Pareto Boundary of the Rate Region for Single-Stream MIMO Interference Channels: Linear Transceiver Design

Pan Cao, *Student Member, IEEE*, Eduard A. Jorswieck, *Senior Member, IEEE* and Shuying Shi

Abstract—We consider a multiple-input multiple-output (MIMO) interference channel (IC), where a single data stream per user is transmitted and each receiver treats interference as noise. The paper focuses on the open problem of computing the outermost boundary (so-called Pareto boundary-PB) of the achievable rate region under linear transceiver design. The Pareto boundary consists of the strict PB and non-strict PB. For the two user case, we compute the non-strict PB and the two ending points of the strict PB exactly. For the strict PB, we formulate the problem to maximize one rate while the other rate is fixed such that a strict PB point is reached. To solve this non-convex optimization problem which results from the hard-coupled two transmit beamformers, we propose an alternating optimization algorithm. Furthermore, we extend the algorithm to the multi-user scenario and show convergence. Numerical simulations illustrate that the proposed algorithm computes a sequence of well-distributed operating points that serve as a reasonable and complete inner bound of the strict PB compared with existing methods.

Index Terms—multiple-input multiple-output (MIMO) interference channel (IC), Pareto boundary, alternating optimization, semidefinite programming, fractional programming.

I. INTRODUCTION

IN wireless cellular systems, multiple sectors of different cells share the same time-frequency resource for communication in order to increase the spectral efficiency and occupancy level, while inter-cell interference brings a strong inter-cell coupling and limits the performance. In this paper, we consider first a two-cell environment, where each cell has a base station (BS) with multiple antennas and a mobile station (MS) with multiple antennas. Each BS is intended to communicate with the MS in its own cell while simultaneously interferes the MS in the other cell. And then a multi-cell scenario consisting of multiple interfering BS-MS links will be also considered later. These scenarios are modeled as a two or multi-user MIMO interference channel (IC). The IC is characterized by its capacity region, defined as the set of largest rates that can be simultaneously achieved by the users in the system while making the error probability arbitrary small. A

pragmatic approach that leads to an achievable region or inner bound of the capacity region is based on two assumptions. i) The class of encoding strategies are constrained to use random Gaussian codebooks; ii) The decoders are restricted to treat the interference as Gaussian noise. Herein, based on these two assumptions, we desire to find the *complete achievable rate region* by linear transceiver design.

Form the perspective of optimization, it is well-known that a bi/multi-objective optimization problem usually admits infinite number of *noninferior* solutions (theoretical limits), which form the outermost boundary of achievable performance region, so-called *Pareto boundary* [1]. A noninferior solution on the Pareto boundary is considered to be Pareto-optimal in the sense that no other solution can improve the performance of some objectives without reducing other objective(s). Generally, it is hard to find the Pareto boundary efficiently, but it is significant to study it in order to determine optimal system operations based on Pareto-optimal rate tuples and their associated strategies. In this paper, we propose an algorithm to compute the complete Pareto boundary[1] for the two/multi-user MIMO single-stream IC through linear transceiver design.

How to design linear transceiver schemes to achieve the Pareto boundary has attracted intensive research for several decades. A brief, comprehensive, yet non-exhaustive review of the related works is given as follows.

Parameterization Approaches: The Pareto boundary is characterized by a few parameters. For a two-user multiple-input single-output (MISO) IC, the authors proposed a necessary condition for Pareto-optimal transmit beamformers, i.e., linear combinations of zero-forcing (ZF) and maximum-ratio transmission (MRT) beamformers with two $[0, 1]$-parameters [2]. This parameterization is later used to derive a characterization with only a single parameter in [3], [4]. A general framework for parameterizing Pareto-optimal transmit strategies for multi-user MISO IC was proposed in [5], which is applicable when the utility functions of the systems are monotonic in the received power gains. In [6], the authors proposed a parametrization to characterize the Pareto boundary of the multi-cell MIMO performance region under an assumption that each receiver has only a single effective antenna, while this limiting assumption, in fact, degrades the MIMO IC to the MISO IC.

Computation Approaches: Different from the parameter-

---

[1]When referring to Pareto boundary, we mean the Pareto boundary of achievable rate region unless otherwise specified hereafter.

ization schemes, another approach is to compute a Pareto boundary point directly. One branch is to maximize the rate of one user for a fixed rate of the other users (e.g., in two-user MISO IC [7] and in two-user MIMO IC [8]). Another important branch is to find the intersection point between the Pareto boundary and a ray from the origin (e.g., so-called rate profiles approach in the multi-user MISO IC [9], [10]). The computations approaches are able to compute the *whole* Pareto boundary, if the optimization problems are solved optimally.

Weighted Sum Maximization Approaches: A standard technique for generating the Pareto-optimal solutions to multi-objective optimization problems is to maximize weighted sums of the different objectives for various different settings of the weights. Generally, the weighted sum rate maximization problem for multi-user IC is non-convex. For the multi-user SISO IC, the MAPEL algorithm proposed in [11] transformed the weighted sum rate maximization into a generalized linear fractional programming problem, which can be solved optimally. In [12], the authors jointly utilized the monotonic optimization and rate profile techniques to solve the weighted sum rate maximization optimally at the cost of computation load in multi-user SISO/MISO/SIMO IC. However, it is NP-hard to obtain a global optimal solution of the weighted sum rate maximization for a multi-user MIMO IC (e.g., in [13]). In particular, most algorithms focus on finding only a single sum-rate maximum point, e.g., for the two-user MIMO IC based on pricing in [14], for the two-user MIMO IC based on approximation of sum rate in [15], for the single-stream MIMO IC based on balancing the egoistic and the altruistic behavior in [16], and for the multi-user MIMO IC based on interference alignment in [17]. However, it is well-known that the weighted sum maximization method has two major drawbacks [18]: i) If the Pareto boundary is not convex, there does not exist any weight corresponding to the points on the nonconvex part. Increasing the number of steps of the weighting factor does not resolve this problem; ii) Even if the Pareto boundary is convex, an even spread of weights does not produce an even spread of points on the Pareto boundary. Therefore, weighted sum rate maximization is not a promising method to perform *the complete Pareto boundary*, especially the non-convex boundary.

Game Theoretic Approaches: Game theory as a useful tool has been widely applied to resource allocation in multi-user IC by studying the conflicting or cooperative behavior of the users. Distributed optimization algorithms based on iterative water-filling for the MIMO IC (e.g. [19], [20]) can be modeled as non-cooperative games, where each user is considered as a player that attempts to maximize its own utility selfishly. Such approaches may not converge in general or may converge to the Nash equilibrium (NE). It is well known that the Nash equilibrium is *often* not Pareto-optimal [21], since the best achievable performance characterized by Pareto boundary represents the set of optimal trade-offs among these conflicting/competing users' objectives. The trade-off of different users needs to be optimized by cooperative algorithms to achieve their joint outcome [22].

A direct improvement from NE to Nash bargaining (NB) by cooperation for the MIMO IC has been studied in [23],

where the case is studied in which the interference-plus-noise covariance matrix of each user approaches an identity matrix and the rate region becomes convex. A main branch of cooperative algorithms is the interference-pricing based method, where each user updates its own strategy to maximize its own utility minus the interference cost determined by the interference prices, which reflect the marginal change in utility per unit interference power. This distributed interference-pricing algorithm has been used to solve (weighted) sum-rate maximization problems for the multi-user SISO and MISO IC [24], multi-user single-stream MIMO IC [25], two-user MIMO IC [14]. A different pricing scheme is to balance the egoistic and altruistical strategies with different weights (i.e., prices), e.g., for the two-user MISO IC [4], [26] and for the multi-user single-stream MIMO IC [16]. In [16], the suboptimal maximum sum rate is achieved, although no convergence analysis is provided. However, most distributed cooperative algorithms for the MIMO IC (e.g., distributed pricing based algorithms [14], [25]) focus on maximizing (weighted) sum-utility rather than computing the whole utility region.

In fact, most approaches of parameterization, computation and weighted sum maximization are coordinated/cooperative algorithms, although they are not described in a game theoretic way.

For the MIMO IC, the achievable rate depends on both transmit and receive strategies involved in a more hard-coupling and complex expression than the MISO IC such that it is not straightforward to extend the *implicit* or *explicit* schemes achieving the complete Pareto boundary for the MISO IC to the MIMO IC. In order to illustrate the *complete Pareto boundary*, we formulate a computation problem to maximize one rate while keeping the rate of the other users unchanged such that one rate can increase always along the same direction in the rate region until the boundary is reached. However in this computation problem, the hard-coupled beamformers exist not only in the objective but also in the constraints, which makes this beamformers optimization problems non-convex (even NP-hard). Therefore in this paper, as most current references on the MIMO IC, we focus on finding high quality sub-optimal operating points efficiently. Our main contributions are described as follows.

i) First, the two-user single-stream MIMO IC is firstly studied: a) We propose an equivalent form of the SINR expression based on the Hermitian angle (Proposition 1 in Section II-B), which gains more insight into the coupling of the transmit beamformers; b) We prove that the strict Pareto-optimal transmit power allocation policy is full power allocation at both the transmitters (Proposition 2 in Section III-A); c) The non-strict Pareto boundary, two ending points of strict Pareto boundary, and certain ZF points are computed exactly. (Section III-B).

ii) To compute the strict Pareto boundary of the two-user single-stream MIMO IC, we formulate a problem to maximize one rate while the other rate is fixed. This non-convex optimization problem is solved by the proposed alternating optimization algorithm [27] such that a convergent point is guaranteed to be achieved (Section IV A-C).

iii) The proposed optimization algorithm can be extended to the multi-user scenario (Section IV-D).

Notation: $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^\dagger$, $\mathrm{Rank}(\cdot)$ and $\mathrm{Tr}(\cdot)$ denote complex conjugate, transpose, Hermitian and Moore-Penrose pseudo inverse, rank and trace, respectively. $|\cdot|$, $\Re(\cdot)$, and j denote the absolute value, the real part of a complex-valued number, and the imaginary unit, respectively. $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^H\boldsymbol{x}}$ and $\overrightarrow{\boldsymbol{x}} = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ denote the vector 2-norm and vector direction, respectively. $\perp$, $\parallel$ and $\nparallel$ denote perpendicularity, parallelity and unparallelity, respectively. $\mathcal{CN}(0, \boldsymbol{X})$ denotes a complex circularly-symmetric jointly-Gaussian probability density function with zero mean and covariance matrix $\boldsymbol{X}$. $\boldsymbol{X} \succeq \boldsymbol{0}$ or $\boldsymbol{X} \succ \boldsymbol{0}$ means $\boldsymbol{X}$ is a positive semidefinite matrix or a positive definite matrix. $\lambda_i(\boldsymbol{X})$ and $\boldsymbol{u}_i(\boldsymbol{X})$ denote the i-th *largest* eigenvalue of $\boldsymbol{X}$ and its corresponding eigenvector, respectively. $\lambda_i(\boldsymbol{X}, \boldsymbol{Y})$ and $\boldsymbol{u}_i(\boldsymbol{X}, \boldsymbol{Y})$ denote the i-th largest generalized eigenvalue of $\boldsymbol{X}, \boldsymbol{Y}$ and its corresponding eigenvector, respectively. $\Pi_{\boldsymbol{X}} \triangleq \boldsymbol{X}(\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H$ denotes the orthogonal projection onto the column space of $\boldsymbol{X}$, and $\Pi_{\boldsymbol{X}}^{\perp} \triangleq \boldsymbol{I} - \Pi_{\boldsymbol{X}}$ denotes the orthogonal projection onto the orthogonal complement of the column space of $\boldsymbol{X}$.

II. SYSTEM MODEL

A. Signal Model

Consider a two-user MIMO IC denoted by $\mathrm{TX}_i \mapsto \mathrm{RX}_i, i = 1, 2$, where each transmitter TX_i and each receiver RX_i are equipped with $N_T \geq 2$ and $N_R \geq 2$ antennas and a single data stream is transmitted in each user. In this IC, the received data at RX_i is modeled as

$$y_i = \boldsymbol{g}_i^H\left(\boldsymbol{H}_{ii}\boldsymbol{w}_i x_i + \boldsymbol{H}_{ki}\boldsymbol{w}_k x_k + \boldsymbol{n}_i\right), \; i, \; k \in \{1, 2\}, k \neq i$$

where $x_i \sim \mathcal{CN}(0, 1)$ is the transmitted symbol of TX_i by the transmit beamformer $\boldsymbol{w}_i \in \mathbb{C}^{N_T \times 1}$. At RX_i, $\boldsymbol{g}_i \in \mathbb{C}^{N_R \times 1}$ is the receive beamformer, and $\boldsymbol{n}_i \in \mathbb{C}^{N_R \times 1} \sim \mathcal{CN}(0, \sigma_i^2 \boldsymbol{I})$ is the additive white Gaussian noise (AWGN) vector. The matrices $\boldsymbol{H}_{ii}, \boldsymbol{H}_{ki} \in \mathbb{C}^{N_R \times N_T}$ denote the flat fading channel-matrix of the direct link $\mathrm{TX}_i \mapsto \mathrm{RX}_i$ and the cross-talk link $\mathrm{TX}_k \mapsto \mathrm{RX}_i$, respectively. Each transmitter has a power constraint that we, without loss of generality, set to 1 and define the set of feasible transmit beamformers as $\mathcal{W} \triangleq \left\{ \boldsymbol{w} \in \mathbb{C}^{N_T \times 1} : \|\boldsymbol{w}\|^2 \leq 1 \right\}$.

B. Rate with MMSE Receiver

Assume that the interference from the other transmitter is treated as additive Gaussian noise at each receiver. The achievable rate of the link $\mathrm{TX}_i \mapsto \mathrm{RX}_i$ is given by:

$$R_i(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{g}_i) = \log_2\left(1 + \mathrm{SINR}_i(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{g}_i)\right), \quad (1)$$

where $\mathrm{SINR}_i(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{g}_i) = \frac{|\boldsymbol{g}_i^H \boldsymbol{H}_{ii}\boldsymbol{w}_i|^2}{\sigma_i^2 + |\boldsymbol{g}_i^H \boldsymbol{H}_{ki}\boldsymbol{w}_k|^2}$. In the linear transceiver design, it is well known that the MMSE filter is the optimal receiver for given transmit strategies. In this paper, the MMSE filter $\boldsymbol{g}_i = \left(\sigma_i^2 \boldsymbol{I} + \boldsymbol{H}_{ii}\boldsymbol{w}_i\boldsymbol{w}_i^H\boldsymbol{H}_{ii}^H + \boldsymbol{H}_{ki}\boldsymbol{w}_k\boldsymbol{w}_k^H\boldsymbol{H}_{ki}^H\right)^{-1}\boldsymbol{H}_{ii}\boldsymbol{w}_i$ is employed at RX_i. Then,

$\mathrm{SINR}_i(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{g}_i)$ becomes

$$\mathrm{SINR}_i(\boldsymbol{w}_1, \boldsymbol{w}_2) = \boldsymbol{w}_i^H \underbrace{\boldsymbol{H}_{ii}^H\left(\sigma_i^2\boldsymbol{I} + \boldsymbol{H}_{ki}\boldsymbol{w}_k\boldsymbol{w}_k^H\boldsymbol{H}_{ki}^H\right)^{-1}\boldsymbol{H}_{ii}}_{\triangleq \boldsymbol{A}_i(\boldsymbol{w}_k)} \boldsymbol{w}_i. \quad (2)$$

The complex mathematical structure (inverse of the sum of matrices and product of matrices) causes a hard-coupling problem of $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ in the SINR in (2), which makes it difficult to analyze the SINR directly. To gain an insight into this coupling problem, we propose an equivalent form of the SINR expression.

Proposition 1 *For the two-user single-beam MIMO IC, the SINR in (2) can be reformulated as*

$$\mathrm{SINR}_i(\boldsymbol{w}_1, \boldsymbol{w}_2) = \sin^2(\theta_{H,i})\frac{\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2}{\sigma_i^2} + \cos^2(\theta_{H,i})\frac{\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2}{\sigma_i^2 + \|\boldsymbol{H}_{ki}\boldsymbol{w}_k\|^2}, \quad (3)$$

where $\cos(\theta_{H,i}) = \left|\overrightarrow{\boldsymbol{H}_{ii}\boldsymbol{w}_i}^H \cdot \overrightarrow{\boldsymbol{H}_{ki}\boldsymbol{w}_k}\right|$ *and* $\theta_{H,i} \in [0, \pi/2]$. $\square$

Proof: See Appendix A. ∎

Note that the $\mathrm{SINR}_i(\boldsymbol{w}_1, \boldsymbol{w}_2)$ can be considered as a combination of $\frac{\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2}{\sigma_i^2}$ and $\frac{\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2}{\sigma_i^2 + \|\boldsymbol{H}_{ki}\boldsymbol{w}_k\|^2}$ with the weights $\sin^2(\theta_{H,i})$ and $\cos^2(\theta_{H,i})$. That is, $\mathrm{SINR}_i(\boldsymbol{w}_1, \boldsymbol{w}_2)$ depends not only on the desired signal power $\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2$ and the interference power $\|\boldsymbol{H}_{ki}\boldsymbol{w}_k\|^2$, but also on the Hermitian angle $\theta_{H,i}$ between the directions $\overrightarrow{\boldsymbol{H}_{ii}\boldsymbol{w}_i}$ and $\overrightarrow{\boldsymbol{H}_{ki}\boldsymbol{w}_k}$. The SINR is coupled in a difficult way because of the existence of $\theta_{H,i}$. This is why it is more difficult to analyze the SINR of a MIMO IC than that of a MISO IC.

III. PARETO BOUNDARY AND COMPUTATION OF SOME KEY POINTS

A. Pareto Boundary

The achievable rate region is defined as a set of the achievable rate pairs with all the feasible beamformers

$$\mathcal{R} \triangleq \bigcup_{\boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathcal{W}} \left(R_1(\boldsymbol{w}_1, \boldsymbol{w}_2), R_2(\boldsymbol{w}_1, \boldsymbol{w}_2)\right).$$

Note that the achievable rate region $\mathcal{R}$ is not the capacity region. Its outermost boundary is called Pareto boundary in this paper, which can be denoted by a set $\mathcal{R}^\star \triangleq \bigcup(R_1^\star, R_2^\star)$ where $(R_1^\star, R_2^\star)$ is a Pareto-optimal point. More precisely, Pareto-optimality is defined as follows.

Definition 1 *A rate pair* $(R_1^\star, R_2^\star) \in \mathcal{R}$ *is [strict] Pareto-optimal iff there does not exist another rate pair* $(R_1, R_2) > (R_1^\star, R_2^\star)$ *[*$(R_1, R_2) \geq (R_1^\star, R_2^\star)$ *and* $(R_1, R_2) \neq (R_1^\star, R_2^\star)$*]* *with* $(R_1, R_2) \in \mathcal{R}$, *where the inequality is component-wise.* $\square$

As shown in Fig. 1, the Pareto boundary consists of the *strict Pareto boundary* (the upperright part graphically, denoted by "strict PB") and the *non-strict Pareto boundary* (including the vertical part and the horizontal part graphically, denoted by "non-strict PB"), divided by "E1" and "E2". "E1" and "E2", "SU1" and "SU2" mean two ending points of the strict Pareto

Fig. 1.   Achievable rate region and Pareto boundary in two-user case

boundary and two single user points, which will be studied in Section III-B. More precisely, for an arbitrary point on the strict Pareto boundary, it is impossible to improve one rate without simultaneously decreasing the other. For a point on the non-strict Pareto boundary, one rate can be further improved while the other rate remains the maximum rate. In particular, the strict Pareto boundary can be characterized as follows.

Proposition 2 *For the two-user single-stream MIMO IC, all the operating points on the strict Pareto boundary can be achieved only when both the transmitters spend the full power, i.e., $\|\boldsymbol{w}_1\|^2 = \|\boldsymbol{w}_2\|^2 = 1$.* □

Proof: See Appendix B.   ∎

Remark 1 *In fact, Proposition 2 has solved a strict Pareto-optimal transmit power allocation problem in this scenario. When the strict Pareto-optimal power allocation policy is employed, i.e., both the transmitters spend the full power, the two strict Pareto-optimal transmit beamformers design reduces to the optimization of two transmit beamforming patterns.* □

Here, we define a set of all the beamformers with full transmit power as $\mathcal{W}_{\mathcal{FP}} \triangleq \left\{ \boldsymbol{w} \in \mathbb{C}^{N_T \times 1} : \|\boldsymbol{w}\|^2 = 1 \right\}$ Note that all the strict Pareto-optimal transmit beamformers should be in the set $\mathcal{W}_{\mathcal{FP}}$.

B. Computation of Some Key Points

In this part, we compute exactly the non-strict Pareto boundary, two ending points of the strict Pareto boundary and certain ZF operating points.

1) Single-User Points $SU1(\overline{R}_1, 0)$ and $SU2(0, \overline{R}_2)$: A single-user point can be easily achieved when only one TX_i works and simultaneously operates "egoistically" to maximize its own rate. The maximum achievable rate $\overline{R}_i$ of the link $\mathrm{TX}_i \mapsto \mathrm{RX}_i$ and its associated "egoistic" strategy $\boldsymbol{w}_i^{Ego}$ are

$$\overline{R}_i = \log_2\left(1 + \frac{\lambda_1(\boldsymbol{H}_{ii}^H \boldsymbol{H}_{ii})}{\sigma_i^2}\right), \ \boldsymbol{w}_i^{Ego} = \boldsymbol{u}_1(\boldsymbol{H}_{ii}^H \boldsymbol{H}_{ii}) \ \forall i. \tag{4}$$

2) Ending Points of Strict Pareto Boundary $E1(\overline{R}_1, \underline{R}_2)$ and $E2(\underline{R}_1, \overline{R}_2)$: Each ending point of the strict Pareto boundary can be achieved when one transmitter employs an "altruistic" strategy to create no interference to the other receiver and simultaneously to maximize its own rate and the other transmitter operates "egoistically". For $E1(\overline{R}_1, \underline{R}_2)$, we easily find from (3) that $\theta_{H,1} = \pi/2$ results in no interference in the cross-talk link $\mathrm{TX}_2 \mapsto \mathrm{RX}_1$. How to find the "altruistic" strategy $\boldsymbol{w}_2^{Alt}$ is shown as follows.

Proposition 3 *$E1(\overline{R}_1, \underline{R}_2)$ can be achieved by $(\boldsymbol{w}_1^{Ego}, \boldsymbol{w}_2^{Alt})$, where $\overline{R}_1$ and $\boldsymbol{w}_1^{Ego}$ are in (4) and*

$$\underline{R}_2 = \log_2\left(1 + \boldsymbol{w}_2^{Alt,H} \boldsymbol{A}_2(\boldsymbol{w}_1^{Ego}) \boldsymbol{w}_2^{Alt}\right),$$

$$\boldsymbol{w}_2^{Alt} = \overrightarrow{\Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp} \boldsymbol{u}_1\left(\boldsymbol{B}_1, \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp}\right)}, \tag{5}$$

with $\boldsymbol{B}_1 \triangleq \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp} \boldsymbol{A}_2(\boldsymbol{w}_1^{Ego}) \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp}$. □

Proof: See Appendix C.   ∎

Similarly, $E2(\underline{R}_1, \overline{R}_2)$ with $(\boldsymbol{w}_1^{Alt}, \boldsymbol{w}_2^{Ego})$ can be easily obtained by interchanging the indices.

3) Non-strict Pareto-Optimal Points $(R_1^\star, R_2^\star)$: For the non-strict Pareto boundary, either the horizontal part or the vertical part starts and ends with a single user point and an ending point. Therefore, an arbitrary point $(R_1^\star, R_2^\star)$ on the non-strict Pareto boundary can be computed as

$$R_i^\star = \gamma \cdot \underline{R}_i \ \text{ and } \ R_k^\star = \overline{R}_k, \quad \forall i, \ k \in \{1, 2\}, \ k \neq i$$

where $i = 1$ and $i = 2$ correspond to the horizontal part and the vertical part, respectively. The scalar γ satisfies $\gamma \in [0, 1)$. The point $(R_1^\star, R_2^\star)$ becomes a single-user point or an ending point when $\gamma = 0$ or $\gamma = 1$, respectively. The associated non-strict Pareto-optimal transmit strategies are

$$\boldsymbol{w}_i^\star = \sqrt{\gamma} \cdot \boldsymbol{w}_i^{Alt} \ \text{ and } \ \boldsymbol{w}_k^\star = \boldsymbol{w}_k^{Ego},$$

from which we find that it is not necessary for both the transmitters to spend full power simultaneously to achieve the non-strict Pareto boundary. Thus, the non-strict Pareto-optimal power allocation policy is different from the strict Pareto-optimal power allocation policy (Proposition 2).

4) Zero-Forcing (ZF) Points $ZF(R_1^{ZF}, R_2^{ZF})$: ZF points are achieved when there is no interference between different users. Although these points are not on the Pareto boundary, it is still interesting to study ZF strategies if there exsits an additional requirement (like interference temperature or secrecy constraints) that each transmitter does not leak its own signal to other receivers.

In (3), we find that $\theta_{H,1} = \theta_{H,2} = \pi/2$ results in no interference in the cross-talk links $\mathrm{TX}_2 \mapsto \mathrm{RX}_1$ and $\mathrm{TX}_1 \mapsto \mathrm{RX}_2$ simultaneously. The ZF conditions are

$$\theta_{H,1} = \pi/2 \Leftrightarrow \overrightarrow{\boldsymbol{H}_{11}\boldsymbol{w}_1}^H \cdot \overrightarrow{\boldsymbol{H}_{21}\boldsymbol{w}_2} = 0$$
$$\Leftrightarrow \boldsymbol{H}_{11}\boldsymbol{w}_1 \perp \boldsymbol{H}_{21}\boldsymbol{w}_2 \Leftrightarrow \boldsymbol{w}_2 \perp \boldsymbol{H}_{21}^H \boldsymbol{H}_{11}\boldsymbol{w}_1,$$
$$\theta_{H,2} = \pi/2 \Leftrightarrow \overrightarrow{\boldsymbol{H}_{22}\boldsymbol{w}_2}^H \cdot \overrightarrow{\boldsymbol{H}_{12}\boldsymbol{w}_1} = 0$$
$$\Leftrightarrow \boldsymbol{H}_{22}\boldsymbol{w}_2 \perp \boldsymbol{H}_{12}\boldsymbol{w}_1 \Leftrightarrow \boldsymbol{w}_2 \perp \boldsymbol{H}_{22}^H \boldsymbol{H}_{12}\boldsymbol{w}_1, \tag{6}$$

from which and under a sufficient (and necessary only when $N_T = N_R = 2$) condition[2], i.e., $\boldsymbol{H}_{21}^H \boldsymbol{H}_{11}\boldsymbol{w}_1 \parallel \boldsymbol{H}_{22}^H \boldsymbol{H}_{12}\boldsymbol{w}_1$,

---

[2]This condition is the same as that in [15], while we derive it from a different perspective (Hermitian angle in (3)).

we obtain some of the ZF transmit strategies

$$\boldsymbol{w}_1^{ZF} = \boldsymbol{u}_\ell(\boldsymbol{H}_{22}^H \boldsymbol{H}_{12}, \boldsymbol{H}_{21}^H \boldsymbol{H}_{11}), \forall \ell \in \{1, 2, ..., N_T\},$$

$$\boldsymbol{w}_2^{ZF} = \sum_{\ell=1}^{N_T-1} c_\ell \boldsymbol{u}_\ell(\Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{ZF}}^\perp), \qquad (7)$$

where $\{c_\ell\}_{\ell=1}^{N_T-1}$ are complex-valued numbers and satisfy $\sum_{\ell=1}^{N_T-1} |c_\ell|^2 = 1$.

The $ZF(R_1^{ZF}, R_2^{ZF})$ can be achieved by $(\boldsymbol{w}_1^{ZF}, \boldsymbol{w}_2^{ZF})$ as

$$R_i^{ZF}(\boldsymbol{w}_i^{ZF}, \boldsymbol{w}_2^{ZF}) = \log_2 \left(1 + \frac{\|\boldsymbol{H}_{ii} \boldsymbol{w}_i^{ZF}\|^2}{\sigma_i^2}\right) \; \forall i. \quad (8)$$

IV. COMPUTATION OF THE STRICT PARETO BOUNDARY

Since the rate region of the two-user single-beam MIMO IC is always a normal region[3] according to Proposition 2, there exists only one intersection point between the line $R_i(\boldsymbol{w}_1, \boldsymbol{w}_2) = R_i^\star$ where $R_i^\star \in (\underline{R}_i, \overline{R}_i)$ and the strict Pareto boundary. Thus, an arbitrary point on the strict Pareto boundary can be uniquely determined when one rate is fixed and the other rate is maximized. This motivates us to propose the following optimization problem

$$(\text{P0}) \begin{cases} \max\limits_{\boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathcal{W}_{\mathcal{FP}}} & \text{SINR}_1(\boldsymbol{w}_1, \boldsymbol{w}_2) \\ \text{s.t.} & \text{SINR}_2(\boldsymbol{w}_1, \boldsymbol{w}_2) = \text{SINR}_2^\star. \end{cases}$$

where $\text{SINR}_2^\star \in (2^{\underline{R}_2} - 1, 2^{\overline{R}_2} - 1)$ is a SINR constraint, and $\boldsymbol{w}_1, \boldsymbol{w}_2$ should be in $\mathcal{W}_{\mathcal{FP}}$ according to Proposition 2. Then, $(R_1^\star, R_2^\star) = (\log_2(1 + \text{SINR}_1(\boldsymbol{w}_1^\star, \boldsymbol{w}_2^\star)), \log_2(1 + \text{SINR}_2^\star(\boldsymbol{w}_1^\star, \boldsymbol{w}_2^\star)))$ is achieved by the optimal solution $(\boldsymbol{w}_1^\star, \boldsymbol{w}_2^\star)$ to (P0).

For (P0), direct joint optimization of $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ is analytically intractable due to the hard-coupling problem of them in both the objective and the constraints. To solve (P0), an alternating optimization algorithm [27] is applied to optimize $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ alternatively by solving two single-beamformer optimization problems at each iteration. In the following, how to solve each single-beamformer problem is studied.

A. Optimization of $\boldsymbol{w}_1$

For a given *feasible* $\boldsymbol{w}_2$ (the feasibility of $\boldsymbol{w}_2$ will be studied in Proposition 4), the problem (P0) becomes a single-beamformer optimization problem w.r.t. $\boldsymbol{w}_1$. Its constraint is

$$\boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \left(\sigma_2^2 \boldsymbol{I} + \boldsymbol{H}_{12} \boldsymbol{w}_1 \boldsymbol{w}_1^H \boldsymbol{H}_{12}^H\right)^{-1} \boldsymbol{H}_{22} \boldsymbol{w}_2 = \text{SINR}_2^\star$$

$$\overset{(a)}{\Leftrightarrow} \boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \boldsymbol{H}_{22} \boldsymbol{w}_2 - \frac{|\boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \boldsymbol{H}_{12} \boldsymbol{w}_1|^2}{\sigma_2^2 + \|\boldsymbol{H}_{12} \boldsymbol{w}_1\|^2} = \sigma_2^2 \text{SINR}_2^\star,$$

$$\overset{(b)}{\Leftrightarrow} \frac{\boldsymbol{w}_1^H \boldsymbol{H}_{12}^H \boldsymbol{H}_{22} \boldsymbol{w}_2 \boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \boldsymbol{H}_{12} \boldsymbol{w}_1}{\boldsymbol{w}_1^H (\sigma_2^2 \boldsymbol{I} + \boldsymbol{H}_{12}^H \boldsymbol{H}_{12}) \boldsymbol{w}_1}$$
$$= \boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \boldsymbol{H}_{22} \boldsymbol{w}_2 - \sigma_2^2 \text{SINR}_2^\star, \quad (9a)$$

$$\overset{(c)}{\Leftrightarrow} \boldsymbol{w}_1^H \boldsymbol{C}(\boldsymbol{w}_2) \boldsymbol{w}_1 = 0 \text{ and } \boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \boldsymbol{H}_{22} \boldsymbol{w}_2 \geq \sigma_2^2 \text{SINR}_2^\star. \quad (9b)$$

The transformation (a) is based on the matrix inverse lemma. The transformation (b) is due to $\|\boldsymbol{w}_1\|^2 = 1$. In the transformation (c), the nonnegative left-hand side of (9a) demands $\boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \boldsymbol{H}_{22} \boldsymbol{w}_2 \geq \sigma_2^2 \text{SINR}_2^\star$, and $\boldsymbol{C}$ is a Hermitian matrix defined as

$$\boldsymbol{C}(\boldsymbol{w}_2) \triangleq \boldsymbol{H}_{12}^H \boldsymbol{H}_{22} \boldsymbol{w}_2 \boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \boldsymbol{H}_{12}$$
$$- (\boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \boldsymbol{H}_{22} \boldsymbol{w}_2 - \sigma_2^2 \text{SINR}_2^\star) \cdot (\sigma_2^2 \boldsymbol{I} + \boldsymbol{H}_{12}^H \boldsymbol{H}_{12}). \quad (10)$$

Then, $\boldsymbol{w}_1$ can be optimized by

$$(\text{P1}) \begin{cases} \max\limits_{\boldsymbol{w}_1 \in \mathcal{W}_{\mathcal{FP}}} & \boldsymbol{w}_1^H \boldsymbol{A}_1(\boldsymbol{w}_2) \boldsymbol{w}_1 \\ \text{s.t.} & \boldsymbol{w}_1^H \boldsymbol{C}(\boldsymbol{w}_2) \boldsymbol{w}_1 = 0 \end{cases}$$

where $\boldsymbol{C}(\boldsymbol{w}_2)$ and $\boldsymbol{A}_1(\boldsymbol{w}_2)$ are Hermitian matrices. Observe that the problem (P1) is a homogeneous quadratically constrained quadratic program (QCQP) and the objective function is convex but the convexity of constraints is unclear. Generally, it is difficult to solve this non-convex problem.

Note that $\boldsymbol{w}_1^H \boldsymbol{X} \boldsymbol{w}_1 = \text{Tr}(\boldsymbol{X} \boldsymbol{W}_1)$ for any matrix $\boldsymbol{X}$, where $\boldsymbol{W}_1 = \boldsymbol{w}_1 \boldsymbol{w}_1^H$ is a rank-one Hermitian positive semidefinite matrix. By the semidefinite programming and rank relaxation (SDR) method, (P1) can be transformed to

$$(\text{P2}) \begin{cases} \max\limits_{\boldsymbol{W}_1 \succeq \boldsymbol{0}} & \text{Tr}(\boldsymbol{A}_1(\boldsymbol{w}_2) \boldsymbol{W}_1) \\ \text{s.t.} & \text{Tr}(\boldsymbol{C}(\boldsymbol{w}_2) \boldsymbol{W}_1) = 0 \\ & \text{Tr}(\boldsymbol{W}_1) = 1. \end{cases}$$

Observe that the SDR (P2) is convex and solvable, i.e., its respective finite optimal solutions exist for a feasible $\boldsymbol{w}_2$, based on Weierstrass' Theorem. Its optimal solution $\boldsymbol{W}_1^\star$ is efficiently obtained by a convex optimization toolbox, e.g., SeDuMi [28] or CVX [29]. However, the rank of $\boldsymbol{W}_1^\star$ to (P2) is usually more than one because we have discarded the rank constraint $\text{Rank}(\boldsymbol{W}_1) = 1$. Therefore, we need to extract an optimal rank-one solution $\boldsymbol{w}_1$ to (P1) from $\boldsymbol{W}_1^\star$. If $\text{Rank}(\boldsymbol{W}_1^\star) = 1$, it is clear $\boldsymbol{w}_1 = \boldsymbol{u}_1(\boldsymbol{W}_1^\star)$. Otherwise, other tight matrix rank-one decomposition methods are needed. In [30], Ai et al. have proven a matrix rank-one decomposition theorem and used it to show that *the SDRs of a large class of complex-valued homogeneous QCQPs with not more than 4 constraints are in fact tight*[4]. Since the problem (P1) as a homogeneous QCQP with 2 constraints, an optimal $\boldsymbol{w}_1$ to the QCQP (P1) can be reconstructed from $\boldsymbol{W}_1^\star$ to the SDR (P2) based on the theorem and algorithm of the matrix rank-one decomposition in [30].

Remark 2 *In (P1), if $\boldsymbol{C}(\boldsymbol{w}_2)$ is a positive/negative semidefinite matrix without full rank, $\boldsymbol{w}_1$ should and must be in the null space of $\boldsymbol{C}(\boldsymbol{w}_2)$ to satisfy $\boldsymbol{w}_1^H \boldsymbol{C}(\boldsymbol{w}_2) \boldsymbol{w}_1 = 0$. According to the proof in Appendix C, $\boldsymbol{w}_1$ can be expressed by $\overrightarrow{\boldsymbol{U}_1 \boldsymbol{U}_1^H \boldsymbol{p}_1}$ where $\boldsymbol{U}_1 \in \mathbb{C}^{N_T \times (N_T - \text{Rank}(\boldsymbol{C}(\boldsymbol{w}_2)))}$ consists of $N_T - \text{Rank}(\boldsymbol{C}(\boldsymbol{w}_2))$ eigenvectors corresponding to zero*

---

[3] A set $\mathcal{G} \subseteq \mathbb{R}_n^+$ is called a normal region if for any two points $\boldsymbol{x} \in \mathcal{G}, \boldsymbol{x}' \in \mathbb{R}_n^+$ such that if $\boldsymbol{x}' \leq \boldsymbol{x}$, then $\boldsymbol{x}' \in \mathcal{G}$, too.

[4] Note that the application of Theorem 2.2 and Theorem 2.3 in [30] needs $N_T \geq 3$. It implies that $\text{TX}_k \; \forall k$ should have $N_T \geq 3$ antennas in our scenario.

eigenvalues of $C(w_2)$. Then, (P1) is equivalent to

$$\max_{p_1 \in \mathbb{C}^{N_T \times 1}} \frac{p_1^H U_1 U_1^H A_1(w_2) U_1 U_1^H p_1}{p_1^H U_1 U_1^H p_1}, \qquad (11)$$

from which it is easy to derive the optimal solution $p_1^{opt} = u_1\big(U_1 U_1^H A_1(w_2) U_1 U_1^H, U_1 U_1^H\big)$. Therefore, the optimal solution to (P1) is $w_1 = \overrightarrow{U_1 U_1^H p_1^{opt}}$. □

B. Optimization of w_2

For a given *feasible* w_1, (P0) becomes another single-beamformer optimization problem w.r.t. w_2. Maximization of its objective function is

$$\max_{w_2 \in \mathcal{W}_{\mathcal{FP}}} w_1^H H_{11}^H \left(\sigma_1^2 I + H_{21} w_2 w_2^H H_{21}^H \right)^{-1} H_{11} w_1$$

$$\iff \max_{w_2 \in \mathcal{W}_{\mathcal{FP}}} w_1^H H_{11}^H H_{11} w_1 - \frac{|w_1^H H_{11}^H H_{21} w_2|^2}{\sigma_1^2 + w_2^H H_{21}^H H_{21} w_2}$$

$$\iff \min_{w_2 \in \mathcal{W}_{\mathcal{FP}}} \frac{|w_1^H H_{11}^H H_{21} w_2|^2}{\sigma_1^2 + w_2^H H_{21}^H H_{21} w_2}$$

$$\iff \min_{w_2 \in \mathcal{W}_{\mathcal{FP}}} \frac{w_2^H C_1(w_1) w_2}{w_2^H C_2 w_2} \qquad (12)$$

where $C_1(w_1) \triangleq H_{21}^H H_{11} w_1 w_1^H H_{11}^H H_{21}$ and $C_2 \triangleq \sigma_1^2 I + H_{21}^H H_{21}$ are Hermitian matrices.

Then, w_2 can be optimized by

$$\text{(P3)} \begin{cases} \min\limits_{w_2 \in \mathcal{W}_{\mathcal{FP}}} & \dfrac{w_2^H C_1(w_1) w_2}{w_2^H C_2 w_2} \\ \text{s.t.} & w_2^H A_2(w_1) w_2 = \text{SINR}_2^\star. \end{cases}$$

Observe that (P3) is a fractional QCQP problem. The objective function is not even a quasi-convex function due to the convexity of both the nominator function and denominator function. To deal with this problem, we transform it by the SDR to

$$\text{(P4)} \begin{cases} \min\limits_{W_2 \succeq 0} & \dfrac{\text{Tr}\big(C_1(w_1) W_2\big)}{\text{Tr}\big(C_2 W_2\big)} \\ \text{s.t.} & \text{Tr}\big(A_2(w_1) W_2\big) = \text{SINR}_2^\star \\ & \text{Tr}\big(W_2\big) = 1, \end{cases}$$

which is still a non-convex problem. Fortunately, the fractional structure can be removed by a variation of the Charnes-Cooper variable transformation [31]. Define the transformed variable $Q = sW_2$ with $s = \frac{1}{\text{Tr}(C_2 W_2)}$. Then, (P4) becomes

$$\text{(P5)} \begin{cases} \min\limits_{Q,\, s} & \text{Tr}\big(C_1(w_1) Q\big) \\ \text{s.t.} & \text{Tr}\big(A_2(w_1) Q\big) = s \cdot \text{SINR}_2^\star \\ & \text{Tr}\big(C_2 Q\big) = 1,\ \text{Tr}\big(Q\big) = s \\ & Q \succeq 0,\ \dfrac{1}{\lambda_1(C_2)} \le s \le \dfrac{1}{\lambda_N(C_2)}. \end{cases}$$

which is a convex problem w.r.t. Q and s and solvable (see the Appendix D). By a convex optimization toolbox, we can obtain the optimal solution $(Q^\star, s^\star)$. Then, the optimal solution to (P4) can be easily obtained by $W_2^\star = \frac{Q^\star}{s^\star}$. Observe that (P3) is equivalent to a homogeneous QCQP with 3 constraints.

Therefore, by the matrix rank-one decomposition method, an optimal rank-one solution w_2 to (P3) can be extracted from $W_2^\star$ when $\text{Rank}(W_2^\star) > 1$.

Remark 3 *In (P3), if $D \triangleq A_2(w_1) - \text{SINR}_2^\star \cdot I$ is a positive/negative semidefinite matrix without full rank, w_1 should and must be in the null space of D to satisfy $w_2^H D w_2 = 0$. According to the proof in Appendix C, w_2 can be expressed by $\overrightarrow{U_2 U_2^H p_2}$ where $U_2 \in \mathbb{C}^{N_T \times (N_T - \text{Rank}(D))}$ consists of $N_T - \text{Rank}(D)$ eigenvectors corresponding to zero eigenvalues of D. Then, (P3) is equivalent to*

$$\max_{p_2 \in \mathbb{C}^{N_T \times 1}} \frac{p_2^H U_2 U_2^H C_1(w_1) U_2 U_2^H p_2}{p_2^H U_2 U_2^H C_2 U_2 U_2^H p_2}, \qquad (13)$$

from which it is easy to derive the optimal solution $p_2^{opt} = u_1\big(U_2 U_2^H C_1(w_1) U_2 U_2^H, U_2 U_2^H C_2 U_2 U_2^H\big)$. Therefore, the optimal solution to (P3) is $w_2 = \overrightarrow{U_2 U_2^H p_2^{opt}}$. □

C. Algorithm

In this section, algorithm discussions are given to gain some insights into the proposed alternating optimization algorithm.

1) Algorithm Description: A feasible initial w_2 for optimization of TX$_1$ can be obtained as follows.

Proposition 4 *For a given $\text{SINR}_2^\star \in (2^{\underline{R}_2} - 1, 2^{\overline{R}_2} - 1)$, (w_1, w_2) is always a feasible solution pair to (P0) if $w_1 \in \mathcal{W}_{\mathcal{FP}}$, $w_2 \in \mathcal{W}_{\mathcal{F}}$ with*

$$\mathcal{W}_{\mathcal{F}} \triangleq \Big\{ w_2 \in \mathcal{W}_{\mathcal{FP}} : w_2^H H_{22}^H H_{22} w_2 \ge \sigma_2^2 \text{SINR}_2^\star,$$
$$\lambda_1\big(C(w_2)\big) \cdot \lambda_{N_T}\big(C(w_2)\big) \le 0 \Big\}, \qquad (14)$$

where $C(w_2)$ is defined in (10). □

Proof: See Appendix E. ∎

That is, (P1) equivalent to (P0) with a fixed $w_2 \in \mathcal{W}_{\mathcal{F}}$ always has at least one feasible point w_1 in $\mathcal{W}_{\mathcal{FP}}$ (more analysis of initialization will be given in Section IV-C-2).

The proposed alternating optimization algorithm with any initial $w_2 \in \mathcal{W}_{\mathcal{F}}$ is described in pseudo-code as Algorithm 1:

Algorithm 1 Two-User Alternating Optimization Algorithm

Input: w_2^{Alt}, w_2^{Ego}, an arbitrary $R_2^\star \in \big(\underline{R}_2, \overline{R}_2\big)$, and $\text{SINR}_2^\star = 2^{R_2^\star} - 1$.

Output: A convergent point $(R_1^{(\ell)}, R_2^\star)$ by $(w_1^{(\ell)}, w_2^{(\ell)})$.

begin

    Initialization: Set a feasible $w_2^{(0)} \in \mathcal{W}_{\mathcal{F}}$, $\ell = 0$.

    while *some convergence criterion is not satisfied* **do**

        $\ell ++$.

        Given $w_2^{(\ell-1)}$, obtain an optimal W_1 to (P2).

        Extract an optimal $w_1^{(\ell)}$ from W_1 to (P1);

        Given $w_1^{(\ell)}$, obtain an optimal (Q, s) to (P5) and an optimal $W_2 = \frac{Q}{s}$ to (P4).

        Extract an optimal $w_2^{(\ell)}$ from W_2 to (P3).

        Compute $R_1^{(\ell)} = \log_2\big(1 + w_1^{(\ell),H} A_1(w_2^{(\ell)}) w_1^{(\ell)}\big)$.

---

2) Algorithm Analysis: In this section, we discuss the proposed algorithm in the following aspects: i) the convergence, ii) the quality of the solution, iii) the implementation, iv) the complexity.

i) Convergence Analysis: Based on the results in Section IV-A and Section IV-B, a global optimal solution to each single-beamformer optimization problem (P1) and (P3) can be obtained at each iteration. We will show that the sequence $\left\{ \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)}) \right\}_{\ell=1}^{\infty}$ by Algorithm 1 monotonically increases and converges, i.e., $\mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell+1)}, \boldsymbol{w}_2^{(\ell+1)}) \geq \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)}) \ \forall \ell$.

Denote the optimization of $\boldsymbol{w}_1$ and the optimization of $\boldsymbol{w}_2$ by the mapping functions $\boldsymbol{w}_1 = \Phi(\boldsymbol{w}_2)$ and $\boldsymbol{w}_2 = \Theta(\boldsymbol{w}_1)$, respectively. Then, the procedure of Algorithm 1 at the $\ell+1$th iteration is shown as

$$\mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)}) \xrightarrow{\boldsymbol{w}_1^{(\ell+1)} = \Phi\left(\boldsymbol{w}_2^{(\ell)}\right)} \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell+1)}, \boldsymbol{w}_2^{(\ell)})$$

$$\xrightarrow{\boldsymbol{w}_2^{(\ell+1)} = \Theta\left(\boldsymbol{w}_1^{(\ell+1)}\right)} \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell+1)}, \boldsymbol{w}_2^{(\ell+1)}),$$

In $\boldsymbol{w}_1^{(\ell+1)} = \Phi(\boldsymbol{w}_2^{(\ell)})$, since the feasible point set of $\boldsymbol{w}_1$ of (P1) includes $\boldsymbol{w}_1^{(\ell)}$ and additionally the global optimal solution to (P1) can be obtained by $\Phi(\cdot)$, it obviously implies $\mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell+1)}, \boldsymbol{w}_2^{(\ell)}) = \mathrm{SINR}_1(\Phi\left(\boldsymbol{w}_2^{(\ell)}\right), \boldsymbol{w}_2^{(\ell)}) \geq \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)})$. Similarly, it is also easily verified $\mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell+1)}, \boldsymbol{w}_2^{(\ell+1)}) = \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell+1)}, \Theta\left(\boldsymbol{w}_1^{(\ell+1)}\right)) \geq \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell+1)}, \boldsymbol{w}_2^{(\ell)})$ by $\boldsymbol{w}_2^{(\ell+1)} = \Theta(\boldsymbol{w}_1^{(\ell+1)})$ due to the optimality of $\Theta(\cdot)$. As a consequence, the sequence of $\left\{ \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)}) \right\}_{\ell=1}^{\infty}$ *monotonically* increases as the iteration number ℓ increases. In addition, since the sequence $\left\{ \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)}) \right\}_{\ell=1}^{\infty}$ is upper-bounded by the single-user SINR, i.e., $\frac{\lambda_1(\boldsymbol{H}_{11}^H \boldsymbol{H}_{11})}{\sigma_1^2}$ in (4), the convergence of the sequence $\left\{ \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)}) \right\}_{\ell=1}^{\infty}$, and thus the convergence of Algorithm 1 is guaranteed for any feasible initial point $\boldsymbol{w}_2^{(0)}$.

Since the *hard-coupled* two beamformers exist not only in the objective but also in the constraints in (P0), the conventional convergence analysis for the block coordinate descent algorithm [32] that requires that the constraints are *separable* among the variables is not applicable to our scenario. Therefore, it is unclear whether the proposed algorithm converges to a stationary point $\left(R_1(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)}), R_2(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)}) \right)$ where $\boldsymbol{w}_1^{(\ell)}$ and $\boldsymbol{w}_2^{(\ell)}$ satisfy the KKT conditions of the original problem (P0).

ii) Quality of Solutions: Due to the convergence of the proposed algorithm, the limit point of sequence of $\left\{ \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)}) \right\}_{\ell=0}^{\infty}$ for an arbitrary feasible initial $\boldsymbol{w}_2^{(0)}$ can be achieved by

$$\lim_{\ell \to \infty} \mathrm{SINR}_1(\boldsymbol{w}_1^{(\ell)}, \boldsymbol{w}_2^{(\ell)})$$
$$= \mathrm{SINR}_1\left(\Phi\left(\Theta\left(... \left(\Phi\left(\boldsymbol{w}_2^{(0)} \right) \right) \right) \right), \Theta\left(... \left(\boldsymbol{w}_2^{(0)} \right) \right) \right).$$

It implies that the performance of the alternating optimization algorithm depends on the initial beamformer $\boldsymbol{w}_2^{(0)}$. Denote

the global optimal solution to (P0) by $(\boldsymbol{w}_1^{\star}, \boldsymbol{w}_2^{\star})$. Take an extreme example, if $\boldsymbol{w}_2^{(0)} = \boldsymbol{w}_2^{\star}$, we can obtain the global optimum $\boldsymbol{w}_1^{\star} = \Phi\left(\boldsymbol{w}_2^{(0)} \right)$ directly due to the optimality of $\Phi(\cdot)$. Therefore, a good initial beamformer could lead to high performance. However generally, it is difficult to find a good initial point *efficiently* for such a complex multi-variable optimization problem. In order to improve the performance, a common way in references is to implement the alternating optimization algorithm with multiple random initializations and then select the one with the best performance. In this paper, we desire to design a scheme to generate a good initialization *efficiently*.

Inspired by the idea in [2], [16], we heuristically propose a transmit beamformer design scheme by balancing the "egoistic" and "altruistic" strategies as

$$\boldsymbol{w}_i(\xi_{i,1}, \xi_{i,2}) = \overrightarrow{\xi_{i,1} \boldsymbol{w}_i^{Ego} + \xi_{i,2} \boldsymbol{w}_i^{Alt}}, \quad i = 1, 2, \quad (15)$$

where $\xi_{i,1}$ and $\xi_{i,2}$ are complex-valued parameters satisfying $|\xi_{i,1}| + |\xi_{i,2}| = 1$. In fact, this tradeoff scheme is reasonable. For instance, it is necessary to be Pareto optimal for the two-user MISO IC [2], and its a similar form still provides a good performance in sum-rate maximization for the multi-user single-stream MIMO IC [16]. The following simulation results show that this characterization cannot exactly achieve the whole strict Pareto boundary for the two-user MIMO IC but still has a promising performance. In particular, the two ending point of strict Pareto boundary $E1$ and $E2$ can be achieved exactly by $(\boldsymbol{w}_1(1,0), \boldsymbol{w}_2(0,1))$ and $(\boldsymbol{w}_1(0,1), \boldsymbol{w}_2(1,0))$, respectively.

If $(\boldsymbol{w}_1^{(0)}, \boldsymbol{w}_2^{(0)})$ corresponds to the bound of (15) or of random beamforming, then the proposed algorithm must improve (or at least keep) the bound of (15) or of random beamforming. However, it is not efficient to find those beamforming pairs achieving the bound achieved by (15) or by the random beamforming. Therefore, there is no guarantee to say that the proposed algorithm with *only one initial beamformer* always achieves an *outer* boundary than the bound of (15) or of random beamforming, but its performance will increase with the number of initial beamformers (i.e., multiple initializations). Since the $2N_T$-dimensional real space of each complex $\boldsymbol{w}_i$ can be *approximately* reduced to 3-dimensional real space (i.e., $|\xi_{i,1}|$, $|\xi_{i,2}|$ and the difference of the phases of $\xi_{i,1}$ and $\xi_{i,2}$ in (15) without significant performance loss. Therefore, the proposed algorithm with the proposed initialization in (15) is *more efficient or likely* to achieve a good performance compared with a random initialization.

To further enhance the efficiency of initialization by (15), a real constant parameter (i.e., the proportion of the "egoistic" strategy) is employed to reduce (15) to

$$\boldsymbol{w}_i = \overrightarrow{\zeta \cdot \boldsymbol{w}_i^{Ego} + (1-\zeta) \cdot \boldsymbol{w}_i^{Alt}}, \quad i = 1, 2, \quad (16)$$

where $\zeta = \frac{R_2^{\star} - \underline{R}_2}{\overline{R}_2 - \underline{R}_2}$ is a constant for a given $R_2^{\star}$. If $\boldsymbol{w}_2 \notin \mathcal{W}_{\mathcal{F}}$, we reset $\zeta \in \frac{R_2^{\star} - \underline{R}_2}{\overline{R}_2 - \underline{R}_2} + [-\nu, \nu]$ with $0 < \nu \leq \min\left\{ \frac{R_2^{\star} - \underline{R}_2}{\overline{R}_2 - \underline{R}_2}, \frac{\overline{R}_2 - R_2^{\star}}{\overline{R}_2 - \underline{R}_2} \right\}$ until $\boldsymbol{w}_2 \in \mathcal{W}_{\mathcal{F}}$. If $\boldsymbol{w}_2$ is still infeasible, we choose a randomly generated $\boldsymbol{w}_2 \in \mathcal{W}_{\mathcal{F}}$ directly.

Note that the characterization in (16) *directly* corresponds to its own bound. Thus, our proposed algorithm with the initialization by (16) always outperforms the bound by (16), which can serve as a lower bound of the proposed algorithm.

iii) Implementation with Transmitter Cooperation: For the purpose of implementation without an authority[5], we assume that each transmitter knows perfect global CSI in a block-fading environment through training and feedback and is willing to cooperate with each other transmitter for information exchange via backhaul links. For the optimization of $\boldsymbol{w}_1$ at TX_1, TX_1 solving (P1) based on the updated $\boldsymbol{w}_2$ from TX_2. Similarly, TX_2 optimizes $\boldsymbol{w}_2$ by solving (P3) based on the updated $\boldsymbol{w}_1$ from TX_1. The algorithm can be terminated once TX_1 finds that convergence criterion is satisfied.

iv) Complexity Analysis: For the proposed algorithm shown in Algorithm 1, each iteration involves solving two convex semidefinite relaxation (SDR) problems (i.e., (P2) and (P5)) and two implementations of the matrix rank-one decomposition of $\boldsymbol{W}_i$. In [33], it is shown that the complexity of solving the SDR is *polynomial* in the problem size (i.e., N_T) and the number of constraints (denoted by m), i,e., $\mathcal{O}(\max{(m, N_T)}^4 N_T^{1/2} \log(1/\epsilon))$ given a solution accuracy $\epsilon > 0$. In this paper, we have $N_T \geq 3$, and $m = 2$ in (P2) and $m = 3$ in (P5). Thus, the complexity of solving (P2) and (P5) is $\mathcal{O}(N_T^{4.5} \log(1/\epsilon))$. In terms of the complexity of the matrix rank-one decomposition [34], the rank-one solution can be extracted in *polynomial-time* if $\text{Rank}(\boldsymbol{W}_i) \geq 3$; If $\text{Rank}(\boldsymbol{W}_i) \geq 2$, it is sufficient to seek for a rank-one solution to a sequence of linear matrix equations within a slightly expanded range space of $\boldsymbol{W}_i$. If $\text{Rank}(\boldsymbol{W}_i) = 1$, only eigen-decomposition of $\boldsymbol{W}_i$ is needed.

In the following simulations, the average time of an iteration of Algorithm 1 is 0.6180 seconds by running the MATLAB 7.10 on the computer with AMD Athlon(TM) 64 Processor 3200+, 2.01 GHZ and 2GB RAM. Additionally, the fast convergent behavior of the proposed algorithm is implied numerically (e.g., Fig. 2(b) with 8.55 iterations on average and Fig. 3(b) with 5.16 iterations on average). Therefore, the proposed algorithm has reasonable complexity.

D. Extension to the multi-user MIMO IC

Consider the K-user single-stream MIMO IC. With the MMSE receiver, the achievable rate of the link $TX_k \mapsto RX_k$ is expressed as $R_k(\{\boldsymbol{w}_k\}_\mathcal{K}) = \log_2\left(1 + \text{SINR}_k(\{\boldsymbol{w}_k\}_\mathcal{K})\right) \forall k \in \mathcal{K} = \{1, ..., K\}$, where

$$\text{SINR}_k(\{\boldsymbol{w}_i\}_\mathcal{K}) = \boldsymbol{w}_k^H \boldsymbol{H}_{kk}^H \underbrace{(\sum_{i \neq k} \boldsymbol{H}_{ik} \boldsymbol{w}_i \boldsymbol{w}_i^H \boldsymbol{H}_{ik}^H + \sigma_k^2 \boldsymbol{I})^{-1}}_{\boldsymbol{A}_k(\boldsymbol{w}_{-k})} \boldsymbol{H}_{kk} \boldsymbol{w}_k. \quad (17)$$

is the SINR expression of the kth user. $\boldsymbol{w}_{-k}$ denotes $\{\boldsymbol{w}_i\}_{\mathcal{K}\setminus\{k\}}$.

Without loss of generality, the optimization problem (P0) in the two-user case can be generalized to[6]

$$(Q0) \begin{cases} \max\limits_{\{\boldsymbol{w}_i\}_\mathcal{K}} & \text{SINR}_1(\{\boldsymbol{w}_i\}_\mathcal{K}) \\ \text{s.t.} & \text{SINR}_k(\{\boldsymbol{w}_i\}_\mathcal{K}) = \text{SINR}_k^\star, \quad \forall k \in \mathcal{K}\setminus\{1\}. \\ & \boldsymbol{w}_i^H \boldsymbol{w}_i \leq 1, \quad \forall i \in \mathcal{K}, \end{cases}$$

where $\{\text{SINR}_i^\star\}_{\mathcal{K}\setminus\{1\}}$ are assumed to be feasible. The problem (Q0) is a non-convex problem of $\{\boldsymbol{w}_i\}_\mathcal{K}$.

To extend the proposed algorithm to (Q0), we derive equivalent expressions of $\text{SINR}_k(\{\boldsymbol{w}_i\}_\mathcal{K})$ in (17) by defining $\boldsymbol{D}(\boldsymbol{w}_{-i-k}) = \sum_{\ell \neq k,i} \boldsymbol{H}_{\ell k} \boldsymbol{w}_\ell \boldsymbol{w}_\ell^H \boldsymbol{H}_{\ell k}^H + \sigma_k^2 \boldsymbol{I}$:

$$\text{SINR}_k(\{\boldsymbol{w}_i\}_\mathcal{K})$$
$$= \boldsymbol{w}_k^H \boldsymbol{H}_{kk}^H \left(\boldsymbol{D}(\boldsymbol{w}_{-i-k}) + \boldsymbol{H}_{ik} \boldsymbol{w}_i \boldsymbol{w}_i^H \boldsymbol{H}_{ik}^H\right)^{-1} \boldsymbol{H}_{kk} \boldsymbol{w}_k$$
$$\overset{(a)}{=} \underbrace{\boldsymbol{w}_k^H \boldsymbol{H}_{kk}^H \boldsymbol{D}(\boldsymbol{w}_{-k-i})^{-1} \boldsymbol{H}_{kk} \boldsymbol{w}_k}_{\boldsymbol{F}_k(\boldsymbol{w}_{-i})} -$$
$$\frac{\boldsymbol{w}_i^H \overbrace{\boldsymbol{H}_{ik}^H \boldsymbol{D}(\boldsymbol{w}_{-k-i})^{-1} \boldsymbol{H}_{kk} \boldsymbol{w}_k \boldsymbol{w}_k^H \boldsymbol{H}_{kk}^H \boldsymbol{D}(\boldsymbol{w}_{-i-k})^{-1} \boldsymbol{H}_{ik}} \boldsymbol{w}_i}{1 + \boldsymbol{w}_i^H \underbrace{\boldsymbol{H}_{ik}^H \boldsymbol{D}(\boldsymbol{w}_{-i-k})^{-1} \boldsymbol{H}_{ik}}_{\boldsymbol{G}_k(\boldsymbol{w}_{-i-k})} \boldsymbol{w}_i}$$
$$(18)$$

where the transformation (a) is based on the Sherman-Morrison Formula [35] and $\boldsymbol{w}_{-i-k} = \{\boldsymbol{w}_\ell\}_{\mathcal{K}\setminus\{i,k\}}$.

Therefore, maximization of the objective function in (Q0) w.r.t. different beamformers is equivalent to

$$\max_{\boldsymbol{w}_1} \text{SINR}_1(\{\boldsymbol{w}_i\}_\mathcal{K}) \iff \max_{\boldsymbol{w}_1} \boldsymbol{w}_1^H \boldsymbol{A}_1(\boldsymbol{w}_{-1}) \boldsymbol{w}_1$$
$$\max_{\boldsymbol{w}_k} \text{SINR}_1(\{\boldsymbol{w}_i\}_\mathcal{K}) \overset{(b)}{\iff} \min_{\boldsymbol{w}_k} \frac{\boldsymbol{w}_k^H \boldsymbol{F}_k(\boldsymbol{w}_{-k}) \boldsymbol{w}_k}{1 + \boldsymbol{w}_k^H \boldsymbol{G}_k(\boldsymbol{w}_{-1-k}) \boldsymbol{w}_k},$$
$$\forall k \in \mathcal{K}\setminus\{1\}, \quad (19)$$

and individual SINR constraint is equivalent to

$$\text{SINR}_k(\{\boldsymbol{w}_i\}_\mathcal{K}) = \text{SINR}_k^\star$$
$$\iff \boldsymbol{w}_k^H \boldsymbol{A}_k(\boldsymbol{w}_{-k}) \boldsymbol{w}_k = \text{SINR}_k^\star,$$
$$\overset{(b)}{\iff} \boldsymbol{w}_i^H \boldsymbol{E}_k(\boldsymbol{w}_{-i}) \boldsymbol{w}_i = \gamma_k(\boldsymbol{w}_{-i}), \quad \forall i \neq k \quad (20)$$

where $\boldsymbol{E}_k(\boldsymbol{w}_{-i})$ and $\gamma_k(\boldsymbol{w}_{-i})$ are defined as

$$\boldsymbol{E}_k(\boldsymbol{w}_{-i}) \triangleq \boldsymbol{F}_k(\boldsymbol{w}_{-i}) - \left(\boldsymbol{w}_k^H \boldsymbol{H}_{kk}^H \boldsymbol{D}(\boldsymbol{w}_{-k-i})^{-1} \boldsymbol{H}_{kk} \boldsymbol{w}_k - \text{SINR}_k^\star\right) \boldsymbol{G}_k(\boldsymbol{w}_{-i-k});$$
$$\gamma_k(\boldsymbol{w}_{-i}) \triangleq \boldsymbol{w}_k^H \boldsymbol{H}_{kk}^H \boldsymbol{D}(\boldsymbol{w}_{-k-i})^{-1} \boldsymbol{H}_{kk} \boldsymbol{w}_k - \text{SINR}_k^\star.$$

The equivalence (b) in both (19) and (20) is based on (18).

1) Optimization of $\boldsymbol{w}_1$: Given the fixed $\boldsymbol{w}_{-1}$ and based on the equivalence results in (19) and (20), (Q0) w.r.t. $\boldsymbol{w}_1$ is equivalent to

$$(Q1) \begin{cases} \max\limits_{\boldsymbol{w}_1} & \boldsymbol{w}_1^H \boldsymbol{A}_1(\boldsymbol{w}_{-1}) \boldsymbol{w}_1 \\ \text{s.t.} & \boldsymbol{w}_1^H \boldsymbol{E}_k(\boldsymbol{w}_{-1}) \boldsymbol{w}_1 = \gamma_k(\boldsymbol{w}_{-1}), \; \forall k \in \mathcal{K}\setminus\{1\}. \\ & \boldsymbol{w}_1^H \boldsymbol{w}_1 \leq 1. \end{cases}$$

Observe that (Q1) is a homogeneous QCQP. By the SDR, (Q1) is relaxed to

$$(\text{Q1}')\begin{cases} \max_{\boldsymbol{W}_1 \succeq \boldsymbol{0}} & \text{Tr}(\boldsymbol{A}_1(\boldsymbol{w}_{-1})\boldsymbol{W}_1) \\ \text{s.t.} & \text{Tr}(\boldsymbol{E}_k(\boldsymbol{w}_{-1})\boldsymbol{W}_1) = \gamma_k(\boldsymbol{w}_{-1}), \ \forall k \in \mathcal{K}\backslash\{1\} \\ & \text{Tr}(\boldsymbol{W}_1) \leq 1 \end{cases}$$

where $\boldsymbol{W}_1 = \boldsymbol{w}_1\boldsymbol{w}_1^H$. Now, (Q1$'$) becomes a convex problem w.r.t. $\boldsymbol{W}_1$. The optimal $\boldsymbol{W}_1^\star$ to (Q1$'$) can be efficiently solved by a convex optimization toolbox.

If $\text{Rank}(\boldsymbol{W}_1^\star) = 1$, the optimal rank-one solution is $\boldsymbol{w}_1 = \boldsymbol{u}_1(\boldsymbol{W}_1^\star)$. Otherwise, we observe that (Q1) as a homogeneous QCQP has K constraints, and thus an optimal $\boldsymbol{w}_1$ to (Q1) can be reconstructed from $\boldsymbol{W}_1^\star$ for $K \leq 4$ by the matrix rank-one decomposition method in [30]. When $K \geq 5$, there exist several approaches (e.g., the eigenvector approximation method and the randomization method) to extract an *approximate* $\boldsymbol{w}_1$ from $\boldsymbol{W}_1^\star$. Although these approximation methods are not tight, intensive research show that they provide promising performance (the interested readers could refer to the analysis of approximation accuracy bounds in [33]).

2) Optimization of $\boldsymbol{w}_k$, $\forall k \neq 1$: Given the fixed $\boldsymbol{w}_{-k}$ and based on the equivalence results in (19) and (20), (Q0) is equivalent to

$$(\text{Qk})\begin{cases} \min_{\boldsymbol{w}_k} & \dfrac{\boldsymbol{w}_k^H \boldsymbol{F}_k(\boldsymbol{w}_{-k})\boldsymbol{w}_k}{1 + \boldsymbol{w}_k^H \boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{w}_k} \\ \text{s.t.} & \boldsymbol{w}_k^H \boldsymbol{A}_k(\boldsymbol{w}_{-k})\boldsymbol{w}_k = \text{SINR}_k^\star \\ & \boldsymbol{w}_k^H \boldsymbol{E}_\ell(\boldsymbol{w}_{-k})\boldsymbol{w}_k = \gamma_\ell(\boldsymbol{w}_{-k}), \ \forall \ell \in \mathcal{K}\backslash\{1,k\} \\ & \boldsymbol{w}_k^H \boldsymbol{w}_k \leq 1. \end{cases}$$

Observe that the objective function belongs to fractional program, while it is not a quasi-convex function due to the convexity of both the nominator function and the denominator function. To deal with this problem, we transform the problem (Qk) via the SDR to

$$(\text{Qk}')\begin{cases} \min_{\boldsymbol{W}_k \succeq \boldsymbol{0}} & \dfrac{\text{Tr}(\boldsymbol{F}_k(\boldsymbol{w}_{-k})\boldsymbol{W}_k)}{1 + \text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k)} \\ \text{s.t.} & \text{Tr}(\boldsymbol{A}_k(\boldsymbol{w}_{-k})\boldsymbol{W}_k) = \text{SINR}_k^\star \\ & \text{Tr}(\boldsymbol{E}_\ell(\boldsymbol{w}_{-k})\boldsymbol{W}_k) = \gamma_\ell(\boldsymbol{w}_{-k}), \forall \ell \in \mathcal{K}\backslash\{1,k\} \\ & \text{Tr}(\boldsymbol{W}_k) \leq 1. \end{cases}$$

where $\boldsymbol{W}_k = \boldsymbol{w}_k\boldsymbol{w}_k^H$. It is known that full power transmission is not always Pareto-optimal for the general multi-user MIMO/MISO IC (related to the number of users and transmit/receive antennas), which is different from the two-user Pareto-optimal full power transmission (Proposition 2). It leads to $\frac{\text{Tr}(\boldsymbol{F}_k(\boldsymbol{w}_{-k})\boldsymbol{W}_k)}{1+\text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k)} \neq \frac{\text{Tr}(\boldsymbol{F}_k(\boldsymbol{w}_{-k})\boldsymbol{W}_k)}{\text{Tr}((\boldsymbol{I}+\boldsymbol{G}_k(\boldsymbol{w}_{-1-k}))\boldsymbol{W}_k)}$. Thus, the Charnes-Cooper variable transformation used in the optimization problem (P4) in the two-user case is not applicable to (Qk$'$) any longer. Nevertheless, we observe that both the nominator function and the denominator function of the objective function are non-negative, differentiable and affine with $\boldsymbol{W}_k$. By introducing a real scalar parameter $\mu_k \geq 0$, the fractional programming problem (Qk$'$) is equivalent to a

parametric programming problem

$$\mathcal{F}(\mu_k) = \min_{\boldsymbol{W}_k \in \mathcal{S}_{\boldsymbol{W}_k}} \left\{ \text{Tr}(\boldsymbol{F}_k(\boldsymbol{w}_{-k})\boldsymbol{W}_k) \right. \\ \left. - \mu_k\left(1 + \text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k)\right) \right\}, \quad (21)$$

where $\mathcal{S}_{\boldsymbol{W}_k}$ denotes the constraint set of $\boldsymbol{W}_k$ consisting of all the constraints in (Qk$'$), and it is obvious that $\mathcal{S}_{\boldsymbol{W}_k}$ is a convex set. Assume the optimal solution to (Qk$'$) is $\boldsymbol{W}_k^\star$. If $\mu_k^\star = \frac{\text{Tr}(\boldsymbol{F}_k(\boldsymbol{w}_{-k})\boldsymbol{W}_k^\star)}{1+\text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k^\star)}$, it implies $\mathcal{F}(\mu_k^\star) = 0$. Thus, solving (Qk$'$) is equivalent to finding the root of the equation $\mathcal{F}(\mu_k) = 0$.

Given μ_k, (21) is a convex optimization problem w.r.t. $\boldsymbol{W}_k$, and its optimal solution $\boldsymbol{W}_k^\star(\mu_k)$ can be efficiently solved. Therefore, $\mathcal{F}(\mu_k) = 0$ can be further formulated as

$$\mathcal{F}(\mu_k^\star) = \text{Tr}(\boldsymbol{F}_k(\boldsymbol{w}_{-k})\boldsymbol{W}_k^\star) \\ - \mu_k^\star \cdot (1 + \text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k^\star)) = 0, \quad (22)$$

From [36], we know that $\mathcal{F}(\mu_k)$ is continuous, concave, strictly decreasing in μ_k and $\mathcal{F}(\mu_k) = 0$ has a unique solution. Additionally, we find that $-(1 + \text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k^\star))$ is a subgradient of $\mathcal{F}(\mu_k)$ for any μ_k. Thus, (22) can be solved by a generalized Newton method (also known as the Dinkelbach algorithm) described in Algorithm 2.

Algorithm 2 The generalized Newton method to solve (21)

Input: $\mu_k^{(0)}$ satisfying $\mathcal{F}(\mu_k^{(0)}) \leq 0$, tolerance ϵ.
Output: Optimal $\mu_k^\star$ and $\boldsymbol{W}_k^\star$.
begin
  $\ell = 0$
  while $|\mathcal{F}(\mu_k^{(\ell)})| > \epsilon$ **do**
    Given $\mu_k^{(\ell)}$, solve optimal $\boldsymbol{W}_k^\star(\mu_k^{(\ell)})$ to (21);
    $\mu_k^{(\ell+1)} = \frac{\text{Tr}(\boldsymbol{F}_k(\boldsymbol{w}_{-k})\boldsymbol{W}_k^\star(\mu_k^{(\ell)}))}{1+\text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k^\star(\mu_k^{(\ell)}))}$ [7];
    $\ell++$.
  $\mu_k^\star = \mu_k^{(\ell)}$ and $\boldsymbol{W}_k^\star = \boldsymbol{W}_k^\star(\mu_k^\star)$.

The algorithm as a Newton procedure to determine the root of the equation $\mathcal{F}(\mu_k^\star) = 0$ has superlinear convergence. By the Algorithm 2, the optimal $\mu_k^\star$ and $\boldsymbol{W}_k^\star(\mu_k^\star)$ to (21) is obtained. Equivalently, $\boldsymbol{W}_k^\star(\mu_k^\star)$ is an optimal solution to (Qk$'$) [36]. Then, a tight (for $K \leq 4$) or an approximate (for $K \geq 5$) solution $\boldsymbol{w}_k$ to (Qk) can be extracted from $\boldsymbol{W}_k^\star$.

Above all, the proposed alternating optimization algorithm extended to solve (Q0) can be described as Algorithm 3.

Remark 4 *The proposed alternating optimization algorithm can be extended to the K-user MIMO IC. For $K \leq 4$, it*

---

[7]This generalized Newton iterative update is from $\mu_k^{(\ell+1)} \triangleq \mu_k^{(\ell)} - \frac{\mathcal{F}(\mu_k^{(\ell)})}{-\left(1+\text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k^\star(\mu_k^{(\ell)}))\right)} = \mu_k^{(\ell)} - \frac{\text{Tr}(\boldsymbol{F}_k(\boldsymbol{w}_{-k})\boldsymbol{W}_k^\star(\mu_k^{(\ell)}))-\mu_k^{(\ell)}\left(1+\text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k^\star(\mu_k^{(\ell)}))\right)}{-\left(1+\text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k^\star(\mu_k^{(\ell)}))\right)} = \frac{\text{Tr}(\boldsymbol{F}_k(\boldsymbol{w}_{-k})\boldsymbol{W}_k^\star(\mu_k^{(\ell)}))}{1+\text{Tr}(\boldsymbol{G}_k(\boldsymbol{w}_{-1-k})\boldsymbol{W}_k^\star(\mu_k^{(\ell)}))}$.
[8]This condition is to make sure that a better (at least the same) solution to (Qk) (only for $K \geq 5$) is always obtained in each iteration such that the objective function's non-decreasing convergence is guaranteed.

Algorithm 3 K-User Alternating Optimization Algorithm

Input: $\{R_k^\star\}_{\mathcal{K}\setminus\{1\}}$ where $R_k^\star = \log_2\left(1 + \text{SINR}_k^\star\right)$ and $\text{SINR}_k^\star \in \left(0, \frac{1}{\sigma_k^2}\lambda_1(\boldsymbol{H}_{kk}\boldsymbol{H}_{kk}^H)\right]$.

Output: A convergent point $(R_1^{(\ell)}, R_2^\star, ..., R_K^\star)$ with $\{\boldsymbol{w}_i^{(\ell)}\}_{\mathcal{K}}$.

begin

Initialization: Set a feasible $\boldsymbol{w}_{-1}^{(0)}$, $\ell = 0$.

while *some convergence criterion is not satisfied* **do**

$\ell + +$.

for $k = 1 \to K$ **do**

Given $\boldsymbol{w}_{-k}^{(\ell-1)}$, obtain an optimal $\boldsymbol{W}_k$ to (Qk');
Extract a tight/approximate $\boldsymbol{w}_k$ from $\boldsymbol{W}_k$ to (Qk).

if $K \geq 5$ and $\text{SINR}_1\left(\boldsymbol{w}_k^{(\ell)}, \boldsymbol{w}_{-k}^{(\ell-1)/(\ell)}\right) < \text{SINR}_1\left(\boldsymbol{w}_k^{(\ell-1)}, \boldsymbol{w}_{-k}^{(\ell-1)/(\ell)}\right)^8$ **then**

$\boldsymbol{w}_k^{(\ell)} = \boldsymbol{w}_k^{(\ell-1)}$;

Compute $R_1^{(\ell)} = \log_2\left(1 + \text{SINR}_1\left(\{\boldsymbol{w}_k^{(\ell)}\}_{\mathcal{K}}\right)\right)$.

is the same as the two-user case that each optimal single-beamformer can be obtained in each iteration. For $K \geq 5$, each approximate optimal single-beamformer is obtained in each iteration. Following the same line of the proof of the Algorithm 1's convergence in Section IV-C-2, the convergence of the Algorithm 3 is also guaranteed. □

V. ILLUSTRATIONS AND DISCUSSIONS

To illustrate the achievable rate region by the proposed algorithm, we consider a two-user Gaussian MIMO IC, where $N_T = 3$ and $N_R = 2$. The transmit power budget is set to 1 for the two users, and noise power $\sigma_1^2 = \sigma_2^2 = 10^{-\frac{\text{SNR}}{10}}$ where SNR=10dB. The channels $\boldsymbol{H}_{11}, \boldsymbol{H}_{12}, \boldsymbol{H}_{21}$ and $\boldsymbol{H}_{22}$ are

$$\begin{pmatrix} -0.3034+1.9096i & -0.3790+0.4201i & 0.0357+0.7337i \\ -0.6358-0.8030i & -0.7881-0.1273i & 0.7534+0.8348i \end{pmatrix},$$

$$\begin{pmatrix} -0.6758+0.1040i & -0.5949-0.0344i & 0.4311+0.9658i \\ -2.1621+0.5404i & -0.0037+0.6627i & 0.8611+1.2318i \end{pmatrix},$$

$$\begin{pmatrix} 0.3999+0.1567i & 0.3798-0.5619i & -0.1005+0.2836i \\ -0.5494-0.4648i & 1.1971-0.5297i & -0.7271+0.2114i \end{pmatrix},$$

$$\begin{pmatrix} -0.0308-0.1133i & 0.0433-0.3313i & 0.3047-1.2157i \\ -1.4947-1.8676i & -0.9430+0.5704i & -1.3328+1.4638i \end{pmatrix}.$$

A. Convergence and Performance of Initialization by Eq.(16)

To study the convergence rate of the alternating optimization algorithm and evaluate the effectiveness of the initialization in (16), we respectively use (16) and 200 randomly generated feasible normalized vectors as initial $\boldsymbol{w}_2$. Then, we run Algorithm 1 until $|R_1^{(\ell)} - R_1^{(\ell-1)}| \leq 10^{-3}$.

Fig. 2 and Fig. 3 show the performance of the alternating optimization algorithm by different initializations. More precisely, Fig. 2 is for a given $R_2^\star = \underline{R}_2 + \frac{2}{19}\cdot(\overline{R}_2-\underline{R}_2) = 5.6398$ (close to the ending point $(\overline{R}_1, \underline{R}_2)$). Fig. 2(a) shows that the achieved R_1 with initialization by (16) nearly always outperforms that with 200 random initializations. Fig. 2(b) implies the convergence rate of the algorithm. Fig. 3 is for a

(a)

(b)

Fig. 2. Performance of Algorithm 1 by different initializations at $R_2^\star = 5.6398$: (a) Performance comparison of 200 random initializations and 1 initialization by Eq. (16); (b) Convergence behavior of Algorithm 1.

given $R_2^\star = \underline{R}_2 + \frac{11}{19}\cdot(\overline{R}_2-\underline{R}_2) = 6.2898$ (corresponding to the middle of strict Pareto boundary). Similarly, Fig. 3(a) and Fig. 3(b) also show that initialization by (16) has a promising performance and fast convergence behavior. Therefore, simulation results imply that (16) is a good choice for initialization.

B. Performance Comparison

Fig. 4 illustrates the achievable boundary by the proposed algorithm, i.e., Algorithm 1, compared with the existing methods. The term "Proposed_1+9" denotes the best result obtained by running Algorithm 1 with 1 initialization in (16) and 9 random initializations, while "Proposed_1" represents the result only with 1 initialization in (16). The SINR targets are $\text{SINR}_2^\star = 2^{R_2^\star-1}$s where $R_2^\star = \underline{R}_2 + \frac{n}{50}\cdot(\overline{R}_2-\underline{R}_2)$, $n = 1, 2, ..., 49$. Similarly, the term "WMMSE_10" denotes the weighted sum rate maximization algorithm [13] with 10 random initializations, where weighted sum rate is expressed as $w \cdot R_1 + (1-w) \cdot R_2$. with the weights ws in $[0.05 : 0.05 : 0.95]$. "WMMSE_1" denotes the result only with 1 initialization by (16), i.e, setting the initial beamformers as $\boldsymbol{w}_1^{(0)} = \overrightarrow{w \cdot \boldsymbol{w}_1^{Ego} + (1-w) \cdot \boldsymbol{w}_1^{Alt}}$ and $\boldsymbol{w}_2^{(0)} = \overrightarrow{w \cdot \boldsymbol{w}_2^{Ego} + (1-w) \cdot \boldsymbol{w}_2^{Alt}}$. The curve denoted by "RandBeam_10mil" means the outermost boundary of the rate region achieved by 10 million random normalized transmit beamformer pairs (each receiver is the MMSE filter).

(a)

(b)

Fig. 3.   Performance of Algorithm 1 by different initializations at $R_2^\star = 6.2898$: (a) Performance comparison of 200 random initializations and 1 initialization by Eq. (16) ; (b) Convergence behavior of Algorithm 1.

Theoretically, if *exhaustive* random beamforming pairs are chosen, its bound is exactly the Pareto boundary. However, there exist infinite random beamforming pairs so that we choose as many as 100 million random beamforming pairs in the simulations to serve as an *approximate Pareto boundary*. The term "SimpleReceiver" is the outermost boundary of the region achieved by [7] where each receiver is fixed as the largest left singular vector of the corresponding direct channel matrix. The curve denoted by "Eq.(15)" is the outermost boundary of the achieved region by (15) with complex-valued parameters by 3-dimensional grid search. The boundary of "Eq.(16)" illustrates (16) with $\zeta = \frac{n-1}{N}$, $n = 1, ..., N+1$ where $N = 100$. "ZF points" denotes two outmost points of the ZF points by Eq. (8).

If we consider the curve by "Proposed_1+9" as a nearly optimal boundary, we find that "Proposed_1" has a promising/robust performance only with one initialization by (16). Also observe that the proposed algorithm "Proposed_1+9" and "Proposed_1", and "WMMSE_10" yield a similar performance at convex parts of boundary and outperform the others under the same accuracy for convergence. However, since the weighted sum maximization method cannot achieve the non-convex boundary, and even the achieved points on the convex boundary are still unevenly distributed. This is why the "WMMSE_10" as a weighted sum maximization method does not achieve the part between "P1" and "P2" in Fig. 4.

To further evaluate the performance of "WMMSE_10" and "Proposed_1+9" on illustrating the Pareto boundary, another simulation is done and shown in Fig. 5. Even with fine weights ws in $[0.05 : 0.005 : 0.95]$, we find that there exists a large jump between the points "P3" and "P4" by "WMMSE_10" so that the rate region cannot be illustrated effectively.

Fig. 4.   Achievable boundary with SNR=10dB and $N_T = 3, N_R = 2$ for a random Gaussian channel data

Fig. 5.   Achievable boundary with SNR=10dB and $N_T = 3, N_R = 2$ for another random Gaussian channel data

In fact, only through the results by "WMMSE_10", we do not know what the part between "P1" and "P2" in Fig. 4 and the part between "P1" and "P2" in Fig. 5 look like. If they are concave parts, the degree of concavity is still unknown. For our proposed algorithm, although its curve denoted by "Proposed_1+9" is not guaranteed to exactly be the strict Pareto boundary, it has a even better performance than the *approximate Pareto boundary* (i.e., "Rand_10mil" curves in Fig. 4 and Fig. 5). Thus, it is able to serve as a more reasonable/complete inner bound of the whole strict Pareto boundary, especially the non-convex part. This is a main advantage of the proposed algorithm to the existing algorithms.

C. Illustration of the multi-user case

In order to evaluate the performance of Algorithm 3 for the multi-user MIMO IC, a three-user MIMO IC example is

simulated. Fig. 6 shows the fast convergence behavior when R_1 is maximized with $(R_3, R_2) = (0.2700, 2.3720)$ and the convergence threshold 10^{-4}. In Fig. 7, a three dimensional rate region is illustrated after computing R_1 for 65 samples of (R_3, R_2) and interpolation.

Fig. 6. Convergence behavior of Algorithm 3 for three-user MIMO IC with SNR=0dB and $N_T = 3, N_R = 2$.

Fig. 7. Achievable rate region for three-user MIMO IC with SNR=0dB and $N_T = 3, N_R = 2$. The color bar shows the sum rate.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, the non-strict Pareto boundary and two ending points of strict Pareto boundary of the two-user scenario are computed exactly. To find the strict Pareto boundary for the two/multi-user single-beam MIMO IC, we formulate an optimization problem to maximize the rate of one user while the rates of the other users are fixed such that a bound point is reached. This problem is different from the traditional problems (e.g., maximize (weighted) sum-rate and max-min rate) and there exists little work on this type of problem in the MIMO IC. We propose an alternating optimization algorithm to solve this non-convex problem. The convergence of the proposed algorithm is guaranteed and a high quality suboptimal solution is obtained. Furthermore, the proposed computation algorithm has two main advantages: 1) This algorithm can be applied to those optimization problems with rate constraints

(e.g. rate requirements of primary links in overlay cognitive radio environments or for private messages); 2) For the strict Pareto boundary, this algorithm is able to compute a high quality suboptimal operating point satisfying arbitrary rate requirements. A series of well-distributed rate requirements can lead to a sequence of *well-distributed* operating points, which can serve as a reasonable and complete inner bound of the strict Pareto boundary.

The proposed algorithm requires that each transmitter knows perfect global CSI, which is challenging for distributed systems. However, this work provides a benchmark for the algorithms that only imperfect/partial CSI is available at transmitters. Future work should focus on developing robust cooperative algorithms for the multi-user MIMO IC.

APPENDIX A
PROOF OF PROPOSITION 1

Proof: Given beamformers $\boldsymbol{w_1}$ and $\boldsymbol{w_2}$, according to the matrix inversion lemma [35], (2) can be rewritten as:

$$\text{SINR}_i(\boldsymbol{w}_1, \boldsymbol{w}_2)$$

$$= (\boldsymbol{H}_{ii}\boldsymbol{w}_i)^H \left(\frac{1}{\sigma_i^2}\boldsymbol{I} - \frac{\boldsymbol{H}_{ki}\boldsymbol{w}_k(\boldsymbol{H}_{ki}\boldsymbol{w}_k)^H}{\sigma_i^2(\sigma_i^2 + \|\boldsymbol{H}_{ki}\boldsymbol{w}_k\|^2)} \right) \boldsymbol{H}_{ii}\boldsymbol{w}_i$$

$$= \frac{\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2}{\sigma_i^2} - \frac{\left|(\boldsymbol{H}_{ii}\boldsymbol{w}_i)^H \boldsymbol{H}_{ki}\boldsymbol{w}_k\right|^2}{\sigma_i^2(\sigma_i^2 + \|\boldsymbol{H}_{ki}\boldsymbol{w}_k\|^2)}$$

$$= \frac{\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2}{\sigma_i^2} \cdot \left(1 - \frac{\left|\overrightarrow{\boldsymbol{H}_{ii}\boldsymbol{w}_i}^H \cdot \overrightarrow{\boldsymbol{H}_{ki}\boldsymbol{w}_k}\right|^2 \cdot \|\boldsymbol{H}_{ki}\boldsymbol{w}_k\|^2}{\sigma_i^2 + \|\boldsymbol{H}_{ki}\boldsymbol{w}_k\|^2} \right)$$

$$= \left(1 - \left|\overrightarrow{\boldsymbol{H}_{ii}\boldsymbol{w}_i}^H \cdot \overrightarrow{\boldsymbol{H}_{ki}\boldsymbol{w}_k}\right|^2 \right) \cdot \frac{\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2}{\sigma_i^2}$$

$$+ \left|\overrightarrow{\boldsymbol{H}_{ii}\boldsymbol{w}_i}^H \cdot \overrightarrow{\boldsymbol{H}_{ki}\boldsymbol{w}_k}\right|^2 \cdot \frac{\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2}{\sigma_i^2 + \|\boldsymbol{H}_{ki}\boldsymbol{w}_k\|^2}. \quad (23)$$

For two complex vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, the cosine of the complex-valued angle between $\boldsymbol{a}$ and $\boldsymbol{b}$ is defined as [37] $\cos(\theta_C) = \frac{\boldsymbol{a}^H\boldsymbol{b}}{\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|}$ where $\cos(\theta_C) = \mu e^{j\psi}$ with $\mu = |\cos(\theta_C)| \le 1$ and $-\pi \le \theta_C \le \pi$ is called pseudo angle between $\boldsymbol{a}$ and $\boldsymbol{b}$.

The Hermitian angle between $\boldsymbol{a}$ and $\boldsymbol{b}$ is defined as

$$\cos(\theta_H) = |\cos(\theta_C)| = \frac{|\boldsymbol{a}^H\boldsymbol{b}|}{\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|}, \quad 0 \le \theta_H \le \pi/2.$$

It implies $\left|\overrightarrow{\boldsymbol{H}_{ii}\boldsymbol{w}_i}^H \cdot \overrightarrow{\boldsymbol{H}_{ki}\boldsymbol{w}_k}\right|^2 = \cos^2(\theta_{H,i})$ because of $\left\|\overrightarrow{\boldsymbol{H}_{ii}\boldsymbol{w}_i}\right\|^2 = \left\|\overrightarrow{\boldsymbol{H}_{ki}\boldsymbol{w}_k}\right\|^2 = 1$. Thus, (23) becomes

$$\text{SINR}_i(\boldsymbol{w}_1, \boldsymbol{w}_2)$$

$$= \sin^2(\theta_{H,i}) \cdot \frac{\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2}{\sigma_i^2} + \cos^2(\theta_{H,i}) \cdot \frac{\|\boldsymbol{H}_{ii}\boldsymbol{w}_i\|^2}{\sigma_i^2 + \|\boldsymbol{H}_{ki}\boldsymbol{w}_k\|^2},$$

where $\theta_{H,i} \in [0, \pi/2]$ denotes the Hermitian angle between the desired signal direction $\overrightarrow{\boldsymbol{H}_{ii}\boldsymbol{w}_i}$ and the interference direction $\overrightarrow{\boldsymbol{H}_{ki}\boldsymbol{w}_k}$ at RX_i. Obviously, when $\overrightarrow{\boldsymbol{H}_{ii}\boldsymbol{w}_i} \parallel \overrightarrow{\boldsymbol{H}_{ki}\boldsymbol{w}_k}$ (or $\overrightarrow{\boldsymbol{H}_{ii}\boldsymbol{w}_i} \perp \overrightarrow{\boldsymbol{H}_{ki}\boldsymbol{w}_k}$), we have $\theta_{H,i} = 0$ (or $\theta_{H,i} = \pi/2$). ∎

APPENDIX B
PROOF OF PROPOSITION 2

Proof: Our idea is to show that it is impossible for a strict Pareto-optimal point achieved by the transmit beamformers with less than full power. The proof works by contradiction.

Assume that a strict Pareto-optimal point $\big(R_1(\boldsymbol{w}_1,\boldsymbol{w}_2), R_2(\boldsymbol{w}_1,\boldsymbol{w}_2)\big)$ is achieved by $(\boldsymbol{w}_1,\boldsymbol{w}_2)$ where $\|\boldsymbol{w}_1\|^2 < 1$ and $\|\boldsymbol{w}_2\|^2 \le 1$. We consider whether there exists an *outer*[9] point $\big(R_1(\hat{\boldsymbol{w}}_1,\boldsymbol{w}_2), R_2(\hat{\boldsymbol{w}}_1,\boldsymbol{w}_2)\big)$ achieved by $(\hat{\boldsymbol{w}}_1,\boldsymbol{w}_2)$ where $\|\boldsymbol{w}_1\|^2 < \|\hat{\boldsymbol{w}}_1\|^2 \le 1$ and $\|\boldsymbol{w}_2\|^2 \le 1$. If it exists, e.g., $R_1(\hat{\boldsymbol{w}}_1,\boldsymbol{w}_2) > R_1(\boldsymbol{w}_1,\boldsymbol{w}_2)$ and $R_2(\hat{\boldsymbol{w}}_1,\boldsymbol{w}_2) = R_2(\boldsymbol{w}_1,\boldsymbol{w}_2)$, we can improve the $R_1(\boldsymbol{w}_1,\boldsymbol{w}_2)$ only by consuming more transmit power while keeping $R_2(\boldsymbol{w}_1,\boldsymbol{w}_2)$ unchanged. Thus, the existence of an *outer* operating point contradicts the assumption that $\big(R_1(\boldsymbol{w}_1,\boldsymbol{w}_2), R_2(\boldsymbol{w}_1,\boldsymbol{w}_2)\big)$ is a strict Pareto-optimal point.

Define $\hat{\boldsymbol{w}}_1 = \boldsymbol{w}_1 + \boldsymbol{\delta}_p$. To guarantee the Pareto improvement, we need to show the existence of a nonzero perturbation vector $\boldsymbol{\delta}_p$ satisfying:

$$(\boldsymbol{w}_1+\boldsymbol{\delta}_p)^H \boldsymbol{A}_1(\boldsymbol{w}_2)(\boldsymbol{w}_1+\boldsymbol{\delta}_p) > \boldsymbol{w}_1^H \boldsymbol{A}_1(\boldsymbol{w}_2)\boldsymbol{w}_1 \quad (24a)$$

$$\boldsymbol{w}_2^H \boldsymbol{A}_2(\boldsymbol{w}_1+\boldsymbol{\delta}_p)\boldsymbol{w}_2 = \boldsymbol{w}_2^H \boldsymbol{A}_2(\boldsymbol{w}_1)\boldsymbol{w}_2 \quad (24b)$$

$$\|\boldsymbol{w}_1+\boldsymbol{\delta}_p\|^2 > \|\boldsymbol{w}_1\|^2 \quad (24c)$$

$$\|\boldsymbol{w}_1+\boldsymbol{\delta}_p\|^2 \le 1. \quad (24d)$$

An arbitrary nonzero $\boldsymbol{\delta}_p$ can be expressed as

$$\boldsymbol{\delta}_p = \|\boldsymbol{\delta}_p\| \cdot e^{j\phi_\delta} \cdot \overrightarrow{\boldsymbol{\delta}_p}. \quad (25)$$

It means that we should find $\|\boldsymbol{\delta}_p\|$, ϕ_δ and $\overrightarrow{\boldsymbol{\delta}_p}$ to satisfy all the conditions in (24) simultaneously. The proof of the existence of ϕ_δ is similar to that for the two-user MISO IC in [22]. However, it is more difficult to find a $\overrightarrow{\boldsymbol{\delta}_p}$ for the MIMO IC because the cross-talk channel matrix (rather than a vector in the MISO IC case) does not always have a null space for $\overrightarrow{\boldsymbol{\delta}_p}$. We give the proof in detail as follows.

1. Existence of $\overrightarrow{\boldsymbol{\delta}_p}$

By the matrix inverse lemma [35], the condition in (24b) is equivalent to

$$\frac{|(\boldsymbol{w}_1+\boldsymbol{\delta}_p)^H \boldsymbol{H}_{12}^H \boldsymbol{H}_{22}\boldsymbol{w}_2|^2}{\sigma_2^2 + \|\boldsymbol{H}_{12}(\boldsymbol{w}_1+\boldsymbol{\delta}_p)\|^2} = \frac{|\boldsymbol{w}_1^H \boldsymbol{H}_{12}^H \boldsymbol{H}_{22}\boldsymbol{w}_2|^2}{\sigma_2^2 + \|\boldsymbol{H}_{12}\boldsymbol{w}_1\|^2}. \quad (26)$$

It is difficult to solve $\boldsymbol{\delta}_p$ directly. In fact, we only need to prove the existence of $\boldsymbol{\delta}_p$ satisfying (26).

Case 1. $N_R < N_T$ or $\text{Rank}(\boldsymbol{H}_{12}) < N_T \le N_R$:
We always have $\boldsymbol{H}_{12}\boldsymbol{\delta}_p = 0$ if

$$\overrightarrow{\boldsymbol{\delta}_p} = \sum_{i=1}^{\text{Rank}(\Pi_{\boldsymbol{H}_{12}^T}^\perp)} a_i \boldsymbol{u}_i(\Pi_{\boldsymbol{H}_{12}^T}^\perp), \quad (27)$$

where $a_i, i=1,...,\text{Rank}(\Pi_{\boldsymbol{H}_{12}^T}^\perp)$ are complex-valued numbers and $\sum_{i=1}^{\text{Rank}(\Pi_{\boldsymbol{H}_{12}^T}^\perp)} |a_i|^2 = 1$. Then, (26) always holds because any $\boldsymbol{\delta}_p$ in the null space of the cross-talk channel $\boldsymbol{H}_{12}$ does not cause extra interference to RX$_2$.

---

[9]A point $\boldsymbol{p}' \in \mathbb{R}_n^+$ is called an outer point than $\boldsymbol{p} \in \mathbb{R}_n^+$, if $\boldsymbol{p}'$ dominates $\boldsymbol{p}$, i.e., $\boldsymbol{p}' \ge \boldsymbol{p}$ and $\boldsymbol{p}' \ne \boldsymbol{p}$ where the inequality is component-wise. The improvement from $\boldsymbol{p}$ to $\boldsymbol{p}'$ is called Pareto improvement.

Case 2. $\text{Rank}(\boldsymbol{H}_{12}) = N_T \le N_R$:

It is impossible to nullify the perturbation directly as Case 1. Define $\boldsymbol{v}_1 \triangleq \boldsymbol{H}_{12}\boldsymbol{w}_1$, $\boldsymbol{v}_\delta \triangleq \boldsymbol{H}_{12}\boldsymbol{\delta}_p$ and $\boldsymbol{v}_2 \triangleq \boldsymbol{H}_{22}\boldsymbol{w}_2$. Then (26) becomes

$$\frac{|(\boldsymbol{v}_1+\boldsymbol{v}_\delta)^H \boldsymbol{v}_2|^2}{\sigma_2^2 + \|\boldsymbol{v}_1+\boldsymbol{v}_\delta\|^2} = \frac{|\boldsymbol{v}_1^H \boldsymbol{v}_2|^2}{\sigma_2^2 + \|\boldsymbol{v}_1\|^2}. \quad (28)$$

Assume that $\boldsymbol{v}_\delta$ is a combination of two orthogonal vectors

$$\boldsymbol{v}_\delta \triangleq \|\boldsymbol{v}_\delta\| \cdot \big(\sqrt{\eta} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2}^\perp \boldsymbol{v}_1} + \sqrt{1-\eta} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2} \boldsymbol{v}_1}\big), \quad (29)$$

Note that $\overrightarrow{\Pi_{\boldsymbol{v}_2}\boldsymbol{v}_1} = \overrightarrow{\boldsymbol{v}_2} \cdot e^{-j\phi_1}$ where $\phi_1 = \arg(\boldsymbol{v}_1^H \boldsymbol{v}_2)$. Now, it remains to find whether there is a $\boldsymbol{v}_\delta$ in the plane spanned by $\overrightarrow{\Pi_{\boldsymbol{v}_2}^\perp \boldsymbol{v}_1}$ and $\overrightarrow{\Pi_{\boldsymbol{v}_2}\boldsymbol{v}_1}$ satisfying (28).

Substituting (29) into (28) yields

$$\frac{\big|\boldsymbol{v}_1^H \boldsymbol{v}_2 + \|\boldsymbol{v}_\delta\| \cdot \sqrt{1-\eta} \cdot e^{j\phi_1} \cdot \|\boldsymbol{v}_2\|\big|^2}{\sigma_2^2 + \big\|\boldsymbol{v}_1 + \|\boldsymbol{v}_\delta\|\big(\sqrt{\eta} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2}^\perp \boldsymbol{v}_1} + \sqrt{1-\eta} \cdot e^{-j\phi_1} \cdot \overrightarrow{\boldsymbol{v}_2}\big)\big\|^2}$$

$$= \frac{|\boldsymbol{v}_1^H \boldsymbol{v}_2|^2}{\sigma_2^2 + \|\boldsymbol{v}_1\|^2}. \quad (30)$$

Define the right-hand side and the left-hand side of (30) as R_{side} and $L_{side}(\eta)$, respectively. It is still hard to get a closed-form solution of η by solving (30) directly. Observe that the denominator of $L_{side}(\eta)$ is always positive for $\eta \in [0,1]$, and $L_{side}(\eta)$ as a function of η is continuous over the interval $[0,1]$. Therefore, if $(R_{side} - L_{side}(1))(R_{side} - L_{side}(0)) \le 0$, there must exist a $\boldsymbol{v}_\delta(\eta)$ with at least a certain $\eta \in [0,1]$ satisfying (30).

When $\eta = 1$, the term $L_{side}(\eta)$ becomes

$$L_{side}(1) = \frac{|\boldsymbol{v}_1^H \boldsymbol{v}_2|^2}{\sigma_2^2 + \big\|\boldsymbol{v}_1 + \|\boldsymbol{v}_\delta\| \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2}^\perp \boldsymbol{v}_1}\big\|^2}$$

$$= \frac{|\boldsymbol{v}_1^H \boldsymbol{v}_2|^2}{\sigma_2^2 + \|\boldsymbol{v}_1\|^2 + \|\boldsymbol{v}_\delta\|^2 + 2\|\boldsymbol{v}_\delta\| \cdot \|\Pi_{\boldsymbol{v}_2}^\perp \boldsymbol{v}_1\|}.$$

Observe that $\|\boldsymbol{v}_\delta\|^2 + 2\|\boldsymbol{v}_\delta\| \cdot \|\Pi_{\boldsymbol{v}_2}^\perp \boldsymbol{v}_1\| > 0$. Thus, we have $L_{side}(1) < R_{side}$.

When $\eta = 0$, the term $L_{side}(\eta)$ becomes

$$L_{side}(0) = \frac{\big|\boldsymbol{v}_1^H \boldsymbol{v}_2 + \|\boldsymbol{v}_\delta\| \cdot e^{j\phi_1} \cdot \|\boldsymbol{v}_2\|\big|^2}{\sigma_2^2 + \big\|\boldsymbol{v}_1 + \|\boldsymbol{v}_\delta\| \cdot e^{-j\phi_1} \cdot \overrightarrow{\boldsymbol{v}_2}\big\|^2}$$

$$= \frac{|\boldsymbol{v}_1^H \boldsymbol{v}_2|^2 + \|\boldsymbol{v}_2\|^2 \cdot (\|\boldsymbol{v}_\delta\|^2 + 2\|\boldsymbol{v}_\delta\| \cdot |\boldsymbol{v}_1^H \overrightarrow{\boldsymbol{v}_2}|)}{\sigma_2^2 + \|\boldsymbol{v}_1\|^2 + (\|\boldsymbol{v}_\delta\|^2 + 2\|\boldsymbol{v}_\delta\| \cdot |\boldsymbol{v}_1^H \overrightarrow{\boldsymbol{v}_2}|)},$$

where $\|\boldsymbol{v}_\delta\|^2 + 2 \cdot \|\boldsymbol{v}_\delta\| \cdot |\boldsymbol{v}_1^H \overrightarrow{\boldsymbol{v}_2}| > 0$. If $L_{side}(0) > R_{side}$, we need $\|\boldsymbol{v}_2\|^2 > R_{side}$. Furthermore, R_{side} is bounded by

$$R_{side} = \frac{|\boldsymbol{v}_1^H \boldsymbol{v}_2|^2}{\sigma_2^2 + \|\boldsymbol{v}_1\|^2} \le \frac{\|\boldsymbol{v}_1\|^2 \cdot \|\boldsymbol{v}_2\|^2}{\sigma_2^2 + \|\boldsymbol{v}_1\|^2} = \frac{\|\boldsymbol{v}_2\|^2}{\frac{\sigma_2^2}{\|\boldsymbol{v}_1\|^2} + 1} < \|\boldsymbol{v}_2\|^2.$$

Thus, we have $L_{side}(0) > R_{side}$.

Due to $L_{side}(1) < R_{side} < L_{side}(0)$, there exists at least one $\eta_0 \in (0,1)$ satisfying $L_{side}(\eta_0) = R_{side}$. In this case, $\boldsymbol{H}_{12}$ has an inverse/Moore-Penrose pseudo-inverse matrix

$\boldsymbol{H}_{12}^{\dagger}$. Then, we have

$$
\begin{aligned}
\boldsymbol{\delta}_p &= \|\boldsymbol{v}_\delta\| \cdot \boldsymbol{H}_{12}^{\dagger}\big(\sqrt{\eta_0} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2}^{\perp} \boldsymbol{v}_1} + \sqrt{1-\eta_0} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2} \boldsymbol{v}_1}\big), \\
&= \|\boldsymbol{\delta}_p\| \cdot e^{j\phi_\delta} \cdot \overrightarrow{\boldsymbol{\delta}_p},
\end{aligned} \tag{31}
$$

where $\|\boldsymbol{v}_\delta\| = \dfrac{\|\boldsymbol{\delta}_p\|}{\left\| \boldsymbol{H}_{12}^{\dagger}\big(\sqrt{\eta_0} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2}^{\perp} \boldsymbol{v}_1} + \sqrt{1-\eta_0} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2} \boldsymbol{v}_1}\big) \right\|}$ depends on but has no requirement for $\|\boldsymbol{\delta}_p\|$. Therefore, $\boldsymbol{v}_\delta$ with any $\|\boldsymbol{\delta}_p\|$ and $\overrightarrow{\boldsymbol{\delta}_p} = e^{-j\phi_\delta} \cdot \boldsymbol{H}_{12}^{\dagger}\big(\sqrt{\eta_0} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2}^{\perp} \boldsymbol{v}_1} + \sqrt{1-\eta_0} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2} \boldsymbol{v}_1}\big)$ satisfies (28).

Therefore, any $\boldsymbol{\delta}_p = \|\boldsymbol{\delta}_p\| \cdot e^{j\phi_\delta} \cdot \overrightarrow{\boldsymbol{\delta}_p}$ with

$$
\overrightarrow{\boldsymbol{\delta}_p} =
\begin{cases}
\displaystyle\sum_{i=1}^{\mathrm{Rank}(\Pi_{\boldsymbol{H}_{12}^T}^{\perp})} a_i \boldsymbol{u}_i(\Pi_{\boldsymbol{H}_{12}^T}^{\perp}), \quad \displaystyle\sum_{i=1}^{\mathrm{Rank}(\Pi_{\boldsymbol{H}_{12}^T}^{\perp})} |a_i|^2 = 1, \\
\quad \text{when } N_R < N_T \text{ or } \mathrm{Rank}(\boldsymbol{H}_{12}) < N_T \leq N_R \\
e^{-j\phi_\delta} \cdot \boldsymbol{H}_{12}^{\dagger}\big(\sqrt{\eta_0} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2}^{\perp} \boldsymbol{v}_1} + \sqrt{1-\eta_0} \cdot \overrightarrow{\Pi_{\boldsymbol{v}_2} \boldsymbol{v}_1}\big), \\
\quad \text{when } \mathrm{Rank}(\boldsymbol{H}_{12}) = N_T \leq N_R,
\end{cases} \tag{32}
$$

satisfies the condition (24b).

2. Existence of ϕ_δ

Substituting (25) into (24a) yields

$$
\begin{aligned}
&(\boldsymbol{w}_1 + \boldsymbol{\delta}_p)^H \boldsymbol{A}_1(\boldsymbol{w}_2)(\boldsymbol{w}_1 + \boldsymbol{\delta}_p) > \boldsymbol{w}_1^H \boldsymbol{A}_1(\boldsymbol{w}_2) \boldsymbol{w}_1 \\
&\Leftrightarrow \|\boldsymbol{\delta}_p\|^2 \overrightarrow{\boldsymbol{\delta}_p}^H \boldsymbol{A}_1(\boldsymbol{w}_2) \overrightarrow{\boldsymbol{\delta}_p} + 2\|\boldsymbol{\delta}_p\| \Re\big(\boldsymbol{w}_1^H \boldsymbol{A}_1(\boldsymbol{w}_2) \overrightarrow{\boldsymbol{\delta}_p} e^{j\phi_\delta}\big) > 0 \\
&\Leftrightarrow \frac{\|\boldsymbol{\delta}_p\|}{2} \overrightarrow{\boldsymbol{\delta}_p}^H \boldsymbol{A}_1(\boldsymbol{w}_2) \overrightarrow{\boldsymbol{\delta}_p} + \big|\boldsymbol{w}_1^H \boldsymbol{A}_1(\boldsymbol{w}_2) \overrightarrow{\boldsymbol{\delta}_p}\big| \cos(\phi_\delta + \phi_2) > 0 \\
&\Leftrightarrow \cos(\phi_\delta + \phi_2) > -\frac{\|\boldsymbol{\delta}_p\|}{2} \cdot \frac{\overrightarrow{\boldsymbol{\delta}_p}^H \boldsymbol{A}_1(\boldsymbol{w}_2) \overrightarrow{\boldsymbol{\delta}_p}}{\big|\boldsymbol{w}_1^H \boldsymbol{A}_1(\boldsymbol{w}_2) \overrightarrow{\boldsymbol{\delta}_p}\big|},
\end{aligned} \tag{33}
$$

where $\phi_2 \triangleq \arg(\boldsymbol{w}_1^H \boldsymbol{A}_1(\boldsymbol{w}_2) \overrightarrow{\boldsymbol{\delta}_p})$.

At the same time, substituting (25) into (24c) yields

$$
\begin{aligned}
&\|\boldsymbol{w}_1 + \boldsymbol{\delta}_p\|^2 > \|\boldsymbol{w}_1\|^2 \\
&\iff \|\boldsymbol{\delta}_p\|^2 + 2\|\boldsymbol{\delta}_p\| \Re\big(\boldsymbol{w}_1^H \overrightarrow{\boldsymbol{\delta}_p} e^{j\phi_\delta}\big) > 0 \\
&\iff \big|\boldsymbol{w}_1^H \overrightarrow{\boldsymbol{\delta}_p}\big| \cos(\phi_\delta + \phi_3) > -\frac{\|\boldsymbol{\delta}_p\|}{2} \\
&\iff \cos(\phi_\delta + \phi_3) > -\frac{\|\boldsymbol{\delta}_p\|}{2|\boldsymbol{w}_1^H \overrightarrow{\boldsymbol{\delta}_p}|},
\end{aligned} \tag{34}
$$

where $\phi_3 \triangleq \arg(\boldsymbol{w}_1^H \overrightarrow{\boldsymbol{\delta}_p})$.

Define $\phi_\delta + \phi_2 \in [\theta_1, \theta_2]$ and $\phi_\delta + \phi_3 \in [\theta_3, \theta_4]$. Since both the right-hand side of (33) and (34) are negative, the range $[\theta_1, \theta_2]$ and $[\theta_3, \theta_4]$ are strictly wider than π. In addition, the intersection of two angular ranges wider than π is nonempty. Then, for arbitrary $\|\boldsymbol{\delta}_p\|$ and $\overrightarrow{\boldsymbol{\delta}_p}$, any $\boldsymbol{\delta}_p = \|\boldsymbol{\delta}_p\| \cdot e^{j\phi_\delta} \overrightarrow{\boldsymbol{\delta}_p}$ in (25) with $\phi_\delta \in [\theta_1, \theta_2] \cap [\theta_3, \theta_4]$ always satisfies the conditions (33) and (34) simultaneously.

3. Existence of $\|\boldsymbol{\delta}_p\|$

The condition (24d) is equivalent to

$$
\begin{aligned}
&\|\boldsymbol{\delta}_p\|^2 + 2|\boldsymbol{w}_1^H \overrightarrow{\boldsymbol{\delta}_p}| \cos(\phi_\delta + \phi_3)\|\boldsymbol{\delta}_p\| + \|\boldsymbol{w}_1\|^2 - 1 \leq 0, \\
&\overset{(a)}{\Longleftrightarrow} \|\boldsymbol{\delta}_p\| \in \Big(0, -|\boldsymbol{w}_1^H \overrightarrow{\boldsymbol{\delta}_p}| \cos(\phi_\delta + \phi_3) \\
&\qquad + \sqrt{|\boldsymbol{w}_1^H \overrightarrow{\boldsymbol{\delta}_p}|^2 \cos^2(\phi_\delta + \phi_3) - (\|\boldsymbol{w}_1\|^2 - 1)}\Big),
\end{aligned} \tag{35}
$$

where the transformation (a) is based on $\|\boldsymbol{w}_1\| \leq 1$ and $\|\boldsymbol{\delta}_p\| > 0$. For arbitrary $\overrightarrow{\boldsymbol{\delta}_p}$ and ϕ_δ, any any $\boldsymbol{\delta}_p = \|\boldsymbol{\delta}_p\| \cdot e^{j\phi_\delta} \overrightarrow{\boldsymbol{\delta}_p}$ with $\|\boldsymbol{\delta}_p\|$ in (35) will satisfy the condition (24d).

Above all, the existence of $\overrightarrow{\boldsymbol{\delta}_p}$, ϕ_δ and $\|\boldsymbol{\delta}_p\|$ has been proved. That is, there always exists some $\boldsymbol{\delta}_p = \|\boldsymbol{\delta}_p\| \cdot e^{j\phi_\delta} \cdot \overrightarrow{\boldsymbol{\delta}_p}$ satisfying all the conditions in (24). Then, $R_1(\boldsymbol{w}_1, \boldsymbol{w}_2)$ can still be improved until $\|\boldsymbol{w}_1\|^2 = 1$, while $R_1(\boldsymbol{w}_1, \boldsymbol{w}_2)$ remains unchanged simultaneously. This contradicts the assumption that $\big(R_1(\boldsymbol{w}_1, \boldsymbol{w}_2), R_2(\boldsymbol{w}_1, \boldsymbol{w}_2)\big)$ is on the strict Pareto boundary. Therefore, Proposition 2 holds. ∎

APPENDIX C
PROOF OF PROPOSITION 3

Proof: For the ending point $E1(\overline{R}_1, \underline{R}_2)$, to achieve the maximum rate of link $\mathrm{TX}_1 \mapsto \mathrm{RX}_1$, i.e., $\overline{R}_1$ in (4a), $(\boldsymbol{w}_1, \boldsymbol{w}_2)$ should satisfy the following conditions

$$
\begin{aligned}
&\boldsymbol{w}_1 = \boldsymbol{w}_1^{Ego} = \boldsymbol{u}_1(\boldsymbol{H}_{11}^H \boldsymbol{H}_{11}), \tag{36a} \\
&\theta_{H,1} = \pi/2 \Leftrightarrow \boldsymbol{w}_2 \perp \boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1. \tag{36b}
\end{aligned}
$$

This means $\boldsymbol{w}_2$ should be in the null space of $\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}$ to cause no interference to RX_1. All $\boldsymbol{w}_2 \in \mathcal{W}_{\mathcal{FP}}$ satisfying (36) form a set of the ZF strategies, defined as $\mathcal{W}_{\mathcal{ZF}}$. Then, any $\boldsymbol{w}_2 \in \mathcal{W}_{\mathcal{ZF}}$ can be expressed by

$$
\boldsymbol{w}_2 = \overrightarrow{\Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp} \boldsymbol{v}_2}, \tag{37}
$$

where $\boldsymbol{v}_2 \in \mathbb{C}^{N_T \times 1}$ and $\boldsymbol{v}_2 \nparallel \boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}$.

To achieve $(\overline{R}_1, \underline{R}_2)$, we need to find $\boldsymbol{v}_2^{opt}$ which maximizes $\mathrm{SINR}_2(\boldsymbol{w}_1^{Ego}, \boldsymbol{w}_2)$ simultaneously. Here, we define the optimal "altruistic" strategy $\boldsymbol{w}_2^{Alt}$ as

$$
\boldsymbol{w}_2^{Alt} \triangleq \arg \max_{\boldsymbol{w}_2 \in \mathcal{W}_{\mathcal{ZF}}} \boldsymbol{w}_2^H \boldsymbol{A}_2(\boldsymbol{w}_1^{Ego}) \boldsymbol{w}_2 \overset{(a)}{\Longleftrightarrow}
$$

$$
\boldsymbol{v}_2^{opt} = \arg \max_{\boldsymbol{v}_2} \frac{\boldsymbol{v}_2^H \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp,H} \boldsymbol{A}_2(\boldsymbol{w}_1^{Ego}) \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp} \boldsymbol{v}_2}{\boldsymbol{v}_2^H \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp,H} \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp} \boldsymbol{v}_2}
$$

$$
\overset{(b)}{=} \arg \max_{\boldsymbol{v}_2} \frac{\boldsymbol{v}_2^H \overbrace{\Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp} \boldsymbol{A}_2(\boldsymbol{w}_1^{Ego}) \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp}}^{\triangleq \boldsymbol{B}_1} \boldsymbol{v}_2}{\boldsymbol{v}_2^H \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp} \boldsymbol{v}_2}
$$

$$
= \boldsymbol{u}_1\big(\boldsymbol{B}_1, \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp}\big), \tag{38}
$$

where transformation (a) is based on (37) and transformation (b) is due to $\Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp,H} = \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp}$ and $\Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp,H} \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp} = \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp}$. Substituting (38) into (37), we obtain the optimal "altruistic" strategy

$$
\boldsymbol{w}_2^{Alt} = \overrightarrow{\Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp} \boldsymbol{u}_1(\boldsymbol{B}_1, \Pi_{\boldsymbol{H}_{21}^H \boldsymbol{H}_{11} \boldsymbol{w}_1^{Ego}}^{\perp})}.
$$

Therefore, $E1(\overline{R}_1, \underline{R}_2)$ is achieved by $(\boldsymbol{w}_1^{Ego}, \boldsymbol{w}_1^{Alt})$. ∎

APPENDIX D
SOLVING THE SDR PROBLEM (P4)

Lemma 1 *Both the fractional problem* (P4) *and the problem* (P5) *are solvable.* □

Proof: For the fractional problem (P4), its constraint set
$$\Omega = \{ \boldsymbol{W}_2 \succeq \boldsymbol{0} : \mathrm{Tr}(\boldsymbol{W}_2) = 1, \ \mathrm{Tr}(\boldsymbol{A}_2(\boldsymbol{w}_1)\boldsymbol{W}_2) = \mathrm{SINR}_2^\star \}$$
is nonempty and compact.

In (P4), the denominator of the objective over Ω satisfies
$$\sigma_1^2 + \lambda_{N_T}(\boldsymbol{H}_{21}^H \boldsymbol{H}_{21}) = \lambda_{N_T}(\boldsymbol{C}_2) \leq \mathrm{Tr}(\boldsymbol{C}_2 \boldsymbol{W}_2)$$
$$\leq \lambda_1(\boldsymbol{C}_2) = \sigma_1^2 + \lambda_1(\boldsymbol{H}_{21}^H \boldsymbol{H}_{21}),$$

and the numerator obviously satisfies
$$0 \leq \mathrm{Tr}(\boldsymbol{C}_1(\boldsymbol{w}_1)\boldsymbol{W}_2) \leq \lambda_1(\boldsymbol{H}_{21}^H \boldsymbol{H}_{11}\boldsymbol{H}_{11}^H \boldsymbol{H}_{21}).$$

This implies that the objective (P4) over Ω is bounded by
$$0 \leq \frac{\mathrm{Tr}(\boldsymbol{C}_1(\boldsymbol{w}_1)\boldsymbol{W}_2)}{\mathrm{Tr}(\boldsymbol{C}_2 \boldsymbol{W}_2)} \leq \frac{\lambda_1(\boldsymbol{H}_{21}^H \boldsymbol{H}_{11}\boldsymbol{H}_{11}^H \boldsymbol{H}_{21})}{\sigma_1^2 + \lambda_{N_T}(\boldsymbol{H}_{21}^H \boldsymbol{H}_{21})}. \tag{39}$$

Based on Weierstrass' Theorem, the problem (P4) always has an optimal solution.

Assume that $\boldsymbol{W}_2^\star$ is an optimal solution to (P4), we know that $s^\star = \frac{1}{\mathrm{Tr}(\boldsymbol{C}_2 \boldsymbol{W}_2^\star)}$ and $\boldsymbol{Q}^\star = s^\star \boldsymbol{W}_2^\star$ are feasible for (P5). Also note that the objective is bounded by (39). Similarly, (P5) is solvable according to Weierstrass' Theorem. ∎

Lemma 2 *The problems* (P4) *and* (P5) *have the same value. Furthermore, if* $\boldsymbol{W}_2^\star$ *solves* (P4), *then* $s^\star = \frac{1}{\mathrm{Tr}(\boldsymbol{C}_2 \boldsymbol{W}_2^\star)}$ *and* $\boldsymbol{Q}^\star = s^\star \cdot \boldsymbol{W}_2^\star$ *solves* (P5); *if* $\boldsymbol{Q}^\star$ *and* $s^\star$ *solve* (P5), *then* $\boldsymbol{W}_2^\star = \frac{\boldsymbol{Q}^\star}{s^\star}$ *solves* (P4). □

Proof: Assume that $\boldsymbol{W}_2^\star$ is an optimal solution to (P4), and $v_{(P4)}^\star$ and $v_{(P5)}^\star$ are the optimal values of the objective of (P4) and (P5), respectively. Thus, $s = \frac{1}{\mathrm{Tr}(\boldsymbol{C}_2 \boldsymbol{W}_2^\star)}$ and $\boldsymbol{Q} = s\boldsymbol{W}_2^\star$ are feasible for (P5). The value of the objective of (P5) at this feasible point is
$$v_{(P5)} = \mathrm{Tr}(\boldsymbol{C}_1(\boldsymbol{w}_1)\boldsymbol{Q})$$
$$= \mathrm{Tr}(\boldsymbol{C}_1(\boldsymbol{w}_1)(s \cdot \boldsymbol{W}_2^\star)) = \frac{\mathrm{Tr}(\boldsymbol{C}_1(\boldsymbol{w}_1)\boldsymbol{W}_2^\star)}{\mathrm{Tr}(\boldsymbol{C}_2 \boldsymbol{W}_2^\star)} = v_{(P4)}^\star$$
$$\geq v_{(P5)}^\star.$$

On the other hand, suppose that $\boldsymbol{Q}^\star$ and $s^\star$ are the optimal solutions to (P5). Since $s^\star$ is always positive, $\boldsymbol{W}_2 = \frac{\boldsymbol{Q}^\star}{s^\star}$ is also feasible for (P4). Then, the value of the objective of (P4) at this feasible point is
$$v_{(P4)} = \frac{\mathrm{Tr}(\boldsymbol{C}_1(\boldsymbol{w}_1)\boldsymbol{W}_2)}{\mathrm{Tr}(\boldsymbol{C}_2 \boldsymbol{W}_2)} = \frac{\mathrm{Tr}(\boldsymbol{C}_1(\boldsymbol{w}_1)\frac{\boldsymbol{Q}^\star}{s^\star})}{\mathrm{Tr}(\boldsymbol{C}_2 \frac{\boldsymbol{Q}^\star}{s^\star})}$$
$$= \frac{\mathrm{Tr}(\boldsymbol{C}_1(\boldsymbol{w}_1)\boldsymbol{Q}^\star)}{\mathrm{Tr}(\boldsymbol{C}_2 \boldsymbol{Q}^\star)} = \mathrm{Tr}(\boldsymbol{C}_1(\boldsymbol{w}_1)\boldsymbol{Q}^\star) = v_{(P5)}^\star$$
$$\geq v_{(P4)}^\star.$$

Above all, we have $v_{(P4)}^\star = v_{(P5)}^\star$. ∎

APPENDIX E
PROOF OF PROPOSITION 4

Proof: We need to find a feasible set $\mathcal{W}_\mathcal{F}$ such that there exists at least one solution $\boldsymbol{w}_1 \in \mathcal{W}_{\mathcal{FP}}$ to problem (P0) by fixing $\boldsymbol{w}_2 \in \mathcal{W}_\mathcal{F}$.

In (9b), we derive that the constraint of (P0) is equivalent to $\boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \boldsymbol{H}_{22} \boldsymbol{w}_2 \geq \sigma_2^2 \mathrm{SINR}_2^\star$ and
$$\boldsymbol{w}_1^H \boldsymbol{C}(\boldsymbol{w}_2) \boldsymbol{w}_1 = 0. \tag{40}$$

To guarantee the existence of $\boldsymbol{w}_1 \in \mathcal{W}_{\mathcal{FP}}$ in (40), a feasible $\boldsymbol{w}_2$ should be determined in a way such that (9b) holds.

By the eigen-decomposition, $\boldsymbol{C}(\boldsymbol{w}_2)$ can be rewritten as $\sum_{i=1}^{N_T} \lambda_i(\boldsymbol{C}(\boldsymbol{w}_2))\boldsymbol{u}_i(\boldsymbol{C}(\boldsymbol{w}_2))\boldsymbol{u}_i^H(\boldsymbol{C}(\boldsymbol{w}_2))$. We analyze $\boldsymbol{C}(\boldsymbol{w}_2)$ for two cases.

Case 1. When $\boldsymbol{C}(\boldsymbol{w}_2)$ is a full rank matrix, i.e., $\lambda_i(\boldsymbol{C}(\boldsymbol{w}_2)) \neq 0, \forall i = 1, ..., N_T$. If $\boldsymbol{C}(\boldsymbol{w}_2)$ is a positive or negative definite matrix, it is clear that there is no nonzero vector $\boldsymbol{w}_1$ satisfying (40). Otherwise, $\boldsymbol{C}(\boldsymbol{w}_2)$ has $\lambda_1(\boldsymbol{C}(\boldsymbol{w}_2)) > 0$ and $\lambda_{N_T}(\boldsymbol{C}(\boldsymbol{w}_2)) < 0$, a sufficient solution to (40) is
$$\boldsymbol{w}_1 = \sqrt{\frac{\lambda_{N_T}(\boldsymbol{C}(\boldsymbol{w}_2))}{\lambda_1(\boldsymbol{C}(\boldsymbol{w}_2)) - \lambda_{N_T}(\boldsymbol{C}(\boldsymbol{w}_2))}} j \cdot \boldsymbol{u}_1(\boldsymbol{C}(\boldsymbol{w}_2))$$
$$+ \sqrt{\frac{\lambda_1(\boldsymbol{C}(\boldsymbol{w}_2))}{\lambda_1(\boldsymbol{C}(\boldsymbol{w}_2)) - \lambda_{N_T}(\boldsymbol{C}(\boldsymbol{w}_2))}} \cdot \boldsymbol{u}_{N_T}(\boldsymbol{C}(\boldsymbol{w}_2)).$$

Case 2. When $\boldsymbol{C}(\boldsymbol{w}_2)$ is not a full rank matrix, i.e., $\lambda_i(\boldsymbol{C}(\boldsymbol{w}_2)) = 0$ for some $i \in \{1, ..., N_T\}$. In this case, $\boldsymbol{C}(\boldsymbol{w}_2)$ always has null space for $\boldsymbol{w}_1$ to satisfy (40).

Above all, the sufficient and necessary condition of $\boldsymbol{w}_2$ satisfying (40) is $\lambda_1(\boldsymbol{C}(\boldsymbol{w}_2)) \cdot \lambda_{N_T}(\boldsymbol{C}(\boldsymbol{w}_2)) \leq 0$. That is, any $\boldsymbol{w}_2 \in \mathcal{W}_\mathcal{F}$ is always feasible for (9b) where $\mathcal{W}_\mathcal{F}$ is
$$\mathcal{W}_\mathcal{F} \triangleq \Big\{ \boldsymbol{w}_2 \in \mathcal{W}_{\mathcal{FP}} : \boldsymbol{w}_2^H \boldsymbol{H}_{22}^H \boldsymbol{H}_{22} \boldsymbol{w}_2 \geq \sigma_2^2 \mathrm{SINR}_2^\star,$$
$$\lambda_1(\boldsymbol{C}(\boldsymbol{w}_2)) \cdot \lambda_{N_T}(\boldsymbol{C}(\boldsymbol{w}_2)) \leq 0 \Big\}.$$
∎

REFERENCES

[1] L. Zadeh, "Optimality and non-scalar-valued performance criteria," *IEEE Trans. Automat. Contr.*, vol. 8, no. 1, pp. 59–60, Jan. 1963.

[2] E. Jorswieck, E. 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, Oct. 2008.

[3] J. Lindblom, E. Karipidis, and E. Larsson, "Closed-form parameterization of the Pareto boundary for the two-user MISO interference channel," in *Proc. IEEE ICASSP*, Prague, Czech Republic, May 2011.

[4] R. Mochaourab and E. Jorswieck, "Walrasian equilibrium in two-user multiple-input single-output interference channel," in *Proc. IEEE ICC*, Kyoto, Japan, 2011.

[5] R. Mochaourab and E. Jorswieck, "Optimal beamforming in interference networks with perfect local channel information," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1128 – 1141, Mar. 2011.

[6] E. Bjornson, M. Bengtsson, and B. Ottersten, "Pareto characterization of the multicell MIMO performance region with simple receivers," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4464–4469, Aug. 2012.

[7] E. Karipidis and E. Larsson, "Efficient computation of the Pareto boundary for the MISO interference channel with perfect CSI," in *Proc. WiOpt 2010*, Avignon, France, June 2010.

[8] P. Cao, S. Shi, and E. Jorswieck, "Efficient computation of the Pareto boundary for the two-user single-stream MIMO interference channel," in *Proc. IEEE SPAWC*, Cesme, Turkey, June 2012.

[9] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5450–5458, Oct. 2010.

[10] J. Qiu, R. Zhang, Z.-Q. Luo, and S. Cui, "Optimal distributed beamforming for MISO interference channels," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5638–5643, Nov. 2011.

[11] L. Qian, A. Y. Zhang, and J. Huang, "MAPEL: achieving global optimality for a non-convex wireless power control problem," *IEEE Trans. Wirel. Commun.*, vol. 8, no. 3, pp. 1553–1563, Mar. 2009.

[12] 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. Wirel. Commun.*, vol. 11, no. 5, pp. 1933–1945, May 2012.

[13] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[14] C. Shi, D. A. Schmidt, R. A. Berry, M. L. Honig, and W. Utschick, "Distributed interference pricing for the MIMO interference channel," in *Proc. IEEE ICC*, Dresden, Germany, 2009.

[15] C.-B. Chae, I. Hwang, Jr. R. W. Heath, and V. Tarokh, "Interference Aware-Coordinated Beamforming in a Multi-Cell System," *IEEE Trans. Wirel. Commun.*, vol. 11, no. 10, pp. 3692–3703, Oct. 2012.

[16] Z. K. M. Ho and D. Gesbert, "Balancing egoism and altruism on the interference channel: The MIMO case," in *Proc. IEEE ICC*, Cape Town, South Africa, 2010.

[17] S. W. Peters and Jr. R. W. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 206–218, Jan. 2011.

[18] I. Das and J.E. Dennis, "A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems," *Struct. and Multidiscip. Optim.*, vol. 14, no. 1, pp. 63–69, 1997.

[19] S. Ye and R. S. Blum, "Optimized signalling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Oct. 2003.

[20] G. Scutari, D. P. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: a unified view," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1089–1103, Sep. 2008.

[21] J. E. Cohen, "Cooperation and self-interest: Pareto-inefficiency of Nash equilibria in finite random games," *Proc. Natl. Acad. Sci. USA*, vol. 95, pp. 9724–9731, Aug. 1998.

[22] E. G. Larsson and E. A. Jorswieck, "Competition versus cooperation on the MISO interference channel," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1059–1069, Sep. 2008.

[23] Z. Chen, S. A. Vorobyov, C.-X. Wang, and J. Thompson, "Pareto region characterization for rate control in MIMO interference systems and Nash bargaining," *IEEE Trans. Automat. Contr.*, vol. 57, no. 12, pp. 3203–3208, Dec. 2012.

[24] C. Shi, R. A. Berry, and M. L. Honig, "Monotonic convergence of distributed interference pricing in wireless networks," in *Proc. IEEE ISIT*, Seoul, Korea, Jun. 2009.

[25] C. Shi, R. A. Berry, and M. L. Honig, "Local interference pricing for distributed beamforming in MIMO networks," in *Proc. IEEE MILCOM*, Boston, MA, Oct. 2009.

[26] E. A. Jorswieck and E. G. Larsson, "The MISO interference channel from a game-theoretic perspective: a combination of selfishness and altruism achieve Pareto optimality," in *Proc. IEEE ICASSP*, Las Vegas, NV, Apr. 2008.

[27] D. P. Bertsekas, "Nonlinear programming," *2nd ed. Athena Scientific*, 1999.

[28] I. Polik, "Addendum to the SeDuMi user guide: Version 1.1," 2005, available from http://sedumi.ie.lehigh.edu.

[29] M. Grant and S. Boyd, "CVX: matlab software for disciplined convex programming," 2012, available from http://cvxr.com/cvx/download/.

[30] W. Ai, Y. Huang, and S. Zhang, "New results on Hermitian matrix rank-one decomposition," *Math. Program: Ser. A*, vol. 128, no. 1-2, pp. 253–283, Jun. 2011.

[31] A. Charnes and W. W. Cooper, "Programming with linear fractional functions," *Naval Res. Logist. Quarter.*, vol. 9, pp. 181–186, 1962.

[32] M. V. Solodov, "On the convergence of constrained parallel variable distribution algorithm," *SIAM J. Optimization*, vol. 8, no. 1, pp. 187–196, Feb. 1998.

[33] Z.-Q. Luo, W.-K. Ma, M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[34] Y. Huang, A. De Maio, and S. Zhang, "Semidefinite programming, matrix decomposition, and radar code design," in *Convex Optimization in Signal Processing and Communications*, Cambridge University Press, New York, 2010, pp. 192–228.

[35] C. D. Meyer, *Matrix analysis and applied linear algebra*, Cambridge University Press, New York, NY, USA, 2000.

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

[37] K. Scharnhorst, "Angles in complex vector spaces," *Acta Appl. Math.*, vol. 69, no. 1, pp. 95–103, Oct. 2001.

Pan Cao (S'12) received his B.Eng. degree (with honors) in Department of Mechano-Electronic Engineering in 2008 and M. Eng. degree in Information and Signal Processing in 2011, both from Xidian University, Xi'an, P. R. China. Now, he is with the Communications Laboratory at Dresden University of Technology (TUD), Germany, to peruse his Ph.D. degree under the supervision of Prof. Eduard A. Jorswieck.

His research interests include resource allocation for MIMO interference channels and relay channels with the application of optimization techniques and game theory. He received China Scholarship Council (CSC) Scholarship from 2010 to 2014, the Best Student Paper Award of the 13th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC) in 2012, and the Qualcomm Innovation Fellowship (QInF) in 2013.

Eduard A. Jorswieck (S'01-M'05-SM'08) received his Diplom-Ingenieur degree and Doktor-Ingenieur (Ph.D.) degree, both in electrical engineering and computer science from the Berlin University of Technology (TUB), Germany, in 2000 and 2004, respectively. He was with the Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institute (HHI) Berlin, from 2001 to 2006. In 2006, he joined the Signal Processing Department at the Royal Institute of Technology (KTH) as a postdoc and became a Assistant Professor in 2007. Since February 2008, he has been the head of the Chair of Communications Theory and Full Professor at Dresden University of Technology (TUD), Germany.

His research interests are within the areas of applied information theory, signal processing and wireless communications. He is senior member of IEEE and elected member of the IEEE SPCOM Technical Committee. From 2008-2011 he served as an Associate Editor and since 2012 as a Senior Associate Editor for IEEE SIGNAL PROCESSING LETTERS. Since 2011 he serves as an Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING. Since 2013, Eduard serves as Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. In 2006, he was co-recipient of the IEEE Signal Processing Society Best Paper Award.

Shuying Shi received her Ph.D. degree in electrical engineering from the Technische Universität Berlin, Germany, in 2009, and her M.Sc. degree in electrical engineering from the Technische Universität Kaiserslautern, Germany, in 2002. She worked at the Fraunhofer German-Sino Lab for Mobile Communications (MCI), Germany, from 2003 to 2006, at the Technische Universität Berlin, Germany, from 2006 to 2009, at the Linköing University, Sweden, from 2009 to 2010, and at the Technische Universität Dresden, Germany, from 2010 to 2011.

Her research interests include multi-user MIMO signal processing, dynamic resource allocation, and the application of optimization techniques to the design of wireless communications.