DoF Analysis of the MIMO Broadcast Channel with Alternating/Hybrid CSIT

Borzoo Rassouli, Chenxi Hao and Bruno Clerckx

***Abstract*—We consider a K-user multiple-input single-output (MISO) broadcast channel (BC) where the channel state information (CSI) of user $i(i = 1, 2, \ldots, K)$ may be instantaneously perfect (P), delayed (D) or not known (N) at the transmitter with probabilities λ_P^i, λ_D^i and λ_N^i, respectively. In this setting, according to the three possible CSIT for each user, knowledge of the joint CSIT of the K users could have at most 3^K states. In this paper, given the marginal probabilities of CSIT (i.e., λ_P^i, λ_D^i and λ_N^i), we derive an outer bound for the DoF region of the K-user MISO BC. Subsequently, we tighten this outer bound by taking into account a set of inequalities that capture some of the 3^K states of the joint CSIT. One of the consequences of this set of inequalities is that for $K \geq 3$, it is shown that the DoF region is not completely characterized by the marginal probabilities in contrast to the two-user case. Afterwards, the tightness of these bounds are investigated through the discussion on the achievability. Finally, a two user MIMO BC having CSIT among P and N is considered in which an outer bound for the DoF region is provided and it is shown that in some scenarios it is tight.**

I. INTRODUCTION

In contrast to point to point multiple-input multiple-output (MIMO) communication where the channel state information at the transmitter (CSIT) does not affect the multiplexing gain, in a multiple-input single-output (MISO) broadcast channel (BC), knowledge of CSIT is crucial for interference mitigation and beamforming purposes [1]. However, the assumption of perfect CSIT may not always be true in practice due to channel estimation error and feedback latency. Therefore, the idea of communication under some sort of imperfection in CSIT has gained more attention recently. The so called MAT algorithm was presented in [2] where it was shown that in terms of the degrees of freedom, even an outdated CSIT can result in significant performance improvement in comparison to the case with no CSIT. Assuming correlation between the feedback information and current channel state (e.g., when the feedback latency is smaller than the coherence time of the channel), the authors in [3] and [4] consider the degrees of freedom in a time correlated MISO BC which is shown to be a combination of zero forcing beamforming (ZFBF) and MAT algorithm. Following these works, the general case of mixed CSIT and the K-user MISO BC with time correlated delayed CSIT are discussed in [5] and [6], respectively. While all these works consider the concept of delayed CSIT in time domain, [7] and [8] deal with the DoF region and its achievable schemes in a frequency correlated MISO BC where there is no delayed CSIT but imperfect CSIT across subbands, which is more inline with practical systems as Long Term Evolution (LTE) [1]. In [9], the synergistic benefits of alternating CSIT over fixed CSIT was presented in a two user MISO BC with two transmit antennas. In [10] and [11], the MISO BC with hybrid CSIT (Perfect or Delayed) was considered. The recent work of [12] investigates the DoF region of the K-user MISO BC with hybrid CSIT and linear encoding at the transmitter. [13] and [14] show that the optimal sum DoF is achievable if the CSIT is not too delayed in broadcast channels and interference networks, respectively.

The complete characterization of the general MISO BC with perfect, delayed or unknown CSIT is an open problem. The main aim of this paper is to investigate this problem and provide some answers toward this goal. To this end, our contributions are as follows.

- Given the marginal probabilities of CSIT in a K-user MISO BC, we derive an outer bound for the DoF region.
- A set of inequalities is proposed that captures not only the marginals, but also the joint CSIT distribution. This shows that for the K-user case ($K \geq 3$), marginal probabilities are not sufficient for characterizing the DoF region.
- The tightness of the outer bounds is investigated in certain cases.
- Finally, a two-user MIMO BC is considered in which the CSI of a user is either perfect or unknown. An outer bound for the DoF region is provided and it is shown to be tight when the joint CSIT probabilities satisfy a certain relationship.

The paper is organized as follows. In section II the system model and preliminaries are presented. An outer bound is provided in section III based on the marginal probabilities and the proof is given in section IV. Section V provides an outer bound that depends on the joint CSIT probabilities. The tightness of the outerbounds will be discussed in section VI. Section VII investigates a two user MIMO BC with CSIT either perfect or unknown, and section VIII concludes the paper.

Throughout the paper, $f \sim o(\log P)$ is equivalent to $\lim_{P\to\infty} \frac{f}{\log P} = 0$. $(.)^T$ and $(.)^H$ denote the transpose and conjugate transpose, respectively. $CN(\mathbf{0}, \mathbf{\Sigma})$ is the circularly

Borzoo Rassouli and Chenxi Hao are with the Communication and Signal Processing group of Department of Electrical and Electronics, Imperial College London, United Kingdom. emails: b.rassouli12@imperial.ac.uk , chenxi.hao10@imperial.ac.uk

Bruno Clerckx is with the Communication and Signal Processing group of Department of Electrical and Electronics, Imperial College London and the School of Electrical Engineering, Korea University, Korea. email: b.clerckx@imperial.ac.uk

This paper was presented in part at the IEEE International Conference on Communications (ICC) 2015, London, UK.

This work was partially supported by the Seventh Framework Programme for Research of the European Commission under grant number HARP-318489.

Time Slots

Users

D	N	P
P	D	N
P	P	P

Fig. 1: A CSIT pattern with $\lambda_{DPP} = \lambda_{NDP} = \lambda_{PNP} = \frac{1}{3}$

symmetric complex Gaussian distribution with covariance matrix $\mathbf{\Sigma}$. For a pair of integers $m \leq q$, the discrete interval is defined as $[m : q] = \{m, m+1, \ldots, q\}$. $Y_{[i:j]} = \{Y_i, Y_{i+1}, \ldots, Y_j\}$, $Y([i : j]) = \{Y(i), Y(i+1), \ldots, Y(j)\}$ and $Y^n = Y([1 : n])$.

II. System Model

We consider a MISO BC, in which a base station with M antennas sends independent messages $W_1, \ldots, W_K$ to K single-antenna users ($M \geq K$). In a flat fading scenario, the discrete-time baseband received signal of user k at channel use (henceforth, time instant) t can be written as

$$Y_k(t) = \mathbf{H}_k^H(t)\mathbf{X}(t) + W_k(t) \ , \ k \in [1 : K] \ , \ t \in [1 : n] \quad (1)$$

where $\mathbf{X}(t) \in C^{(M\times 1)}$ is the transmitted signal at time instant t satisfying the (per codeword) power constraint $\sum_{t=1}^{n} \|\mathbf{x}(t)\|^2 \leq nP$. $W_k(t)$ and $\mathbf{H}_k(t)$ are the additive noise and channel vector of user k, respectively, and are also assumed i.i.d. over the time instants and the users. We assume global perfect Channel State Information at Receivers (CSIR).

The rate tuple $(R_1, R_2, \ldots, R_K)$, in which $R_i = \frac{\log(|W_i|)}{n}$, is achievable if there exists a coding scheme such that the probability of error in decoding W_i at user $i (i \in [1 : K])$ can be made arbitrarily small with sufficiently large coding block length. The DoF region is defined as $\{(d_1, \ldots, d_K) | \exists (R_1, R_2, \ldots, R_K) \in C(P) \text{ such that } d_i = \lim_{P\to\infty} \frac{R_i}{\log P}, \ \forall i\}$ where $C(P)$ is the capacity region (i.e., the closure of the set of achievable rate tuples).

The probabilistic model used in this paper for CSIT availability allows the transmitter to have a Perfect (P) instantaneous knowledge of the CSI of a particular user at some time instants, whereas at some other time instants it receives the CSI with Delay (D) and finally, for the remaining time instants the CSI of the user is Not known (N) at the transmitter. The CSIT model can be fixed (i.e., as in the hybrid model), alternating or both (i.e., fixed for a subset of the users and alternating for the remaining subset.) When there is delayed CSIT, we assume that the feedback delay is much larger than the coherence time of the channel making the feedback information completely independent of the current channel state. In this configuration, the joint CSIT of all the K users has at most 3^K states. For example, in a 3 user MISO BC, they will be $PPP, PPD, PPN, PDP, \ldots$ with corresponding probabilities $\lambda_{PPP}, \lambda_{PPD}, \lambda_{PPN}, \lambda_{PDP}, \ldots$ and, as an example, the marginal probability of perfect CSIT for user 1 is $\lambda_P^1 = \sum_{Q,Q'\in\{P,D,N\}} \lambda_{PQQ'}$.

P	D	D
D	P	D
D	D	P

Fig. 2: A symmetric CSIT pattern for the 3-user MISO BC with the marginals $\lambda_P = \frac{1}{3}, \lambda_D = \frac{2}{3}$.

By CSIT pattern we refer to the knowledge of CSIT represented in a space-time matrix where the rows and columns represent users and time slots, respectively. The channel remains fixed within each time slot, while it changes independently from one slot to another. For simplicity, we assume the delayed CSI arrives at the transmitter after one time slot. Figure 1 shows an example of a CSIT pattern, in which the transmitter knows the channels of users 2 and 3 perfectly at time slot 1 and has no information about the channel of user 1. The CSI of user 1 will be known in the next time slot due to feedback delay and is completely independent of the channel in time slot 2.

Finally, a symmetric CSIT pattern means that the marginal probabilities of perfect, delayed and unknown CSIT are the same across the users, i.e. $\lambda_Q^i = \lambda_Q, \ \forall i \in [1 : K], Q \in \{P, D, N\}$. As an example figure 2 shows a symmetric CSIT pattern for the 3-user MISO BC in which $\lambda_P = \frac{1}{3}, \lambda_D = \frac{2}{3}$.

III. An outer bound given the marginals

Theorem 1. Let $\pi^j(.)$ be an arbitrary permutation of size j over the indices $(1, 2, \ldots, K)$, and $\alpha_{\pi^j}(.)$ be a permutation of π^j satisfying[1]

$$(\lambda_P^{\alpha_{\pi^j}(i)} + \lambda_D^{\alpha_{\pi^j}(i)}) \leq (\lambda_P^{\alpha_{\pi^j}(i+1)} + \lambda_D^{\alpha_{\pi^j}(i+1)}) \ , \ i \in [1 : j-1]. \quad (2)$$

Given the marginal probabilities of CSIT for user i (which can be any two of λ_P^i, λ_D^i and λ_N^i, since $\lambda_P^i + \lambda_D^i + \lambda_N^i = 1$), an outer bound for the DoF region of the K-user MISO BC with M transmit antennas at the transmitter ($M \geq K$) is defined by the following sets of inequalities

$$\sum_{i=1}^{j} \frac{d_{\pi^j(i)}}{i} \leq 1 + \sum_{i=2}^{j} \frac{\sum_{r=1}^{i-1} \lambda_P^{\pi^j(r)}}{i(i-1)} \quad (3)$$

$$\sum_{i=1}^{j} d_{\pi^j(i)} \leq 1 + \sum_{i=1}^{j-1} (\lambda_P^{\alpha_{\pi^j}(i)} + \lambda_D^{\alpha_{\pi^j}(i)}) \ , \quad \forall \pi^j, \ j \in [1 : K]. \quad (4)$$

For the symmetric scenario, the sets of inequalities are simplified as

$$\sum_{i=1}^{j} \frac{d_{\pi^j(i)}}{i} \leq 1 + \lambda_P \sum_{i=2}^{j} \frac{1}{i} \quad (5)$$

$$\sum_{i=1}^{j} d_{\pi^j(i)} \leq 1 + (j-1)(\lambda_P + \lambda_D) \ , \quad \forall \pi^j, \ j \in [1 : K]. \quad (6)$$

For $K = 2$, the outer bound boils down to the optimal DoF region in [9].

[1]The reason for arranging the users according to the sum of the perfect and delayed CSIT probabilities becomes clear in (28).

IV. Proof of Theorem 1

For simplicity, we assume $j = K$, since it is obvious that each subset of users with cardinality j ($j < K$) can be regarded as a j-user BC. Also, we assume the identity permutation (i.e., $\pi^K(i) = i$) while the results could be easily applied to any other arbitrary permutation.

A. Proof of $\sum_{i=1}^{K} \frac{d_i}{i} \leq 1 + \sum_{i=2}^{K} \frac{\sum_{r=1}^{i-1} \lambda_P^r}{i(i-1)}$

First, we improve the channel by giving the message and observation of user i to users $[i+1 : K]$ ($i \in [1 : K-1]$). Hence, from Fano's inequality,

$$nR_i \leq I(W_i; Y_{[1:i]}^n | W_{[1:i-1]}, \Omega^n) + n\epsilon_n \tag{7}$$

where Ω^n denotes the global CSIR up to time instant n, $W_0 = \emptyset$ and ϵ_n goes to zero as n goes to infinity. By this improvement, channel input and outputs (i.e., the enhanced observations of users) form a Markov chain which results in a physically degraded broadcast channel [15]. Therefore, according to [16], since feedback does not increase the capacity of physically degraded broadcast channels, we can ignore the delayed CSIT (D) and replace them with No CSIT (N). This is equivalent to having the channel of user i perfectly known with probability λ_P^i and not known otherwise. From now on, we ignore the term $n\epsilon_n$ for simplicity (since later it will be divided by n and $n \to \infty$) and write

$$\sum_{i=1}^{K} \frac{nR_i}{i} \leq \sum_{i=1}^{K} \frac{I(W_i; Y_{[1:i]}^n | W_{[1:i-1]}, \Omega^n)}{i} \tag{8}$$

$$\leq h(Y_1^n|\Omega^n) + \sum_{i=2}^{K} \left[\frac{h(Y_{[1:i]}^n | W_{[1:i-1]}, \Omega^n)}{i} - \frac{h(Y_{[1:i-1]}^n | W_{[1:i-1]}, \Omega^n)}{i-1} \right] + no(\log P) \tag{9}$$

where $Y_0 = \emptyset$ and we have used the fact that $\frac{h(Y_{[1:K]}^n | W_{[1:K]}, \Omega^n)}{nK} \sim o(\log P)$, since with the knowledge of $W_{[1:K]}$ and Ω^n, the observations $Y_{[1:K]}^n$ can be reconstructed within the noise distortion. Before going further, the following lemma is needed.

Lemma 1. Let $\Gamma_N = \{Y_1, Y_2, \ldots, Y_N\}$ be a set of $N (\geq 2)$ arbitrary random variables and $\Psi_i^j(\Gamma_N)$ be a sliding window of size j over Γ_N ($1 \leq i, j \leq N$) starting from Y_i i.e.,

$$\Psi_i^j(\Gamma_N) = Y_{(i-1)_N+1}, Y_{(i)_N+1}, \ldots, Y_{(i+j-2)_N+1}$$

where $(.)_N$ defines the modulo N operation. Then, the following inequality holds for $\forall m \in [1 : N-1]$

$$(N-m) h(Y_{[1:N]}|A) \leq \sum_{i=1}^{N} h(\Psi_i^{N-m}(\Gamma_N)|A) \tag{10}$$

where A is an arbitrary condition.

Proof: We prove the lemma by showing that for every fixed $m (\geq 1)$, (10) holds for all $N (\geq m+1)$ using induction. It is obvious that for every $m (\geq 1)$, (10) holds for $N = m+1$. In other words, $h(Y_{[1:N]}|A) \leq \sum_{i=1}^{N} h(Y_i|A)$. Now, considering that (10) is valid for $N (\geq m+1)$, we show that it also holds for $N+1$. Replacing N with $N+1$, we have

$$(N+1-m) h(Y_{[1:N+1]}|A)$$

$$= h(Y_{[1:N+1]}|A) + (N-m) h(Y_{[1:N-1]}, \overbrace{Y_N, Y_{N+1}}^{Z}|A)$$

$$\leq h(Y_{[1:N+1]}|A) + \sum_{i=1}^{N} h(\Psi_i^{N-m}(\Phi_N)|A) \tag{11}$$

$$= h(Y_{[1:N+1]}|A) + \sum_{i=1}^{m} h(\Psi_i^{N-m}(\Phi_N)|A) + \sum_{i=m+1}^{N} h(\Psi_i^{N+1-m}(\Gamma_{N+1})|A) \tag{12}$$

$$= h(Y_{[N-m+1:N]}|Y_{N+1}, Y_{[1:N-m]}, A) + \sum_{i=1}^{m} h(\Psi_i^{N-m}(\Phi_N)|A) + h(Y_{N+1}, Y_{[1:N-m]}|A) + \sum_{i=m+1}^{N} h(\Psi_i^{N+1-m}(\Gamma_{N+1})|A) \tag{13}$$

$$= h(Y_{[N-m+1:N]}|Y_{N+1}, Y_{[1:N-m]}, A) + \sum_{i=1}^{m} h(\Psi_i^{N-m}(\Phi_N)|A) + \sum_{i=m+1}^{N+1} h(\Psi_i^{N+1-m}(\Gamma_{N+1})|A)$$

$$= \sum_{i=1}^{m} h(Y_{N-m+i}|Y_{N+1}, Y_{[1:N-m+i-1]}, A) + \sum_{i=1}^{m} h(Y_{[i:N-m+i-1]}|A) + \sum_{i=m+1}^{N+1} h(\Psi_i^{N+1-m}(\Gamma_{N+1})|A) \tag{14}$$

$$\leq \sum_{i=1}^{m} h(Y_{N-m+i}|Y_{[i:N-m+i-1]}, A) + \sum_{i=1}^{m} h(Y_{[i:N-m+i-1]}|A) + \sum_{i=m+1}^{N+1} h(\Psi_i^{N+1-m}(\Gamma_{N+1})|A) \tag{15}$$

$$= \sum_{i=1}^{m} h(\Psi_i^{N+1-m}(\Gamma_{N+1})|A) + \sum_{i=m+1}^{N+1} h(\Psi_i^{N+1-m}(\Gamma_{N+1})|A)$$

$$= \sum_{i=1}^{N+1} h(\Psi_i^{N+1-m}(\Gamma_{N+1})|A) \tag{16}$$

where in (11), $\Phi_N = \{Y_{[1:N-1]}, Z\}$ and we have used the validity of (10) for N. In (12), we have used the fact that $\Psi_i^{N+1-m}(\Gamma_{N+1}) = \Psi_i^{N-m}(\Phi_N)$ for $i \in [m+1 : N]$. In (13), the chain rule of entropies is used and in (14), the sliding window is written in terms of its elements. Finally, in (15), the fact that conditioning reduces the differential entropy is used. Therefore, since $m (\geq 1)$ was chosen arbitrarily and (10) is valid for $N = m+1$ and from its validity for $N (\geq m+1)$ we could show it also holds for $N+1$, we conclude that (10) holds for all values of m and N satisfying $1 \leq m \leq N-1$. ■

Each term in the summation of (9) can be rewritten as

$$\frac{(i-1) h(Y_{[1:i]}^n | W_{[1:i-1]}, \Omega^n) - i h(Y_{[1:i-1]}^n | W_{[1:i-1]}, \Omega^n)}{i(i-1)}$$

$$\leq \frac{\sum_{r=1}^{i}\left[h(\Psi_r^{i-1}(\Gamma_i)|T_{i,n}) - h(Y_{[1:i-1]}^n|T_{i,n})\right]}{i(i-1)} \quad (17)$$

$$= \frac{\sum_{r=1}^{i-1}\left[h(Y_i^n|E_{r,i},T_{i,n}) - h(Y_r^n|E_{r,i},T_{i,n})\right]}{i(i-1)} \quad (18)$$

where $\Gamma_i = \{Y_{[1:i]}^n\}$, $T_{i,n} = \{W_{[1:i-1]}, \Omega^n\}$ and $E_{r,i} = \{Y_{[1:i-1]}^n\} - \{Y_r^n\}$. (17) is from the application of lemma 1 ($m = 1$) and (18) is from the chain rule of entropies. Before going further, the following lemma is needed. This lemma, which is based on [17], is the key part in the proof.

Lemma 2. In the K-user MISO BC defined in (1), for the users $m, q \in [1:K]$ $(m \neq q)$, we have

$$\lim_{n,P\to\infty} \frac{h(Y_m^n|A) - h(Y_q^n|A)}{n\log P} \leq \begin{cases} 1 & \text{CSIT of } q \text{ is } P \\ 0 & \text{CSIT of } q \text{ is } N \end{cases} \quad (19)$$

where A is a condition such as the condition of entropies in (18) or later in (25). Interestingly, (19) is only a function of the CSIT of the second user.

Proof: Based on the four possible states for the joint CSIT of m and q, we have

1) CSIT of m is N or P and CSIT of q is P:

$$h(Y_m^n|A) - h(Y_q^n|A) \leq \underbrace{h(Y_m^n|A)}_{\leq n\log(P)} - \underbrace{h(Y_q^n|A, W_{[1:K]})}_{no(\log P)} \quad (20)$$

A Gaussian input with the conditional covariance matrix of $\Sigma_{X|A} = P\mathbf{u}_q^{\perp}\mathbf{u}_q^{\perp H}$ achieves the upper bound, where $\mathbf{u}_q^{\perp}$ is a unit vector in the direction orthogonal to $\mathbf{H}_q$ (since $\mathbf{H}_q$ is known).

2) CSIT of m is N and CSIT of q is N: In this case both Y_m^n and Y_q^n are statistically equivalent (i.e., having the same probability density functions, and subsequently, the same entropies.) Therefore,

$$h(Y_m^n|A) - h(Y_q^n|A) = 0 \quad (21)$$

3) CSIT of m is P and CSIT of q is N: This is the second result of Theorem 1 in [17][2]. ■

From (9) and (18), we have

$$\begin{aligned}\sum_{i=1}^{K}\frac{nR_i}{i} &\leq \sum_{i=2}^{K}\sum_{r=1}^{i-1}\frac{h(Y_i^n|A_{r,i}) - h(Y_r^n|A_{r,i})}{i(i-1)} \\ &\quad + n\log P + no(\log P) \\ &\leq n\log P + \sum_{i=2}^{K}\sum_{r=1}^{i-1}\frac{n\lambda_P^r}{i(i-1)}\log P + no(\log P)\end{aligned} \quad (22)$$

where $A_{r,i}$ is the condition of the entropies in (18) and (22) is from the application of lemma 2 and the fact that n is sufficiently large. Therefore,

$$\sum_{i=1}^{K}\frac{d_i}{i} \leq 1 + \sum_{i=2}^{K}\frac{\sum_{r=1}^{i-1}\lambda_P^r}{i(i-1)}. \quad (23)$$

[2]The differential entropy terms in the left hand side of (19) can be written in terms of the expectation of the difference of entropies conditioned on the realizations of A. Since the conditional probability density functions exist and have a bounded peak, the same steps of [17] as discretization, considering the cannonical form and bounding the cardinality of aligned image set can be applied.

It is obvious that the same approach can be applied to any other permutations on $(1, 2, \ldots, K)$ which results in (3). In addition to the mentioned proof, an alternative proof is provided in Appendix A.

B. Proof of $\sum_{i=1}^{K} d_i \leq 1 + \sum_{i=1}^{K-1}(\lambda_P^{\alpha_{\pi^K}(i)} + \lambda_D^{\alpha_{\pi^K}(i)})$

We enhance the channel in two ways:

1) Like the approach in [9], whenever there is delayed CSIT (D), we assume that it is perfect instantaneous CSIT (P), but we keep the probability of delayed CSIT. In other words, the CSIT of user i is perfect with probability $\lambda_P^i + \lambda_D^i$ and unknown otherwise.
2) We give the message of user i to users $[i+1:K]$.

Therefore,

$$nR_i \leq I(W_i; Y_i^n|W_{[1:i-1]}, \Omega^n) + n\epsilon_n \ , \ \forall i \in [1:K]. \quad (24)$$

By summing (40) over users and writing the mutual information in terms of differential entropies,

$$\begin{aligned}\sum_{i=1}^{K} nR_i &\leq \overbrace{h(Y_1^n|\Omega^n)}^{\leq n\log P} + no(\log P) \\ &\quad + \sum_{i=2}^{K}\left[h(Y_i^n|W_{[1:i-1]}, \Omega^n) - h(Y_{i-1}^n|W_{[1:i-1]}, \Omega^n)\right].\end{aligned} \quad (25)$$

By applying the results of lemma 2 to (25), we have

$$\sum_{i=1}^{K} d_i \leq 1 + \sum_{i=2}^{K}(\lambda_P^{i-1} + \lambda_D^{i-1}) = 1 + \sum_{i=1}^{K-1}(\lambda_P^i + \lambda_D^i). \quad (26)$$

Let $\pi^K(.)$ be an arbitrary permutation of size K on $(1, \ldots, K)$. Applying the same reasoning, we have

$$\sum_{i=1}^{K} d_i \leq 1 + \sum_{i=1}^{K-1}(\lambda_P^{\pi^K(i)} + \lambda_D^{\pi^K(i)}) \ , \ \forall \pi^K(.). \quad (27)$$

(27) results in K inequalities all having the same left hand side. Therefore,

$$\sum_{i=1}^{K} d_i \leq 1 + \min_{\pi^K(.)}\sum_{i=1}^{K-1}(\lambda_P^{\pi^K(i)} + \lambda_D^{\pi^K(i)}) \quad (28)$$

This is due to the possible orders of channel enhancements and it is obvious that $\alpha_{\pi^K}(.)$ will minimize (28) if it satisfies (2) (for $j = K$.)

V. An outer bound capturing the joint CSIT probabilities

In the previous section, an outer bound was provided in terms of the marginal probabilities. In this section, we tighten the outer bound by introducing a set of inequalities that captures the joint CSIT probabilities. We start with simple motivating examples. Consider the pattern shown in figure 3. By Fano's inequality, we write,

$$nR_1 \leq I(W_1; Y_1^n|\Omega^n) \quad (29)$$

$$nR_1 \leq I(W_1; Y_1^n|\Omega^n, W_2). \quad (30)$$

P	N	N
N	P	N
N	N	P

Fig. 3: A symmetric CSIT pattern for the 3-user MISO BC.

Adding (29) and (30) results in

$$2nR_1 \leq I(W_1; Y_1^n|\Omega^n) + I(W_1; Y_1^n|\Omega^n, W_2). \tag{31}$$

By doing the same for R_2, we have

$$2nR_2 \leq I(W_2; Y_2^n|\Omega^n) + I(W_2; Y_2^n|\Omega^n, W_1). \tag{32}$$

Finally, the rate of user 3 is written as

$$nR_3 \leq I(W_3; Y_3^n|\Omega^n, W_1, W_2). \tag{33}$$

Therefore,

$$\begin{aligned}
&2nR_1 + 2nR_2 + nR_3 \\
&\leq \underbrace{h(Y_2^n|\Omega^n, W_1) - h(Y_1^n|\Omega^n, W_1)}_{\leq \frac{n}{3}\log P} + h(Y_3^n|\Omega^n, W_1, W_2) \\
&\quad \underbrace{+h(Y_1^n|\Omega^n, W_2) - h(Y_2^n|\Omega^n, W_2)}_{\leq \frac{n}{3}\log P} + \underbrace{h(Y_1^n|\Omega^n)}_{\leq n\log P} + \underbrace{h(Y_2^n|\Omega^n)}_{\leq n\log P} \\
&\quad \underbrace{-h(Y_1^n|\Omega^n, W_1, W_2) - h(Y_2^n|\Omega^n, W_1, W_2)}_{\leq -h(Y_1^n, Y_2^n|\Omega^n, W_1, W_2)} \qquad (34) \\
&\leq \frac{8n}{3}\log P + h(Y_3^n|\Omega^n, W_1, W_2) - h(Y_1^n, Y_2^n|\Omega^n, W_1, W_2) \qquad (35) \\
&= \frac{8n}{3}\log P + \underbrace{h(Y_3^n|\Theta) - h(Y_{2,PNN}^n, Y_{1,NPN}^n, Y_{1,NNP}^n|\Theta)}_{o(\log P)} \\
&\quad \underbrace{-h(Y_{1,PNN}^n, Y_{2,NPN}^n, Y_{2,NNP}^n|\Theta, Y_{2,PNN}^n, Y_{1,NPN}^n, Y_{1,NNP}^n)}_{\leq -h(Y_{1,PNN}^n, Y_{2,NPN}^n, Y_{2,NNP}^n|\Theta, Y_{2,PNN}^n, Y_{1,NPN}^n, Y_{1,NNP}^n, W_3) \sim o(\log P)} \qquad (36) \\
&\leq \frac{8n}{3}\log P \qquad (37)
\end{aligned}$$

where in (34), lemma 2 is applied to the differences resulting in the values written under the braces and in (36), $\Theta = \{\Omega^n, W_1, W_2\}$. We have split the observations of users 1 and 2 in terms of the joint CSIT, i.e., $Y_1^n = (Y_{1,PNN}^n, Y_{1,NPN}^n, Y_{1,NNP}^n)$ and $Y_2^n = (Y_{2,PNN}^n, Y_{2,NPN}^n, Y_{2,NNP}^n)$. (36) is due to the fact that there is at least one unknown CSIT (N) in the joint states of user 1 and user 2 (i.e., PN, NP and NN. see rows 1 and 2 of the CSIT pattern shown in figure 3). Therefore, we have the following inequalities for the pattern shown in figure 3

$$\begin{aligned}
2d_1 + 2d_2 + d_3 &\leq \frac{8}{3} \\
2d_1 + d_2 + 2d_3 &\leq \frac{8}{3} \\
d_1 + 2d_2 + 2d_3 &\leq \frac{8}{3}.
\end{aligned} \tag{38}$$

From (38), the sum DoF of the pattern in figure 3 has the upper bound of $\frac{8}{5}$, while it can be easily verified that for the pattern with PPP in the first slot and NNN in the next two slots, which has the same marginals as in figure 3, the sum DoF is $\frac{5}{3} (> \frac{8}{5})$. This simple example confirms that for the K-user MISO BC ($K \geq 3$), the marginal probabilities are not sufficient in characterizing the DoF region[3]. Motivated by this simple example, we can have the following set of inequalities for the 3-user MISO BC with P and N

$$\begin{aligned}
2d_1 + 2d_2 + d_3 &\leq 2 + 2\lambda_P + \lambda_{PP-} \\
2d_1 + d_2 + 2d_3 &\leq 2 + 2\lambda_P + \lambda_{P-P} \\
d_1 + 2d_2 + 2d_3 &\leq 2 + 2\lambda_P + \lambda_{-PP}
\end{aligned} \tag{39}$$

where a dashed line in the above means that the CSIT of the corresponding user is not important (for example, $\lambda_{PP-} = \lambda_{PPP} + \lambda_{PPN}$ which is a summation over all the possible states for the CSIT of user 3). By looking at the difference of entropies in (35), it is observed that this difference is of order $o(\log P)$ when there is at least one N in the joint CSIT of users 1 and 2 (i.e., PNN, PNP, NPN, NPP, NNP and NNN) and, therefore, is upperbounded by $n(\lambda_{PPP} + \lambda_{PPN}) \log P$. This results in the first inequality of (39) and the same reasoning applies to the remaining two inequalities. (39) is a set of inequalities that captures the joint CSIT probabilities and is not only a function of the marginals.

P	N	N	N
N	P	N	N
N	N	P	N
N	N	N	P

Fig. 4: A symmetric CSIT pattern for the 4-user MISO BC.

Now consider the pattern shown in figure 4 for the 4-user MISO BC. From (31), (32) and (33), we can write

$$\begin{aligned}
&2n(R_1 + R_2 + R_3) \\
&\leq \underbrace{h(Y_2^n|\Omega^n, W_1) - h(Y_1^n|\Omega^n, W_1)}_{\leq \frac{n}{4}\log P} \\
&\quad \underbrace{+h(Y_1^n|\Omega^n, W_2) - h(Y_2^n|\Omega^n, W_2)}_{\leq \frac{n}{4}\log P} + \underbrace{h(Y_1^n|\Omega^n)}_{\leq n\log P} + \underbrace{h(Y_2^n|\Omega^n)}_{\leq n\log P} \\
&\quad \underbrace{+h(Y_3^n|\Omega^n, W_1, W_2) - h(Y_1^n|\Omega^n, W_1, W_2)}_{\leq \frac{n}{4}\log P} \\
&\quad \underbrace{+h(Y_3^n|\Omega^n, W_1, W_2) - h(Y_2^n|\Omega^n, W_1, W_2)}_{\leq \frac{n}{4}\log P} \\
&\quad - 2h(Y_3^n|\Omega^n, W_1, W_2, W_3) \\
&\leq 3n\log P - 2h(Y_3^n|\Omega^n, W_1, W_2, W_3)
\end{aligned} \tag{40}$$

Alternatively, we can change the role of users 1 and 3 and

[3]It is important to emphasize on the difference between the following two statements

a) Two CSIT patterns with **different marginals** can have the **same** DoF regions.

b) Two CSIT patterns with **the same marginals** can have **different** DoF regions.

The first statement is already known in literature. For example, by comparing the original 2-user MAT (i.e., $\lambda_D = 1$) and the scheme DN,ND,NN in [9], it is concluded that both of them have the sum DoF of 4/3, while having different marginal prbabilities (for the latter, $\lambda_D = \frac{1}{3}$). However, the set of inequalities proposed in this section addresses the second statement which is a new problem and cannot result from the first statement.

write

$$\begin{aligned}2nR_1 &\le I(W_1; Y_1^n|\Omega^n, W_2, W_3) + I(W_1; Y_1^n|\Omega^n, W_2, W_3)\\ 2nR_2 &\le I(W_2; Y_2^n|\Omega^n) + I(W_2; Y_2^n|\Omega^n, W_3)\\ 2nR_3 &\le I(W_3; Y_3^n|\Omega^n) + I(W_3; Y_3^n|\Omega^n, W_2).\end{aligned}$$

Following the same reasoning in (40), we have

$$2n(R_1 + R_2 + R_3) \le 3n \log P - 2h(Y_1^n|\Omega^n, W_1, W_2, W_3). \tag{41}$$

Adding (40) and (41), we have

$$\begin{aligned}&4n(R_1 + R_2 + R_3)\\ &\le 6n \log P\\ &\quad - 2\left(h(Y_1^n|\Omega^n, W_1, W_2, W_3) + h(Y_3^n|\Omega^n, W_1, W_2, W_3)\right)\\ &\le 6n \log P - 2h(Y_1^n, Y_3^n|\Omega^n, W_1, W_2, W_3). \qquad (42)\end{aligned}$$

For the rate of user 4, we can write

$$\begin{aligned}2nR_4 &\le 2I(W_4; Y_4^n|\Omega^n, W_1, W_2, W_3)\\ &= 2h(Y_4^n|\Omega^n, W_1, W_2, W_3)\\ &\quad \underbrace{-2h(Y_4^n|\Omega^n, W_1, W_2, W_3, W_4)}_{o(\log P)}. \qquad (43)\end{aligned}$$

Adding (42) and (43), we get

$$\begin{aligned}&4n(R_1 + R_2 + R_3) + 2nR_4\\ &\le 6n \log P + 2\left(h(Y_4^n|\Psi) - h(Y_1^n, Y_3^n|\Psi)\right) \qquad (44)\\ &\le 6n \log P + \underbrace{2\left(h(Y_4^n|\Psi) - h(T_n|\Psi)\right)}_{o(\log P)} - 2h(T'_n|T_n, \Psi) \qquad (45)\\ &\le 6n \log P \underbrace{-2h(T'_n|T_n, \Psi, W_4)}_{o(\log P)}. \qquad (46)\end{aligned}$$

where in (44), $\Psi = \{\Omega^n, W_1, W_2, W_3\}$, and in (45), $T_n = \{Y^n_{3,PNNN}, Y^n_{1,NPNN}, Y^n_{1,NNPN}, Y^n_{1,NNNP}\}$, $T'_n = \{Y_1^n, Y_3^n\} - T_n$. Therefore, we have

$$2d_1 + 2d_2 + 2d_3 + d_4 \le 3 \tag{47}$$

In the left hand side of (47), user 4 has the coefficient of 1 and the remaining 3 users have the coefficient of 2. Also, instead of changing the roles of user 1 and 3, roles of user 2 and 3 or roles of user 1 and 2 could have been changed. Although this $\binom{3}{2}$ changes would not result in a new inequality due to the structure of the pattern shown in figure 4, these changes of the roles of the remaining 3 users (with coefficient 2) are necessary in general. Therefore, motivated by this simple example, we can have a set of inequalities for the 4-user MISO BC with P and N.

$$\begin{aligned}2d_1 + 2d_2 + 2d_3 + d_4 &\le 2 + 4\lambda_P\\ &\quad + \min\{\lambda_{PP--}, \lambda_{P-P-}, \lambda_{-PP-}\}\\ d_1 + 2d_2 + 2d_3 + 2d_4 &\le 2 + 4\lambda_P\\ &\quad + \min\{\lambda_{-PP-}, \lambda_{-P-P}, \lambda_{--PP}\}\\ 2d_1 + d_2 + 2d_3 + 2d_4 &\le 2 + 4\lambda_P\\ &\quad + \min\{\lambda_{P-P-}, \lambda_{P--P}, \lambda_{--PP}\}\\ 2d_1 + 2d_2 + d_3 + 2d_4 &\le 2 + 4\lambda_P\\ &\quad + \min\{\lambda_{P--P}, \lambda_{PP--}, \lambda_{-P-P}\}\end{aligned} \tag{48}$$

where each inequality in (48) is obtained from $\binom{3}{2}$ inequalities each of which with the same left hand side. The general K-user MISO BC can be addressed by using the following definition

$$\begin{aligned}\lambda(a, b) = &\text{ The probability that the CSIT of users } a \text{ and } b \text{ is P.}\\ &a, b \in [1 : K]\ ,\ a \ne b\end{aligned} \tag{49}$$

Theorem 2. Let $\pi^j(.)$ be an arbitrary permutation of size j over $[1 : K]$. For the K-user symmetric MISO BC with no delayed CSIT[4], we have

$$\begin{aligned}2\sum_{i=1}^{j-1} d_{\pi^j(i)} + d_{\pi^j(j)} &\le 2 + 2(j-2)\lambda_P\\ &\quad + \min_{a,b\in[1:j-1]:a<b}\{\lambda(\pi^j(a), \pi^j(b))\}\\ &\quad \forall \pi^j, j \in [3 : K].\end{aligned} \tag{50}$$

Proof. The proof is a straightforward generalization of the previous examples.

VI. On the achievability

In this section, we consider the bounds in (6) for the symmetric scenario.[5] For $K \ge 3$, we show that given the marginal probabilities of CSIT, there exists at least one CSIT pattern that achieves the outer bound in the following two scenarios.

A. $\lambda_D = 0$

In this case, $2^K - 1$ inequalities are active and the remaining inequalities become inactive. The reason can be easily verified from the inequalities, however, a simpler intuitive way is to consider that when there is no delayed CSIT, those inequalities derived from the degraded broadcast channel are inactive. In this case, the region is defined by $2^K - 1$ hyperplanes in R_+^K and has the following K corner points

$$(1, \lambda_P, \ldots, \lambda_P), (\lambda_P, 1, \lambda_P, \ldots, \lambda_P), \ldots, (\lambda_P, \ldots, \lambda_P, 1) \tag{51}$$

[4] The assumptions of symmetric scenario and no delayed CSIT are only used for the readability of formulations. It is important to note that the approach in this section can be applied to the general asymmetric scenario including the delayed CSIT (in this case, the delay is enhanced to perfect instantaneous as in subsection IV-B).

[5] The main goal of this section is to show that these bounds can become tight and are not always loose.

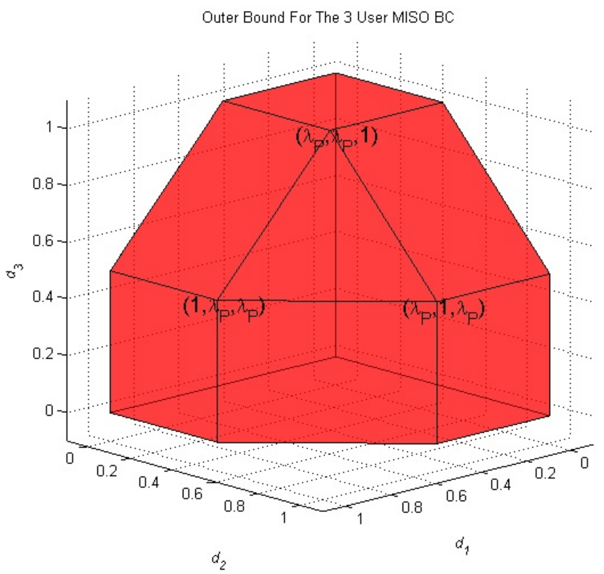

Fig. 5: Region in case A for 3 user BC

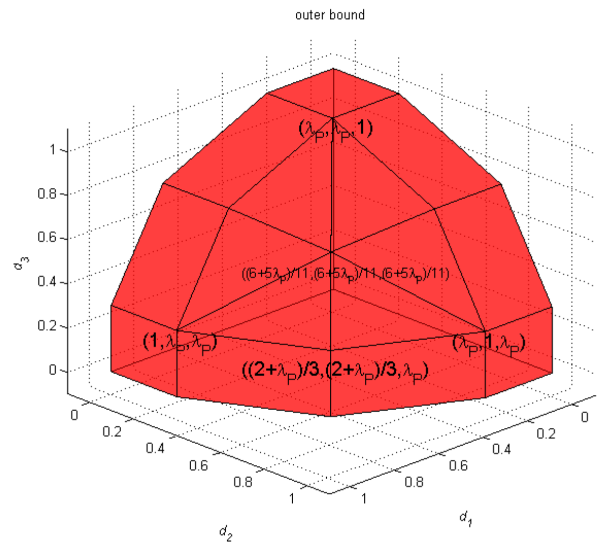

Fig. 6: Region in case B for 3 user BC

The corner points have the unique characteristic that the whole region can be constructed by time sharing between them. Therefore, the achievability of these points is equivalent to the achievability of the whole region. Figure 5 shows the region for the 3 user broadcast channel. The corner points are simply achieved by a scheme that has N time slots and consists of two parts: in the first $\lambda_P N$ time slots, zero forcing beamforming (ZFBF) is carried out where each user receives one interference-free symbol. In the remaining $\lambda_N N$ time slots, only one particular user (depending on the corner point of interest) is scheduled.

B. $\lambda_N \leq \frac{\lambda_D}{\sum_{j=2}^{K}\frac{1}{j}}$

Before going further, we need the following simple lemma.

Lemma 3. The minimum probability of delayed CSIT for sending order-j symbols in the K-user MAT is

$$\lambda_D^{min}(K,j) = 1 - \frac{K-j+1}{K\sum_{i=j}^{K}\frac{1}{i}}. \tag{52}$$

Proof: From [2], the MAT algorithm is based on a concatenation of K phases. Phase j takes $(K-j+1)\binom{K}{j}$ order-j messages as its input, takes $\binom{K}{j}$ time slots and produces $j\binom{K}{j+1}$ order-$j+1$ messages as its output. In each time slot of phase j, the transmitter sends a random linear combination of the $(K-j+1)$ symbols to a subset S of receivers , $|S| = j$. Sending the overheard interferences from the remaining $(K-j)$ receivers to receivers in subset S enables them to successfully decode their $(K-j+1)$ symbols by constructing a set of $(K-j+1)$ linearly independent equations. Therefore, the transmitter needs to know the channel of only $(K-j)$ receivers. In other words, at each time slot of phase j, the feedback of $(K-j)$ CSI is enough. In the MAT algorithm the number of output symbols that phase j produces should match the number of input symbols of phase $j+1$. The ratio between the input of phase $j+1$ and output of phase j is:

$$\frac{(K-j)\binom{K}{j+1}}{j\binom{K}{j+1}} = \frac{(K-j)}{j}.$$

This means that $(K-j)$ repetition of phase j will produce the inputs needed by j repetition of phase $j+1$. In general, in order to have an integer number for repetitions, we multiply phase 1 by $K!$ (i.e., repeat it $K!$ times), phase 2 by $\frac{K!}{(K-1)}$, and so on. Therefore, phase j will be repeated $((j-1)!(K-j)!)K$ times which takes $((j-1)!(K-j)!)K\binom{K}{j}$ time slots. Since $(K-j)$ feedbacks from each time slot is sufficient, the number of feedbacks will be $((j-1)!(K-j)!)K\binom{K}{j}(K-j)$. For a successive decoding or order-j symbols, all the higher order symbols must be decoded successfully. Therefore, instead of having delayed CSIT at all time instants from all users, the minimum probability of delayed CSIT is the number of feedbacks from phase j to K divided by the whole number of time slots multiplied by the number of users,

$$\begin{aligned}\lambda_D^{min}(K,j) &= \frac{\sum_{i=j}^{K}(i-1)!(K-i)!K\binom{K}{i}(K-i)}{\sum_{i=j}^{K}(i-1)!(K-i)!K\binom{K}{j+1}K} \\ &= 1 - \frac{K-j+1}{K\sum_{i=j}^{K}\frac{1}{i}}.\end{aligned}$$

■

In this case (i.e., $\lambda_N \leq \frac{\lambda_D}{\sum_{j=2}^{K}\frac{1}{j}}$), the $2^K - K - 1$ inequalities having $\sum_i d_i$ (summation with equal weights) in the left-hand side become inactive and the remaining $\sum_{j=1}^{K} j!\binom{K}{j}$ inequalities are active which construct $\sum_{j=1}^{K} j!\binom{K}{j}$ hyperplanes in R_+^K. The region has $2^K - 1$ corner points. In other words, if the coordinates of a point are shown as $(p_1, p_2, \ldots, p_K)$, there are $\binom{K}{j}$ $(j \in [1:K])$ points where j of its K coordinates are $\frac{1+\lambda_P\sum_{i=2}^{j}\frac{1}{i}}{\sum_{i=1}^{j}\frac{1}{i}}$ and the remaining $K-j$ coordinates are λ_P. The region for the 3 user broadcast channel and the achievable scheme are shown in figure 6 and figure 7, respectively. The

$$\overbrace{\hphantom{P \cdots P N D \cdots N}}^{11N_1N_2}$$

P	...	P	N	D	...	N
P	...	P	D	N	...	N
P	...	P	D	D	...	N

$\underbrace{\hphantom{P \cdots P}}_{11M_1N_2}$ ZFBF $\quad$ $\underbrace{\hphantom{N D \cdots N}}_{11(N_1N_2 - M_1N_2) \geq 11M_2N_1}$ MAT

Fig. 7: Achievable scheme in case B for 3 user BC

achievable scheme is based on a concatenation of ZFBF and MAT as follows. For the $\binom{K}{j}$ corner points, we write

$$\lambda_P = \frac{M_1}{N_1}, \lambda_D = \frac{M_2}{N_2}, \lambda_D^{min}(j,1) = \frac{m}{n} \tag{53}$$

where m, n, M_i and N_i $(i = 1, 2)$ are integers. Making a common denominator between λ_P and λ_D we have

$$\lambda_P = \frac{nM_1N_2}{nN_1N_2}, \lambda_D = \frac{nN_1M_2}{nN_1N_2}. \tag{54}$$

We construct nN_1N_2 time slots where the CSIT of each user can be Perfect (P) or Delayed (D) in nM_1N_2 or nN_1M_2 time slots, respectively. In the first nM_1N_2 time slots, ZFBF is carried out. In the remaining $n(N_1N_2 - M_1N_2)$ time slots, j-user MAT algorithm is done. At each time slot of the ZFBF part, 1 interference-free symbol is received by each user and in the MAT part, $\frac{n(N_1N_2 - M_1N_2)}{1+\frac{1}{2}+\cdots+\frac{1}{j}}$ symbols are sent to each of the users in subset S (with $|S| = j$) where S depends on the corner point of interest. In order to do the MAT algorithm in the second part, the minimum probability of delayed CSIT should be met

$$nN_1M_2 \geq \lambda_D^{min}(j)n(N_1N_2 - M_1N_2) \tag{55}$$

Dividing both sides by nN_1N_2,

$$\lambda_D \geq \lambda_D^{min}(j,1)(1 - \lambda_P) = \lambda_D^{min}(j,1)(\lambda_D + \lambda_N) \tag{56}$$

which results in

$$\lambda_N \leq \frac{\lambda_D}{\sum_{i=2}^{j} \frac{1}{i}}. \tag{57}$$

Since it should be valid for all j, we have

$$\lambda_N \leq \frac{\lambda_D}{\sum_{i=2}^{K} \frac{1}{i}}. \tag{58}$$

which is the condition assumed in this case.

Finally, through an example, we show that the bounds in Theorem 2 can be tight. Consider the pattern shown in figure 8. According to sections III and V, the DoF region has the following outer bound

$$0 \leq d_1, d_2, d_3 \leq 1 \quad , \quad d_1 + d_2 \leq \frac{3}{2} \tag{59}$$

$$2d_1 + d_2 + 2d_3 \leq 3 \tag{60}$$

$$d_1 + 2d_2 + 2d_3 \leq 3. \tag{61}$$

The achievable point $(d_1, d_2, d_3) = (\frac{1}{2}, \frac{1}{2}, \frac{3}{4})$ makes the inequalities in (60) and (61) tight therefore, it is on the boundary of DoF region. This point is achievable as shown in figure 9 where the receivers are called A,B and C. Symbols are shown in red where the transmitter has perfect CSIT and those received signals that are not important in the achievability scheme are shown as "...".

P	N
P	N
N	P

Fig. 8: An example.

u_C	$L_2(u_C, u_A)$	v_A	...
v_C	...	v_B	$L_3(v_C, u_B)$
$L_1(u_C, v_C, w_C)$	u_C	...	v_C

Fig. 9: The achievable scheme for the boundary point $(\frac{1}{2}, \frac{1}{2}, \frac{3}{4})$.

VII. Two user MIMO

In previous sections, the K-user MISO BC was considered. The general MIMO BC is more challenging due to the mismatch between the number of receive antennas[6]. In this section, we consider a two user MIMO BC where each user is equipped with N_k $(k \in [1:2])$ antennas and a base station with $M (\geq N_1 + N_2)$ antennas wishes to send two independent messages W_1 and W_2 to their corresponding receivers. The received signal of user k is given by

$$\mathbf{Y}_k(t) = \mathbf{H}_k^H(t)\mathbf{X}(t) + \mathbf{W}_k(t) \ , \ k \in [1:2] \ , \ t \in [1:n] \tag{62}$$

where the channel matrices are assumed to be full rank almost surely. We assume that the CSI of a particular user is either instantaneously Perfect (P) or Not known (N) resulting in the four possible states PP, PN, NP and NN with corresponding probabilities $\lambda_{PP}, \lambda_{PN}, \lambda_{NP}$ and λ_{NN}. Let $Y_{i,j}$ denote the received signal at the j^{th} antenna of user i $(i \in [1:2], j \in [1:N_i])$. Without loss of generality, we assume $N_1 \geq N_2$. An outer bound on the DoF region is provided in Theorem 3 and its achievability is discussed afterwards.

Theorem 3. An outer bound for the DoF region of the channel in (62) is given by

$$\frac{d_1}{N_1} + \frac{d_2}{N_2} \leq 1 + \lambda_{PP} + \lambda_{NP} = 1 + \lambda_P^2 \tag{63}$$

$$d_1 + d_2 \leq N_1 + N_2(\lambda_{PP} + \lambda_{PN}) = N_1 + N_2\lambda_P^1 \tag{64}$$

Proof: By enhancing user 1 with the message of user 2,

[6]It is important to note that with different number of antennas, as stated in [18], the dimensions of useful signals and interference signals are not the same in contrast to the symmetric case. Furthermore, the users have different capabilities of decoding which must be taken into account in the achievability schemes.

Fano's inequality (ignoring $n\epsilon_n$) results in

$$nR_1 \le I(W_1; \mathbf{Y}_1^n|\Omega^n, W_2) = h(\mathbf{Y}_1^n|\Omega^n, W_2) - \underbrace{h(\mathbf{Y}_1^n|\Omega^n, W_1, W_2)}_{no(\log P)} \quad (65)$$

$$nR_2 \le I(W_2; \mathbf{Y}_2^n|\Omega^n) = \underbrace{h(\mathbf{Y}_2^n|\Omega^n)}_{\le nN_2 \log P} - h(\mathbf{Y}_2^n|\Omega^n, W_2). \quad (66)$$

Ignoring $o(\log P)$, we have

$$\begin{aligned} & n(N_2R_1 + N_1R_2) \\ & \le nN_1N_2 \log P + N_2h(\mathbf{Y}_1^n|\Omega^n, W_2) - N_1h(\mathbf{Y}_2^n|\Omega^n, W_2) && (67) \\ & \le nN_1N_2 \log P + \sum_{i=1}^{N_1} h(\Psi_i^{N_2}(\Gamma_{N_1})|F) - N_1h(\mathbf{Y}_2^n|F) && (68) \\ & = nN_1N_2 \log P + \sum_{i=1}^{N_1} \Big[h(\Psi_i^{N_2}(\Gamma_{N_1})|F) - h(\mathbf{Y}_2^n|F)\Big] && (69) \\ & \le nN_1N_2 \log P + nN_1N_2(\lambda_{PP} + \lambda_{NP}) \log P && (70) \end{aligned}$$

where in (68), $F = \{\Omega^n, W_2\}$ and lemma 1 has been applied with Γ_{N_1} denoting the N_1 elements of $\mathbf{Y}_1^n$ (i.e., $Y_{1,[1:N_1]}^n$) and $m = N_1 - N_2$. Applying the same procedure of section III and lemma 2 to each term of the summation in (69) results in (70). By dividing both sides of (70) by $n \log P$ and taking the limit $n, P \to \infty$, (63) is obtained.

For the inequality in (64), we have

$$\begin{aligned} nR_1 & \le I(W_1; \mathbf{Y}_1^n|\Omega^n) \\ & = I(W_1; Y_{1,[1:N_2]}^n|\Omega^n) + I(W_1; Y_{1,[N_2+1:N_1]}^n|\Omega^n, Y_{1,[1:N_2]}^n) \\ & = h(Y_{1,[1:N_2]}^n|\Omega^n) - h(Y_{1,[1:N_2]}^n|\Omega^n, W_1) \\ & \quad + h(Y_{1,[N_2+1:N_1]}^n|\Omega^n, Y_{1,[1:N_2]}^n) \\ & \quad - h(Y_{1,[N_2+1:N_1]}^n|\Omega^n, Y_{1,[1:N_2]}^n, W_1) \\ & \le \underbrace{h(Y_{1,[1:N_2]}^n)}_{\le nN_2 \log P} - h(Y_{1,[1:N_2]}^n|\Omega^n, W_1) + \underbrace{h(Y_{1,[N_2+1:N_1]}^n)}_{\le n(N_1-N_2) \log P} \\ & \quad - \underbrace{h(Y_{1,[N_2+1:N_1]}^n|\Omega^n, Y_{1,[1:N_2]}^n, W_1, W_2)}_{no(\log P)} && (71) \\ & \le nN_1 \log P - h(Y_{1,[1:N_2]}^n|\Omega^n, W_1) - no(\log P) && (72) \end{aligned}$$

where in (71), we used the fact that conditioning reduces the entropy. We enhance user 2 with the message of user 1. Therefore,

$$nR_2 \le I(W_2; \mathbf{Y}_2^n|\Omega^n, W_1) = h(\mathbf{Y}_2^n|\Omega^n, W_1) - \underbrace{h(\mathbf{Y}_2^n|\Omega^n, W_1, W_2)}_{no(\log P)}. \quad (73)$$

By adding (72) and (73), we get

$$nR_1 + nR_2 \le \underbrace{h(\mathbf{Y}_2^n|\Omega^n, W_1) - h(Y_{1,[1:N_2]}^n|\Omega^n, W_1)}_{\le nN_2(\lambda_{PP}+\lambda_{PN}) \log P} + nN_1 \log P - 2no(\log P) \quad (74)$$

where the same procedure of section III has been applied to the difference in (74). Therefore,

$$d_1 + d_2 \le N_1 + N_2(\lambda_{PP} + \lambda_{PN}) = N_1 + N_2\lambda_P^1. \quad (75)$$

■

In the sequel, we show that when $\lambda_{PN} \le \frac{N_2}{N_1}\lambda_{NP}$, the outer bound, which is defined by (63) and (64), is tight. Specifically, we show the achievability of the inner bound defined by the following inequalities

$$\frac{d_1}{N_1} + \frac{d_2}{N_2} \le 1 + \lambda_P^2 \quad (76)$$

$$d_1 + d_2 \le N_1 + N_2(\lambda_{PP} + \min(\lambda_{PN}, \frac{N_2}{N_1}\lambda_{NP})). \quad (77)$$

It is obvious that when $\lambda_{PN} \le \frac{N_2}{N_1}\lambda_{NP}$, the inner bound coincides with the outer bound. We consider a block of n (sufficiently large) time instants. In this block, there are $n\lambda_{PN}$ time instants in the PN state (i.e., where the CSI of user 1 is perfectly known and CSI of user 2 is unknown), $n\lambda_{NP}$ time instants in the NP state, $n\lambda_{PP}$ time instants in the PP state and $n\lambda_{NN}$ time instants in the NN state. Without loss of generality, we assume n is chosen in such a way that all these numbers are integers. From now on, whenever it is said that N symbols are sent orthogonal to the matrix $\mathbf{H}$, it is meant that these N symbols are precoded by a matrix whose columns are chosen from the null space of $\mathbf{H}^H$.

The following achievable schemes are based on a simple interference cancellation scheme. In other words, if at each of the m time instants in the PN state, N_1 private symbols are sent to user 1 and N_2 private symbols are sent (orthogonal to the channel of user 1) to user 2, user 2 needs to get rid of nN_1 interfering symbols from user 1 to decode its own symbols. If we pick $m\frac{N_1}{N_2}$ time instants in the NP state, at each of these time instants, N_2 interfering symbols can be sent to user 2 and since these interfering symbols are already known at user 1, N_1 new private symbols can be sent (orthogonal to the channel of user 2) to user 1. This could be viewed as a generalization of the $S_3^{\frac{3}{2}}$ [9] to the MIMO case where the mismatch between the number of receiving antennas across the users is taken into account. The achievability is divided into two scenarios.

A. $N_1\lambda_{PN} \le N_2\lambda_{NP}$

In this case, the region is shown in figure 10.

A.1: When $N_1 - N_2 + N_2\lambda_P^1 \le N_1\lambda_P^2$, the region (figure 10 (a)) has the corner points $A_1(N_1, N_2\lambda_P^1)$ and $A_2(N_1 - N_2 + N_2\lambda_P^1, N_2)$.

The achievability of A_1 is as follows.

Phase 1: At each of the $n\lambda_{PN}$ time instants, N_1 and N_2 private symbols are sent to user 1 and user 2, respectively. These N_2 private symbols are sent orthogonal to $\mathbf{H}_1(t)$. Therefore, user 1 receives its intended $nN_1\lambda_{PN}$ symbols and user 2 receives $n(N_1 + N_2)\lambda_{PN}$ symbols. User 1 can decode its symbols immediately, while user 2 has to get rid of $nN_1\lambda_{PN}$ interfering symbols. **Phase 2**: Among the $n\lambda_{NP}$ time instants in the NP state, $\frac{N_1}{N_2}n\lambda_{PN}(\le n\lambda_{NP})$ time instants are selected. At each of these selected time instants, N_2 interfering symbols of phase 1 are sent to user 2 and N_1

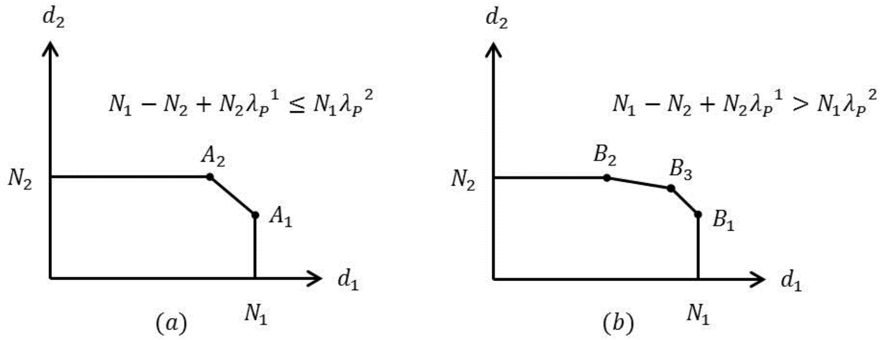

Fig. 10: The DoF region when $N_1\lambda_{PN} \le N_2\lambda_{NP}$.

new private symbols are sent to user 1. These N_1 private symbols are sent orthogonal to $\mathbf{H}_2(t)$. User 2 receives the $nN_1\lambda_{PN}$ interfering symbols which enables it to decode its private symbols in phase 1. The interfering symbols of user 2 are already known at user 1, therefore, user 1 can successfully decode its private symbols in this phase.

Phase 3: In the remaining time instants in the NP state (i.e., $n\lambda_{NP} - \frac{N_1}{N_2}n\lambda_{PN}$) and all the $n\lambda_{NN}$ time instants, N_1 private symbols are sent to user 1.

Phase 4: In all the $n\lambda_{PP}$ time instants, N_1 and N_2 private messages orthogonal to $\mathbf{H}_2(t)$ and $\mathbf{H}_1(t)$, respectively are sent to user 1 and user 2.

Therefore, user 1 and user 2 can, respectively, decode nN_1 and $nN_2(\lambda_{PP} + \lambda_{PN})$ private symbols in the block of n time instants which achieves the first corner point $(N_1, N_2\lambda_P^1)$.

The achievability of A_2 is as follows.

Phase 1: Among the $n\lambda_{NP}$ time instants in the NP state, $n\frac{N_1}{N_2}\lambda_{PN}$ time instants are selected. At each of these selected time instants, N_2 and N_1 private symbols are sent to user 2 and user 1, respectively. These N_1 private symbols are sent orthogonal to $\mathbf{H}_2(t)$. Therefore, user 2 can decode $nN_1\lambda_{PN}$ private symbols and user 1 receives $n(N_1 + N_2)\frac{N_1}{N_2}\lambda_{PN}$ symbols of which $nN_1\lambda_{PN}$ symbols are interferers.

Phase 2: At each of the $n\lambda_{PN}$ time instants, N_1 interfering symbols in phase 1 are sent to user 1 and N_2 private symbols to user 2. These N_2 private symbols are sent orthogonal to $\mathbf{H}_1(t)$. Therefore, user 1 is able to decode its private symbols in phase 1.

Phase 3: There are $n\lambda_{NP} - \frac{N_1}{N_2}n\lambda_{PN}$ remaining time instants in the NP state. $n\lambda_{NN}\frac{(N_1-N_2)}{N_2}$ of them are selected (note that $n\lambda_{NN}\frac{(N_1-N_2)}{N_2} \le n\lambda_{NP} - \frac{N_1}{N_2}n\lambda_{PN}$ due to the condition in the figure 10(a)). At each of these selected time instants, N_2 and N_1 private symbols are sent to user 2 and user 1, respectively. These N_1 private symbols are sent orthogonal to $\mathbf{H}_2(t)$. Therefore, user 1 has to get rid of $n\lambda_{NN}(N_1 - N_2)$ interfering symbols.

Phase 4: At each of the $n\lambda_{NN}$ time instants, N_2 private symbols are sent to user 2 and $N_1 - N_2$ interfering symbol from phase 3 are sent to user 1. The interfering symbols are already known at user 2, therefore user 2 successfully decodes its symbols. User 1, having N_1 antennas, is capable of decoding all the sent symbols in this phase.

Phase 5: In the remaining time instants in the NP states, N_2 and $N_1 - N_2$ private symbols are sent to user 2 and user 1, respectively. These $N_1 - N_2$ private symbols are sent orthogonal to $\mathbf{H}_2(t)$.

Phase 6: The same as phase 4 for the achievability of A_1

Therefore, user 1 and user 2 can, respectively, decode $n(N_1 - N_2 + N_2\lambda_P^1)$ and nN_2 private symbols in the block of n time instants which achieves the second corner point.

A.2: When $N_1 - N_2 + N_2\lambda_P^1 > N_1\lambda_P^2$, the region (figure 10 (b)) has the corner points $B_1(N_1, N_2\lambda_P^1)$, $B_2(N_1\lambda_P^2, N_2)$ and $B_3(N_1 - \frac{N_1N_2(\lambda_P^2-\lambda_P^1)}{N_1-N_2}, \frac{N_1N_2\lambda_P^2-N_2^2\lambda_P^1}{N_1-N_2})$.

The achievability of B_1 is the same as that of A_1 and the achievability of B_2 is as follows.

Phase 1 and 2: Similar to the phase 1 and phase 2 in the achievability of A_2.

Phase 3: There are $n\lambda_{NP} - \frac{N_1}{N_2}n\lambda_{PN}$ remaining time instants in the NP state. At each of these remaining time instants, N_2 and N_1 private symbols are sent to user 2 and user 1, respectively. These N_1 private symbols are sent orthogonal to $\mathbf{H}_2(t)$. Therefore, user 1 has to get rid of $nN_2\lambda_{NP} - nN_1\lambda_{PN}$ interfering symbols.

Phase 4: There are $n\lambda_{NN}$ time instant in the NN state. $\frac{nN_2\lambda_{NP}-nN_1\lambda_{PN}}{N_1-N_2}$ of them are selected (note that $n\lambda_{NN} > \frac{nN_2\lambda_{NP}-nN_1\lambda_{PN}}{N_1-N_2}$ due to the condition in the figure 10(b)) At each of these selected time instants, N_2 private symbols are sent to user 2 and $N_1 - N_2$ interfering symbols from phase 3 are sent to user 1. Therefore, with N_1 antennas, user 1 can decode its private symbols in phase 3. Since these interfering symbols are already known at user 2, it can successfully decode its N_2 private symbols in this phase.

Phase 5: In the remaining time instants in the NN states, N_2 private symbols are sent to user 2.

Phase 6: The same as phase 4 in the achievability A_1.

Therefore, user 1 and user 2 can, respectively, decode $nN_1\lambda_P^2$ and nN_2 private symbols in the block of n time instants which achieves the second corner point.

The achievability of B_3 follows the same lines as the achievability of B_2 except that in phase 5, in the remaining NN time instants, instead of sending N_2 private symbols to user 2, N_1 private symbols are sent to user 1.

In conclusion, the outer bound in Theorem 3 is the optimal DoF region in this case (i.e., $N_1\lambda_{PN} \le N_2\lambda_{NP}$).

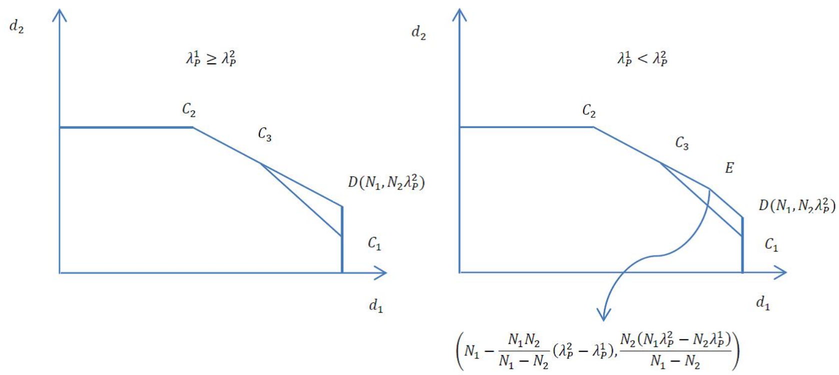

Fig. 11: The achievable DoF region (i.e., the inner bound) and the outer bound when $N_1\lambda_{PN} > N_2\lambda_{NP}$.

B. $N_1\lambda_{PN} > N_2\lambda_{NP}$

In this case, the achievable region has three corner points $C_1(N_1, N_2\lambda_{PP} + \frac{N_2^2}{N_1}\lambda_{NP})$, $C_2(N_1\lambda_P^2, N_2)$ and $C_3(N_1 - N_2\lambda_{NP}, N_2\lambda_P^2 + \frac{N_2^2}{N_1}\lambda_{NP})$. This is shown in figure 11 along with the outer bound where the outer bound has two corner points (C_2 and D) when $\lambda_P^1 \geq \lambda_P^2$ and three corner points otherwise (i.e., C_2, E and D).

The achievability of C_1 is as follows.

Phase 1: There are $n\lambda_{PN}$ time instants in the PN state and $n\frac{N_2}{N_1}\lambda_{NP}(\leq n\lambda_{PN})$ of them are selected. At each of these selected time instants, N_1 and N_2 private symbols are sent to user 1 and user 2, respectively. These N_2 symbols are sent orthogonal to $\mathbf{H}_1(t)$. Therefore, user 1 receives its intended $nN_2\lambda_{NP}$ symbols and user 2 receives $n(N_1+N_2)\frac{N_2}{N_1}\lambda_{NP}$ symbols. User 1 can decode its symbols immediately, while user 2 has to get rid of $nN_2\lambda_{NP}$ interfering symbols.

Phase 2: At each of the $n\lambda_{NP}$ time instants in the NP state, N_2 interfering symbols of phase 1 are sent to user 2 and N_1 private symbols are sent to user 1. These N_1 symbols are sent orthogonal to $\mathbf{H}_2(t)$. User 2 receives the $nN_2\lambda_{NP}$ interfering symbols which enables it to decode its private symbols in phase 1. Since these interfering symbols are already known at user 1, it can successfully decode its N_1 private symbols in this phase.

Phase 3: In the remaining time instants in the PN state (i.e., $n\lambda_{PN} - \frac{N_2}{N_1}n\lambda_{NP}$) and all the $n\lambda_{NN}$ time instants, N_1 private symbols are sent to user 1.

Phase 4: The same as phase 4 in the achievability of A_1.

Therefore, user 1 and user 2 can, respectively, decode nN_1 and $n(N_2\lambda_{PP} + \frac{N_2^2}{N_1}\lambda_{NP})$ private symbols in the block of n time instants which achieves the first corner point.

The achievability of C_2 is as follows.

Phase 1: At each of $n\lambda_{NP}$ time instants, N_2 and N_1 private symbols are sent to user 2 and user 1, respectively. These N_1 symbols are sent orthogonal to $\mathbf{H}_2(t)$. Therefore, user 2 can decode its intended $nN_2\lambda_{NP}$ symbols and user 1 receives $n(N_1+N_2)\lambda_{NP}$ symbols of which $nN_2\lambda_{NP}$ are interferes.

Phase 2: Among the $n\lambda_{PN}$ time instants in the PN state, $n\frac{N_2}{N_1}\lambda_{NP}(\leq n\lambda_{PN})$ time instants are selected. At each of these selected time instants, N_1 interfering symbols of phase 1 are sent to user 2 and N_2 private symbols are sent to user 2. These N_2 symbols are sent orthogonal to $\mathbf{H}_1(t)$. Therefore, user 1 can decode its private symbols in phase 1.

Phase 3: In the remaining time instants in the PN state (i.e., $n\lambda_{PN} - \frac{N_2}{N_1}n\lambda_{NP}$) and all the $n\lambda_{NN}$ time instants, N_2 private symbols are sent to user 2.

Phase 4: The same as phase 4 in the achievability of A_1.

Therefore, user 1 and user 2 can, respectively, decode $nN_1\lambda_P^2$ and nN_2 private symbols in the block of n time instants which achieves the second corner point.

The achievability of C_3 follows the same lines as the achievability of C_2 with the difference that in phase 3, in the remaining time instants in the PN state and all the $n\lambda_{NN}$ time instants, instead of sending N_2 private symbols to user 2, N_1 private symbols are sent to user 1.

As an example, figure 12 shows the achievability of the corner point B_3 in figure 10(b). In this example, $\lambda_{PN} = \lambda_{PP} = \frac{1}{6}, \lambda_{NP} = \lambda_{NN} = \frac{1}{3}$. u and v are private symbols from (independently) Gaussian encoded codewords for user 1 and user 2, respectively and $n = 12$. When the CSI of a user is known at the transmitter, it is shown in red.

VIII. Conclusion

Given the marginal probabilities of CSIT, an outer bound was derived for the DoF region of the K-user MISO BC with alternating/hybrid CSIT . This outer bound was shown to be achievable by specific CSIT patterns in certain regions. A set of inequalities was provided based on the joint CSIT distribution which shows that in general, the DoF region of the K-user MISO BC (when $K \geq 3$) cannot be characterized completely by the marginal probabilities. Finally, an outer bound for the DoF region of a two user MIMO BC in which the CSIT of a user is either perfect or unknown was derived which was shown to be tight in some scenarios.

Appendix A
An alternative proof of $\sum_{i=1}^{K}\frac{d_i}{i} \leq 1 + \sum_{i=2}^{K}\frac{\sum_{r=1}^{i-1}\lambda_P^r}{i(i-1)}$

The proof is based on the approach used in [3], therefore the following definitions are necessary. The channel vector of user k at time instant t can be written as

$$\mathbf{H}_k(t) = \widehat{\mathbf{H}}_k(t) + \widetilde{\mathbf{H}}_k(t) \tag{78}$$

where $\widehat{\mathbf{H}}_k(t)$ and $\widetilde{\mathbf{H}}_k(t)$ are the estimate of the channel and estimation error with distributions $CN(\mathbf{0}, (1-\sigma_k^2(t))\mathbf{I})$ and $CN(\mathbf{0}, \sigma_k^2(t)\mathbf{I})$, respectively. The variance of error is

$$\sigma_k^2(t) = E\left[\|\widetilde{\mathbf{H}}_k(t)\|^2\right].$$

As observed from the above, although the channel is assumed stationary, the estimate is a non-stationary process meaning that the quality of estimation varies over time. The quality of CSIT for user k at time instant t is

$$\alpha_k(t) = -\lim_{P\to\infty} \frac{\log\left(\sigma_k^2(t)\right)}{\log P}. \tag{79}$$

From the results of [19], if the rate of feedback scales linearly with $\log P$ (or equivalently, the variance of estimation error decrease as $o(P^{-1})$ or faster), perfect CSIT multiplexing gain can be obtained. Therefore, the effective range of $\alpha_k(t)$ will be $[0,1]$ where in terms of DoF, $\alpha_k(t) = 1$ could be interpreted as perfect CSIT of user k at time instant t. We also define $\hat{\Omega}^t$ as the set of all channel estimates up to time instant t. Again, for simplicity, we show the inequalities for a fixed permutation of the users while the results could be easily extended to any arbitrary permutations. As in part A of the first proof, the same channel improvement is done here. The only difference is that we assume the users not only have perfect global CSIR, but also they know the channel estimates at the transmitter. From the chain rule of entropies, each of the terms in the summation in (9) can be written as

$$\sum_{t=1}^{n}\left[\frac{h(Y_{[1:i]}(t)|W_{[1:i-1]}, Y_{[1:i]}^{t-1}, \Omega^t, \hat{\Omega}^t)}{i} - \frac{h(Y_{[1:i-1]}(t)|W_{[1:i-1]}, Y_{[1:i-1]}^{t-1}, \Omega^t, \hat{\Omega}^t)}{i-1}\right]. \tag{80}$$

By adding Y_i^{t-1} to the conditions of the second entropy, (80) will be increased. Therefore,

$$\begin{aligned}\sum_{i=1}^{K}\frac{nR_i}{i} \le{}& \overbrace{h(Y_1^n|\Omega^n, \hat{\Omega}^n)}^{\le n\log P} \\ &+ \sum_{i=2}^{K}\sum_{t=1}^{n}\left[\frac{h(Y_{[1:i]}(t)|U_{i,t}, \Omega(t))}{i} - \frac{h(Y_{[1:i-1]}(t)|U_{i,t}, \Omega(t))}{i-1}\right] + no(\log P). \end{aligned} \tag{81}$$

where $U_{i,t} = (W_{[1:i-1]}, Y_{[1:i]}^{t-1}, \Omega^{t-1}, \hat{\Omega}^t)$ and $\Omega(t)$ is the global CSIR at time instant t. In what follows, we find an upper bound for the term in the brackets of (81). Following the same approach as in [3], (82) to (89) are obtained in which we have the Markov chain $\mathbf{X}(t) \leftrightarrow U_{i,t} \leftrightarrow \hat{\Omega}(t) \leftrightarrow \Omega(t)$. In (85), we have written the signals of the users in terms of their concatenated channels and noise terms. (87) is the application of extremal inequality [20], [21] where the Gaussian distribution maximizes a specific difference between two differential entropies. The last inequality (89) comes from (101) in [22], in which

$$\Sigma^2 = \text{diag}\left(\sigma_{[1:i-1]}^2(t)\right).$$

Therefore, we can write

$$\begin{aligned}\sum_{i=1}^{K}\frac{nR_i}{i} &\le n\log P + \sum_{i=2}^{K}\sum_{t=1}^{n}\left[-\frac{\log\det\left(\Sigma^2\right)}{i(i-1)} + o(\log P)\right] \\ &\quad + no(\log P) \\ &= n\log P + \sum_{i=2}^{K}\frac{\sum_{t=1}^{n}[\alpha_1(t)+\cdots+\alpha_{i-1}(t)]}{i(i-1)}\log P \\ &\quad + nKo(\log P). \end{aligned} \tag{90}$$

Since the channel is degraded and D is replaced with N, the CSIT is either P or N. Therefore, the α's are either 1 with probability λ_P^i or 0 otherwise. Hence, for n large enough, we have

$$\lim_{n\to\infty}\sum_{t=1}^{n}[\alpha_1(t)+\cdots+\alpha_{i-1}(t)] = n\sum_{r=1}^{i-1}\lambda_P^r$$

which results in

$$\sum_{i=1}^{K}\frac{nR_i}{i} \le n\log P + \sum_{i=2}^{K}\sum_{r=1}^{i-1}\frac{n\lambda_P^r}{i(i-1)}\log P + nKo(\log P) \tag{91}$$

at large n. Dividing both sides by $n\log P$ and taking the limit of (91) as $n, P\to\infty$, we get

$$\sum_{i=1}^{K}\frac{d_i}{i} \le 1 + \sum_{i=2}^{K}\frac{\sum_{r=1}^{i-1}\lambda_P^r}{i(i-1)}. \tag{92}$$

It is obvious that the same approach can be applied to any other permutations of $(1,2,\ldots,K)$.

References

[1] B. Clerckx and C. Oestges, *MIMO Wireless Networks, 2nd Edition*. Academic Press, 2013.

[2] M. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, July 2012.

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

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

[5] J. Chen and P. Elia, "Degrees-of-freedom region of the MISO broadcast channel with general mixed-CSIT," available on arxiv:1205.3474.

[6] P. de Kerret, X. Yi, and D. Gesbert, "On the degrees of freedom of the K-user time correlated broadcast channel with delayed CSIT," in *IEEE ISIT*, July 2013, pp. 624–628.

[7] C. Hao and B. Clerckx, "Imperfect and unmatched CSIT is still useful for the frequency correlated MISO broadcast channel," in *IEEE ICC*, Budapest, Hungary, June 2013, pp. 3181–3186.

[8] ——, "MISO broadcast channel with imperfect and (un)matched CSIT in the frequency domain: DoF region and transmission strategies," in *IEEE PIMRC*, Sep. 2013, pp. 1–6.

[9] R. Tandon, S. Jafar, S. Shamai Shitz, and H. Poor, "On the synergistic benefits of alternating CSIT for the MISO broadcast channel," *IEEE Trans. Inf. Theory.*, vol. 59, no. 7, pp. 4106–4128, July 2013.

[10] K. Mohanty and M. Varanasi, "On the DoF region of the K-user MISO broadcast channel with hybrid CSIT." available on arXiv:1312.1309.

[11] S. Amuru, R. Tandon, and S. Shamai, "On the degrees-of-freedom of the 3-user MISO broadcast channel with hybrid CSIT," in *IEEE ISIT*, 2014, pp. 2137–2141.

$$\max_{P_{U_{i,t}} P_{\mathbf{X}(t)|U_{i,t}}} \left[\frac{h(Y_{[1:i]}(t)|U_{i,t},\Omega(t))}{i} - \frac{h(Y_{[1:i-1]}(t)|U_{i,t},\Omega(t))}{i-1}\right] \tag{82}$$

$$\leq \max_{P_{U_{i,t}}} E_{U_{i,t}} \left[\max_{P_{\mathbf{X}(t)|U_{i,t}}} \left(\frac{h(Y_{[1:i]}(t)|U_{i,t}=U,\Omega(t))}{i} - \frac{h(Y_{[1:i-1]}(t)|U_{i,t}=U,\Omega(t))}{i-1}\right)\right] \tag{83}$$

$$= \max_{P_{U_{i,t}}} E_{U_{i,t}} \left[\max_{P_{\mathbf{X}(t)|U_{i,t}}} E_{\Omega(t)|U_{i,t}} \left(\frac{h(Y_{[1:i]}(t)|U_{i,t}=U,\Omega(t)=\mathbf{H})}{i} - \frac{h(Y_{[1:i-1]}(t)|U_{i,t}=U,\Omega(t)=\mathbf{H})}{i-1}\right)\right] \tag{84}$$

$$= \max_{P_{U_{i,t}}} E_{U_{i,t}} \left[\max_{P_{\mathbf{X}(t)|U_{i,t}}} E_{\Omega(t)|\hat{\Omega}(t)} \left(\frac{h(\mathbf{H}_{[1:i]}(t)\mathbf{X}(t)+\mathbf{W}_{[1:i]}(t)|U_{i,t}=U)}{i} - \frac{h(\mathbf{H}_{[1:i-1]}(t)\mathbf{X}(t)+\mathbf{W}_{[1:i-1]}(t)|U_{i,t}=U)}{i-1}\right)\right] \tag{85}$$

$$= \max_{P_{U_{i,t}}} E_{U_{i,t}} \left[\max_{\boldsymbol{C}:\boldsymbol{C}\succeq 0, tr(\boldsymbol{C})\leq P} \max_{\substack{P_{\mathbf{X}(t)|U_{i,t}} \\ Cov(\mathbf{X}(t)|U_{i,t})\preceq \boldsymbol{C}}} E_{\Omega(t)|\hat{\Omega}(t)} \left(\frac{h(\mathbf{H}_{[1:i]}(t)\mathbf{X}(t)+\mathbf{W}_{[1:i]}(t)|U)}{i} - \frac{h(\mathbf{H}_{[1:i-1]}(t)\mathbf{X}(t)+\mathbf{W}_{[1:i-1]}(t)|U)}{i-1}\right)\right] \tag{86}$$

$$= \max_{P_{U_{i,t}}} E_{U_{i,t}} \left[\max_{\boldsymbol{C}:\boldsymbol{C}\succeq 0, tr(\boldsymbol{C})\leq P} E_{\Omega(t)|\hat{\Omega}(t)} \left(\frac{\log\det\left(I_i + \mathbf{H}_{[1:i]}(t)\mathbf{K}_*\mathbf{H}^H_{[1:i]}(t)\right)}{i} - \frac{\log\det\left(I_{i-1} + \mathbf{H}_{[1:i-1]}(t)\mathbf{K}_*\mathbf{H}^H_{[1:i-1]}(t)\right)}{i-1}\right)\right] \tag{87}$$

$$\leq E_{\hat{\Omega}(t)} \left[\max_{\boldsymbol{K}:\boldsymbol{K}\succeq 0, tr(\boldsymbol{K})\leq P} E_{\Omega(t)|\hat{\Omega}(t)} \left(\frac{\log\det\left(I_i + \mathbf{H}_{[1:i]}(t)\mathbf{K}\mathbf{H}^H_{[1:i]}(t)\right)}{i} - \frac{\log\det\left(I_{i-1} + \mathbf{H}_{[1:i-1]}(t)\mathbf{K}\mathbf{H}^H_{[1:i-1]}(t)\right)}{i-1}\right)\right] \tag{88}$$

$$\leq -\frac{\log\det\left(\Sigma^2\right)}{i(i-1)} + o(\log P) \tag{89}$$

[12] S. Lashgari, R. Tandon, and S. Avestimehr, "MISO broadcast channel with hybrid CSIT: Beyond two users," available on arXiv:1504.04615.

[13] N. Lee and R. Heath Jr., "Space-time interference alignment and degrees of freedom regions for the MISO broadcast channel with periodic CSI feedback," *IEEE Trans. Inf. Theory.*, vol. 60, no. 1, pp. 515–528, Jan. 2014.

[14] N. Lee, R. Tandon, and R. Heath Jr., "Distributed space-time interference alignment with moderately delayed CSIT," *IEEE Trans. Wireless Comm.*, vol. 14, no. 2, pp. 1048–1059, Feb. 2015.

[15] T. M. Cover and J. A. Thomas, *"Elements of Information Theory, second edition*. New York: Wiley-Intersicence, 2006.

[16] A. Gamal, "The feedback capacity of degraded broadcast channels (corresp.)," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 379 – 381, May 1978.

[17] A. Davoodi and S. Jafar, "Aligned image sets under channel uncertainty: Settling a conjecture by Lapidoth, Shamai and Wigger on the collapse of degrees of freedom under finite precision CSIT," available on arxiv:1403.1541.

[18] X. Yi, S. Yang, D. Gesbert, and M. Kobayashi, "The degree of freedom region of temporally correlated MIMO networks with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 494 –514, Jan. 2014.

[19] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045 –5060, Nov. 2006.

[20] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1839–1851, May 2007.

[21] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936 –3964, Sep. 2006.

[22] J. Chen and P. Elia, "Toward the performance vs. feedback tradeoff for the two-user MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, pp. 8336–8356, Dec. 2013.

Borzoo Rassouli received his M.Sc. degree in communication systems engineering from University of Tehran, Iran, in 2012. He is currently pursuing the Ph.D. degree in the Communication and Signal Processing Group, Imperial College, London, UK. His research interests are in the general area of information theory, wireless communications, detection and estimation theory.

Chenxi Hao received the B.Sc. and M.Sc. degrees from Beijing University of Posts and Telecommunications in 2010 and University of Southampton in 2011, respectively. He is currently pursuing the Ph.D. degree in the Communication and Signal Processing Group, Imperial College London. He is also with Beijing Samsung Telecom R&D Center. His current research interests lie in communication theory, network information theory and MIMO systems with limited feedback.

Bruno Clerckx received the M.S. and Ph.D. degrees in applied science from Universite catholique de Louvain, Belgium. He is now a Senior Lecturer (Associate Professor) at Imperial College London. He held visiting research positions at Stanford University and EURECOM and was with Samsung Electronics from 2006 to 2011. He actively contributed to 3GPP LTE/LTE-A and IEEE802.16m. He is the author or coauthor of two books on MIMO wireless communications and networks and numerous research papers, standard contributions and patents. He received the Best Student Paper Award at the IEEE Symposium on Communications and Vehicular Technology in 2002 and several awards from Samsung in recognition of special achievements. Dr. Clerckx has served as an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS and is currently an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

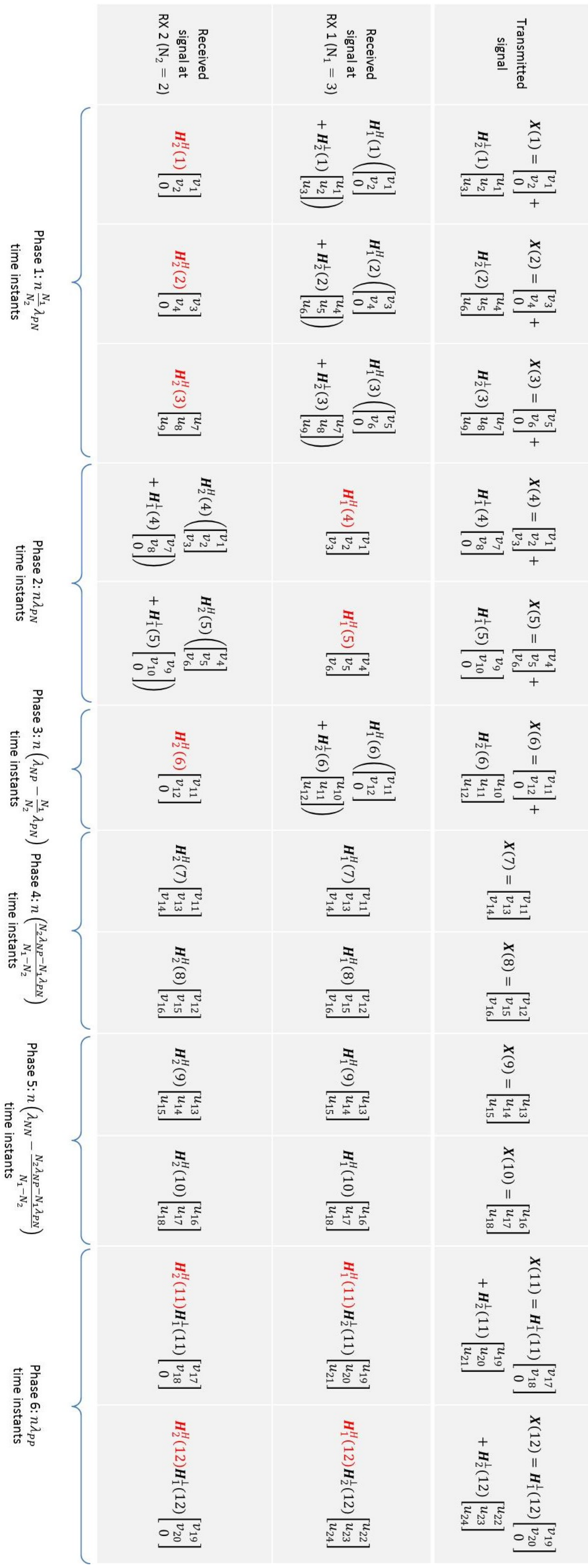

Fig. 12: An example for achieving the corner point $B_3(N_1 - \frac{N_1N_2(\lambda_P^2-\lambda_P^1)}{N_1-N_2}, \frac{N_1N_2\lambda_P^2-N_2^2\lambda_P^1}{N_1-N_2}) = (2, \frac{5}{3})$. In this example $\lambda_{PN} = \lambda_{PP} = \frac{1}{6}, \lambda_{NP} = \lambda_{NN} = \frac{1}{3}$.