Rate Region of the K-user MIMO Interference Channel with Imperfect Transmitters

Mohammad Soleymani*, Ignacio Santamaria†, Behrouz Maham‡, and Peter J. Schreier*

*Signal & System Theory Group, Universität Paderborn, Germany
†Dept. Communications Engineering, Universidad de Cantabria, Spain
‡ Dept. of Electrical and Electronic Engineering, School of Engineering, Nazarbayev University, Astana, Kazakhstan
Email: {mohammad.soleymani,peter.schreier}@sst.upb.de,i.santamaria@unican.es, behrouz.maham@nu.edu.kz

Abstract—**This paper studies the rate region of a multiple-input, multiple-output (MIMO) system with imperfect transmitters when interference is treated as noise at the receiver side. We consider a K-user MIMO interference channel (IC) in which the transmitters suffer from an additive hardware distortion (HWD) modeled as spatially uncorrelated Gaussian noise with covariance matrix proportional to the transmit covariance matrix. We employ the difference of convex programming (DCP) technique to solve the rate-region optimization problem and obtain its stationary points. Our proposed HWD-aware algorithm outperforms the HWD-unaware design that disregards HWD. Our results show that the performance of the K-user MIMO IC is highly affected by HWD, especially in high signal-to-noise-ratio scenarios.**

Index Terms—**Achievable rate region, additive hardware distortions, difference of convex programming, interference channel, MIMO systems.**

I. INTRODUCTION

Hardware nonidealities have always been among the main bottlenecks for wireless communications. In 5G and Beyond systems, which use millimeter-wave bands, the impact of hardware distortions (HWD) on the system performance is even more critical [1]. Due to hardware imperfections like quantization noise, phase noise, amplifier nonlinearities, and I/Q imbalance, the transmitters might suffer from additive HWD, which can limit the system performance [2]–[13].

Another performance limitation in modern wireless communications comes from interference. Indeed, interference management techniques are among the main concerns for 5G [14]. Due to limited resources and the great need for high data rates, modern wireless systems are mostly interference-limited. Treating interference as noise (TIN) is one of the main interference management techniques due to its simplicity. For this reason, in this work we will consider that all users apply TIN at the receiver side.

The performance of various single antenna interference-limited systems with imperfect devices has been studied in several works [3]–[5]. For example, the paper [3] derived the rate region of a 2-user SISO interference channel (IC) in the presence of additive asymmetric HWD. The performance of dual-hop relaying with additive HWD is studied in [4]. In [5], the authors studied a device-to-device millimeter wave communication system with HWD under Nakagami-m fading channel and derived the outage probability of the system.

HWD can be more harmful in multiple-antenna systems, and HWD-aware schemes are anticipated to play a key role in next generations of multiple-antenna systems [6]–[13], [15]. The performance of multiple-antenna systems with imperfect devices has been investigated in [6]–[13], [15] for different scenarios. The work [9] studied the outage performance of a MIMO amplify-and-forward relay channel with HWD. The paper [8] considered the uplink and downlink of a cellular system with massive MIMO and proposed a system model for HWD at the transceivers of a massive MIMO system. The paper [11] proposed a framework for rate analysis in a multicell massive MIMO system with Rician fading channels and different types of HWD. In [15], the authors studied the achievable rate of a point-to-point MIMO system with HWD and I/Q imbalance. When the HWD includes I/Q imbalance, the received signal is a function of the widely linear transform of the transmitted signal and additive noise, which results in improper or circularly asymmetric noise [15].

In this paper, we derive the achievable rate region of the K-user MIMO IC with additive HWD at the transmitters. To the best of our knowledge, this problem has not been considered before in the literature. We assume non-ideal transmitters that produce additive HWD modeled as a Gaussian distribution [2]–[4], [6]–[8], [10]. We obtain a stationary point of the rate region by means of a difference of convex programming (DCP) technique. DCP is an iterative optimization technique suited to optimization problems in which the objective function and/or constraints are a difference of two convex/concave functions [16]. The users rates are a difference of two concave/convex functions when interference is treated as noise. Hence, DCP can be applied to obtain a stationary point of the achievable rate region.

Our numerical results show that the benefits of employing our HWD-aware scheme increase with the number of users K and the level of HWD. We also observe that HWD may significantly degrade the rate performance of the K-user MIMO IC, especially at high SNR regimes.

The rest of this paper is organized as follows. In Section II, we describe the system model and formulate the rate region problem. In Section III, we propose our HWD-aware scheme based on DCP. Section IV presents some numerical results.

II. SYSTEM MODEL

A. Network Scenario

We consider a K-user MIMO IC with non-ideal transmitters, as shown in Fig. 1. Without loss of generality, we assume
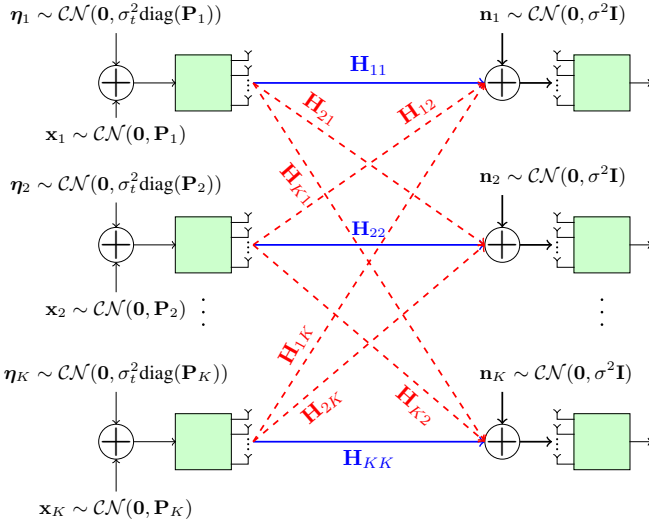
Fig. 1: K-user MIMO IC with transmit noise provoked by HWD.

that each transceiver pair has the same number of antennas and generates HWD with similar statistics. Each user has N_t transmit antennas and N_r receive antennas. We further assume that the non-ideal transmitter of each user produces a transmit noise that is modeled by zero-mean additive proper Gaussian noise [3]–[8], [10]. Hence, the received signal of user k is

$$\mathbf{y}_k = \sum_{i=1}^{K} \mathbf{H}_{ki}(\mathbf{x}_i + \boldsymbol{\eta}_i) + \mathbf{n}_k \qquad (1)$$

$$= \underbrace{\mathbf{H}_{kk}\mathbf{x}_k}_{\text{desired signal}} + \underbrace{\sum_{i=1,i\neq k}^{K} \mathbf{H}_{ki}\mathbf{x}_i}_{\text{interference}} + \underbrace{\sum_{i=1}^{K} \mathbf{H}_{ki}\boldsymbol{\eta}_i + \mathbf{n}_k}_{\text{aggregated noise}}, \qquad (2)$$

where $\mathbf{x}_i \in \mathbb{C}^{N_t \times 1}$, $\boldsymbol{\eta}_i \in \mathbb{C}^{N_t \times 1}$, $\mathbf{H}_{ki} \in \mathbb{C}^{N_t \times N_r}$, and $\mathbf{n}_k \in \mathbb{C}^{N_r \times 1}$ are, respectively, the transmit symbols of user i, the additive HWD noise, the channel matrix between transmitter i and receiver k, and the additive noise at receiver k.

We represent the covariance matrix of the transmitted signal of user k by $\mathbf{P}_k = \mathbb{E}\{\mathbf{x}_k\mathbf{x}_k^H\}$. We consider $\mathbb{E}\{\mathbf{n}_k\mathbf{n}_k^H\} = \sigma^2\mathbf{I}_{N_r}$ for $k = 1, \cdots, K$, where σ^2 is the variance of each noise component. We adhere to the HWD model used in [6]–[8], [10], which assumes the transmit noise is a zero-mean complex Gaussian random variable with covariance matrix $\sigma_t^2\mathrm{diag}(\mathbf{P})$: $\boldsymbol{\eta}_t \sim \mathcal{CN}(\mathbf{0}, \sigma_t^2\mathrm{diag}(\mathbf{P}))$. It is worth emphasizing that the covariance matrix of the HWD noise is diagonal, which means that there is no correlation between noise components. Moreover, the variance of each noise component is a linear function of the transmit power at the corresponding antenna. Thus, the higher transmit power at an antenna, the higher HWD at the corresponding antenna.

When interference is treated as noise, the rate of user k as a function of the transmit covariance matrices is [17], [18]

$$R_k = \underbrace{\log_2 \det\left(\boldsymbol{\Phi}_{k,1}\left(\{\mathbf{P}_i\}_{i=1}^K\right)\right)}_{\triangleq\, r_{k,1}\left(\{\mathbf{P}_i\}_{i=1}^K\right)} - \underbrace{\log_2 \det\left(\boldsymbol{\Phi}_{k,2}\left(\{\mathbf{P}_i\}_{i=1}^K\right)\right)}_{\triangleq\, r_{k,2}\left(\{\mathbf{P}_i\}_{i=1}^K\right)}, \qquad (3)$$

where

$$\boldsymbol{\Phi}_{k,1}\left(\{\mathbf{P}_i\}_{i=1}^K\right) = \sum_{i=1}^{K} \mathbf{H}_{ki}\mathbf{P}_i\mathbf{H}_{ki}^H + \sigma_t^2 \sum_{i=1}^{K} \mathbf{H}_{ki}\mathrm{diag}(\mathbf{P}_i)\mathbf{H}_{ki}^H$$
$$+ \sigma^2\mathbf{I}_{N_r}, \qquad (4)$$

$$\boldsymbol{\Phi}_{k,2}\left(\{\mathbf{P}_i\}_{i=1}^K\right) = \boldsymbol{\Phi}_{k,1}\left(\{\mathbf{P}_i\}_{i=1}^K\right) - \mathbf{H}_{kk}\mathbf{P}_k\mathbf{H}_{kk}^H. \qquad (5)$$

Note that the rate of user k is a difference of two concave functions in $\{\mathbf{P}_i\}_{i=1}^K$: $r_{k,1}\left(\{\mathbf{P}_i\}_{i=1}^K\right)$ and $r_{k,2}\left(\{\mathbf{P}_i\}_{i=1}^K\right)$ in (3).

B. Problem Statement

In this paper, we aim at obtaining the achievable rate region for the described K-user IC. To this end, we employ the rate profile technique, which leads to the following optimization problem [3]

$$\max_{R, \{\mathbf{P}_k\}_{k=1}^K} \quad R \qquad (6a)$$

$$\text{s.t.} \quad R_k \geq \lambda_k R, \qquad \forall k, \qquad (6b)$$

$$0 \leq \mathrm{Trace}(\mathbf{P}_k) \leq p_k, \quad \mathbf{P}_k \succcurlyeq \mathbf{0}, \quad \forall k, \qquad (6c)$$

where p_k is the power budget for user k. Moreover, the rate profile parameters are fixed positive values, $\lambda_k \geq 0$, satisfying $\sum_{k=1}^K \lambda_k = 1$. We can obtain the boundary of the rate region by solving (6) for different rate-profile parameters, i.e., $\{\lambda_k\}_{k=1}^K$. Unfortunately, (6) is not a convex optimization problem, and it is complicated to obtain its global optimal solution in polynomial time. However, since the user rate in (3) is a difference of two concave functions, we are able to derive a stationary point of (6) by applying the difference of convex programming (DCP) technique [16], as will be explained in the next section.

III. PROPOSED ALGORITHM

In this section, we employ DCP to obtain a stationary point of (6). The optimization problem (6) is not convex due to the non-concave part of the rates in (3), i.e., $-r_{k,2}(\cdot)$. DCP is an iterative optimization algorithm and may be applied to optimization problems in which the objective function and/or the constraints are a difference of two convex/concave functions [16]. It solves such a non-convex optimization problem by solving a sequence of convex optimization problems. In each iteration, the original optimization problem is approximated by a convex surrogate optimization problem. To this end, DCP employs convex-concave procedure (CCP), in which the non-concave part of the rates is approximated by a linear function. The CCP employs the first-order Taylor expansion to approximate a convex function by a linear (affine) function.

In this paper, we apply DCP to (6) and solve it iteratively. In each iteration, we approximate the rate in (3) by a concave lower bound. In the following, we first present a lemma, which is used to derive the lower bound for the achievable rate in Theorem 1.

Lemma 1. *An affine upper bound for* $\log \det(\mathbf{W})$ *is*

$$\log \det(\mathbf{W}) \leq \log \det \left(\mathbf{W}^{(l)}\right) + \text{Trace}((\mathbf{W}^{(l)})^{-1}(\mathbf{W} - \mathbf{W}^{(l)})), \quad (7)$$

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

Theorem 1. *A concave lower-bound for the rate of user* k *at the* l*th iteration is*

$$R_k \geq \tilde{R}_k^{(l)} = r_{k,1} \left(\{\mathbf{P}_i\}_{i=1}^K\right) - r_{k,2} \left(\{\mathbf{P}_i^{(l-1)}\}_{i=1}^K\right)$$
$$- \sum_{i=1}^K \text{Trace} \left(\mathbf{B}_i^{(l-1)^H} \left(\mathbf{P}_i - \mathbf{P}_i^{(l-1)}\right)\right), \quad (8)$$

where $\{\mathbf{P}_k^{(l-1)}\}_{k=1}^K$ *and* $r_{k,2} \left(\{\mathbf{P}_i^{(l-1)}\}_{i=1}^K\right)$ *are, respectivly, the covariance matrices of the users and* $r_{k,2}(\cdot)$ *at the previous iteration. Moreover,* $\mathbf{B}_i^{(l-1)}$ *is the derivative of* $r_{k,2}\left(\{\mathbf{P}_i\}_{i=1}^K\right)$ *with respect to* $\mathbf{P}_i$ *at the* $(l-1)$*th iteration, i.e.,* $\mathbf{B}_i^{(l-1)} = \frac{\partial r_{k,2}\left(\{\mathbf{P}_i\}_{i=1}^K\right)}{\partial \mathbf{P}_i}|_{\{\mathbf{P}_k^{(l-1)}\}_{k=1}^K}$, *where* $\frac{\partial r_{k,2}(\cdot)}{\partial \mathbf{P}_i}$ *is*

$$\frac{\partial r_{k,2}(\cdot)}{\partial \mathbf{P}_i} = \begin{cases} \frac{\sigma_t^2}{\ln 2} \text{diag} \left(\mathbf{H}_{kk}^H \boldsymbol{\Phi}_{k,2}^{-1} \left(\{\mathbf{P}_i\}_{i=1}^K\right) \mathbf{H}_{kk}\right) & \text{for } i = k, \\ \frac{\sigma_t^2}{\ln 2} \text{diag} \left(\mathbf{H}_{ki}^H \boldsymbol{\Phi}_{k,2}^{-1} \left(\{\mathbf{P}_i\}_{i=1}^K\right) \mathbf{H}_{ki}\right) \\ + \frac{1}{\ln 2} \mathbf{H}_{ki}^H \boldsymbol{\Phi}_{k,2}^{-1} \left(\{\mathbf{P}_i\}_{i=1}^K\right) \mathbf{H}_{ki} & \text{for } i \neq k. \end{cases}$$
$$(9)$$

Proof. Using CCP, we approximate the convex part of R_k, i.e., $-r_{k,2}$, by an affine function in each iteration, while the concave part remains unchanged. In order to approximate $-r_{k,2}$, we employ the bound in Lemma 1 as:

$$- r_{k,2} \left(\{\mathbf{P}_i\}_{i=1}^K\right) \geq -r_{k,2} \left(\{\mathbf{P}_i^{(l-1)}\}_{i=1}^K\right)$$
$$- \sum_{k=1}^K \text{Trace} \left((\frac{\partial r_{k,2}(\cdot)}{\partial \mathbf{P}_i}|_{\{\mathbf{P}_i^{(l-1)}\}_{i=1}^K})^H \left(\mathbf{P}_i - \mathbf{P}_i^{(l-1)}\right)\right).$$
$$(10)$$

The theorem can be proved by substituting (10) in (3). Please refer to [16] for more details on DCP and CCP. □

In order to obtain a stationary point of (6), we solve the following sequence of surrogate optimization problems:

$$\max_{R, \{\mathbf{P}_k\}_{k=1}^K} R \quad \text{s.t.} \quad \tilde{R}_k^{(l)} \geq \lambda_k R \quad \forall k, \quad \text{and (6c)}. \quad (11)$$

The optimization problem in (11) is convex, and its global optimum solution can be derived efficiently. As indicated in [16], DCP converges to the stationary point of the original problem.

IV. NUMERICAL RESULTS

In this section, we provide some numerical results. Here, we assume $\sigma^2 = 1$ and equal power budget for all users, i.e., $p_1 = p_2 = \cdots = p_K = p$. We define the signal-to-noise-ratio (SNR) as SNR$= \frac{p}{\sigma^2}$. The maximum number of iterations of the DCP algorithm is 30.

There are generally two types of performance metrics for rate-region analysis in the literature [19]. One is to consider a
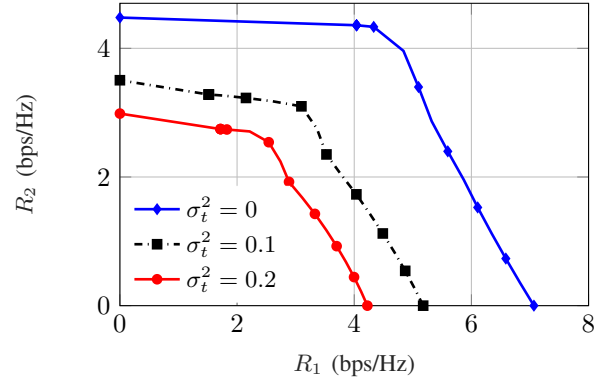

Fig. 2: The achievable rate region of the 2-user IC 2×2 MIMO with SNR$= 10$ dB and different σ_t^2.

specific channel realization and derive the whole rate region for the given channel realization, which is presented in Section IV-A. The other type of performance evaluation is to employ Monte Carlo simulations for a specific point of the rate region, in which the results are averaged over a large number of channel realizations for the specific point of the region as described in Section IV-B.

A. Rate region

In Fig. 2, we show the achievable rate region for a 2-user 2×2 MIMO IC with SNR$= 10$ dB, $\sigma_t^2 = 0, 0.1, 0.2$ and the following channel realization

$$\mathbf{H}_{11} = \begin{bmatrix} -0.0582 - 1.0619i & 0.6289 + 0.4246i \\ -0.6971 - 0.4074i & -1.4091 - 0.5512i \end{bmatrix},$$

$$\mathbf{H}_{12} = \begin{bmatrix} 0.0168 + 0.1373i & 0.5545 + 0.3165i \\ 0.6714 + 0.4017i & 0.6933 + 0.3237i \end{bmatrix},$$

$$\mathbf{H}_{21} = \begin{bmatrix} 0.4203 + 0.7962i & 1.1189 - 0.6048i \\ -0.7680 + 0.1541i & -1.3321 - 0.0836i \end{bmatrix},$$

$$\mathbf{H}_{22} = \begin{bmatrix} 0.5857 + 0.2984i & 0.5236 - 1.1194i \\ -0.1130 + 0.3857i & -0.1024 - 0.0475i \end{bmatrix}.$$

As can be observed, the achievable rate region significantly shrinks when the HWD level σ_t^2 increases.

B. Fairness (symmetric) rate

The results in this subsection have been obtained by averaging 100 independent channel realizations. Every element of each channel matrix is drawn from a complex zero-mean Gaussian distribution with unit variance. We consider the fairness rate, which is obtained for $\lambda_k = \frac{1}{K}$ for $k = 1, \cdots, K$. In this case, each user receives the same achievable rate, and we indeed maximize the minimum achievable rate of the system. That is why this point of the rate region is referred to as the "fairness point". We compare our proposed HWD-aware algorithm (labeled "HWD-A") with the scheme that does not consider HWD (labeled "HWD-U").

In Fig. 3, we show the effect of σ_t^2 on the average achievable fairness rate of the 2-user 4×4 MIMO IC. In order to make the figure clearer, plots (a) and (b) use a different scale for different values of σ_t^2, and we consider only the 2-user MIMO IC. We consider the 3-user IC with a different number
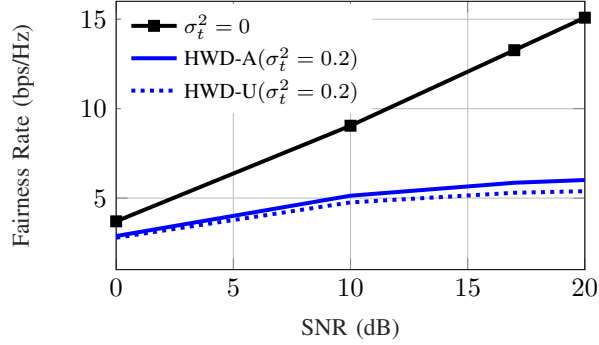
(a) $\sigma_t^2 = 0, 0.2$.
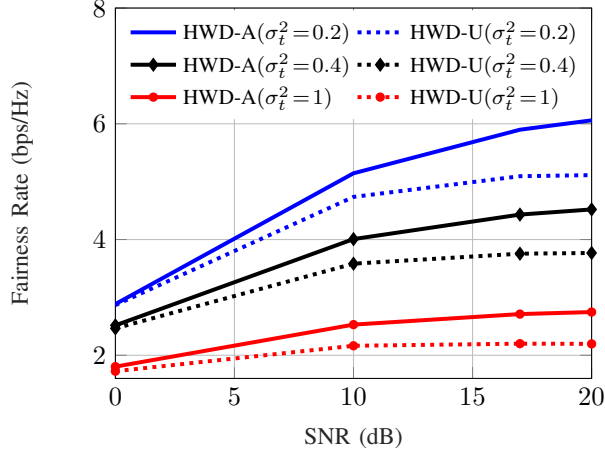
(b) $\sigma_t^2 = 0.2, 0.4, 1$.

Fig. 3: The average fairness rate versus SNR for the 2-user 4×4 MIMO IC with different σ_t^2.
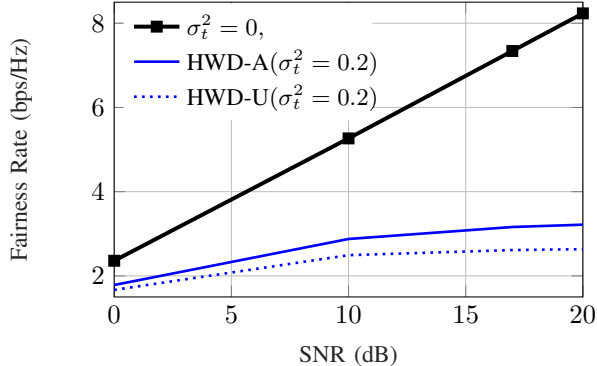

Fig. 4: The average fairness rate versus SNR for the 3-user 3×3 MIMO IC with different σ_t^2.

of antennas later in Fig. 4. As can be observed in Fig. 3, the achievable fairness rate drastically decreases when the devices are not ideal especially at high SNR, where the average fairness rate by our scheme is 15.08 bps/Hz for ideal devices and 6 bps/Hz for $\sigma_t^2 = 0.2$ in this example, which means more than 70 % performance loss by HWD. The reason is that the HWD level increases with the transmission power, which causes a severe performance degradation in high SNR.

Figure 4 shows the average fairness rate of the 3-user 3×3 MIMO IC with $\sigma_t^2 = 0, 0.2$. As can be observed, the average
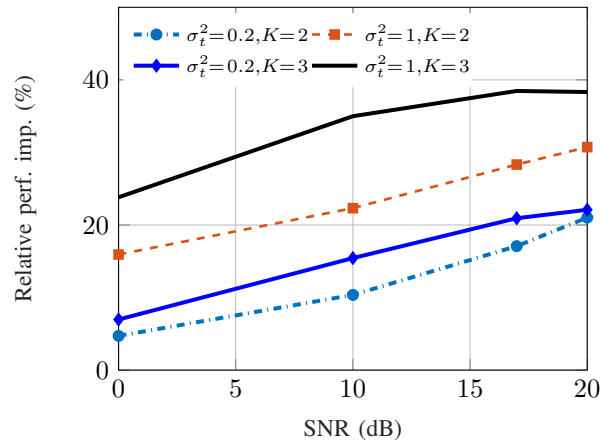

Fig. 5: The relative performance improvement of our algorithm versus SNR for the 2-user and 3-user 3×3 MIMO ICs with different σ_t^2 with respect to the scheme "HWD-U".
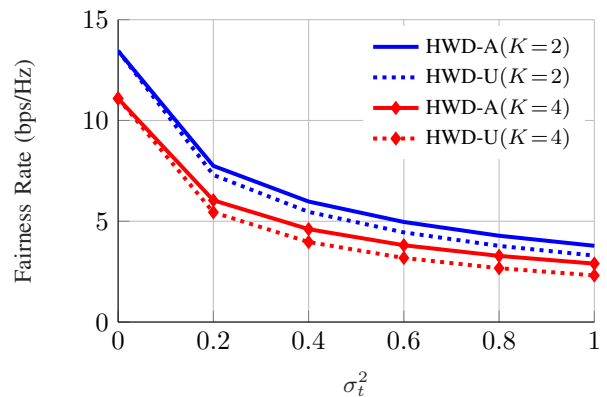

Fig. 6: The average fairness rate versus σ_t^2 for the 2-user and 3-user 6×6 MIMO IC with SNR= 10 dB.

fairness rate significantly decreases with additive HWD similar to the 2-user 4×4 MIMO IC, shown in Fig. 3a. Additionally, our HWD-aware scheme provides more benefits in the 3-user IC. In other words, we observe that when the number of users or antennas increase, to account for HWD is more important.

In Fig. 5, we depict the relative performance improvement by our scheme for the 2-user and 3-user 3×3 MIMO ICs with different σ_t^2. In this figure, both the 2-user and 3-user ICs employ the same number of antennas. As can be observed, the benefits of our scheme increase with power budget and σ_t^2. The reason is that the HWD at each antenna is a linear function of the transmission power at the corresponding antenna, as indicated in Section II-A. Thus, HWD level increases with the transmission power, and the higher HWD is, the higher performance improvement is achieved by our scheme. Furthermore, as indicated in Fig. 4, when K increases it is more important to account for HWD. This is due to the fact that when there are more transmitters with a similar HWD, the receiver experiences more HWD.

In Fig. 6, we show the impact of the HWD level on the average fairness rate of the 2-user and 3-user 6×6 MIMO ICs with SNR= 10 dB. As can be observed, HWD can significantly degrade the system performance. Moreover,
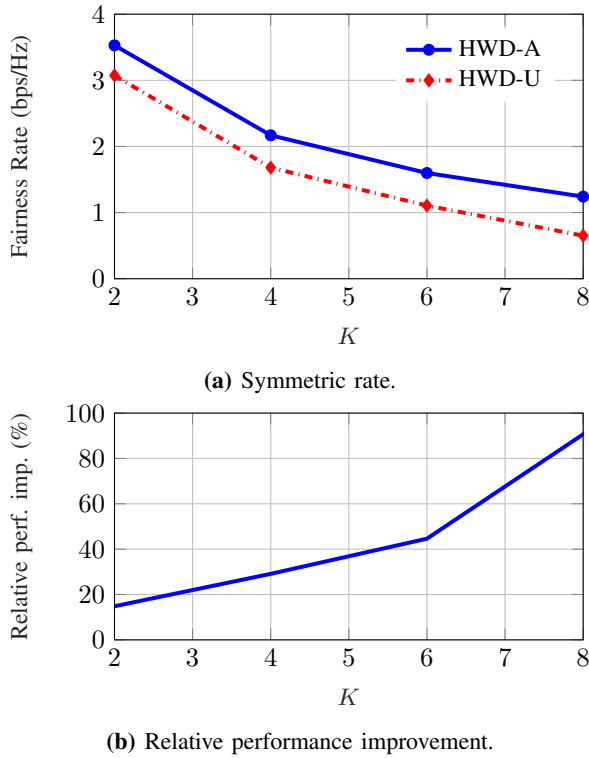
(a) Symmetric rate.

(b) Relative performance improvement.

Fig. 7: The average fairness rate and relative performance of our proposed scheme versus K for the K-user 5×3 MIMO IC with $\sigma_t^2 = 0.5$, and SNR= 0 dB.

we observe again that considering HWD in design is more important when the number of users increases.

In Fig. 7, we show the effect of K on the average fairness rate of the K-user 3×5 MIMO IC with $\sigma_t^2 = 0.5$, and SNR= 0 dB. As can be observed, the fairness rate is decreasing in K. However, the benefits of our HWD-aware scheme increase with K, which shows the importance of our scheme in practical scenarios. As indicated, the HWD at each receiver increases with K for a fixed $\sigma_t^2 > 0$. Hence, our HWD-aware scheme can provide more benefits when K grows due to the increase in HWD level.

V. Conclusion

In this paper, we derived a stationary point of the rate region for the K-user MIMO IC with imperfect transmitters. To this end we employed DCP, which is an iterative optimization technique. We showed that the performance of the system is highly affected by HWD at transmitters. Our proposed scheme outperforms the scheme that does not consider non-idealities at transmitters. Moreover, the benefit of employing our HWD-aware scheme increases with the number of users. As future work, it is interesting to verify how close is the solution of this algorithm to the global optimal solution of the achievable rate region.

References

[1] 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.

[2] E. Boshkovska, D. W. K. Ng, L. Dai, and R. Schober, "Power-efficient and secure WPCNs with hardware impairments and non-linear EH circuit," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2642–2657, 2018.

[3] M. Soleymani, C. Lameiro, I. Santamaria, and P. J. Schreier, "Improper signaling for SISO two-user interference channels with additive asymmetric hardware distortion," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8624–8638, 2019.

[4] 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.

[5] L. Tlebaldiyeva, B. Maham, and T. A. Tsiftsis, "Device-to-device mmwave communication in the presence of interference and hardware distortion noises," *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1607–1610, 2019.

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

[10] 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.

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

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

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

[14] 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.

[15] 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.

[16] 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.

[17] 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.

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

[19] 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.