On the Synergistic Benefits of Reconfigurable Antennas and Partial Channel Knowledge for the MIMO Interference Channel

Bofeng Yuan, Nilab Ismailoglu, and Syed A. Jafar

***Abstract*—Blind Interference Alignment (BIA) schemes create and exploit channel coherence patterns without the knowledge of channel realizations at transmitters, while beamforming schemes rely primarily on channel knowledge available to the transmitters without regard to channel coherence patterns. In order to explore the compatibility of these disparate ideas and the possibility of synergistic gains, this work studies the Degrees of Freedom (DoF) of the 2-user $(M_1 \times N_1)(M_2 \times N_2)$ Multiple-Input Multiple-Output (MIMO) Interference Channel (IC) where Transmitter 1 is equipped with reconfigurable antennas and has no channel knowledge, while Transmitter 2 has partial channel knowledge but no reconfigurable antennas. Taking a fundamental dimensional analysis perspective, the main question is to identify which antenna configurations allow synergistic DoF gains. The main results of this work are two-fold. The first result identifies antenna configurations where both reconfigurable antennas and partial channel knowledge are individually beneficial, as those where $M_1 < N_1 < \min(M_2, N_2)$. The second result shows that synergistic gains exist in each of these settings, over the best known solutions that rely on either reconfigurable antennas or partial channel knowledge alone. Coding schemes that jointly exploit partial channel knowledge and reconfigurable antennas emerge as a byproduct of the analysis.**

***Index Terms*—Blind Interference Alignment, Degrees of Freedom, Interference Channel, MIMO, Partial CSIT**

I. Introduction

MUCH of the recent progress in understanding the fundamental limits of interference management has come from Degrees of Freedom (DoF) studies. In spite of the inherent limitations of first order asymptotic analysis, DoF studies have been instrumental in finding solutions to a number of fundamental challenges. This includes ways to exploit delayed channel knowledge [1], [2], alternating channel knowledge [3], mixed channel knowledge [4], [5], and key breakthroughs in the pursuit of robust interference management principles — e.g., Topological Interference Management (TIM) [6], [7], [8], Rate-splitting [9], [10], Treating Interference as Noise (TIN) [11], [12], [13] and Blind Interference Alignment (BIA) [14], [15], [16]. In practice, multiple challenges may be faced simultaneously. The solutions to each of those challenges individually identify key pieces of a puzzle. Once the individual pieces of the puzzle are identified, it is important to understand how they fit together. In particular, are there ways to combine disparate ideas such that they yield new synergistic gains? This potential for the discovery of synergistic gains is the main motivation of this work.

The focus of this work is on the 2-user $(M_1 \times N_1)(M_2 \times N_2)$ Multiple-Input Multiple-Output (MIMO) Interference Channel (IC), where the transmitter and receiver of User i are equipped with M_i and N_i antennas, respectively, $i \in \{1, 2\}$. In particular, we are interested in the compatibility of BIA schemes with beamforming schemes. BIA schemes rely on the presence of reconfigurable antennas, to create and exploit channel coherence patterns without the knowledge of channel realizations at transmitters [14], [16], while beamforming schemes rely primarily on the channel knowledge available at transmitters without regard to channel coherence patterns [17]. In order to explore the potential of synergistic gains, we study the DoF of the 2-user MIMO IC where Transmitter 1 (Tx1) is equipped with reconfigurable antennas and has no channel knowledge, while Transmitter 2 (Tx2) has partial channel knowledge but no reconfigurable antennas.

Let $\mathcal{D}_{\alpha\beta}$ denote the DoF region for this setting, where α is a binary parameter whose value is 1 if Tx1 is equipped with reconfigurable antennas, and 0 otherwise. Similarly, let β be a continuous parameter whose value increases continuously from 0 to 1 as the Channel State Information at the Transmitter (CSIT) level of User 2 increases continuously from 'no CSIT' to 'perfect CSIT'. Note that in general, the characterization of $\mathcal{D}_{\alpha\beta}$ remains an open problem, partly due to the difficulty of finding outer bounds under channel uncertainty and channel coherence patterns as stated in [18]. In fact, even $\mathcal{D}_{01}$ and $\mathcal{D}_{10}$ are not fully known. To circumvent this barrier, our search for synergistic gains relies on the best known *achievable* DoF regions, denoted by $\underline{\mathcal{D}}_{\alpha\beta}$ when the actual regions $\mathcal{D}_{\alpha\beta}$ remain unknown.

We allow arbitrary numbers of antennas at each node. For certain antenna configurations, the CSIT at Tx2 is not useful in the DoF sense, i.e., $\mathcal{D}_{01} = \mathcal{D}_{00}$. Note that, this also means that $\mathcal{D}_{0\beta} = \mathcal{D}_{01} = \mathcal{D}_{00}$ for all $\beta \in [0, 1]$. Similarly, for some antenna configurations, BIA is not useful, i.e., $\mathcal{D}_{10} = \mathcal{D}_{00}$. Even though $\mathcal{D}_{01}$ and $\mathcal{D}_{10}$ are both unknown in general, in this work we are able to identify precisely which antenna configurations have both $\mathcal{D}_{01} \neq \mathcal{D}_{00}$ and $\mathcal{D}_{10} \neq \mathcal{D}_{00}$, i.e., which antenna configurations benefit from both reconfigurable

This paper was presented in part at IEEE SPAWC 2019.

This work is supported in part by the grants NSF CNS 1731384, NSF CCF 1907053, ONR N00014-21-1-2386 and ARO W911NF-19-1-0344.

B. Yuan is with Qualcomm Technologies, Santa Clara, CA 95051 USA (e-mail: bofengy@uci.edu).

N. Ismailoglu and S. Jafar are with the Center for Pervasive Communications and Computing, Department of Electrical Engineering and Computer Science, University of California at Irvine, Irvine, CA 92697 USA (e-mail: nismailo@uci.edu, syed@uci.edu).

antennas and partial CSIT, individually. The main result of this work is to show that for these antenna configurations, there always exists a synergistic DoF gain over the best known achievable schemes, i.e., there exists a partial CSIT level $\beta \in (0,1)$, such that $\text{co}(\mathcal{D}_{10} \cup \underline{\mathcal{D}}_{0\beta}) \subsetneq \mathcal{D}_{1\beta}$, where $\text{co}(\cdot)$ is the convex hull operation, and $A \subsetneq B$ means that A is a subset of B but A is not equal to B. In other words, we can achieve *more* by jointly exploiting reconfigurable antennas and partial CSIT than what we could by exploiting each of them individually. The key to the synergistic gain is a new achievable scheme that allows User 2 to carefully design the transmitted signal, so that BIA and partial Zero-Forcing (ZF) approaches can be used simultaneously. In addition to the general achievable scheme, it is demonstrated through two representative examples that rate-splitting [19], [9], [20], [10] can also be combined with BIA and partial ZF schemes in non-trivial ways.

II. System Model

Consider the 2-user $(M_1 \times N_1)(M_2 \times N_2)$ MIMO IC, where each transmitter sends an independent message to its corresponding receiver. The M_1 antennas at Tx1 are assumed to be reconfigurable antennas, each of which can switch among N_1 preset modes. All antennas at the other transmitters and receivers are conventional antennas. Tx2 is assumed to have partial CSIT for its channel to Receiver 1 (Rx1). No other CSIT is assumed. All receivers have perfect channel knowledge.

For simplicity, we assume all symbols take only real values. At time slot $t \in \mathbb{Z}^+$, the channel input-output equations are given by,

$$\mathbf{y}_1(t) = \mathbf{H}_{11}^m(t)\mathbf{x}_1(t) + \mathbf{H}_{12}(t)\mathbf{x}_2(t) + \mathbf{z}_1(t), \quad (1)$$
$$\mathbf{y}_2(t) = \mathbf{H}_{21}^m(t)\mathbf{x}_1(t) + \mathbf{H}_{22}(t)\mathbf{x}_2(t) + \mathbf{z}_2(t). \quad (2)$$

Here, $\mathbf{x}_k(t) = [x_{k1}(t), x_{k2}(t), ..., x_{kM_k}(t)]^T \in \mathbb{R}^{M_k \times 1}$ is the real signal vector sent from Transmitter k, which satisfies an average power constraint $\mathbb{E}(\|\mathbf{x}_k(t)\|^2) \le P$. $\mathbf{y}_k(t) \in \mathbb{R}^{N_k \times 1}$ is the received signal vector at Receiver k. $\mathbf{z}_k(t) \in \mathbb{R}^{N_k \times 1}$ is the independent and identically distributed (i.i.d.) real Additive White Gaussian Noise (AWGN) at Receiver k, each entry of which is an i.i.d. Gaussian random variable with zero-mean and unit-variance. $\mathbf{H}_{k1}^m(t) \in \mathbb{R}^{N_k \times M_1}$ denotes the m^{th} mode of the $N_k \times M_1$ channel matrix between Tx1 and Receiver k at time slot t, where $m \in \mathcal{M} = \{1, ..., N_1\}$. $\mathbf{H}_{22}(t) \in \mathbb{R}^{N_2 \times M_2}$ is the channel matrix from Tx2 to Receiver 2 (Rx2). To model partial CSIT, we assume that

$$\mathbf{H}_{12}(t) = \hat{\mathbf{H}}_{12}(t) + \sqrt{P^{-\beta}}\tilde{\mathbf{H}}_{12}(t) \quad (3)$$

where $\hat{\mathbf{H}}_{12}(t)$ is the channel estimate which is known to Tx2, while $\tilde{\mathbf{H}}_{12}(t)$ is the estimation error matrix which is unknown to Tx2. We assume that the coefficients of $\hat{\mathbf{H}}_{12}(t)$, $\tilde{\mathbf{H}}_{12}(t)$ and all other channel matrices are drawn from continuous distributions, independent of each other with zero-mean and unit-variance. The parameter β measures the quality of the channel estimate. The value β=0 corresponds to the case of no CSIT, while β=1 corresponds to perfect CSIT. It is also assumed that the coherence time T_c is long enough so that the channels stay constant across each block of N_1 time slots.

Codebooks, probability of error, achievable rates (R_1, R_2), capacity region $C(P)$, and DoF region $\mathcal{D}_{\alpha\beta}$ are all defined in the standard Shannon theoretic sense.

Finally, the antenna configuration (M_1, M_2, N_1, N_2) is said to belong to **Case I**, **II**, **III** or **IV** based on the following criteria.

1) **Case I** (Reconfigurable antennas are useful): $\mathcal{D}_{00} \neq \mathcal{D}_{10}$.
2) **Case II** (CSIT is useful): $\mathcal{D}_{00} \neq \mathcal{D}_{01}$.
3) **Case III** (Both are useless): $\mathcal{D}_{00} = \mathcal{D}_{11}$.
4) **Case IV** (Both are useful): Intersection of Cases I and II.

III. Known Results

From [17], it follows that the DoF region, denoted with $\widehat{\mathcal{D}}_{01}$, with perfect CSIT at *both* transmitters is,

$$\{(d_1, d_2) \in \mathbb{R}_+^2 : d_1 \le \min(M_1, N_1), d_2 \le \min(M_2, N_2), \quad (4)$$
$$d_1 + d_2 \le \min(M_1 + M_2, N_1 + N_2, \max(M_1, N_2), \max(M_2, N_1)\}.$$

Following the proof in [17], it is evident that the region remains the same with or without reconfigurable antennas, regardless of the length of coherence time. Note that, $\widehat{\mathcal{D}}_{01}$ is different from $\mathcal{D}_{01}$ as the latter allows CSIT only at Tx2. Obviously, $\mathcal{D}_{01} \subset \widehat{\mathcal{D}}_{01}$, but $\mathcal{D}_{01}$ is not known from prior work. In fact, while various inner bounds (i.e., achievable DoF) are presented for $\mathcal{D}_{0\beta}$ in [21], [22] for all $\beta \in [0,1]$, and a matching outer bound is presented in [23] for coherence time $T_c = 1$, the characterization of optimal $\mathcal{D}_{0\beta}$ for larger coherence times in general remains a challenging open problem due to the difficulty of extending outer bounds to larger coherence intervals [18].

From [24], [25], it follows that with no CSIT at any transmitter and no reconfigurable antennas, regardless of the network coherence time, the DoF region $\mathcal{D}_{00}$ is,

$$\mathcal{D}_{00} = \{(d_1, d_2) \in \mathbb{R}_+^2 : d_1 \le \min(M_1, N_1), d_2 \le \min(M_2, N_2),$$
$$d_1 + \frac{\min(N_1, M_2) - L}{\min(N_2, M_2) - L}(d_2 - L) \le \min(M_1, N_1)\}, \quad (5)$$

where $L = \min(M_1 + M_2, N_1) - \min(M_1, N_1)$.

From [21], [22], [23], it follows that with partial CSIT at only one transmitter and no reconfigurable antennas, the achievable DoF region $\underline{\mathcal{D}}_{0\beta}$ is equal to $\mathcal{D}_{00}$ if $N_1 > \min(M_2, N_2)$. If $N_1 \le \min(M_2, N_2)$, then we have two subcases. First, if $N_1 \le \min(M_2, N_2)$ and $N_1 + M_1 < N_2$, then,

$$\underline{\mathcal{D}}_{0\beta} = \{(d_1, d_2) \in \mathbb{R}^{2+} : (8), (9), (10), (11)\}, \quad (6)$$

and, second, if $N_1 \le \min(M_2, N_2)$ and $N_1 + M_1 \ge N_2$, then,

$$\underline{\mathcal{D}}_{0\beta} = \{(d_1, d_2) \in \mathbb{R}^{2+} : (8), (9), (10), (12), (13)\}, \quad (7)$$

where the conditions (8)-(13) are defined as follows,

$$d_1 \le \min(M_1, N_1), \tag{8}$$

$$d_1 + d_2 \le \min(M_2, N_2), \tag{9}$$

$$\frac{d_1}{\min(M_1, N_1)} + \frac{d_2}{\min(M_2, N_2) - (N_1 - M_1)^+} \le \frac{\min(M_2, N_2) + (M_2 - N_1)\beta}{\min(M_2, N_2) - (N_1 - M_1)^+}, \tag{10}$$

$$\frac{d_1}{\min(M_1, N_1)} + \frac{d_2}{N_2 - \max(M_1, N_1) + \min(M_1, N_1)} \le \frac{N_2 + (N_2 - M_1)\beta}{N_2 - \max(M_1, N_1) + \min(M_1, N_1)}, \tag{11}$$

$$d_1 + \frac{d_2}{2} \le \frac{1}{2}[M_1 + N_1 + (N_2 - M_1)\beta], \tag{12}$$

$$\frac{d_1}{M_1} + \frac{d_2}{N_2+M_1-N_1} \le \frac{N_2}{N_2+M_1-N_1} + \left[\frac{N_2 - M_1}{N_2+M_1-N_1} + \frac{(M_1+N_1-N_2)\min(M_1, M_1+M_2-N_1-N_2)}{M_1(N_2 + M_1 - N_1)}\right]\beta. \tag{13}$$

Lastly, while the DoF region with reconfigurable antennas and no CSIT is not solved completely, it follows from [16] that when $N_1 < \min(M_2, N_2)$, then,

$$\mathcal{D}_{10} = \{(d_1, d_2) \in \mathbb{R}_+^2 : d_1 \le \min(M_1, N_1), d_2 \le \min(M_2, N_2), \quad d_1 + \frac{\min(N_1, N_2, M_2)}{\min(N_2, M_2)} d_2 \le \min(M_1 + M_2, N_1) \quad \frac{\min(N_1, N_2, M_1)}{\min(N_1, M_1)} d_1 + d_2 \le \min(M_1 + M_2, N_2)\}. \tag{14}$$

Note that, although [16] allows reconfigurable antennas at either transmitter and we allow it only at Tx1, this region still holds because the achievable scheme in [16] does not require reconfigurable antennas at Tx2 in this case.

IV. Main Results

Our first result identifies antenna configurations for which both CSIT and reconfigurable antennas are individually useful.

Theorem 1: $(M_1, M_2, N_1, N_2) \in$ **Case IV** if and only if $M_1 < N_1 < \min(M_2, N_2)$.

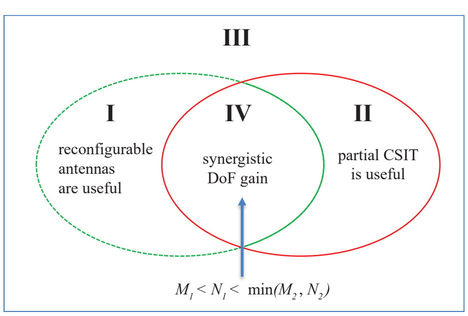

Fig. 1: DoF benefits of reconfigurable antennas and partial CSIT.

Theorem 1 is proved in Section V. The following theorem shows that synergistic DoF gain exists for every antenna configuration in **Case IV**.

Theorem 2: Let $(M_1, M_2, N_1, N_2) \in$ **Case IV**, and

$$(d_1^*, d_2^*) \triangleq \left(M_1, \frac{(N_1-M_1)\min(M_2,N_2)}{N_1} + \beta\frac{M_1(\min(M_2,N_2)-N_1)}{N_1}\right), \tag{15}$$

then $\exists \beta \in (0,1)$ such that $(d_1^*, d_2^*) \in \mathcal{D}_{1\beta}$, but $(d_1^*, d_2^*) \notin \text{co}(\mathcal{D}_{10} \cup \underline{\mathcal{D}}_{0\beta})$.

The key idea of the achievable scheme used in Theorem 2 is presented first in a toy example in Section VI-A. A complete proof of Theorem 2 is presented in Section VI-B.

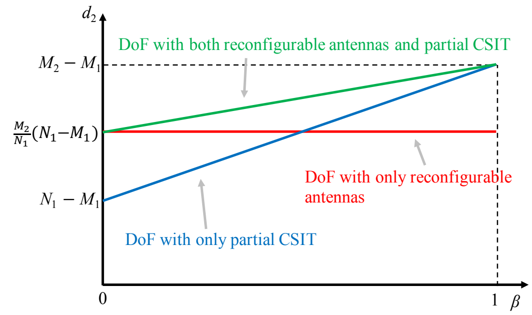

Fig. 2: User 2's DoF d_2 when $d_1 = M_1$ and $M_1 < N_1 < M_2 \le N_2$.

Remark 1: Consider d_2 to be the DoF achieved by User 2 while User 1 achieves its maximum possible DoF, $d_1 = M_1$. With no CSIT and no reconfigurable antennas ($\alpha = \beta = 0$), [25] shows that the optimal value of d_2 is $N_1 - M_1$. By using reconfigurable antennas alone, i.e., without any other CSIT, [16] shows that the optimal value of d_2 is $\frac{M_2(N_1-M_1)}{N_1}$, hence reconfigurable antennas are useful. With partial CSIT and no reconfigurable antennas, as a function of β, the best known achievable DoF value is $d_2 = N_1 - M_1 + (M_2 - N_1)\beta$ from [21], [22], which is better than with no CSIT meaning that partial CSIT is also useful. By combining reconfigurable antennas and partial CSIT, User 2 can achieve $\frac{(N_1-M_1)M_2}{N_1} + \beta\frac{M_1(M_2-N_1)}{N_1}$. The synergistic gain is illustrated in Fig. 2 for the case $M_1 < N_1 < M_2 \le N_2$. The synergistic DoF gain for the case of $M_1 < N_1 < N_2 < M_2$ is similarly verified.

V. Proof of Theorem 1

To prove Theorem 1, we first prove that $N_1 < \min(M_2, N_2)$ is the condition for **Case II**, i.e., partial CSIT is useful regardless of the network coherence time. Then, it is sufficient to prove that $M_1 < N_1 < \min(M_2, N_2)$ is the only subcase where reconfigurable antennas are useful in **Case II**.

First, for the 2-user IC with partial CSIT at Tx2, and without reconfigurable antennas, one can classify such a channel into two cases according to the antenna configuration: 1) $N_1 < N_2$ and 2) $N_1 \ge N_2$.

Case 1 can be further divided into two subcases: a) $M_2 \le N_1 < N_2$, b) $N_1 < \min(M_2, N_2)$. Any CSIT is not useful for subcase 1.a, since the DoF region with perfect CSIT is the same as with no CSIT. For subcase 1.b, partial CSIT at Tx2 is useful [21], [22].

Case 2 can be further divided into three subcases: a) $M_1 \le N_2 \le N_1$, b) $N_2 < M_1$ and $N_2 \le \min(M_2, N_1)$, c) $N_2 \le N_1$ and $M_2 < N_2 < M_1$. Any CSIT is not useful for subcase 2.a, since the DoF region with perfect CSIT is the same with no

CSIT For subcase 2.b, partial CSIT at Tx2 is not useful, since the bound for no CSIT in [24] still holds here as long as Tx1 has no CSIT. For subcase 2.c, partial CSIT at Tx2 is also not useful, since any 2-user channel with partial CSIT only at Tx2 and $M_2 < \min(N_1, N_2)$ is equivalent to a channel without any CSIT.

The only subcase where partial CSIT at Tx2 is useful, is subcase 1.b. In subcase 1.b, if there is no CSIT anywhere, and Tx1 has reconfigurable antennas, then reconfigurable antennas are useful only when $M_1 \geq N_1$ according to [16].

Therefore, both reconfigurable antennas and partial CSIT are useful if and only if $M_1 < N_1 < \min(M_2, N_2)$.

VI. Proof of Theorem 2

The most interesting aspect of the reconfigurable antennas and partial CSIT problem is the synergistic DoF gain. First, a representative toy example of this phenomenon is presented.

A. *Example 1:* $(M_1, M_2, N_1, N_2) = (1, 3, 2, 3)$

As a special case of Theorem 2, let us start with the setting in Fig. 3, $(M_1, M_2, N_1, N_2) = (1, 3, 2, 3)$, where User 1 has a reconfigurable antenna and achieves $d_1 = 1$, i.e., its maximum DoF. Suppose Tx2 has partial CSIT level of $\beta = \frac{1}{4}$ for its interference carrying link to Rx1. As shown in Fig. 3, User 2 can achieve 1.625 DoF without hurting User 1.

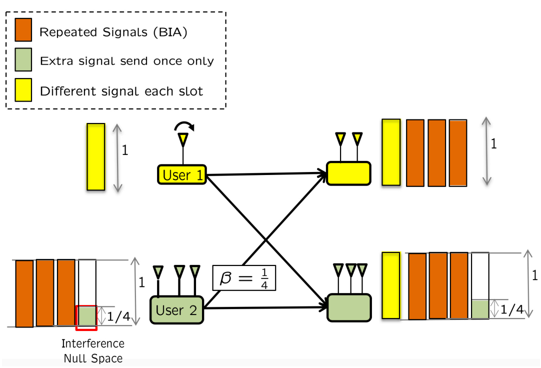

Fig. 3: Illustration of the achievable scheme for Example 1.

1) Employing the original BIA scheme with reconfigurable antennas: We first introduce the scheme based on BIA that can achieve $d_2 = 1.5$ with reconfigurable antennas alone [16]. It operates over two channel uses. During the first channel use, User 1 sends 1 symbol, User 2 sends 3 symbols (signal subspace denoted by orange color in Fig. 3). Rx1 sees as many linear combinations of these symbols as the number of receive antennas. During the second channel use, User 1 switches its antenna mode and sends a new symbol, while User 2 repeats the same 3 symbols. The repetition of symbols aligns interference at Rx1, since Rx1 is able to subtract the output of one channel use from the other to eliminate interference, leaving it with 2 interference free observations of the 2 desired symbols which can therefore be resolved. The reconfigurable antenna at Tx1 is important to make sure that the 2 desired symbols do not align with each other. User 2 has enough receive antennas to decode all symbols. Thus, User 1 achieves 1 DoF and User 2 achieves 1.5 DoF.

2) Sending extra symbols with partial CSIT: Tx2 can send an additional stream X_{2b} with $\frac{1}{4}$ DoF (signal subspace denoted by green color in Fig. 3) at first channel use with partial CSIT. This stream is not heard by Rx1, i.e., it is transmitted through the null-space of the estimated channel $\hat{\mathbf{H}}_{12}$, in which Tx2 must not exceed the power level $P^{\frac{1}{4}}$, so that User 1 still achieves 1 DoF.

Meanwhile X_{2b} can be decoded by Rx2. This is possible because Rx2 can first zero-force X_{2b} at each time slot, then Rx2 is equivalent to a receiver with 2 antennas. If Rx1 can decode its desired message, this equivalent receiver with the same number of antennas as Rx1 and statistically equivalent channel must also be able to decode the same message. By subtracting the decoded message sent by Tx1 from the received signal at Rx2, and then by subtracting the remaining received signal of one channel use from the other, the repeated streams are eliminated, and X_{2b} can be decoded. At the end, by subtracting X_{2b} from the received signal at Rx2 all the three repeated streams can be decoded. Thus, User 2 achieves 3.25 DoF over two channel uses, and $d_2 = 1.625$ per channel use without hurting User 1.

B. *General Proof*

We now prove the achievable DoF in (15). Consider N_1 time slots. During these N_1 time slots, Tx1 switches its antenna mode each time to go through all N_1 modes. Tx2 sends messages through the same channel during N_1 time slots. It is sufficient to first show that for a channel with $M_2 = N_2 = M$, $(N_1 M_1, (N_1 - M_1)M + \beta M_1(M - N_1))$ DoF can be achieved in N_1 time slots. Then, the scheme for the channel with arbitrary M_2 and N_2 is the same as that for the $(M_1, N_1, \min(M_2, N_2), \min(M_2, N_2))$ channel.

1) Employing the original BIA scheme in [16]: The only difference from Example 1 is that for the general scheme, Tx2 does not simply repeat the transmitted signals. Instead, at each time slot, Tx2 transmits a different linear combination of the overall transmitted signals. Rx1 is able to eliminate the interference by performing a linear transformation among received signals over all the time slots.

The overall transmitted signals in N_1 time slots are $\mathbf{X}_1 = \overline{P}\widehat{\mathbf{x}}_1$ and $\mathbf{X}_2 = \overline{P}\widehat{\mathbf{V}}^1\widehat{\mathbf{x}}_{21}$. Here, $\overline{P} = \sqrt{P}$, $\widehat{\mathbf{x}}_1 = [x_1^{[1]}, x_1^{[2]}, ..., x_1^{[N_1 M_1]}]^T$, and $\widehat{\mathbf{x}}_{21} = [x_{21}^{[1]}, x_{21}^{[2]}, ..., x_{21}^{[(N_1-M_1)M]}]^T$. $x_1^{[i]}$ and $x_{21}^{[j]}$ are independent Gaussian codewords from unit power codebooks. Each $x_1^{[i]}$ and $x_{21}^{[j]}$ carries 1 DoF. $\widehat{\mathbf{V}}^1 = \mathbf{V}^1_{N_1 \times (N_1 - M_1)} \otimes \mathbf{I}_M$ is the same as in [16]. Rx1 chooses $\widehat{\mathbf{Q}}_1 = \mathbf{Q}_{M_1 \times N_1} \otimes \mathbf{I}_{N_1}$ to eliminate interference such that $\mathbf{QV}^1 = \mathbf{0}$ [16].

In Example 1, $\widehat{\mathbf{V}}^1 = [1, 1]^T \otimes \mathbf{I}_3$ and $\widehat{\mathbf{Q}} = [1, -1] \otimes \mathbf{I}_2$.

2) Sending extra symbols with partial CSIT: Now, based on the original scheme, $\widehat{\mathbf{x}}_{22}$ is transmitted through the null space of the estimated channel matrix.

$$\mathbf{X}_1 = \overline{P}\widehat{\mathbf{x}}_1 \tag{16}$$

$$\mathbf{X}_2 = \overline{P}\widehat{\mathbf{V}}^1\widehat{\mathbf{x}}_{21} + \overline{P}^{\beta}\widehat{\mathbf{V}}^2\widehat{\mathbf{x}}_{22} \tag{17}$$

Here, $\widehat{\mathbf{x}}_{22} = [x_{22}^{[1]}, x_{22}^{[2]}, \ldots, x_{22}^{M_1(M-N_1)}]^T$ is independent Gaussian codewords from unit power codebooks. Each $x_{22}^{[j]}$ carries β DoF.

$\widehat{\mathbf{V}}^2$ is the partial ZF matrix for $\widehat{\mathbf{x}}_{22}$ at Tx2 with following structures,

$$\widehat{\mathbf{V}}^2_{MN_1\times M_1(M-N_1)} = \begin{bmatrix}\mathbf{I}_{M_1}\\ \mathbf{0}_{(N_1-M_1)\times M_1}\end{bmatrix} \otimes \mathbf{V}_{M\times(M-N_1)} \tag{18}$$

$$= \mathbf{I}' \otimes \mathbf{V}, \tag{19}$$

where $\mathbf{V}_{M\times(M-N_1)}$ is chosen from the right null space of $\hat{\mathbf{H}}_{12}$, i.e., $\hat{\mathbf{H}}_{12}\mathbf{V} = \mathbf{0}$.

3) Decodability: Note that, $\widehat{\mathbf{x}}_{22}$ is transmitted with power level $\sim P^{\beta}$, it is received at Rx1 below the noise floor. Thus, Rx1 is equivalent to a receiver in [16] with only reconfigurable antennas at transmitter. Therefore, Rx1 is able to remove the interference term caused by $\widehat{\mathbf{x}}_{21}$ through BIA, and decode $\widehat{\mathbf{x}}_1$.

Considering Rx2, we first show that it can decode $\widehat{\mathbf{x}}_1$. Rx2 first zero-forces $\widehat{\mathbf{x}}_{22}$ at each time slot with a receiving ZF matrix, $\mathbf{R}_{(M-N_1)\times N_2}$. $\mathbf{R}$ is chosen such that $\mathbf{R}\mathbf{H}_{22}\mathbf{V} = \mathbf{0}$. Now, Rx2 is equivalent to a receiver in [16] with N_1 antennas. Since this equivalent receiver has the same number of antennas as Rx1 and their channel are statistically equivalent, if Rx1 can decode $\widehat{\mathbf{x}}_1$, then Rx2 must also be able to decode the same messages.

Then, by subtracting the decoded message $\widehat{\mathbf{x}}_1$, the remaining received signals at Rx2 are desired signals $\widehat{\mathbf{x}}_{21}$ and $\widehat{\mathbf{x}}_{22}$. After applying nulling matrix $\widehat{\mathbf{Q}}_2$ for $\widehat{\mathbf{x}}_{21}$, $\widehat{\mathbf{Q}}_2 = \mathbf{Q} \otimes \mathbf{I}_M$, the overall received signals at Rx2 is,

$$\mathbf{Y}_2' = \overline{P}\widehat{\mathbf{Q}}_2\widehat{\mathbf{H}}_{22}\widehat{\mathbf{V}}^1\widehat{\mathbf{x}}_{21} + \overline{P}^{\beta}\widehat{\mathbf{Q}}_2\widehat{\mathbf{H}}_{22}\widehat{\mathbf{V}}^2\widehat{\mathbf{x}}_{22} + \widehat{\mathbf{Q}}_2\mathbf{z}_2, \tag{20}$$

where $\widehat{\mathbf{H}}_{22} = \mathbf{I}_{N_1} \otimes \mathbf{H}_{22}$. To ensure Rx2 can decode $\widehat{\mathbf{x}}_{22}$, the following matrix $\mathbf{R}$ needs to be full rank,

$$\begin{aligned}\mathbf{R} &= \widehat{\mathbf{Q}}_2\widehat{\mathbf{H}}_{22}\widehat{\mathbf{V}}^2\\ &= (\mathbf{Q}\otimes\mathbf{I}_M)(\mathbf{I}_{N_1}\otimes\mathbf{H}_{22})(\mathbf{I}'\otimes\mathbf{V})\\ &= (\mathbf{Q}\mathbf{I}')\otimes(\mathbf{H}_{22}\mathbf{V}).\end{aligned} \tag{21}$$

Since $\mathbf{Q}\mathbf{I}'$ and $\mathbf{H}_{22}\mathbf{V}$ have rank M_1 and $M - N_1$, respectively, and $\mathbf{R}$ has rank $M_1(M-N_1)$, $\widehat{\mathbf{x}}_{22}$ can be decoded and subtracted from the received signals at Rx2. At last, Rx2 can decode $\widehat{\mathbf{x}}_{21}$ because $\widehat{\mathbf{H}}_{22}\widehat{\mathbf{V}}^1$ is full rank as in [16].This concludes the proof.

VII. Non-trivial scheme to combine BIA, rate-splitting and partial ZF

The previous section shows that BIA and partial ZF schemes can be combined in general to obtain a synergistic DoF gain. By incorporating rate-splitting along with BIA and partial ZF, the synergistic DoF gain can be further improved under the setting $M_1 < N_1 < N_2 < M_2$. In this section, we illustrate how these ideas can be combined through two challenging examples.

A. Example 2: $(M_1, M_2, N_1, N_2) = (1, 4, 2, 3)$

Consider the setting $(M_1, M_2, N_1, N_2) = (1, 4, 2, 3)$ as shown in Fig. 4. Suppose that Tx1 achieves its maximum possible DoF, i.e., $d_1 = \min(M_1, N_1) = 1$ in this case. Then, how many DoF can Tx2 achieve simultaneously without

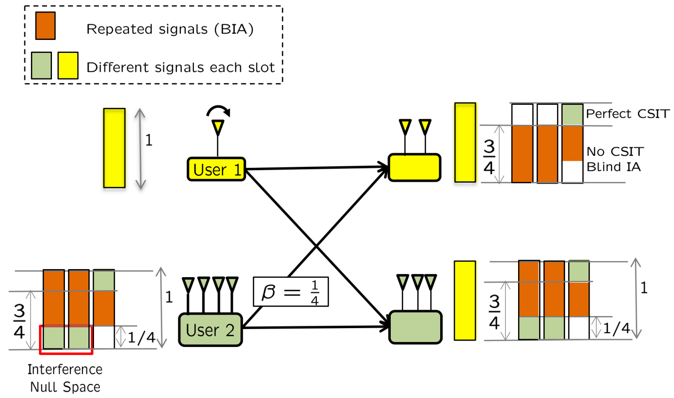

Fig. 4: Illustration of the achievable scheme for Case 1.

hurting User 1? If we ignore the reconfiguring capability of the antenna at Tx1, then as shown in [21], [22], based on partial CSIT alone, it is possible for Tx2 to achieve $d_2 = 1.5$ DoF without hurting User 1. On the other hand, if we ignore partial CSIT at Tx2, then by exploiting only the reconfigurable antenna at Tx1, BIA schemes also achieve $d_2 = 1.5$ without hurting Tx1, as shown in [14]. Thus, either partial CSIT or BIA can individually achieve $(d_1, d_2) = (1, 1.5)$. According to Theorem 2, by jointly exploiting both BIA and partial ZF, one can achieve $(d_1, d_2) = (1, 1.625)$. Is it possible to achieve even more? Indeed as it turns out, it is possible to achieve $(d_1, d_2) = (1, 1.75)$ by further exploiting rate-splitting schemes. The achievable scheme uses a combination of rate-splitting, partial ZF and BIA as illustrated in Fig. 4, and is explained next.

1) Designing transmitted signals: The scheme involves applying precoding over two channel uses. The transmitted signals over the first channel use are,

$$\mathbf{x}_1 = c_1(\overline{P}X_{11} + \overline{P}^{\frac{3}{4}}X_{12} + \overline{P}^{\frac{1}{4}}X_{13}), \tag{22}$$

$$\mathbf{x}_2 = c_2(\mathbf{v}_{21}X_{21} + \mathbf{v}_{22}X_{22} + \mathbf{v}_{23}X_{23}), \tag{23}$$

where X_{1i} are independent Gaussian codewords from unit power codebooks, and c_i are bounded, i.e., $O(1)$, normalization factors needed to satisfy the transmit power constraint (e.g., $c_1 = \overline{P}/(\sqrt{P + P^{3/4} + P^{1/4}}) \in [1/3, 1]$) that are inconsequential for DoF, and henceforth omitted for compact notation. X_{11} and X_{13} each carries $\frac{1}{4}$ DoF, X_{12} carries $\frac{1}{2}$ DoF. Each of the precoding vectors $\mathbf{v}_{2j}$ at Tx2 has size 4×1. $\mathbf{v}_{21}$ and $\mathbf{v}_{22}$ are chosen from the right null space of $\hat{\mathbf{H}}_{12}$, i.e., $\hat{\mathbf{H}}_{12}[\mathbf{v}_{21}\ \mathbf{v}_{22}] = \mathbf{0}$. $\mathbf{v}_{23}$ is generic, e.g., the all ones vector. X_{2i} has a layered structure as follows,

$$X_{21} = \overline{P}^{\frac{3}{4}}X_{21}^{b} + \overline{P}^{\frac{1}{4}}X_{21}^{p} + \overline{P}X_{24}^{b}, \tag{24}$$

$$X_{22} = \overline{P}^{\frac{3}{4}}X_{22}^{b} + \overline{P}^{\frac{1}{4}}X_{22}^{p} + \overline{P}X_{25}^{b}, \tag{25}$$

$$X_{23} = \overline{P}^{\frac{3}{4}}X_{23}^{b} + \overline{P}X_{23}^{p}, \tag{26}$$

where X_{2j}^{b} and X_{2i}^{p} are all independent Gaussian codewords from unit power codebooks, with the difference that X_{2j}^{b} (used for BIA) is repeated by Tx2 at the second channel use. X_{21}^{b},

X_{22}^b and X_{23}^b each carries $\frac{1}{2}$ DoF and all other codewords each carries $\frac{1}{4}$ DoF.

Now the received signals are,

$$
\begin{aligned}
&\mathbf{y}_1(1) = \\
&\mathbf{H}_{11}^1\mathbf{x}_1 + \bar{P}^{-\frac{1}{4}}\tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{21} + \mathbf{v}_{22}X_{22}) + \mathbf{H}_{12}\mathbf{v}_{23}X_{23} + \mathbf{z}_1, \\
&= \bar{P}[\mathbf{H}_{11}^1 X_{11} + \mathbf{H}_{12}\mathbf{v}_{23}X_{23}^p] \\
&+ \bar{P}^{\frac{3}{4}}[\mathbf{H}_{11}^1 X_{12} + \tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{24}^b + \mathbf{v}_{22}X_{25}^b) + \mathbf{H}_{12}\mathbf{v}_{23}X_{23}^b] \\
&+ \bar{P}^{\frac{1}{2}}[\tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{21}^b + \mathbf{v}_{22}X_{22}^b)] \\
&+ \bar{P}^{\frac{1}{4}}\mathbf{H}_{11}^1 X_{13} + \tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{21}^p + \mathbf{v}_{22}X_{22}^p) + \mathbf{z}_1, \qquad (27)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{y}_2(1) &= \mathbf{H}_{21}^1\mathbf{x}_1 + \mathbf{H}_{22}(\mathbf{v}_{21}X_{21} + \mathbf{v}_{22}X_{22} + \mathbf{v}_{23}X_{23}) + \mathbf{z}_2, \\
&= \bar{P}[\mathbf{H}_{21}^1 X_{11} + \mathbf{H}_{22}(\mathbf{v}_{21}X_{24}^b + \mathbf{v}_{22}X_{25}^b + \mathbf{v}_{23}X_{23}^p)] \\
&+ \bar{P}^{\frac{3}{4}}[\mathbf{H}_{21}^1 X_{12} + \mathbf{H}_{22}(\mathbf{v}_{21}X_{21}^b + \mathbf{v}_{22}X_{22}^b + \mathbf{v}_{23}X_{23}^b)] \\
&+ \bar{P}^{\frac{1}{4}}[\mathbf{H}_{21}^1 X_{13} + \mathbf{H}_{22}(\mathbf{v}_{21}X_{21}^p + \mathbf{v}_{22}X_{22}^p)] + \mathbf{z}_2, \qquad (28)
\end{aligned}
$$

At the second channel use, Tx1 switches its antenna mode. Tx1 and Tx2 send new codewords, X_{1i}' and $X_{2i}'^p$, with the same power levels, precoding vectors and DoF values as X_{1i} and X_{2i}^p in the first channel use. Tx2 also repeats codewords X_{2j}^b with the same power levels and precoding vectors as in the first channel use. Then, the received signals are,

$$
\begin{aligned}
\mathbf{y}_1(2) &= \bar{P}[\mathbf{H}_{11}^2 X_{11}' + \mathbf{H}_{12}\mathbf{v}_{23}X_{23}'^p] \\
&+ \bar{P}^{\frac{3}{4}}[\mathbf{H}_{11}^2 X_{12}' + \tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{24}^b + \mathbf{v}_{22}X_{25}^b) + \mathbf{H}_{12}\mathbf{v}_{23}X_{23}^b] \\
&+ \bar{P}^{\frac{1}{2}}[\tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{21}^b + \mathbf{v}_{22}X_{22}^b)] \\
&+ \bar{P}^{\frac{1}{4}}\mathbf{H}_{11}^2 X_{13}' + \tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{21}'^p + \mathbf{v}_{22}X_{22}'^p) + \mathbf{z}_1', \qquad (29) \\
\mathbf{y}_2(2) &= \bar{P}[\mathbf{H}_{21}^2 X_{11}' + \mathbf{H}_{22}(\mathbf{v}_{21}X_{24}^b + \mathbf{v}_{22}X_{25}^b + \mathbf{v}_{23}X_{23}'^p)] \\
&+ \bar{P}^{\frac{3}{4}}[\mathbf{H}_{21}^2 X_{12}' + \mathbf{H}_{22}(\mathbf{v}_{21}X_{21}^b + \mathbf{v}_{22}X_{22}^b + \mathbf{v}_{23}X_{23}^b)] \\
&+ \bar{P}^{\frac{1}{4}}[\mathbf{H}_{21}^2 X_{13}' + \mathbf{H}_{22}(\mathbf{v}_{21}X_{21}'^p + \mathbf{v}_{22}X_{22}'^p)] + \mathbf{z}_2'. \qquad (30)
\end{aligned}
$$

2) Decoding at Rx1: Rx1 can first decode X_{11} and X_{23}^p from $\mathbf{y}_1(1)$ by treating all other signals as noise. This is possible because a) Rx1 has two antennas to separate the two streams, b) X_{11} and X_{23}^p are each received at Rx1 with power level $\sim P$, each carrying $\frac{1}{4}$ DoF. The equivalent noise floor due to all other signals is $\sim P^{\frac{3}{4}}$ (refer to the top $\frac{1}{4}$ level at Rx1 in Fig. 4). Once X_{11} and X_{23}^p are decoded, Rx1 can subtract them from $\mathbf{y}_1(1)$. Note that, Rx1 can also decode X_{11}' and $X_{23}'^p$ in a same manner and subtract them from $\mathbf{y}_1(2)$. The remaining received signals at Rx1 are,

$$
\begin{aligned}
\mathbf{y}_1'(1) &= \bar{P}^{\frac{3}{4}}[\mathbf{H}_{11}^1 X_{12} + \tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{24}^b + \mathbf{v}_{22}X_{25}^b) + \mathbf{H}_{12}\mathbf{v}_{23}X_{23}^b] \\
&+ \bar{P}^{\frac{1}{2}}[\tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{21}^b + \mathbf{v}_{22}X_{22}^b)] + \bar{P}^{\frac{1}{4}}\mathbf{H}_{11}^1 X_{13} \\
&+ \tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{21}^p + \mathbf{v}_{22}X_{22}^p) + \mathbf{z}_1, \qquad (31)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{y}_1'(2) &= \bar{P}^{\frac{3}{4}}[\mathbf{H}_{11}^2 X_{12}' + \tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{24}^b + \mathbf{v}_{22}X_{25}^b) + \mathbf{H}_{12}\mathbf{v}_{23}X_{23}^b] \\
&+ \bar{P}^{\frac{1}{2}}[\tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{21}^b + \mathbf{v}_{22}X_{22}^b)] + \bar{P}^{\frac{1}{4}}\mathbf{H}_{11}^2 X_{13}' \\
&+ \tilde{\mathbf{H}}_{12}(\mathbf{v}_{21}X_{21}'^p + \mathbf{v}_{22}X_{22}'^p) + \mathbf{z}_1'. \qquad (32)
\end{aligned}
$$

Note that, Rx1 observes the same linear combinations of X_{2j}^b over both channel uses, and X_{21}^p, X_{22}^p, $X_{21}'^p$ and $X_{22}'^p$ are all received at the noise floor level (i.e., with constant power level P^0). Therefore, Rx2 can cancel the interference, and thus decode X_{12}, X_{13}, X_{12}' and X_{13}' from,

$$
\begin{aligned}
&\mathbf{y}_1'(1) - \mathbf{y}_1'(2) = \\
&\bar{P}^{\frac{3}{4}}[\mathbf{H}_{11}^1 X_{12} - \mathbf{H}_{11}^2 X_{12}'] + \bar{P}^{\frac{1}{4}}[\mathbf{H}_{11}^1 X_{13} - \mathbf{H}_{11}^2 X_{13}'] + \mathbf{z}, \qquad (33)
\end{aligned}
$$

where $\mathbf{z}$ is the abbreviation for all the noise and interference terms with power level $\sim P^0$. Thus, Rx1 achieves a total of 2 DoF in two channel uses, and $d_1 = 1$ per channel use.

3) Decoding at Rx2: Rx2 is able to decode all the signals transmitted from both transmitters. Specifically, the received signals at Rx2 have a three-layered structure. X_{11}, X_{11}', X_{24}^b, X_{25}^b, X_{23}^p and $X_{23}'^p$ are received at Rx2 with power level $\sim P$, and each carries $\frac{1}{4}$ DoF. They are the top layer of $\mathbf{y}_2(1)$ and $\mathbf{y}_2(2)$ (refer to the top $\frac{1}{4}$ level at Rx2 in Fig. 4). $[X_{11}\ X_{11}'\ X_{24}^b\ X_{25}^b\ X_{23}^p\ X_{23}'^p]^T$ can be jointly decoded by Rx2 treating all other signals as noise. This is possible because the equivalent noise floor is $\sim P^{\frac{3}{4}}$, and the following overall 6×6 matrix across two channel uses has full rank almost surely,

$$
\begin{bmatrix}
\mathbf{H}_{21}^1 & \mathbf{0} & \mathbf{H}_{22}\mathbf{v}_{21} & \mathbf{H}_{22}\mathbf{v}_{22} & \mathbf{H}_{22}\mathbf{v}_{23} & \mathbf{0} \\
\mathbf{0} & \mathbf{H}_{21}^2 & \mathbf{H}_{22}\mathbf{v}_{21} & \mathbf{H}_{22}\mathbf{v}_{22} & \mathbf{0} & \mathbf{H}_{22}\mathbf{v}_{23}
\end{bmatrix}. \qquad (34)
$$

This is because the determinant polynomial of this matrix in terms of the channel random variables is not identically zero, which is easily verified by substituting, for example, $\mathbf{H}_{21}^1 = [1\ \ 0\ \ 0]^T$, $\mathbf{H}_{21}^2 = [0\ \ 1\ \ 0]^T$, $\mathbf{v}_{21} = [1\ \ 0\ \ 0\ \ 0]^T$, $\mathbf{v}_{22} = [0\ \ 1\ \ 0\ \ 0]^T$, $\mathbf{v}_{23} = [0\ \ 0\ \ 1\ \ 0]^T$ and

$$
\mathbf{H}_{22} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.
$$

Since the determinant polynomial is not identically zero, it must be non-zero almost surely for random channel realizations, i.e., the matrix has full rank almost surely.

After Rx2 decodes codewords on its first layer, it subtracts them from the received signals (let $\mathbf{y}_2'(1)$ and $\mathbf{y}_2'(2)$ denote the remaining received signals), and decodes all the codewords on the middle layer (refer to the middle $\frac{1}{2}$ level at Rx2 in Fig. 4) by treating the signals on the bottom layer as white noise. Note that, all the desired codewords on the middle layer are repeated over two channel uses. Therefore, in a same manner as Rx1's bottom layer, Rx2 can first cancel the repeated codewords and decode X_{12} and X_{12}'. Then, by subtracting them from $\mathbf{y}_2'(1)$, Rx2 can decode X_{21}^b, X_{22}^b and X_{23}^b.

After subtracting the codewords decoded on the middle layer, Rx2 has enough antennas to separate all the remaining streams for each channel use, and thus, can decode the remaining codewords on the bottom layer (refer to the bottom $\frac{1}{4}$ level at Rx2 in Fig. 4). Therefore, Rx2 achieves a total of

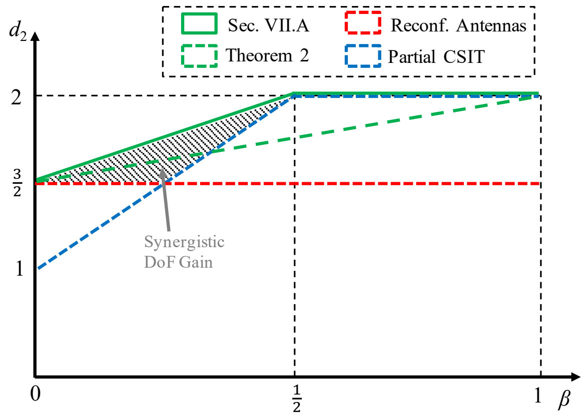

Fig. 5: Synergistic DoF gain for Case 1.

3.5 DoF over two channel uses, and $d_2 = 1.75$ per channel use.

To extend this scheme to the arbitrary partial CSIT level, i.e., $0 \leq \beta \leq 1$, we only need to adjust the power proportion of each layer in the transmitted signal (24)-(26) to $\min(0.5, \beta)$, $1-2\min(0.5, \beta)$ and $\min(0.5, \beta)$, respectively. Therefore, the DoF achieved by Tx2 as a function of β is $d_2 = 1.5 + \max(\beta, 0.5)$, as shown in Fig. 5. To emphasize the synergistic benefit, the DoF achieved with only partial CSIT and with only reconfigurable antennas are also presented in Fig. 5.

B. Example 3: $(M_1, M_2, N_1, N_2) = (1, 6, 3, 4)$

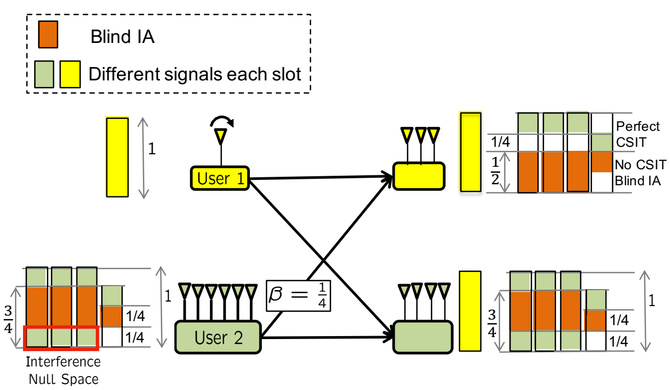

Fig. 6: Illustration of the achievable scheme for Case 2.

We now consider the setting $(M_1, M_2, N_1, N_2) = (1, 6, 3, 4)$ as illustrated in Fig. 6. In this case, a more sophisticated scheme is needed for Tx2 to design its transmitted signal as a four-layer structure. Elevated multiplexing [21] is utilized at Tx2. Furthermore, the BIA scheme at Tx2 is not merely a simple repetition of transmitted signals. The proposed scheme is as follows.

1) Designing transmitted signals: The scheme operates over three channel uses. The transmitted signals at each channel use are,

$$\mathbf{x}_1(t) = \bar{P} X_{14}(t) + \bar{P}^{\frac{3}{4}} X_{13}(t) + \bar{P}^{\frac{1}{2}} X_{12}(t) + \bar{P}^{\frac{1}{4}} X_{11}(t), \tag{35}$$

$$\begin{aligned}\mathbf{x}_2(t) =& \mathbf{v}_{21} X_{21}(t) + \mathbf{v}_{22} X_{22}(t) + \mathbf{v}_{23} X_{23}(t) + \mathbf{v}_{24} X_{24}(t) \\ &+ \mathbf{v}_{25} X_{25}(t) + \mathbf{v}_{26} X_{26}(t) + \mathbf{v}_{27} X_{27}(t),\end{aligned} \tag{36}$$

where $t \in \{1, 2, 3\}$, $X_{1i}(t)$, $i \in \{1, 2, 3, 4\}$ and $X_{2j}(t)$, $j \in \{5, 6, 7\}$ are independent Gaussian codewords from unit power codebooks, each carrying $\frac{1}{4}$ DoF. Each of the precoding vectors $\mathbf{v}_{2j}$ at Tx2 has size 6×1. $\mathbf{v}_{21}$, $\mathbf{v}_{22}$ and $\mathbf{v}_{23}$ are chosen from the right null space of $\hat{\mathbf{H}}_{12}$, i.e., $\hat{\mathbf{H}}_{12}[\mathbf{v}_{21}\ \mathbf{v}_{22}\ \mathbf{v}_{23}] = \mathbf{0}$. $\mathbf{v}_{2j}$, $j \in \{4, 5, 6, 7\}$ can be chosen as generic vectors. $X_{2i}(t), i \in \{1, 2, 3, 4\}$ has a layered structure as follows,

$$\begin{aligned}X_{2k}(t) =& \bar{P}^{\frac{3}{4}} X^b_{2k3}(t) + \bar{P}^{\frac{1}{2}} X^b_{2k2}(t) + \bar{P}^{\frac{1}{4}} X_{2k1}(t), \\ & \forall k \in \{1, 2, 3\}\end{aligned} \tag{37}$$

$$X_{24}(t) = \bar{P}^{\frac{3}{4}} X_{243}(t) + \bar{P}^{\frac{1}{2}} X^b_{242}(t). \tag{38}$$

Here, X^b_{2ij} and X_{2ij} are independent Gaussian codewords from unit power codebooks, each carrying $\frac{1}{4}$ DoF, with the difference that X^b_{2ij} sent by Tx2 at the first two channel uses are repeated at the third channel use as $X^b_{2ij}(3) = X^b_{2ij}(1) + X^b_{2ij}(2)$.

The received signals at each time slot are,

$$\begin{aligned}\mathbf{y}_1(t) =& \mathbf{H}^t_{11}\mathbf{x}_1(t) + \mathbf{H}_{12}\mathbf{v}_{24} X_{24}(t) + \mathbf{H}_{12}\mathbf{v}_{25} X_{25}(t) \\ &+ \bar{P}^{-\frac{1}{4}} \tilde{\mathbf{H}}_{12}[\mathbf{v}_{21} X_{21}(t) + \mathbf{v}_{22} X_{22}(t) + \mathbf{v}_{23} X_{23}(t)] \\ &+ \mathbf{H}_{12}\mathbf{v}_{26} X_{26}(t) + \mathbf{H}_{12}\mathbf{v}_{27} X_{27}(t) + \mathbf{z}_1(t)\end{aligned} \tag{39}$$

$$\begin{aligned}=& \mathbf{H}^t_{11}\mathbf{x}_1(t) + \mathbf{H}_{12}\mathbf{v}_{25} X_{25}(t) + \mathbf{H}_{12}\mathbf{v}_{26} X_{26}(t) + \mathbf{H}_{12}\mathbf{v}_{27} X_{27}(t) \\ &+ \bar{P}^{\frac{1}{2}} \tilde{\mathbf{H}}_{12}[\mathbf{v}_{21} X^b_{213}(t) + \mathbf{v}_{22} X^b_{223}(t) + \mathbf{v}_{23} X^b_{233}(t)] \\ &+ \bar{P}^{\frac{1}{4}} \tilde{\mathbf{H}}_{12}[\mathbf{v}_{21} X^b_{212}(t) + \mathbf{v}_{22} X^b_{222}(t) + \mathbf{v}_{23} X^b_{232}(t)] \\ &+ \tilde{\mathbf{H}}_{12}[\mathbf{v}_{21} X_{211}(t) + \mathbf{v}_{22} X_{221}(t) + \mathbf{v}_{23} X_{231}(t)] \\ &+ \bar{P}^{\frac{3}{4}} \mathbf{H}_{12}\mathbf{v}_{24} X_{243}(t) + \bar{P}^{\frac{1}{2}} \mathbf{H}_{12}\mathbf{v}_{24} X^b_{242}(t) + \mathbf{z}_1(t),\end{aligned} \tag{40}$$

$$\begin{aligned}\mathbf{y}_2(t) =& \mathbf{H}^t_{21}\mathbf{x}_1(t) + \mathbf{z}_2(t) \\ &+ \mathbf{H}_{22}\,[\mathbf{v}_{21} X_{21}(t) + \mathbf{v}_{22} X_{22}(t) + \mathbf{v}_{23} X_{23}(t) + \mathbf{v}_{24} X_{24}(t) \\ &+ \mathbf{v}_{25} X_{25}(t) + \mathbf{v}_{26} X_{26}(t) + \mathbf{v}_{27} X_{27}(t)]\,.\end{aligned} \tag{41}$$

2) Decoding at Rx1: At time slot t, Rx1 can first jointly decode $X_{14}(t)$, $X_{13}(t)$, $X_{243}(t)$ and $X_{2j}(t), j \in \{5, 6, 7\}$ from $\mathbf{y}_1(t)$ as a Multi-Access Channel (MAC) channel by treating all other signals as white noise. This is possible because $X_{2j}(t), j \in \{5, 6, 7\}$ are received at Rx1 with an elevated power level $\sim P$ and the equivalent noise floor is $\sim P^{\frac{1}{2}}$ (refer to the top $\frac{1}{2}$ level at Rx1 in Fig. 6). Once $X_{14}(t)$, $X_{13}(t)$, $X_{243}(t)$

and $X_{2j}(t), j \in \{5,6,7\}$ are decoded, Rx1 can subtract them from $\mathbf{y}_1(t)$. The remaining received signal at Rx1 is,

$$\begin{aligned}\mathbf{y}_1'(t) &= \mathbf{H}_{11}^t[\overline{P}^{\frac{1}{2}} X_{12}(t) + \overline{P}^{\frac{1}{4}} X_{11}(t)]\\ &+ \overline{P}^{\frac{1}{2}} \tilde{\mathbf{H}}_{12}[\mathbf{v}_{21}X_{213}^b(t) + \mathbf{v}_{22}X_{223}^b(t) + \mathbf{v}_{23}X_{233}^b(t)]\\ &+ \overline{P}^{\frac{1}{4}} \tilde{\mathbf{H}}_{12}[\mathbf{v}_{21}X_{212}^b(t) + \mathbf{v}_{22}X_{222}^b(t) + \mathbf{v}_{23}X_{232}^b(t)]\\ &+ \tilde{\mathbf{H}}_{12}[\mathbf{v}_{21}X_{211}(t) + \mathbf{v}_{22}X_{221}(t) + \mathbf{v}_{23}X_{231}(t)]\\ &+ \overline{P}^{\frac{1}{2}} \mathbf{H}_{12}\mathbf{v}_{24}X_{242}^b(t) + \mathbf{z}_1(t). \end{aligned} \tag{42}$$

The remaining interference at Rx1 over the three channel uses is aligned into a six-dimensional subspace (refer to the bottom $\frac{1}{2}$ level at Rx1 in Fig. 6) corresponding to the 3 receive dimensions (antennas) over the first two channel uses. This is because at the third channel use, Rx1 observes only linear combinations of the interference from the first two channel uses. Therefore, Rx1 can use the BIA approach to decode $X_{12}(t)$, $X_{11}(t)$, i.e., it can cancel the interference via the following operation,

$$\begin{aligned}\mathbf{y}_1'(3) - \mathbf{y}_1'(2) - \mathbf{y}_1'(1) &= \mathbf{H}_{11}^3[\overline{P}^{\frac{1}{2}} X_{12}(3) + \overline{P}^{\frac{1}{4}} X_{11}(3)]\\ &- \mathbf{H}_{11}^1[\overline{P}^{\frac{1}{2}} X_{12}(1) + \overline{P}^{\frac{1}{4}} X_{11}(1)]\\ &- \mathbf{H}_{11}^2[\overline{P}^{\frac{1}{2}} X_{12}(2) + \overline{P}^{\frac{1}{4}} X_{11}(2)] + \mathbf{z}, \end{aligned} \tag{43}$$

where $\mathbf{z}$ is the abbreviation for all the noise and interference terms with power level $\sim P^0$. $X_{12}(t)$, $X_{11}(t)$ can be decoded because the matrix $[\mathbf{H}_{11}^3\ -\mathbf{H}_{11}^1\ -\mathbf{H}_{11}^2]$ is full rank almost surely. Then, Rx1 achieves a total of 3 DoF over three channel uses, and $d_1 = 1$ per channel use.

3) Decoding at Rx2: Rx2 is able to decode all the signals from both transmitters. Specifically, the received signals at Rx2 have a four-layered structure. $X_{14}(t)$ and $X_{2j}(t), j \in \{5,6,7\}$ are received at Rx2 with power level $\sim P$, and each carries $\frac{1}{4}$ DoF (refer to the top $\frac{1}{4}$ level at Rx2 in Fig. 6). Rx2 can decode them at each time slot by treating all other signals as noise, since the equivalent noise floor is $\sim P^{\frac{3}{4}}$.

After Rx2 decodes codewords on its first layer, it subtracts them from the received signals, and then proceeds to decode all the codewords jointly across the three time slots on the second layer (refer to the second $\frac{1}{4}$ level at Rx2 in Fig. 6) by treating all other signals as noise. This is possible because the equivalent noise floor is $\sim P^{\frac{1}{2}}$, and similar to the previous case, it is easily verified that the following overall matrix has full rank almost surely,

$$\begin{bmatrix}\widehat{\mathbf{H}}_{21} & \widehat{\mathbf{H}}_{22}^1 & \widehat{\mathbf{H}}_{22}^2\end{bmatrix}_{12\times 12}, \tag{44}$$

where

$$\widehat{\mathbf{H}}_{21} = \begin{bmatrix}\mathbf{H}_{21}^1 & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{H}_{21}^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{H}_{21}^3\end{bmatrix}_{12\times 3}, \tag{45}$$

$$\widehat{\mathbf{H}}_{22}^1 = \begin{bmatrix}\mathbf{H}_{22}\mathbf{v}_{21} & \mathbf{H}_{22}\mathbf{v}_{22} & \mathbf{H}_{22}\mathbf{v}_{23} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{H}_{22}\mathbf{v}_{21} & \mathbf{H}_{22}\mathbf{v}_{22} & \mathbf{H}_{22}\mathbf{v}_{23}\\ \mathbf{H}_{22}\mathbf{v}_{21} & \mathbf{H}_{22}\mathbf{v}_{22} & \mathbf{H}_{22}\mathbf{v}_{23} & \mathbf{H}_{22}\mathbf{v}_{21} & \mathbf{H}_{22}\mathbf{v}_{22} & \mathbf{H}_{22}\mathbf{v}_{23}\end{bmatrix}, \tag{46}$$

$$\widehat{\mathbf{H}}_{22}^2 = \begin{bmatrix}\mathbf{H}_{22}\mathbf{v}_{24} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{H}_{22}\mathbf{v}_{24} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{H}_{22}\mathbf{v}_{24}\end{bmatrix}_{12\times 3}. \tag{47}$$

$\mathbf{0}$ is zero matrix with size 4×1 in all above.

By subtracting the codewords decoded on the second layer (let $\mathbf{y}_2'(t)$ denote the remaining received signals), Rx2 can decode the remaining codewords on the last two layers successively (refer to the bottom $\frac{1}{2}$ level at Rx2 in Fig. 6). Note that ,all the desired codewords on the third layer are repeated in the third channel use. Therefore, in a similar manner to Rx1's bottom layer, Rx2 can first cancel its desired codewords and decode $X_{12}(t)$. Then, by subtracting them from $\mathbf{y}_2'(t)$, Rx2 can decode the remaining codewords on the third layer, and later subtract them as well.

Lastly, at the bottom layer, Rx2 has enough antennas to separate all the remaining streams over each channel use. Thus, Rx2 achieves $\frac{35}{4}$ DoF over three channel uses, and $d_2 = \frac{35}{12}$ per channel use. By extending this scheme to the arbitrary partial CSIT level, i.e., $0 \le \beta \le 1$, the DoF achieved by Tx2 as a function of β is found to be $d_2 = \frac{8}{3} + \max(\beta, \frac{1}{3})$. To demonstrate the synergistic benefit, the DoF achieved with only partial CSIT and with only reconfigurable antennas are shown in Fig. 7.

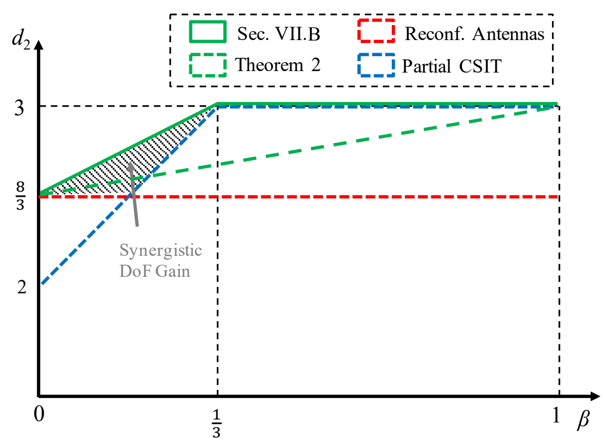

Fig. 7: Synergistic DoF gain for Case 2.

VIII. Conclusion and Discussions

In this work, the 2-user MIMO IC with reconfigurable antennas and partial CSIT was considered. A novel achievable scheme was proposed to jointly exploit both reconfigurable antennas and partial CSIT. This revealed synergistic DoF gains that cannot be seen through the study of each individual element alone. Antenna configurations that allow synergistic DoF gains were explicitly identified. We now provide some intuition behind our scheme and the condition for the synergistic DoF gain stated in Theorem 2.

Theorem 2 shows that synergistic DoF gains exist for a 2-user MIMO IC, if both reconfigurable antennas and partial CSIT are individually useful. This condition can be understood as the existence of two separate signal subspaces at Tx2, i.e., one subspace for BIA and the other for partial ZF. Particularly, for the cases where reconfigurable antennas are useful, $N_1 > M_1$ is always satisfied, i.e., Rx1 can tolerate some interference from Tx2. Thus, Tx2 can use BIA to reduce the dimension of this interference subspace at Rx1, and achieve a higher DoF compared to no CSIT. On the other hand, when partial CSIT at Tx2 is useful, there is always an interference null space from Tx2 to Rx1. Thereby, Tx2 can achieve a higher DoF compared to no CSIT by sending part of its signals through the null space so that they are not heard by Rx1.

In conclusion, we can see why BIA and partial ZF approaches can be simultaneously beneficial in a 2-user MIMO IC. BIA utilizes the signal subspace that is seen by Rx1, and partial ZF utilizes the signal subspace that is not seen by Rx1. Intuitively, if any of the two signal subspaces does not exist, these two approaches cannot be combined.

To support the claims of Theorem 2, in Fig. 8, we present numerical results in order to observe whether the DoF analysis is meaningful at finite Signal-to-Noise Ratio (SNR) despite the limitations of the first order asymptotic analysis. To obtain the numerical results, the setting given in *Example 1* is considered, $(M_1, M_2, N_1, N_2) = (1, 3, 2, 3)$. Rates of both users are evaluated numerically and averaged over many iterations, when only BIA is used, when only partial ZF is used, and when both are used simultaneously, at finite SNR and at $\beta = 0.5$. Channel realizations in this setting have i.i.d. complex Gaussian entries with zero-mean and unit-variance. Note that, the normalization factors of the input signals, which were omitted in the DoF analysis, are included here to meet the power constraints, and hence, to obtain the accurate rate results at finite SNR. To compare the rate gain of each scheme for User 2, the transmit signal power of User 1 is adjusted in each scheme to level the rate of User 1 across all schemes. Here, the advantage of the proposed scheme over using the schemes individually, and the gain of synergistic benefits can be observed at finite SNR. As stated in Theorem 2, the achievable rate for the proposed scheme increases as β increases. At $\beta = 0.5$, as it is also indicated in Fig. 2 the proposed scheme has the highest gain in the achievable rate and in DoF, compared to BIA only and partial ZF only.

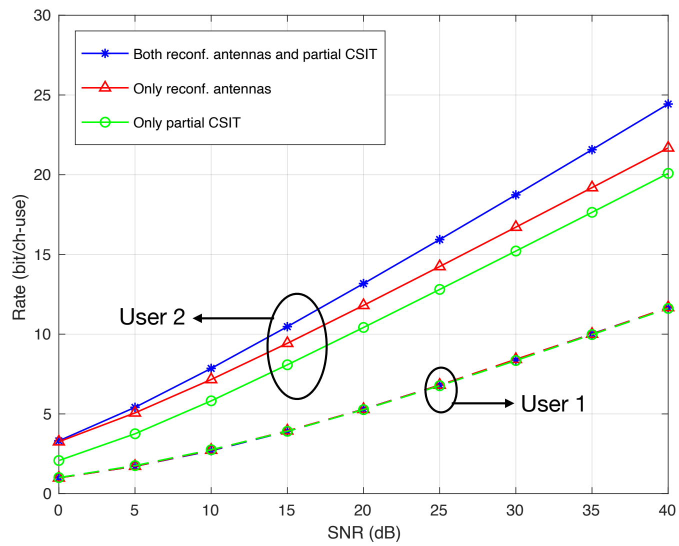

Fig. 8: Rate vs. SNR of User 1 and 2 at $\beta = 0.5$

References

[1] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4418–4431, 2012.

[2] H. Maleki, S. A. Jafar, and S. Shamai, "Retrospective interference alignment over interference networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 3, pp. 228–240, 2012.

[3] R. Tandon, S. A. Jafar, S. Shamai, and H. V. Poor, "On the synergistic benefits of alternating CSIT for the MISO broadcast channel," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4106–4128, 2013.

[4] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 315–328, 2013.

[5] T. Gou and S. A. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MISO BC with mixed CSIT," *IEEE Communications Letters*, vol. 16, no. 7, pp. 1084–1087, 2012.

[6] S. A. Jafar, "Topological interference management through index coding," *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 529–568, 2014.

[7] Xinping Yi and D. Gesbert, "Topological Interference Management with Transmitter Cooperation," *Information Theory, IEEE Transactions on*, vol. 61, no. 11, pp. 6107–6130, 2015.

[8] Xinping Yi, H. Sun, S. Jafar, and D. Gesbert, "TDMA is optimal for all-unicast DoF region of TIM if and only if topology is chordal bipartite," *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 2065 – 2076, March 2018.

[9] B. Clerckx, H. Joudeh, C. Hao, D. Mingbo, and B. Rassouli, "Rate splitting for MIMO wireless networks: A promising PHY-layer strategy for LTE evolution," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 98–105, May 2016.

[10] H. Joudeh and B. Clerckx, "Robust Transmission in Downlink Multiuser MISO Systems: A Rate-Splitting Approach," *IEEE Transactions on Signal Processing*, vol. 64, no. 23, pp. 6227–6242, December 2016.

[11] C. Geng, N. Naderializadeh, A. S. Avestimehr, and S. A. Jafar, "On the optimality of treating interference as noise," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1753–1767, 2015.

[12] H. Joudeh, Xinping Yi, B. Clerckx, and G. Caire, "On the Optimality of Treating Inter-Cell Interference as Noise: Downlink Cellular Networks and Uplink-Downlink Duality," *IEEE Transactions on Information Theory*, vol. 66, no. 11, pp. 6939–6961, November 2020.

[13] Xinping Yi and G. Caire, "Optimality of treating interference as noise: A combinatorial perspective," *IEEE Transactions on Information Theory*, vol. 62, no. 8, pp. 4654–4673, 2016.

[14] S. A. Jafar, "Blind interference alignment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 3, pp. 216–227, 2012.

[15] T. Gou, C. Wang, and S. A. Jafar, "Aiming perfectly in the dark-blind interference alignment through staggered antenna switching," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2734–2744, 2011.

[16] L. Ke and Z. Wang, "Degrees of freedom regions of two-user MIMO Z and full interference channels: The benefit of reconfigurable antennas," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3766–3779, 2012.

[17] S. A. Jafar and M. J. Fakhereddin, "Degrees of freedom for the MIMO interference channel," *IEEE Transactions on Information Theory*, vol. 53, no. 7, pp. 2637–2642, 2007.

[18] A. G. Davoodi and S. A. Jafar, "Network coherence time matters—aligned image sets and the degrees of freedom of interference networks with finite precision CSIT and perfect CSIR," *IEEE Transactions on Information Theory*, vol. 64, no. 12, pp. 7780–7791, 2018.

[19] H. Joudeh and B. Clerckx, "Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A rate-splitting

approach," *IEEE Transactions on Communications*, vol. 64, no. 11, pp. 4847–4861, 2016.

[20] H. Joudeh and B. Clerckx, "Rate-splitting for max-min fair multigroup multicast beamforming in overloaded systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7276–7289, 2017.

[21] B. Yuan and S. A. Jafar, "Elevated multiplexing and signal space partitioning in the 2 user MIMO IC with partial CSIT," in *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2016, pp. 1–6.

[22] C. Hao, B. Rassouli, and B. Clerckx, "Achievable DoF regions of MIMO networks with imperfect CSIT," *IEEE Transactions on Information Theory*, vol. 63, no. 10, pp. 6587–6606, 2017.

[23] B. Yuan, A. G. Davoodi, and S. A. Jafar, "DoF region of the MIMO interference channel with partial CSIT," in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2017, pp. 1–5.

[24] C. Huang, S. A. Jafar, S. Shamai, and S. Vishwanath, "On degrees of freedom region of MIMO networks without channel state information at transmitters," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 849–857, 2012.

[25] Y. Zhu and D. Guo, "The degrees of freedom of isotropic MIMO interference channels without state information at the transmitters," *IEEE Transactions on Information Theory*, vol. 58, no. 1, pp. 341–352, 2012.