.

It is to be noted that, in practice, discrete rather than Gaussian signaling is employed (see, e.g., [16], [42], [43]), which will lead to performance degradation with respect to IGS. The significance of studying improper Gaussian signals is that it shows us whether *improper* signaling may in principle achieve performance improvements over *proper* signaling. In this paper, we focus on IGS and leave the analysis and design of improper discrete constellations for future work.

D. Problem Statement

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 ([19], [44]). 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., $\mathcal{R} = \bigcup_{\{\mathbf{p}, \mathbf{q}\} \in \Omega} (R_1, R_2)$, and its boundary can be derived by the rate profile technique as in the following optimization problem [15]

$$\underset{R, \mathbf{p}, \mathbf{q}}{\text{maximize}} \quad R \tag{10a}$$

$$\text{s.t.} \quad R_k \geq \lambda_k R, \quad\quad k = 1, 2, \tag{10b}$$

$$0 \leq p_k \leq P_k, \quad\quad k = 1, 2, \tag{10c}$$

$$|q_k| \leq p_k, \quad\quad k = 1, 2, \tag{10d}$$

where $\lambda_1, \lambda_2 \geq 0$ are fixed and $\lambda_1 + \lambda_2 = 1$. We can obtain the boundary of the rate region by solving (10) for different rate-profile parameters, i.e., λ_1 and λ_2. Note there are efficient algorithms to derive the global optimal solution of convex optimization problems [39], [45], [46]. However, we are unable to apply these algorithms to (10) due to the fact that the rates are not concave functions of the optimization variables, which makes (10) non-convex [39], [45], [46]. Hence, in this paper we propose numerical algorithms to derive suboptimal solutions to (10).

The paper [30] proposed an algorithm based on DCP to maximize the achievable rate of a multihop relay system with additive asymmetric HWD, in which all nodes transmit with maximum power, by optimizing over the complementary variances. Such algorithms cannot be applied for a joint optimization of powers and complementary variances since, in this case, the rates are not a difference of two jointly concave/convex functions in $\mathbf{p}$ and $\mathbf{q}$. Hence, we solve (10) by MM and FP. In MM, the objective function and constraints of an optimization problem are not required to follow a very specific structure such as being a difference of two convex/concave functions, which makes it more powerful than DCP.

To solve (10), we rewrite it such that it is more suitable to be solved with MM and FP. To this end, we employ the PSINR profile technique in [47], [48] to write an optimization problem that results in the solution of (10). We define

$$R_k = \frac{1}{2}\log_2\left(\frac{\left(\sigma^2 + \sum_{j=1}^{2} p_j |h_{jk}|^2(1+\sigma_{\eta_{jk}}^2)\right)^2 - \left|\sum_{j=1}^{2}(q_j + p_j \tilde{\sigma}_{\eta_{jk}}^2)h_{jk}^2\right|^2}{\left(\sigma^2 + \sum_{j=1}^{2} p_j |h_{jk}|^2(1+\sigma_{\eta_{jk}}^2) - p_k|h_{kk}|^2\right)^2 - \left|\sum_{j=1}^{2}(q_j + p_j\tilde{\sigma}_{\eta_{jk}}^2)h_{jk}^2 - q_k h_{kk}^2\right|^2}\right), \tag{3}$$

$$\mathbf{a}_k = \left[\, |h_{1k}|^2(1+\sigma_{\eta_{1k}}^2) \quad |h_{2k}|^2(1+\sigma_{\eta_{2k}}^2)\,\right]^T, \qquad\qquad \mathbf{f}_k = \left[\, h_{1k}^2 \quad h_{2k}^2 \,\right]^H, \tag{5}$$

$$\tilde{\mathbf{f}}_k = \left[\, h_{1k}^2\tilde{\sigma}_{\eta_{1k}}^2 \quad h_{2k}^2\tilde{\sigma}_{\eta_{2k}}^2 \,\right]^H, \tag{6}$$

$$\mathbf{b}_1 = \left[\, |h_{11}|^2\sigma_{\eta_{11}}^2 \quad |h_{21}|^2(1+\sigma_{\eta_{21}}^2) \,\right]^T, \qquad\qquad \mathbf{g}_1 = \left[\, 0 \quad h_{21}^2 \,\right]^H, \tag{7}$$

$$\mathbf{b}_2 = \left[\, |h_{12}|^2(1+\sigma_{\eta_{12}}^2) \quad |h_{22}|^2\sigma_{\eta_{22}}^2 \,\right]^T, \qquad\qquad \mathbf{g}_2 = \left[\, h_{12}^2 \quad 0 \,\right]^H, \tag{8}$$

$$\mathbf{q} = \left[\, q_1 \quad q_2 \,\right]^T, \qquad\qquad \mathbf{p} = \left[\, p_1 \quad p_2 \,\right]^T. \tag{9}$$

---

the PSINR profile as

$$\underset{E,\mathbf{p},\mathbf{q}}{\text{maximize}} \quad E \tag{11a}$$

$$\text{s.t.} \quad E_k(\mathbf{p},\mathbf{q}) \geq 1 + \alpha_k E, \quad k = 1,2, \tag{11b}$$

$$0 \leq p_k \leq P_k, \quad k = 1,2, \tag{11c}$$

$$|q_k| \leq p_k, \quad k = 1,2, \tag{11d}$$

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

$$E_k(\mathbf{p},\mathbf{q}) \triangleq \frac{(\sigma^2 + \mathbf{a}_k^T\mathbf{p})^2 - |\mathbf{f}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p}|^2}{(\sigma^2 + \mathbf{b}_k^T\mathbf{p})^2 - |\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p}|^2} = \frac{u_k(\mathbf{p},\mathbf{q})}{v_k(\mathbf{p},\mathbf{q})}. \tag{12}$$

We can derive the boundary of the PSINR region by varying $\alpha_1 \in [0,1]$. Note that $E_k(\mathbf{p},\mathbf{q}) \geq 1$ for $k = 1,2$ since the rates are non-negative. Moreover, the numerator and denominator of $E_k(\mathbf{p},\mathbf{q})$ are strictly positive because the rates are bounded and non-negative. In the following lemma, we show that this technique results in the boundary of the rate region in (10).

Lemma 1 ([47], [48]). *Every point in the boundary of the rate region corresponds to a point in the boundary of the PSINR region, and vice versa.*

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 PSINR region. In other words, the pair $(E_1 = 2^{R_1}, E_2 = 2^{R_2})$, which is a feasible PSINR pair, is not on the boundary of the PSINR 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. Similarly, it can be shown that every point in the boundary of the PSINR region associates with a point in the boundary of the rate region. ∎

Note that we can rewrite (11) as the following maximin optimization problem by removing the variable E

$$\underset{0 \leq p_k \leq P_k, |q_k| \leq p_k}{\text{maximize}} \quad \min_{k=1,2}\left\{\frac{E_k(\mathbf{p},\mathbf{q}) - 1}{\alpha_k}\right\}. \tag{13}$$

III. BOUNDARY OF THE RATE REGION BY FRACTIONAL PROGRAMMING

In this section, we solve the PSINR profile problem in (11) by MM, which results in solving a sequence of fractional optimization problems. We solve each fractional optimization problem by FP and the generalized Dinkelbach

algorithm [35]–[37]. Our proposed algorithm converges to a stationary point of (11). We first provide preliminaries on generalized Dinkelbach's algorithm in Section III-A and then propose our algorithm in III-B.

A. Preliminaries of generalized Dinkelbach's algorithm

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 [34]. 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 ([34], [35]). *Consider the fractional functions $\frac{u_i(\mathbf{x})}{v_i(\mathbf{x})}$, where $u_i(\mathbf{x})$ and $v_i(\mathbf{x})$ are continuous in $\mathbf{x}$, $v_i(\mathbf{x})$ is strictly positive in $\mathbf{x}$, and $\mathbf{x}$ is a vector with dimension n that belongs to a compact set $\mathcal{X}$. Let us define*

$$V(\mu) = \max_{\mathbf{x}} \ \min_{i}\left(u_i(\mathbf{x}) - \mu v_i(\mathbf{x})\right), \tag{14}$$

$$\bar{\mu} = \max_{\mathbf{x}} \ \min_{i}\left(\frac{u_i(\mathbf{x})}{v_i(\mathbf{x})}\right), \tag{15}$$

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

1) *$V(\mu)$ is continuous and strictly decreasing in μ.*
2) *The optimization problems (14) and (15) 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 (14) 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 μ to obtain $\bar{\mu}$. Assuming $u_i(\mathbf{x}) \geq 0$, which is the case we consider in this paper, $V(0) = \max_{\mathbf{x}} \min_{i}(u_i(\mathbf{x})) > 0$. Since $V(\mu)$ is continuous and strictly decreasing in μ, μ is chosen monotonically increasing at each iteration ($\mu^{(l)} > \mu^{(l-1)}$) until $V(\mu)$ approaches 0. At the lth iteration, $\mu^{(l)}$ is

$$\mu^{(l)} = \min\left(\frac{u_1(\mathbf{x}^{(l-1)})}{v_1(\mathbf{x}^{(l-1)})}, \frac{u_2(\mathbf{x}^{(l-1)})}{v_2(\mathbf{x}^{(l-1)})}\right) > 0, \tag{16}$$

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

$$\mathbf{x}^{(l-1)} = \arg\max_{\mathbf{x}} \ \min_{i}\left(u_i(\mathbf{x}) - \mu^{(l-1)}v_i(\mathbf{x})\right). \tag{17}$$

The generalized Dinkelbach algorithm updates $\mu^{(l)}$ and $\mathbf{x}^{(l-1)}$ based on (16) and (17), respectively, until a convergence metric is met, e.g., $V(\mu^{(l)}) < \epsilon$, where $\epsilon > 0$. This algorithm converges linearly to the optimal solution [34].

Note that in order to apply the generalized Dinkelbach algorithm, it is not required that $u_i(\mathbf{x})$ and $v_i(\mathbf{x})$ fulfill any other condition (except those in the lemma), which makes this algorithm a powerful tool to solve different types of fractional problems. If $u_i(\mathbf{x})$ and $v_i(\mathbf{x})$ 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.

B. Proposed algorithm

We can apply the generalized Dinkelbach algorithm to derive the boundary of the PSINR region since the optimization problem can be written as a maximin weighted problem as indicated in (13). However, since $u_k(\mathbf{p}, \mathbf{q})$ and $v_k(\mathbf{p}, \mathbf{q})$ are not, respectively, concave and convex in optimization variables, the corresponding optimization problem in each iteration of the generalized Dinkelbach algorithm is not convex. Indeed, $u_k(\mathbf{p}, \mathbf{q})$ and $v_k(\mathbf{p}, \mathbf{q})$ are a difference of two convex/concave functions:

$$u_k(\mathbf{p}, \mathbf{q}) = \underbrace{-|\mathbf{f}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}|^2}_{\text{concave part}} + \underbrace{(\sigma^2 + \mathbf{a}_k^T \mathbf{p})^2}_{\text{convex part}}, \quad (18)$$

$$v_k(\mathbf{p}, \mathbf{q}) = \underbrace{-|\mathbf{g}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}|^2}_{\text{concave part}} + \underbrace{(\sigma^2 + \mathbf{b}_k^T \mathbf{p})^2}_{\text{convex part}}. \quad (19)$$

Hence, to solve (13), we employ a sequential optimization approach by approximating $E_k(\mathbf{p}, \mathbf{q})$ with a lower bound $\tilde{E}_k(\mathbf{p}, \mathbf{q}, \mu)$ in each iteration [33], [39]. Then, we obtain the global optimal solution of each surrogate optimization problem by the generalized Dinkelbach algorithm. To this end, in each iteration, we first approximate $u_k(\mathbf{p}, \mathbf{q})$ by a lower bound concave function $\tilde{u}_k(\mathbf{p}, \mathbf{q})$ and $v_k(\mathbf{p}, \mathbf{q})$ by an upper bound convex function $\tilde{v}_k(\mathbf{p}, \mathbf{q})$ as in the following lemma.

Lemma 3. *A concave lower bound for* $u_k(\mathbf{p}, \mathbf{q})$ *in the* mth *iteration is*

$$\tilde{u}_k^{(m)}(\mathbf{p}, \mathbf{q}) = -|\mathbf{f}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}|^2 + (\sigma^2 + \mathbf{a}_k^T \mathbf{p}^{(m)})^2$$
$$+ 2(\sigma^2 + \mathbf{a}_k^T \mathbf{p}^{(m)})\mathbf{a}_k^T(\mathbf{p} - \mathbf{p}^{(m)}), \quad (20)$$

Moreover, a convex upper bound for $v_k(\mathbf{p}, \mathbf{q})$ *in the* mth *iteration is*

$$\tilde{v}_k^{(m)}(\mathbf{p}, \mathbf{q}) = (\sigma^2 + \mathbf{b}_k^T \mathbf{p})^2 - |\mathbf{g}_k^H \mathbf{q}^{(m)} + \tilde{\mathbf{f}}_k^H \mathbf{p}^{(m)}|^2$$
$$- 2\Re\left[\tilde{\mathbf{f}}_k^H(\mathbf{g}_k^H \mathbf{q}^{(m)} + \tilde{\mathbf{f}}_k^H \mathbf{p}^{(m)})^*\right](\mathbf{p} - \mathbf{p}^{(m)})$$
$$- 2\Re\left[(\mathbf{g}_k^H \mathbf{q}^{(m)} + \tilde{\mathbf{f}}_k^H \mathbf{p}^{(m)})^* \mathbf{g}_k^H(\mathbf{q} - \mathbf{q}^{(m)})\right], \quad (21)$$

where $\mathbf{p}^{(m)}$ *and* $\mathbf{q}^{(m)}$ *are the power and complementary variances at the* mth *iteration, which are the solution of the previous iteration. Furthermore,* $\Re[x]$ *takes the real part of* x.

Proof: Please refer to Appendix A. ∎

Algorithm I Proposed sequential optimization algorithm.

Initialization
Set ϵ, M, $\mathbf{p}^{(0)} = \mathbf{0}$, $\mathbf{q}^{(0)} = \mathbf{0}$, $m = 1$, convergence=0
While convergence=0 and $m \leq M$ **do**
  Construct $\tilde{E}_k^{(m)}(\mathbf{p}, \mathbf{q}) = \tilde{u}_k^{(m)}(\mathbf{p}, \mathbf{q})/\tilde{v}_k^{(m)}(\mathbf{p}, \mathbf{q})$ for $k = 1, 2$ using Lemma 3
  Obtain $\mathbf{p}^{(m+1)}$ and $\mathbf{q}^{(m+1)}$ by solving (22), i.e., run algorithm II
  If $\|\mathbf{p}^{(m)} - \mathbf{p}^{(m+1)}\|/\|\mathbf{p}^{(m)}\| < \epsilon$
  and $\|\mathbf{q}^{(m)} - \mathbf{q}^{(m+1)}\|/\|\mathbf{q}^{(m)}\| < \epsilon$
    convergence=1
    $\mathbf{p}^\star = \mathbf{p}^{(m+1)}$ and $\mathbf{q}^\star = \mathbf{q}^{(m+1)}$
  End (If)
  $m = m + 1$
End (While)
Return $\mathbf{p}^\star$ and $\mathbf{q}^\star$.

Now, we are able to write the surrogate optimization problem in mth iteration as

$$\underset{E', \mathbf{p}, \mathbf{q}}{\text{maximize}} \quad E \quad (22a)$$

$$\text{s.t.} \quad \tilde{E}_k^{(m)}(\mathbf{p}, \mathbf{q}) \geq 1 + \alpha_k E', \quad k = 1, 2, \quad (22b)$$

$$0 \leq p_k \leq P_k, \quad k = 1, 2, \quad (22c)$$

$$|q_k| \leq p_k, \quad k = 1, 2, \quad (22d)$$

where $\tilde{E}_k^{(m)}(\mathbf{p}, \mathbf{q}) = \frac{\tilde{u}_k^{(m)}(\mathbf{p}, \mathbf{q})}{\tilde{v}_k^{(m)}(\mathbf{p}, \mathbf{q})}$ and $E_k(\mathbf{p}, \mathbf{q})$ fulfill the following conditions:

1) $\tilde{E}_k^{(m)}(\mathbf{p}, \mathbf{q}) \leq E_k(\mathbf{p}, \mathbf{q})$ for all feasible $\mathbf{p}, \mathbf{q}$ and $k = 1, 2$.
2) $\tilde{E}_k^{(m)}(\mathbf{p}^{(m)}, \mathbf{q}^{(m)}) = E_k(\mathbf{p}^{(m)}, \mathbf{q}^{(m)})$ for $k = 1, 2$.
3) $\frac{\partial \tilde{E}_k^{(m)}(\mathbf{p}^{(m)}, \mathbf{q}^{(m)})}{\partial \mathbf{p}} = \frac{\partial E_k(\mathbf{p}^{(m)}, \mathbf{q}^{(m)})}{\partial \mathbf{p}}$ and $\frac{\partial \tilde{E}_k^{(m)}(\mathbf{p}^{(m)}, \mathbf{q}^{(m)})}{\partial \mathbf{q}} = \frac{\partial E_k(\mathbf{p}^{(m)}, \mathbf{q}^{(m)})}{\partial \mathbf{q}}$ for $k = 1, 2$.

These properties guarantee that the algorithm converges to a stationary point of (11) [39, Section II.B]. To solve (22) and obtain $\mathbf{p}^{(m+1)}$ and $\mathbf{q}^{(m+1)}$, we employ the generalized Dinkelbach algorithm, which gives the global optimal solution of (22), as in the following. We summarize this procedure in Algorithm I.

Now we solve (22) and obtain its global optimal solution by FP, which is also an iterative algorithm as explained in Section III-A. To this end, we introduce the following functions, which are the corresponding surrogate functions of $\frac{\tilde{E}_k^{(m)} - 1}{\alpha_k}$ for $k = 1, 2$:

$$\hat{E}_k(\mathbf{p}, \mathbf{q}, \mu^{(l)}) \triangleq u_k^{(m)}(\mathbf{p}, \mathbf{q}) - (\mu^{(l)}\alpha_k + 1)v_k^{(m)}(\mathbf{p}, \mathbf{q}), \quad (23)$$

where $\mu^{(l)} \in \mathbb{R}$ is fixed and given by

$$\mu^{(l)} = \min_{k=1,2}\left(\frac{\tilde{E}_k(\mathbf{p}^{(l-1)}, \mathbf{q}^{(l-1)}) - 1}{\alpha_k}\right). \quad (24)$$

It is worth mentioning that the generalized Dinkelbach algorithm requires an initial point $\mu^{(0)}$, which can be obtained by substituting $\mathbf{p}^{(m)}$ and $\mathbf{q}^{(m)}$ in (24). By substituting (23) in (22), the optimization problem at each iteration of the generalized Dinkelbach algorithm is

$$\underset{E', \mathbf{p}, \mathbf{q}}{\text{maximize}} \quad E' \quad (25a)$$

$$\text{s.t.} \quad \hat{E}_k(\mathbf{p}, \mathbf{q}, \mu^{(l)}) \geq E', \quad k = 1, 2, \quad (25b)$$

$$(22c), (22d). \quad (25c)$$

We solve (25) for the given $\mu^{(l)}$, which results in $\mathbf{p}^{(l)}$ and $\mathbf{q}^{(l)}$. Then, we update $\mu^{(l)}$ by (24) and repeat the procedure

Algorithm II Generalized Dinkelbach algorithm.

Initialization
Set ϵ, L, $l = 0$, $\mu^{(l)} = \min_{k=1,2} \left(\frac{\tilde{E}_k(\mathbf{p}^{(m)}, \mathbf{q}^{(m)}) - 1}{\alpha_k} \right)$
Compute $\hat{E}_k(\mathbf{p}, \mathbf{q}, \mu^{(l)})$ for $k = 1, 2$ by (23)
While $\min_{k=1,2} \{\hat{E}_k(\mathbf{p}, \mathbf{q}, \mu^{(l)})\} \geq \epsilon$ and $l \leq L$ **do**
$l = l + 1$
Obtain $\mathbf{p}^{(l)}$ and $\mathbf{q}^{(l)}$ by solving (25)
If $\min_{k=1,2} \{\tilde{E}_k(\mathbf{p}, \mathbf{q}, \mu^{(l)})\} < \epsilon$
$\mathbf{p}^\star = \mathbf{p}^{(l)}$ and $\mathbf{q}^\star = \mathbf{q}^{(l)}$
Else
Update $\mu^{(l)}$ by (24)
End (If)
End (While)
Return $\mathbf{p}^\star$ and $\mathbf{q}^\star$.

until a convergence metric is met. As indicated in Section III-A, the convergence rate of the generalized Dinkelbach algorithm is linear. The optimization problem (25) is convex, and its global optimal solution can be efficiently obtained [45]. We summarize this procedure in Algorithm II.

To sum up, the proposed algorithm works as follows. We solve the PSINR profile in (11) by solving a sequence of fractional optimization problems. Indeed, we employ a sequential optimization approach and approximate the PSINR term of each user by a lower bound. In order to derive the global optimal solution of each fractional optimization problem, we perform another iterative algorithm, i.e., the generalized Dinkelbach algorithm. It is worth mentioning that this algorithm does not converge to the Pareto-optimal solution; however, it obtains a stationary point of (11).

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 $\mathbf{p}$ for PGS, i.e., for $\mathbf{q} = \mathbf{0}$. This problem is addressed in Section IV-A. Then, in Section IV.B, we optimize the complementary variances for the resulting transmit power $\mathbf{p}$ such that the rates of all users is simultaneously increased.

A. Power optimization

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

$$\underset{E,\mathbf{p}}{\text{maximize}} \quad E \tag{26a}$$

$$\text{s.t.} \quad \frac{(\sigma^2 + \mathbf{a}_i^T\mathbf{p})^2 - |\tilde{\mathbf{f}}_i^H\mathbf{p}|^2}{(\sigma^2 + \mathbf{b}_i^T\mathbf{p})^2 - |\tilde{\mathbf{f}}_i^H\mathbf{p}|^2} \geq 1 + \alpha_k E, \quad k = 1, 2, \tag{26b}$$

$$0 \leq p_k \leq P_k, \qquad k = 1, 2, \tag{26c}$$

for $\alpha_1, \alpha_2 \geq 0$ and $\alpha_1 + \alpha_2 = 1$. Unfortunately, the optimization problem in (26) is not convex due to (26b). In the following lemma, we derive a lower bound for (26b), which allows us to simplify (26) and derive a low-complexity algorithm.

Lemma 4. *A lower bound for the left-hand side of* (26b) *is*

$$\frac{(\sigma^2 + \mathbf{a}_i^T\mathbf{p})^2 - |\tilde{\mathbf{f}}_i^H\mathbf{p}|^2}{(\sigma^2 + \mathbf{b}_i^T\mathbf{p})^2 - |\tilde{\mathbf{f}}_i^H\mathbf{p}|^2} \geq \frac{(\sigma^2 + \mathbf{a}_i^T\mathbf{p})^2}{(\sigma^2 + \mathbf{b}_i^T\mathbf{p})^2}, \tag{27}$$

where the equality in (27) *holds if and only if the HWD noise is proper, i.e.,* $\tilde{\mathbf{f}}_i = \mathbf{0}$.

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

$$f(t) = \frac{\beta_1 - t}{\beta_2 - t}, \tag{28}$$

where $0 \leq t < \beta_2 < \beta_1$. The lower bound in (27) is then satisfied if $f(t)$ is increasing in t. This function is strictly increasing in $t \in [0, \beta_2)$ since

$$\frac{\partial f(t)}{\partial t} = \frac{\beta_1 - \beta_2}{(\beta_2 - t)^2} > 0, \tag{29}$$

Thus, we have

$$\frac{\beta_1 - t}{\beta_2 - t} \geq \frac{\beta_1}{\beta_2}, \tag{30}$$

with equality if and only if $t = 0$. ∎

For each point characterized by α_1 and α_2, we solve (26) for the lower bound in (27) as the optimization problem

$$\underset{E,\mathbf{p}}{\text{maximize}} \quad E \tag{31a}$$

$$\text{s.t.} \quad \frac{\sigma^2 + \mathbf{a}_k^T\mathbf{p}}{\sigma^2 + \mathbf{b}_k^T\mathbf{p}} \geq \sqrt{1 + \alpha_k E}, \quad k = 1, 2, \tag{31b}$$

$$0 \leq p_k \leq P_k, \qquad k = 1, 2. \tag{31c}$$

It is worth mentioning that the lower bound in Lemma 4 is employed to simplify (26) and obtain the powers, and the actual rates are derived by substituting the obtained powers in (3). Note that the region achieved by solving (26) includes the region achieved by solving (31). If the additive HWD noise is proper, (31) is equivalent to (26)[3]. The global optimum solution of (31) can be derived by employing a bisection method and solving a sequence of feasibility problems [39]. That is, we fix E as E' and consider the feasibility problem (32), shown at the top of the next page. If (32) is feasible for a given E', the optimal solution of (31) is greater than or equal to E', i.e., $E^\star \geq E'$. Otherwise, $E^\star < E'$. In order to find $E^\star$, we employ the well-known bisection method over E' solving (32) at each iteration, which yields, upon convergence, the global optimal solution of (31) [45]. Constraints (32b) and (32c) are linear in $\mathbf{p}$, which permits deriving a closed-form expression for a feasible point, as presented in the following theorem. It is worth mentioning that this algorithm does not attain the global optimal solution of (26). There might be optimization approaches to obtain its global optimal solution such as the monotonic optimization framework [50]–[52], although the computational complexity of these approaches is high.

Theorem 1. *The optimization problem in* (32) *is feasible for a given* E' *if and only if* $0 \leq p'_k \leq P_k$, *for* $k = 1, 2$, *where*

$$\begin{bmatrix} p'_1 \\ p'_2 \end{bmatrix} = \mathbf{A}^{-1} \begin{bmatrix} (\sqrt{1 + \alpha_1 E'} - 1)\sigma^2 \\ (\sqrt{1 + \alpha_2 E'} - 1)\sigma^2 \end{bmatrix}. \tag{33}$$

Moreover, $\mathbf{A}$ *is given by* (34), *shown at the top of the next page.*

---

[3]This is in line with [49], where it was shown that proper Gaussian noise is the worst case in a K-user MIMO IC with ideal devices.

$$\text{find} \quad \mathbf{p} \in \mathbb{R}^2, \tag{32a}$$

$$\text{s.t.} \quad (\mathbf{a}_k^T - \sqrt{1 + \alpha_k E'} \mathbf{b}_k^T)\mathbf{p} \geq (\sqrt{1 + \alpha_k E'} - 1)\sigma^2, \quad k = 1, 2, \tag{32b}$$

$$0 \leq p_k \leq P_k, \quad k = 1, 2. \tag{32c}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T - \sqrt{1 + \alpha_1 E'} \mathbf{b}_1^T \\ \mathbf{a}_2^T - \sqrt{1 + \alpha_2 E'} \mathbf{b}_2^T \end{bmatrix} = \begin{bmatrix} |h_{11}|^2 \left(1 - \sigma_{\eta_{11}}^2(\sqrt{1 + \alpha_1 E'} - 1)\right) & -|h_{21}|^2(1 + \sigma_{\eta_{21}}^2)(\sqrt{1 + \alpha_1 E'} - 1) \\ -|h_{12}|^2(1 + \sigma_{\eta_{12}}^2)(\sqrt{1 + \alpha_2 E'} - 1) & |h_{22}|^2 \left(1 - \sigma_{\eta_{22}}^2(\sqrt{1 + \alpha_2 E'} - 1)\right) \end{bmatrix}. \tag{34}$$

$$\underset{t,\mathbf{q}}{\text{maximize}} \quad t \tag{35a}$$

$$\text{s.t.} \quad \frac{(\sigma^2 + \mathbf{a}_k^T \mathbf{p}^\star)^2 - |\mathbf{f}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2}{(\sigma^2 + \mathbf{b}_k^T \mathbf{p}^\star)^2 - |\mathbf{g}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2} \geq E_{p,k} + \alpha_k t, \quad k = 1, 2, \tag{35b}$$

$$|q_k| \leq p_k^\star, \quad k = 1, 2. \tag{35c}$$

$$\underset{t_1,t_2,\mathbf{q}}{\text{maximize}} \quad \min(t_1, t_2) \tag{38a}$$

$$\text{s.t.} \quad \frac{(\sigma^2 + \mathbf{a}_k^T \mathbf{p}^\star)^2 - |\mathbf{f}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2 - \alpha_k t_k}{(\sigma^2 + \mathbf{b}_k^T \mathbf{p}^\star)^2 - |\mathbf{g}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2} \geq E_{p,k}, \quad k = 1, 2, \tag{38b}$$

$$|q_k| \leq p_k^\star, \quad k = 1, 2. \tag{38c}$$

Proof: Please refer to Appendix B. ∎

We note that this algorithm leads to the optimal PGS only when HWD noise is proper. Note that PGS is sub-optimal, in *point-to-point* communications, in the presence of asymmetric HWD [4], [31]. Thus, the users may improve the performance by employing IGS in additive asymmetric HWD. It is worth noting that, in this paper, we aim at proposing PGS and IGS schemes for the two-user IC with additive asymmetric HWD, but we do not derive sufficient and necessary conditions for the optimality of IGS or PGS in the two-user IC with additive asymmetric HWD, which remains an open problem.

B. Complementary variance design

In this subsection, we optimize the complementary variances $\mathbf{q}$ for a given $\mathbf{p}^\star$, which has been obtained by solving (31). We obtain $\mathbf{q}$ such that the rates of both users exceed the rates achieved by PGS, which are the rates achievable with $\mathbf{q} = \mathbf{0}$ and the power vector $\mathbf{p}^\star$ obtained by solving (31). In other words, we want to solve the optimization problem (35), shown at the top of this page, where $p_k^\star$ is the kth element of $\mathbf{p}^\star$. Moreover, $E_{p,k}$ is fixed and given by

$$E_{p,k} = \frac{(\sigma^2 + \mathbf{a}_k^T \mathbf{p}^\star)^2 - |\tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2}{(\sigma^2 + \mathbf{b}_k^T \mathbf{p}^\star)^2 - |\tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2}. \tag{36}$$

Unfortunately, (35) is not convex due to (35b). Hence, in order to efficiently solve (35), we first rewrite (35b) as

$$\frac{(\sigma^2 + \mathbf{a}_k^T \mathbf{p}^\star)^2 - |\mathbf{f}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2 - \alpha_k t_k}{(\sigma^2 + \mathbf{b}_k^T \mathbf{p}^\star)^2 - |\mathbf{g}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2} \geq E_{p,k}, \tag{37}$$

where $t_k = t \left[(\sigma^2 + \mathbf{b}_k^T \mathbf{p}^\star)^2 - |\mathbf{g}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2\right]$. We then relax the relation between t_1, t_2, and $\mathbf{q}$ and treat $\tilde{t}_1$ and t_2 as new optimization variables. In other words, we approximate (35) as (38), shown at the top of this page. If $\min(t_1, t_2) > 0$, the rates of both users are simultaneously increased by

employing IGS. Otherwise, we set $\mathbf{q} = \mathbf{0}$ and employ PGS. Note that the constraint (38b) can be rewritten as

$$E_{p,k} |\mathbf{g}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2 - |\mathbf{f}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2$$
$$+ (\sigma^2 + \mathbf{a}_k^T \mathbf{p}^\star)^2 - E_{p,k}(\sigma^2 + \mathbf{b}_k^T \mathbf{p}^\star)^2 \geq \alpha_k t_k, \quad (39)$$

which is a difference of two convex functions. Thus, (38) is not a convex optimization problem, but it can be efficiently solved by difference of convex programming and a convex-concave procedure similar to (25) [33], [38], [53]–[55]. Hence, we employ difference of convex programming (DCP) and solve (38) iteratively. At each iteration, we approximate the left-hand side of (39) by a concave function. To this end, we employ the first-order Taylor expansion and approximate the convex part of (39) around the point $\mathbf{q}^{(l)}$ by an affine function as

$$|\mathbf{g}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2 \simeq |\mathbf{g}_k^H \mathbf{q}^{(l)} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2$$
$$+ 2\Re \left((\mathbf{g}_k^H \mathbf{q}^{(l)} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star)^* \mathbf{g}_k^H (\mathbf{q} - \mathbf{q}^{(l)})\right), \quad (40)$$

where $\mathbf{q}^{(l)}$ contains the complementary variances of the users in the lth iteration. It is worth mentioning that $|\mathbf{g}_i^H \mathbf{q} + \tilde{\mathbf{f}}_i^H \mathbf{p}^\star|^2$ is always greater than or equal to the right-hand side of (40), and consequently, no trust region is required in DCP [53]–[55]. Finally, in the lth each iteration, (39) can be approximated by

$$- |\mathbf{f}_k^H \mathbf{q} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2 + E_{p,k} |\mathbf{g}_k^H \mathbf{q}^{(l)} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star|^2$$
$$+ 2E_{p,k} \Re \left((\mathbf{g}_k^H \mathbf{q}^{(l)} + \tilde{\mathbf{f}}_k^H \mathbf{p}^\star)^H \mathbf{g}_k^H (\mathbf{q} - \mathbf{q}^{(l)})\right)$$
$$+ (\sigma^2 + \mathbf{a}_k^T \mathbf{p}^\star)^2 - E_{p,k}(\sigma^2 + \mathbf{b}_k^T \mathbf{p}^\star)^2 \geq \alpha_k t_k. \quad (41)$$

Finally, the convex optimization problem in the lth iteration is

$$\underset{t_1,t_2,\mathbf{q}}{\text{maximize}} \quad \min(t_1, t_2) \tag{42a}$$

$$\text{s.t.} \quad (41), (38c). \tag{42b}$$

This problem can be easily solved by standard numerical tools [45]. Moreover, the proposed DCP algorithm converges to a stationary point of (38) [33], [38], [53]–[55]. It is worth mentioning that a stationary point of (38) is not necessarily a stationary point of (35).

The proposed simplified algorithm can be summarized as follows. The joint optimization problem for $\mathbf{p}$ and $\mathbf{q}$ 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 ϵ, 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 σ^2, i.e., SNR$= \frac{P}{\sigma^2}$. We compare our proposed algorithms with PGS and the joint variance and complementary variance optimization algorithm in [15] for IGS, which is designed for ideal devices. To the best of our knowledge, there exists no PGS algorithm for additive asymmetric HWD 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 [15] for ideal devices,
- **S-TS**: our proposed design in Section IV with time sharing,
- **F-TS**: our proposed design in Section III with time sharing,
- **P-TS**: the proposed PGS design in Section IV-A with time sharing.

A. Ideal devices

In this subsection, we compare the performance of our proposed algorithms with the joint variance and covariance IGS algorithm in [15] when there is no HWD. 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 [15], especially at high SNR. Our simplified algorithm performs similarly to the proposed algorithm in [15] for low SNR. However, the algorithm in [15] 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.
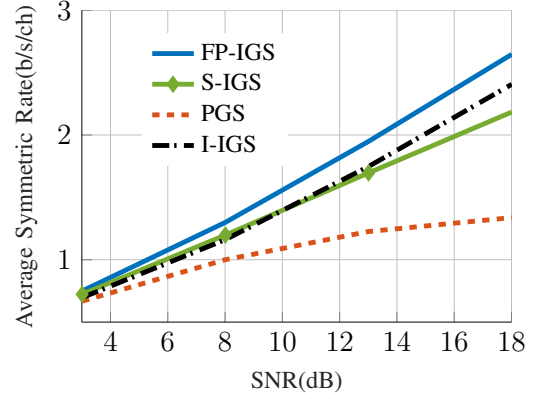
**Fig. 2:** Average symmetric rate for ideal devices versus the SNR.

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

$$\mathbf{H}_1 = \begin{bmatrix} 1.4070e^{i0.2721} & 0.9288e^{i1.8320} \\ 0.9288e^{i1.8320} & 1.7367e^{i1.1136} \end{bmatrix}, \quad (43)$$

where $[\mathbf{H}_1]_{ij} = h_{ij}$ for $i, j \in \{1, 2\}$. 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 averaged 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 [15] perform very closely to each other. In Fig. 3b, we also consider the effect of time sharing[4] on the achievable rate region. As can be observed, IGS with time sharing outperforms PGS with time sharing for this example. Since the IGS designs perform similarly, for this example, we provide only the time sharing for our proposed IGS design in Section III.

The joint variance and covariance IGS algorithm in [15] 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. Since the solution of the SDR in [15] is not ensured to be rank-one, it does not necessarily obtain a valid solution, and a Gaussian randomization procedure is employed to obtain a rank-one solution. The solution obtained by the randomization procedure is not ensured to fulfill any optimality condition, which is in contrast with our proposed algorithm, which converges to a stationary point of (11). That may be the reason why our algorithm provides a better average symmetric rate than SDR for high SNR in this scenario. It is also worth mentioning that our proposed algorithms are more general since they consider additive asymmetric HWD, while the algorithm in [15] can only be applied for ideal devices.

---

[4]We derive the achievable rate region with TS by taking the convex hull operation over the corresponding achievable rate regions [15]. It is worth mentioning that time sharing results in the convex hull operation when power constraint is considered for each operational point. The achievable rate region with time sharing might be enlarged if an average power constraint over different operational point is considered [56]. However, this analysis is outside of the scope of this paper.
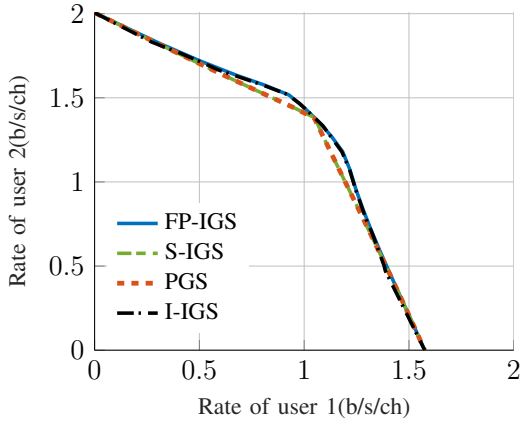
(a) $P = 1$ (SNR= 0 dB)
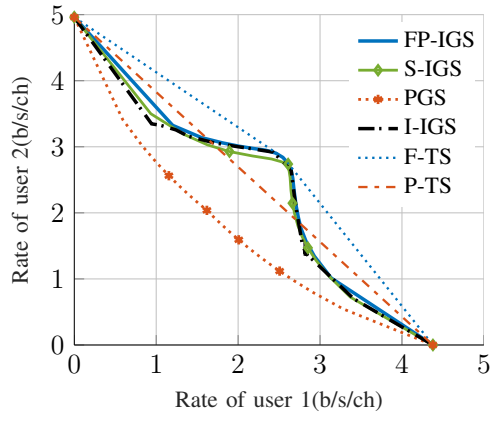
(b) $P = 10$ (SNR= 10 dB)

Fig. 3: Achievable rate region for ideal devices and channel realization $\mathbf{H}_1$ in (43).
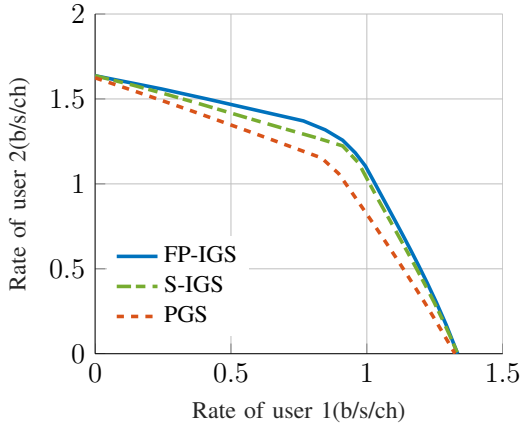
(a) $\sigma_\eta^2 = 0.2$
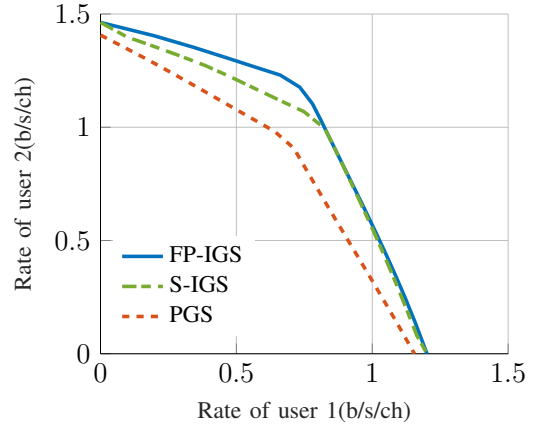
(b) $\sigma_\eta^2 = 0.5$

Fig. 4: Achievable rate region for $|\tilde{\sigma}_\eta^2| = \sigma_\eta^2$, $P = 1$, and channel realization $\mathbf{H}_1$ in (43).

B. Non-ideal devices

In this subsection, we consider the effect of HWD on the performance of the two-user IC. Throughout this subsection, we consider the same statistics for HWD in all devices. In Fig. 4, we show the rate region for $\mathbf{H}_1$ and $P = 1$ under maximally improper HWD[5] noise. As shown in Fig. 3a, 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 additive asymmetric HWD. Note that even in point-to-point communications, PGS is in general suboptimal for asymmetric HWD, 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 $\tilde{\sigma}_\eta^2 = 0$, $P = 1$ (SNR= 0 dB), and channel realization

$$\mathbf{H}_2 = \begin{bmatrix} 0.3764e^{i1.4381} & 0.4029e^{i0.9486} \\ 1.8542e^{i2.8153} & 0.6277e^{i2.3697} \end{bmatrix}. \quad (44)$$

We take $\tilde{\sigma}_\eta^2 = 0$, i.e., symmetric (proper) HWD. We can observe that IGS enlarges the rate region even for proper HWD with high noise variance, i.e., $\sigma_\eta^2 = 0.5$ and $\sigma_\eta^2 = 1$. It is worth mentioning that the PGS design is Pareto-optimal in the presence of additive symmetric HWD. As can be observed, our IGS design in Section III with time sharing

[5]Maximally improper HWD happens when the in-phase and quadrature-phase noises are completely correlated [30].

outperforms the Pareto-optimal PGS with time sharing for these examples.

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., $\mathcal{CN}(0, 1, 0)$.

In Fig. 6, we consider the effect of the variance of the HWD noise on the average symmetric rate of users ($\alpha_1 = \alpha_2 = 0.5$) for $P = 20$. In this figure, we consider proper ($\tilde{\sigma}_\eta^2 = 0$) and maximally improper ($\tilde{\sigma}_\eta^2 = \sigma_\eta^2$) HWD noise. We observe that our proposed algorithm with FP outperforms the other algorithms for maximally improper HWD noise. Moreover, in Fig. 6a, our proposed IGS algorithms perform better than PGS for proper HWD noise with different variances, which is Pareto-optimal PGS in this case. Furthermore, our simplified algorithm outperforms the IGS algorithm in [15] in the presence of HWD. However, the performance improvement by our algorithms is minor for proper HWD with high noise variance, where our algorithms only provide 5% improvement over PGS when $\sigma_\eta^2 = 1$ for this example.

Figure 7 shows the effect of the circularity coefficient of the HWD noise on the symmetric rate for $P = 20$. As can be observed, the benefits of employing IGS increase
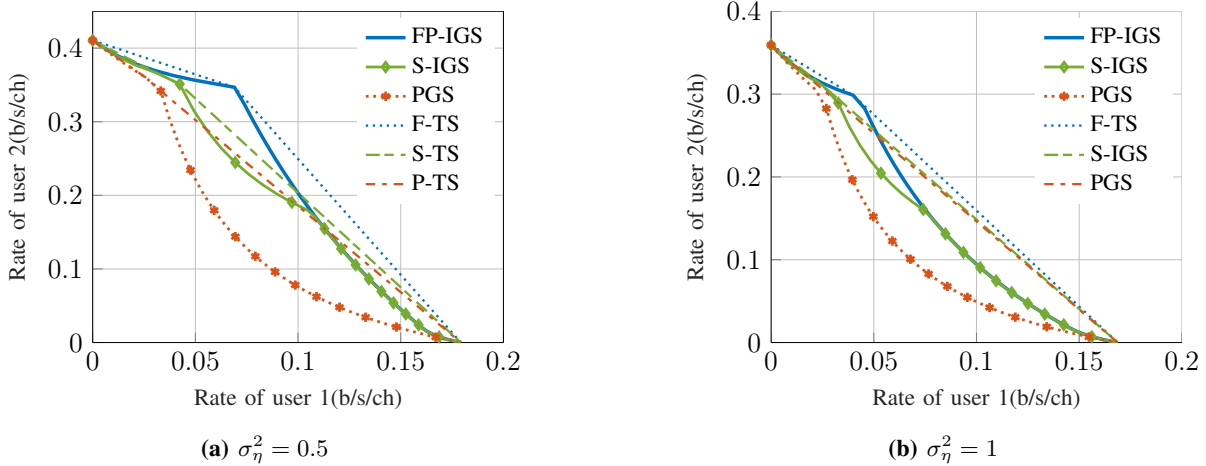
(a) $\sigma_\eta^2 = 0.5$

(b) $\sigma_\eta^2 = 1$

Fig. 5: Achievable rate region for $\tilde{\sigma}_\eta^2 = 0$, $P = 1$, and channel realization $\mathbf{H}_2$ in (44).

(a) $\tilde{\sigma}_\eta^2 = 0$
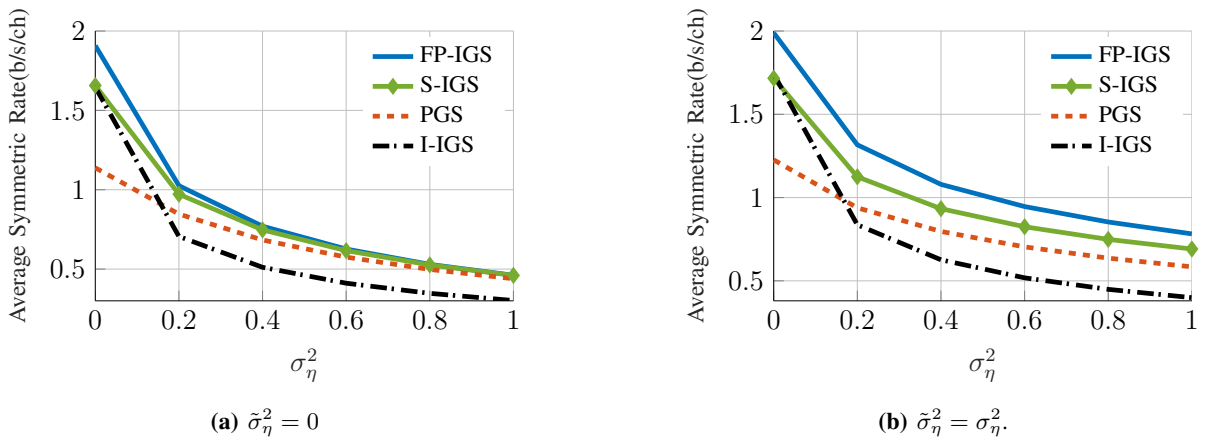
(b) $\tilde{\sigma}_\eta^2 = \sigma_\eta^2$.

Fig. 6: Average symmetric rate versus the variance of the HWD noise for $P = 20$.

(a) $\sigma_\eta^2 = 0.5$
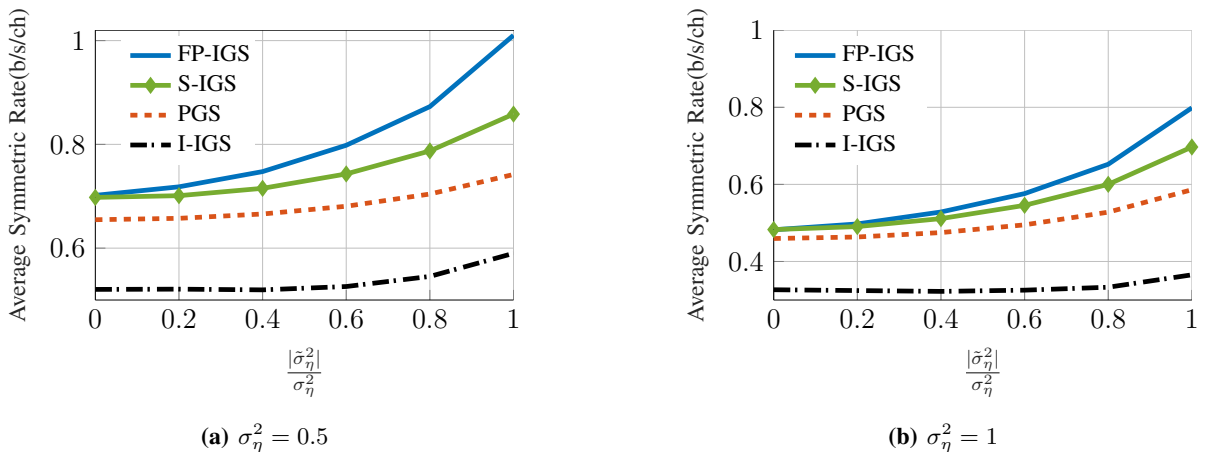
(b) $\sigma_\eta^2 = 1$

Fig. 7: Average symmetric rates versus the circularity coefficient of HWD noise for $P = 20$.

with the circularity coefficient of the HWD noise, and there is a considerable performance improvement by IGS in maximally improper HWD noise. We emphasize that PGS is suboptimal, even in interference-free communications, under asymmetric HWD. Our proposed IGS design with FP outperforms the other algorithms, especially in highly asymmetric HWD noise. When the variance of the HWD 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 HWD. Since the simplified algorithm

has less computational cost, it can be employed for proper HWD noise when the variance of the HWD noise is high, i.e., $\sigma_\eta^2 \geq 0.5$. However, our proposed algorithm based on FP outperforms the other algorithm in low-power HWD noise and/or highly asymmetric HWD noise. Note that, since the IGS algorithm in [15] is proposed for ideal devices and does not consider HWD, it performs worse than the proposed PGS, which considers additive symmetric HWD, from the average symmetric rate point of view, even when the HWD noise is maximally improper.

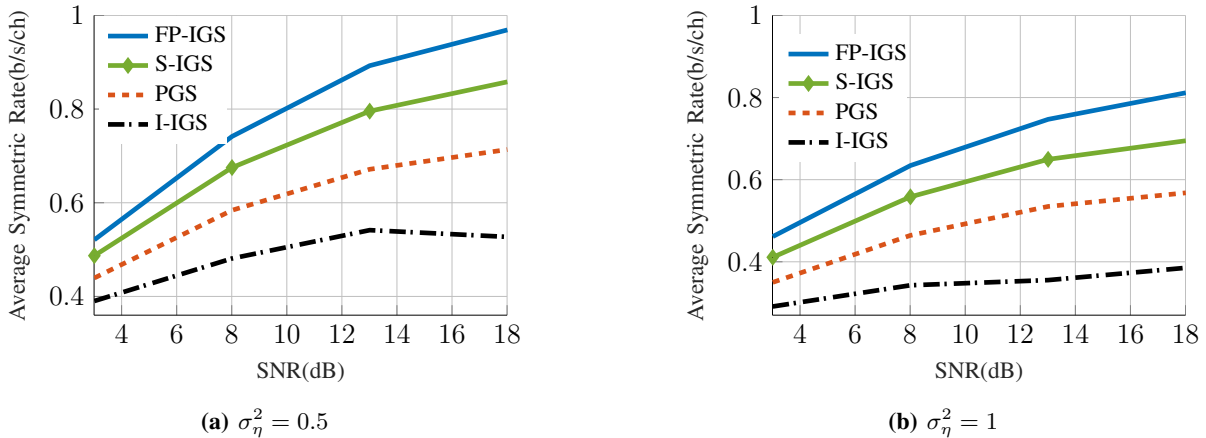In Fig. 8, we consider the effect of the power budget on

(a) $\sigma_\eta^2 = 0.5$

(b) $\sigma_\eta^2 = 1$

Fig. 8: Average symmetric rate versus SNR for $\tilde{\sigma}_\eta^2 = 0.9\sigma_\eta^2$.

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.

VI. CONCLUSION

In this paper, we considered a two-user IC with additive asymmetric HWD 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 algorithm, which is based on MM and the generalized Dinkelbach algorithm, obtains a stationary point of the PSINR region. In this algorithm, we jointly optimize the powers and complementary variances. We also proposed a simplified algorithm that has lower computational complexity. This simplified algorithm is based on the separate optimization of the powers and complementary variances. Through numerical examples, we showed that the proposed approaches enlarge the achievable rate region and outperform PGS and existing IGS algorithms, especially as the HWD becomes more asymmetric.

APPENDIX A
PROOF OF LEMMA 3

In order to approximate $u_k(\mathbf{p}, \mathbf{q})$ and $v_k(\mathbf{p}, \mathbf{q})$, we employ convex-concave (or concave-convex) procedure (CCP), in which the convex (concave) 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 $u(\mathbf{x})$ around the point $\mathbf{x}_0$ is obtained through its Taylor expansion as [27], [57]

$$\Gamma(\mathbf{x}) \approx \Gamma(\mathbf{x}_0) + 2\Re\left[\left(\frac{\partial\Gamma(\mathbf{x})}{\partial\mathbf{x}}\Big|_{\mathbf{x}=\mathbf{x}_0}\right)^T(\mathbf{x}-\mathbf{x}_0)\right], \quad (45)$$

where $\mathbf{x}$ is a complex vector. In order to apply the CCP to $u_k(\mathbf{p}, \mathbf{q})$, we have to differentiate the convex part in (18) with respect to $\mathbf{p}$, which is straightforward since it is a real function on a real domain and consequently, analytic in $\mathbf{p}$. The derivative of $(\sigma^2 + \mathbf{a}_k^T\mathbf{p})^2$ with respect to $\mathbf{p}$ is

$$\frac{\partial(\sigma^2 + \mathbf{a}_k^T\mathbf{p})^2}{\partial\mathbf{p}} = 2\mathbf{a}_k(\sigma^2 + \mathbf{a}_k^T\mathbf{p}), \quad (46)$$

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

$$(\sigma^2 + \mathbf{a}_k^T\mathbf{p})^2 \simeq (\sigma^2 + \mathbf{a}_k^T\mathbf{p}^{(m)})^2 + 2(\sigma^2 + \mathbf{a}_k^T\mathbf{p}^{(m)})\mathbf{a}_k^T(\mathbf{p} - \mathbf{p}^{(m)}). \quad (47)$$

By substituting (47) in (18), we can derive $\tilde{u}_k(\mathbf{p}, \mathbf{q})$.

In order to convexify $v_k(\mathbf{p}, \mathbf{q})$, we have to differentiate the concave part in (19) with respect to $\mathbf{p}$ and $\mathbf{q}$. The derivative of $|\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p}|^2$ with respect to $\mathbf{p}$ is also straightforward since it is analytic in $\mathbf{p}$:

$$\frac{\partial|\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p}|^2}{\partial\mathbf{p}} = 2\Re\left[\tilde{\mathbf{f}}_k(\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p})\right]. \quad (48)$$

The term $|\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p}|^2$, on the other hand, is not analytic in $\mathbf{q}$ since it is a real-valued function while $\mathbf{q}$ is a complex vector [27], [57]. Thus, we have to employ Wirtinger calculus to obtain the derivative of $|\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p}|^2$ with respect to $\mathbf{q}$. By Wirtinger calculus, we treat $\mathbf{q}$ and $\mathbf{q}^*$ as two independent complex variables [27], [57]. Thus, we take the derivative of $|\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p}|^2$ with respect to $\mathbf{q}$ while treating $\mathbf{q}^*$ as a constant, which results in

$$\frac{\partial|\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p}|^2}{\partial\mathbf{q}} = \mathbf{g}_k^*(\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p})^*. \quad (49)$$

Now by (45), we can approximate $|\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p}|^2$ as an affine function as

$$|\mathbf{g}_k^H\mathbf{q} + \tilde{\mathbf{f}}_k^H\mathbf{p}|^2 \simeq |\mathbf{g}_k^H\mathbf{q}^{(m)} + \tilde{\mathbf{f}}_k^H\mathbf{p}^{(m)}|^2 + 2\Re\left(\tilde{\mathbf{f}}_k^H(\mathbf{g}_k^H\mathbf{q}^{(m)} + \tilde{\mathbf{f}}_k^H\mathbf{p}^{(m)})^*\right)(\mathbf{p} - \mathbf{p}^{(m)}) + 2\Re\left((\mathbf{g}_k^H\mathbf{q}^{(m)} + \tilde{\mathbf{f}}_k^H\mathbf{p}^{(m)})^*\mathbf{g}_k^H(\mathbf{q} - \mathbf{q}^{(m)})\right). \quad (50)$$

By substituting (50) in (19), we can obtain $\tilde{v}_k(\mathbf{p}, \mathbf{q})$.

APPENDIX B
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 (32). Let us first consider the two linear constraints in (32b), which can be written as (51) and (52), shown at the top of the next page. We can construct $\mathbf{A}$ in (34) by the coefficients of p_1 and p_2 in (51) and (52). It is worth mentioning that the non-diagonal elements of $\mathbf{A}$ in (34) are non-positive since $\sqrt{1 + \alpha_1 E'} \geq 1$ and $\sqrt{1 + \alpha_2 E'} \geq 1$. Thus, if the diagonal
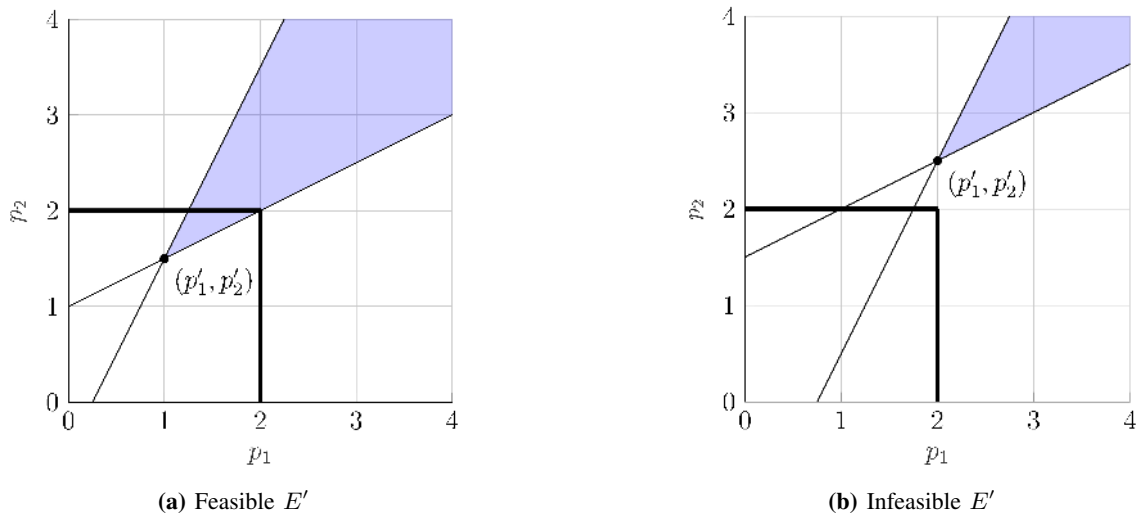
(a) Feasible E'

(b) Infeasible E'

Fig. 9: The constraints of (32) in the power plane.

$$|h_{11}|^2\Big(1-\sigma_{\eta_{11}}^2(\sqrt{1+\alpha_1 E'}-1)\Big)p_1 - |h_{21}|^2(1+\sigma_{\eta_{21}}^2)(\sqrt{1+\alpha_1 E'}-1)p_2 \geq (\sqrt{1+\alpha_1 E'}-1)\sigma^2, \tag{51}$$

$$-|h_{12}|^2(1+\sigma_{\eta_{12}}^2)(\sqrt{1+\alpha_2 E'}-1)p_1 + |h_{22}|^2\Big(1-\sigma_{\eta_{22}}^2(\sqrt{1+\alpha_2 E'}-1)\Big)p_2 \geq (\sqrt{1+\alpha_2 E'}-1)\sigma^2. \tag{52}$$

elements of $\mathbf{A}$ are not positive, there is no positive power pair that satisfies (51) and (52) simultaneously. Hence, in the following, we assume without loss of generality that $\mathbf{A}$ has strictly positive diagonal elements and strictly negative non-diagonal elements.

We can rewrite (51) and (52) as

$$[\mathbf{A}]_{11}p_1 \geq -[\mathbf{A}]_{12}p_2 + y_1, \tag{53}$$

$$[\mathbf{A}]_{22}p_2 \geq -[\mathbf{A}]_{21}p_1 + y_2, \tag{54}$$

where $\mathbf{y} = \left[\ (\sqrt{1+\alpha_1 E'}-1)\sigma^2 \quad (\sqrt{1+\alpha_2 E'}-1)\sigma^2 \ \right]^T$. Moreover, $[\mathbf{A}]_{ij}$, and y_i for $i,j \in \{1,2\}$ are the ijth element of $\mathbf{A}$, and the ith element of $\mathbf{y}$, respectively. If we decouple the inequalities, we end up with

$$\det(\mathbf{A})p_1 \geq [\mathbf{A}]_{22}y_1 - [\mathbf{A}]_{12}y_2, \tag{55}$$

$$\det(\mathbf{A})p_2 \geq -[\mathbf{A}]_{21}y_1 + [\mathbf{A}]_{11}y_2. \tag{56}$$

The right-hand sides (RHS) in (55) and (56) are positive for a feasible E' as mentioned before. Note that if $\det(\mathbf{A}) < 0$, there are no positive power pairs that satisfy (55) and (56) for the given structure of $\mathbf{A}$ in (34). Thus, we consider $\det(\mathbf{A}) > 0$, which yields

$$p_1 \geq p_1' = \frac{[\mathbf{A}]_{22}y_1 - [\mathbf{A}]_{12}y_2}{\det(\mathbf{A})}, \tag{57}$$

$$p_2 \geq p_2' = \frac{-[\mathbf{A}]_{21}y_1 + [\mathbf{A}]_{11}y_2}{\det(\mathbf{A})}, \tag{58}$$

or equivalently $\mathbf{p} = \left[\ p_1 \quad p_2 \ \right]^T \geq \mathbf{A}^{-1}\mathbf{y}$, where p_1' and p_2' are the intersecting point given in (33). Hence, the intersecting point provides the minimum positive power pairs that satisfy (51) and (52). If p_1' and p_2' satisfy the power constraint, E' is feasible (Fig. 9.a). Otherwise, E' is infeasible (Fig. 9.b).

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] 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.

[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] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

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

[21] O. Amin, W. Abediseid, and M.-S. Alouini, "Overlay spectrum sharing using improper Gaussian signaling," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 1, pp. 50–62, 2017.

[22] M. Soleymani, C. Lameiro, P. J. Schreier, and I. Santamaria, "Energy-efficient design for underlay cognitive radio using improper signaling," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Processing (ICASSP)*, 2019, pp. 4769–4773.

[23] 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.

[24] ——, "Energy-Efficient improper signaling for K-User interference channels," in *Proc. IEEE European Signal Processing Conference (EUSIPCO)*, 2019.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

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

[30] 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.

[31] 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.

[32] 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.

[33] 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.

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

[35] 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.

[36] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492–498, 1967.

[37] 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.

[38] 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.

[39] 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.

[40] T. Schenk, *RF imperfections in high-rate wireless systems: impact and digital compensation*. Springer Science & Business Media, 2008.

[41] 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.

[42] ——, "Asymmetric modulation for hardware impaired systems-error probability analysis and receiver design," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1723–1738, 2019.

[43] I. Santamaria, P. Crespo, C. Lameiro, and P. Schreier, "Information-theoretic analysis of a family of improper discrete constellations," *Entropy*, vol. 20, no. 1, p. 45, 2018.

[44] 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.

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

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

[47] 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.

[48] I. Santamaría, 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.

[49] C. Hellings and W. Utschick, "On the worst-case noise in Gaussian MIMO systems with proper and with improper signaling," in *WSA 2017; 21th International ITG Workshop on Smart Antennas*. VDE, 2017, pp. 1–7.

[50] A. Rubinov, H. Tuy, and H. Mays, "An alogrithm for monotonic global optimization problems," *Optimization*, vol. 49, no. 3, pp. 205–221, 2001.

[51] L. Liu, R. Zhang, and K.-C. Chua, "Achieving global optimality for weighted sum-rate maximization in the k-user Gaussian interference channel with multiple antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1933–1945, 2012.

[52] A. Zappone, E. Björnson, L. Sanguinetti, and E. Jorswieck, "Globally optimal energy-efficient power control and receiver design in wireless networks," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2844–2859, 2017.

[53] J. Duchi, "Sequential convex programming," *Lecture Notes, EE364b, Stanford University*, 2018.

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

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

[56] C. Hellings and W. Utschick, "Improper signaling versus time-sharing in the two-user Gaussian interference channel with TIN," *arXiv preprint arXiv:1808.01611*, 2018.

[57] 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.