# Transmitter Cooperation under Finite Precision CSIT: A GDoF Perspective

Arash Gholami Davoodi and Syed A. Jafar
Center for Pervasive Communications and Computing (CPCC)
University of California Irvine, Irvine, CA 92697
*Email: {gholamid, syed}@uci.edu*

**Abstract**

The benefits of partial and full transmitter cooperation are evaluated for a two user interference channel under finite precision channel state information at the transmitters (CSIT), using the generalized degrees of freedom (GDoF) metric. Under finite precision CSIT, the benefits of interference alignment are completely lost, so that the $X$ channel obtained by partial transmitter cooperation does no better than the underlying interference channels. Full transmitter cooperation produces a vector broadcast channel (BC) which has a strict GDoF advantage over partial cooperation (X channel) and whose GDoF are fully achieved by interference enhancement. [1]

## 1 Introduction

Cooperation is widely regarded as the panacea for countering interference in wireless networks. The benefits of cooperation are known to be quite powerful under ideal channel knowledge assumptions, but are not as well understood in the presence of channel uncertainty. This is especially critical for *transmitter* side cooperation, because the quality of channel state information at the transmitters (CSIT) is typically much more limited in practice. It is therefore of great interest to understand the fundamental limits of transmitter cooperation under finite precision CSIT.

In spite of much research activity aimed at limited CSIT settings (summarized in Section 3.2), a fundamental understanding of the *finite precision* CSIT setting has remained rather elusive. This is the case even from the very coarse, degrees of freedom (DoF) perspective. As a representative example, the 2005 conjecture of Lapidoth, Shamai and Wigger (in short, the LSW conjecture) which predicts a collapse of the degrees of freedom of a vector broadcast channel under finite precision CSIT [1], remained open for nearly a decade, and was finally settled (in the affirmative) only in 2014 [2].

In the settling of the LSW conjecture, there is cause for both hope and despair. On the one hand, it takes away some of the optimism behind transmitter cooperation, because it shows that the benefits of transmitter cooperation are entirely lost from a DoF perspective under finite precision CSIT. On the other hand, it does so by introducing a new tool for finite precision CSIT settings — a novel combinatorial argument limiting the size of aligned image sets, i.e., the sets of codewords that are distinguishable at one receiver but not at another receiver — under all possible (including non-linear) coding schemes. The new tool, if it can be generalized, offers hope of further refining our

[1]This work was supported in part by ONR grant N00014-15-1-2557, and by NSF grants CCF-1319104 and CCF-1161418. This work was presented in part at IEEE Globecom 2015.

understanding of the finite precision CSIT setting, beyond the coarse DoF perspective. Specifically, it points the way to the next logical step, a *generalized* degrees of freedom (GDoF) characterization, which is goal of this paper.

### 1.1 Generalized Degrees of Freedom (GDoF)

Much of the recent progress on the capacity of wireless networks has come about from a progressive refinement approach that pursues the path:

$$\text{DoF} \ \rightarrow \ \text{GDoF} \ \rightarrow O(1) \text{ gap} \ \rightarrow \ \text{constant gap} \ \rightarrow \ \text{capacity}$$

The coarse DoF metric serves as the starting point, but suffers from severe limitations, e.g., it essentially treats all non-zero channels as equally strong (capable of carrying one DoF each) in the high SNR limit. Distinctions in the strength of various signals, which are extremely important in practice, are essentially ignored in the DoF perspective. The GDoF perspective refines the picture by adopting a model that maintains the ratio of signal strengths in the dB scale (essentially the ratio of channel capacities) constant as the high SNR limit is approached. It is therefore able to explore weak and strong interference regimes, which are hidden in the DoF perspective, and offer insights into optimal schemes for those regimes. GDoF characterizations tend to be stepping stones to capacity characterizations within an O(1) gap, i.e., a gap that does not depend on SNR, but may depend on the channel realizations. The next progressive refinement goal tends to be a capacity characterization within a constant gap, i.e., a gap that does not depend on SNR or channel realizations. Not surprisingly, the ultimate refinement goal is the capacity itself.

Following this approach, since the DoF of the finite precision CSIT setting with transmitter cooperation are now settled, the logical next goal is to pursue a GDoF characterization.

## 2 System Model

### 2.1 Interference Channel: IC($W_{11}, W_{22}$)

As the underlying channel model, consider the 2 user interference channel defined by the input-output equations:

$$Y_1(t) = \sqrt{P^{\alpha_{11}}} G_{11}(t) X_1(t) + \sqrt{P^{\alpha_{12}}} G_{12}(t) X_2(t) + Z_1(t) \tag{1}$$

$$Y_2(t) = \sqrt{P^{\alpha_{21}}} G_{21}(t) X_1(t) + \sqrt{P^{\alpha_{22}}} G_{22}(t) X_2(t) + Z_2(t) \tag{2}$$

Here, over the $t^{th}$ channel use, $X_k(t)$ is the symbol sent from transmitter $k$, normalized so that it is subject to unit power constraint. $Y_k(t)$ is the symbol observed at receiver $k$. $Z_k(t) \sim \mathcal{N}(0,1)$ is the zero mean unit variance additive white Gaussian noise. $G_{ij}(t)$ is the channel coefficient from transmitter $j$ to receiver $i$, whose value is assumed to be bounded away from zero and infinity, i.e., there exist constants $0 < \Delta_1 < \Delta_2 < \infty$ such that $|G_{ij}(t)| \in [\Delta_1, \Delta_2]$, $\forall i, j \in \{1,2\}, \forall t \in \mathbb{N}$.

For ease of exposition, we will start with the assumption that all symbols take only real values. The results do extend to complex channels as well, as shown subsequently in this work.

The 2 user interference channel has messages $W_{kk}$ that originate at transmitter $k$ and are intended for receiver $k$, $k = 1, 2$. Since codebooks, probability of error, achievable rate tuples $(R_1, R_2)$, and capacity region $\mathcal{C}$, are all defined in the standard Shannon theoretic sense, their definitions will not be repeated here.

The channel model is parameterized by $P$. The GDoF region is defined as

$$\mathcal{D} = \left\{(d_1, d_2) : \exists (R_1(P), R_2(P)) \in \mathcal{C}(P) \text{ s. t. } d_1 = \lim_{P\to\infty} \frac{R_1(P)}{C_o(P)}, d_2 = \lim_{P\to\infty} \frac{R_2(P)}{C_o(P)}\right\} \quad (3)$$

where $C_o(P)$ is the baseline reference capacity of an additive white Gaussian noise channel $Y = X+N$ with transmit power $P$ and unit variance noise. For real settings it is $1/2\log(P)+o(\log(P)) = \log(\bar{P}) + o(\log(P))$ where for notational convenience we define

$$\bar{P} \triangleq \sqrt{P}.$$

*Remark: Note that unlike DoF, the scaling with $P$ in the GDoF framework does not correspond to a physical scaling of powers in a given channel, because of the different power scaling exponents $\alpha_{ij}$. Rather, each $P$ value defines a new channel. Intuitively, this class of channels belong together because, normalized by* $\log(P)$*, they have (approximately) the same capacity, so that a GDoF characterization simultaneously characterizes the capacity of all the channels in this class within a gap of $o(\log(P))$.*

For the interference channel, quantities of interest for user $k$ include the signal (interference) to noise power ratio, $\text{SNR}_k$ ($\text{INR}_k$) defined here (in logarithmic scale) as follows.

$$\begin{aligned} \log \text{SNR}_k &= \alpha_{kk} \log(\bar{P}) + 2\log(|G_{kk}|) \\ \log \text{INR}_k &= \alpha_{k\bar{k}} \log(\bar{P}) + 2\log(|G_{k\bar{k}}|) \end{aligned}$$

where $\bar{k}$ is defined to be 1 if $k = 2$, and 2 if $k = 1$. Note that

$$\alpha_{kk} = \lim_{P\to\infty} \frac{\log \text{SNR}_k}{\log(\bar{P})} \quad (4)$$

$$\alpha_{k\bar{k}} = \lim_{P\to\infty} \frac{\log \text{INR}_k}{\log(\bar{P})} \quad (5)$$

### 2.2 Partial Cooperation: X Channel

A partial cooperation scenario of interest is to allow each transmitter to serve independent messages to both users. This produces the X channel setting, with four independent messages: $W_{11}, W_{12}, W_{21}, W_{22}$, such that message $W_{ij}$ originates at transmitter $j$ and is intended for receiver $i$.

### 2.3 Full Cooperation: MISO BC

Allowing full cooperation between the two transmitters produces the MISO BC (multiple input single output broadcast channel) setting where the effective transmitter has two antennas, and there are two independent messages $W_1, W_2$ intended for receivers $1, 2$, respectively, each of which is equipped with one antenna.

## 3 Background

### 3.1 Perfect CSIT

The perfect CSIT assumption implies that the channel knowledge at the transmitters is infinitely precise, instantaneous, and globally available.[2] In terms of DoF results ($\alpha_{ij} = 1, \forall i, j \in \{1, 2\}$), with

[2]For all the discussion in this paper, please note that perfect channel knowledge is always assumed at the receivers.

perfect CSIT, full cooperation (BC) enables 2 DoF, partial cooperation (X channel) enables 4/3 DoF [3], whereas no cooperation (interference channel) allows only 1 DoF. GDoF region characterizations are also known under perfect CSIT. For ease of exposition in this section, let us focus on sum-GDoF in the symmetric setting, and compare the interference channel without cooperation, with partial cooperation (X), and with full cooperation (BC).

$$\textbf{symmetric setting:} \quad \begin{array}{l} \alpha_{11} = \alpha_{22} = 1 \\ \alpha_{12} = \alpha_{21} = \alpha \end{array} \tag{6}$$

If perfect CSIT is assumed to be available, then the sum-GDoF of the 2 user interference channel are represented by the so-called "W" curve [4], shown in Fig. 1 by the green line segments. Starting from the left, the different segments correspond to very weak, weak, moderately weak, strong, and very strong interference scenarios.

The most interesting aspect of partial cooperation, i.e., the $X$ channel setting, is the possibility of interference alignment, which does not arise in the 2 user interference channel. In the symmetric setting (6) with perfect CSIT, the GDoF of the X channel [5] are represented in Fig. 1 by the red line segments. To identify the gains from interference alignment in the X channel, as opposed to the GDoF of the underlying interference channel, it is important to note that the X channel contains another interference channel, with messages $W_{12}, W_{21}$, whose sum-GDoF in the symmetric setting are shown in Fig. 1 in blue. From Fig. 1, it is evident that the X channel has a GDoF advantage over the best of the two underlying interference channels only in the regime $\frac{2}{3} < \alpha < \frac{3}{2}$ (shaded in Fig. 1) [5]. This is the regime where the red plot strictly dominates the best of blue and green plots — the regime where interference alignment is useful. Outside this regime, in order to achieve the optimal sum-GDoF, it suffices to operate the $X$ channel as the weak interference channel. Remarkably, the GDoF characterization for the X channel has also been further refined all the way to an exact capacity characterization in the very weak (also known as "noisy") interference regime [5].

With full cooperation, if perfect CSIT is available, then zero forcing suffices to achieve the sum-GDoF of the resulting BC, which, in the symmetric setting, are easily seen to be $2\max(1, \alpha)$, and are shown in Fig. 1 at the top of the figure. Clearly, the benefits of full cooperation are quite significant under perfect CSIT.

### 3.2 Limited CSIT

Given the difficulty of achieving near perfect channel knowledge at the transmitters in practice, there has been much work aimed at relaxing this assumption. It is known that under no CSIT (isotropic fading) the DoF of the BC setting collapse [6], so there is no DoF benefit of cooperation. If channels are drawn from generic sets of finite cardinality, with the specific realization unknown to the transmitter, then under this limited CSIT model (also known as the compound setting), the DoF of the BC setting collapse to those of the X channel [7]. Remarkably, the X channel in the compound setting does not lose any DoF relative to perfect CSIT. Thus, under the compound channel uncertainty model full cooperation does not allow any more DoF benefits beyond that of partial cooperation as represented by the X channel. Other models of limited CSIT include delayed CSIT [8] where full cooperation allows 4/3 DoF, while the optimal DoF of partial cooperation (X channel) remain open. The DoF of mixed CSIT models where imperfect current CSIT and perfect delayed CSIT are both available, have been characterized for the full cooperation scenario (BC) in [9]. The DoF of alternating CSIT models where CSIT can vary across users between perfect, delayed and none, have been explored in [10] which also identifies synergistic benefits.

### 3.3 Finite Precision CSIT

Under the finite precision CSIT model, the transmitters are assumed to be aware of the $\alpha_{ij}$ values, i.e., the coarse channel strength parameters, but not the precise $G_{ij}$ values. For the $G_{ij}$ the transmiters are only aware of the joint probability density function (PDF). Define the set of channel coefficient variables $\mathcal{G} = \{G_{ij}(t) : t \in \mathbb{N}, i, j \in \{1, 2\}\}$. Finite precision CSIT corresponds to the existence of bounded density functions. Precisely, the finite precision CSIT model assumes that there exists a finite positive constant $f_{\max}$,

$$0 < f_{\max} < \infty$$

such that for all finite cardinality disjoint subsets $\mathbf{G}_1, \mathbf{G}_2$ of $\mathcal{G}$,

$$\mathbf{G}_1 \subset \mathcal{G}, \mathbf{G}_2 \subset \mathcal{G}, \mathbf{G}_1 \cap \mathbf{G}_2 = \phi, |\mathbf{G}_1| < \infty, |\mathbf{G}_2| < \infty$$

the conditional PDF

$$\forall G_1, G_2, \quad f_{\mathbf{G_1}|\mathbf{G_2}}(G_1|G_2) \quad \leq \quad f_{\max}^{|\mathbf{G}_1|}.$$

Despite being investigated extensively over the past decade, the DoF with transmitter cooperation remained an open problem under finite precision CSIT, until recently it was shown that there is no DoF advantage of full or partial cooperation, i.e., the BC (and therefore also the X channel) has only 1 DoF under finite precision CSIT [2]. In fact, this was shown to be true even if perfect CSIT for one of the two users, say user 1, was consistently available to the transmitter.

## 4 Results: GDoF under Finite Precision CSIT

In this section, we provide an overview of the results of this work and place them in perspective with prior work.

### 4.1 Main Result

The main result of this work is a complete GDoF characterization for full transmitter cooperation (BC) under finite precision CSIT as stated in the following theorem.

**Theorem 1** *The GDoF region of the 2 user MISO broadcast channel under finite precision CSIT is:*

$$\mathcal{D} = \left\{ (d_1, d_2) \in \mathbb{R}_2^+ : \begin{array}{c} d_1 \leq A, \\ d_2 \leq B, \\ d_1 + d_2 \leq \min(A + C, B + D) \end{array} \right\} \tag{7}$$

*where*

$$A = \max(\alpha_{11}, \alpha_{12}) \tag{8}$$

$$B = \max(\alpha_{21}, \alpha_{22}) \tag{9}$$

$$C = \max((\alpha_{21} - \alpha_{11})^+, (\alpha_{22} - \alpha_{12})^+) \tag{10}$$

$$D = \max((\alpha_{11} - \alpha_{21})^+, (\alpha_{12} - \alpha_{22})^+) \tag{11}$$

Note that Theorem 1 provides a full GDoF region (as opposed to only sum-GDoF) characterization, and for all values of $\alpha_{ij}$ (as opposed to only symmetric setting).

### 4.2 Comparison: Perfect vs Finite Precision CSIT

For ease of comparison with other related GDoF characterizations, we focus on the sum-GDoF under the symmetric setting as illustrated in Fig. 1. For the MISO BC with finite precision CSIT, the sum-GDoF value is $\max(2-\alpha, 2\alpha-1)$ and is shown in black in the figure.

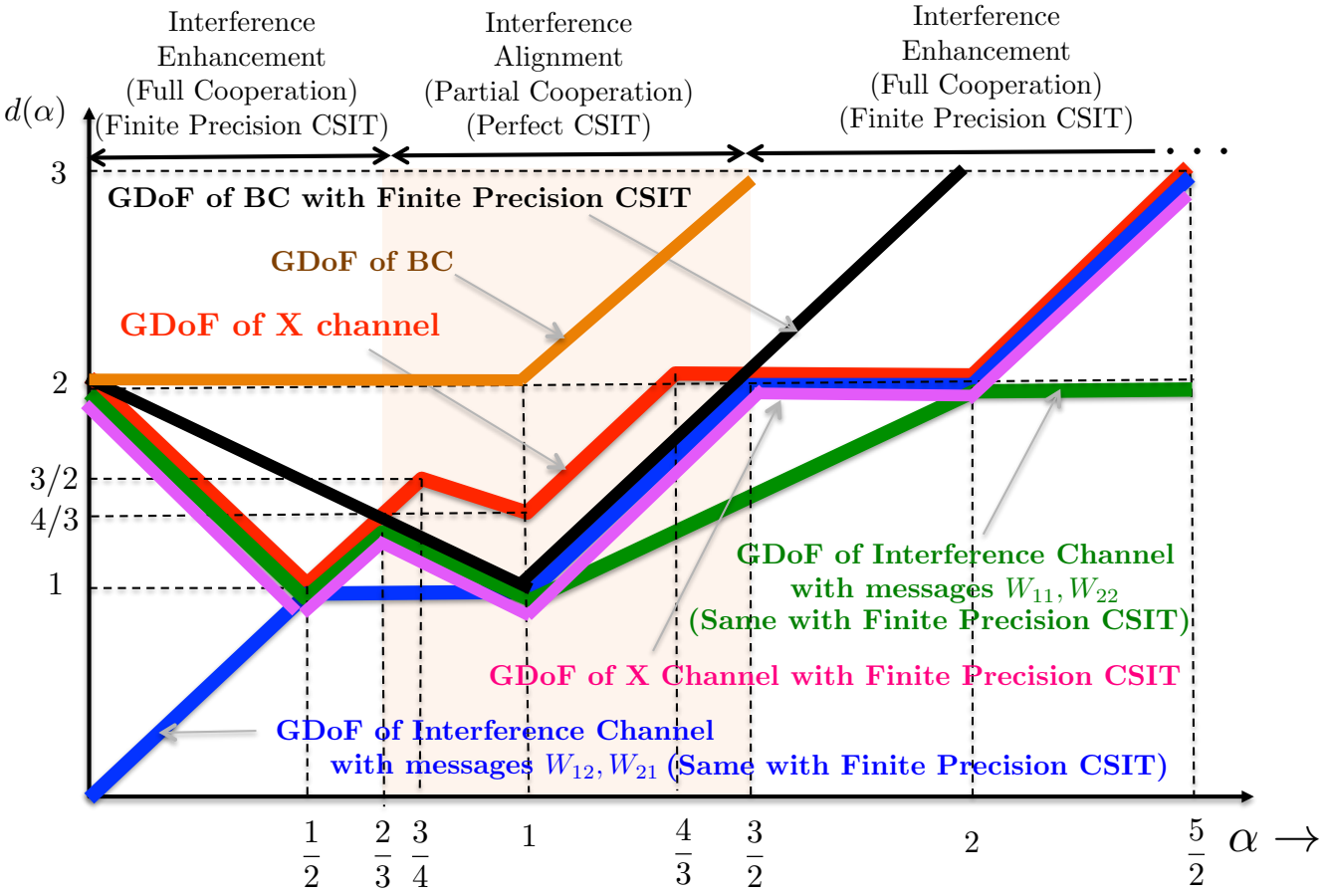


Figure 1: Sum-GDoF in the symmetric case $\alpha_{11} = \alpha_{22} = 1, \alpha_{12} = \alpha_{21} = \alpha$.

Here we list the key observations.

1. [**No cooperation (IC) – No loss**]: It is straightforward to verify that the GDoF of the interference channel under finite precision CSIT are the same as with perfect CSIT. This is also true for the other interference channel with messages $W_{12}, W_{21}$.

2. [**Full Cooperation (BC) – Loss of** $\min(1, \alpha)$]: As long as $\alpha \neq 0$, there is always a loss in the BC GDoF due to finite precision CSIT compared to perfect CSIT, and the loss is equal to $\min(1, \alpha)$.

3. [**Partial Cooperation (X) – Reduced to Trivial**]: Recall that the X channel had a strict advantage over the underlying interference channels only in the regime $2/3 < \alpha < 3/2$ where interference alignment allowed higher GDoF under perfect CSIT. Under finite precision CSIT, the sum-GDoF of the X channel are bounded above by that of the BC under finite precision CSIT in the regime $2/3 < \alpha < 3/2$, and by the X channel GDoF under perfect

CSIT everywhere outside. However, these bounds always correspond to one of the underlying interference channels. Thus, there is no benefit of partial cooperation relative to using the best of the underlying interference channels under finite precision CSIT.

4. [**Interference Alignment Benefits Disappear**]: Consider the regime where interference alignment was useful under perfect CSIT and partial cooperation (X channel), i.e., $2/3 < \alpha < 3/2$. Under finite precision CSIT, the sum GDoF (even with full cooperation) in this regime collapse to the best of the underlying interference channels. In other words, the benefits of interference alignment are entirely lost under finite precision CSIT.

5. [**Interference Enhancement offers the only Robust Advantage**]: Remarkably, while the regime where interference alignment was useful sees a collapse to underlying interference channels, the opposite happens everywhere interference alignment was not useful. Everywhere outside the regime $2/3 < \alpha < 3/2$, note that the sum-GDoF of the BC under finite precision CSIT strictly dominate the best of the interference channels. Since in this regime there was no additional advantage of partial cooperation even with perfect CSIT, the BC with finite precision CSIT also dominates the X channel (even with perfect CSIT!) in this regime. Indeed, as we will see, the advantage is not due to interference alignment, but rather due to interference enhancement[3] [11], i.e., strengthening the interference so that it can be decoded and subtracted by the undesired receiver. Thus, remarkably, *under finite precision CSIT, interference enhancement emerges as the only scheme with a robust GDoF advantage relative to the underlying interference channels, and this advantage is accessible only through full cooperation (not through partial cooperation).*

While a pictorial representation is difficult for the fully asymmetric setting because of the abundance of parameters, the following theorem presents the corresponding generalization.

**Theorem 2** *For all parameters $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}$, we have*

$$\max(S(IC_1), S(IC_2)) \quad = S^o(X) \quad = \min(S^o(BC), S^1(X))$$

*where $IC_1$ is the interference channel with messages $W_{11}, W_{22}$, $IC_2$ is the interference channel with messages $W_{21}, W_{12}$, and for any channel $S^o$ is the sum-GDoF value with finite precision CSIT, and $S^1$ is the sum-GDoF value with perfect CSIT. Note that for the interference channels the CSIT superscripts are omitted because $S^1(IC_1) = S^o(IC_1)$ and $S^1(IC_2) = S^o(IC_2)$.*

# 5 Proof of Theorem 1: Real Setting

For ease of exposition we first present the proof for the real setting.

## 5.1 Outer Bound

Outer bounds $d_1 \leq A$ and $d_2 \leq B$ are trivial bounds for single user capacity. We will prove the remaining bound $d_1 + d_2 \leq \min(A + C, B + D)$. Note that the outer bound argument is a generalization of the combinatorial argument introduced in [2]. To avoid repetition, and due to space limitations, we will omit some of the detailed explanations for similar steps in [2].

[3]Interference enhancement is also sometimes referred to as "interference forwarding" [11] when relays are involved. We prefer the terminology interference enhancement for the broadcast channel without relays per se.

**Step 1: Deterministic Channel Model.** The deterministic channel model has inputs $\bar{X}_1(t), \bar{X}_2(t) \in \mathbb{Z}$ and outputs $\bar{Y}_1(t), \bar{Y}_2(t) \in \mathbb{Z}$, defined as

$$\begin{aligned} \bar{Y}_1(t) &= \lfloor \bar{P}^{\alpha_{11}-\max(\alpha_{11},\alpha_{21})} G_{11}(t)\bar{X}_1(t) \rfloor + \lfloor \bar{P}^{\alpha_{12}-\max(\alpha_{12},\alpha_{22})} G_{12}(t)\bar{X}_2(t) \rfloor & (12) \\ \bar{Y}_2(t) &= \lfloor \bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})} G_{21}(t)\bar{X}_1(t) \rfloor + \lfloor \bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})} G_{22}(t)\bar{X}_2(t) \rfloor & (13) \end{aligned}$$

such that

$$\begin{aligned} \bar{X}_1(t) &\in \{0, 1, \cdots, \lceil \bar{P}^{\max(\alpha_{11},\alpha_{21})} \rceil\}, \forall t \in \mathbb{N} & (14) \\ \bar{X}_2(t) &\in \{0, 1, \cdots, \lceil \bar{P}^{\max(\alpha_{12},\alpha_{22})} \rceil\}, \forall t \in \mathbb{N} & (15) \end{aligned}$$

Recall that $\bar{P} = \sqrt{P}$.

**Step 2: Fano's Inequality.**

$$\begin{aligned} n(R_1 + R_2) &\leq n \max(\alpha_{21}, \alpha_{22}) \log(\bar{P}) + \left[ H(\bar{Y}_1^{[n]} | W_2, G^{[n]}) - H(\bar{Y}_2^{[n]} | W_2, G^{[n]}) \right] \\ &\quad + n\ o\log(\bar{P})) + o(n) & (16) \end{aligned}$$

**Step 3: Functional Dependence.** As in [2], without loss of generality

$$\begin{aligned} (\bar{X}_1^{[n]}, \bar{X}_2^{[n]}) &= f(\bar{Y}_1^{[n]}, W_2, G_{11}^{[n]}, G_{12}^{[n]}) & (17) \\ \Rightarrow \bar{Y}_2^{[n]} &= f(\bar{Y}_1^{[n]}, W_2, G^{[n]}) & (18) \end{aligned}$$

where $a = f(b)$ denotes that $a$ is *some* function of $b$.

**Step 4: Aligned Image Sets.** For given $W_2$ and channel realization $G^{[n]}$, define $S_{\bar{Y}_1{}^{[n]}}(G^{[n]}, W_2)$ as the set of all codewords $(\bar{X}_1^{[n]}, \bar{X}_2^{[n]})$ that produce the same output, $\bar{Y}_2{}^{[n]}$, at receiver 2, as is produced at receiver 2 by the codeword that produces $\bar{Y}_1^{[n]}$ at receiver 1.

$$\begin{aligned} & H(\bar{Y}_1{}^{[n]}, S_{\bar{Y}_1{}^{[n]}} | W_2, G^{[n]}) & (19) \\ &= H(\bar{Y}_1{}^{[n]} | W_2, G^n) + H(S_{\bar{Y}_1{}^{[n]}} | W_2, G^{[n]}, \bar{Y}_1{}^{[n]}) & (20) \\ &= H(\bar{Y}_1{}^{[n]} | W_2, G^n) & (21) \\ &= H(S_{\bar{Y}_1{}^{[n]}} | W_2, G^{[n]}) + H(\bar{Y}_1{}^{[n]} | S_{\bar{Y}_1{}^{[n]}}, W_2, G^{[n]}) \\ &= H(\bar{Y}_2{}^{[n]} | W_2, G^{[n]}) + H(\bar{Y}_1{}^{[n]} | S_{\bar{Y}_1{}^{[n]}}, W_2, G^{[n]}) & (22) \\ &\leq H(\bar{Y}_2{}^{[n]} | W_2, G^{[n]}) + \mathrm{E}[\log |S_{\bar{Y}_1{}^{[n]}}|] & (23) \\ &\leq H(\bar{Y}_2{}^{[n]} | W_2, G^{[n]}) + \log \left( \mathrm{E}[|S_{\bar{Y}_1{}^{[n]}}|] \right) & (24) \end{aligned}$$

From (21) and (24) we have

$$H(\bar{Y}_1{}^{[n]} | W_2, G^n) - H(\bar{Y}_2{}^{[n]} | W_2, G^{[n]}) \leq \log \left( \mathrm{E}[|S_{\bar{Y}_1{}^{[n]}}|] \right)$$

So it only remains to bound the average size of an aligned image set $\mathrm{E}[|S_{\bar{Y}_1{}^{[n]}}|]$.

$$\mathrm{E}[|S_{\bar{y}_1^{[n]}}|] = \sum_{\bar{y'}_1^n \in \{\bar{Y}_1{}^{[n]}\}} \mathbb{P}\left( \bar{y'}_1^n \in S_{\bar{y}_1^{[n]}} \right) \quad (25)$$

**Step 5. Probability that Images Align.** Given $G_{11}^{[n]}, G_{12}^{[n]}$, consider two distinct realizations of user 1's output sequence $\bar{Y}_1^{[n]}$, denoted as $\lambda^{[n]}$ and $\nu^{[n]}$, which are produced by the corresponding two realizations of the codeword $(X_1^{[n]}, X_2^{[n]})$ denoted by $(\lambda_1^{[n]}, \lambda_2^{[n]})$ and $(\nu_1^{[n]}, \nu_2^{[n]})$, respectively.

$$\begin{aligned}\lambda(t) &= \lfloor \bar{P}^{\alpha_{11}-\max(\alpha_{11},\alpha_{21})} G_{11}(t)\lambda_1(t) \rfloor + \lfloor \bar{P}^{\alpha_{12}-\max(\alpha_{12},\alpha_{22})} G_{12}(t)\lambda_2(t) \rfloor \quad (26)\\ \nu(t) &= \lfloor \bar{P}^{\alpha_{11}-\max(\alpha_{11},\alpha_{21})} G_{11}(t)\nu_1(t) \rfloor + \lfloor \bar{P}^{\alpha_{12}-\max(\alpha_{12},\alpha_{22})} G_{12}(t)\nu_2(t) \rfloor \quad (27)\end{aligned}$$

We wish to bound the probability that the images of these two codewords align at user 2, i.e., $\nu^{[n]} \in S_{\lambda^{[n]}}$. For simplicity, consider first the single channel use setting, $n = 1$. For $\nu \in S_\lambda$ we must have

$$\begin{aligned}&\lfloor \bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})} G_{21}\nu_1 \rfloor + \lfloor \bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})} G_{22}\nu_2 \rfloor \\ &= \lfloor \bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})} G_{21}\lambda_1 \rfloor + \lfloor \bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})} G_{22}\lambda_2 \rfloor \quad (28)\end{aligned}$$

So for fixed value of $G_{22}$ the random variable $\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})} G_{21}(\nu_1 - \lambda_1)$ must take values within an interval of length no more than 4. If $\nu_1 \neq \lambda_1$, then $G_{21}$ must takes values in an interval of length no more than $\frac{4}{\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}|\nu_1-\lambda_1|}$, the probability of which is no more than $\frac{4f_{\max}}{\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}|\nu_1-\lambda_1|}$. Similarly, for fixed value of $G_{21}$ the random variable $\bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})} G_{22}(\nu_2 - \lambda_2)$ must take values within an interval of length no more than 4. If $\nu_1 = \lambda_1$ then, because $\nu \neq \lambda$, we must have $\nu_2 \neq \lambda_2$, then the probability of alignment is similarly bounded by $\frac{4f_{\max}}{\bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})}|\nu_2-\lambda_2|}$.

Next we will bound the max of $\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}|\nu_1 - \lambda_1|$ and $\bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})}|\nu_2 - \lambda_2|$. From (26) and (27) we have

$$\begin{aligned}|\lambda - \nu| &\leq 2 + \bar{P}^{\alpha_{11}-\max(\alpha_{11},\alpha_{21})}|G_{11}||\lambda_1 - \nu_1| + \bar{P}^{\alpha_{12}-\max(\alpha_{12},\alpha_{22})}|G_{12}||\lambda_2 - \nu_2| \quad (29)\\ &\leq 2 + 2\Delta_2 \bar{P}^{\max(\alpha_{11}-\alpha_{21},\alpha_{12}-\alpha_{22},0)} \\ &\quad \times \max(\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}|\nu_1 - \lambda_1|, \bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})}|\nu_2 - \lambda_2|) \quad (30)\end{aligned}$$

So, if $|\lambda - \nu| > 2$, the probability of $\nu \in S_\lambda$ is no more than

$$\frac{8\Delta_2 f_{\max} \bar{P}^{\max(\alpha_{11}-\alpha_{21},\alpha_{12}-\alpha_{22},0)}}{|\lambda - \nu| - 2} \quad (31)$$

Now let us return to the case of general $n$, where we similarly have,

$$\mathbb{P}(\lambda^{[n]} \in S_{\nu^{[n]}}) \leq (8\Delta_2 f_{\max})^n \bar{P}^{\max(\alpha_{11}-\alpha_{21},\alpha_{12}-\alpha_{22},0)} \times \prod_{t:|\lambda(t)-\nu(t)|>2} \frac{1}{|\lambda(t) - \nu(t)| - 2} \quad (32)$$

**Step 6. Bounding the Expected Size of Aligned Image Sets.**

$$\begin{aligned}\mathrm{E}(|S_{\nu^{[n]}}|) &= \sum_{\lambda^n \in \{\bar{Y_1}^{[n]}\}} \mathbb{P}\left(\lambda^n \in S_{\nu^{[n]}}\right) \quad (33)\\ &\leq (8\Delta_2 f_{\max})^n \bar{P}^{n\max(\alpha_{11}-\alpha_{21},\alpha_{12}-\alpha_{22},0)} \\ &\quad \prod_{t=1}^{n} \left( \sum_{\lambda(t):|\lambda(t)-\nu(t)|\leq 2} 1 + \sum_{\lambda(t):2<|\lambda(t)-\nu(t)|\leq Q_y} \frac{1}{|\lambda(t) - \nu(t)| - 2} \right) \\ &\leq (8\Delta_2 f_{\max})^n \bar{P}^{n\max(\alpha_{11}-\alpha_{21},\alpha_{12}-\alpha_{22},0)} \times \left(\max(\alpha_{11}, \alpha_{12}) \log(\bar{P}) + o(\log(\bar{P}))\right)^n \quad (34)\end{aligned}$$

where $Q_y \leq (2\Delta_2+2)\lceil \bar{P}^{\max(\alpha_{11},\alpha_{12})}\rceil$. Substituting these bounds back into (16) we have

$$\begin{aligned} n(R_1+R_2) &\leq n\max(\alpha_{21},\alpha_{22})\log(\bar{P}) \\ &\quad + \left[H(\bar{Y}_1^{[n]}|W_2,G^{[n]}) - H(\bar{Y}_2^{[n]}|W_2,G^{[n]})\right] + n\ o\log(\bar{P})) + o(n) \qquad (35) \\ &\leq n\max(\alpha_{21},\alpha_{22})\log(\bar{P}) + \log \mathrm{E}|S_{\nu^{[n]}}| + n\ o\log(\bar{P})) + o(n) \\ &\leq n\left(\max(\alpha_{21},\alpha_{22}) + \max(\alpha_{11}-\alpha_{21},\alpha_{12}-\alpha_{22},0)\right)\log(\bar{P}) \\ &\qquad + n\ o\log(\bar{P})) + o(n) \qquad (36) \end{aligned}$$

So that we obtain the GDoF bound

$$d_1 + d_2 \leq B + D \qquad (37)$$

By symmetry we also have the GDoF bound

$$d_1 + d_2 \leq A + C \qquad (38)$$

Together these two bounds give us $d_1 + d_2 \leq \min(A+C, B+D)$, completing the proof of the outer bounds for Theorem 1.

### 5.2 Achievability

The key idea for achievability is interference enhancement [11]. Before presenting the general proof, let us convey the main insights through a simple example. Consider the symmetric setting with $\alpha = 0.5$, where we wish to achieve the sum-GDoF value of $d_1 + d_2 = 1.5$ through the tuple $d_1 = 1, d_2 = 0.5$. To do this, let us split user 1's message as $W_1 = (W_c, W_{1p})$ and user 2's message as $W_2 = W_{2p}$, where $W_{1p}$ acts as a private sub-message to be decoded only by user 1, $W_{2p}$ acts as a private sub-message to be decoded only by user 2, while $W_c$ acts as a common sub-message that can be decoded by both users. Each sub-message carries 0.5 GDoF. Messages $W_c, W_{1p}, W_{2p}$ are encoded into independent Gaussian codebooks $X_c, X_{1p}, X_{2p}$, with powers $\mathrm{E}|X_c|^2 = 1 - P^{-0.5}$, $\mathrm{E}|X_{1p}|^2 = P^{-0.5}$ and $\mathrm{E}|X_{2p}|^2 = P^{-0.5}$. From the first transmit antenna, we send $X_1 = X_c + X_{1p}$ and from the second transmit antenna, $X_2 = X_c + X_{2p}$. Suppressing the time index for clarity, the received signals are:

$$\begin{aligned} Y_1 &= \bar{P}G_{11}(X_c + X_{1p}) + \bar{P}^{0.5}G_{12}(X_c + X_{2p}) + Z_1 \\ Y_2 &= \bar{P}^{0.5}G_{21}(X_c + X_{1p}) + \bar{P}G_{22}(X_c + X_{2p}) + Z_2 \end{aligned}$$

Receiver 1 first decodes the codeword $X_c$ for the message $W_c$, treating everything else as noise. The SINR value for this decoding is

$$\frac{|\bar{P}G_{11} + \bar{P}^{0.5}G_{12}|^2(1-P^{-0.5})}{1 + P(P^{-0.5})|G_{11}|^2 + P^{0.5}P^{-0.5}|G_{12}|^2} \geq \frac{P((\Delta_1 - \bar{P}^{-0.5}\Delta_2)^+)^2(1-P^{-0.5})}{1 + P^{0.5}\Delta_2^2 + \Delta_2^2} \qquad (39)$$

and the achievable rate (for real channels) is $0.5\log(1+\mathrm{SINR}) = 0.25\log(P) + o(\log(P)) = 0.5\log(\bar{P}) + o(\log(P))$, which gives us the GDoF value $d_c = 0.5$.

$$\frac{P(1-P^{-0.5})|G_{11}(t)|^2}{(1 + P(P^{-0.5})|G_{11}(t)|^2 + P^{0.5}((1-P^{-0.5})|G_{12}(t)|^2 + P^{-0.5}|G_{12}|^2))}$$

$$\geq \frac{(P - P^{0.5})\Delta_1^2}{1 + P^{0.5}\Delta_2^2 + (P^{0.5} - 1)\Delta_2^2 + |G_{12}|^2} \tag{40}$$

After decoding $W_c$, receiver 1 is able to reconstruct codeword $X_c$ and subtract its contribution from the received signal. After this, receiver 1 decodes the codeword $X_{1p}$ for its desired message $W_1$, while treating the remaining signals as noise. The rate that is supported for this message is:

$$\begin{aligned} &= \frac{1}{2}\log\left(\frac{P(P^{-0.5})|G_{11}|^2}{1 + P^{0.5}(P^{-0.5})|G_{12}|^2}\right) \\ &\geq \frac{1}{2}\log\left(\frac{P^{0.5}\Delta_1^2}{1 + \Delta_2^2}\right) \\ &= \frac{1}{4}\log(P) + o(\log(P)) = \frac{1}{2}\log(\bar{P}) + o(\log(P)) \end{aligned}$$

which gives us the GDoF value $d_1 = 0.5$. Receiver 2 proceeds similarly, first decoding $X_c$ for $W_c$ while treating all other signals as noise, which is feasible for $d_c = 0.5$, and then reconstructs and subtracts the contribution of $X_c$ from its received signal. It finally decodes $X_{2p}$ for $W_2$ while treating the remaining signals as noise, which is feasible for $d_2 = 0.5$. Thus, the GDoF achieved is $(d_1, d_2) = (d_{1p} + d_c, d_2) = (0.5 + 0.5, 0.5) = (1, 0.5)$. Note the key role of interference enhancement, in the encoding of $W_c$ into $X_c$. This is interference for receiver 2, and yet by also sending it from the stronger antenna (antenna 2) for user 2, the power of the interference at user 2 is enhanced enough so that it can be decoded and subtracted by receiver 2, before proceeding to decode its desired signal.

Now, for the general case of $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}$ we present the achievability scheme for the point $d_1 = A, d_2 = \min(A+C, B+D) - A$. Note that the other point $d_1 = \min(A+C, B+D) - B, d_2 = B$ is derived similarly, and the whole region is derived by time sharing. The four cases summarized below cover all possibilities.

Consider the four cases of ($\alpha_{12} \leq \alpha_{11}$ and $\alpha_{21} \leq \alpha_{22}$), ($\alpha_{11} \leq \alpha_{12}$ and $\alpha_{22} \leq \alpha_{21}$), ($\alpha_{12} \leq \alpha_{11}$ and $\alpha_{22} \leq \alpha_{21}$), and ($\alpha_{11} \leq \alpha_{12}$ and $\alpha_{21} \leq \alpha_{22}$).

1. $\alpha_{12} \leq \alpha_{11}$ and $\alpha_{21} \leq \alpha_{22}$.

   (a) $\alpha_{11} \leq \alpha_{21}$. Here $W_1 = W_c, W_2 = W_{2p}$. $\mathrm{E}|X_c|^2 = 1 - P^{-\alpha_{11}}$ and $\mathrm{E}|X_{2p}|^2 = P^{-\alpha_{11}}$. $X_1 = X_c, X_2 = X_c + X_{2p}$. Achieves $(d_1 = \alpha_{11}, d_2 = \alpha_{22} - \alpha_{11})$.
   User 1's message $W_1$ acts as a common message to be decoded by both receivers and is encoded into a Gaussian codebook $X_c$ with power $\mathrm{E}|X_c|^2 = 1 - P^{-\alpha_{11}}$. User 2's message acts as a private message decodable only by user 2, and is encoded into an independent Gaussian codebook $X_p$ with power $\mathrm{E}|X_{2p}|^2 = P^{-\alpha_{11}}$. From the first transmit antenna we send $X_1 = X_c$ and from the second transmit antenna, $X_2 = X_c + X_{2p}$. Receiver 1 decodes $X_c$ treating everything else as noise to get its desired message $W_1$ with $d_1 = \alpha_{11}$ GDoF. Receiver 2 decodes and subtracts $X_c$ while treating its own desired signal as noise, and then decodes its desired signal to achieve $d_2 = \alpha_{22} - \alpha_{11}$ GDoF. Suppressing the time index for clarity, the received signals are:

$$\begin{aligned} Y_1 &= \bar{P}^{\alpha_{11}} G_{11}(X_c) + \bar{P}^{\alpha_{12}} G_{12}(X_c + X_{2p}) + Z_1 \\ Y_2 &= \bar{P}^{\alpha_{21}} G_{21}(X_c) + \bar{P}^{\alpha_{22}} G_{22}(X_c + X_{2p}) + Z_2 \end{aligned}$$

Receiver 1 first decodes the codeword $X_c$ for the message $W_c$, treating everything else as noise. The SINR value for this decoding is

$$\begin{aligned} & \log\left(\frac{P^{\alpha_{11}}((\Delta_1 - \bar{P}^{\alpha_{12}-\alpha_{11}}\Delta_2)^+)^2(1-P^{-\alpha_{11}})}{1+P^{\alpha_{12}}P^{-\alpha_{11}}\Delta_2^2}\right) \\ \geq \quad & \alpha_{11}\log(P) + o(\log(P)) \end{aligned} \tag{41}$$

and,

$$\begin{aligned} & \log\left(\frac{P^{\alpha_{22}}((\Delta_1 - \bar{P}^{\alpha_{21}-\alpha_{22}}\Delta_2)^+)^2(1-P^{-\alpha_{11}})}{1+P^{\alpha_{22}}P^{-\alpha_{11}}\Delta_2^2}\right) \\ \geq \quad & \alpha_{11}\log(P) + o(\log(P)) \end{aligned} \tag{42}$$

so, both of the recievers can decode $X_c$. $\log(1+\text{SINR}) \geq \alpha_{11}\log(P) + o(\log(P))$, which gives us the GDoF value $d_c = \alpha_{11}$.
After decoding $W_c$, receiver 2 is able to reconstruct codeword $X_c$ and subtract its contribution from the received signal. After this, receiver 2 decodes the codeword $X_{2p}$ for its desired message $W_2$, while treating the remaining signals as noise. The rate that is supported for this message is:

$$\begin{aligned} = \quad & \log\left(\frac{P^{\alpha_{22}}(P^{-\alpha_{11}})|G_{22}|^2}{1}\right) \\ \geq \quad & (\alpha_{22}-\alpha_{11})\log(P) + o(\log(P)) \end{aligned} \tag{43}$$

Receiver 2 proceeds similarly, first decoding $X_c$ for $W_c$ while treating all other signals as noise, which is feasible for $d_c = \alpha_{11}$, and then reconstructs and subtracts the contribution of $X_c$ from its received signal. It finally decodes $X_{2p}$ for $W_2$ while treating the remaining signals as noise, which is feasible for $d_2 = \alpha_{22} - \alpha_{11}$. Thus, the GDoF achieved is $(d_1, d_2) = (d_c, d_{2p}) = (\alpha_{11}, \alpha_{22} - \alpha_{11})$.
Note that when $\alpha_{11} = \alpha_{12}$, the SINR value for decoding the message $W_c$ by treating everything as noise will be

$$\begin{aligned} & \log\left(E_{G_{11},G_{12}}\frac{P^{\alpha_{11}}|G_{11}+G_{12}|^2(1-P^{-\alpha_{11}})}{1+P^{\alpha_{12}}P^{-\alpha_{11}}|G_{12}|^2}\right) \\ & \log\left(\frac{P^{\alpha_{11}}(1-P^{-\alpha_{11}})}{1+P^{\alpha_{12}}P^{-\alpha_{11}}\Delta_2^2}E_{G_{11},G_{12}}|G_{11}+G_{12}|^2\right) \\ \geq \quad & \alpha_{11}\log(P) + o(\log(P)) \end{aligned} \tag{44}$$

where we used that for all $m$, $E_{G_{11}}|G_{11}+m|^2 \geq E_{\{G_{11},|G_{11}+m|>\frac{1}{2f_{\max}}\}}|G_{11}+m|^2 \geq \frac{1}{4f_{\max}^2}Pr(|G_{11}+m| > \frac{1}{2f_{\max}}) \geq \frac{1}{8f_{\max}^2}$.

(b) $\alpha_{21} \leq \alpha_{12} \leq \alpha_{22}$. Here $W_1 = (W_c, W_{1p})$, $W_2 = (W_{2p})$. $\text{E}|X_c|^2 = 1 - P^{-\alpha_{12}}$, $\text{E}|X_{1p}|^2 = P^{-\alpha_{12}}$, and $\text{E}|X_{2p}|^2 = P^{-\alpha_{12}}$. $X_1 = X_c + X_{1p}$ and $X_2 = X_c + X_{2p}$. Achieves $(d_1, d_2) = (\alpha_{11}, \alpha_{22} - \alpha_{12})$.
Split messages $W_1 = (W_c, W_{1p})$, and $W_2 = (W_{2p})$ which are encoded into independent Gaussian codebooks $X_c, X_{1p}, X_{2p}$ with powers $\text{E}|X_c|^2 = 1 - P^{-\alpha_{12}}$, $\text{E}|X_{1p}|^2 = P^{-\alpha_{12}}$,

and $\mathrm{E}|X_{2p}|^2 = P^{-\alpha_{12}}$. Send $X_1 = X_c + X_{1p}$ and $X_2 = X_c + X_{2p}$. Both receivers decode $X_c$ and subtract it from their received signal to decode they private codewords. The GDoF achieved is $(d_1, d_2) = (\alpha_{11}, \alpha_{22} - \alpha_{12})$.

Similarly, let us split user 1's message as $W_1 = (W_c, W_{1p})$ and split user 2's message as $W_1 = (W_c, W_{2p})$ where sub-messages $W_c$, $W_{1p}$ and $W_{2p}$ carries $\alpha_{12}$, $\alpha_{11} - \alpha_{12}$ and $\alpha_{22} - \alpha_{12}$ GDoF respectively. Messages $W_c, W_{1p}, W_{2p}$ are encoded into independent Gaussian codebooks $X_c, X_{1p}, X_{2p}$, with powers $\mathrm{E}|X_c|^2 = 1 - P^{-\alpha_{12}}$, $\mathrm{E}|X_{1p}|^2 = P^{-\alpha_{12}}$, and $\mathrm{E}|X_{2p}|^2 = P^{-\alpha_{12}}$. From the first transmit antenna, we send $X_1 = X_c + X_{1p}$ and from the second transmit antenna, $X_2 = X_c + X_{2p}$. similar to the first case, both of the recievers can decode $X_c$. $\log(1 + \mathrm{SINR}) \geq \alpha_{12} \log(P) + o(\log(P))$, which gives us the GDoF value $d_c = \alpha_{12}$.

After decoding $W_c$, receiver 1 and 2 are able to reconstruct codeword $X_c$ and subtract its contribution from the received signal. After this, receiver 1 decodes the codewords $X_{1p}$ for its desired message $W_1$ and receiver 2 decodes the codewords $X_{2p}$ for its desired message $W_2$, while treating the remaining signals as noise. Similar to the first case, $d_{1p} = \alpha_{11} - \alpha_{12}$, and, $d_{2p} = \alpha_{22} - \alpha_{12}$. Thus, the GDoF achieved is $(d_1, d_2) = (d_{1p} + d_c, d_{2p}) = (\alpha_{11}, \alpha_{22} - \alpha_{12})$.

(c) $\alpha_{12} \leq \alpha_{21} \leq \alpha_{11}$. Here $W_1 = (W_c, W_{1p}), W_2 = W_{2p}$. $\mathrm{E}|X_c|^2 = 1 - P^{-\alpha_{21}}$, $\mathrm{E}|X_{1p}|^2 = P^{-\alpha_{21}}$, and $\mathrm{E}|X_{2p}|^2 = P^{-\alpha_{21}}$. $X_1 = X_c + X_{1p}, X_2 = X_c + X_{2p}$. Achieves $(d_1 = \alpha_{11}, d_2 = \alpha_{22} - \alpha_{21})$.

In this case DoF pair of $(d_1 = \alpha_{11}, d_2 = \alpha_{22} - \alpha_{21})$ is achievable similar to the case $\alpha_{21} \leq \alpha_{12} \leq \alpha_{22}$, except that sub-messages $W_c$, $W_{1p}$ and $W_{2p}$ carries $\alpha_{21}$, $\alpha_{11} - \alpha_{21}$ and $\alpha_{22} - \alpha_{21}$ GDoF respectively. Messages $W_c, W_{1p}, W_{2p}$ are encoded into independent Gaussian codebooks $X_c, X_{1p}, X_{2p}$, with powers $\mathrm{E}|X_c|^2 = 1 - P^{-\alpha_{21}}$, $\mathrm{E}|X_{1p}|^2 = P^{-\alpha_{21}}$, and $\mathrm{E}|X_{2p}|^2 = P^{-\alpha_{21}}$.

(d) $\alpha_{22} \leq \alpha_{12}$. $(d_1 = \alpha_{11}, d_2 = 0)$ is achievable trivially.

2. $(\alpha_{11} \leq \alpha_{12}$ and $\alpha_{22} \leq \alpha_{21})$. This case is similar to the case $(\alpha_{12} \leq \alpha_{11}$ and $\alpha_{21} \leq \alpha_{22})$ except the user indices are switched.

3. $(\alpha_{11} \leq \alpha_{12}$ and $\alpha_{21} \leq \alpha_{22})$. In this case, DoF pair of $(d_1 = \alpha_{12}, d_2 = \max(\alpha_{12}, \alpha_{22}) - \alpha_{12})$ is achievable as follows. If $\alpha_{22} \leq \alpha_{12}$ then $(d_1 = \alpha_{12}, d_2 = 0)$ is trivial. If $\alpha_{22} \geq \alpha_{12}$,$(d_1 = \alpha_{12}, d_2 = \alpha_{22} - \alpha_{12})$ is achievable by only transmitting at the second antenna with the full power $\alpha_{22}$. Here $W_2 = (W_{1p}, W_{2p})$. $\mathrm{E}|X_{1p}|^2 = 1 - P^{-\alpha_{12}}$, $\mathrm{E}|X_{2p}|^2 = P^{-\alpha_{12}}$, and $X_2 = X_{1p} + X_{2p}$.

$$\begin{aligned} Y_1 &= \bar{P}^{\alpha_{12}} G_{12}(X_{1p} + X_{2p}) + Z_1 \\ Y_2 &= \bar{P}^{\alpha_{22}} G_{22}(X_{1p} + X_{2p}) + Z_2 \end{aligned}$$

Its easy to check that receiver 1 and 2 can decode the codeword $X_{1p}$ for the message $W_{1p}$, treating everything else as noise. Moreover, after decoding $W_{1p}$, receiver 2 is able to reconstruct codeword $X_{1p}$ and subtract its contribution from the received signal. After this, receiver 2 decodes the codeword $X_{2p}$ for its desired message $W_{2p}$, while treating the remaining signals as noise.

4. $(\alpha_{12} \leq \alpha_{11}$ and $\alpha_{22} \leq \alpha_{21})$. If $\alpha_{21} \leq \alpha_{11}$ then $(d_1 = \alpha_{11}, d_2 = 0)$ is trivial. If $\alpha_{21} \geq \alpha_{11}$, then, $(d_1 = \alpha_{11}, d_2 = \alpha_{21} - \alpha_{11})$ is achievable by only transmitting at the first antenna with

the full power $\alpha_{21}$ similar to the case ($\alpha_{11} \leq \alpha_{12}$ and $\alpha_{21} \leq \alpha_{22}$). Here $W_1 = (W_{1p}, W_{2p})$. $\mathrm{E}|X_{1p}|^2 = 1 - P^{-\alpha_{11}}$, $\mathrm{E}|X_{2p}|^2 = P^{-\alpha_{11}}$, and $X_1 = X_{1p} + X_{2p}$.

This completes the proof of achievability.

# 6 Proof of Theorem 1: Complex Setting

Achievability for the complex setting is essentially identical to the real setting. Here we describe the outer bound proof for the complex setting. The channel model for the complex setting is the identical to the real setting described in Section 2, except that all symbols are complex and instead of (3), The GDoF region is defined as

$$\mathcal{D} = \left\{(d_1, d_2) : \exists (R_1(P), R_2(P)) \in \mathcal{C}(P) \text{ s. t. } d_1 = \lim_{P\to\infty} \frac{R_1(P)}{C_o(P)}, d_2 = \lim_{P\to\infty} \frac{R_2(P)}{C_o(P)}\right\} \quad (45)$$

where $C_o(P)$ is the baseline reference capacity of an additive white Gaussian noise complex channel $Y = X + N$ with transmit power $P$ and unit variance noise where for complex settings it is $\log(P) + o(\log(P))$. The deterministic channel model is described similar to the real setting as follows.

The deterministic channel model has inputs $\bar{X}_1(t), \bar{X}_2(t) \in \mathbb{C}$ and outputs $\bar{Y}_1(t), \bar{Y}_2(t) \in \mathbb{C}$, defined as,

$$\bar{Y}_1(t) = \lfloor \bar{P}^{\alpha_{11}-\max(\alpha_{11},\alpha_{21})} G_{11}(t)\bar{X}_1(t)\rfloor + \lfloor \bar{P}^{\alpha_{12}-\max(\alpha_{12},\alpha_{22})} G_{12}(t)\bar{X}_2(t)\rfloor \quad (46)$$

$$\bar{Y}_2(t) = \lfloor \bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})} G_{21}(t)\bar{X}_1(t)\rfloor + \lfloor \bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})} G_{22}(t)\bar{X}_2(t)\rfloor \quad (47)$$

where the real and imaginary parts of the inputs, i.e. $\bar{X}_{kR}(t)$ and $\bar{X}_{kI}(t)$ are integers and satisfy the following per-symbol power constraint

$$\bar{X}_{1R}(t), \bar{X}_{1I}(t) \in \{0, 1, \cdots, \lceil \bar{P}^{\max(\alpha_{11},\alpha_{21})}\rceil\}, \forall t \in \mathbb{N} \quad (48)$$

$$\bar{X}_{2R}(t), \bar{X}_{2I}(t) \in \{0, 1, \cdots, \lceil \bar{P}^{\max(\alpha_{12},\alpha_{22})}\rceil\}, \forall t \in \mathbb{N} \quad (49)$$

Similar to the real setting, define the set of channel coefficient variables

$$G_{jk}^{[n]} = \{G_{jk,R}(t), t \in [n]\} \cup \{G_{jk,I}(t), t \in [n]\}, \forall j, k \in \{1, 2\}$$

where $G_{jk,R}(t), G_{jk,I}(t)$ are the real and imaginary parts of $G_{jk}(t)$, respectively. Define $G^{[n]} = \bigcup_{j,k\in\{1,2\}} G_{jk}^{[n]}$. Similar to the real setting, $\forall n \in \mathbb{N}$, and for all finite cardinality disjoint subsets $\mathbf{G}_1, \mathbf{G}_2$ of $G^{[n]}$,

$$\mathbf{G}_1 \subset G^{[n]}, \mathbf{G}_2 \subset G^{[n]}, \mathbf{G}_1 \cap \mathbf{G}_2 = \phi, |\mathbf{G}_1| < \infty, |\mathbf{G}_2| < \infty$$

the conditional PDF

$$\forall G_1, G_2, \quad f_{\mathbf{G_1}|\mathcal{T},\mathbf{G_2}}(G_1|\mathcal{T}, G_2) \leq f_{\max}^{|\mathbf{G}_1|}.$$

The generalization of the proof to the complex channel coefficients setting is, for the most part, straightforward based on the real case studied earlier. To avoid repetition, here we focus only on the differences.

**Step 2: Fano's Inequality.**

$$\begin{aligned} n(R_1+R_2) &\leq n\max(\alpha_{21},\alpha_{22})\log(P)+\Big[H(\bar{Y}_1^{[n]}|W_2,G^{[n]})-H(\bar{Y}_2^{[n]}|W_2,G^{[n]})\Big] \\ &\quad +n\ o\log(\bar{P}))+o(n) \end{aligned} \tag{50}$$

**Step 3: Functional Dependence.** and **Step 4: Aligned Image Sets.** are exactly the same as the real setting.

**Step 5. Probability that Images Align.** Similarly, given $G_{11}^{[n]}, G_{12}^{[n]}$, consider two distinct realizations of user 1's output sequence $\bar{Y}_1^{[n]}$, denoted as $\lambda^{[n]}$ and $\nu^{[n]}$, which are produced by the corresponding two realizations of the codeword $(X_1^{[n]}, X_2^{[n]})$ denoted by $(\lambda_1^{[n]}, \lambda_2^{[n]})$ and $(\nu_1^{[n]}, \nu_2^{[n]})$, respectively.

$$\begin{aligned} \lambda(t) &= \lfloor \bar{P}^{\alpha_{11}-\max(\alpha_{11},\alpha_{21})}\{G_{11,R}(t)+jG_{11,I}(t)\}\{\lambda_{1,R}(t)+j\lambda_{1,I}(t)\}\rfloor \\ &\quad +\lfloor \bar{P}^{\alpha_{12}-\max(\alpha_{12},\alpha_{22})}\{G_{12,R}(t)+jG_{12,I}(t)\}\{\lambda_{2,R}(t)+j\lambda_{2,I}(t)\}\rfloor \end{aligned} \tag{51}$$

$$\begin{aligned} \nu(t) &= \lfloor \bar{P}^{\alpha_{11}-\max(\alpha_{11},\alpha_{21})}\{G_{11,R}(t)+jG_{11,I}(t)\}\{\nu_{1,R}(t)+j\nu_{1,I}(t)\}\rfloor \\ &\quad +\lfloor \bar{P}^{\alpha_{12}-\max(\alpha_{12},\alpha_{22})}\{G_{12,R}(t)+jG_{12,I}(t)\}\{\nu_{2,R}(t)+j\nu_{2,I}(t)\}\rfloor \end{aligned} \tag{52}$$

We wish to bound the probability that the images of these two codewords align at user 2, i.e., $\nu^{[n]} \in S_{\lambda^{[n]}}$. For simplicity, consider first the single channel use setting, $n=1$. For $\nu \in S_\lambda$ we must have

$$\begin{aligned} &\lfloor \bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}\{G_{21,R}(t)+jG_{21,I}(t)\}\{\nu_{1,R}(t)+j\nu_{1,I}(t)\}\rfloor \\ &+\lfloor \bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})}\{G_{22,R}(t)+jG_{22,I}(t)\}\{\nu_{2,R}(t)+j\nu_{2,I}(t)\}\rfloor \\ =\ &\lfloor \bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}\{G_{21,R}(t)+jG_{21,I}(t)\}\{\lambda_{1,R}(t)+j\lambda_{1,I}(t)\}\rfloor \\ &+\lfloor \bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})}\{G_{22,R}(t)+jG_{22,I}(t)\}\{\lambda_{2,R}(t)+j\lambda_{2,I}(t)\}\rfloor \end{aligned} \tag{53}$$

so, both the real and imaginary part of the two sides of the equality should be equal, or in the other words,

$$\begin{aligned} &\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}\{G_{21,R}(t)\{\nu_{1,R}(t)-\lambda_{1,R}(t)\}-G_{21,I}(t)\{\nu_{1,I}(t)-\lambda_{1,I}(t)\}\} \\ &+\bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})}\{G_{22,R}(t)\{\nu_{2,R}(t)-\lambda_{2,R}(t)\}-G_{22,I}(t)\{\nu_{2,I}(t)-\lambda_{2,I}(t)\}\} \in [-2,+2] \end{aligned} \tag{54}$$

$$\begin{aligned} &\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}\{G_{21,R}(t)\{\nu_{1,I}(t)-\lambda_{1,I}(t)\}+G_{21,I}(t)\{\nu_{1,R}(t)-\lambda_{1,R}(t)\}\} \\ &+\bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})}\{G_{22,R}(t)\{\nu_{2,I}(t)-\lambda_{2,I}(t)\}+G_{22,I}(t)\{\nu_{2,R}(t)-\lambda_{2,R}(t)\}\} \in [-2,+2] \end{aligned} \tag{55}$$

Without loss of generality assume,

$$\begin{aligned} M \ \stackrel{\text{def}}{=}\ &\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}|\nu_{1,R}(t)-\lambda_{1,R}(t)| \geq \max\{\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}|\nu_{1,I}(t)-\lambda_{1,I}(t)| \\ &,\bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})}|\nu_{2,R}(t)-\lambda_{2,R}(t)|,\bar{P}^{\alpha_{22}-\max(\alpha_{12},\alpha_{22})}|\nu_{2,I}(t)-\lambda_{2,I}(t)|\} \end{aligned} \tag{56}$$

Note that from (56), if $\nu_1(t) \neq \lambda_1(t)$ then $\nu_{1,R}(t) \neq \lambda_{1,R}(t)$. So, from (54) and (55) for fixed values of $G_{22,R}, G_{22,I}$,

$$\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}\{G_{21,R}(t)\{\nu_{1,R}(t)-\lambda_{1,R}(t)\}-G_{21,I}(t)\{\nu_{1,I}(t)-\lambda_{1,I}(t)\}\} \in [a-2,a+2] \tag{57}$$

$$\bar{P}^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}\{G_{21,R}(t)\{\nu_{1,I}(t)-\lambda_{1,I}(t)\}+G_{21,I}(t)\{\nu_{1,R}(t)-\lambda_{1,R}(t)\}\} \in [b-2,b+2] \tag{58}$$

for some numbers $a, b$ which are independent from $G_{21,R}(t), G_{21,I}(t)$. So, by multiplying (57) and (58) to $|\nu_{1,R}(t) - \lambda_{1,R}(t)|$ and $|\nu_{1,I}(t) - \lambda_{1,I}(t)|$ respectively and then summing together, it can be concluded that both the $G_{21,R}(t)$ and $G_{21,I}(t)$ can get value within an interval with the length of $\frac{8}{M}$. So, the probability of alignment is computed similar to the real case, and is bounded by $\frac{64 f_{\max}^2}{M^2}$. Next, we will bound $\Re|\lambda(t) - \nu(t)|$ and, $\Im|\lambda(t) - \nu(t)|$ from above.

$$\begin{aligned}
\Re|\lambda(t) - \nu(t)| &\leq 2 + \bar{P}^{\alpha_{11} - \max(\alpha_{11}, \alpha_{21})} |G_{11,R}(t)||\nu_{1,R}(t) - \lambda_{1,R}(t)| \\
&\quad + \bar{P}^{\alpha_{11} - \max(\alpha_{11}, \alpha_{21})} |G_{11,I}(t)||\nu_{1,I}(t) - \lambda_{1,I}(t)| \\
&\quad + \bar{P}^{\alpha_{12} - \max(\alpha_{12}, \alpha_{22})} |G_{12,R}(t)||\nu_{2,R}(t) - \lambda_{2,R}(t)| \\
&\quad + \bar{P}^{\alpha_{12} - \max(\alpha_{12}, \alpha_{22})} |G_{12,I}(t)||\nu_{2,I}(t) - \lambda_{2,I}(t)| \\
&\leq 2 + M \times \bar{P}^{\max(\alpha_{11} - \alpha_{21}, \alpha_{12} - \alpha_{22})} \times \{|G_{11,R}(t)| + |G_{11,I}(t)| + |G_{12,R}(t)| + |G_{12,I}(t)|\}
\end{aligned} \tag{59}$$

$$\begin{aligned}
\Im|\lambda(t) - \nu(t)| &\leq 2 + \bar{P}^{\alpha_{11} - \max(\alpha_{11}, \alpha_{21})} |G_{11,I}(t)||\nu_{1,R}(t) - \lambda_{1,R}(t)| \\
&\quad + \bar{P}^{\alpha_{11} - \max(\alpha_{11}, \alpha_{21})} |G_{11,R}(t)||\nu_{1,I}(t) - \lambda_{1,I}(t)| \\
&\quad + \bar{P}^{\alpha_{12} - \max(\alpha_{12}, \alpha_{22})} |G_{12,I}(t)||\nu_{2,R}(t) - \lambda_{2,R}(t)| \\
&\quad + \bar{P}^{\alpha_{12} - \max(\alpha_{12}, \alpha_{22})} |G_{12,R}(t)||\nu_{2,I}(t) - \lambda_{2,I}(t)| \\
&\leq 2 + M \times \bar{P}^{\max(\alpha_{11} - \alpha_{21}, \alpha_{12} - \alpha_{22})} \times \{|G_{11,R}(t)| + |G_{11,I}(t)| + |G_{12,R}(t)| + |G_{12,I}(t)|\}
\end{aligned} \tag{60}$$

so for $\max(\Re|\lambda(t) - \nu(t)|, \Im|\lambda(t) - \nu(t)|) > 2$ using (59) and (60), the probability of alignment which was bounded by $\frac{64 f_{\max}^2}{M^2}$ can be bounded by

$$\begin{aligned}
&\frac{64 f_{\max}^2}{M^2} \\
\leq\ & \frac{1024 f_{\max}^2 \Delta_2^2 \bar{P}^{2\max(\alpha_{11} - \alpha_{21}, \alpha_{12} - \alpha_{22})}}{\max(\Re|\lambda(t) - \nu(t)| - 2, \Im|\lambda(t) - \nu(t)| - 2)^2} \\
\leq\ & \frac{1024 f_{\max}^2 \Delta_2^2 \bar{P}^{2\max(\alpha_{11} - \alpha_{21}, \alpha_{12} - \alpha_{22})}}{\max(\Re|\lambda(t) - \nu(t)| - 2, 1) \max(\Im|\lambda(t) - \nu(t)| - 2, 1)}
\end{aligned} \tag{61}$$

For $\max(\Re|\lambda(t) - \nu(t)|, \Im|\lambda(t) - \nu(t)|) \leq 2$ the probability of alignment is bounded by one. Now let us return to the case of general $n$, where we similarly have,

$$\begin{aligned}
\mathbb{P}(\lambda^{[n]} \in S_{\nu^{[n]}}) \leq\ & \max(1, 1024 f_{\max}^2 \Delta_2^2 \bar{P}^{2\max(\alpha_{11} - \alpha_{21}, \alpha_{12} - \alpha_{22})})^n \times \prod_{t: \max(\Re|\lambda(t) - \nu(t)|, \Im|\lambda(t) - \nu(t)|) > 2} \\
& \frac{1}{\max(\Re|\lambda(t) - \nu(t)| - 2, 1) \max(\Im|\lambda(t) - \nu(t)| - 2, 1)}
\end{aligned} \tag{62}$$

**Step 5. Bounding the Expected Size of Aligned Image Sets.**

$$\begin{aligned}
\mathrm{E}(|S_{\nu^{[n]}}|) &= \sum_{\lambda^n \in \{\bar{Y}_1^{[n]}\}} \mathbb{P}\left(\lambda^n \in S_{\nu^{[n]}}\right) \qquad (63)\\
&\leq \max(1024 f_{\max}^2 \Delta_2^2, 1)^n \bar{P}^{2n\max(\alpha_{11}-\alpha_{21},\alpha_{12}-\alpha_{22},0)}\\
&\quad \times \prod_{t=1}^{n}\left(\sum_{\Re|\lambda(t)-\nu(t)|:\Re|\lambda(t)-\nu(t)|\leq 2} 1 + \sum_{\Re|\lambda(t)-\nu(t)|:2<\Re|\lambda(t)-\nu(t)|\leq Q_y} \frac{1}{\Re|\lambda(t)-\nu(t)|-2}\right)\\
&\quad \times \prod_{t=1}^{n}\left(\sum_{\Im|\lambda(t)-\nu(t)|:\Im|\lambda(t)-\nu(t)|\leq 2} 1 + \sum_{\Im|\lambda(t)-\nu(t)|:2<\Im|\lambda(t)-\nu(t)|\leq Q_y} \frac{1}{\Im|\lambda(t)-\nu(t)|-2}\right)\\
&\leq \max(1024 f_{\max}^2 \Delta_2^2, 1)^n \bar{P}^{2n\max(\alpha_{11}-\alpha_{21},\alpha_{12}-\alpha_{22},0)} \times \left(\max(\alpha_{11},\alpha_{12})\log(\bar{P}) + o(\log(\bar{P}))\right)^{2n}
\end{aligned}$$

where $Q_y \leq (2\Delta_2+2)\lceil \bar{P}^{\max(\alpha_{11},\alpha_{12})}\rceil$. Substituting these bounds back into (50) we have

$$\begin{aligned}
n(R_1+R_2) &\leq n\max(\alpha_{21},\alpha_{22})\log(P)\\
&\quad + \left[H(\bar{Y}_1^{[n]}|W_2,G^{[n]}) - H(\bar{Y}_2^{[n]}|W_2,G^{[n]})\right] + n\ o\log(P)) + o(n) \qquad (64)\\
&\leq n\max(\alpha_{21},\alpha_{22})\log(P) + \log \mathrm{E}|S_{\nu^{[n]}}| + n\ o\log(P)) + o(n)\\
&\leq n\left(\max(\alpha_{21},\alpha_{22}) + \max(\alpha_{11}-\alpha_{21},\alpha_{12}-\alpha_{22},0)\right)\log(P)\\
&\quad + n\ o(\log(P)) + o(n) \qquad (65)
\end{aligned}$$

So that we obtain the GDoF bound

$$d_1 + d_2 \leq B + D \qquad (66)$$

By symmetry we also have the GDoF bound

$$d_1 + d_2 \leq A + C \qquad (67)$$

Together these two bounds give us $d_1 + d_2 \leq \min(A+C, B+D)$, completing the proof of the outer bounds for Theorem 1.

# 7 Proof of Theorem 2

For this proof, let us consider the three regimes of weak interference channel, mixed interference channel, strong interference channel based on $IC_1$

**Weak interference channel regime ($\alpha_{11} \geq \alpha_{21}$ and $\alpha_{22} \geq \alpha_{12}$).** Within this regime, consider the following three cases

1. $\max(\alpha_{12},\alpha_{21}) \leq \min(\alpha_{11},\alpha_{22})$. In this case, from the [4] and [12] we have,

$$\begin{aligned}
S(IC_1) &= \min(\alpha_{11}+\alpha_{22}-\max(\alpha_{12},\alpha_{21}), \alpha_{11}+\alpha_{22}-\alpha_{12}-\alpha_{21}+D_1) \qquad (68)\\
S(IC_2) &= \min(\alpha_{11},\alpha_{22}) \qquad (69)\\
S^1(X) &= \alpha_{11}+\alpha_{22}-\alpha_{12}-\alpha_{21}+\min(D_1,D_2,D_3,D_4) \qquad (70)\\
S^o(BC) &= \alpha_{11}+\alpha_{22}-\max(\alpha_{12},\alpha_{21}) \qquad (71)
\end{aligned}$$

where $D_1, D_2, D_3, D_4$ are defined in [12]. Note that, from (68), (70), and (71) we have, $\min(S^o(BC), S^1(X)) \leq S(IC_1) \leq \max(S(IC_1), S(IC_2)) \leq S^o(BC)$. Moreover, we also know that, for any arbitrary parameters $\max(S_{IC,1}, S_{IC,2}) \leq S_X$. Therefore (12) holds.

2. $\alpha_{12} \leq \alpha_{22} \leq \alpha_{21} \leq \alpha_{11}$. We have $\max(S(IC_1), S(IC_2)) = S^o(BC) = \alpha_{11}$, i.e., (12) holds.

3. $\alpha_{21} \leq \alpha_{11} \leq \alpha_{12} \leq \alpha_{22}$. We have $\max(S(IC_1), S(IC_2)) = S^o(BC) = \alpha_{22}$. So (12) holds.

**Mixed interference channel regime.** Without loss of generality, assume $\alpha_{11} \geq \alpha_{21}, \alpha_{12} \geq \alpha_{22}$. So, in this setting, we have $\max(S(IC_1), S(IC_2)) = S^o(BC) = \max(\alpha_{11}, \alpha_{12})$, i.e., (12) holds.

**Strong interference channel regime.** The strong interference regime maps to the weak interference regime by a relabeling of parameters. So, (12) holds here as it does in the weak interference regime.

# 8 Conclusion

The approach of [2] is developed further to fully characterize the GDoF region of the two user interference channel with partial (X channel) and full (BC) transmitter cooperation, under finite precision CSIT. While the benefits of interference alignment disappear, and along with it the non-trivial benefits of partial cooperation, full cooperation shows a remarkable benefit, which is shown to be due entirely to interference enhancement. While interference alignment (under perfect CSIT) was useful mainly when channels were of comparable strength, interference enhancement becomes more powerful as the disparity between channel strengths increases.

# References


[1] A. Lapidoth, S. Shamai, and M. Wigger, "On the capacity of fading MIMO broadcast channels with imperfect transmitter side-information," in *Proceedings of 43rd Annual Allerton Conference on Communications, Control and Computing*, Sep. 28-30, 2005.

[2] A. G. Davoodi and S. Jafar, "Settling conjectures on the collapse of degrees of freedom under finite precision CSIT," in *Proceedings of Globecom*, Dec. 2014.

[3] S. Jafar and S. Shamai, "Degrees of freedom region for the MIMO X channel," *IEEE Trans. on Information Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.

[4] R. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5534–5562, 2008.

[5] C. Huang, V. Cadambe, and S. Jafar, "Interference alignment and the generalized degrees of freedom of the X channel," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5130–5150, August 2012.

[6] S. Jafar and A. Goldsmith, "Isotropic fading vector broadcast channels: the scalar upperbound and loss in degrees of freedom," *IEEE Trans. Inform. Theory*, vol. 51, no. 3, pp. 848–857, March 2005.

[7] T. Gou, S. Jafar, and C. Wang, "On the degrees of freedom of finite state compound wireless networks," *IEEE Transactions on Information Theory*, vol. 57, no. 6, pp. 3268–3308, June 2011.

[8] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *48th Annual Allerton Conference on Communication, Control and Computing*, 2010.

[9] 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 Communication Letters*, vol. 16, no. 7, pp. 1084 – 1087, July 2012.

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

[11] A. G. I. Maric, R. Dabora, "Relaying in the presence of interference: Achievable rates, interference forwarding and outer bounds," *IEEE Transactions on Information Theory*, no. 7, July 2012.

[12] U. Niesen and M. A. Maddah-Ali, "Interference alignment: From degrees-of-freedom to constant-gap capacity approximations," *CoRR*, vol. abs/1112.4879, 2011.