UNIVERSITY OF CALIFORNIA,
IRVINE

Fundamental Limits of Robust Interference Management

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Electrical Engineering

by

Junge Wang

Dissertation Committee:
Professor Syed Jafar, Chair
Professor Hamid Jafarkhani
Professor Yanning Shen

2022

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Professor Syed Jafar. I have benefited enormously from his profound insight, immense knowledge and rigorous attitude in the journey. Without his omnipresent guidance, countless discussions and valuable feedback, this dissertation would not have been possible. His enthusiasm for research motivated and inspired me for the past four years and will be continuous in my future work. It has been a great honor to be his Ph.D student.

I would like to thank Professor Hamid Jafarkhani and Professor Yanning Shen for serving on my dissertation committee, and Professor Lee Swindlehurst, Professor Zhaoxia Yu for serving on my qualifying examination committee. I appreciate all valuable and insightful suggestions from my committee members.

Let me express my gratitude to Yao-Chia Chan, Zhuqing Jia, Bofeng Yuan and Lexiang Huang for collaborating and co-authoring papers with me. I would also like to thank my former and current colleagues including Zhen Chen, Yuxiang Lu, Nilab Ismailoglu and Yuhang Yao. It is a great pleasure to be part of this wonderful team.

Last but not the least, special thanks to my girlfriend Yue and my parents for their unconditional support and encouragement. They are the most important people in my world and I dedicate this dissertation to them.

VITA

Junge Wang

EDUCATION

Doctor of Philosophy in Electrical Engineering **2022**
University of California, Irvine *Irvine, CA*

Master of Science in Electrical Engineering **2019**
University of California, Irvine *Irvine, CA*

Bachelor of Engineering in Electrical Engineering **2017**
Xi'an Jiaotong University *Xi'an, China*

RESEARCH EXPERIENCE

Graduate Research Assistant **2017–2022**
University of California, Irvine *Irvine, California*

ABSTRACT OF THE DISSERTATION

Fundamental Limits of Robust Interference Management

By

Junge Wang

Doctor of Philosophy in Electrical Engineering

University of California, Irvine, 2022

Professor Syed Jafar, Chair

Generalized degrees of freedom (GDoF) analysis of wireless networks has contributed significantly to recent advances in the fundamental understanding of their capacity. A variety of schemes have been developed that are GDoF-optimal under the assumption that the Channel State Information at Transmitter(s) (CSIT) is perfect. However, these schemes become fragile in practice because channel uncertainty is unavoidable. Motivated by the need to understand the robust information-theoretic limits of wireless networks, we explore three settings under the assumption that the CSIT is limited to finite precision, and present their GDoF characterization.

First, the optimality of Treating Interference as Noise (TIN) for K-user Interference Channel under finite precision CSIT is investigated. TIN is found to be GDoF optimal by Geng et al. in a parameter regime called the "TIN regime" under perfect CSIT. Our result shows that TIN is GDoF optimal in a much larger regime, called the CTIN regime, under finite precision CSIT.

Next, the GDoF of the two-user interference channel are characterized for all parameter regimes under finite precision CSIT, when a limited amount of (half-duplex or full-duplex) cooperation is allowed between the transmitters. In all cases, the number of over-the-air bits that each cooperation bit buys is shown to be equal to either 0, 1, 1/2 or 1/3. The most

interesting aspect of the result is the 1/3 slope, which appears only under finite precision CSIT and strong interference, and as such has not been encountered in previous studies that invariably assumed perfect CSIT.

Finally, we explore the multi-hop interference channel. We consider the sum-GDoF of the symmetric multi-hop interference channel under finite precision CSIT. The sum-GDoF value is first characterized for the $2 \times 2 \times 2$ setting that is comprised of 2 sources, 2 relays, and 2 destinations. The result is then generalized to the $2 \times 2 \times \cdots \times 2$ setting that is comprised of L hops. Remarkably, for large L, the sum-GDoF value approaches that of the one-hop broadcast channel that is obtained by full cooperation among the two transmitters of the last hop, under finite precision CSIT. Under finite precision CSIT, a combination of classical random coding schemes that are simpler and much more robust, namely a rate-splitting between decode-and-forward and amplify-and-forward, is shown to be GDoF optimal.

Chapter 1

Introduction

1.1 Background

Information theory studies the fundamental limits of communication. The central question is: what is maximum data rate of reliable communication? Channel capacity introduced by Shannon is the answer to this question: for all communication rates above the capacity it is impossible to have reliable communication, while all rates below are indeed achievable. The capacity of the point-to-point channel is found to be the mutual information between the input and output. For the mostly studied channel setting, i.e., Additive White Gaussian Noise (AWGN) Channel, the capacity is $C = \frac{1}{2}\log(1+SNR)$ in the unit of bits/channel use.

The study of capacity is extended to multi-terminal networks, which are the focus of network information theory. The capacity of multi-terminal wireless networks remains generally open, except in very few cases such as multiple access channel and the Gaussian broadcast channel. Some of the most powerful ideas for the multi-terminal wireless networks that have come out of information-theoretic studies, e.g., dirty paper coding and interference alignment, have had very little practical impact so far in wireless networks. This is at least partially

because of the critical assumption of perfect channel knowledge at transmitters. Because channel uncertainty is unavoidable in practice, and fine aspects of structured codes are inherently combinatorial, perfect CSIT settings are not only theoretically intractable but also too optimistic to be practically useful. On the other hand, we have robust schemes that do not rely on fine channel structure. We will refer them as random coding schemes, for example, Rate-splitting, TIN, superposition coding and successive decoding. Optimal rates achieved by these schemes often have closed-form (single-letter) characterizations. These schemes are more tolerant of channel uncertainty. This motivates us to study the fundamental limits under channel uncertainty and to see if these random coding schemes are optimal under channel uncertainty.

Since exact capacity is intractable, the focus has been on capacity approximation through asymptotic (high SNR) analysis. Most relevant to our purpose are the two metrics: the degrees of freedom (DoF) and generalized degrees of freedom (GDoF). The DoF metric relies on the asymptotic proportional scaling of the transmitted powers, while GDoF metric relies on asymptotic proportional scaling of the link capacities.

To characterize the fundamental limits, it is required to prove two parts: the achievable scheme, which stands for the inner bound, and the converse, which stands for the outer bound. In the perfect CSIT setting, the achievability part is much more challenging because it requires the study of structured codes like interference alignment schemes, while the converse part is relatively straightforward. However, in the robust setting when the CSIT is available with finite precision, the achievable scheme is relatively straightforward because we may only require the random codes. The converse proof becomes challenging and the recently developed aligned images bounds [16, 17, 21, 20] are necessary.

In this dissertation, we will present the GDoF characterization for three channel settings with finite precision CSIT assumptions: the optimality of TIN for K-user interference channel, two user interference channel with limited cooperation and multi-hop Interference channel.

The aligned images bounds are essential to our converse proof. The corresponding achievable coding schemes that match the converse bound are given for the three settings.

1.2 Overview of the Dissertation

In Chapter 2, we characterize the optimality regime of Treating Interference as Noise for K-user Interference Channel under finite precision CSIT setting in the GDoF sense. The optimality regime is found to be the Convex TIN regime, which is larger than the TIN regime [28] under perfect CSIT setting. The aligned images set bounds [17] are essential in the derivations.

In chapter 3, the GDoF of 2 user Interference Channel with limited transmitter cooperation for all parameter regimes under finite precision CSIT is studied. In all cases, the number of over-the-air bits that each cooperation bit buys is shown to be equal to either $0, 1, 1/2$ or $1/3$. The most interesting aspect of the result is the $1/3$ slope, which appears only under finite precision CSIT and strong interference and was never encountered in previous studies under perfect CSIT. Indeed, the converse relies on non-trivial applications of Aligned Images bounds. The result is also extended to the 2-user X channel setting.

In chapter 4, we explore the sum-GDoF of the symmetric layered multi-hop interference channel under finite precision CSIT setting. The sum-GDoF value is first characterized for the $2 \times 2 \times 2$ channel. It is shown that the sum-GDoF does not improve even if perfect CSIT is allowed in the first hop, as long as the CSIT in the second hop is limited to finite precision. The sum-GDoF characterization is then generalized to the $2 \times 2 \times \cdots \times 2$ setting that is comprised of L hops. In terms of optimal solutions, under finite precision CSIT we found that ideas such as Interference Neutralization [11, 61, 47, 38, 59], Aligned Interference Neutralization [34, 64, 33] and Network Diagonalization [58] are too fragile to retain their

GDoF benefits, and instead rate-splitting solutions that combine amplify-and-forward and decode-and-forward principles, along with careful layering (superposition) of messages that allows each successive stage of relays to acquire more common information, are sum-GDoF optimal.

1.3 Notations and Abbreviations

The notation $(x)^+$ represents $\max(x, 0)$. For integers i, j, the notation $[i : j]$ represents the set $\{i, i+1, \cdots, j\}$ if $i \leq j$ and the empty set otherwise. $X^{[N]}$ denotes the sequence $\{X(1), X(2), \cdots, X(N)\}$. $f(x) = o(g(x))$ denotes that $\limsup_{x\to\infty} \frac{|f(x)|}{|g(x)|} = 0$. Define $\lfloor x \rfloor$ as the largest integer that is smaller than or equal to x when x is non-negative, and the smallest integer that is larger than or equal to x when x is negative.

The following table lists the abbreviations used in this dissertation.

Table 1.1: Table of abbreviations

DoF	Degrees of Freedom
GDoF	Generalized Degrees of Freedom
SNR	Signal to Noise Ratio
SINR	Signal to Interference and Noise Ratio
MISO	Multiple Input Single Output
CSIT	Channel State Information at Transmitters
CSIR	Channel State Information at Receivers
IC	Interference Channel
BC	Broadcast Channel
AIS	Aligned Images Set
TIN	Treating Interference as Noise
CTIN	Convex Treating Interference as Noise
AWGN	Additive White Gaussian Noise

Chapter 2

Optimality of Treating Interference as Noise under Finite Precision CSIT

Under finite precision CSIT setting, the optimal condition of treating interference as noise is found. Specifically, it is shown that the TIN is optimal in the CTIN regime under finite precision CSIT.

2.1 Introduction

Treating interference as noise is very attractive from both practical and theoretical perspective[28, 26, 25, 29, 27, 69]. Practically, TIN is very simple to implement and robust to the channel uncertainty. It only requires coarse knowledge of signal to interference and noise power ratio (SINR) at the transmitters. In the TIN scheme, each transmitter chooses the power control value and each receiver decodes its desired message by treating all other interference as noise. Theoretically, [49, 57, 2] found that in the noisy interference regime, Treating interference as noise achieves the sum capacity of the interference channel. For multiple input

multiple-output (MIMO) Gaussian interference channels, the extension of the noisy interference regime is found in [3]. Later, a surprising result in [28] shows that the simple scheme of TIN is GDoF optimal in the weak interference regime even with perfect CSIT. The result has been extended to the compound channel setting [26], multiple messages set [29] and security setting [25]. These results characterize the optimality of TIN under perfect CSIT setting. However, the assumption of perfect CSIT is too optimistic to be hold in practice. Thus, in this chapter we would like to characterize the optimality of TIN in the robust sense, i.e., under finite precision CSIT setting. Our new result shows that under finite precision CSIT, TIN is GDoF optimal in the larger regime called CTIN regime. Therefore, the advantages of the clever schemes (IA) disappear in this regime. Notably, TIN is not optimal in this new regime if CSIT is perfect.

2.2 Problem Statements

For GDoF studies, the K user interference channel is modeled as [16, 17]

$$Y_k(t) = \sum_{i=1}^{K} \bar{P}^{\alpha_{ki}} G_{ki}(t) X_i(t) + Z_k(t), \qquad \forall k \in [K]. \tag{2.1}$$

During the t^{th} channel use, $X_i(t), Y_k(t), Z_k(t) \in \mathbb{C}$ are, respectively, the symbol transmitted by Transmitter i subject to a normalized unit transmit power constraint, the symbol received by User k, and the zero mean unit variance additive white Gaussian noise (AWGN) at User k. $\bar{P} \triangleq \sqrt{P}$, is a nominal parameter that approaches infinity to define the GDoF limit. The exponent $\alpha_{ki} \geq 0$ is referred to as the channel strength of the link between Transmitter i and Receiver k, and is known to all transmitters and receivers. The channel coefficients $G_{ki}(t)$ are known perfectly to the receivers but only available to finite precision at the transmitters. The finite precision CSIT assumption implies that from the transmitter's perspective, the joint and conditional probability density functions of the channel coefficients

exist and the peak values of these distributions are bounded, i.e., they do not grow with P . Note that the transmitters know the distributions but not the actual realizations of $G_{ki}(t)$, therefore the transmitted symbols $X_i(t)$ are independent of the realizations of $G_{ki}(t)$. In the K user IC, there are K independent messages, one for each user, and each message is independently encoded by its corresponding transmitter. The definitions of achievable rate tuples and capacity region, $\mathcal{C}_{\text{IC}}(P)$ are standard, see e.g., [16]. The GDoF region of the K user interference channel is defined as

$$\mathcal{D}_{\text{IC}} = \left\{ (d_k)_{k\in[K]} \;\middle|\; \begin{array}{l} d_k = \lim_{P\to\infty} \frac{R_k(P)}{\log(P)}, \\ (R_k(P))_{k\in[K]} \in \mathcal{C}_{\text{IC}}(P) \end{array} \right\}. \tag{2.2}$$

The maximum sum-GDoF value is denoted $\mathcal{D}_{\Sigma,\text{IC}}$. As shown in [16] the GDoF of the channel model in (2.1) are bounded above by the GDoF of the corresponding deterministic model with inputs $\bar{X}_k(t)$ and outputs $\bar{Y}_k(t)$, defined as

$$\bar{Y}_k(t) = \sum_{i=1}^{K} \left\lfloor \bar{P}^{\alpha_{ki}-\alpha_{\max,i}} G_{ki}(t) \bar{X}_i(t) \right\rfloor, \tag{2.3}$$

where $\bar{X}_i(t) = \bar{X}_i^R(t) + j\bar{X}_i^I(t)$ with $\bar{X}_i^R(t), \bar{X}_i^I(t) \in \{0, 1, 2, \cdots, \lceil \bar{P}^{\alpha_{\max,i}} \rceil\}$, and $\alpha_{\max,i} = \max_{j\in[K]} \alpha_{ji}$. For all the parameter regimes considered in this work, $\alpha_{\max,i} = \alpha_{ii}$. The assumptions regarding channel coefficients $G_{ki}(t)$, channel knowledge at transmitters and receivers, and definitions of messages, codebooks, achievable rates, and GDoF are the same as before. Let us also recall a very useful bound for our current purpose, a special case of Lemma 1 in [17].

Lemma 2.1 (Lemma 1 in [17])**.**

$$H\left(\left(\sum_{i=1}^{K}\lfloor\bar{P}^{\lambda_i-\alpha_{\max,i}}G_{ki}(t)\bar{X}_i(t)\rfloor\right)^{[1:T]}\,\middle|\,\mathcal{G},W_S\right)$$
$$-H\left(\left(\sum_{i=1}^{K}\lfloor\bar{P}^{\nu_i-\alpha_{\max,i}}G_{k'i}(t)\bar{X}_i(t)\rfloor\right)^{[1:T]}\,\middle|\,\mathcal{G},W_S\right)$$
$$\leq \max_{i\in[K]}(\lambda_i-\nu_i)^+T\log(P)+T\ o(\log(P)), \tag{2.4}$$

where $H(Z)$ is the entropy of Z, the notation $(A(t))^{[1:T]}$ stands for $(A(1),A(2),\cdots,A(T))$, $\mathcal{G}$ is a random vector containing the values of all channel coefficients $G_{ki}(t),G_{k'i}(t)$ for $k,k',i\in[K],t\in[1:T]$, the constants λ_i,ν_i are arbitrary values between 0 and $\alpha_{\max,i}$, the set $S\subset[K]$ is an arbitrary (possibly empty) subset of users, say $S=\{i_1,i_2,\cdots,i_M\}$, and $W_S=(W_{i_1},W_{i_2},\cdots,W_{i_M})$ is comprised of the corresponding users' desired messages.

The significance of Lemma 2.1 may be intuitively understood as follows. Suppose there are K transmitters, transmitting symbols $\bar{X}_i(t)$, $i\in[K]$, independent of the realizations of the bounded density channel coefficients $G_{ki}(t),G_{k'i}(t)$, for all $i,k,k'\in[K],t\in[1:T]$, and the transmitted symbols $\bar{X}_i(t)$ can be heard at two receivers, k and k' with power levels up to λ_i and ν_i respectively. Then the maximum difference of entropies in the GDoF sense, that can exist between the signals received at the two receivers is no more than the maximum of the difference of the corresponding values of λ_i and ν_i (or zero if the maximum difference is negative). In other words, the greatest difference in the GDoF sense that can be created between the entropies of received signals at two receivers can be achieved by simply transmitting from only one antenna, which is the antenna that experiences the largest difference of channel strengths between the two receivers.

The GDoF model is essentially a generalization of the deterministic model of [5]. The significance of the GDoF model may be intuitively understood as follows. The channel

strength parameters represent the arbitrary and finite values of corresponding link SNRs and INRs in dB scale for a given network setting, i.e., $\alpha_{ii} = \log(\text{SNR}_{ii})$ and $\alpha_{ij} = \log(\text{INR}_{ij})$ (see, for example [28] for a more detailed explanation). Note that α_{ii} and α_{ij} may also be understood to be the approximate capacities of the corresponding links in isolation. Unlike the degrees of freedom (DoF) metric which proportionately scales all the transmit *powers*, the GDoF model proportionately scales all the link *capacities*. The exponential scaling of powers in the GDoF model corresponds to a linear scaling of all of the corresponding link capacities by the same factor, and this factor is $\log(P)$ (note that the isolated link with signal strength $P^{\alpha_{ij}}$ has capacity $\approx \alpha_{ij} \log(P)$, thus the scaling factor is $\log(P)$). The linear scaling of powers in the DoF model causes the ratios of capacities of any two non-zero links to approach 1 as $P \to \infty$. Thus, a very weak channel and a very strong channel become essentially *equally* strong in the DoF limit, thereby fundamentally changing the character of the original network of interest. The GDoF model on the other hand keeps the ratios of all capacities unchanged as $P \to \infty$, so that strong channels remain strong, and weak channels remain weak. The intuition behind GDoF is that if the capacities of all the individual links in a network are scaled by the *same* factor, then the overall network capacity region should scale by approximately the same factor as well — essentially a principle of scale invariance.[1] If so, then normalizing by the scaling factor $\log(P)$ should produce an approximation to the capacity region of the original finite SNR network setting. This is precisely how GDoF are measured, note the normalization by $\log(P)$ in (2.2). Indeed, the validity of this intuition is borne out by numerous bounded-gap capacity approximations that have been enabled by GDoF characterizations (e.g., [62, 32, 55, 41, 6]), starting with the original result – the capacity characterization of the 2 user interference channel within a 1 bit gap in [24].

[1]While the scaling of P may be interpreted as a physical scaling of transmit powers in the DoF metric (which unfortunately changes the character of the given network), P does not have the same interpretation of physical transmit power in GDoF. Instead, in the GDoF setting, P is just a nominal parameter, such that each value of P identifies a new network according to (2.1). These distinct networks are lumped together by the GDoF metric based on the intuition that comes from the principle of scale invariance, i.e., when normalized by $\log(P)$ all of these networks should have *approximately* the same capacity region (see also the discussion in [40]).

Asymptotic analysis under perfect CSIT often leads to fragile schemes that are difficult to translate into practice, for example the DoF of the K user interference channel have been shown in [23, 48] to depend on whether the channels take rational or irrational values – a distinction of no practical significance. Zero forcing schemes that rely on precise channel phase knowledge to cancel signals can fail catastrophically due to relatively small phase perturbations. Robust schemes are much more valuable in practice. Restricting the CSIT to finite precision naturally shifts the focus to robust schemes that rely primarily on a coarse knowledge of channel strengths at the transmitters. While the finite precision CSIT model [44, 16] allows arbitrary fading distributions subject to bounded densities, it is instructive to consider in particular the model $G_{ki}(t) = g_{ki}^R(t) + jg_{ki}^I(t)$ where $g_{ki}^R(t), g_{ki}^I(t)$ are independent and uniformly distributed over $(1-\epsilon, 1+\epsilon)$ for some arbitrarily small but positive ϵ. Interpreted this way, $G_{ki}(t)$ are seen as arbitrarily small *perturbations* in the channel state that serve primarily to limit CSIT in the channel model to ϵ-precision, while the coarse knowledge of channel strengths remains available to the transmitters in the form of the parameters α_{ij}. From a GDoF perspective, these perturbations filter out fragile schemes that rely on highly precise CSIT. Indeed, the GDoF benefits of most sophisticated interference alignment and zero forcing schemes disappear under finite precision CSIT [16]. However, the benefits of robust schemes that rely only on the knowledge of channel strengths, such as rate-splitting [14], elevated multiplexing [71], layered superposition coding [15, 5], and treating interference as noise [1, 57, 49, 28] remain accessible. Thus, GDoF characterizations under finite precision CSIT provide approximately optimal solutions for power control, rate-splitting, layered superposition based schemes that are quite robust in practice. The approximately optimal solutions serve as good initialization points for finer numerical optimizations needed at finite SNR, and inspire approximately optimal resource allocation schemes such as ITLinQ [50] and ITLinQ+ [68]. As such GDoF characterizations under finite precision CSIT are tremendously useful in bringing theory closer to practice.

2.3 Definitions

Definition 2.1 (TIN Regime)**.** *Define*

$$\mathcal{A}_{\text{TIN}} = \left\{ \begin{array}{l|l} [\alpha]_{K\times K} & \alpha_{ii} \geq \alpha_{il} + \alpha_{mi}, \\ \quad \in \mathbb{R}_+^{K\times K} & \quad \forall i,l,m \in [K], i \notin \{l,m\} \end{array} \right\}. \tag{2.5}$$

The significance of the TIN regime is that in this regime, it was shown by Geng et al. in [28] that TIN is GDoF-optimal.

Definition 2.2 (CTIN Regime)**.** *Define*

$$\mathcal{A}_{\text{CTIN}} = \left\{ \begin{array}{l|l} [\alpha]_{K\times K} & \alpha_{ii} \geq \max \left(\begin{array}{c} \alpha_{ij} + \alpha_{ji}, \\ \alpha_{ik} + \alpha_{ji} - \alpha_{jk} \end{array} \right), \\ \quad \in \mathbb{R}_+^{K\times K} & \qquad \forall i,j,k \in [K], i \notin \{j,k\} \end{array} \right\}. \tag{2.6}$$

The significance of the CTIN regime is that in this regime, it was shown by Yi and Caire in [69] that the GDoF region achievable with TIN (also known as $\mathcal{D}_{\text{TINA}}$, see Definition 2.7), is convex, without the need for time-sharing, and equal to the polyhedral TIN region over the set of all K users. The optimal GDoF region was heretofore unknown in the CTIN regime for the K user interference channel, both under perfect CSIT and under finite precision CSIT. In this chapter we settle the GDoF region in the CTIN regime under finite precision CSIT, and show that it is achieved by TIN.

Let us show that the CTIN parameter regime is significantly larger than the TIN parameter regime through the example illustrated in Fig. 2.1. This is a 3 user cyclic symmetric interference channel parameterized by channel strengths a, b. We can easily see from the right figure that TIN regime is included in the CTIN regime. If $a = 1/3, b = 2/3$, the network is inside the CTIN regime but outside the TIN regime. we could easily get that TIN

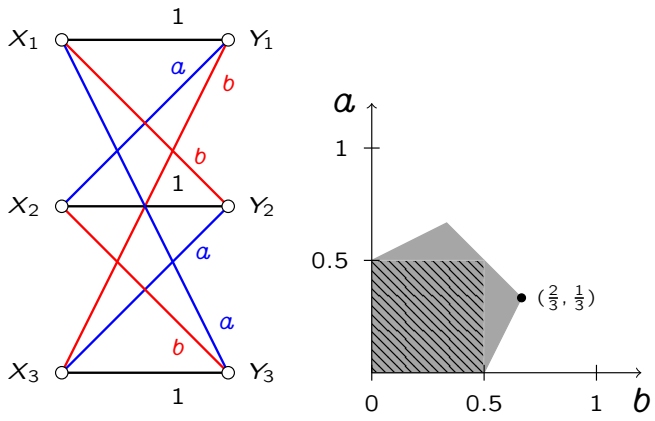

Figure 2.1: *For the* 3 *user symmetric setting shown here, the TIN regime is marked by the slanted line pattern, the CTIN regime includes the TIN regime and the region shaded in dark gray, and the SLS regime includes the CTIN regime and the region shaded in light gray.*

can achieve sum-GDoF equal to 1. However, interference alignment achieves 1.5 sum-GDoF under perfect CSIT. Thus, TIN is not GDoF optimal in the CTIN regime under perfect CSIT. We will show that TIN is GDoF optimal in the CTIN regime under finite precision CSIT, in other words, the sum-GDoF value for the example is 1 when $a = 1/3, b = 2/3$ under finite precision CSIT.

Definition 2.3 (Cycle π)**.** *A cycle* π *of length* $M > 1$ *denoted as*

$$\pi = (i_1 \to i_2 \to \cdots \to i_M \hookleftarrow) \tag{2.7}$$

is an ordered collection of links in the $K \times K$ *interference network, that includes the desired link between Transmitter* i_m *and Receiver* i_m*, and the interfering link between Transmitter* i_m *and Receiver* i_{m+1}*, for all* $m \in [1 : M]$*, where we set* $i_{M+1} = i_1$*, and the indices* $i_1, i_2, \cdots, i_M \in [K]$ *are all distinct.*

Definition 2.4 (δ_{ij})**.** *For* $i, j \in [K]$*, define*

$$\delta_{ij} = \begin{cases} \alpha_{ii} - \alpha_{ji}, & i \neq j, \\ 0, & i = j. \end{cases} \tag{2.8}$$

Definition 2.5 (Δ_π)**.** *For any cycle π of length M, $\pi = (i_1 \to i_2 \to \cdots \to i_M \leftarrow)$, define*

$$\Delta_\pi = \begin{cases} \delta_{i_1 i_2} + \delta_{i_2 i_3} + \cdots + \delta_{i_{M-1} i_M} + \delta_{i_M i_1}, & \text{if } M > 1, \\ \alpha_{i_1 i_1}, & \text{if } M = 1. \end{cases} \tag{2.9}$$

Definition 2.6 ($\mathcal{D}_{\text{P-TIN}}(S)$)**.** *For any subset of users, $S \subset [K]$, the polyhedral-TIN region [28] is defined as*

$$\mathcal{D}_{\text{P-TIN}}(S) = \left\{ (d_k : k \in [K]) \;\middle|\; \begin{array}{l} 0 = d_k, \quad \forall k \in [K] \backslash S, \\ 0 \le d_k, \qquad \forall k \in S, \\ \sum_{k \in \{\pi\}} d_k \le \Delta_\pi, \\ \qquad \forall \pi \in \Pi, \{\pi\} \subset S \end{array} \right\}. \tag{2.10}$$

The bounds, $\sum_{k\in\{\pi\}} d_k \le \Delta_\pi$, are called cycle-bounds. Note that these are not bounds on the general GDoF region, rather these are only bounds on the polyhedral TIN region for a given subset S. The sum-GDoF value of polyhedral-TIN over the set S is defined as

$$\mathcal{D}_{\Sigma,\text{P-TIN}}(S) = \max_{\mathcal{D}_{\text{P-TIN}}(S)} \sum_{k \in S} d_k. \tag{2.11}$$

If $S = [K]$, then we will simply write $\mathcal{D}_{\Sigma,\text{P-TIN}}([K]) = \mathcal{D}_{\Sigma,\text{P-TIN}}$.

A remarkable fact about the polyhedral TIN region is that even if $S_1 \subset S_2$, it is possible that the polyhedral region for S_1 is strictly larger than the polyhedral region for S_2. See the simple example at the end of this section.

Definition 2.7 ($\mathcal{D}_{\text{TINA}}$)**.** *The TINA region [28, 69] is defined as*

$$\mathcal{D}_{\text{TINA}} = \bigcup_{S : S \subset [K]} \mathcal{D}_{\text{P-TIN}}(S). \tag{2.12}$$

The sum-GDoF over the TINA region are defined as

$$\mathcal{D}_{\Sigma,\text{TINA}} = \max_{\mathcal{D}_{\text{TINA}}} \sum_{k\in[K]} d_k. \tag{2.13}$$

Thus the TINA region is a union of polyhedral TIN regions. In general this union does not produce a convex region.

2.4 Results

Our result settles the GDoF of the K user interference channel in the CTIN regime. Note that all our result is under the assumption of finite precision CSIT.

Theorem 2.1. *In the CTIN regime, TIN is GDoF optimal for the K user interference channel.*

$$[\alpha]_{K\times K} \in \mathcal{A}_{\text{CTIN}} \Rightarrow \qquad \mathcal{D}_{\text{IC}} = \mathcal{D}_{\text{TINA}}. \tag{2.14}$$

2.5 Proof

The proof for Theorem 2.1 is as follows: Consider any subset of $M > 1$ users, $S \subset [K]$, $|S| = M$, and let π be a cycle of length M, involving these M users. We will prove that the corresponding cycle bound is a valid information theoretic GDoF bound. Since we are proving an outer bound for an interference channel, without loss of generality, let us eliminate all users other than these M users. This cannot hurt the M users that remain. Now, for each of the users $\pi(m), m \in [M]$, let us apply Fano's inequality within the deterministic model of the K user interference channel as follows. As usual $o(\log(P))$ terms that are inconsequential

for GDoF are ignored for cleaner notation.

$$
\begin{aligned}
&TR_{\pi(m)} \\
\leq& I((\bar{X}_{\pi(m)}(t))^{[1:T]}; (\bar{Y}_{\pi(m)})^{[1:T]} \mid \mathcal{G}) \\
=& H((\bar{Y}_{\pi(m)})^{[1:T]} \mid \mathcal{G}) - H((\bar{Y}_{\pi(m)})^{[1:T]} \mid \mathcal{G}, (\bar{X}_{\pi(m)}(t))^{[1:T]}). \qquad (2.15)
\end{aligned}
$$

Adding these inequalities for all M users,

$$
\begin{aligned}
&T\sum_{m=1}^{M} R_{\pi(m)} \\
\leq& \sum_{m=1}^{M} \left(\begin{array}{l} H((\bar{Y}_{\pi(m)})^{[1:T]} \mid \mathcal{G}) \\ \qquad -H((\bar{Y}_{\pi(m)})^{[1:T]} \mid \mathcal{G}, (\bar{X}_{\pi(m)}(t))^{[1:T]}) \end{array} \right) \qquad (2.16) \\
=& \sum_{m=1}^{M} \left(\begin{array}{l} H((\bar{Y}_{\pi(m)})^{[1:T]} \mid \mathcal{G}) \\ \quad -H((\bar{Y}_{\pi(m+1)})^{[1:T]} \mid \mathcal{G}, (\bar{X}_{\pi(m+1)}(t))^{[1:T]}) \end{array} \right) \qquad (2.17) \\
=& \sum_{m=1}^{M} \left(\mathbb{H}_g \left(\left[\begin{array}{l} \alpha_{\pi(m)\pi(1)}, \alpha_{\pi(m)\pi(2)}, \cdots, \\ \alpha_{\pi(m)\pi(m+1)}, \cdots, \alpha_{\pi(m)\pi(M)} \end{array} \right] \right) \right. \\
&\left. - \mathbb{H}_g \left(\left[\begin{array}{l} \alpha_{\pi(m+1)\pi(1)}, \alpha_{\pi(m+1)\pi(2)}, \cdots, \\ \underbrace{\cancel{\alpha_{\pi(m+1)\pi(m+1)}}}_{\text{replace with } 0}, \cdots, \alpha_{\pi(m+1)\pi(M)} \end{array} \right] \right) \right) \qquad (2.18) \\
\leq& \sum_{m=1}^{M} \max \Big(\max_{\ell \in [M], \ell \neq m+1} \left(\alpha_{\pi(m)\pi(\ell)} - \alpha_{\pi(m+1)\pi(\ell)} \right)^{+}, \\
&\qquad\qquad \alpha_{\pi(m)\pi(m+1)} \Big) T \log(P) \qquad (2.19) \\
\leq& \sum_{m=1}^{M} \left(\alpha_{\pi(m)\pi(m)} - \alpha_{\pi(m+1)\pi(m)} \right) T \log(P) \qquad (2.20) \\
=& \sum_{m=1}^{M} \delta_{\pi(m)\pi(m+1)} T \log(P) \qquad (2.21) \\
=& \Delta_{\pi} T \log(P). \qquad (2.22)
\end{aligned}
$$

Thus, in the GDoF limit we have the bound,

$$\sum_{m=1}^{M} d_{\pi(m)} \leq \Delta_{\pi}. \tag{2.23}$$

Recall that for a cycle of length M the user indices are modulo M, i.e., $\pi(M+1) = \pi(1)$. In (2.18) we used the fact that the contribution to $\bar{Y}_{\pi(m+1)}$ from $\bar{X}_{\pi(m+1)}$ can be subtracted due to the conditioning on $\bar{X}_{\pi(m+1)}$, after which the conditioning on $\bar{X}_{\pi(m+1)}$ can be dropped because in an interference channel the inputs from different transmitters are independent of each other, i.e., $\bar{X}_{\pi(m+1)}$ is independent of all remaining inputs $\bar{X}_{\pi(j)}, j \in [M], j \neq m+1$. Removing $\bar{X}_{\pi(m+1)}$ from $\bar{Y}_{\pi(m+1)}$ is equivalent to replacing the channel strength $\alpha_{\pi(m+1)\pi(m+1)}$ with zero. In (2.19) we used the result of Lemma 2.1. In (2.20) we used the definition of the CTIN regime, which implies that,

$$\alpha_{\pi(m)\pi(m)} + \alpha_{\pi(m+1)\pi(\ell)} \geq \alpha_{\pi(m+1)\pi(m)} + \alpha_{\pi(m)\pi(\ell)}, \tag{2.24}$$

$$\alpha_{\pi(m)\pi(m)} \geq \alpha_{\pi(m+1)\pi(m)} + \alpha_{\pi(m)\pi(m+1)}. \tag{2.25}$$

Finally, it is trivial that for cycles of length $M = 1$, the cycle bound is also an information theoretic GDoF bound. Thus, we have shown that in the CTIN regime, under finite precision CSIT, for every cycle π in the K user interference channel the cycle bound is an information theoretic GDoF bound. The region described by these bounds is the polyhedral TIN region $\mathcal{D}_{\text{P-TIN}}([K])$. Therefore, $\mathcal{D}_{\text{IC}} \subset \mathcal{D}_{\text{P-TIN}}([K])$. However, $\mathcal{D}_{\text{P-TIN}}([K]) \subset \mathcal{D}_{\text{TINA}}$, and $\mathcal{D}_{\text{TINA}} \subset \mathcal{D}_{\text{IC}}$. Therefore, the TIN achievable region must be the optimal GDoF region, $\mathcal{D}_{\text{TINA}} = \mathcal{D}_{\text{IC}}$. $\square$

2.6 Summary

The optimality of TIN is studied in this chapter for K-user Interference channel. Through Aligned Images sets it is proved that TIN is optimal in the CTIN regime under finite precision

CSIT, which is larger than the optimal TIN regime under perfect CSIT.

Chapter 3

Sum-GDoF of 2-user Interference Channel with Limited Cooperation under Finite Precision CSIT

In this chapter, the GDoF of the two user interference channel are characterized for all parameter regimes under the assumption of finite precision CSIT, when a limited amount of (half-duplex or full-duplex) cooperation is allowed between the transmitters in the form of π DoF of shared messages. In all cases, the number of over-the-air bits that each cooperation bit buys is shown to be equal to either $0, 1, 1/2$ or $1/3$. The most interesting aspect of the result is the $1/3$ slope, which appears only under finite precision CSIT and strong interference, and as such has not been encountered in previous studies that invariably assumed perfect CSIT. Indeed, the achievability and converse for the parameter regimes with $1/3$ slope are the most challenging aspects of this work. In particular, the converse relies on non-trivial applications of Aligned Images bounds.

3.1 Introduction

As distributed computing applications become increasingly practical there is renewed interest in fundamental limits of cooperative communication in *robust* settings. Partially overlapping message sets naturally arise as computing tasks are distributed with some redundancy, e.g., to account for straggling nodes and adverse channel conditions [70, 51, 45, 42]. Studies of cellular communication with limited backhaul [56], unreliable cooperating links [37], and variable delay constrained messages [54] lead to similar scenarios as well. An elementary model for information theoretic analysis of such settings is an interference network with a limited amount of shared messages between the transmitters. While the body of literature on information theoretic benefits of cooperative communication is too vast to survey here (e.g., see [60]), it is notable that robust settings with finite precision CSIT remain underexplored, especially with limited cooperative capacities. Most closely related to this work are degrees of freedom (DoF) and generalized degrees of freedom (GDoF) studies in [43, 16, 21, 12, 18, 24, 65]. Connections to these prior works are explained in the remainder of this section.

Since exact capacity limits tend to be intractable, DoF and GDoF studies have emerged as an alternative path to progress for understanding the fundamental limits of wireless networks. Both metrics rely on an asymptotic scaling of signal strengths, which de-emphasizes the role of additive noise and allows these metrics to focus on the interference aspects. However, there is a fundamental distinction between these two metrics. While the DoF metric proportionately scales all *transmit powers* to infinity, the GDoF metric proportionately scales all *link capacities* to infinity. Since link capacities are essentially logarithmic functions of signal powers, the linear scaling of capacities in the GDoF framework corresponds to an exponential scaling of powers. The different approaches to the asymptotic regime naturally impart a distinct character to DoF and GDoF results. In particular, in the DoF setting where all powers are scaled linearly, all links become asymptotically equally strong from the capacity perspective (in the sense that every link carries 1 DoF), so that the distinction of strong

versus weak channels is lost in the DoF framework. On the other hand, the proportional scaling of link capacities in the GDoF metric preserves the diversity of signal strengths. Intuitively, a proportional scaling of capacities is the right scaling because of the idea of approximate scale-invariance — if all the link capacities in a network are scaled by the same factor, then the capacity of the network should scale approximately by that same factor as well. Thus, normalizing the asymptotic capacity of a network by the scaling factor should yield an approximation to the capacity of the network in the original finite SNR setting (see additional discussion of this aspect following equation (3.2) in Section 3.2). This is also the intuitive reason why capacity approximations invariably rely on the GDoF metric rather than the DoF metric.

Robustness is enforced in GDoF studies by limiting the channel state information at the transmitters (CSIT) to finite precision. Until recently, a stumbling block for robust GDoF characterizations has been the difficulty of obtaining tight converse bounds under finite precision CSIT (cf. Lapidoth-Shamai-Wigger conjecture in [43] and the PN conjecture in [63]). However, the introduction of aligned images bounds in [16] has made it possible to circumvent this challenge. Building upon this opportunity, in this work we pursue the the GDoF of the interference channel under finite precision CSIT with limited cooperation between the transmitters.

Perhaps the most powerful regime for cooperative communication is the strong interference regime, because the sharing of messages among transmitters allows essentially a re-routing of messages through stronger channels. However, this regime turns out to be also the most challenging regime for information theoretic GDoF characterizations under finite precision CSIT. For example, in [21] the GDoF are characterized for the K user broadcast channel obtained by full transmitter cooperation in a K user symmetric interference channel with partial CSIT levels. Remarkably, while the GDoF are characterized for the weak interference regime, the strong interference regime remains open. More recently, the extremal GDoF

benefits of transmitter cooperation under finite precision CSIT were characterized in [12] for large interference networks. The benefits of cooperation are shown to be substantial, but the extremal analysis is again limited to weak interference settings. Evidently the strong interference regime poses some challenges. To gauge the difficulty of robust GDoF characterizations in different parameter regimes with limited cooperation, especially the strong interference regime, in this work we explore the 2-user setting.

The main result of this work is the exact GDoF characterization of the 2-user interference channel under finite precision CSIT, when a limited amount of cooperation is allowed between the transmitters in the form of π DoF of shared messages. The following table places this work in perspective. The GDoF region (in fact the capacity region) for the 2-user

Table 3.1: *GDoF results for various levels of CSIT and transmitter cooperation.*

	Perfect CSIT	Finite Precision CSIT	Intermediate CSIT Levels
Interference Channel (no cooperation)	[24]	[17]	implied by[24, 17]
Broadcast Channel (full cooperation)	[9, 67]	[18]	[21]
Interference Channel with limited cooperation	[65]	This chapter	Open

broadcast channel (where all messages are shared) under perfect CSIT is known from [9, 67]. Under finite-precision CSIT, the GDoF region of the 2-user broadcast channel is found in [18]. It is noted in [17] that the GDoF region of 2-user interference channel (where no messages are shared) under finite-precision CSIT is the same as that under perfect CSIT [24]. The gap between full transmitter cooperation (broadcast channel) and no transmitter cooperation (interference channel) is bridged under perfect CSIT by Wang and Tse in [65], where they characterize the GDoF region of the 2-user interference channel with limited cooperation among transmitters under the assumption of perfect CSIT. This chapter bridges the corresponding gap under finite precision CSIT. Incidentally, it is also possible to bridge the gap between finite precision CSIT and perfect CSIT, by adopting 'intermediate' CSIT models where the CSIT precision gets finer with power asymptotically according to $P^{-\beta}$. By

covering the range $\beta \in (0, \infty)$, such a model bridges the gap between finite precision CSIT ($\beta = 0$) and perfect CSIT ($\beta = \infty$). For the interference channel it is implied by [24, 17], that the GDoF under intermediate CSIT levels are the same as under perfect CSIT. For the broadcast channel, the GDoF under intermediate CSIT levels are characterized in [21]. Notably, with limited cooperation and intermediate levels of CSIT, the GDoF remain open.

The significance of finite precision CSIT under limited cooperation may be noted as follows. Under perfect CSIT and limited cooperation, Wang and Tse found in [65] that each bit of cooperation buys either $0, 1$ or $1/2$ bit over-the-air. In this work, with finite precision CSIT, for all parameter regimes we show that the number of over-the-air bits that each bit of transmitter cooperation buys is either $0, 1, 1/2$ or $1/3$. Remarkably, the $1/3$ factor shows up only in the strong interference regime and only under finite precision CSIT. Indeed, the central contribution of this work, is the strong interference regime which requires the most sophisticated converse and achievability arguments.

3.2 System Model

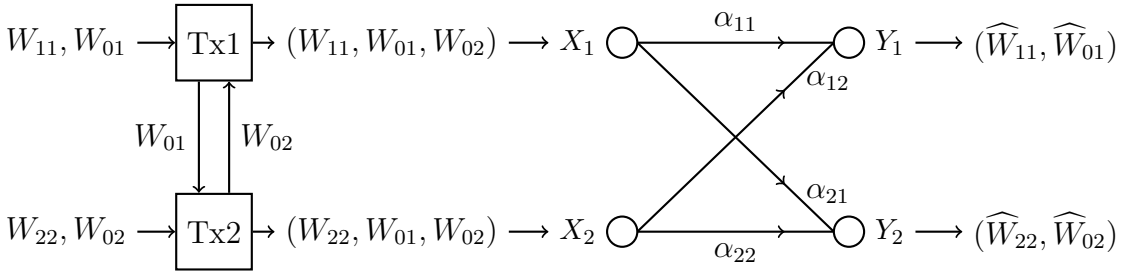

Figure 3.1: *Interference Channel with Limited Cooperation. The rates of cooperative messages W_{01}, W_{02} are limited by the cooperation capability π.*

The interference channel with limited cooperation is comprised of 4 independent messages: $W_{11}, W_{22}, W_{01}, W_{02}$. Messages W_{11}, W_{22} are the noncooperative messages that originate at Transmitters $1, 2$, and are intended for Receivers $1, 2$, respectively. Messages W_{01}, W_{02} are the cooperative messages intended for Receivers $1, 2$, respectively, with the distinction that these

messages are assumed to be known to both transmitters because they are shared among the transmitters through the limited conference link. Specifically, message W_{01} is sent through the cooperation link by Transmitter 1 to Transmitter 2, and message W_{02} is sent through the cooperation link by Transmitter 2 to Transmitter 1.

For GDoF studies, the 2-user interference channel with limited cooperation is described by the following input-output relationship.

$$Y_1(t) = \sqrt{P^{\alpha_{11}}} G_{11}(t) X_1(t) + \sqrt{P^{\alpha_{12}}} G_{12}(t) X_2(t) + Z_1(t) \tag{3.1}$$

$$Y_2(t) = \sqrt{P^{\alpha_{21}}} G_{21}(t) X_1(t) + \sqrt{P^{\alpha_{22}}} G_{22}(t) X_2(t) + Z_2(t) \tag{3.2}$$

During the t^{th} use of the channel, $X_i(t) = f_{i,t}(W_{ii}, W_{01}, W_{02}) \in \mathbb{C}$ is the symbol sent from Transmitter i, and is subject to unit transmit power constraint. The symbol observed by Receiver i is denoted $Y_i(t) \in \mathbb{C}$, and $Z_i(t) \sim \mathcal{N}_{\mathcal{C}}(0,1)$ is the zero mean unit variance additive white Gaussian noise (AWGN) at Receiver i. The variable P is referred to as power and represents a nominal parameter that approaches infinity to define the GDoF limit. The parameters $\alpha_{ki} \in \mathbb{R}^+$ represent the coarse channel strength between Transmitter i and Receiver k, respectively. To understand the intuition behind the GDoF model, it is useful to think of α_{ki} as the (approximate) capacity of the physical channel between Transmitter i and Receiver k in a given finite SNR setting that we wish to study. The GDoF model scales the capacity of every link by the same factor $\gamma = \log(P)$. Note that in the GDoF model the capacity of the physical channel between Transmitter i and Receiver k is approximately $\alpha_{ki} \log(P)$. Intuitively, the reason for this proportional scaling of capacities is the expectation of approximate scale invariance, i.e., when the capacity of every link in a network is scaled by the same factor γ, then we expect that the capacity of the overall network should scale approximately by the same factor γ as well. So normalizing the capacity of the network by $\gamma = \log(P)$ yields an approximation to the capacity of the original finite SNR network; hence the normalization by $\log(P)$ of the rates in the GDoF definition (see (3.3)).

The power P and the channel strengths α_{ki} are known to all transmitters and receivers. $G_{ki}(t) \in \mathbb{C}$ are the channel coefficient values, known perfectly to the receivers. Robustness is enforced by the assumption that the channel coefficients are only available to transmitters with *finite precision*. Recall that under the finite precision CSIT assumption, as defined in [16], the transmitters are only aware of the probability density functions of the channel coefficients, and it is assumed that all joint and conditional probability density functions of channel coefficients exist and are bounded. Precisely, it is assumed[1] that there exists a finite positive constant $f_{\max}$, such that for all finite cardinality disjoint subsets $\mathcal{G}_1, \mathcal{G}_2$ of $\mathcal{G}^{[n]}$, where $\mathcal{G}^{[n]} \triangleq \{G_{ij}(\tau) : i,j \in [2], \tau \in [n]\}$, the conditional PDF $f_{\mathcal{G}_1|\mathcal{G}_2}(g_1 \mid g_2)$ exists and is bounded above by $f_{\max}^{|\mathcal{G}_1|}$. As in [16], to avoid degenerate conditions, the channel coefficients are assumed to be bounded away from 0 and infinity, i.e., all $|G_{ki}(t)| \in [1/\Delta, \Delta]$ for some positive finite constant Δ. For example, a basic model for finite precision CSIT is of the form $G_{ij}(t) \in (1-\epsilon, 1+\epsilon)$, where ϵ is a small fixed value that represents the level of CSIT precision. Since $|G_{ij}|$ is limited to values close to 1, it is bounded away from zero and infinity. The set of all channel coefficient random variables is denoted $\mathcal{G} = \{G_{ki}(t) \mid i,k \in \{1,2\}, t \in \mathbb{Z}_+\}$.

The rates associated with messages $W_{11}, W_{22}, W_{01}, W_{02}$ are denoted as $R_{11}, R_{22}, R_{01}, R_{02}$, respectively. The definitions of probability of error, achievable rate tuples $(R_{11}, R_{22}, R_{01}, R_{02})$, codebooks and capacity region $\mathcal{C}$ are all in the standard Shannon-theoretic sense (see for example [22]). The GDoF region is defined as,

$$\mathcal{D} = \left\{ \begin{array}{l} (d_{11}, d_{22}, d_{01}, d_{02}) : \\ \\ \exists((R_{11}(P), R_{22}(P), R_{01}(P), R_{02}(P)) \in \mathcal{C}(P) \\ \\ s.t. \quad d_{11} = \lim\limits_{P\to\infty} \frac{R_{11}(P)}{\log(P)}, \quad d_{22} = \lim\limits_{P\to\infty} \frac{R_{22}(P)}{\log(P)}, \\ \qquad d_{01} = \lim\limits_{P\to\infty} \frac{R_{01}(P)}{\log(P)}, \quad d_{02} = \lim\limits_{P\to\infty} \frac{R_{02}(P)}{\log(P)}. \end{array} \right\} \tag{3.3}$$

[1]While the finite precision CSIT model allows much more general settings, a basic example of this model is where $G_{ij}(t)$ are all i.i.d. uniform in the interval $(1-\epsilon, 1+\epsilon)$, with ϵ representing a small constant that defines the fixed precision-level of available CSIT.

The total cooperation capability of the system is fixed by a given parameter π. We focus in particular on two models for cooperation, half-duplex and full-duplex, represented by the following assumptions.

$$\text{Half-duplex Assumption:} \qquad d_{01} + d_{02} \leq \pi, \tag{3.4}$$

$$\text{Full-duplex Assumption:} \qquad d_{01} \leq \frac{\pi}{2}, \quad d_{02} \leq \frac{\pi}{2}. \tag{3.5}$$

Thus, the half-duplex assumption implies that the capacity of the cooperation link is limited to π GDoF, which can be divided arbitrarily between the two one-way modes, while the full-duplex assumption implies that the capacity of the cooperation link is limited to $\frac{\pi}{2}$, which can be simultaneously utilized in both directions without mutual interference. The sum-GDoF value is the maximum value of $d_{11} + d_{22} + d_{01} + d_{02}$ across all $(d_{11}, d_{22}, d_{01}, d_{02})$ tuples in the GDoF region (3.3). It is denoted as $\mathcal{D}_{\Sigma,\text{ICLC}}$ for the half-duplex model while as $\mathcal{D}'_{\Sigma,\text{ICLC}}$ for the full-duplex model.

3.2.1 Interference Channel

The interference channel corresponds to the setting with no cooperation, i.e., $\pi = 0$, so there are no cooperative messages W_{01}, W_{02}. In [24], the GDoF region of the interference channel is characterized under perfect CSIT. As noted in [17], for the 2-user interference channel, GDoF under finite precision CSIT are the same as that under perfect CSIT. The sum-GDoF value, denoted $\mathcal{D}_{\Sigma,\text{IC}}$ is found to be,

$$\begin{aligned}\mathcal{D}_{\Sigma,\text{IC}} = \min \Big(& \max(\alpha_{11} - \alpha_{21}, \alpha_{12}) + \max(\alpha_{22} - \alpha_{12}, \alpha_{21}), \\ & \max(\alpha_{11}, \alpha_{12}) + (\alpha_{22} - \alpha_{12})^+, \\ & \max(\alpha_{21}, \alpha_{22}) + (\alpha_{11} - \alpha_{21})^+, \\ & \alpha_{11} + \alpha_{22} \Big).\end{aligned} \tag{3.6}$$

3.2.2 Broadcast Channel

The broadcast channel corresponds to unlimited cooperation, i.e., $\pi \to \infty$, so that only cooperative messages W_{01}, W_{02} are needed for the sum-GDoF characterization. The sum-GDoF value, denoted $\mathcal{D}_{\Sigma,\text{BC}}$ under finite-precision CSIT is found in [18] as,

$$\mathcal{D}_{\Sigma,\text{BC}} = \min\Big(\max(\alpha_{11}, \alpha_{12}) + \max(\alpha_{21} - \alpha_{11}, \alpha_{22} - \alpha_{12})^+, \\ \max(\alpha_{21}, \alpha_{22}) + \max(\alpha_{11} - \alpha_{21}, \alpha_{12} - \alpha_{22})^+\Big). \tag{3.7}$$

Note that unlike the interference channel, the broadcast channel suffers a loss in GDoF due to finite precision CSIT as compared to perfect CSIT.

3.2.3 Weak, Mixed and Strong Interference Regimes

The range of values of α_{ki} parameters is partitioned into three regimes, labeled weak, mixed and strong interference. These regimes are defined as follows.

$$\text{Weak interference: } \max(\alpha_{12}, \alpha_{21}) \leq \min(\alpha_{11}, \alpha_{22}) \tag{3.8}$$

$$\text{Mixed interference: } \min(\alpha_{12}, \alpha_{21}) \leq \max(\alpha_{11}, \alpha_{22}), \\ \max(\alpha_{12}, \alpha_{21}) \geq \min(\alpha_{11}, \alpha_{22}) \tag{3.9}$$

$$\text{Strong interference: } \max(\alpha_{11}, \alpha_{22}) \leq \min(\alpha_{12}, \alpha_{21}) \tag{3.10}$$

The boundaries between regimes may be considered to belong to either regime.

3.2.4 Sub-Messages

In the description of the achievable scheme, we partition messages into sub-messages, and in labeling these sub-messages we use subscripts to indicate transmitter cooperation, while the superscripts are associated with the decodability of the message. Specifically, if the subscript contains a 0 then that part of the message is shared between the two transmitters, otherwise it is not. Similarly, if the superscript is a p then that part of the message is private, i.e., only decodable at its desired receiver, otherwise it is common, i.e., decodable by both receivers. Specifically, the noncooperative message W_{ii} and cooperative message W_{0i} are split into common and private parts, so that $W_{ii} = (W_{ii}^c, W_{ii}^p), W_{0i} = (W_{0i}^c, W_{0i}^p)$, and we have the following sub-messages:

W_{11}^p: Noncooperative private message, encoded by Transmitter 1 and decoded by Receiver 1.

W_{22}^p: Noncooperative private message, encoded by Transmitter 2 and decoded by Receiver 2.

W_{11}^c: Noncooperative common message, encoded by Transmitter 1, decoded by both receivers.

W_{22}^c: Noncooperative common message, encoded by Transmitter 2, decoded by both receivers.

W_{01}^p: Cooperative private message, private part of W_{01}, encoded[2] by Transmitter 2, decoded by Receiver 1.

W_{02}^p: Cooperative private message, private part of W_{02}, encoded by Transmitter 1, decoded by Receiver 2.

W_{01}^c: Cooperative common message, common part of W_{01}, encoded by both transmitters, decoded by both receivers.

W_{02}^c: Cooperative common message, common part of W_{02}, encoded by both transmitters, decoded by both receivers.

W_0^c: Combination[3] of common parts of W_{01}, W_{02}, i.e., $W_0^c = (W_{01}^c, W_{02}^c)$.

[2]Note that even though W_{01}^p is a cooperative message, i.e., it is known to both transmitters and as such could be jointly encoded by both transmitters, our achievable schemes only require it to be encoded by Transmitter 2. Similar observation holds for W_{02}^p as well.

[3]Note that since the cooperative common message W_0^c can be encoded by both transmitters, the scope of achievable schemes exceeds traditional interference-channel schemes [35].

3.3 Results

Under the half-duplex model, the sum-GDoF value for the interference channel with limited transmitter cooperation under finite precision CSIT is characterized in the following theorem.

Theorem 3.1. *Under the half-duplex model, in the weak and mixed interference regime, we have*

$$\mathcal{D}_{\Sigma,\mathrm{ICLC}} = \min\Big(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi, \mathcal{D}_{\Sigma,\mathrm{BC}}\Big), \tag{3.11}$$

and in the strong interference regime,

$$\mathcal{D}_{\Sigma,\mathrm{ICLC}} = \min\Big(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi, \frac{\mathcal{D}_{2e}+\pi}{2}, \frac{\mathcal{D}_{3e}+\pi}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\Big), \tag{3.12}$$

where

$$\mathcal{D}_{2e} = \alpha_{12} + \alpha_{21}, \tag{3.13}$$

$$\mathcal{D}_{3e} = \alpha_{21} + \max(\alpha_{21} - \alpha_{11}, \alpha_{22}) + \alpha_{12} + \max(\alpha_{12} - \alpha_{22}, \alpha_{11}). \tag{3.14}$$

As an immediate corollary, we obtain the minimum value of π needed for the interference channel to achieve the same sum-GDoF value as the broadcast channel.

Corollary 3.1. *Let π^*_{half} denote the minimum half-duplex cooperation GDoF needed to achieve the broadcast channel bound, i.e., $\pi^*_{half} = \min_{\mathcal{D}_{\Sigma,\mathrm{ICLC}}=\mathcal{D}_{\Sigma,\mathrm{BC}}} \pi$. In the strong interference regime with an assumption $\alpha_{12} \geq \alpha_{21}$, $\pi^*_{half} > \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}}$, and its value is given*

below

$$\pi^*_{half} = \begin{cases} N - 2\max(\alpha_{11}, \alpha_{22}), & \alpha_{12}, \alpha_{21} \leq M, N \leq M + \alpha_{11}, \\ 2N - M - 3\max(\alpha_{11}, \alpha_{22}), & \alpha_{12}, \alpha_{21} \leq M, N \geq M + \alpha_{11}, \\ N + \alpha_{21} - 3\max(\alpha_{11}, \alpha_{22}), & \alpha_{12} \geq M, \alpha_{21} \leq M, \\ N + M - 3\max(\alpha_{11}, \alpha_{22}), & \alpha_{12} \geq M, \alpha_{21} \geq M, \end{cases} \tag{3.15}$$

where $M = \alpha_{11} + \alpha_{22}, N = \alpha_{12} + \alpha_{21}$. *In all other parameter regimes,* $\pi^*_{half} = \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}}$.

Our next result is the sum-GDoF characterization of the interference channel with limited *full-duplex* transmitter cooperation, under finite precision CSIT, as presented in the following theorem.

Theorem 3.2. *Under the full-duplex model, in the weak interference regime, we have*

$$\mathcal{D}'_{\Sigma,\mathrm{ICLC}} = \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi, \mathcal{D}_{\Sigma,\mathrm{BC}}\right), \tag{3.16}$$

in the mixed interference regime we have

$$\mathcal{D}'_{\Sigma,\mathrm{ICLC}} = \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \frac{\pi}{2}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right), \tag{3.17}$$

and in the strong interference regime we have

$$\mathcal{D}'_{\Sigma,\mathrm{ICLC}} = \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi, \min(\alpha_{12}, \alpha_{21}) + \frac{\pi}{2}, \frac{\mathcal{D}_{3e} + \pi}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right), \tag{3.18}$$

where $\mathcal{D}_{3e}$ *is the same as in* (3.14).

Similarly, as a corollary we obtain the minimum value of π needed for the interference channel with full-duplex cooperation to achieve the same sum-GDoF value as the broadcast channel.

Corollary 3.2. *Let* π^*_{full} *denote the minimum full-duplex cooperation GDoF needed to achieve*

the broadcast channel bound, i.e., $\pi^*_{full} = \min_{\mathcal{D}'_{\Sigma,\mathrm{ICLC}}=\mathcal{D}_{\Sigma,\mathrm{BC}}} \pi$. *In the weak interference regime,* $\pi^*_{full} = (\mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}})$. *In the mixed interference regime,* $\pi^*_{full} = 2(\mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}})$. *In the strong interference regime, where we assume* $\alpha_{21} \leq \alpha_{12}$ *without loss of generality, the value of* π^*_{full} *is given below*

$$\pi^*_{full} = \begin{cases} 2N - M - 3\max(\alpha_{11}, \alpha_{22}), & \alpha_{12}, \alpha_{21} \leq M, 2\alpha_{21} \geq M + \max(\alpha_{11}, \alpha_{22}), \\ N + \alpha_{21} - 3\max(\alpha_{11}, \alpha_{22}), & \alpha_{12} \geq M, \alpha_{21} \leq M, \alpha_{12} \leq 2\alpha_{21} - \max(\alpha_{11}, \alpha_{22}), \\ N + M - 3\max(\alpha_{11}, \alpha_{22}), & \alpha_{12} \geq M, \alpha_{21} \geq M, \alpha_{12} \leq \alpha_{21} + \min(\alpha_{11}, \alpha_{22}), \\ 2\alpha_{12} - 2\max(\alpha_{11}, \alpha_{22}), & otherwise, \end{cases} \tag{3.19}$$

where $M = \alpha_{11} + \alpha_{22}, N = \alpha_{12} + \alpha_{21}$.

To place the results in perspective, let us present some observations and examples.

1. A comparison of Theorem 3.1 with Theorem 3.2 reveals that the sum-GDoF of the full-duplex setting are identical to the half-duplex setting, i.e., for the same amount of total cooperation capability, with only two exceptions – the mixed interference regime where the full-duplex bound $\mathcal{D}_{\Sigma,\mathrm{IC}} + \frac{\pi}{2}$ is different from the half-duplex bound $\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi$, and the strong interference regime where the full-duplex bound $\min(\alpha_{12}, \alpha_{21}) + \frac{\pi}{2}$ is different from the half-duplex bound $0.5(\alpha_{12} + \alpha_{21} + \pi)$. A notable insight here is that when either of these bounds is active in the full-duplex setting, then only one-way cooperation is needed, i.e., half of the cooperation capability is wasted in the full-duplex setting.

2. The slope of sum-GDoF with respect to π for full-duplex and half-duplex, respectively, represents how many over-the-air bits are bought with each bit of total cooperation capability. Based on Theorem 3.1 and Theorem 3.2 the slope only takes values $0, 1, 1/2$, or $1/3$. Figure 3.2 shows an example where the slopes $0, 1, 1/2, 1/3$ can all be seen. Note

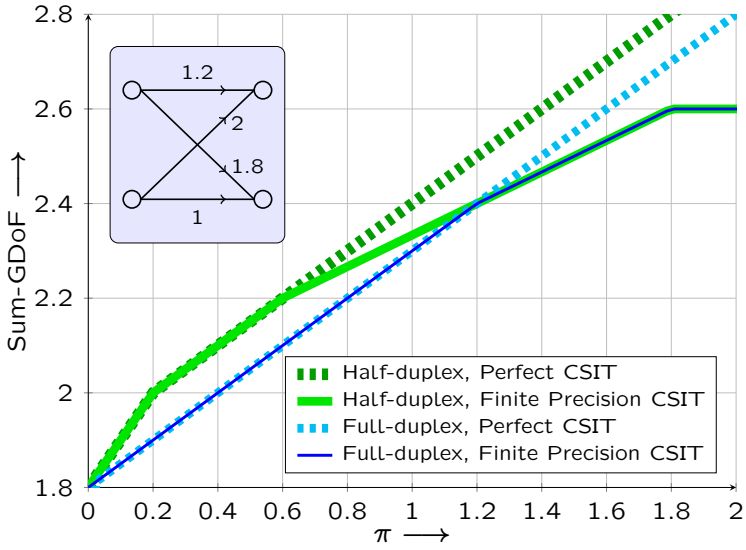

Figure 3.2: *Sum-GDoF of the interference channel* ($\alpha_{11} = 1.2, \alpha_{22} = 1, \alpha_{12} = 2, \alpha_{21} = 1.8$) *with limited cooperation for half-duplex and full-duplex settings, under perfect CSIT [65] and finite precision CSIT (this work).*

that in Figure 3.2, half-duplex cooperation has greater slope than full-duplex cooperation for $0 \leq \pi \leq 0.2$, and smaller slope than full-duplex cooperation for $0.6 \leq \pi \leq 1.2$. Thus, the incremental benefit from each additional bit of cooperation capability may be greater for either half-duplex or full-duplex cooperation in different regimes. Also note that the benefits of cooperation saturate much more quickly under finite precision CSIT.

3. In general, for both full-duplex and half-duplex settings, each incremental bit of cooperation capability buys either $0, 1, 1/2$ or $1/3$ additional over-the-air bit. Compare this to the findings in [65] for perfect CSIT, where each incremental bit of cooperation capability buys either $0, 1$, or $1/2$ additional bit over-the-air. The $1/3$ slope appears only under finite precision CSIT and only under strong interference. In fact, the GDoF bounds with slope $1/3$ are the only[4] bounds in Theorem 3.2 and Theorem 3.1 that do not appear in the perfect CSIT setting studied in [65]. Indeed, the converse and

[4]Of course, the bound corresponding to the sum-GDoF of the broadcast channel takes different values under perfect CSIT and finite precision CSIT. Under perfect CSIT, we have $\mathcal{D}_{\Sigma,\mathrm{BC}} = \max(\alpha_{11} + \alpha_{22}, \alpha_{12} + \alpha_{21})$, while under finite precision CSIT, the value is given by (3.7).

achievability for the parameter regimes where the 1/3 slope appears are the central contributions of this work.

4. A notable insight here is that when the 1/3 slope appears, it is because each incremental ϵ increase in GDoF corresponds to an ϵ increase in the GDoF of each of the three cooperative messages $W_{01}^p, W_{02}^p, W_0^c$, and a simultaneous ϵ decrease in the GDoF of each of the two noncooperative messages W_{11}, W_{22}. Therefore, the total increase in GDoF is ϵ, while the total increase in the required cooperation capability is 3ϵ, which gives us the 1/3 slope.

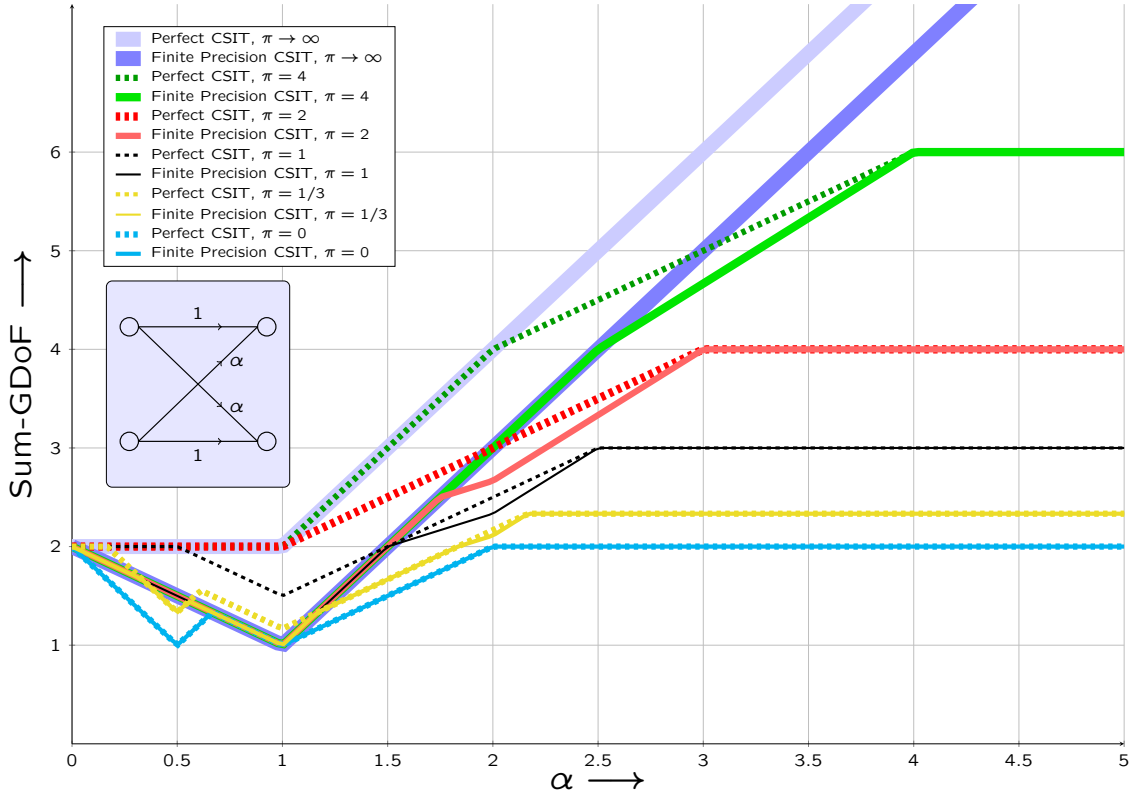

Figure 3.3: *Sum-GDoF of the symmetric interference channel ($\alpha_{11} = \alpha_{22} = 1, \alpha_{12} = \alpha_{21} = \alpha$) with limited cooperation for various half-duplex and full-duplex settings, under perfect CSIT [65] and finite precision CSIT (this work).*

5. Figure 3.3 plots the sum-GDoF value of the 2-user interference channel with limited cooperation under the symmetric setting ($\alpha_{11} = \alpha_{22} = 1, \alpha_{12} = \alpha_{21} = \alpha$) for both half-duplex and full-duplex cooperation models, under both perfect CSIT [65] and

finite-precision CSIT (this work). Note that in this symmetric setting, full-duplex cooperation and half-duplex cooperation have identical sum-GDoF as a function of π. This is because the mixed interference regime does not appear in the symmetric setting, and in the strong interference regime the full-duplex bound $\min(\alpha_{12}, \alpha_{21}) + \frac{\pi}{2}$ matches the half-duplex bound $0.5(\alpha_{12} + \alpha_{21} + \pi)$. Note that there is no cooperation gain for $2/3 \leq \alpha \leq 1$, which recovers the results in [18]. Furthermore, for any fixed cooperation capability π, as α increases, eventually the sum-GDoF with perfect CSIT match the sum-GDoF of finite precision CSIT, as they both converge to $2+\pi$. In fact, this is true more generally (even with asymmetric settings) in the following sense. For any fixed values of $(\pi, \alpha_{11}, \alpha_{22})$, as the cross-channels α_{12}, α_{21} become stronger, the sum-GDoF for both finite precision CSIT and perfect CSIT must converge to $\alpha_{11} + \alpha_{22} + \pi$. This is because as the cross-channels become stronger, each receiver is able to decode all interference and desired signals without interference, so the sum-GDoF for each user are only limited by the min-cut between its transmitter and receiver. Thus, the total GDoF of User 1, d_1 is only limited by $\alpha_{11} + d_{01}$, and similarly d_2 is only limited by $\alpha_{22} + d_{02}$, so that the sum-GDoF are only limited by $\alpha_{11} + \alpha_{22} + \pi$.

6. From Theorems 1 and 2 we can find the minimum amount of cooperation capability needed to achieve any given sum-GDoF value. In particular, the minimum amount of cooperation needed to achieve the same sum-GDoF as with unlimited cooperation, i.e., $\mathcal{D}_{\Sigma,\mathrm{BC}}$, is specified in Corollaries 1 and 2. Since the $\mathcal{D}_{\Sigma,\mathrm{BC}}$ was characterized previously in [18], a natural question is to gauge the efficiency of the achievable schemes used in [18]. Since cooperation efficiency is not a concern in [18], understandably the schemes from [18] that achieve $\mathcal{D}_{\Sigma,\mathrm{BC}}$ are in general not the most efficient in terms of the amount of cooperation needed. This is shown explicitly through the examples in Figure 3.7 and Figure 3.8. Evidently, even for settings where the sum-GDoF are already known, the most efficient solution in terms of the minimum required level of cooperation is a non-trivial question that is answered by Theorems 3.1 and 3.2.

The extension from the 2-user interference channel to 2-user X channel turns out to be quite straightforward. It is characterized by the following proposition.

Proposition 3.1. *For the 2-user X channel with limited cooperation (either half-duplex or full-duplex), the sum-GDoF value is*

$$\mathcal{D}_{\Sigma,\mathrm{XLC}} = \min(\mathcal{D}_{\Sigma,\mathrm{X}} + \pi, \mathcal{D}_{\Sigma,\mathrm{BC}}), \tag{3.20}$$

where $\mathcal{D}_{\Sigma,\mathrm{XLC}}$ denotes the sum-GDoF value of X channel with limited cooperation, $\mathcal{D}_{\Sigma,\mathrm{X}}$ denotes the sum-GDoF of X channel.

Proof. The converse is trivial because the sum-GDoF of the noncooperative messages are bounded by $\mathcal{D}_{\Sigma,\mathrm{X}}$ and the sum-GDoF of the cooperative messages are bounded by π, so the total sum-GDoF cannot exceed $\mathcal{D}_{\Sigma,\mathrm{X}} + \pi$. Also, the broadcast channel is still an outer bound. For achievability, let us first consider the weak interference channel regime, where $\max(\alpha_{12}, \alpha_{21}) \leq \min(\alpha_{11}, \alpha_{22})$. From Theorem 2 in [18] we know that in the weak interference regime, $\mathcal{D}_{\Sigma,\mathrm{X}} = \mathcal{D}_{\Sigma,\mathrm{IC}}$, which means X channel boils down to the interference channel with message W_{11}, W_{22} from the sum-GDoF perspective. Therefore $\mathcal{D}_{\Sigma,\mathrm{XLC}} = \min(\mathcal{D}_{\Sigma,\mathrm{X}} + \pi, \mathcal{D}_{\Sigma,\mathrm{BC}}) = \min(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi, \mathcal{D}_{\Sigma,\mathrm{BC}})$ whose achievability is implied by Theorem 3.1 and Theorem 3.2 in this work. The strong interference regime maps to the weak interference regime by relabeling the parameters so the sum-GDoF are established for that as well. This leaves just the mixed interference regime. But from Theorem 2 in [18], we know that $\mathcal{D}_{\Sigma,\mathrm{X}} = \mathcal{D}_{\Sigma,\mathrm{BC}}$ in the mixed interference regime, i.e., cooperation has no gain in the mixed interference regime. Thus, the sum-GDoF of the 2-user X channel with limited cooperation are easily characterized and turn out to be much simpler than the 2-user interference channel.

3.4 Converse

3.4.1 Preliminaries

Let us recall some definitions that are needed for aligned images bounds.

Definition 3.1 (Power Levels)**.** *Consider the integer valued random variables X_i over alphabet $\mathcal{X}_{\lambda_i}$*

$$\mathcal{X}_{\lambda_i} \triangleq \{0, 1, 2, \cdots, \lfloor \sqrt{P^{\lambda_i}} \rfloor - 1\}. \tag{3.21}$$

We are primarily interested in limits as $P \to \infty$, where $P \in \mathbb{R}_+$ is denoted as power. The constant λ_i refers to the power level of X_i.

Definition 3.2. *Consider integer valued random variables $X \in \mathcal{X}_\lambda$, and any non-negative real numbers λ_1, λ_2 such that $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda$, define*

$$(X)^{\lambda_2} \triangleq \Big\lfloor \frac{X}{\sqrt{P^{\lambda-\lambda_2}}} \Big\rfloor, \tag{3.22}$$

$$(X)_{\lambda_1} \triangleq X - \sqrt{P^{\lambda_1}} \Big\lfloor \frac{X}{\sqrt{P^{\lambda_1}}} \Big\rfloor, \tag{3.23}$$

$$(X)_{\lambda_1}^{\lambda_2} \triangleq \Big\lfloor \frac{(X)_{\lambda_2}}{\sqrt{P^{\lambda_1}}} \Big\rfloor. \tag{3.24}$$

In other words, $(X)^{\lambda_2}$ retrieves the top λ_2 power levels of X, $(X)_{\lambda_1}$ retrieves the bottom λ_1 power levels of X, $(X)_{\lambda_1}^{\lambda_2}$ retrieves the partition of X between levels λ_1 and λ_2. Intuitively, and in a very coarse sense, one may think of $X \in \mathcal{X}_\lambda$ as a non-negative integer value represented in $\sqrt{P}$-ary alphabet as $X = x_\lambda x_{\lambda-1} \cdots x_2 x_1$, which would require a string of λ such $\sqrt{P}$-ary symbols. In this sense, $(X)^{\lambda_2}$ corresponds to the integer value represented by the λ_2 most significant symbols of that string, i.e., $(X)^{\lambda_2} = x_\lambda x_{\lambda-1} \cdots x_{\lambda-\lambda_2+1}$, and $(X)_{\lambda_1}^{\lambda_2}$ is the integer value represented by the sub-string $x_{\lambda_2} x_{\lambda_2-1} \cdots x_{\lambda_1}$. While intuitively helpful, this

coarse understanding is also a rather extreme oversimplification, because indeed λ_1, λ_2 can take arbitrary real (non-integer) values. Thus, the partitioning with real valued λ is a non-trivial generalization of the original partitions of binary representations that appear in ADT models [5]. Note that while ADT models are used to study GDoF under perfect CSIT, the partitioning with real λ values is necessary here because of the focus on finite precision CSIT.

Definition 3.3 (Sub-section, Interval, Level, Size, Disjoint)**.** *For $X \in \mathcal{X}_\lambda$, we say that $(X)_{\lambda_1}^{\lambda_2}$ is a sub-section of X if $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda$. We refer to (λ_1, λ_2) as the interval corresponding to sub-section $(X)_{\lambda_1}^{\lambda_2}$. Furthermore, we define the lower end of the interval (λ_1, λ_2) as the 'level' of the partition, denoted as $\ell((X)_{\lambda_1}^{\lambda_2}) = \lambda_1$, and the length of the interval (λ_1, λ_2) as the 'size' of the partition, denoted as $\mathcal{T}((X)_{\lambda_1}^{\lambda_2}) = \lambda_2 - \lambda_1$. Sub-sections $(X)_{\lambda_1}^{\lambda_2}$ and $(X)_{\nu_1}^{\nu_2}$ of the same $X \in \mathcal{X}_\lambda$ are said to be disjoint if the intervals (λ_1, λ_2) and (ν_1, ν_2) are disjoint.*

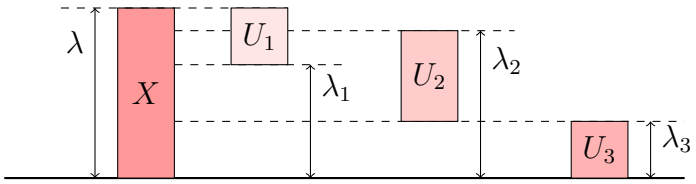

Figure 3.4: $U_1 = (X)_{\lambda_1}^{\lambda}, U_2 = (X)_{\lambda_3}^{\lambda_2}, U_3 = (X)_0^{\lambda_3}$ *are sub-sections of* X*. The corresponding levels are* $\ell(U_1) = \lambda_1, \ell(U_2) = \lambda_3, \ell(U_3) = 0$*. Additionally, the size of these sub-sections are* $\mathcal{T}(U_1) = \lambda - \lambda_1, \mathcal{T}(U_2) = \lambda_2 - \lambda_3, \mathcal{T}(U_3) = \lambda_3$*.* U_1 *and* U_3 *are disjoint.*

Our proofs will rely on sum-set inequalities based on Aligned Image sets, from [20]. While [20] presents these sum-set inequalities in generalized forms, certain simplified forms of those inequalities will suffice for our purpose. In particular, the following lemma can be obtained as a special case of Theorem 3 in [20]. Details of this simplification are presented in Appendix A.1.

Lemma 3.1. *Let $\overline{Y}(t) = \sum_{k=1}^{K} \lfloor G_k(t) \overline{X}_k(t) \rfloor$ for $\overline{X}_k \in \mathcal{X}_{\mu_k}$, and $G_k \in \mathcal{G}$ for all $k \in [K]$. For all $k \in [K]$, let S_k be a set of finitely many disjoint sub-sections of $\overline{X}_k$ (the same partitioning is applied to $\overline{X}_k(t)$ for every t), and let $\{U_1, U_2, \cdots, U_m\}$ be a subset of $\cup_{k=1}^{K} S_k$. The following*

sum-set inequality holds,

$$H(\overline{Y}^{[n]} \mid W_S, \mathcal{G}) \geq H(U_1^{[n]}, U_2^{[n]}, \cdots, U_m^{[n]} \mid W_S, \mathcal{G}) + n \times o(\log(P)), \tag{3.25}$$

if all of the following conditions (3.26)-(3.30) *are satisfied.*

$$0 = I\Big((\overline{X}_k(t), k \in [K], t \in [n]), W_S \ ; \ \mathcal{G}\Big), \tag{3.26}$$

$$\ell(U_2) \geq \mathcal{T}(U_1), \tag{3.27}$$

$$\ell(U_3) \geq \mathcal{T}(U_1) + \mathcal{T}(U_2), \tag{3.28}$$

$$\vdots \tag{3.29}$$

$$\ell(U_m) \geq \mathcal{T}(U_1) + \mathcal{T}(U_2) + \cdots + \mathcal{T}(U_{m-1}). \tag{3.30}$$

Let us intuitively explain this lemma. For our purpose, $\overline{X}_k$ are transmitted codewords (in this chapter since we only consider 2 transmitters, it suffices to set $K = 2$), and W_S is some subset of messages. Condition (3.26) requires that the messages and transmitted codewords must be independent of channel coefficient realizations (because of finite precision CSIT). Conditions (3.27)-(3.30) are quite simple as well. Imagine that we have boxes labeled $U_1, \cdots, U_m$, with the size of each box U equal to $\mathcal{T}(U)$. Each box U has a certain original position in $\overline{Y}$, which is $\ell(U)$, representing how high above the ground that box appears in $\overline{Y}$. Let's call this the original height of U. Now, suppose we stack the boxes vertically on top of each other, starting with U_1 at the bottom, and proceeding in that order until we place U_m on top of the vertical stack. In this stack, box U_j appears at a height above the ground equal to $\mathcal{T}(U_1) + \mathcal{T}(U_2) + \cdots + \mathcal{T}(U_{j-1})$ because it sits on top of boxes $U_1, U_2, \cdots, U_{j-1}$. Let's call this the new height of U_j. Then the conditions (3.27)-(3.30) simply mean that the new height of each box must be no higher than its original height. In other words, if we can vertically stack all the sub-sections without elevating the height of any sub-section above its original height in $\overline{Y}$, then the sum-set inequality (3.25) holds. Next, note that while the conditions

(3.27)-(3.30) seem to fix the stacking order from the bottom to the top as $U_1, U_2, \cdots, U_m$, the entropy on the RHS of (3.25) does not depend on the ordering of these sub-sections. So one can equivalently rearrange the ordering of $U_1, U_2, \cdots, U_m$ and apply the conditions (3.27)-(3.30) to the permuted ordering. In other words, if there exists *any* ordering in which we can vertically stack all of the sub-sections without elevating the height of any sub-section above its original height in $\overline{Y}$, then the sum-set inequality (3.25) holds.

Figure 3.5 presents a few examples that illustrate the use of Lemma 3.1.

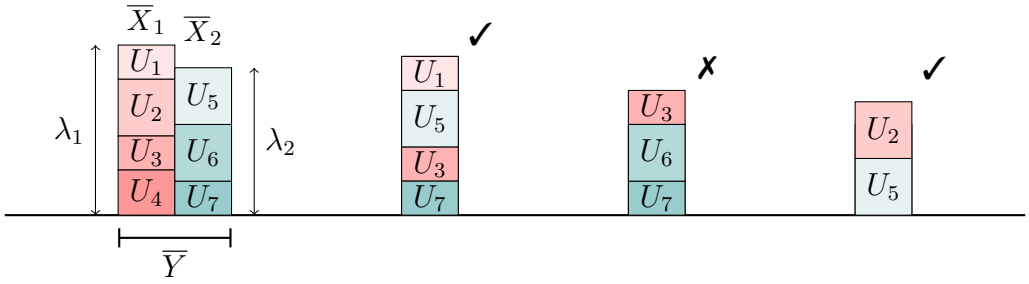

Figure 3.5: *Lemma 3.1 implies that the sum-set inequality* $H(\overline{Y}^{[n]} \mid W_S, \mathcal{G}) \geq H(U_1^{[n]}, U_5^{[n]}, U_3^{[n]}, U_7^{[n]} \mid W_S)$ *holds in the GDoF sense, because sub-sections* U_1, U_5, U_3, U_7 *can be vertically stacked without elevating any sub-section above its original height in* $\overline{Y}$*. However, Lemma 3.1 does not imply that* $H(\overline{Y}^{[n]} \mid W_S, \mathcal{G}) \geq H(U_3^{[n]}, U_6^{[n]}, U_7^{[n]} \mid W_S)$*, because it is impossible to vertically stack* U_3, U_6, U_7 *in any order without elevating at least one of them above its original position in* $\overline{Y}$*. As another example, Lemma 3.1 does imply that* $H(\overline{Y}^{[n]} \mid W_S, \mathcal{G}) \geq H(U_2^{[n]}, U_5^{[n]} \mid W_S)$ *in the GDoF sense, because* U_2 *and* U_5 *can be vertically stacked as shown, without elevating either of them above its original position in* $\overline{Y}$*.*

It is worth highlighting that the sum-set inequalities of [20] are quite powerful because of their generality. In particular, $\overline{X}_k(t)$ can have any distribution, i.e., no assumption of independence is made among the codewords sent from different transmitters. Thus, these sum-set inequalities apply not only to interference channels, but also to broadcast channel settings as well, where all the codewords can be jointly designed. On the other hand, these sum-set inequalities are specialized to finite precision CSIT. They do not apply under perfect CSIT, i.e., if the codewords can be designed with infinitely precise knowledge of the channel coefficients. For example, it is easy to see that these inequalities can be violated

if interference alignment was possible. In essence, the sum-set inequality proves that in the GDoF sense, interference alignment is not possible under finite precision CSIT for any possible design of the codebooks. Since the goal is to understand *robust* fundamental limits of wireless networks, finite precision CSIT is an absolutely essential aspect of this chapter. As such, this sum-set inequality is also essential to this work. Additional lemmas (Lemma A.1, Lemma A.2, Lemma A.3) that justify some basic manipulations that do not affect GDoF are included in Appendix A.2.

3.4.2 Proof of Converse

Let us now prove the outer bounds on the GDoF region of the interference channel with limited cooperation under finite precision CSIT, for arbitrary levels of cooperation, $d_{01} \leq \pi_{01}, d_{02} \leq \pi_{02}$. These bounds can then be specialized to obtain the tight converse for both half-duplex and full-duplex models. As noted previously, with the exception of the bounds that have slope 1/3 (as a function of π), all other bounds that we need for Theorem 3.1 and Theorem 3.2 also hold under perfect CSIT, so they can be obtained from [65]. However, for the sake of completeness we will prove all the bounds here.

The bound $\mathcal{D}_{\Sigma,\text{ICLC}} \leq \mathcal{D}_{\Sigma,\text{BC}}$ is trivial because full cooperation cannot reduce GDoF. The bound $\mathcal{D}_{\Sigma,\text{ICLC}} \leq \mathcal{D}_{\Sigma,\text{IC}} + \pi_{01} + \pi_{02}$ is also trivial because $d_{11} + d_{22} \leq \mathcal{D}_{\Sigma,\text{IC}}$ and $d_{01} \leq \pi_{01}, d_{02} \leq \pi_{02}$ by assumption. These bounds suffice for the weak interference regime in both half-duplex and full-duplex settings.

Next, let us consider the bounds that are needed for the mixed and strong interference regimes. Here we will use the Aligned Images bounds approach, starting with the deterministic model of [16] whose GDoF region contains the GDoF region of the original channel

model from above.

$$\overline{Y}_1(t) = \lfloor \sqrt{P^{\alpha_{11}-\max(\alpha_{11},\alpha_{21})}} G_{11}(t)\overline{X}_1(t) \rfloor + \lfloor \sqrt{P^{\alpha_{12}-\max(\alpha_{12},\alpha_{22})}} G_{12}(t)\overline{X}_2(t) \rfloor \tag{3.31}$$

$$\overline{Y}_2(t) = \lfloor \sqrt{P^{\alpha_{21}-\max(\alpha_{11},\alpha_{21})}} G_{21}(t)\overline{X}_1(t) \rfloor + \lfloor \sqrt{P^{\alpha_{12}-\max(\alpha_{12},\alpha_{22})}} G_{22}(t)\overline{X}_2(t) \rfloor \tag{3.32}$$

where $\overline{X}_i(t) = \overline{X}_{iR}(t) + j\overline{X}_{iI}(t)$, $i \in \{1,2\}$, and $\overline{X}_{1R}(t), \overline{X}_{1I}(t) \in \{0,1,2,\cdots,\lceil\sqrt{P^{\max(\alpha_{11},\alpha_{21})}}\rceil\}$, while $\overline{X}_{2R}(t), \overline{X}_{2I}(t) \in \{0,1,2,\cdots,\lceil\sqrt{P^{\max(\alpha_{12},\alpha_{22})}}\rceil\}$. Note that according to Lemma A.2, we can also write,

$$\overline{Y}_1(t) = \underbrace{\lfloor G_{11}(t)\left(\overline{X}_1(t)\right)^{\alpha_{11}} \rfloor + \lfloor G_{12}(t)\left(\overline{X}_2(t)\right)^{\alpha_{12}} \rfloor}_{\overline{Y}'_1(t)} + \zeta_1(t) \tag{3.33}$$

$$\overline{Y}_2(t) = \underbrace{\lfloor G_{21}(t)\left(\overline{X}_1(t)\right)^{\alpha_{21}} \rfloor + \lfloor G_{22}(t)\left(\overline{X}_2(t)\right)^{\alpha_{22}} \rfloor}_{\overline{Y}'_2(t)} + \zeta_2(t) \tag{3.34}$$

where $\zeta_1(t), \zeta_2(t)$ are integer valued random variables whose magnitude is bounded, so it does not scale with P. Specifically, according to Lemma A.2, we know that $\max(|\zeta_1(t)|, |\zeta_2(t)|) \leq 2(2+\Delta) = o(\log(P))$.

Applying Fano's inequality and ignoring the $o(\log(P))$ terms that are inconsequential for GDoF,

$$nR_{22} + nR_{02} \leq I(W_{22}, W_{02}; \overline{Y}_2^{[n]} \mid \mathcal{G}) \tag{3.35}$$

$$\leq I(W_{22}, W_{02}; \overline{Y}_2^{[n]} \mid W_{01}, \mathcal{G}) \tag{3.36}$$

$$= H(\overline{Y}_2^{[n]} \mid W_{01}, \mathcal{G}) - H(\overline{Y}_2^{[n]} \mid W_{22}, W_{01}, W_{02}, \mathcal{G}) \tag{3.37}$$

$$= H(\overline{Y}_2^{[n]} \mid W_{01}, \mathcal{G}) - H(\overline{Y}_2^{\prime[n]} \mid W_{22}, W_{01}, W_{02}, \mathcal{G}) \tag{3.38}$$

$$\leq H(\overline{Y}_2^{[n]} \mid W_{01}, \mathcal{G}) - H(\lfloor G_{21}^{[n]}(\overline{X}_1^{[n]})^{\alpha_{21}} \rfloor \mid W_{22}, W_{01}, W_{02}, \mathcal{G}) \tag{3.39}$$

$$\leq H(\overline{Y}_2^{[n]} \mid W_{01}, \mathcal{G}) - H((\overline{X}_1^{[n]})^{\alpha_{21}} \mid W_{22}, W_{01}, W_{02}, \mathcal{G}). \tag{3.40}$$

(3.38) holds because of Lemma A.1, (3.39) holds because $\overline{X}_2^{[n]}$ is a function of (W_{01}, W_{02}, W_{22}),

and (3.40) holds because of Lemma A.3.

Similarly, starting with Fano's inequality again,

$$
\begin{aligned}
nR_{11} &\leq I(W_{11}; \overline{Y}_1^{[n]} \mid \mathcal{G}) && (3.41)\\
&\leq I(W_{11}; \overline{Y}_1^{[n]} \mid W_{01}, W_{02}, W_{22}, \mathcal{G}) && (3.42)\\
&\leq H(\overline{Y}_1^{[n]} \mid W_{01}, W_{02}, W_{22}, \mathcal{G}) && (3.43)\\
&= H(\overline{Y}_1^{\prime[n]} \mid W_{01}, W_{02}, W_{22}, \mathcal{G}) && (3.44)\\
&= H\left(\lfloor G_{11}^{[n]} \left(\overline{X}_1^{[n]}\right)^{\alpha_{11}} \rfloor \mid W_{01}, W_{02}, W_{22}, \mathcal{G}\right) && (3.45)\\
&\leq H((\overline{X}_1^{[n]})^{\alpha_{11}} \mid W_{22}, W_{01}, W_{02}, \mathcal{G}), && (3.46)
\end{aligned}
$$

where (3.42) holds because $I(A;B) \leq I(A;B \mid C)$ if C is independent of A, (3.43) holds because we dropped a negative term, (3.44) holds because of Lemma A.1, (3.45) holds because $\overline{X}_2^{[n]}$ is a function of (W_{01}, W_{02}, W_{22}), and (3.46) holds because of Lemma A.3.

Adding (3.39) and (3.46), we have,

$$
\begin{aligned}
nR_{11} + nR_{22} + nR_{02} &\leq H(\overline{Y}_2^{[n]} | W_{01}, \mathcal{G}) + \\
&\qquad \left[H((\overline{X}_1^{[n]})^{\alpha_{11}} \mid W_{22}, W_{01}, W_{02}, \mathcal{G}) - H((\overline{X}_1^{[n]})^{\alpha_{21}} \mid W_{22}, W_{01}, W_{02}, \mathcal{G})\right] && (3.47)\\
&\leq H(\overline{Y}_2^{[n]} | W_{01}, \mathcal{G}) + H((\overline{X}_1^{[n]})^{\alpha_{11}} \mid (\overline{X}_1^{[n]})^{\alpha_{21}}, W_{22}, W_{01}, W_{02}, \mathcal{G}) && (3.48)\\
&\leq n \max(\alpha_{21}, \alpha_{22}) \log(P) + n(\alpha_{11} - \alpha_{21})^+ \log(P). && (3.49)
\end{aligned}
$$

Similarly,

$$
nR_{11} + nR_{22} + nR_{01} \leq n \max(\alpha_{12}, \alpha_{11}) \log(P) + n(\alpha_{22} - \alpha_{12})^+ \log(P). \quad (3.50)
$$

Dividing both sides in (3.49) and (3.50) by $n\log(P)$, and applying the GDoF limit, we obtain the following GDoF bounds:

$$d_{11} + d_{22} + d_{02} \leq \max(\alpha_{21}, \alpha_{22}) + (\alpha_{11} - \alpha_{21})^+, \tag{3.51}$$

$$d_{11} + d_{22} + d_{01} \leq \max(\alpha_{12}, \alpha_{11}) + (\alpha_{22} - \alpha_{12})^+. \tag{3.52}$$

Thus, the following bound is obtained:

$$\mathcal{D}_{\Sigma,\text{ICLC}} \leq \min(\max(\alpha_{21}, \alpha_{22}) + (\alpha_{11} - \alpha_{21})^+ + \pi_{01}, \max(\alpha_{12}, \alpha_{11}) + (\alpha_{22} - \alpha_{12})^+ + \pi_{02}). \tag{3.53}$$

This bound is useful in the mixed and strong regimes. Note that in the strong interference regime, the bound can be simplified as

$$\mathcal{D}_{\Sigma,\text{ICLC}} \leq \min(\alpha_{21} + \pi_{01}, \alpha_{12} + \pi_{02}). \tag{3.54}$$

Finally, consider the strong interference regime, and in particular, the case $\alpha_{12} \geq \alpha_{21}$. The alternative setting of $\alpha_{12} \leq \alpha_{21}$ will follow similarly. For ease of notation, define

$$A \triangleq \begin{cases} (\overline{X}_1^{[n]})_{\alpha_{22}}^{\alpha_{21}}, & \alpha_{21} \leq \alpha_{11} + \alpha_{22}, \\ (\overline{X}_1^{[n]})_{\alpha_{21}-\alpha_{11}}^{\alpha_{21}}, & \alpha_{21} \geq \alpha_{11} + \alpha_{22}, \end{cases} \tag{3.55}$$

$$B \triangleq \begin{cases} (\overline{X}_1^{[n]})_{\alpha_{21}-\alpha_{11}}^{\alpha_{22}}, & \alpha_{21} \leq \alpha_{11} + \alpha_{22}, \\ 0, & \alpha_{21} \geq \alpha_{11} + \alpha_{22}, \end{cases} \tag{3.56}$$

$$C \triangleq (\overline{X}_2^{[n]})_{\alpha_{12}-\alpha_{22}}^{\alpha_{12}}. \tag{3.57}$$

Figure 3.6 illustrates the definitions for the case $\alpha_{21} \leq \alpha_{11} + \alpha_{22}$, where the notation $[n]$ is omitted for simplicity. The case $\alpha_{21} \geq \alpha_{11} + \alpha_{22}$ can be shown similarly. Note that if

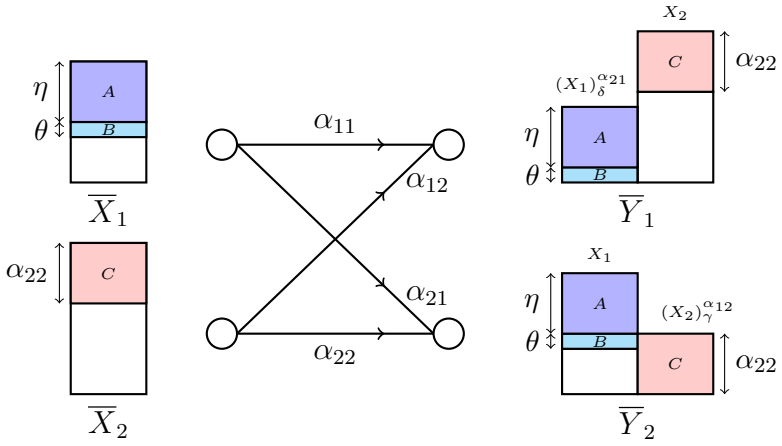

Figure 3.6: *Power level partitions* A, B, C *where* $\eta = \alpha_{21} - \alpha_{22}, \theta = \alpha_{11} + \alpha_{22} - \alpha_{21}, \delta = \alpha_{21} - \alpha_{11}, \gamma = \alpha_{12} - \alpha_{22}, \overline{X}_1 \in \mathcal{X}_{\alpha_{21}}, \overline{X}_2 \in \mathcal{X}_{\alpha_{12}}$ *and* $\alpha_{21} \leq \alpha_{11} + \alpha_{22}$.

$\alpha_{21} \leq \alpha_{11} + \alpha_{22}$, then A represents the top $\alpha_{21} - \alpha_{22}$ power levels of $\overline{X}_1^{[n]}$, and B represents the remaining power level partition of $\overline{X}_1^{[n]}$ that appears above the noise floor at Receiver 1. Otherwise, if $\alpha_{21} \geq \alpha_{11} + \alpha_{22}$, then A represents the top α_{11} levels of $\overline{X}_1^{[n]}$ and B is zero. The combination of A, B is the partition of $\overline{X}_1^{[n]}$ that is heard by Receiver 1 above the noise floor. Note that in both cases, A represents the power level partition of $\overline{X}_1^{[n]}$ that is heard above the signal due to $\overline{X}_2^{[n]}$ at Receiver 2, i.e.,

$$H(A \mid \overline{Y}_2^{[n]}, \mathcal{G}) = n \times o(\log(P)). \tag{3.58}$$

C represents the top α_{22} power levels of $\overline{X}_2^{[n]}$, which is all that Receiver 2 is able to hear from Transmitter 2. Note that the sum of power levels of A and C is always less than α_{12}, which will be important when applying Lemma 3.1.

Because C is a function of W_{22}, W_{01}, W_{02},

$$\begin{aligned} &H(C \mid W_{22}, W_{02}, \mathcal{G}) \\ &= I(C; W_{01} \mid W_{22}, W_{02}, \mathcal{G}) \\ &\leq I(A, C; W_{01} \mid W_{22}, W_{02}, \mathcal{G}) \end{aligned} \tag{3.59}$$

$$= I(A; W_{01} \mid W_{22}, W_{02}, \mathcal{G}) + I(C; W_{01} \mid W_{22}, W_{02}, A, \mathcal{G}) \tag{3.60}$$

$$\leq I(A; W_{01} \mid W_{22}, W_{02}, \mathcal{G}) + H(C \mid W_{22}, W_{02}, A, \mathcal{G}). \tag{3.61}$$

Thus,

$$I(A; W_{01} \mid W_{22}, W_{02}, \mathcal{G}) \geq H(C \mid W_{22}, W_{02}, \mathcal{G}) - H(C \mid W_{22}, W_{02}, A, \mathcal{G}). \tag{3.62}$$

At the same time, we also have the following bound,

$$H(C \mid W_{22}, W_{02}, \mathcal{G}) \geq H(C \mid W_{02}, \mathcal{G}) - H(W_{22}|W_{02}, \mathcal{G}) \tag{3.63}$$

$$= H(C \mid W_{02}, \mathcal{G}) - H(W_{22} \mid \mathcal{G}) \tag{3.64}$$

$$\geq H(C \mid W_{02}, \mathcal{G}) - I(C; W_{22} \mid W_{11}, W_{01}, W_{02}, \mathcal{G}) \tag{3.65}$$

$$\geq H(C \mid W_{02}, \mathcal{G}) - H(C \mid W_{11}, W_{01}, W_{02}, \mathcal{G}) \tag{3.66}$$

$$= I(C; W_{11}, W_{01} \mid W_{02}, \mathcal{G}) \tag{3.67}$$

$$= I(C, \overline{Y}_1^{[n]}; W_{11}, W_{01} \mid W_{02}, \mathcal{G}) - I(\overline{Y}_1^{[n]}; W_{11}, W_{01} \mid C, W_{02}, \mathcal{G}) \tag{3.68}$$

$$= I(\overline{Y}_1^{[n]}; W_{11}, W_{01} \mid W_{02}, \mathcal{G}) + I(C; W_{11}, W_{01} \mid \overline{Y}_1^{[n]}, \mathcal{G}) - H(\overline{Y}_1^{[n]} \mid C, W_{02}, \mathcal{G}) + H(\overline{Y}_1^{[n]} \mid C, W_{11}, W_{01}, W_{02}, \mathcal{G}) \tag{3.69}$$

$$\geq I(\overline{Y}_1^{[n]}; W_{11}, W_{01} \mid W_{02}, \mathcal{G}) - H(\overline{Y}_1^{[n]} \mid C, W_{02}, \mathcal{G}) \tag{3.70}$$

$$\geq I(\overline{Y}_1^{[n]}; W_{11}, W_{01} \mid \mathcal{G}) - H(\overline{Y}_1^{[n]} \mid C, W_{02}, \mathcal{G}) \tag{3.71}$$

$$\geq I(\overline{Y}_1^{[n]}; W_{11}, W_{01} \mid \mathcal{G}) - H(\overline{Y}_1^{[n]} \mid C, \mathcal{G}) \tag{3.72}$$

$$\geq nR_{11} + nR_{01} - H(\overline{Y}_1^{[n]} \mid C, \mathcal{G}), \tag{3.73}$$

where (3.70) is because mutual information and entropy are no less than zero. (3.71) is because (W_{11}, W_{01}) is independent from W_{02}. (3.72) is because conditioning cannot increase entropy.

Next, from Fano's inequality, we have

$$
\begin{aligned}
nR_{11} + nR_{01} &\leq I(W_{11}, W_{01}; \overline{Y}_1^{[n]} \mid \mathcal{G}) && (3.74)\\
&= H(\overline{Y}_1^{[n]} \mid \mathcal{G}) - H(\overline{Y}_1^{[n]} \mid W_{11}, W_{01}, \mathcal{G}) && (3.75)\\
&\leq H(\overline{Y}_1^{[n]} \mid \mathcal{G}) - H(A, C \mid W_{11}, W_{01}, \mathcal{G}) && (3.76)\\
&= H(\overline{Y}_1^{[n]} \mid \mathcal{G}) - H(A \mid W_{11}, W_{01}, \mathcal{G}) - H(C \mid W_{11}, W_{01}, A, \mathcal{G}) && (3.77)\\
&\leq H(\overline{Y}_1^{[n]} \mid \mathcal{G}) - H(A \mid W_{11}, W_{01}, \mathcal{G}) - H(C \mid W_{11}, W_{01}, A, W_{02}, \mathcal{G}) && (3.78)\\
&= H(\overline{Y}_1^{[n]} \mid \mathcal{G}) - H(A \mid W_{11}, W_{01}, \mathcal{G}) - H(C \mid W_{11}, W_{01}, W_{02}, \mathcal{G}) && (3.79)\\
&= H(\overline{Y}_1^{[n]} \mid \mathcal{G}) - H(A \mid W_{11}, W_{01}, \mathcal{G}) - nR_{22} && (3.80)\\
&= H(\overline{Y}_1^{[n]} \mid \mathcal{G}) - I(A; W_{22}, W_{02} \mid W_{11}, W_{01}, \mathcal{G}) - nR_{22} && (3.81)\\
&\leq H(\overline{Y}_1^{[n]} \mid \mathcal{G}) - I(A; W_{22}, W_{02} \mid \mathcal{G}) - nR_{22}, && (3.82)
\end{aligned}
$$

where (3.76) is due to Lemma 3.1: A,C can be stacked vertically (See Figure 3.6) without elevating either of them above their original height in $\overline{Y}_1$. (3.78) is because conditioning cannot increase entropy, (3.79) is because A is a function of W_{11}, W_{01}, W_{02}. (3.80) is because message W_{22} can only be transmitted through C as it is the partition above the noise floor that is sent from Transmitter 2 to Receiver 2. Rearranging the above inequality we get

$$
I(A; W_{22}, W_{02} \mid \mathcal{G}) \leq H(\overline{Y}_1^{[n]} \mid \mathcal{G}) - n(R_{11} + R_{22} + R_{01}). \tag{3.83}
$$

Next, applying Fano's inequality at Receiver 2, we have

$$
\begin{aligned}
&nR_{22} + nR_{02}\\
&\quad \leq I(W_{22}, W_{02}; \overline{Y}_2^{[n]} \mid \mathcal{G}) && (3.84)\\
&\quad \leq I(W_{22}, W_{02}; \overline{Y}_2^{[n]}, A \mid \mathcal{G}) && (3.85)\\
&\quad = I(W_{22}, W_{02}; A \mid \mathcal{G}) + I(W_{22}, W_{02}; \overline{Y}_2^{[n]} \mid A, \mathcal{G}) && (3.86)
\end{aligned}
$$

$$= I(W_{22}, W_{02}; A \mid \mathcal{G}) + H(\overline{Y}_2^{[n]} \mid A, \mathcal{G}) - H(\overline{Y}_2^{[n]} \mid A, W_{22}, W_{02}, \mathcal{G}) \tag{3.87}$$

$$= I(W_{22}, W_{02}; A \mid \mathcal{G}) + H(\overline{Y}_2^{[n]} \mid A, \mathcal{G}) - H(\overline{Y}_2^{[n]} \mid W_{22}, W_{02}, \mathcal{G}) + I(\overline{Y}_2^{[n]}; A \mid W_{22}, W_{02}, \mathcal{G}) \tag{3.88}$$

$$\leq I(W_{22}, W_{02}; A \mid \mathcal{G}) + H(\overline{Y}_2^{[n]} \mid A, \mathcal{G}) - H(A, C \mid W_{22}, W_{02}, \mathcal{G}) + I(\overline{Y}_2^{[n]}; A \mid W_{22}, W_{02}, \mathcal{G}) \tag{3.89}$$

$$\leq I(W_{22}, W_{02}; A \mid \mathcal{G}) + H(\overline{Y}_2^{[n]} \mid A, \mathcal{G}) - H(C \mid A, W_{22}, W_{02}, \mathcal{G}), \tag{3.90}$$

where (3.89) is implied by Lemma 3.1, and (3.90) is because of (3.58). Combining (3.83) and (3.90),

$$H(C \mid A, W_{22}, W_{02}, \mathcal{G}) \leq H(\overline{Y}_1^{[n]} \mid \mathcal{G}) + H(\overline{Y}_2^{[n]} \mid A, \mathcal{G}) - n(R_{11} + 2R_{22} + R_{01} + R_{02}). \tag{3.91}$$

Combining (3.62), (3.73), (3.91), we have

$$I(A; W_{01} \mid W_{22}, W_{02}, \mathcal{G}) \geq n(2R_{11} + 2R_{22} + 2R_{01} + R_{02}) - H(\overline{Y}_1^{[n]} \mid C, \mathcal{G}) - H(\overline{Y}_1^{[n]} \mid \mathcal{G}) - H(\overline{Y}_2^{[n]} \mid A, \mathcal{G}). \tag{3.92}$$

Using again Lemma 3.1 we have

$$H(\overline{Y}_2^{[n]} \mid W_{22}, W_{02}, \mathcal{G}) \geq H(A, B \mid W_{22}, W_{02}, \mathcal{G}). \tag{3.93}$$

Therefore, from Fano's inequality, we have

$$nR_{22} + nR_{02} \leq I(W_{22}, W_{02}; \overline{Y}_2^{[n]} \mid \mathcal{G}) \tag{3.94}$$

$$= H(\overline{Y}_2^{[n]} \mid \mathcal{G}) - H(\overline{Y}_2^{[n]} \mid W_{22}, W_{02}, \mathcal{G}) \tag{3.95}$$

$$\leq H(\overline{Y}_2^{[n]} \mid \mathcal{G}) - H(A, B \mid W_{22}, W_{02}, \mathcal{G}) \tag{3.96}$$

$$\leq H(\overline{Y}_2^{[n]} \mid \mathcal{G}) - H(A \mid W_{22}, W_{02}, \mathcal{G}) - H(B \mid W_{22}, W_{02}, A, \mathcal{G}). \tag{3.97}$$

Rearranging the terms we get

$$H(B \mid W_{22}, W_{02}, A, \mathcal{G}) \leq H(\overline{Y}_2^{[n]} \mid \mathcal{G}) - n(R_{22} + R_{02}) - H(A \mid W_{22}, W_{02}, \mathcal{G}). \tag{3.98}$$

Message W_{11} can only be transmitted through A, B as it is the partition above the noise floor that is sent from Transmitter 1 to Receiver 1, such that W_{11} can be successfully decoded by User 1. Therefore,

$$nR_{11} \leq H(A, B \mid W_{22}, W_{02}, W_{01}, \mathcal{G}) \tag{3.99}$$

$$= H(A \mid W_{22}, W_{02}, W_{01}, \mathcal{G}) + H(B \mid W_{22}, W_{02}, W_{01}, A, \mathcal{G}) \tag{3.100}$$

$$\leq H(A \mid W_{22}, W_{02}, W_{01}, \mathcal{G}) + H(B \mid W_{22}, W_{02}, A, \mathcal{G}). \tag{3.101}$$

Combining (3.98) and (3.101), we get

$$I(A; W_{01} \mid W_{02}, W_{22}, \mathcal{G}) \leq H(\overline{Y}_2^{[n]} \mid \mathcal{G}) - n(R_{11} + R_{22} + R_{02}). \tag{3.102}$$

Because (3.102) and (3.92) are upper and lower bound on the same mutual information, combining them we have

$$3n(R_{11} + R_{22}) + 2n(R_{01} + R_{02}) \leq H(\overline{Y}_2^{[n]} \mid \mathcal{G}) + H(\overline{Y}_2^{[n]} \mid A, \mathcal{G}) + H(\overline{Y}_1^{[n]} \mid \mathcal{G}) + H(\overline{Y}_1^{[n]} \mid C, \mathcal{G}). \tag{3.103}$$

Note that the following bounds hold, with $o(\log(P))$ terms omitted.

$$H(\overline{Y}_1^{[n]} \mid \mathcal{G}) \leq n\alpha_{12} \log(P), \tag{3.104}$$

$$H(\overline{Y}_1^{[n]} \mid C, \mathcal{G}) \leq n \max(\alpha_{12} - \alpha_{22}, \alpha_{11}) \log(P), \tag{3.105}$$

$$H(\overline{Y}_2^{[n]} \mid \mathcal{G}) \leq n\alpha_{21} \log(P), \tag{3.106}$$

$$H(\overline{Y}_2^{[n]} \mid A, \mathcal{G}) \leq n \max(\alpha_{21} - \alpha_{11}, \alpha_{22}) \log(P). \tag{3.107}$$

$$\tag{3.108}$$

Thus, (3.103) yields the GDoF bound

$$3d_{11} + 3d_{22} + 2d_{01} + 2d_{02} \leq D_{3e}. \tag{3.109}$$

Combining it with the assumption $d_{01} \leq \pi_{01}, d_{02} \leq \pi_{02}$, we get the bound

$$\mathcal{D}_{\Sigma,\text{ICLC}} \leq \frac{D_{3e} + \pi_{01} + \pi_{02}}{3}. \tag{3.110}$$

Proceeding similarly, the same bound is obtained for $\alpha_{21} \geq \alpha_{12}$.

At this point, let us list the bounds that we have shown along with the regimes where they are useful.

Weak Interference Regime:

$$\mathcal{D}_{\Sigma,\text{ICLC}} \leq \min\Big(\mathcal{D}_{\Sigma,\text{IC}} + \pi_{01} + \pi_{02}, \mathcal{D}_{\Sigma,\text{BC}}\Big), \tag{3.111}$$

Mixed Interference Regime:

$$\mathcal{D}_{\Sigma,\text{ICLC}} \leq \min\Big(\mathcal{D}_{\Sigma,\text{IC}} + \pi_{01} + \pi_{02}, \max(\alpha_{21}, \alpha_{22}) + (\alpha_{11} - \alpha_{21})^+ + \pi_{01},$$

$$\max(\alpha_{12}, \alpha_{11}) + (\alpha_{22} - \alpha_{12})^+ + \pi_{02}, \mathcal{D}_{\Sigma,\text{BC}}\Big), \tag{3.112}$$

Strong Interference Regime:

$$\mathcal{D}_{\Sigma,\text{ICLC}} \leq \min\Big(\mathcal{D}_{\Sigma,\text{IC}} + \pi_{01} + \pi_{02}, \alpha_{12} + \pi_{02}, \alpha_{21} + \pi_{01}, \frac{\mathcal{D}_{3e} + \pi_{01} + \pi_{02}}{3}, \mathcal{D}_{\Sigma,\text{BC}}\Big). \tag{3.113}$$

Next, we show how these bounds provide a tight converse for Theorem 3.1 as well as Theorem 3.2. Combining the above bounds with our assumptions (3.4) for half-duplex setting and (3.5) for full-duplex setting, the converse bounds for all regimes can be found with a few simple derivations, as shown in the following section.

3.4.3 Converse for Theorem 3.1 and Theorem 3.2

Weak Interference

First consider the weak interference regime where we apply the bound (3.111). Setting $\pi_{01} + \pi_{02} \leq \pi$ for the half-duplex setting we recover the tight converse bound, $\mathcal{D}_{\Sigma,\text{ICLC}} \leq \min(\mathcal{D}_{\Sigma,\text{BC}}, \mathcal{D}_{\Sigma,\text{IC}} + \pi)$. Similarly, setting $\pi_{01} \leq \frac{\pi}{2}, \pi_{02} \leq \frac{\pi}{2}$ for the full-duplex setting, we obtain the tight converse bound, $\mathcal{D}'_{\Sigma,\text{ICLC}} \leq \min(\mathcal{D}_{\Sigma,\text{BC}}, \mathcal{D}_{\Sigma,\text{IC}} + \pi)$.

Mixed Interference

Next consider the mixed interference regime. The converse for the half-duplex case with mixed interference is trivial because the bounds are identical to the weak interference regime. So let us focus on the full-duplex case. It follows from the sum-GDoF under finite precision

CSIT of the broadcast channel (reference [18], summarized in (3.7)), and the interference channel without cooperation (reference [65, 18], summarized in (3.6)), that in the mixed interference regime, there is no cooperation gain, i.e., $\mathcal{D}_{\Sigma,\mathrm{BC}} = \mathcal{D}_{\Sigma,\mathrm{IC}}$, when either of the following conditions holds.

1. $\alpha_{11} + \alpha_{22} \geq \alpha_{12} + \alpha_{21}$
2. $\min(\alpha_{11}, \alpha_{22}) \leq \min(\alpha_{12}, \alpha_{21}) \leq \max(\alpha_{12}, \alpha_{21}) \leq \max(\alpha_{11}, \alpha_{22})$

In both cases the trivial bound $\mathcal{D}_{\Sigma,\mathrm{ICLC}} \leq \mathcal{D}_{\Sigma,\mathrm{BC}}$ is tight. Henceforth in this section we will only consider the remainder of the mixed interference regime, which excludes $\alpha_{11} + \alpha_{22} \geq \alpha_{12} + \alpha_{21}$ and $\min(\alpha_{11}, \alpha_{22}) \leq \min(\alpha_{12}, \alpha_{21}) \leq \max(\alpha_{12}, \alpha_{21}) \leq \max(\alpha_{11}, \alpha_{22})$.

Let us assume without loss of generality that $\alpha_{22} \leq \alpha_{11}$. Next, let us define $\max(\alpha_{21}, \alpha_{22}) + (\alpha_{11} - \alpha_{21})^+$ as Λ_1, and similarly $\max(\alpha_{12}, \alpha_{11}) + (\alpha_{22} - \alpha_{12})^+$ as Λ_2, so the bound (3.112) can be written as:

$$\mathcal{D}_{\Sigma,\mathrm{ICLC}} \leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi_{01} + \pi_{02}, \Lambda_1 + \pi_{01}, \Lambda_2 + \pi_{02}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right).$$

Now let us show that one of Λ_1, Λ_2 is equal to $\mathcal{D}_{\Sigma,\mathrm{IC}}$ and the other is equal to $\mathcal{D}_{\Sigma,\mathrm{BC}}$. This will be useful to simplify the bound later. We have the following four cases.

- $\alpha_{21} \leq \alpha_{22} \leq \alpha_{11} \leq \alpha_{12}$

  Λ_1 is $\alpha_{11} + \alpha_{22} - \alpha_{21} = \mathcal{D}_{\Sigma,\mathrm{IC}}$, and Λ_2 is $\alpha_{12} = \mathcal{D}_{\Sigma,\mathrm{BC}}$.

- $\alpha_{12} \leq \alpha_{22} \leq \alpha_{11} \leq \alpha_{21}$

  Λ_1 is $\alpha_{21} = \mathcal{D}_{\Sigma,\mathrm{BC}}$, and Λ_2 is $\alpha_{11} + \alpha_{22} - \alpha_{12} = \mathcal{D}_{\Sigma,\mathrm{IC}}$.

- $\alpha_{22} \leq \alpha_{12} \leq \alpha_{11} \leq \alpha_{21}$

  Λ_1 is $\alpha_{21} = \mathcal{D}_{\Sigma,\mathrm{BC}}$, and Λ_2 is $\alpha_{11} = \mathcal{D}_{\Sigma,\mathrm{IC}}$.

- $\alpha_{22} \leq \alpha_{21} \leq \alpha_{11} \leq \alpha_{12}$

  Λ_1 is $\alpha_{11} = \mathcal{D}_{\Sigma,\mathrm{IC}}$, and Λ_2 is $\alpha_{12} = \mathcal{D}_{\Sigma,\mathrm{BC}}$.

Next, let us apply the bound (3.112) to the full-duplex setting (3.5) which corresponds to $\pi_{01} \leq \frac{\pi}{2}, \pi_{02} \leq \frac{\pi}{2}$. In the mixed interference regime with $\alpha_{11} + \alpha_{22} \leq \alpha_{12} + \alpha_{21}$, we have,

$$\begin{aligned}
\mathcal{D}'_{\Sigma,\mathrm{ICLC}} &\leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi_{01} + \pi_{02}, \Lambda_1 + \pi_{01}, \Lambda_2 + \pi_{02}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right) \\
&\leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi, \mathcal{D}_{\Sigma,\mathrm{IC}} + \frac{\pi}{2}, \mathcal{D}_{\Sigma,\mathrm{BC}} + \frac{\pi}{2}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right) \\
&\leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \frac{\pi}{2}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right).
\end{aligned}$$

Thus, a tight converse for the full-duplex setting is obtained in the mixed interference regime.

Strong Interference

Let us apply the bound (3.113) to the half-duplex setting (3.4) which corresponds to $\pi_{01} + \pi_{02} \leq \pi$. Here we have,

$$\begin{aligned}
\mathcal{D}_{\Sigma,\mathrm{ICLC}} &\leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi_{01} + \pi_{02}, \alpha_{12} + \pi_{02}, \alpha_{21} + \pi_{01}, \frac{\mathcal{D}_{3e} + \pi_{01} + \pi_{02}}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right) \\
&\leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi, \frac{\alpha_{12} + \alpha_{21} + \pi_{01} + \pi_{02}}{2}, \frac{\mathcal{D}_{3e} + \pi}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right) \\
&\leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi, \frac{\alpha_{12} + \alpha_{21} + \pi}{2}, \frac{\mathcal{D}_{3e} + \pi}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right) \\
&\leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi, \frac{D_{2e} + \pi}{2}, \frac{\mathcal{D}_{3e} + \pi}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right),
\end{aligned}$$

which is the tight converse bound for the half-duplex setting in the strong interference regime.

Next, let us apply the bound (3.113) to the full-duplex setting (3.5) which corresponds to $\pi_{01} \leq \frac{\pi}{2}, \pi_{02} \leq \frac{\pi}{2}$. Here we have,

$$\mathcal{D}'_{\Sigma,\mathrm{ICLC}} \leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}} + \pi_{01} + \pi_{02}, \alpha_{12} + \pi_{02}, \alpha_{21} + \pi_{01}, \frac{\mathcal{D}_{3e} + \pi_{01} + \pi_{02}}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right)$$

$$\leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}}+\pi, \alpha_{12}+\frac{\pi}{2}, \alpha_{21}+\frac{\pi}{2}, \frac{\mathcal{D}_{3e}+\pi}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right)$$
$$\leq \min\left(\mathcal{D}_{\Sigma,\mathrm{IC}}+\pi, \min(\alpha_{12},\alpha_{21})+\frac{\pi}{2}, \frac{\mathcal{D}_{3e}+\pi}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right)$$

which is the tight converse bound for the full-duplex setting in the strong interference regime. This completes the proof of converse for both Theorem 3.1 and Theorem 3.2.

3.5 Achievability for Weak and Mixed Interference

In this section, we specify the achievable schemes for the weak and mixed interference regimes, for both the half-duplex setting and the full-duplex setting. Without loss of generality, we will assume throughout this section that

$$\alpha_{11} \geq \alpha_{22}. \tag{3.114}$$

3.5.1 Weak Interference Regime: $\max(\alpha_{12},\alpha_{21}) \leq \min(\alpha_{11},\alpha_{22})$

We will assume $\pi \leq \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}}$. There is no loss of generality in this assumption because the achievability for $\pi > \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}}$ is the same as that for $\pi = \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}}$, i.e., the upperbound of $\mathcal{D}_{\Sigma,\mathrm{BC}}$ is achieved without need for further cooperation. The achievable schemes for both half-duplex and full-duplex settings are shown in the Table 3.2. This is because in the weak interference regime the cooperative messages W_{01}, W_{02} are combined into one common message $W_0^c = (W_{01}, W_{02})$, which carries d_0^c DoF and can be decoded by both users. Therefore, without loss of generality we can assume $d_{01} = d_{02} = d_0^c/2$. Since the total cooperation capability is shared equally in the two directions, there is no distinction between the half-duplex and full-duplex settings in the weak interference regime.

Table 3.2: *The achievability for weak interference regime under both half-duplex and full-duplex settings, where* $M \triangleq \alpha_{11}+\alpha_{22}, N \triangleq \alpha_{12}+\alpha_{21}$, *and* $\pi \le \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}}$. *The received powers of different codewords at each receiver are specified in decreasing order, which also corresponds to the successive decoding order at that receiver.*

Sub-cases		Codewords' GDoF and Power	Received Power User 1	Received Power User 2
$\alpha_{11},\alpha_{22} \ge N$, $\pi \le \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}}$ $= \min(\alpha_{12},\alpha_{21})$		$X_{11}: \begin{cases} d_{11} = \alpha_{11}-\alpha_{21} \\ E\lvert X_{11}\rvert^2 = P^{-\alpha_{21}} \end{cases}$	$X_0^c :\sim P^{\alpha_{11}}$	$X_0^c :\sim P^{\alpha_{22}}$
		$X_{22}: \begin{cases} d_{22} = \alpha_{22}-\alpha_{12} \\ E\lvert X_{22}\rvert^2 = P^{-\alpha_{12}} \end{cases}$	$X_{11} :\sim P^{\alpha_{11}-\alpha_{21}}$	$X_{22} :\sim P^{\alpha_{22}-\alpha_{12}}$
		$X_0^c: \begin{cases} d_0^c = \pi \\ E\lvert X_0^c\rvert^2 = \mathtt{Diag}(1-P^{-\alpha_{21}}, 1-P^{-\alpha_{12}}) \end{cases}$	$X_{22} :\sim P^0$	$X_{11} :\sim P^0$
$\alpha_{11} \ge N$, $\alpha_{22} \le N$, $\pi \le \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}}$ $= \alpha_{22} - \max(\alpha_{12},\alpha_{21})$	$\alpha_{12} \ge \alpha_{21}$	$X_{11}^p: \begin{cases} d_{11}^p = \alpha_{11}-\alpha_{21} \\ E\lvert X_{11}^p\rvert^2 = P^{-\alpha_{21}} \end{cases}$	$X_{11}^c :\sim P^{\alpha_{11}}$	$X_0^c :\sim P^{\alpha_{22}}$
		$X_{11}^c: \begin{cases} d_{11}^c = \alpha_{12}+\alpha_{21}-\alpha_{22} \\ E\lvert X_{11}^c\rvert^2 = 1-P^{-d_{11}^c} \end{cases}$	$X_0^c :\sim P^{-d_{11}^c+\alpha_{11}}$	$X_{11}^c :\sim P^{\alpha_{21}}$
		$X_{22}: \begin{cases} d_{22} = \alpha_{22}-\alpha_{12} \\ E\lvert X_{22}\rvert^2 = P^{-\alpha_{12}} \end{cases}$	$X_{11}^p :\sim P^{\alpha_{11}-\alpha_{21}}$	$X_{22} :\sim P^{\alpha_{22}-\alpha_{12}}$
		$X_0^c: \begin{cases} d_0^c = \pi \\ E\lvert X_0^c\rvert^2 = \mathtt{Diag}(P^{-d_{11}^c}-P^{-\alpha_{21}}, 1-P^{-\alpha_{12}}) \end{cases}$	$X_{22} :\sim P^0$	$X_{11}^p :\sim P^0$
	$\alpha_{12} \le \alpha_{21}$	$X_{11}^p: \begin{cases} d_{11}^p = \alpha_{11}-\alpha_{21} \\ E\lvert X_{11}^p\rvert^2 = P^{-\alpha_{21}} \end{cases}$	$X_{11}^c :\sim P^{\alpha_{11}}$	$X_0^c :\sim P^{\alpha_{22}}$
		$X_{11}^c: \begin{cases} d_{11}^c = 2\alpha_{21}-\alpha_{22} \\ E\lvert X_{11}^c\rvert^2 = 1-P^{-d_{11}^c} \end{cases}$	$X_0^c :\sim P^{-d_{11}^c+\alpha_{11}}$	$X_{11}^c :\sim P^{\alpha_{21}}$
		$X_{22}: \begin{cases} d_{22} = \alpha_{22}-\alpha_{21} \\ E\lvert X_{22}\rvert^2 = P^{-\alpha_{21}} \end{cases}$	$X_{11}^p :\sim P^{\alpha_{11}-\alpha_{21}}$	$X_{22} :\sim P^{\alpha_{22}-\alpha_{21}}$
		$X_0^c: \begin{cases} d_0^c = \pi \\ E\lvert X_0^c\rvert^2 = \mathtt{Diag}(P^{-d_{11}^c}-P^{-\alpha_{21}}, 1-P^{-\alpha_{12}}) \end{cases}$	$X_{22} :\sim P^0$	$X_{11}^p :\sim P^0$
$\alpha_{11},\alpha_{22} \le N$, $N+\max(\alpha_{12},\alpha_{21}) \le M$, $\pi \le \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}}$ $= M-N-\max(\alpha_{12},\alpha_{21})$		$X_{11}^p: \begin{cases} d_{11}^p = \alpha_{11}-\alpha_{21} \\ E\lvert X_{11}^p\rvert^2 = P^{-\alpha_{21}} \end{cases}$	$X_{11}^c :\sim P^{\alpha_{11}}$	$X_{22}^c :\sim P^{\alpha_{22}}$
		$X_{11}^c: \begin{cases} d_{11}^c = \alpha_{12}+\alpha_{21}-\alpha_{22} \\ E\lvert X_{11}^c\rvert^2 = 1-P^{-d_{11}^c} \end{cases}$	$X_0^c :\sim P^{-d_{11}^c+\alpha_{11}}$	$X_0^c :\sim P^{-d_{22}^c+\alpha_{22}}$
		$X_{22}^p: \begin{cases} d_{22}^p = \alpha_{22}-\alpha_{12} \\ E\lvert X_{22}^p\rvert^2 = P^{-\alpha_{12}} \end{cases}$	$X_{22}^c :\sim P^{\alpha_{12}}$	$X_{11}^c :\sim P^{\alpha_{21}}$
		$X_{22}^c: \begin{cases} d_{22}^c = \alpha_{12}+\alpha_{21}-\alpha_{11} \\ E\lvert X_{22}^c\rvert^2 = 1-P^{-d_{22}^c} \end{cases}$	$X_{11}^p :\sim P^{\alpha_{11}-\alpha_{21}}$	$X_{22}^p :\sim P^{\alpha_{22}-\alpha_{12}}$
		$X_0^c: \begin{cases} d_0^c = \pi \\ E\lvert X_0^c\rvert^2 = \mathtt{Diag}(P^{-d_{11}^c}-P^{-\alpha_{21}}, P^{-d_{22}^c}-P^{-\alpha_{12}}) \end{cases}$	$X_{22}^p :\sim P^0$	$X_{11}^p :\sim P^0$

As shown in Table 3.2 the achievable schemes are partitioned into three sub-cases. To complement Table 3.2, let us explicitly note the sum-GDoF of the interference channel [24] and the broadcast channel [18] for each -case as follows.

- $\alpha_{11},\alpha_{22} \ge N$

  $\mathcal{D}_{\Sigma,\mathrm{IC}} = M-N$, $\mathcal{D}_{\Sigma,\mathrm{BC}} = M - \max(\alpha_{12},\alpha_{21})$.

- $\alpha_{11} \ge N, \alpha_{22} \le N$

$\mathcal{D}_{\Sigma,\mathrm{IC}} = \alpha_{11}$, $\mathcal{D}_{\Sigma,\mathrm{BC}} = M - \max(\alpha_{12}, \alpha_{21})$.

- $\alpha_{11}, \alpha_{22} \leq N$

  $\mathcal{D}_{\Sigma,\mathrm{IC}} = \min(N, M - \max(\alpha_{12}, \alpha_{21}))$, $\mathcal{D}_{\Sigma,\mathrm{BC}} = M - \max(\alpha_{12}, \alpha_{21})$. Note that if $N + \max(\alpha_{12}, \alpha_{21}) > M$ then there is no cooperation gain as $\mathcal{D}_{\Sigma,\mathrm{IC}} = \mathcal{D}_{\Sigma,\mathrm{BC}} = M - \max(\alpha_{12}, \alpha_{21})$. This is why we have the constraint $N + \max(\alpha_{12}, \alpha_{21}) \leq M$ in the last row of the table.

In order to illustrate how the entries in the table describe the achievable scheme for each case, let us explain the last row of the table. The achievability for all other cases follows from the description in Table 3.2 in a similar fashion.

In the sub-case corresponding to the last row of Table 3.2, the noncooperative messages W_{11}, W_{22} are both split into private and common components, $W_{11} = (W_{11}^c, W_{11}^p)$, $W_{22} = (W_{22}^c, W_{22}^p)$. The sub-messages W_{11}^c, W_{11}^p, W_{22}^c, W_{22}^p carry $\alpha_{12} + \alpha_{21} - \alpha_{22}$, $\alpha_{11} - \alpha_{21}$, $\alpha_{12} + \alpha_{21} - \alpha_{11}$, and $\alpha_{22} - \alpha_{12}$ GDoF respectively. $W_{11}^p, W_{11}^c, W_{22}^p, W_{22}^c$ are encoded into independent Gaussian codebooks producing codewords $X_{11}^p, X_{11}^c, X_{22}^p, X_{22}^c$ with power levels $P^{-\alpha_{21}}$, $1 - P^{d_{11}^c}$, $P^{-\alpha_{12}}$, and $1 - P^{d_{22}^c}$ respectively. The cooperative common message W_0^c carries π GDoF and is encoded into the vector Gaussian codeword $X_0^c = (X_{01}^c, X_{02}^c)$ with covariance matrix $\mathtt{Diag}(P^{-d_{11}^c} - P^{-\alpha_{21}}, P^{-d_{22}^c} - P^{-\alpha_{12}})$. The transmitted symbols are $X_1 = X_{11}^c + X_{01}^c + X_{11}^p, X_2 = X_{22}^c + X_{02}^c + X_{22}^p$. Next let us describe the decoding. User 1 (resp. User 2) decodes $X_{11}^c, X_0^c, X_{22}^c, X_{11}^p$ (resp. $X_{22}^c, X_0^c, X_{11}^c, X_{22}^p$) successively. Specifically, for User 1, the received power of X_{11}^c is $\sim P^{\alpha_{11}}$ while the interference power is $\sim P^{-d_{11}^c + \alpha_{11}}$, so that the SINR is $\sim P^{d_{11}^c}$. Therefore X_{11}^c for message W_{11}^c can be successfully decoded. Then User 1 reconstructs and subtracts the contribution of X_{11}^c and starts to decode X_0^c. The desired power for X_0^c is $\sim P^{-d_{11}^c + \alpha_{11}}$ while the interference power is $P^{\alpha_{12}}$, so that SINR is $\sim P^{-d_{11}^c + \alpha_{11} - \alpha_{12}} = P^{\alpha_{11} + \alpha_{22} - 2\alpha_{12} - \alpha_{21}} = P^{M - N - \alpha_{12}}$. Since $d_0^c = \pi \leq M - N - \max(\alpha_{12}, \alpha_{21}) \leq M - N - \alpha_{12}$, it follows that message W_0^c can be successfully decoded. Proceeding similarly, by using successive interference cancellation, User 1 can decode $X_{22}^c, X_{11}^p, X_{22}^p$ in that order

Table 3.3: *The achievability for mixed interference regime under half-duplex setting and full-duplex setting. It is assumed that $\pi \leq \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}}$, because any further cooperation is redundant. The table also applies to the full-duplex setting, provided that π is replaced by $\frac{\pi}{2}$. This is because one of W_{01}, W_{02} is wasted.*

Sub-cases	Codewords' GDoF and Power	Received Power	
		User 1	User 2
$\alpha_{21} \leq \alpha_{22} \leq \alpha_{11} \leq \alpha_{12}$, $\pi \leq \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}} = N - M$	$X_{11} : d_{11} = \alpha_{11} - \alpha_{21}, E\lvert X_{11}\rvert^2 = P^{-\alpha_{21}}$ $X_{22} : d_{22} = \alpha_{22}, E\lvert X_{22}\rvert^2 = 1 - P^{-\alpha_{22}}$ $X_{01} : d_{01} = \pi, E\lvert X_{01}\rvert^2 = P^{-\alpha_{22}}$	$X_{22} :\sim P^{\alpha_{12}}$ $X_{01} :\sim P^{\alpha_{12}-\alpha_{22}}$ $X_{11} :\sim P^{\alpha_{11}-\alpha_{21}}$	$X_{22} :\sim P^{\alpha_{22}}$ $X_{01} :\sim P^{0}$ $X_{11} :\sim P^{0}$
$\alpha_{12} \leq \alpha_{22} \leq \alpha_{11} \leq \alpha_{21}$, $\pi \leq \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}} = N - M$	$X_{11} : d_{11} = \alpha_{11}, E\lvert X_{11}\rvert^2 = 1 - P^{-\alpha_{11}}$ $X_{22} : d_{22} = \alpha_{22} - \alpha_{12}, E\lvert X_{22}\rvert^2 = P^{-\alpha_{12}}$ $X_{02} : d_{02} = \pi, E\lvert X_{02}\rvert^2 = P^{-\alpha_{11}}$	$X_{11} :\sim P^{\alpha_{11}}$ $X_{02} :\sim P^{0}$ $X_{22} :\sim P^{0}$	$X_{11} :\sim P^{\alpha_{21}}$ $X_{02} :\sim P^{\alpha_{21}-\alpha_{11}}$ $X_{22} :\sim P^{\alpha_{22}-\alpha_{12}}$
$\alpha_{22} \leq \alpha_{12} \leq \alpha_{11} \leq \alpha_{21}$, $\pi \leq \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}} = \alpha_{21} - \alpha_{11}$	$X_{11} : d_{11} = \alpha_{11}, E\lvert X_{11}\rvert^2 = 1 - P^{-\alpha_{11}}$ $X_{02} : d_{02} = \pi, E\lvert X_{02}\rvert^2 = P^{-\alpha_{11}}$	$X_{11} :\sim P^{\alpha_{11}}$ $X_{02} :\sim P^{0}$	$X_{11} :\sim P^{\alpha_{21}}$ $X_{02} :\sim P^{\alpha_{21}-\alpha_{11}}$
$\alpha_{22} \leq \alpha_{21} \leq \alpha_{11} \leq \alpha_{12}$, $\pi \leq \mathcal{D}_{\Sigma,\mathrm{BC}} - \mathcal{D}_{\Sigma,\mathrm{IC}} = \alpha_{12} - \alpha_{11}$	$X_{11} : d_{11} = \alpha_{11}, E\lvert X_{11}\rvert^2 = 1$ $X_{01} : d_{01} = \pi, E\lvert X_{01}\rvert^2 = 1$	$X_{01} :\sim P^{\alpha_{12}}$ $X_{11} :\sim P^{\alpha_{11}}$	$X_{11} :\sim P^{\alpha_{21}}$ $X_{01} :\sim P^{\alpha_{22}}$

for messages $W_{22}^c, W_{11}^p, W_{22}^p$. Note that the decoding order corresponds to the decreasing order of power levels at the receiver, which is also the order in which the codewords are listed in the last two columns of Table 3.2.

3.5.2 Mixed Interference Regime:

$$\min(\alpha_{12}, \alpha_{21}) \leq \max(\alpha_{11}, \alpha_{22}), \max(\alpha_{12}, \alpha_{21}) \geq \min(\alpha_{11}, \alpha_{22})$$

In the mixed interference regime, as explained in Section 3.4.3, there are only four cases where a cooperation gain exists. Therefore, the description of achievable schemes in Table 3.3 shows only these four sub-cases.

The achievability for the full-duplex setting follows by replacing π with $\pi/2$. This is because only one-sided cooperation is needed in the mixed interference regime, i.e., either W_{01} or

W_{02} is not used, thereby wasting one-half of the cooperation capability.

To illustrate how the table describes the achievable scheme, let us consider the first row. In this regime, User 1 is strictly stronger than User 2. Messages W_{11}, W_{22}, W_{01} carry $\alpha_{11} - \alpha_{21}, \alpha_{22}, \pi$ GDoF, and they are encoded into independent Gaussian codebooks X_{11}, X_{22}, X_{01} with powers $P^{-\alpha_{21}}, 1 - P^{-\alpha_{22}}$, and $P^{-\alpha_{22}}$, respectively. The transmitted signals are $X_1 = X_{11}, X_2 = X_{22} + X_{01}$. User 1 decodes X_{22} for W_{22} first, while treating everything else as noise. For this decoding stage, the desired signal power is $\sim P^{\alpha_{12}}$ while the interference power is $\sim P^{\alpha_{12}-\alpha_{22}}$ so that SINR is $\sim P^{\alpha_{22}}$, which gives us the GDoF value $d_{22} = \alpha_{22}$. After successfully decoding W_{22}, Receiver 1 is able to reconstruct the codeword X_{22} and subtract its contribution from the received signal. Then it decodes the codeword X_{01} for its message W_{01} while treating the remaining signal as noise. The desired power is $\sim P^{\alpha_{12}-\alpha_{22}}$ while interference power is $\sim P^{\alpha_{11}-\alpha_{21}}$, so that the SINR for this decoding is $\sim P^{N-M}$. Since $\pi \leq N - M$, W_{01} can be successfully decoded. After reconstructing and subtracting the contribution of codeword X_{01}, User 1 decodes X_{11} for its desired message W_{11}, while treating the remaining signal as noise. The desired signal power is $\sim P^{\alpha_{11}-\alpha_{21}}$ while interference power is $\sim P^0$. Since $d_{11} = \alpha_{11} - \alpha_{21}$, message W_{11} can be successfully decoded. Receiver 2 is able to decode X_{22} by treating everything else as noise.

3.6 Achievability for Strong interference: $\min(\alpha_{12}, \alpha_{21}) \geq \max(\alpha_{11}, \alpha_{22})$

In this section, we describe the achievable schemes for the strong interference regime, which are separated into half-duplex and full-duplex settings. The broadcast channel bound for the strong interference, which is found in [18], is $\mathcal{D}_{\Sigma,\mathrm{BC}} = \alpha_{12} + \alpha_{21} - \max(\alpha_{11}, \alpha_{22})$. In this section, we no longer assume that $\alpha_{11} \geq \alpha_{22}$. Instead, in the strong interference regime, it is

more convenient to assume $\alpha_{12} \geq \alpha_{21}$ without loss of generality.

3.6.1 Half-duplex Setting

Let us begin with an illustrative example where $\alpha_{11} = \alpha_{22} = 2, \alpha_{12} = 5, \alpha_{21} = 3$. For this setting, $\mathcal{D}_{\Sigma,\mathrm{BC}} = 6$ according to [18] and $D_{\Sigma,\mathrm{IC}} = 3$ according to [24]. Let us consider how much cooperation is needed in this case to achieve $\mathcal{D}_{\Sigma,\mathrm{BC}}$. The achievable scheme of [18] summarized in Figure 3.7, requires $\pi = 6$ GDoF of cooperation, i.e., all messages must be shared between the two transmitters. This is because in order to take advantage of the strong interference links, the private messages of Users 1 and 2, are sent from opposing transmitters, i.e., Transmitters 2 and 1, respectively. These are messages W_{01}^p, W_{02}^p in Figure 3.7. The common message W_o^c that is decoded by both users is sent from both transmitters, so it is shared as well. However, as shown in Theorem 3.1 in this chapter, the sum-GDoF of

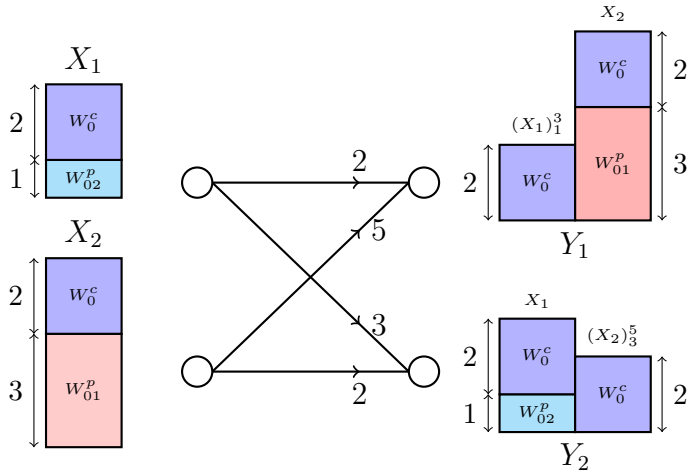

Figure 3.7: *The scheme from [18] requires $\pi = 6$ GDoF of cooperation to achieve the broadcast channel bound.*

limited cooperation interference channel for this example is $D_{\Sigma,\mathrm{ICLC}} = \min(3+\pi, \frac{8+\pi}{2}, \frac{13+\pi}{3}, 6)$. Therefore, $\pi^*_{half} = 5$ is the minimum value of cooperative GDoF needed to achieve the BC bound. The optimally efficient scheme is shown in Figure 3.8. The improvement in efficiency comes from the observation that part of the common message (in this case, W_{22}) can be

transmitted from only one transmitter (in this case, Transmitter 2), and therefore requires no cooperation.

The achievable scheme is described as follows: The cooperative messages W_{01}, W_{02} are split into a cooperative common[5] message $W_0^c = (W_{01}^c, W_{02}^c)$ and the cooperative private messages W_{01}^p, W_{02}^p. Messages $W_{22}, W_0^c, W_{01}^p, W_{02}^p$ carry $1, 1, 3, 1$ GDoF respectively such that $\pi = 5$. $W_{22}, W_{01}^p, W_{02}^p$ are encoded into independent Gaussian codebooks $X_{22}, X_{01}^p, X_{02}^p$ respectively with powers $\mathrm{E}|X_{22}|^2 = 1 - P^{-1}$, $\mathrm{E}|X_{01}^p|^2 = P^{-2}$, $\mathrm{E}|X_{02}^p|^2 = P^{-2}$. Message W_0^c carries 1 GDoF and is encoded to a vector Gaussian codebook $X_0^c = (X_{01}^c, X_{02}^c)$ with power covariance matrix $\mathtt{Diag}(1 - P^{-2}, P^{-1} - P^{-2})$. The transmitted symbols are $X_1 = X_{01}^c + X_{02}^p, X_2 = X_{22} + X_{02}^c + X_{01}^p$. Suppressing the time index for clarity, the received signals are:

$$Y_1 = \sqrt{P^2} G_{11}(X_{01}^c + X_{02}^p) + \sqrt{P^5} G_{12}(X_{22} + X_{02}^c + X_{01}^p) + Z_1$$
$$Y_2 = \sqrt{P^3} G_{21}(X_{01}^c + X_{02}^p) + \sqrt{P^2} G_{22}(X_{22} + X_{02}^c + X_{01}^p) + Z_2$$

When decoding, User 1 first decodes X_{22} for W_{22} while treating everything else as Gaussian noise. Since X_{22} is received at power level $\sim P^5$ while all other signals are received with power levels $\sim P^4$ or lower, the SINR for decoding W_{22} is $\sim P^1$, which gives us the GDoF value $d_{22} = 1$. After decoding W_{22}, Receiver 1 is able to reconstruct codeword X_{22} and subtract its contribution from the received signal. After this, Receiver 1 decodes the codeword X_0^c for message W_0^c, while treating the remaining signals as Gaussian noise. Since the desired signal for this decoding is received with power level $\sim P^4$ while all other signals are received with power levels $\sim P^3$ or less, the SINR for decoding W_0^c is $\sim P^1$ which gives GDoF value $d_0^c = 1$. Then Receiver 1 subtracts the contribution of X_0^c and decodes message W_{01}^p while treating all other remaining signals as Gaussian noise. As evident from Figure 3.8, the SINR for this decoding is $\sim P^3$ which gives us GDoF value $d_{01}^p = 3$. Receiver 2 proceeds similarly by successively decoding W_0^c, W_{22}, W_{02}^p.

[5]The cooperative common message may be arbitrarily divided among the two users, e.g., without loss of generality, we can assume that half of W_0^c is the desired message for User 1 and the other half of W_0^c is the desired message for User 2.

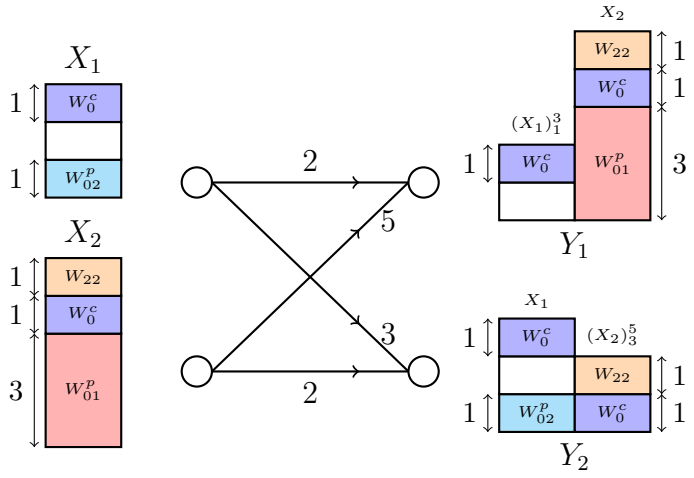

Figure 3.8: *The optimally efficient achievable scheme achieves the broadcast channel bound with only $\pi = 5$ GDoF of cooperation.*

In general, to prove the achievability for the strong interference regime completely, there are 4 sub-cases, which cover all possibilities. Note that we assume $\pi \leq \pi^*_{half}$ because the achievable scheme for $\pi > \pi^*_{half}$ is the same as $\pi = \pi^*_{half}$, since π^*_{half} already achieves the broadcast channel bound.

Case 1: $\alpha_{12} \leq \alpha_{11} + \alpha_{22}, \alpha_{21} \leq \alpha_{11} + \alpha_{22}, \alpha_{12} + \alpha_{21} \leq \alpha_{11} + \alpha_{22} + \max(\alpha_{11}, \alpha_{22})$

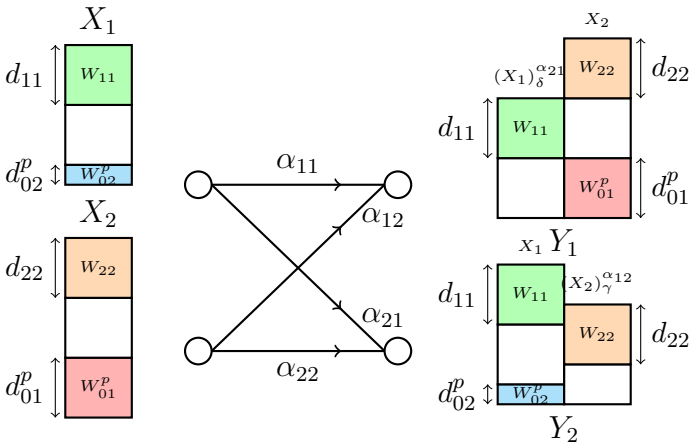

Figure 3.9: *Signal partition in the regime $\alpha_{12} \leq \alpha_{11} + \alpha_{22}, \alpha_{21} \leq \alpha_{11} + \alpha_{22}, \alpha_{12} + \alpha_{21} \leq \alpha_{11} + \alpha_{22} + \max(\alpha_{11}, \alpha_{22}), \delta = \alpha_{21} - \alpha_{11}, \gamma = \alpha_{12} - \alpha_{22}$.*

The sum-GDoF value in this case is characterized as:

$$\mathcal{D}_{\Sigma,\text{ICLC}} = \min\Big(\alpha_{21}+\pi, \frac{\alpha_{12}+\alpha_{21}+\pi}{2}, \mathcal{D}_{\Sigma,\text{BC}}\Big). \tag{3.115}$$

- When $\pi \leq \alpha_{12}-\alpha_{21}$, the first bound is tight, which is achieved by having W_{11}, W_{22}, W_{01}^p carry $\alpha_{21}-\alpha_{22}, \alpha_{22}, \pi$ GDoF respectively. They are encoded into independent Gaussian codebooks producing codewords X_{11}, X_{22}, X_{01}^p with powers $\mathrm{E}|X_{11}|^2 = 1, \mathrm{E}|X_{22}|^2 = 1 - P^{-\alpha_{21}}, \mathrm{E}|X_{01}^p|^2 = P^{-\alpha_{21}}$. The transmitted signals are $X_1 = X_{11}, X_2 = X_{22} + X_{01}^p$. When decoding, User 1 first *jointly* (acting as the receiver in a multiple access channel) decodes X_{11} and X_{22} while treating everything else as noise, while the noise floor due to X_{01}^p is $\sim P^{\alpha_{12}-\alpha_{21}}$. The GDoF region for this multiple access channel is the following.

$$\Big\{(d_{11}, d_{22}) : d_{11} \leq \alpha_{11}+\alpha_{21}-\alpha_{12}, d_{11}+d_{22} \leq \alpha_{21}\Big\}. \tag{3.116}$$

  Since $d_{11} = \alpha_{21}-\alpha_{22} \leq \alpha_{11}+\alpha_{21}-\alpha_{12}, d_{11}+d_{22} = \alpha_{21}$ belongs to the GDoF region of the multiple access channel, User 1 is able to decode X_{11}, X_{22} for messages W_{11}, W_{22}. After this, User 1 subtracts the reconstructed codewords X_{11}, X_{22} and then decodes X_{01}^p. The SINR for this decoding is $\sim P^{\alpha_{12}-\alpha_{21}}$, such that $d_{01}^p = \pi \leq \alpha_{12}-\alpha_{21}$ and the decoding is successful. User 2 decodes X_{11}, X_{22} successively. The SINR values for X_{11}, X_{22} are $\sim P^{\alpha_{21}-\alpha_{22}}, \sim P^{\alpha_{22}}$ respectively, which give us $d_{11} = \alpha_{21}-\alpha_{22}, d_{22} = \alpha_{22}$. Therefore X_{11}, X_{22} are successfully decoded at User 2.

- When $\alpha_{12}-\alpha_{21} \leq \pi \leq \pi_{half}^*$, where $\pi_{half}^* = \alpha_{12}+\alpha_{21}-2\max(\alpha_{11}, \alpha_{22})$ according to Corollary 1, the second bound is tight and is achieved as follows: $W_{11}, W_{22}, W_{01}^p, W_{02}^p$ carry $d_{11} = (\alpha_{21}+2\alpha_{11}-\alpha_{12}-\pi)/2,\ d_{22} = \alpha_{12}-\alpha_{11},\ d_{01}^p = (\alpha_{12}-\alpha_{21}+\pi)/2,\ d_{02}^p = (\alpha_{21}-\alpha_{12}+\pi)/2$ GDoF respectively. They are encoded into independent Gaussian codewords $X_{11}, X_{22}, X_{01}^p, X_{02}^p$ with powers $\mathrm{E}|X_{11}|^2 = 1-P^{-d_{11}-d_{22}}, \mathrm{E}|X_{22}|^2 = 1-P^{-d_{11}-d_{22}}, \mathrm{E}|X_{01}^p|^2 = P^{-d_{11}-d_{22}}, \mathrm{E}|X_{02}^p|^2 = P^{-d_{11}-d_{22}}$. The transmitted symbols are $X_1 = X_{11}+X_{02}^p, X_2 = X_{22}+X_{01}^p$. When decoding, User 1 decodes X_{22}, X_{11}, X_{01}^p successively, The SINRs for these

codewords are $\sim P^{\alpha_{12}-\alpha_{11}}, \sim P^{\alpha_{11}-\alpha_{12}+d_{11}+d_{22}} = P^{d_{11}}, \sim P^{\alpha_{12}-d_{11}-d_{22}} = P^{d_{01}^p}$ respectively. User 2 acts as a multiple access receiver, it jointly decodes X_{11} and X_{22} while treating everything else as noise, where the noise floor due to X_{02}^p is $\alpha_{21} - d_{11} - d_{22} = \frac{\alpha_{21}-\alpha_{12}+\pi}{2}$. Hence the GDoF region for this multiple access channel is the following.

$$\left\{(d_{11}, d_{22}): \quad d_{22} \le \alpha_{22} - \frac{\alpha_{21}-\alpha_{12}+\pi}{2}, \right. \\ \left. d_{11}+d_{22} \le \frac{\alpha_{12}+\alpha_{21}-\pi}{2}\right\}. \tag{3.117}$$

Since $d_{22} = \alpha_{12} - \alpha_{11} \le \alpha_{22} + \max(\alpha_{11}, \alpha_{22}) - \alpha_{21} = \alpha_{22} - \frac{\alpha_{21}-\alpha_{12}+\pi_{half}^*}{2} \le \alpha_{22} - \frac{\alpha_{21}-\alpha_{12}+\pi}{2}, d_{11}+d_{22} = \frac{\alpha_{12}+\alpha_{21}-\pi}{2}$ belongs to the GDoF region of the multiple access channel, the messages W_{11}, W_{22} can be jointly decoded successfully by User 2. After this, User 2 subtracts the contribution of X_{11}, X_{22} and decodes X_{02}^p, whose SINR is $\sim P^{\alpha_{21}-d_{11}-d_{22}} = P^{d_{02}^p}$, such that X_{02}^p for W_{02}^p can be successfully decoded. The signal partitioning is shown in Figure 3.9. The cooperation capability beyond π_{half}^* is redundant because with π_{half}^* cooperation the broadcast GDoF are already achieved.

Case 2: $\alpha_{12} \le \alpha_{11}+\alpha_{22}, \alpha_{21} \le \alpha_{11}+\alpha_{22}, \alpha_{12}+\alpha_{21} \ge \alpha_{11}+\alpha_{22}+\max(\alpha_{11}, \alpha_{22})$

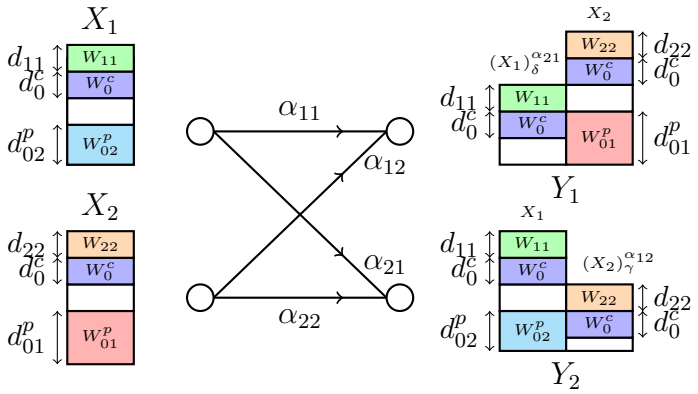

Figure 3.10: *Signal partition in the regime* $\alpha_{12}, \alpha_{21} \le \alpha_{11}+\alpha_{22}, \alpha_{12}+\alpha_{21} \ge \alpha_{11}+\alpha_{22}+\max(\alpha_{11}, \alpha_{22})$*, where* $\delta = \alpha_{21}-\alpha_{11}, \gamma = \alpha_{12}-\alpha_{22}$.

In this regime, the sum-GDoF value, as characterized in (3.12), is:

$$\mathcal{D}_{\Sigma,\mathrm{ICLC}} = \min\Big(\alpha_{21}+\pi, \frac{\alpha_{12}+\alpha_{21}+\pi}{2}, \frac{\alpha_{11}+\alpha_{12}+\alpha_{21}+\alpha_{22}+\pi}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\Big). \tag{3.118}$$

- When $\pi \leq \alpha_{12}-\alpha_{21}$, the first bound is active. The achievable scheme is the same as the achievable scheme in Case 1 which achieves the first bound for the corresponding π value.

- When $\alpha_{21}-\alpha_{12} \leq \pi \leq 2\alpha_{11}+2\alpha_{22}-\alpha_{12}-\alpha_{21}$, the second bound is active and also achieved with the same scheme as in Case 1 for corresponding π value.

- When $2\alpha_{11}+2\alpha_{22}-\alpha_{12}-\alpha_{21} \leq \pi \leq \pi^*_{half}$, where according to Corollary 1 we have $\pi^*_{half} = 2\alpha_{12}+2\alpha_{21}-\alpha_{11}-\alpha_{22}-3\max(\alpha_{11},\alpha_{22})$, the third bound is tight. It is achieved by the following: $W_{11}, W_{22}, W^p_{01}, W^p_{02}$ carry $(2\alpha_{21}-\alpha_{12}+2\alpha_{11}-\alpha_{22}-\pi)/3$, $(2\alpha_{12}-\alpha_{21}+2\alpha_{22}-\alpha_{11}-\pi)/3$, $(\alpha_{11}+\alpha_{22}+\alpha_{12}-2\alpha_{21}+\pi)/3$, $(\alpha_{11}+\alpha_{22}+\alpha_{21}-2\alpha_{12}+\pi)/3$ GDoF respectively. They are encoded into independent Gaussian codebooks $X_{11}, X_{22}, X^p_{01}, X^p_{02}$ with powers $\mathrm{E}|X_{11}|^2 = 1-P^{-d_{11}}$, $\mathrm{E}|X_{22}|^2 = 1-P^{-d_{22}}$, $\mathrm{E}|X^p_{01}|^2 = P^{-d_{11}-d_{22}-d^c_0} = P^{(\alpha_{11}+\alpha_{22}-2\alpha_{12}-2\alpha_{21}+\pi)/3}$, $\mathrm{E}|X^p_{02}|^2 = P^{-d_{11}-d_{22}-d^c_0} = P^{(\alpha_{11}+\alpha_{22}-2\alpha_{12}-2\alpha_{21}+\pi)/3}$. W^c_0 carries $(\alpha_{12}+\alpha_{21}-2\alpha_{11}-2\alpha_{22}+\pi)/3$ GDoF and it is encoded to a vector Gaussian codebook $X^c_0 = (X^c_{01}, X^c_{02})$ with power covariance matrix $\mathtt{Diag}(P^{-d_{11}}-P^{(\alpha_{11}+\alpha_{22}-2\alpha_{12}-2\alpha_{21}+\pi)/3}, P^{-d_{22}}-P^{(\alpha_{11}+\alpha_{22}-2\alpha_{12}-2\alpha_{21}+\pi)/3})$. The transmitted symbols are $X_1 = X_{11}+X^c_{01}+X^p_{02}$, $X_2 = X_{22}+X^c_{02}+X^p_{01}$. When decoding, User 1 decodes $X_{22}, X^c_0, X_{11}, X^p_{01}$ for messages $W_{22}, W^c_0, W_{11}, W^p_{01}$ successively, whose SINR values are $\sim P^{d_{22}}$, $\sim P^{-d_{22}+\alpha_{12}-\alpha_{11}} = P^{d^c_0}$, $\sim P^{\alpha_{11}-(\alpha_{11}+\alpha_{22}+\alpha_{12}-2\alpha_{21}+\pi)/3} = P^{d_{11}}$, $\sim P^{\alpha_{12}+(\alpha_{11}+\alpha_{22}-2\alpha_{12}-2\alpha_{21}+\pi)/3} = P^{d^p_{01}}$ respectively. User 2 proceeds similarly by successively decoding $W_{11}, W^c_0, W_{22}, W^p_{02}$. See Figure 3.10 for an illustration.

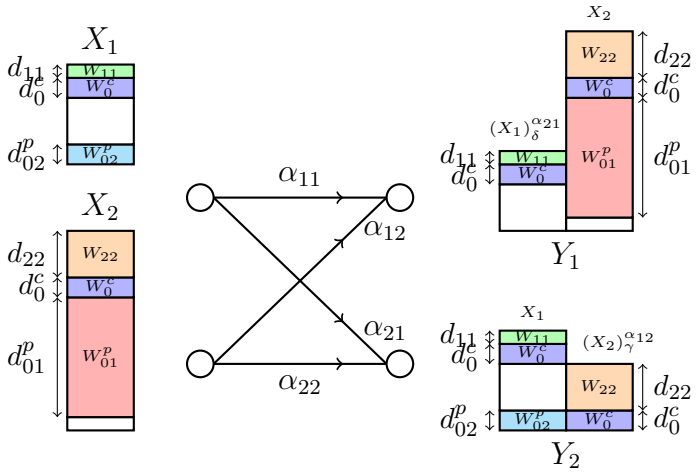

Figure 3.11: *Signal partition depiction for* $\alpha_{12} \geq \alpha_{11} + \alpha_{22}, \alpha_{21} \leq \alpha_{11} + \alpha_{22}$*, where where* $\delta = \alpha_{21} - \alpha_{11}, \gamma = \alpha_{12} - \alpha_{22}$*.*

Case 3: $\alpha_{12} \geq \alpha_{11} + \alpha_{22}, \alpha_{21} \leq \alpha_{11} + \alpha_{22}$

In this regime, the sum-GDoF value is

$$\mathcal{D}_{\Sigma,\text{ICLC}} = \min\left(\alpha_{21} + \pi, \frac{2\alpha_{12} + \alpha_{21} + \pi}{3}, \mathcal{D}_{\Sigma,\text{BC}}\right). \tag{3.119}$$

- When $\pi \leq \alpha_{12} - \alpha_{21}$, the first bound is tight. The achievable scheme is as follows: W_{11}, W_{22}, W_{01}^p carry $\alpha_{21} - \alpha_{22}, \alpha_{22}, \pi$ GDoF respectively and they are encoded into independent Gaussian codebooks producing codewords X_{11}, X_{22}, X_{01}^p with power $\mathrm{E}|X_{11}|^2 = 1, \mathrm{E}|X_{22}|^2 = 1 - P^{-\alpha_{22}}, \mathrm{E}|X_{01}^p|^2 = P^{-\alpha_{22}}$. When decoding, User 1 decodes X_{22} first with SINR value $\sim P^{\alpha_{22}}$. Then, it subtracts the reconstructed codeword X_{22} and acts as a multiple access receiver to jointly decode X_{11} and X_{01}^p. The GDoF region for this multiple access channel is the following.

$$\left\{(d_{11}, d_{01}^p) : d_{11} \leq \alpha_{11}, d_{11} + d_{01}^p \leq \alpha_{12} - \alpha_{22}\right\}. \tag{3.120}$$

Since $d_{11} = \alpha_{21} - \alpha_{22} \leq \alpha_{11}, d_{11} + d_{01}^p = \alpha_{21} - \alpha_{22} + \pi \leq \alpha_{12} - \alpha_{22}$ belongs to the GDoF region, X_{11}, X_{01}^p can be successfully decoded at User 1. For User 2, it successively decodes

X_{11}, X_{22}, whose SINR values are $\sim P^{\alpha_{21}-\alpha_{22}} = P^{d_{11}}, \sim P^{\alpha_{22}} = P^{d_{22}}$ respectively. Therefore W_{11}, W_{22} are decoded successfully decoded at User 2.

- When $\alpha_{12} - \alpha_{21} \leq \pi \leq \pi^*_{half}$, where according to Corollary 1 we have $\pi^*_{half} = \alpha_{12} + 2\alpha_{21} - 3\max(\alpha_{11}, \alpha_{22})$, the second bound is tight. The achievable scheme is as follows: Messages $W_{11}, W_{22}, W^p_{01}, W^p_{02}$ carry $(2\alpha_{21} + \alpha_{12} - 3\alpha_{22} - \pi)/3, (3\alpha_{22} + \alpha_{12} - \alpha_{21} - \pi)/3, (2\alpha_{12} - 2\alpha_{21} + \pi)/3, (\alpha_{21} - \alpha_{12} + \pi)/3$ GDoF respectively. They are encoded into independent Gaussian codebooks producing codewords $X_{11}, X_{22}, X^p_{01}, X^p_{02}$ with powers $\mathrm{E}|X_{11}|^2 = 1 - P^{-d_{11}}, \mathrm{E}|X_{22}|^2 = 1 - P^{-d_{22}}, \mathrm{E}|X^p_{01}|^2 = P^{-\alpha_{22}}$, $\mathrm{E}|X^p_{02}|^2 = P^{-d_{11}-d_{22}-d^c_0} = P^{-(2\alpha_{21}+\alpha_{12}-\pi)/3}$. W^c_0 carries $(\alpha_{21} - \alpha_{12} + \pi)/3$ GDoF and is encoded into a vector Gaussian codebook $X^c_0 = (X^c_{01}, X^c_{02})$ with power covariance matrix $\mathtt{Diag}(P^{-d_{11}} - P^{-(2\alpha_{21}+\alpha_{12}-\pi)/3}, P^{-d_{22}} - P^{-\alpha_{22}})$. The transmitted symbols are $X_1 = X_{11} + X^c_{01} + X^p_{02}, X_2 = X_{22} + X^c_{02} + X^p_{01}$. When decoding, User 1 decodes W_{22}, W^c_0 successively while treating everything else as noise. Their SINR values are $\sim P^{d_{22}}, \sim P^{\alpha_{22}-d_{22}} = P^{d^c_0}$. After this, User 1 subtracts the reconstructed codewords X_{22}, X^c_0. Then it acts as a multiple access receiver to jointly decode W_{11} and W^p_{01} while treating the remaining signal as noise. The GDoF region for this multiple access channel is the following.

$$\left\{(d_{11}, d^p_{01}) : d_{11} \leq \alpha_{11}, d_{11} + d^p_{01} \leq \alpha_{12} - \alpha_{22}\right\}. \tag{3.121}$$

Since $d_{11} = (2\alpha_{21} + \alpha_{12} - 3\alpha_{22} - \pi)/3 \leq \alpha_{21} - \alpha_{22} \leq \alpha_{11}, d_{11} + d^p_{01} \leq \alpha_{12} - \alpha_{22}$ belongs to this GDoF region, it follows that W_{22}, W^c_0 can be successfully decoded. User 2 decodes $X_{11}, X^c_0, X_{22}, X^p_{02}$ successively, whose SINR values are $\sim P^{d_{11}}, \sim P^{\alpha_{21}-d_{11}-\alpha_{22}} = P^{d^c_0}, \sim P^{\alpha_{22}-\alpha 21+(2\alpha_{21}+\alpha_{12}-\pi)/3} = P^{d_{22}}, \sim P^{\alpha_{21}-(2\alpha_{21}+\alpha_{12}-\pi)/3} = P^{d^p_{02}}$ respectively. The signal partition is shown in Figure 3.11. Note that cooperation capability beyond π^*_{half} is redundant because with π^*_{half} cooperation the broadcast GDoF are already achieved.

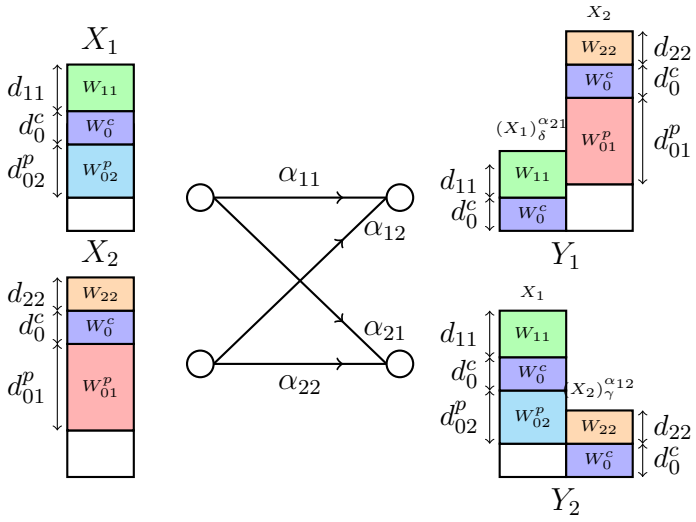

Figure 3.12: *Signal partition in the regime* $\alpha_{12}, \alpha_{21} \geq \alpha_{11}+\alpha_{22}$*, where* $\delta = \alpha_{21}-\alpha_{11}, \gamma = \alpha_{12}-\alpha_{22}$.

Case 4: $\alpha_{12} \geq \alpha_{11} + \alpha_{22}, \alpha_{21} \geq \alpha_{11} + \alpha_{22}$

In this regime, we have

$$\begin{aligned}&\mathcal{D}_{\Sigma,\mathrm{ICLC}}\\&= \min\left(\alpha_{11}+\alpha_{22}+\pi, \frac{2\alpha_{12}+2\alpha_{21}-\alpha_{11}-\alpha_{22}+\pi}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right).\end{aligned} \tag{3.122}$$

- When $\pi \leq \alpha_{21} - \alpha_{11} - \alpha_{22}$, the first bound is active, which is achieved by letting W_{11}, W_{22}, W_{02}^p carry $\alpha_{11}, \alpha_{22}, \pi$ GDoF respectively. They are encoded into independent Gaussian codebooks X_{11}, X_{22}, X_{02}^p with power $\mathrm{E}|X_{11}|^2 = 1-P^{-\alpha_{11}}, \mathrm{E}|X_{22}|^2 = 1, \mathrm{E}|X_{02}^p|^2 = P^{-\alpha_{11}}$. The transmitted symbols are $X_1 = X_{11} + X_{02}^p, X_2 = X_{22}$. When decoding, User 1 decodes X_{22}, X_{11} successively, whose SINR values are $\sim P^{\alpha_{12}-\alpha_{11}}, \sim P^{\alpha_{11}}$ respectively. Since $d_{22} = \alpha_{22} \leq \alpha_{12} - \alpha_{11}, d_{11} = \alpha_{11}$, messages W_{22}, W_{11} can be decoded successfully. User 2 decodes X_{11}, X_{02}^p, X_{22} successively, whose SINR values are $\sim P^{\alpha_{11}}, \sim P^{\alpha_{21}-\alpha_{11}-\alpha_{22}}, \sim P^{\alpha_{22}}$ respectively. Since $d_{11} = \alpha_{11}, d_{02}^p = \pi \leq \alpha_{21} - \alpha_{11} - \alpha_{22}, d_{22} = \alpha_{22}$, messages W_{11}, W_{02}^p, W_{22} can be decoded successfully.

- When $\alpha_{21} - \alpha_{11} - \alpha_{22} \leq \pi \leq \alpha_{12} + \alpha_{21} - 2\alpha_{11} - 2\alpha_{22}$, the first bound is still active

and is achieved by letting $W_{11}, W_{22}, W_{01}^p, W_{02}^p$ carry $\alpha_{11}, \alpha_{22}, \pi + \alpha_{11} + \alpha_{22} - \alpha_{21}, \alpha_{21} - \alpha_{11} - \alpha_{22}$ GDoF respectively. They are encoded into independent Gaussian codebooks $X_{11}, X_{22}, X_{01}^p, X_{02}^p$ with power $\mathrm{E}|X_{11}|^2 = 1 - P^{-\alpha_{11}}, \mathrm{E}|X_{22}|^2 = 1 - P^{-\alpha_{22}}, \mathrm{E}|X_{01}^p|^2 = P^{-\alpha_{22}}, \mathrm{E}|X_{02}^p|^2 = P^{-\alpha_{11}}$. The transmitted symbols are $X_1 = X_{11} + X_{02}^p, X_2 = X_{22} + X_{01}^p$. When decoding, User 1 decodes X_{22}, X_{01}^p, X_{11} successively, whose SINR values are $\sim P^{\alpha_{22}}, \sim P^{\alpha_{12}-\alpha_{22}-\alpha_{11}}, \sim P^{\alpha_{11}}$ respectively. Since $d_{22} = \alpha_{22}, d_{01}^p = \pi + \alpha_{11} + \alpha_{22} - \alpha_{21} \leq \alpha_{12} - \alpha_{22} - \alpha_{11}, d_{11} = \alpha_{11}$, X_{22}, X_{01}^p, X_{11} can be decoded successfully. User 2 proceeds similarly by decoding X_{11}, X_{02}^p, X_{22} successively.

- When $\alpha_{12} + \alpha_{21} - 2\alpha_{11} - 2\alpha_{22} \leq \pi \leq \pi_{half}^*$, where according to Corollary 1 we have $\pi_{half}^* = \alpha_{12} + \alpha_{21} + \alpha_{11} + \alpha_{22} - 3\max(\alpha_{11}, \alpha_{22})$, the second bound is tight. It is achieved as follows: $W_{11}, W_{22}, W_{01}^p, W_{02}^p$ carry $(\alpha_{12} + \alpha_{21} - 2\alpha_{22} + \alpha_{11} - \pi)/3, (\alpha_{12} + \alpha_{21} - 2\alpha_{11} + \alpha_{22} - \pi)/3, (2\alpha_{12} + \pi - \alpha_{11} - \alpha_{22} - \alpha_{21})/3, (2\alpha_{21} + \pi - \alpha_{11} - \alpha_{22} - \alpha_{12})/3$ GDoF respectively and are encoded into independent Gaussian codebooks $X_{11}, X_{22}, X_{01}^p, X_{02}^p$ with powers $\mathrm{E}|X_{11}|^2 = 1 - P^{-d_{11}}, \mathrm{E}|X_{22}|^2 = 1 - P^{-d_{22}}, \mathrm{E}|X_{01}^p|^2 = P^{-\alpha_{22}}, \mathrm{E}|X_{02}^p|^2 = P^{-\alpha_{11}}$. W_0^c carries $(2\alpha_{11} + 2\alpha_{22} + \pi - \alpha_{12} - \alpha_{21})/3$ GDoF and is encoded into a vector Gaussian codebook $X_0^c = (X_{01}^c, X_{02}^c)$ with power covariance matrix $\mathtt{Diag}(P^{-d_{11}} - P^{-\alpha_{11}}, P^{-d_{22}} - P^{-\alpha_{22}})$. The transmitted symbols are $X_1 = X_{11} + X_{01}^c + X_{02}^p, X_2 = X_{22} + X_{02}^c + X_{01}^p$. User 1 decodes W_{22}, W_0^c successively while treating everything else as noise, whose SINR values are $\sim P^{d_{22}}, \sim P^{\alpha_{22}-d_{22}} = P^{d_0^c}$ respectively. After this, User 1 subtracts the contribution of codewords X_{22}, X_0^c and then acts as a multiple access receiver by jointly decoding W_{11} and W_{01}^p while treating the remaining signals as noise. The GDoF region for this multiple access channel is the following.

$$\left\{(d_{11}, d_{01}^p) : d_{11} \leq \alpha_{11}, d_{11} + d_{01}^p \leq \alpha_{12} - \alpha_{22}\right\}. \tag{3.123}$$

Since $d_{11} = (\alpha_{12} + \alpha_{21} - 2\alpha_{22} + \alpha_{11} - \pi)/3 \leq (\alpha_{12} + \alpha_{21} - 2\alpha_{22} + \alpha_{11} - (\alpha_{12} + \alpha_{21} - 2\alpha_{11} - 2\alpha_{22}))/3 \leq \alpha_{11}, d_{11} + d_{01}^p = \alpha_{12} - \alpha_{22}$ belongs to the GDoF region, W_{11}, W_{01}^p can be

decoded successfully. User 2 proceeds similarly. See Figure 3.12 for an illustration.

3.6.2 Full-duplex Setting

In this section we consider the achievability for the full-duplex setting. Before presenting the complete proof, let us use our example ($\alpha_{11} = \alpha_{22} = 2, \alpha_{12} = 5, \alpha_{21} = 3$) to convey the main insights. Here we have $\mathcal{D}'_{\Sigma,\text{ICLC}} = \min(3+\pi, 3+\frac{\pi}{2}, \frac{13+\pi}{3}, 6)$. The bounds $\mathcal{D}_{\Sigma,\text{IC}} + \pi = 3+\pi$ and $\frac{\mathcal{D}_{3e}+\pi}{3} = \frac{13+\pi}{3}$ are redundant. To achieve the broadcast channel bound ($\mathcal{D}_{\Sigma,\text{BC}} = 6$), the GDoF in the conference link is $\pi^*_{full} = 6$, which means our proposed scheme is no more efficient than [18]. This is because in our scheme, $d^c_0 = 1, d^p_{01} = 3, d^p_{02} = 1$, which requires $\frac{\pi}{2} \geq d^p_{01} = 3$. Hence 1 DoF in the W_{02} conference link is wasted because $d_{02} \leq d^p_{02} + d^c_0 = 2$. We can see that under full-duplex setting, one cooperation link is fully wasted in the mixed interference regime, but in the strong interference regime, one cooperation link is partially wasted.

In the full-duplex setting, first of all, the achievable schemes even for one cooperative bit to buy one over-the-air bit or half over-the-air bit become non-trivial as one of the cooperation links is partially wasted for some π values. Hence we will discuss it in a bit more detail. On the other hand, the achievable scheme for achieving the 1/3 bound (when the bound is active) in the full-duplex setting is the same as the corresponding scheme for half-duplex setting. In general, we also consider the 4 cases. Similarly, $\pi \leq \pi^*_{full}$ is assumed.

Case 1: $\alpha_{12} \leq \alpha_{11} + \alpha_{22}, \alpha_{21} \leq \alpha_{11} + \alpha_{22}, \alpha_{12} + \alpha_{21} \leq \alpha_{11} + \alpha_{22} + \max(\alpha_{11}, \alpha_{22})$

In this regime, the sum-GDoF is

$$\mathcal{D}'_{\Sigma,\text{ICLC}} = \min\left(\alpha_{21} + \frac{\pi}{2}, \mathcal{D}_{\Sigma,\text{BC}}\right). \tag{3.124}$$

- When $\frac{\pi}{2} \leq \alpha_{12} - \alpha_{21}$, the first bound is active. The achievability is the same as Case 1 in the half-duplex setting to achieve the first bound, except $d_{01}^p = \pi/2$ here.

- When $\alpha_{12} - \alpha_{21} \leq \frac{\pi}{2} \leq \frac{\pi_{full}^*}{2}$, where according to Corollary 2 we have $\pi_{full}^* = 2\alpha_{12} - 2\max(\alpha_{11}, \alpha_{22})$, the first bound is still active and is achieved by letting $W_{11}, W_{22}, W_{01}^p, W_{02}^p$ carry $\alpha_{11} - \frac{\pi}{2}, \alpha_{12} - \alpha_{11}, \frac{\pi}{2}, \frac{\pi}{2} + \alpha_{21} - \alpha_{12}$ GDoF respectively. They are encoded into independent Gaussian codebooks producing codewords $X_{11}, X_{22}, X_{01}^p, X_{02}^p$ with powers $\mathrm{E}|X_{11}|^2 = 1 - P^{-d_{11}-d_{22}}, \mathrm{E}|X_{22}|^2 = 1 - P^{-d_{11}-d_{22}}, \mathrm{E}|X_{01}^p|^2 = P^{-d_{11}-d_{22}}, \mathrm{E}|X_{02}^p|^2 = P^{-d_{11}-d_{22}}$. The transmitted signals are $X_1 = X_{11} + X_{02}^p, X_2 = X_{22} + X_{01}^p$. When decoding, Receiver 1 uses successive interference cancellation to decode X_{22}, X_{11}, X_{01}^p successively, whose SINR values are $\sim P^{\alpha_{12}-\alpha_{11}} = P^{d_{22}}, \sim P^{\alpha_{11}-\alpha_{12}+d_{11}+d_{22}} = P^{d_{11}}, \sim P^{\alpha_{12}-d_{11}-d_{22}} = P^{\frac{\pi}{2}} = P^{d_{01}^p}$. Therefore, W_{22}, W_{11}, W_{01}^p can be successfully decoded. User 2 acts as a multiple access receiver, it jointly decodes X_{11} and X_{22}, while the noise floor due to X_{02}^p is $P^{\alpha_{21}-\alpha_{12}+\frac{\pi}{2}}$. The GDoF region for this multiple access channel is the following.

$$\left\{(d_{11}, d_{22}): \quad d_{22} \leq \alpha_{22} - (\alpha_{21} - \alpha_{12} + \frac{\pi}{2}), \right. \\ \left. d_{11} + d_{22} \leq \alpha_{12} - \frac{\pi}{2}\right\}. \tag{3.125}$$

Since $d_{22} = \alpha_{12} - \alpha_{11} \leq \alpha_{22} - \alpha_{21} + \max(\alpha_{11}, \alpha_{22}) = \alpha_{22} - (\alpha_{21} - \alpha_{12} + \frac{\pi_{full}^*}{2}) \leq \alpha_{22} - (\alpha_{21} - \alpha_{12} + \frac{\pi}{2}), d_{11} + d_{22} = \alpha_{12} - \frac{\pi}{2}$ belongs to the GDoF region, W_{22}, W_{11} can be decoded successfully. Then User 2 subtracts the contribution of X_{11}, X_{22} and decodes X_{02}^p, whose SINR is $\sim P^{\alpha_{21}-d_{11}-d_{22}} = P^{d_{02}^p}$, so W_{02}^p is decoded successfully.

Case 2: $\alpha_{12} \leq \alpha_{11}+\alpha_{22}, \alpha_{21} \leq \alpha_{11}+\alpha_{22}, \alpha_{12}+\alpha_{21} \geq \alpha_{11}+\alpha_{22}+\max(\alpha_{11},\alpha_{22})$

- $2\alpha_{21} \leq \alpha_{11}+\alpha_{22}+\max(\alpha_{11},\alpha_{22})$

  In this regime, the sum-GDoF value is

$$\mathcal{D}'_{\Sigma,\text{ICLC}} = \min\left(\alpha_{21}+\frac{\pi}{2}, \mathcal{D}_{\Sigma,\text{BC}}\right). \tag{3.126}$$

  - When $\frac{\pi}{2} \leq \alpha_{11}+\alpha_{22}-\alpha_{21}$, the first bound is active. The achievability is the same as in Case 1 above to achieve the first bound for the corresponding π value.
  - When $\alpha_{11}+\alpha_{22}-\alpha_{21} \leq \frac{\pi}{2} \leq \frac{\pi^*_{full}}{2}$, where according to Corollary 2 we have $\pi^*_{full} = 2\alpha_{12}-2\max(\alpha_{11},\alpha_{22})$, the first bound is active and is achieved by letting $W_{11}, W_{22}, W^p_{01}, W^p_{02}, W^c_0$ carry $\alpha_{11}-\frac{\pi}{2}, \alpha_{12}-\alpha_{21}+\alpha_{22}-\frac{\pi}{2}, \frac{\pi}{2}, \alpha_{21}-\alpha_{12}+\frac{\pi}{2}, \alpha_{21}-\alpha_{11}-\alpha_{22}+\frac{\pi}{2}$ GDoF respectively. Messages $W_{11}, W_{22}, W^p_{01}, W^p_{02}$ are encoded into independent Gaussian codebooks producing codewords $X_{11}, X_{22}, X^p_{01}, X^p_{02}$ with powers $\mathrm{E}|X_{11}|^2 = 1-P^{-d_{11}}, \mathrm{E}|X_{22}|^2 = 1-P^{-d_{22}}, \mathrm{E}|X^p_{01}|^2 = P^{\alpha_{11}-\alpha_{12}-d_{11}}, \mathrm{E}|X^p_{02}|^2 = P^{\alpha_{22}-\alpha_{21}-d_{22}}$, respectively. Message W^c_0 is encoded into a vector Gaussian codebook producing codeword $X^c_0 = (X^c_{01}, X^c_{02})$ with the covariance matrix $\texttt{Diag}(P^{-d_{11}}-P^{\alpha_{22}-\alpha_{21}-d_{22}}, P^{-d_{22}}-P^{\alpha_{11}-\alpha_{12}-d_{11}})$. The transmitted symbols are $X_1 = X_{11}+X^c_{01}+X^p_{02}, X_1 = X_{22}+X^c_{02}+X^p_{01}$. When decoding, User 1 decodes $X_{22}, X^c_0, X_{11}, X^p_{01}$ successively, with SINR values $\sim P^{d_{22}}, \sim P^{\alpha_{12}-\alpha_{11}-d_{22}} = P^{d^c_0}, \sim P^{d_{11}}, \sim P^{\alpha_{11}-d_{11}} = P^{\frac{\pi}{2}} = P^{d^p_{01}}$ respectively. Therefore $W_{22}, W^c_0, W_{11}, W^p_{01}$ can be successfully decoded at User 1. User 2 proceeds similarly by decoding $W_{11}, W^c_0, W_{22}, W^p_{02}$ successively. It can be checked that $d_{01} = d^p_{01} = \frac{\pi}{2}, d_{02} = d^p_{02}+d^c_0 = 2\alpha_{21}-\alpha_{12}-\alpha_{11}-\alpha_{22}+\pi \leq 2\alpha_{21}-\alpha_{12}-\alpha_{11}-\alpha_{22}+\frac{\pi^*_{full}}{2}+\frac{\pi}{2} \leq 2\alpha_{21}-\alpha_{11}-\alpha_{22}-\max(\alpha_{11},\alpha_{22})+\frac{\pi}{2} \leq \frac{\pi}{2}$.

- $2\alpha_{21} \geq \alpha_{11}+\alpha_{22}+\max(\alpha_{11},\alpha_{22})$

  The sum-GDoF value in this regime is

$$\mathcal{D}'_{\Sigma,\text{ICLC}} = \min\left(\alpha_{21}+\frac{\pi}{2}, \frac{\alpha_{11}+\alpha_{12}+\alpha_{21}+\alpha_{21}+\pi}{3}, \mathcal{D}_{\Sigma,\text{BC}}\right). \tag{3.127}$$

– When $\frac{\pi}{2} \leq \alpha_{11} + \alpha_{22} + \alpha_{12} - 2\alpha_{21}$, the first bound is active. The achievable scheme is identical to $2\alpha_{21} \leq \alpha_{11} + \alpha_{22} + \max(\alpha_{11}, \alpha_{22})$ for the same π value.

– When $\alpha_{11} + \alpha_{22} - 2\alpha_{12} + \alpha_{21} \leq \frac{\pi}{2} \leq \frac{\pi^*_{full}}{2}$, where according to Corollary 2 we have $\pi^*_{full} = 2\alpha_{12} + 2\alpha_{21} - \alpha_{11} - \alpha_{22} - 3\max(\alpha_{11}, \alpha_{22})$, the second bound is active, whose achievability is the same as the achievable scheme for the corresponding bound in Case 2 in the half-duplex setting.

Case 3: $\alpha_{12} \geq \alpha_{11} + \alpha_{22}, \alpha_{21} \leq \alpha_{11} + \alpha_{22}$

- $\alpha_{12} \geq 2\alpha_{21} - \max(\alpha_{11}, \alpha_{22})$

  In this regime, the sum-GDoF value is

$$\mathcal{D}'_{\Sigma,\mathrm{ICLC}} = \min\left(\alpha_{21} + \frac{\pi}{2}, \mathcal{D}_{\Sigma,\mathrm{BC}}\right). \tag{3.128}$$

  – When $\frac{\pi}{2} \leq \alpha_{12} - \alpha_{21}$, the first bound is tight, and its achievability is identical to first bound in Case 3 under the half-duplex setting except $d^p_{01} = \frac{\pi}{2}$ here.

  – When $\alpha_{12} - \alpha_{21} \leq \frac{\pi}{2} \leq \frac{\pi^*_{full}}{2}$, where according to Corollary 2 we have $\pi^*_{full} = 2\alpha_{12} - 2\max(\alpha_{11}, \alpha_{22})$, the first bound is still active and is achieved as follows: Messages W_{11}, W_{22}, W^p_{01}, W^p_{02} carry $\alpha_{12} - \alpha_{22} - \frac{\pi}{2}, \alpha_{22} + \alpha_{12} - \alpha_{21} - \frac{\pi}{2}, \frac{\pi}{2}, \alpha_{21} - \alpha_{12} + \frac{\pi}{2}$ GDoF respectively. They are encoded into independent Gaussian codewords $X_{11}, X_{22}, X^p_{01}, X^p_{02}$ with powers $\mathrm{E}|X_{11}|^2 = 1 - P^{-d_{11}}, \mathrm{E}|X_{22}|^2 = 1 - P^{-d_{22}}, \mathrm{E}|X^p_{01}|^2 = P^{-\alpha_{22}}, \mathrm{E}|X^p_{02}|^2 = P^{-d_{11}-d_{22}-d^c_0} = P^{-\alpha_{12}+\frac{\pi}{2}}$. W^c_0 carries $\alpha_{21} - \alpha_{12} + \frac{\pi}{2}$ GDoF and is encoded into a vector Gaussian codebook $X^c_0 = (X^c_{01}, X^c_{02})$ with covariance matrix $\mathrm{E}|X^c_0|^2 = \mathtt{Diag}(P^{-d_{11}} - P^{-\alpha_{12}+\frac{\pi}{2}}, P^{-d_{22}} - P^{-\alpha_{22}})$. The transmitted symbols are $X_1 = X_{11} + X^c_{01} + X^p_{02}, X_2 = X_{22} + X^c_{02} + X^p_{01}$. When decoding, User 1 decodes X_{22}, X^c_0 successively while treating everything else as noise. The SINR values for X_{22}, X^c_0 are $\sim P^{d_{22}}, \sim P^{\alpha_{22}-d_{22}} = P^{d^c_0}$ respectively. After subtracting the contribution of X_{22}, X^c_0, it jointly decodes X_{11} and X^p_{01}. The GDoF region for this multiple access channel is

the following.

$$\left\{(d_{11}, d_{01}^p) : d_{11} \leq \alpha_{11}, d_{22} + d_{01}^p \leq \alpha_{12} - \alpha_{22}\right\}. \tag{3.129}$$

Since $d_{11} = \alpha_{12} - \alpha_{22} - \frac{\pi}{2} \leq \alpha_{12} - \alpha_{22} - \frac{\pi^*_{full}}{2} = \max(\alpha_{11}, \alpha_{22}) - \alpha_{22} \leq \alpha_{11}, d_{11} + d_{01}^p = \alpha_{12} - \alpha_{22}$ belongs to the GDoF region of the multiple access channel, W_{11}, W_{01}^p can be decoded successfully. User 2 successively decodes $X_{11}, X_0^c, X_{22}, X_{02}^p$, whose SINR values are $\sim P^{d_{11}}, \sim P^{\alpha_{21}-d_{11}-\alpha_{22}} = P^{\alpha_{21}-\alpha_{12}+\frac{\pi}{2}} = P^{d_0^c}, \sim P^{\alpha_{22}-\alpha_{21}+\alpha_{12}-\frac{\pi}{2}} = P^{d_{22}}, \sim P^{\alpha_{21}-\alpha_{12}+\frac{\pi}{2}} = P^{d_{02}^p}$ respectively. Note that $d_{01} = d_{01}^p = \frac{\pi}{2}, d_{02} = d_{02}^p + d_0^c = 2\alpha_{21} - 2\alpha_{12} + \pi \leq 2\alpha_{21} - 2\alpha_{12} + \frac{\pi^*_{full}}{2} + \frac{\pi}{2} = 2\alpha_{21} - \alpha_{12} - \max(\alpha_{11}, \alpha_{22}) + \frac{\pi}{2} \leq \frac{\pi}{2}$, so the link for W_{02} is partially wasted.

- $\alpha_{12} \leq 2\alpha_{21} - \max(\alpha_{11}, \alpha_{22})$

  In this regime, the sum-GDoF value is

$$\mathcal{D}'_{\Sigma,\text{ICLC}} = \min\left(\alpha_{21} + \frac{\pi}{2}, \frac{2\alpha_{12} + \alpha_{21} + \pi}{3}, \mathcal{D}_{\Sigma,\text{BC}}\right). \tag{3.130}$$

  - When $\frac{\pi}{2} \leq 2\alpha_{12} - 2\alpha_{21}$, the first bound is active. The achievable schemes are identical to those for $\alpha_{12} \geq 2\alpha_{21} - \max(\alpha_{11}, \alpha_{22})$ for the same π value.
  - When $2\alpha_{12} - 2\alpha_{21} \leq \frac{\pi}{2} \leq \frac{\pi^*_{full}}{2}$, where according to Corollary 2 we have $\pi^*_{full} = \alpha_{12} + 2\alpha_{21} - 3\max(\alpha_{11}, \alpha_{22})$, the second bound is active, and its achievability is the same as the corresponding 1/3 factor bound scheme in Case 3 of the half-duplex setting.

Case 4: $\alpha_{12} \geq \alpha_{11} + \alpha_{22}, \alpha_{21} \geq \alpha_{11} + \alpha_{22}$

- $\alpha_{12} \geq \alpha_{21} + \min(\alpha_{11}, \alpha_{22})$

  The achievable scheme for this sub-case is shown in Table 3.4. In this regime the sum-

Table 3.4: *The achievable scheme for Case 4 under the condition* $\alpha_{12} \geq \alpha_{21} + \min(\alpha_{11}, \alpha_{22})$*, where* $M = \alpha_{11} + \alpha_{22}$.

Sub-cases		Codewords' GDoF and Power	Received Power User 1	Received Power User 2
$\alpha_{12} \geq M$, $\alpha_{21} \geq M$, $\alpha_{12} \geq \alpha_{21} + \min(\alpha_{11}, \alpha_{22})$ ($\alpha_{12} \geq \alpha_{21}$ is assumed)	$\frac{\pi}{2} \leq \alpha_{21} - M$, $\mathcal{D}'_{\Sigma,\text{ICLC}} = M + \pi$	$X_{11}: \begin{cases} d_{11} = \alpha_{11} \\ E\lvert X_{11}\rvert^2 = 1 - P^{-\alpha_{11}} \end{cases}$	$X_{22} :\sim P^{\alpha_{12}}$	$X_{11} :\sim P^{\alpha_{21}}$
		$X_{22}: \begin{cases} d_{22} = \alpha_{22} \\ E\lvert X_{22}\rvert^2 = 1 - P^{-\alpha_{22}} \end{cases}$	$X^p_{01} :\sim P^{\alpha_{12}-\alpha_{22}}$	$X^p_{02} :\sim P^{\alpha_{21}-\alpha_{11}}$
		$X^p_{01}: \begin{cases} d^p_{01} = \frac{\pi}{2} \\ E\lvert X^p_{01}\rvert^2 = P^{-\alpha_{22}} \end{cases}$	$X_{11} :\sim P^{\alpha_{11}}$	$X_{22} :\sim P^{\alpha_{22}}$
		$X^p_{02}: \begin{cases} d^p_{02} = \frac{\pi}{2} \\ E\lvert X^p_{02}\rvert^2 = P^{-\alpha_{11}} \end{cases}$	$X^p_{02} :\sim P^0$	$X^p_{01} :\sim P^0$
	$\alpha_{21} - M \leq \frac{\pi}{2}$ $\leq \alpha_{12} - M$, $\mathcal{D}'_{\Sigma,\text{ICLC}} = \alpha_{21} + \frac{\pi}{2}$	$X_{11}: \begin{cases} d_{11} = \alpha_{11} \\ E\lvert X_{11}\rvert^2 = 1 - P^{-\alpha_{11}} \end{cases}$	$X_{22} :\sim P^{\alpha_{12}}$	$X_{11} :\sim P^{\alpha_{21}}$
		$X_{22}: \begin{cases} d_{22} = \alpha_{22} \\ E\lvert X_{22}\rvert^2 = 1 - P^{-\alpha_{22}} \end{cases}$	$X^p_{01} :\sim P^{\alpha_{12}-\alpha_{22}}$	$X^p_{02} :\sim P^{\alpha_{21}-\alpha_{11}}$
		$X^p_{01}: \begin{cases} d^p_{01} = \frac{\pi}{2} \\ E\lvert X^p_{01}\rvert^2 = P^{-\alpha_{22}} \end{cases}$	$X_{11} :\sim P^{\alpha_{11}}$	$X_{22} :\sim P^{\alpha_{22}}$
		$X^p_{02}: \begin{cases} d^p_{02} = \alpha_{21} - M \\ E\lvert X^p_{02}\rvert^2 = P^{-\alpha_{11}} \end{cases}$	$X^p_{02} :\sim P^0$	$X^p_{01} :\sim P^0$
	$\alpha_{12} - M \leq \frac{\pi}{2}$ $\leq \alpha_{12} - \max(\alpha_{11}, \alpha_{22})$, $\mathcal{D}'_{\Sigma,\text{ICLC}} = \alpha_{21} + \frac{\pi}{2}$	$X_{11}: \begin{cases} d_{11} = \alpha_{12} - \alpha_{22} - \frac{\pi}{2} \\ E\lvert X_{11}\rvert^2 = 1 - P^{-d_{11}} \end{cases}$	$X_{22} :\sim P^{\alpha_{12}}$	$X_{11} :\sim P^{\alpha_{21}}$
		$X_{22}: \begin{cases} d_{22} = \alpha_{12} - \alpha_{11} - \frac{\pi}{2} \\ E\lvert X_{22}\rvert^2 = 1 - P^{-d_{22}} \end{cases}$	$X^c_0 :\sim P^{\alpha_{12}-d_{22}}$	$X^c_0 :\sim P^{\alpha_{21}-d_{11}}$
		$X^p_{01}: \begin{cases} d^p_{01} = \frac{\pi}{2} \\ E\lvert X^p_{01}\rvert^2 = P^{-\alpha_{22}} \end{cases}$	$X^p_{01} :\sim P^{\alpha_{12}-\alpha_{22}}$	$X^p_{02} :\sim P^{\alpha_{21}-\alpha_{11}}$
		$X^p_{02}: \begin{cases} d^p_{02} = \alpha_{21} - \alpha_{12} + \frac{\pi}{2} \\ E\lvert X^p_{02}\rvert^2 = P^{-\alpha_{11}} \end{cases}$	$X_{11} :\sim P^{\alpha_{11}}$	$X_{22} :\sim P^{\alpha_{22}}$
		$X^c_0: \begin{cases} d^c_0 = M - \alpha_{12} + \frac{\pi}{2} \\ E\lvert X^c_0\rvert^2 = \texttt{Diag}(P^{-d_{11}} - P^{-\alpha_{11}}, P^{-d_{22}} - P^{-\alpha_{22}}) \end{cases}$	$X^p_{02} :\sim P^0$	$X^p_{01} :\sim P^0$

GDoF value is

$$\mathcal{D}'_{\Sigma,\text{ICLC}} = \min\left(\alpha_{11} + \alpha_{22} + \pi, \alpha_{21} + \frac{\pi}{2}, \mathcal{D}_{\Sigma,\text{BC}}\right). \tag{3.131}$$

– When $\frac{\pi}{2} \leq \alpha_{21} - \alpha_{11} - \alpha_{22}$, the first bound is active. Messages $W_{11}, W_{22}, W^p_{01}, W^p_{02}$ carry $\alpha_{11}, \alpha_{22}, \frac{\pi}{2}, \frac{\pi}{2}$ GDoF respectively. They are encoded into independent Gaussian codebooks producing codewords $X_{11}, X_{22}, X^p_{01}, X^p_{02}$ with powers $\mathrm{E}|X_{11}|^2 = 1 - P^{-\alpha_{11}}, \mathrm{E}|X_{22}|^2 = 1 - P^{-\alpha_{22}}, \mathrm{E}|X^p_{01}|^2 = P^{-\alpha_{22}}, \mathrm{E}|X^p_{02}|^2 = P^{-\alpha_{11}}$. The transmitted signals are $X_1 = X_{11} + X^p_{02}, X_2 = X_{22} + X^p_{01}$. When decoding, User 1 decodes X_{22}, X^p_{01}, X_{11} successively, whose SINR values are $\sim P^{\alpha_{22}}, \sim P^{\alpha_{12}-\alpha_{11}-\alpha_{22}}, \sim P^{\alpha_{11}}$. Since $d_{22} = \alpha_{22}, d^p_{01} = \frac{\pi}{2} \leq \alpha_{21} - \alpha_{11} - \alpha_{22} \leq \alpha_{12} - \alpha_{11} - \alpha_{22}$, messages W_{22}, W^p_{01}, W_{11} can be decoded successfully. User 2 proceeds similarly by decoding X_{11}, X^p_{02}, X_{22}

successively.

– When $\alpha_{21}-\alpha_{11}-\alpha_{22} \leq \frac{\pi}{2} \leq \alpha_{12}-\alpha_{11}-\alpha_{22}$, the second bound is active. $W_{11}, W_{22}, W_{01}^p, W_{02}^p$ carry $\alpha_{11}, \alpha_{22}, \frac{\pi}{2}, \alpha_{21}-\alpha_{11}-\alpha_{22}$ GDoF respectively. They are encoded into independent Gaussian codebooks producing codewords $X_{11}, X_{22}, X_{01}^p, X_{02}^p$ with powers $\mathrm{E}|X_{11}|^2 = 1-P^{-\alpha_{11}}, \mathrm{E}|X_{22}|^2 = 1-P^{-\alpha_{22}}, \mathrm{E}|X_{01}^p|^2 = P^{-\alpha_{22}}, \mathrm{E}|X_{02}^p|^2 = P^{-\alpha_{11}}$. For decoding, User 1 decodes X_{22}, X_{01}^p, X_{11} successively while User 2 decodes X_{11}, X_{02}^p, X_{22} successively. The distinction between the two regimes $\alpha_{21}-\alpha_{11}-\alpha_{22} \leq \frac{\pi}{2} \leq \alpha_{12}-\alpha_{11}-\alpha_{22}$ and $\frac{\pi}{2} \leq \alpha_{21}-\alpha_{11}-\alpha_{22}$ is that message W_{02}^p carries different GDoF values.

– When $\alpha_{12}-\alpha_{11}-\alpha_{22} \leq \frac{\pi}{2} \leq \frac{\pi_{full}^*}{2}$, where according to Corollary 2 we have $\pi_{full}^* = 2\alpha_{12}-2\max(\alpha_{11},\alpha_{22})$, the second bound is still active. Messages W_{11}, W_{22}, W_{01}^p, W_{02}^p carry $\alpha_{12}-\alpha_{22}-\frac{\pi}{2}, \alpha_{12}-\alpha_{11}-\frac{\pi}{2}, \frac{\pi}{2}, \alpha_{21}-\alpha_{12}+\frac{\pi}{2}$ GDoF respectively. They are encoded into independent Gaussian codebooks producing codewords $X_{11}, X_{22}, X_{01}^p, X_{02}^p$ with powers $\mathrm{E}|X_{11}|^2 = 1-P^{-d_{11}}, \mathrm{E}|X_{22}|^2 = 1-P^{-d_{22}}, \mathrm{E}|X_{01}^p|^2 = P^{-\alpha_{22}}, \mathrm{E}|X_{02}^p|^2 = P^{-\alpha_{11}}$. W_0^c carries $\alpha_{11}+\alpha_{22}-\alpha_{12}+\frac{\pi}{2}$ GDoF and is encoded into a vector Gaussian codebook $X_0^c = (X_{01}^c, X_{02}^c)$ with power covariance $\mathrm{E}|X_0^c|^2 = \mathtt{Diag}(P^{-d_{11}}-P^{-\alpha_{11}}, P^{-d_{22}}-P^{-\alpha_{22}})$. The transmitted symbols are $X_1 = X_{11}+X_{01}^c+X_{02}^p, X_2 = X_{22}+X_{02}^c+X_{01}^p$. User 1 decodes W_{22}, W_0^c successively while treating everything else as noise, the SINR values are $\sim P^{d_{22}}, \sim P^{\alpha_{22}-d_{22}} = P^{\alpha_{11}+\alpha_{22}+\frac{\pi}{2}-\alpha_{12}} = P^{d_0^c}$ respectively. After this User 1 subtracts the contribution of X_{22}, X_0^c, and it acts as a multiple access receiver by jointly decoding W_{11} and W_{01}^p while treating the remaining signals as noise. The GDoF region of this multiple access channel is the following.

$$\left\{(d_{11}, d_{01}^p) : d_{11} \leq \alpha_{11}, d_{11}+d_{01}^p \leq \alpha_{12}-\alpha_{22}\right\}. \tag{3.132}$$

Since $d_{11} = \alpha_{12}-\alpha_{22}-\frac{\pi}{2} \leq \alpha_{11}, d_{11}+d_{01}^p = \alpha_{12}-\alpha_{22}$ belongs to the GDoF region, W_{11}, d_{01}^p can be decoded successfully. User 2 proceeds similarly. It can be checked

that $d_{01} = d_{01}^p = \frac{\pi}{2}, d_{02} = d_{02}^p + d_0^c = \alpha_{21} - 2\alpha_{12} + \alpha_{11} + \alpha_{22} + \pi \leq \alpha_{21} - 2\alpha_{12} + \alpha_{11} + \alpha_{22} + \frac{\pi^*_{full}}{2} + \frac{\pi}{2} = \alpha_{21} - \alpha_{12} + \alpha_{11} + \alpha_{22} - \max(\alpha_{11}, \alpha_{22}) + \frac{\pi}{2} = \alpha_{21} - \alpha_{12} + \min(\alpha_{11}, \alpha_{22}) + \frac{\pi}{2} \leq \frac{\pi}{2}$, so the link for W_{02} is partially wasted.

- $\alpha_{12} \leq \alpha_{21} + \min(\alpha_{11}, \alpha_{22})$

  The sum-GDoF value is

$$\mathcal{D}'_{\Sigma,\mathrm{ICLC}} = \min\Big(\alpha_{11} + \alpha_{22} + \pi, \alpha_{21} + \frac{\pi}{2}, \frac{2\alpha_{12} + 2\alpha_{21} - \alpha_{11} - \alpha_{22} + \pi}{3}, \mathcal{D}_{\Sigma,\mathrm{BC}}\Big). \tag{3.133}$$

  - When $\frac{\pi}{2} \leq 2\alpha_{12} - \alpha_{21} - \alpha_{11} - \alpha_{22}$, the achievability for the first and second bounds are the corresponding scheme as the regime $\alpha_{12} \geq \alpha_{21} + \min(\alpha_{11}, \alpha_{22})$ for the same $\frac{\pi}{2}$ value.
  - When $2\alpha_{12} - \alpha_{21} - \alpha_{11} - \alpha_{22} \leq \frac{\pi}{2} \leq \frac{\pi^*_{full}}{2}$, where according to Corollary 2 we have $\pi^*_{full} = \alpha_{12} + \alpha_{21} + \alpha_{11} + \alpha_{22} - 3\max(\alpha_{11}, \alpha_{22})$ the third bound is active. The achievability is the same as the corresponding scheme in Case 4 of the half-duplex setting.

3.7 Summary

The sum-set inequalities of [20] are utilized to characterize the sum-GDoF of the 2-user interference channel with limited cooperation, both in half-duplex setting and full-duplex setting, which bridges the gap between the interference channel and broadcast channel. The sum-GDoF value is characterized for arbitrary parameter regimes. The result is also extended to the 2-user X channel setting.

Chapter 4

Sum-GDoF of Symmetric Multi-hop Interference Channel under Finite Precision CSIT using Aligned-Images Sum-set Inequalities

Aligned-Images Sum-set Inequalities are used in this chapter to study the GDoF of the symmetric layered multi-hop interference channel under the robust assumption that CSIT is limited to finite precision. First, the sum-GDoF value is characterized for the $2 \times 2 \times 2$ setting that is comprised of 2 sources, 2 relays, and 2 destinations. It is shown that the sum-GDoF does not improve even if perfect CSIT is allowed in the first hop, as long as the CSIT in the second hop is limited to finite precision. The sum GDoF characterization is then generalized to the $2 \times 2 \times \cdots \times 2$ setting that is comprised of L hops. Remarkably, for large L, the sum-GDoF value approaches that of the one-hop broadcast channel that is obtained by full cooperation among the two transmitters of the last hop, with finite precision CSIT. Previous studies of multi-hop interference networks either identified sophisticated GDoF

optimal schemes under perfect CSIT, such as aligned interference neutralization and network diagonalization, that are powerful in theory but too fragile to be practical, or studied robust achievable schemes like classical amplify/decode/compress-and-forward without claims of information-theoretic optimality. In contrast, under finite precision CSIT, we show that the benefits of fragile schemes are lost, while a combination of classical random coding schemes that are simpler and much more robust, namely a rate-splitting between decode-and-forward and amplify-and-forward, is shown to be GDoF optimal. As such, this chapter represents another step towards bridging the gap between theory (optimality) and practice (robustness) with the aid of Aligned-Images Sum-set Inequalities.

4.1 Introduction

There is much interest in multi-hop interference networks due to their essential role in expanding coverage and enabling high data rates over underutilized (e.g., mm-wave/THz) frequency bands that suffer from high path loss and blockages. However, an information theoretic understanding of the robust fundamental limits of such networks remains elusive, even in the approximate or asymptotic (high SNR) sense. Information theoretic studies of multi-hop interference networks, such as those in [11, 61, 47, 38, 59, 34, 64, 33, 58, 46, 39, 7, 8], have focused primarily on the idealized setting where the channel state information at the transmitters (CSIT) is perfect. The search for optimal solutions under idealized assumptions leads to ideas like Interference Neutralization [11, 61, 47, 38, 59], Aligned Interference Neutralization [34, 64, 33] and Network Diagonalization [58] that are powerful in theory (e.g., everyone gets all the cake), but too fragile to be relevant in practice, where CSIT is only available to finite precision. For example, Gou et al. introduced in [34] an aligned interference neutralization scheme for the layered $2 \times 2 \times 2$ interference channel which is comprised of two source nodes, two relay nodes and two destination nodes, that achieves the sum De-

grees of Freedom (DoF) value of 2 under perfect CSIT. This is trivially optimal because even if all interference is eliminated, each user by itself cannot achieve more than 1 DoF — a straightforward consequence of the min-cut max-flow bound. The result is generalized to the $K \times K \times K$ setting in [58] where a network diagonalization scheme is shown to achieve K DoF, also trivially optimal for the same reason. Such schemes, that are based on precise alignment and/or neutralization of signals, are difficult to translate to practice because the residual interference due to imperfections in CSIT can be severely detrimental. Under perfect CSIT, even constrained alternatives like decode-and-forward, which can achieve 4/3 DoF by treating each hop as a 2×2 X channel [46, 39, 7, 8], are too fragile as they rely strongly on infinitely precise CSIT to achieve perfect interference alignment. Besides the assumption of perfect CSIT, another limitation of many of these works, e.g., [34, 64, 33, 58, 46, 39, 7, 8], is that their focus is limited to the DoF metric which implicitly assumes that all non-zero channels are equally strong (every non-zero link can carry exactly 1 DoF). To overcome this limitation, the Generalized Degrees of Freedom (GDoF) framework was introduced in [24], which is capable of representing weak and strong interference conditions and is the critical stepping stone to approximate capacity characterizations [24, 5, 4, 28, 6, 41, 53]. Evidently, for a robust information-theoretic understanding of multi-hop interference networks it is important to study their GDoF under finite precision CSIT.

Despite the early recognition of their importance [44], network GDoF characterizations under finite precision CSIT have been generally intractable until recently, mainly due to the difficulty of obtaining tight information theoretic outer bounds under CSIT limitations. Indeed, the highest performing achievable schemes under finite precision CSIT tend to be robust random coding schemes that are relatively well understood. Note that this is in sharp contrast to GDoF studies under *perfect* CSIT, where the outer bounds tend to be relatively straightforward (e.g., min-cut max-flow bounds) and the main challenge is the construction of sophisticated achievable schemes based on alignment and neutralization of signals. Under finite precision CSIT, the outer bounds tend to be challenging because they need

to rule out the potential benefits of all forms of signal alignments that are possible under perfect CSIT but fail under limited CSIT. Since received signals in an interference network are sums (linear combinations) of transmitted signals up to noise distortion, bounding the potential benefits of signal alignments amounts to bounding the size (entropy) of sum-sets (received signals), an inherently combinatorial endeavor that marks a seemingly necessary departure from the elegance of classical information theoretic arguments. This is indeed the approach taken by the so called Aligned Images (AI, in short) bounds that were introduced in [16] and recently expanded significantly in scope to a broad class of sum-set inequalities in [20]. AI Sum-set Inequalities have been applied successfully to find GDoF characterizations under finite precision CSIT for a variety of single-hop interference and broadcast settings [17, 19, 30, 21, 31, 18, 12, 66]. On the other hand, recent observations in [13] indicate that further generalizations of the AI Sum-set Inequalities may be needed beyond [20]. Given this relatively new but limited set of tools that have yet to be applied to multi-hop settings, the extent of their utility for multi-hop interference networks in particular remains an interesting open question. It is this open question that motivates our work in this chapter. An overview of our results is provided next.

To avoid the curse of dimensionality we begin our GDoF study with a symmetric, layered, 2-hop interference network, denoted as a $2 \times 2 \times 2$ setting, which is comprised of 2 source nodes, 2 relay nodes, and 2 destination nodes. Each hop is a 2×2 network, where the direct links are capable of carrying 1 GDoF, and the cross-links are capable of carrying α GDoF. Since well-designed networks invariably operate in the weak interference regime, our primary focus is on the weak interference regime ($\alpha < 1$), although extensions to strong interference are straightforward in this case. As we apply AI Sum-set Inequalities to this setting, an immediate challenge manifests in the critical first step. All prior applications of AI bounds begin by transforming the channel to a deterministic model by a sequence of steps that include removing the Additive White Gaussian Noise (AWGN) and quantizing the noiseless received signals. This transformation works well for one-hop settings because it can be

shown that all the steps involved can collectively only contribute a bounded distortion that is inconsequential in the GDoF sense. However, the same deterministic transformation is difficult to justify in a multi-hop setting. This is because the relays are free to choose *arbitrary* mappings from their input signals to their output signals, but for arbitrary mappings, a bounded distortion of inputs does not necessarily correspond to a bounded distortion of their corresponding outputs. Fortunately, we are able to overcome this obstacle by realizing that a valid outer bound is obtained if we allow perfect CSIT for the first hop and finite precision CSIT for just the second hop. This requires the deterministic transformation only for the second hop, i.e., only the *outputs* of the relays are distorted and not their inputs. Surprisingly, this outer bound is found to be achievable even with only finite precision CSIT for both hops. Specifically, using the compact GDoF expression available from the outer bound for insights, we are able to construct an achievable scheme that uses rate-splitting between amplify-and-forward and decode-and-forward strategies to match the outer bound. This settles the GDoF of the $2 \times 2 \times 2$ setting with finite precision CSIT in both hops, and also shows as a byproduct that the sum-GDoF do not improve even if perfect CSIT is allowed in the first hop. While the proof is non-trivial, it is notable the AI Sum-set Inequalities of [20] turn out to be sufficient for a tight GDoF characterization in this case. The optimal sum-GDoF value appears in Theorem 4.1 in Section 4.3.1. Also notable is that the results automatically translate to strong interference settings simply by switching the labels of the relays. This extension appears as Corollary 4.1 in Section 4.3.1.

Next we generalize the setting to a symmetric layered L-hop interference network, denoted as a $2 \times 2 \times \cdots \times 2$ setting. Here the idea of allowing perfect CSIT in all but the last hop does not work because the resulting bound would be loose for $L > 2$. Instead, a recursive approach is taken, that bounds the maximum mutual information that can be delivered from the source nodes to the nodes in the ℓ^{th} hop, given the maximum mutual information that can be delivered to the nodes in the $(\ell - 1)^{th}$ hop. As ℓ increases from 2 to L, at each stage of this recursive expansion, the deterministic transformation is used only for the last (ℓ^{th})

hop for that stage. This recursive approach, combined with the AI Sum-set Inequalities and the insights from the $L = 2$ setting, turns out to be sufficient to characterize the sum-GDoF value for the L-hop setting. As before, our focus is on the weak interference setting, for which the sum GDoF value is presented in Theorem 4.2 in Section 4.3.2. The result can be immediately extended to strong interference by switching the labels of the relays in every other hop, provided that the number of hops, L is even, thus giving us Corollary 4.2 in Section 4.3.2. Another remarkable aspect of this result is that as L approaches infinity, the sum-GDoF value approaches the sum-GDoF of the corresponding one-hop broadcast channel where the 2 sources are allowed to cooperate fully, under finite precision CSIT. From the achievability perspective, this happens because of a successive onion peeling approach that allows the relays in each successive stage to decode one more layer of interference, so that the common information accumulated asymptotically at the relays as L approaches infinity, is enough to match the broadcast channel where the transmitters cooperate fully.

4.2 System Model

Figure 4.1 depicts the layered symmetric L-hop interference channel model. Each hop is a 2×2 topology, comprised of 2 transmitters and 2 receivers. For the ℓ^{th} hop, $\ell \in [1 : L]$, the two transmitters are labeled as $\text{Tx}_{1[\ell]}$, $\text{Tx}_{2[\ell]}$, and the corresponding receivers are labeled as $\text{Rx}_{1[\ell]}, \text{Rx}_{2[\ell]}$ respectively. The receivers for the ℓ^{th} hop are the same physical nodes that act as the transmitters for the $(\ell+1)^{th}$ hop, i.e., $\text{Rx}_{i[\ell]} \equiv \text{Tx}_{i[\ell+1]}$, $\ell \in [1 : L-1], i \in [1 : 2]$. The transmitters for the first hop, $\text{Tx}_{1[1]}, \text{Tx}_{2[1]}$ are also referred to as *sources*, the receivers for the last hop, $\text{Rx}_{1[L]}, \text{Rx}_{2[L]}$ are also referred to as *destinations*, and the remaining nodes are also referred to as *relays*.

Suppose the communication takes place over N channel uses. There are two independent messages, $W_1 \in [1 : \lceil 2^{NR_1} \rceil], W_2 \in [1 : \lceil 2^{NR_2} \rceil]$, such that for $i \in [1 : 2]$, message W_i

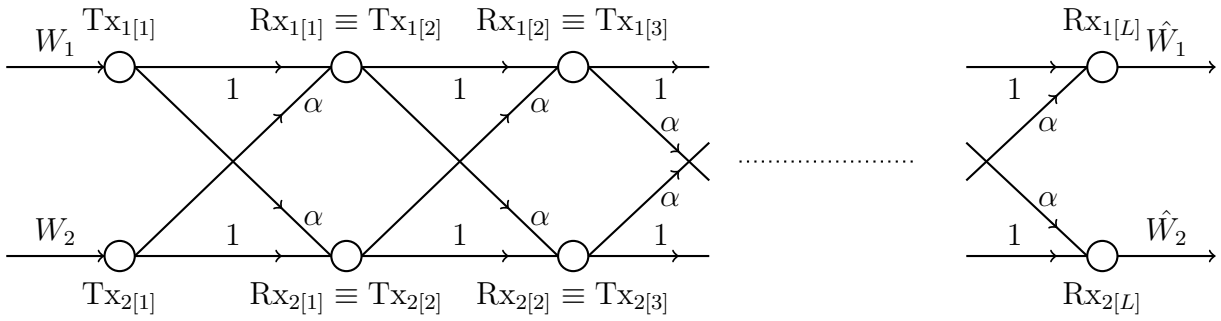

Figure 4.1: *Layered Symmetric L-hop Interference Channel model.*

originates from Source $\mathrm{Tx}_{i[1]}$ and is intended for Destination $\mathrm{Rx}_{i[L]}$ respectively. Following the GDoF formulation, under the n^{th} channel use, $n \in [1:N]$, the inputs and outputs of the ℓ^{th} hop are related as follows,

$$Y_{1[\ell]}(n) = \sqrt{P} G_{11[\ell]}(n) X_{1[\ell]}(n) + \sqrt{P^{\alpha}} G_{12[\ell]}(n) X_{2[\ell]}(n) + Z_{1[\ell]}(n) \tag{4.1}$$

$$Y_{2[\ell]}(n) = \sqrt{P^{\alpha}} G_{21[\ell]}(n) X_{1[\ell]}(n) + \sqrt{P} G_{22[\ell]}(n) X_{2[\ell]}(n) + Z_{2[\ell]}(n) \tag{4.2}$$

such that the signal sent from the transmitter $\mathrm{Tx}_{i[\ell]}$ is denoted as $X_{i[\ell]}(n)$, the signal observed by the receiver $\mathrm{Rx}_{i[\ell]}$ is denoted as $Y_{i[\ell]}(n)$, the channel coefficient between $\mathrm{Tx}_{i[\ell]}$ and $\mathrm{Rx}_{k[\ell]}$ is denoted as $G_{ki[\ell]}(n)$, the additive noise observed by the receiver $\mathrm{Rx}_{i[\ell]}$ is denoted as $Z_{i[\ell]}(n)$, and $i, k \in [1:2], n \in [1:N]$. All symbols are complex, the noise terms represent i.i.d. zero mean unit variance circularly symmetric Additive White Gaussian Noise (AWGN), and the transmitted symbols $X_{1[\ell]}(n)$, $X_{2[\ell]}(n)$ are each subject to a unit transmit power constraint.

We assume that the channel coefficients $G_{ki[\ell]}(n)$ follow the *bounded density assumption* of [16], i.e., all joint and conditional probability density functions exist and are bounded. To make this precise, let $\mathcal{G}$ be a set of real-valued random variables, that satisfies the following two conditions:

- All random variables in $\mathcal{G}$ are bounded away from zero and infinity, i.e., $g \in \mathcal{G} \implies |g| \in [1/\Delta, \Delta]$ for some positive finite constant Δ.

- There exists a finite positive constant f_{max}, such that for all finite cardinality disjoint subsets $\mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{G}$, the conditional probability density function $f_{\mathcal{G}_1|\mathcal{G}_2}$ exists and is bounded above by $f_{max}^{|\mathcal{G}_1|}$.

Now, if we represent each channel coefficient in terms of its real and imaginary components, $G_{ki[\ell]}(n) = G_{ki[\ell],R}(n) + jG_{ki[\ell],I}(n)$, then the bounded density assumption means that we require that $G_{ki[\ell],R}(n), G_{ki[\ell],I}(n)$ are distinct elements of $\mathcal{G}$ for $k, i \in [1:2], \ell \in [1:L], n \in [1:N]$.

Next, in order to specify the channel knowledge assumptions, let us define

$$\mathcal{G}_{[\ell]} = \{G_{uv[\ell]}(n) : u, v \in [1:2], n \in [1:N]\} \tag{4.3}$$

as the subset of $\mathcal{G}$ comprised of only the channel coefficients associated with the ℓ^{th} hop. For simplicity,[1] and since this is a common assumption in practice, let us assume that the channels across different hops are independent. Also, channels are independent of messages and additive noise terms.

Similarly, define $\mathcal{G}_{[\ell_1:\ell_2]} = \bigcup_{l=\ell_1}^{\ell_2} \mathcal{G}_{[l]}$ as the subset of $\mathcal{G}$ comprised of all channel coefficients across hops $[\ell_1 : \ell_2]$. We assume that precise channel state information is available at the receivers (CSIR) for all channels in the *preceding* hops. Specifically, the receivers in the ℓ^{th} hop, $\mathrm{Rx}_{1[\ell]}, \mathrm{Rx}_{2[\ell]}$, have precise knowledge of the realizations of all random variables in $\mathcal{G}_{[1:\ell]}$. Furthermore, since the receivers in the ℓ^{th} hop are the same as the transmitters in the $(\ell+1)^{th}$ hop, $\mathrm{Rx}_{k[\ell]} \equiv \mathrm{Tx}_{k[\ell+1]}$, we allow that the same precise channel knowledge of $\mathcal{G}_{[1:\ell]}$ is available to $\mathrm{Tx}_{1[\ell+1]}, \mathrm{Tx}_{2[\ell+1]}$. The knowledge of all remaining channel coefficients is limited to their joint probability density functions. The assumption that these probability density functions satisfy the bounded density assumption is what limits the CSIT to finite precision.

[1] This assumption is not strictly necessary for our results, but it will simplify the analysis.

Note that the CSIT assumptions imply that

$$I(X_{1[\ell]}^{[N]}, X_{2[\ell]}^{[N]}, \mathcal{G}_{[1:\ell-1]}; \mathcal{G}_{[\ell:L]}) = 0. \tag{4.4}$$

This is because the transmitters over the ℓ^{th} hop have no knowledge of channel *realizations* beyond what can be passed to them from preceding hops.

Remark: The assumption that CSIT is available for preceding hops at each node strengthens the GDoF converse bounds, because additional channel knowledge cannot hurt, but it is noteworthy that the achievable schemes presented in this chapter that meet those bounds do not make use of this CSIT at any encoder. The receivers do utilize the corresponding CSIR of all preceding hops. Similarly, let us note that while we allow a receiver to have perfect CSIR for the channels associated with the *other* receiver in the same hop, e.g., $\text{Rx}_{1[\ell]}$ has perfect knowledge of $G_{22[\ell]}(n)$, such knowledge is not used by the achievable scheme either. As such this assumption also serves mainly to strengthen the converse, and our GDoF results hold both with and without it.

While the bounded density assumption allows fairly general distributions for the channel coefficients, an interesting perspective of the channel coefficients is to view them as small *perturbations*, say i.i.d. uniform in a small interval around 1, such that the length of that interval corresponds to the finite precision constraint — the shorter the length of the perturbation interval, the more precisely the channels are revealed by their statistics, and the larger the peak value of the probability density function. This also explains the need for density functions to be bounded in order to limit CSIT to finite precision. The channel coefficients typically represent physical phenomena like channel fading, but viewed as perturbations they can also represent artifacts that are deliberately introduced into the GDoF model in order to filter out or eliminate the possibility of fragile schemes emerging as optimal solutions, thus allowing us to explore information theoretic optimality of random coding based solutions

that are also practically appealing for their robustness.

Recall that P is a nominal parameter that approaches infinity to define the GDoF limit, and the exponents that appear with P in (4.1),(4.2) represent coarse channel strength parameters that are assumed globally known (equivalently, channel strengths in the absence of perturbations). Specifically, for our symmetric model, the direct links (between $\mathrm{Tx}_{i[\ell]}$ and $\mathrm{Rx}_{i[\ell]}$) have channel strength corresponding to the exponent 1 and cross links have coarse channel strength corresponding to the exponent $\alpha \in \mathbb{R}^+$. Because well-designed networks tend to operate in the weak interference regime, our focus is on the setting $\alpha < 1$, although some of our results generalize to strong interference settings in a straightforward manner. Intuitively, we may think of each of these channel strength parameters as the (approximate) capacity of the corresponding point to point link in its original finite SNR setting, and think of $\log(P)$ as a uniform scaling factor that is simultaneously applied to the capacities of all the links. Since each link capacity is logarithmic in P, linear scaling of capacity corresponds to exponential scaling of P, and the original channel capacities α appear as exponents. The fundamental intuition behind GDoF is that if the capacity of every link in a network is scaled by the same constant factor ($\log(P)$), then the network capacity should also scale (approximately) by the same factor. So normalizing the sum-capacity of the network by $\log(P)$ should produce an approximation to the capacity of the original network. This is indeed why we see normalizations by $\log(P)$ in the definition of GDoF, as presented next.

The rate pair (R_1, R_2) is said to be achievable if there exists a scheme, comprised of encoding functions at the sources, mappings from inputs to outputs at each of the relays, and decoding functions at the destinations, under which $\mathrm{Rx}_{1[L]}, \mathrm{Rx}_{2[L]}$ can decode W_1, W_2 respectively with arbitrarily small error probability in the standard Shannon-theoretic sense[22]. The closure

of achievable rate tuples is the capacity region $\mathcal{C}(P)$. The GDoF region is defined as

$$\mathcal{D}^{f.p.} = \left\{ \begin{array}{ll} (d_1, d_2): & \exists (R_1(P), R_2(P)) \in \mathcal{C}(P) \\ & s.t. \quad d_1 = \lim\limits_{P\to\infty} \frac{R_1(P)}{\log(P)}, \quad d_2 = \lim\limits_{P\to\infty} \frac{R_2(P)}{\log(P)} \end{array} \right\}. \tag{4.5}$$

The superscript '*f.p.*' highlights the finite precision CSIT constraint. Finally, the sum-GDoF value is defined as $\mathcal{D}_{\Sigma}^{f.p.} = \max\limits_{(d_1,d_2)\in\mathcal{D}^{f.p.}} (d_1 + d_2)$.

4.3 Results

Following the information theoretic mindset of starting from the elemental scenarios, the simplest multi-hop setting, where $L = 2$, i.e., the 2-hop interference channel (especially in the weak interference regime, $\alpha \leq 1$) is our main focus in this chapter. Our main result is the sum-GDoF characterization for this channel under finite precision CSIT, presented in Section 4.3.1. Due to its relative simplicity the 2-hop setting is also instructive to introduce the main ideas in their simplest form, whose generalizations eventually allow us to find the sum-GDoF for arbitrary L, as presented in Section 4.3.2.

4.3.1 Sum-GDoF of the 2-hop Layered Symmetric Interference Channel under Finite Precision CSIT

Weak Interference Regime: $\alpha \leq 1$

Theorem 4.1. *For the 2-hop layered symmetric interference channel under finite precision CSIT, in the weak interference regime $\alpha \leq 1$, the sum-GDoF value is given by,*

$$\mathcal{D}_{\Sigma}^{f.p.} = \begin{cases} 2 - 4\alpha/3, & 0 \leq \alpha \leq 1/2, \\ 2/3 + 4\alpha/3, & 1/2 \leq \alpha \leq 4/7, \\ 2 - \alpha, & 4/7 \leq \alpha \leq 1. \end{cases} \tag{4.6}$$

The converse proof for Theorem 4.1 is provided in Section 4.4.3. The converse for the regime $4/7 \leq \alpha \leq 1$ is already available because it corresponds to the sum-GDoF value established in [16] under finite precision CSIT for the MISO broadcast channel[2] that is obtained by allowing full cooperation (which cannot hurt) among all nodes except the two destination nodes. However, the converse for the remaining regime, $0 \leq \alpha \leq \frac{4}{7}$, is non-trivial and is obtained in this chapter based on the sum-set inequalities of [20]. One of the challenging aspects of the converse is that the deterministic transformation that is the starting point of all prior applications of Aligned Images bounds [17, 19, 30, 21, 31, 18, 12, 66], is not directly applicable to the multi-hop setting as explained in the introduction. This challenge is overcome essentially by allowing perfect CSIT in the first hop (which cannot hurt) and only using the deterministic transformation for the second hop. Since this produces a tight converse bound that is achievable with finite precision CSIT in both hops, evidently the sum-GDoF value is the same (given by Theorem 4.1) whether the CSIT in the first hop is perfect or restricted to finite precision, as long as the CSIT in the second hop is limited to

[2]The MISO broadcast channel refers to a two user broadcast channel, where the transmitter is equipped with two antennas and each user is equipped with one antenna.

finite precision.

The achievability for Theorem 4.1 is proved in Section 4.5.1. The achievable scheme is straightforward when $2/3 \leq \alpha \leq 1$, because it corresponds to a concatenation of two interference channels [24, 18] where the intermediate nodes (the relays) simply employ a decode-and-forward strategy. The achievable scheme is non-trivial for the remaining regime $0 \leq \alpha \leq \frac{2}{3}$ and relies on a rate-splitting approach that is comprised of amplify-and-forward and decode-and-forward schemes. Specifically, the sources split their messages into sub-messages, the relays are able to decode-and-forward some of the sub-messages, while they amplify-and-forward the remaining superposition of codewords that they are not able to decode. The relays further split the sub-messages that they are able to decode and then use a different superposition approach (assigning different powers) to transmit the decoded sub-messages. With proper choice of rate-splitting and superposition parameters, the destinations are able to decode their desired messages by a successive decoding approach.

Extension to Strong Interference Regime: $\alpha \geq 1$

As noted previously our main focus is on the weak interference regime, $\alpha \leq 1$. However, the extension of Theorem 4.1 to the strong interference regime, where $\alpha \geq 1$, turns out to be straightforward for the 2-hop setting, as stated in the following corollary.

Corollary 4.1. *For the 2-hop layered symmetric interference channel under finite precision CSIT, in the strong interference regime $\alpha \geq 1$, the sum-GDoF value is given by,*

$$\mathcal{D}_{\Sigma}^{f.p.} = \begin{cases} 2\alpha - 1, & 1 \leq \alpha \leq 7/4, \\ 2\alpha/3 + 4/3, & 7/4 \leq \alpha \leq 2, \\ 2\alpha - 4/3, & \alpha \geq 2. \end{cases} \tag{4.7}$$

The corollary follows from Theorem 4.1 directly, because switching the labels of the two

relays immediately converts the weak interference setting into a strong interference setting. Specifically, switching the relays gives us a channel where the direct channels have strength α and the cross-channels have strength 1. Now, let us scale all channel strength parameters by $1/\alpha$, so that we have direct channels with strength $\alpha \times 1/\alpha = 1$ and cross-channels with strength $\alpha' = 1 \times 1/\alpha > 1$. It follows from the definition of GDoF that if all channel strength parameters are scaled by the same constant,[3] then the GDoF value will be scaled by precisely the same constant as well. Thus, if we denote the sum-GDoF as a function of α as $\mathcal{D}_{\Sigma}^{f.p.}(\alpha)$, then we must have $\mathcal{D}_{\Sigma}^{f.p.}(1/\alpha) = 1/\alpha \mathcal{D}_{\Sigma}^{f.p.}(\alpha)$ for $\alpha \leq 1$, or equivalently, $\mathcal{D}_{\Sigma}^{f.p.}(\alpha') = \alpha' \mathcal{D}_{\Sigma}^{f.p.}(1/\alpha')$ for $\alpha' > 1$, which gives us Corollary 4.1. Thus, Theorem 4.1 and Corollary 4.1 together fully characterize the sum-GDoF value of the 2-hop layered symmetric interference channel under finite precision CSIT, for all α.

Comparisons

To place the sum-GDoF result in perspective, let us compare it against a few benchmarks, as illustrated in Figure 4.2. These benchmarks are explained below.

- **Optimal Sum-GDoF under Perfect CSIT**: Recall the aligned interference neutralization scheme introduced by Gou et al. in [34], which was originally used to show that a sum-DoF value of 2 is achievable for the 2-hop interference channel under perfect CSIT. It is not difficult to apply the same scheme to find the sum-*GDoF* value, $\mathcal{D}_{\Sigma}^{p}$ under perfect CSIT (the 'p' in the superscript stands for '*perfect*' CSIT), which turns out to be equal to the sum-GDoF value of the one-hop MISO broadcast channel with

[3]Essentially this corresponds to defining $P' = P^{\alpha}$ and substituting for P with P', so that $P \to {P'}^{1/\alpha} = P'^{\alpha'}$, and $P^{\alpha} \to P'$. The normalization factor in the definition of GDoF similarly maps as $\log(P) \to 1/\alpha \log(P')$.

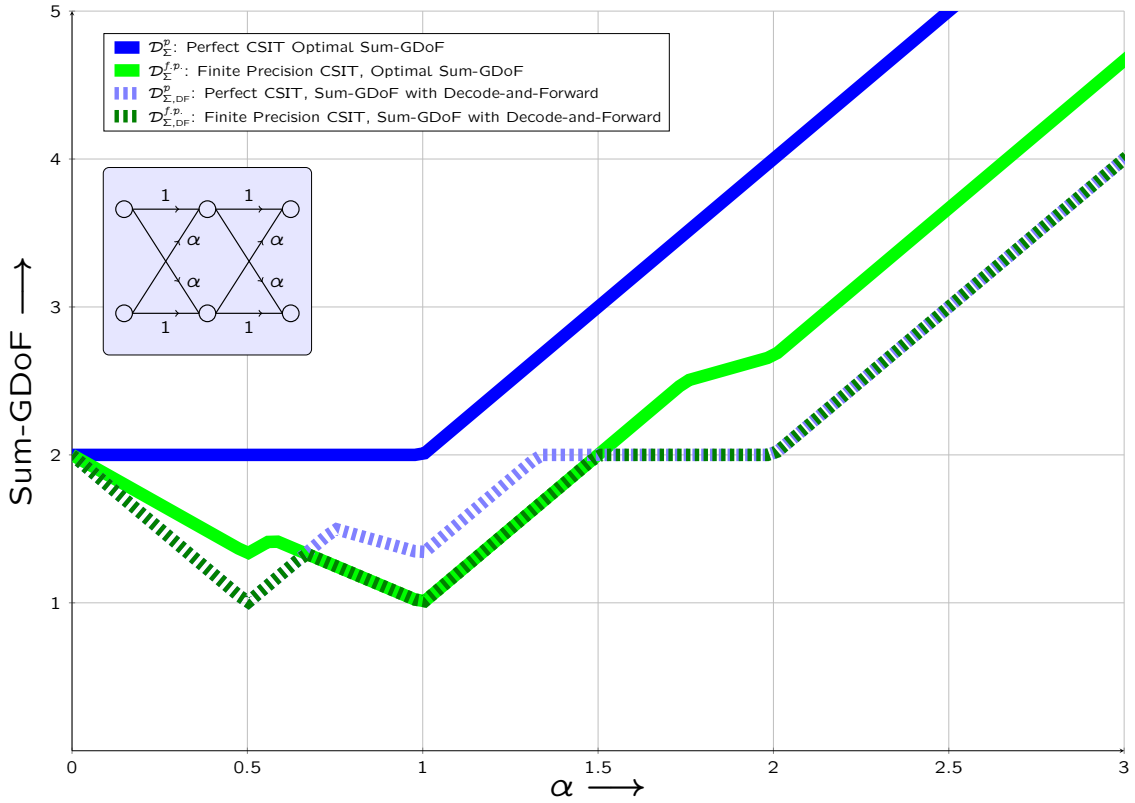

Figure 4.2: *Sum-GDoF comparisons for the layered symmetric 2-hop interference channel.*

perfect CSIT.

$$\mathcal{D}_{\Sigma}^{p} = \begin{cases} 2, & \alpha \leq 1, \\ 2\alpha, & \alpha \geq 1. \end{cases} \tag{4.8}$$

- **Sum-GDoF with Decode-and-Forward under Finite Precision CSIT**: A decode-and-forward solution for the 2-hop interference channel corresponds to treating each hop as an X channel. Recall that a 2×2 X channel is a one-hop setting with two transmitters, two receivers, and 4 independent messages, one from each transmitter to each receiver. By treating each hop as an X channel, each transmitter in the 2-hop interference channel is able to split its message into two independent sub-messages, that are then decoded by different relays, and re-encoded for the X channel on the next hop so that they can be finally decoded by their desired destination. Using the

sum-GDoF of the X channel under finite precision CSIT as characterized in [18] we obtain the sum-GDoF value of the 2-hop layered symmetric interference channel with decode-and-forward under finite precision CSIT as follows.

$$\mathcal{D}^{f.p.}_{\Sigma,\mathrm{DF}} = \begin{cases} 2-2\alpha, & \alpha \leq \frac{1}{2}, \\ 2\alpha, & \frac{1}{2} \leq \alpha \leq \frac{2}{3}, \\ 2-\alpha, & \frac{2}{3} \leq \alpha \leq 1, \\ 2\alpha-1, & 1 \leq \alpha \leq \frac{3}{2}, \\ 2, & \frac{3}{2} \leq \alpha \leq 2, \\ 2\alpha-2, & \alpha \geq 2. \end{cases} \tag{4.9}$$

- **Sum-GDoF with Decode-and-Forward under Perfect CSIT**: Using the sum-GDoF value of the X channel under perfect CSIT as characterized in [36, 7, 39, 46], we obtain the sum-GDoF value of the 2-hop layered symmetric interference channel with decode-and-forward under perfect CSIT as follows.

$$\mathcal{D}^{p}_{\Sigma,\mathrm{DF}} = \begin{cases} 2-2\alpha, & \alpha \leq \frac{1}{2}, \\ 2\alpha, & \frac{1}{2} \leq \alpha \leq \frac{3}{4}, \\ \frac{6-2\alpha}{3}, & \frac{3}{4} \leq \alpha \leq 1, \\ \frac{6\alpha-2}{3}, & 1 \leq \alpha \leq \frac{4}{3}, \\ 2, & \frac{4}{3} \leq \alpha \leq 2, \\ 2\alpha-2, & \alpha \geq 2. \end{cases} \tag{4.10}$$

From Figure 4.2 we note that except for the degenerate case of $\alpha = 0$, there is always a significant loss of sum-GDoF relative to its optimal value under perfect CSIT, i.e., the GDoF benefits of aligned interference neutralization [34] are pervasive and powerful under perfect CSIT but too fragile to survive under finite precision CSIT. Remarkably, we note that $\mathcal{D}^{f.p.}_{\Sigma,\mathrm{DF}} = \min\left(\mathcal{D}^{p}_{\Sigma,\mathrm{DF}}, \mathcal{D}^{f.p.}_{\Sigma}\right)$ even though $\mathcal{D}^{p}_{\Sigma,\mathrm{DF}}$ and $\mathcal{D}^{f.p.}_{\Sigma}$ are almost always different values

(with the exception of cross-overs that occur at $\alpha = 2/3, 3/2$). Thus, relative to the baseline of robust (finite precision) decode-and-forward, the robust (finite precision) gains of multi-hopping appear in the regimes where cross-channels are significantly weaker or stronger, i.e., $\alpha \leq 2/3, \alpha \geq 3/2$, whereas the fragile gains of interference alignment under perfect CSIT appear precisely in the complementary regime $2/3 \leq \alpha \leq 3/2$ where the cross-channels are relatively of similar strength as direct channels. Remarkably, when $\alpha \leq 2/3$ and $\alpha \geq 3/2$, the optimal scheme takes advantage of partial decoding: the relays partially decode part of the information to construct the common information, and they forward the remaining interfered information in its mixed form, leaving it for the destination nodes to resolve the interfered information at a later stage with the help of the common information.

4.3.2 Sum-GDoF of the L-hop Layered Symmetric Interference Channel under Finite Precision CSIT

Building on the insights from the 2-hop setting, in this section we generalize the sum-GDoF results to the L-hop case under finite precision CSIT. As before we start with the weak interference regime.

Weak Interference Regime: $\alpha \leq 1$

Theorem 4.2. *For the L-hop layered symmetric interference channel under finite precision CSIT,*[4] *in the weak interference regime $\alpha \leq 1$, the sum-GDoF value is given by,*

$$\mathcal{D}_{\Sigma}^{f.p.} = \begin{cases} 2 - \alpha - \alpha/(2^L - 1), & 0 \leq \alpha \leq 1/2, \\ 1 + \alpha - (1-\alpha)/(2^L - 1), & 1/2 \leq \alpha \leq 2^L/(2^{L+1} - 1), \\ 2 - \alpha, & 2^L/(2^{L+1} - 1) \leq \alpha \leq 1. \end{cases} \tag{4.11}$$

[4]The sum-GDoF value with perfect CSIT for any $L \geq 2$ is the same as $L = 2$, as discussed in [34].

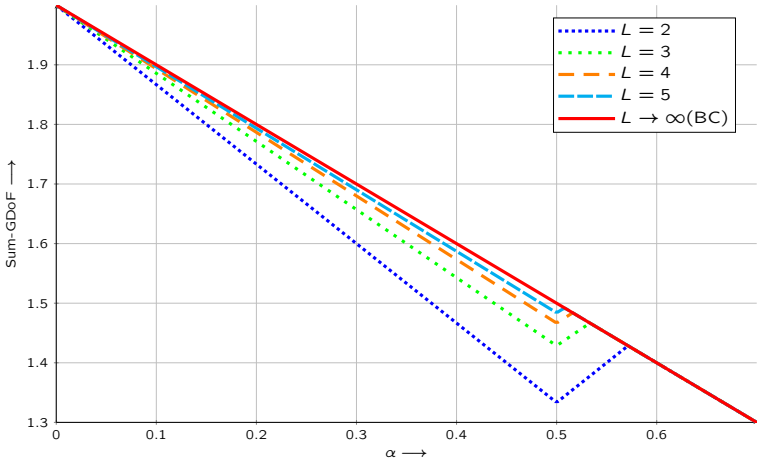

Figure 4.3: *Sum-GDoF of the layered symmetric L-hop interference channel in a subinterval of the weak interference regime.*

The sum-GDoF value specified in Theorem 4.2 is illustrated in Figure 4.3 over a subinterval of the weak interference regime (since the plots are close together, the figure is zoomed in for clarity) for various L. Evidently, as L approaches infinity, the sum-GDoF value of the L-hop interference channel approaches that of the corresponding one-hop broadcast channel with the same cross link strength under finite precision CSIT. The intuition is that a successive onion peeling approach allows the relays in each successive stage to decode one more layer of interference. Therefore, more common information can be decoded at the successive relays and they have more cooperation capability compared to the precedent stage. As $L \to \infty$, the common information accumulated at the relays is enough to match the broadcast channel, in which the transmitters cooperate fully.

The converse for Theorem 4.2 is proved in Section 4.4.4. For the converse proof, the regime $2^L/(2^{L+1}-1) \leq \alpha \leq 1$ is straightforward because this is simply the GDoF value of the broadcast channel [16] that is obtained by allowing full cooperation among all nodes except the destination nodes. For the remaining regimes, as with the 2-hop case, a challenging aspect is the deterministic transformation. Whereas in the 2-hop case it was sufficient to enforce finite precision CSIT only in the last hop, the same idea does not work directly in the L-hop setting. Instead the problem is circumvented by first considering only ℓ hops at a time, as in Lemma 4.1 that appears in Section 4.4.2, and enforcing finite precision CSIT in

the 'last' (i.e., the ℓ^{th} hop) to bound the mutual information that can be transferred from the sources to the receivers in the ℓ^{th} hop. Then a recursive argument is developed in Lemma 4.4 in Section 4.4.2 to obtain a bound for ℓ hops based on the bound for $\ell - 1$ hops.

The proof of achievability for Theorem 4.2 appears in Section 4.5.2. The regime $2/3 \leq \alpha \leq 1$ is straightforward as in the 2-hop case, because it corresponds to a concatenation of L interference channels [24, 18] and a simple decode-and-forward strategy suffices. In other regimes however, the achievable scheme for the L-hop setting is a non-trivial extension of the 2-hop case. While in principle the construction is still based on rate-splitting between amplify-and-forward and decode-and-forward schemes, there is an important element of onion-peeling which allows each successive relay stage to decode one more layer of interference, so that with each hop the nodes acquire more common information and are closer to acting as a broadcast channel. Indeed, as the number of hops $L \to \infty$, the sum-GDoF value does approach that of a broadcast channel where all information is shared between the two transmitters of the last hop.

Extension to Strong Interference Regime: $\alpha \geq 1$

As in the 2-hop case, the sum-GDoF result in Theorem 4.2 for the weak interference regime immediately implies an extension to the strong interference regime by the same argument of switching relay positions, and is presented in the following corollary.

Corollary 4.2. *If L is even, then for the L-hop layered symmetric interference channel under finite precision CSIT, in the strong interference regime $\alpha \geq 1$, the sum-GDoF value is given by,*

$$\mathcal{D}_{\Sigma}^{f.p.} = \begin{cases} 2\alpha - 1, & 1 \leq \alpha \leq 2 - 2^{-L}, \\ \alpha + 1 - (\alpha - 1)/(2^L - 1), & 2 - 2^{-L} \leq \alpha \leq 2, \\ 2\alpha - 1 - 1/(2^L - 1), & \alpha \geq 2. \end{cases} \tag{4.12}$$

A new constraint appears in Corollary 4.2, that L must be even. This is because the idea of interchanging the positions of the relay nodes to convert weak interference into strong interference only works when we switch relays in every *other* hop, which can only be done if the number of hops is even. To see this explicitly, consider Figure 4.4 which shows an $L = 4$ hop setting. The bold edges represent strong channels while the dashed edges represent weak channels. The original network topology is shown on the left side of Figure 4.4, where the cross-channels are weak and the direct channels are strong. Now, if we re-draw the *same* network but switch the positions of the dark red relay with the dark blue relay, and the light red relay with the light blue relay, then we obtain the representation shown on the right side of Figure 4.4, where the direct channels are weak and cross channels are strong. However, this idea of switching the positions of relays in every alternate hop only works when L is even.

Sum-GDoF vs L: Non-Monotonicity

Corollary 4.2 establishes the sum-GDoF in the strong interference regime when L is even, but leaves the sum-GDoF open for odd L in the same regime. One might expect that the GDoF values for odd L may be sandwiched between their even neighbors. The expectation is supported by the observation that the expressions in (4.11) and (4.12) as well as the illustration in Figure 4.3 all seem to show that the Sum-GDoF value monotonically increases with the number of hops, L. In fact, the gap between plots is rather small in Figure 4.3, which suggests that the sum-GDoF values for odd L may be estimated quite accurately from

Figure 4.4: *Two representations of the same network. Interchanging the positions of relay nodes in every* other *hop changes the representation of the network from a weak interference setting to a strong interference setting. This works only when the number of hops, L, is even.*

the neighboring even L values. Somewhat surprisingly, this is not the case, as we show in this section. To highlight the non-monotonic behavior of sum-GDoF vs L, we characterize the sum-GDoF for odd L in the *very* strong interference regime, in the following theorem.

Theorem 4.3. *If L is odd, then for the L-hop layered symmetric interference channel under finite precision CSIT, in the very strong interference regime where $\alpha \geq L+1$, the sum-GDoF value is given by,*

$$\mathcal{D}_{\Sigma}^{f.p.} = 2L = \mathcal{D}_{\Sigma}^{p}. \tag{4.13}$$

The converse proof of Theorem 4.3 is presented in Section 4.4.5 and the achievability is proved in Section 4.5.3. Both are relatively straightforward. The converse is simply the min-cut bound, and achievability is a rate-splitting partitioning of multiple decode-and-forward schemes that require some filtering and rearrangement of the superposition order of codewords as they pass through the relays. Since the min-cut bound applies equally under perfect CSIT, the result of Theorem 4.3 also holds under perfect CSIT.

From Corollary 4.2 and Theorem 4.3 we note that as $\alpha \to \infty$ the sum-GDoF value of the L-hop layered symmetric interference channel approaches infinity if L is even, but is only $2L$ if L is odd, thus proving that the sum-GDoF value is not a monotonic function of L. To see this intuitively, consider again the network shown on the right side of Figure 4.4 and for this intuitive understanding assume that the dashed links are extremely weak (say, strength 0) while the solid links are extremely strong (say, strength approaching infinity). In this $L = 4$ hop network, consider the communication from Source 1 to Destination 1, for which there exists a very strong path, so the GDoF of this communication approach infinity. However, suppose the network had only 3 hops, so Destination 1 was the light blue node. Note that in this $L = 3$ hop network there exists no path from Source 1 to Destination 1, i.e., the GDoF of this communication is 0. This toy example intuitively shows why we notice abrupt drops

of sum-GDoF for odd L in the very strong interference regime. Figure 4.5 illustrates this fact as we note the different behaviors of sum-GDoF vs L in the weak (monotonic) and very strong (non-monotonic) regimes.

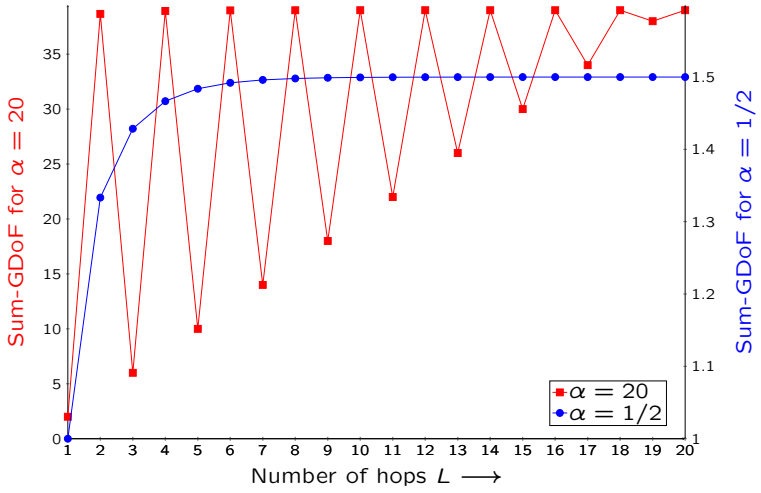

Figure 4.5: *Sum-GDoF of the layered symmetric L-hop interference channel vs the number of hops L for $\alpha = 1/2$ (weak interference) shown in blue, and $\alpha = 20$ (very strong interference) shown in red.*

To summarize our results for the L-hop layered symmetric interference channel, we have found the robust sum-GDoF for arbitrary number of hops L if the network is in the weak interference regime ($\alpha \leq 1$) which is our main focus, or the very strong interference regime ($\alpha \geq L + 1$). For the remaining strong interference regime ($1 \leq \alpha \leq L + 1$), we have found the sum-GDoF if L is even, but the sum-GDoF value for odd L remains a non-trivial open problem in the sense that the answer may not even be approximated by sandwiching between adjacent even values of L.

4.4 Converse Proofs

In this section we provide the converse proofs for Theorem 4.1, Theorem 4.2 and Theorem 4.3. We start with the basic definitions, inherited from [16][20], that are essential for Aligned

Images bounds.

4.4.1 Definitions

Definition 4.1 (Power Levels)**.** *An integer valued random variable X_i with power level λ_i takes values over the alphabet set $\mathcal{X}_{\lambda_i}$ defined as*

$$\mathcal{X}_{\lambda_i} \triangleq \{0, 1, 2, \cdots, \overline{P}^{\lambda_i} - 1\} \tag{4.14}$$

where $\overline{P}^{\lambda_i} \triangleq \lfloor \sqrt{P^{\lambda_i}} \rfloor$. We are primarily interested in limits as $P \to \infty$, where $P \in \mathbb{R}_+$ is referred to as power.

Definition 4.2. *For an integer valued random variable $X \in \mathcal{X}_\lambda$, and any non-negative real numbers λ_1, λ_2 such that $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda$, define*

$$(X)^{\lambda_2} \triangleq \Big\lfloor \frac{X}{\overline{P}^{\lambda - \lambda_2}} \Big\rfloor, \tag{4.15}$$

$$(X)_{\lambda_1} \triangleq X - \overline{P}^{\lambda_1} \Big\lfloor \frac{X}{\overline{P}^{\lambda_1}} \Big\rfloor, \tag{4.16}$$

$$(X)_{\lambda_1}^{\lambda_2} \triangleq \Big\lfloor \frac{(X)_{\lambda_2}}{\overline{P}^{\lambda_1}} \Big\rfloor. \tag{4.17}$$

In other words, $(X)^{\lambda_2}$ retrieves the top λ_2 power levels of X, $(X)_{\lambda_1}$ retrieves the bottom λ_1 power levels of X and $(X)_{\lambda_1}^{\lambda_2}$ retrieves the partition of X between power levels λ_1 and λ_2. As a somewhat oversimplified interpretation for intuitive purposes, X can be thought as a non-negative integer value represented in $\sqrt{P}$-ary alphabet expansion, as $X = x_\lambda x_{\lambda-1} \cdots x_2 x_1$, and $(X)^{\lambda_2}$ retrieves the most significant λ_2 symbols, i.e., $(X)^{\lambda_2} = x_\lambda \cdots x_{\lambda-\lambda_2+1}$. Similarly, $(X)_{\lambda_1}^{\lambda_2}$ is the sub-string $x_{\lambda_2} \cdots x_{\lambda_1}$, $(X)_{\lambda_1}$ is the sub-string $x_{\lambda_1 - 1} \cdots x_1$. This is oversimplified because $\lambda_1, \lambda_2, \lambda$ are not restricted to take only integer values. This is a generalization to the ADT models [5], where binary expansions are used to study GDoF under perfect CSIT.

Definition 4.3 (Sub-section, Interval, Level, Size, Disjoint)**.** *For $X \in \mathcal{X}_\lambda$, we define $(X)_{\lambda_1}^{\lambda_2}$ as a 'sub-section' of X if $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda$, where (λ_1, λ_2) is the corresponding 'interval'. Furthermore, we define the lower end of the interval (λ_1, λ_2) as the 'level' of the partition, denoted as $\ell((X)_{\lambda_1}^{\lambda_2}) = \lambda_1$. The length of the interval (λ_1, λ_2), denoted as $\mathcal{T}((X)_{\lambda_1}^{\lambda_2}) = \lambda_2 - \lambda_1$ is called the 'size' of the partition. Sub-sections $(X)_{\lambda_1}^{\lambda_2}$ and $(X)_{\nu_1}^{\nu_2}$ of the same $X \in \mathcal{X}_\lambda$ are 'disjoint' if the two intervals (λ_1, λ_2) and (ν_1, ν_2) are disjoint.*

Next we recall the definition of the particular deterministic transformation [16] that is used for Aligned Images bounds. The transformation has thus far been used only in single-hop settings, and as noted previously, extensions to multi-hop settings are not immediate. Fortunately, for our purpose and for all our arguments we only need to apply the deterministic transformation to one of the L hops at any time, say the ℓ^{th} hop. This transformation for the ℓ^{th} hop is defined next.

Definition 4.4 (Deterministic Transformation of the ℓ^{th} hop)**.** *In the ℓ^{th} hop, define the mapping from the original input $X_{i[\ell]}$ to the deterministic input $\overline{X}_{i[\ell]}$ as*

$$\overline{X}_{i[\ell]} = \lfloor \sqrt{P^{\max(1,\alpha)}} X_{i[\ell]} \rfloor \mod \lceil \sqrt{P^{\max(1,\alpha)}} \rceil \tag{4.18}$$

such that $\overline{X}_{i[\ell]}(n) = \overline{X}_{iR[\ell]}(n) + j\overline{X}_{iI[\ell]}(n)$, $i \in [1:2]$ and $\overline{X}_{iR[\ell]}(n), \overline{X}_{iI[\ell]}(n) \in \{0, 1, 2, \cdots, \lceil \sqrt{P^{\max(1,\alpha)}} \rceil - 1\}$ for all $n \in [1:N]$. Then the deterministic transformation for the ℓ^{th} hop is represented as follows:

$$\overline{Y}_{1[\ell]}(n) = \lfloor \sqrt{P^{1-\max(1,\alpha)}} G_{11[\ell]}(n) \overline{X}_{1[\ell]}(n) \rfloor + \lfloor \sqrt{P^{\alpha-\max(1,\alpha)}} G_{12[\ell]}(n) \overline{X}_{2[\ell]}(n) \rfloor \tag{4.19}$$

$$\overline{Y}_{2[\ell]}(n) = \lfloor \sqrt{P^{\alpha-\max(1,\alpha)}} G_{21[\ell]}(n) \overline{X}_{1[\ell]}(n) \rfloor + \lfloor \sqrt{P^{1-\max(1,\alpha)}} G_{22[\ell]}(n) \overline{X}_{2[\ell]}(n) \rfloor \tag{4.20}$$

Note that $|\overline{Y}_{i[\ell]}(n)| \leq 4\sqrt{P}\Delta$, and since the real and imaginary parts of $\overline{Y}_{i[\ell]}(n)$ are both integer valued, we must have $H(\overline{Y}_{i[\ell]}(n)) \leq 2\log(8\sqrt{P}\Delta) = \log(P) + o(\log(P))$. Similarly

over N channel uses, we must have

$$H(\overline{Y}_{i[\ell]}^{[N]}) \leq N\log(P) + No(\log(P)). \tag{4.21}$$

As noted in [66], we can also represent this as:

$$\overline{Y}_{1[\ell]}(n) = \lfloor G_{11[\ell]}(n)\left(\overline{X}_{1[\ell]}(n)\right)^1 \rfloor + \lfloor G_{12[\ell]}(n)\left(\overline{X}_{2[\ell]}(n)\right)^\alpha \rfloor + \zeta_{1[\ell]}(n) \tag{4.22}$$

$$\overline{Y}_{2[\ell]}(n) = \lfloor G_{21[\ell]}(n)\left(\overline{X}_{1[\ell]}(n)\right)^\alpha \rfloor + \lfloor G_{22[\ell]}(n)\left(\overline{X}_{2[\ell]}(n)\right)^1 \rfloor + \zeta_{2[\ell]}(n) \tag{4.23}$$

where $\zeta_{1[\ell]}(n), \zeta_{2[\ell]}(n)$ are complex random variables whose real and imaginary parts are integer valued and whose magnitude is bounded, so it does not scale with P. Specifically, $\max(|\zeta_{1[\ell]}(n)|, |\zeta_{2[\ell]}(n)|) \leq 2(2+\Delta) = o(\log(P))$, and Δ corresponds to the bounded range of the channel coefficients, i.e., $|G_{ij[\ell],R}(n)|, |G_{ij[\ell],I}(n)| \in [1/\Delta, \Delta], i, j \in [1:2], \ell \in [1:L]$, as previously stated in Section 4.2. Note that the superscript, e.g., $(\overline{X}_{i[\ell]})^\alpha$, refers to the top α power levels of $\overline{X}_{i[\ell]}$, according to Definition 2.

4.4.2 Lemmas

Our first three lemmas are inherited from prior works [16, 18, 20, 66] on Aligned Images bounds in single-hop scenarios, and specialized to our multi-hop setting where the deterministic transformation has been applied only to the ℓ^{th} hop.

Lemma 4.1 (Deterministic Bound [16])**.** *With the deterministic transformation applied only to the ℓ^{th} hop, we have the following bound for $i \in [1:2]$,*

$$I(W_i; Y_{i[\ell]}^{[N]} \mid \mathcal{G}_{[1:\ell]}) \leq I(W_i; \overline{Y}_{i[\ell]}^{[N]} \mid \mathcal{G}_{[1:\ell]}) + No(\log(P)). \tag{4.24}$$

Lemma 4.1 above is obtained from Lemma 1 of [16] as applied to our setting. Recall that in

[16] the same deterministic transformation that we apply to the ℓ^{th} hop, is applied to the one-hop MISO BC, and it is shown that this cannot reduce the mutual information in the GDoF sense between the messages and the corresponding deterministic outputs, conditioned on the one-hop channels for which only finite precision CSIT is available to the transmitters. There are two key distinctions in our setting. First, unlike the MISO BC where the transmitters cooperate fully, because the transmitters in the ℓ^{th} hop do not directly have access to the messages, their coding functions are more restricted. Second, because transmitters in the ℓ^{th} hop have knowledge of channel realizations of preceding hops, the coding functions may utilize this knowledge, which means that the transmitted symbols need not be independent of the $\mathcal{G}_{[1:\ell-1]}$ terms that are included in the conditioning in (4.24), unlike the one-hop MISO BC where the transmitted symbols are independent of the channels that appear in the conditioning. However, neither of these distinctions affects the validity of Lemma 4.1 because upon inspection of the proof of Lemma 1 of [16] it becomes evident that the proof holds for *all* feasible coding functions in the MISO BC, which includes the restricted class of coding functions available to the transmitters in the ℓ^{th} hop in the multi-hop setting. Furthermore, it turns out that the proof of Lemma 1 of [16] also holds under the additional conditioning on the channels $\mathcal{G}_{[1:\ell-1]}$ which are not necessarily independent of the transmitted symbols; what matters for the proof is that these additional conditioning terms are independent of the additive noise encountered by the receivers of the ℓ^{th} hop. Thus, the proof of Lemma 1 of [16] carries over to Lemma 4.1 in this chapter. For the sake of completeness, the proof is summarized in the Appendix section.

Remark: Note that because $\mathcal{G}_{[\ell+1:L]}$ are independent of all terms that appear in (4.24)*, the result of Lemma 4.1 can also be stated with additional conditioning on $\mathcal{G}_{[\ell+1:L]}$ as:*

$$I(W_i; Y_{i[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \leq I(W_i; \overline{Y}_{i[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + No(\log(P)). \tag{4.25}$$

Next let us recall two sum-set inequalities that will be critical to our converse proofs, as applied to our setting. The first sum-set inequality, namely Sum-set Inequality 1, originally shown in [18, Theorem 1], is used to bound the entropy difference of two received signals in the GDoF sense. Intuitively, this sum-set inequality says that in the GDoF sense the entropy difference is upper bounded by the maximum difference of the corresponding link strengths. The inequality applies to our setting because, as explained for Lemma 4.1, the original version in [18] is proved for the MISO broadcast channel which allows arbitrary coding functions, including the ones available to the transmitters in the ℓ^{th} hop.

Lemma 4.2. *(Sum-set Inequality 1) Let* $\overline{U}_{i[\ell]}^{[N]} = \lfloor G_{i1[\ell]}^{[N]}(\overline{X}_{1[\ell]}^{[N]})^{\mu_i} \rfloor + \lfloor G_{i2[\ell]}^{[N]}(\overline{X}_{2[\ell]}^{[N]})^{\nu_i} \rfloor$*, then for* $i, j \in \{1, 2\}, i \neq j$*,*

$$H(\overline{U}_{i[\ell]}^{[N]} \mid \mathcal{W}_S, \mathcal{G}_{[\ell]}) - H(\overline{U}_{j[\ell]}^{[N]} \mid \mathcal{W}_S, \mathcal{G}_{[\ell]}) \leq \max(\mu_i - \mu_j, \nu_i - \nu_j)^+ N \log(P) + No(\log(P)), \tag{4.26}$$

where $\mathcal{W}_S$ *is a set of random variables satisfying*

$$I(\overline{X}_{1[\ell]}^{[N]}, \overline{X}_{2[\ell]}^{[N]}, \mathcal{W}_S; \mathcal{G}_{[\ell]}) = 0. \tag{4.27}$$

The next sum-set inequality, Sum-set Inequality 2, appeared originally in a generalized form in [20, Theorem 4]. The following simplified form, taken from [66, Lemma 1] and specialized to our setting, is sufficient for our purpose.

Lemma 4.3. *(Sum-set Inequality 2) Let* $\overline{Y}_{[\ell]}(n) = \sum_{k=1}^{2} \lfloor G_{k[\ell]}(n) \overline{X}_{k[\ell]}(n) \rfloor$ *for* $\overline{X}_{k[\ell]}(n) \in \mathcal{X}_{\mu_{k[\ell]}}$*, and let* $G_{k[\ell]}(n)$ *be distinct elements of* $\mathcal{G}$ *for all* $k \in [2], n \in [1 : N]$*. For* $k \in [2]$*, let* S_k *be a set of finitely many disjoint sub-sections of* $\overline{X}_{k[\ell]}$ *(the same partitioning is applied to* $\overline{X}_{k[\ell]}(n)$ *for every* n*), and let* $\{U_1, U_2, \cdots, U_m\}$ *be a subset of* $S_1 \cup S_2$*. The following sum-set*

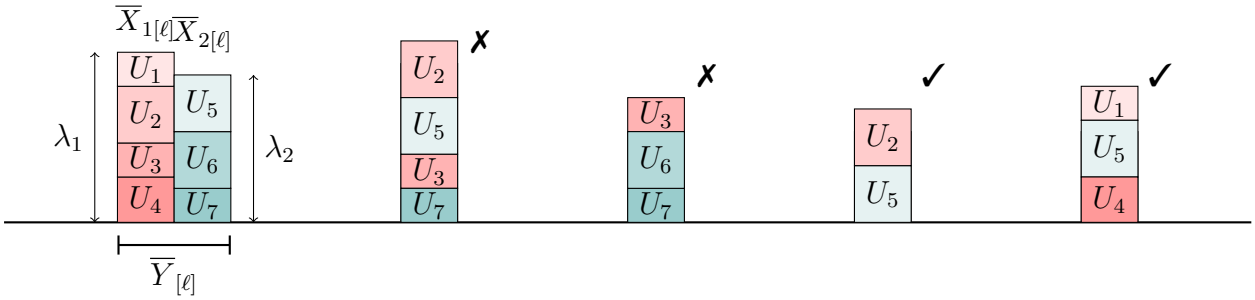

Figure 4.6: *An illustration of Lemma 4.3. Lemma 4.3 implies the sum-set inequalities* $H(\overline{Y}_{[\ell]}^{[N]} \mid W_S, \mathcal{G}_{[\ell]}) \geq H(U_2^{[N]}, U_5^{[N]} \mid W_S, \mathcal{G}_{[\ell]})$ *and* $H(\overline{Y}_{[\ell]}^{[N]} \mid W_S, \mathcal{G}_{[\ell]}) \geq H(U_1^{[N]}, U_4^{[N]}, U_5^{[N]} \mid W_S, \mathcal{G}_{[\ell]})$ *in the GDoF sense because the boxes in these inequalities can be vertically stacked without elevating any sub-section of them above its original height in* $\overline{Y}_{[\ell]}$*. However, Lemma 4.3 implies neither* $H(\overline{Y}_{[\ell]}^{[N]} \mid W_S, \mathcal{G}_{[\ell]}) \geq H(U_2^{[N]}, U_3^{[N]}, U_5^{[N]}, U_7^{[N]} \mid W_S, \mathcal{G}_{[\ell]})$ *nor* $H(\overline{Y}_{[\ell]}^{[N]} \mid W_S, \mathcal{G}_{[\ell]}) \geq H(U_3^{[N]}, U_6^{[N]}, U_7^{[N]} \mid W_S, \mathcal{G}_{[\ell]})$*, because it is impossible to vertically stack the corresponding boxes in any order without elevating at least one of them above its original position in* $\overline{Y}_{[\ell]}$*.*

inequality holds,

$$H(\overline{Y}_{[\ell]}^{[N]} \mid \mathcal{W}_S, \mathcal{G}_{[\ell]}) \geq H(U_1^{[N]}, U_2^{[N]}, \cdots, U_m^{[N]} \mid \mathcal{W}_S, \mathcal{G}_{[\ell]}) + No(\log(P)), \tag{4.28}$$

if both of the following conditions are satisfied.

$$I\left(\overline{X}_{1[\ell]}^{[N]}, \overline{X}_{2[\ell]}^{[N]}, \mathcal{W}_S \; ; \; \mathcal{G}_{[\ell]}\right) = 0, \tag{4.29}$$

$$\sum_{j=1}^{i-1} \mathcal{T}(U_j) \leq \ell(U_i), \qquad \forall i \in [2:m]. \tag{4.30}$$

Condition (4.30) can be visualized in terms of a vertical stacking of m boxes $U_1, \cdots, U_m$ in that order from bottom to top where the j^{th} box has height $\mathcal{T}(U_j)$. Conditions (4.30) simply mean that the height at which the i^{th} box appears in the vertical stacking (the LHS of (4.30)) should not be higher than its original level in $\overline{Y}_{[\ell]}$, i.e., $\ell(U_i)$. In other words, if there exists *any* ordering such that we can vertically stack all of the sub-sections without lifting up any one of them above its original height in $\overline{Y}_{[\ell]}$, then the sum-set inequality (4.28) holds. Figure 4.6 presents a few examples that satisfy or violate Lemma 4.3.

Next we present our main lemma that is developed in this chapter specifically for the multi-hop setting, to capture a recursive bounding argument which will allow us to use the deterministic bounds for each hop, one-hop at a time.

Lemma 4.4. *The following inequality holds for any $\ell \in [2:L]$, and any $\alpha \in [0,1]$.*

$$\begin{aligned} &NR_1 + NR_2 + 2I(W_1; Y_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + 2I(W_2; Y_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \\ &\leq (2 + 2\max(1-\alpha, \alpha))N\log(P) + I(W_1; Y_{1[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) \\ &\qquad + No(\log(P)). \end{aligned} \tag{4.31}$$

Corollary 4.1. *Let $\mathcal{D}_{\Sigma}^{f.p.}(n)$ represent the sum-GDoF for the n hop setting. Intuitively, one may loosely interpret the terms on the LHS of (4.31) as $\mathcal{D}_{\Sigma}^{f.p.}(L) + 2\mathcal{D}_{\Sigma}^{f.p.}(\ell)$ and the terms on the RHS as $2 + \mathcal{D}_{\Sigma}^{f.p.}(1) + \mathcal{D}_{\Sigma}^{f.p.}(\ell-1)$. Note that $\mathcal{D}_{\Sigma}^{f.p.}(1) = 2\max(1-\alpha,\alpha)$ in the regime $\alpha < 2/3$ where Lemma 4.4 will be primarily used.*

Proof. For compact notation, in this proof we will occasionally suppress $No(\log(P))$ terms that are inconsequential for GDoF. Starting with (4.25), we have

$$I(W_1; Y_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \leq I(W_1; \overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}), \tag{4.32}$$

$$I(W_2; Y_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \leq I(W_2; \overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}). \tag{4.33}$$

Next, we have the Markov chain,

$$(W_1, W_2) \leftrightarrow (Y_{1[\ell]}^{[N]}, Y_{2[\ell]}^{[N]}, \mathcal{G}_{[1:\ell]}) \leftrightarrow (Y_{1[L]}^{[N]}, Y_{2[L]}^{[N]}, \mathcal{G}_{[1:L]}). \tag{4.34}$$

Using this Markov chain and the data processing inequality, we proceed as follows.

$$NR_1 + NR_2 \leq I\left(W_1, W_2; Y_{1[L]}^{[N]}, Y_{2[L]}^{[N]}, \mathcal{G}_{[1:L]}\right) \tag{4.35}$$

$$\leq I\left(W_1, W_2; Y_{1[\ell]}^{[N]}, Y_{2[\ell]}^{[N]}, \mathcal{G}_{[1:\ell]}\right) \tag{4.36}$$

$$= I\left(W_1, W_2; Y_{1[\ell]}^{[N]}, Y_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:\ell]}\right) \tag{4.37}$$

$$\leq I(W_1, W_2; \overline{Y}_{1[\ell]}^{[N]}, \overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:\ell]}) \tag{4.38}$$

$$\leq I(W_1, W_2; \overline{X}_{1[\ell]}^{[N]}, \overline{X}_{2[\ell]}^{[N]}, \overline{Y}_{1[\ell]}^{[N]}, \overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:\ell]}) \tag{4.39}$$

$$= I(W_1, W_2; \overline{X}_{1[\ell]}^{[N]}, \overline{X}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:\ell]}) \tag{4.40}$$

$$\leq H(\overline{X}_{1[\ell]}^{[N]}, \overline{X}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:\ell]}) \tag{4.41}$$

$$= H(\overline{X}_{1[\ell]}^{[N]}, \overline{X}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \tag{4.42}$$

$$= H(\overline{X}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + H(\overline{X}_{2[\ell]}^{[N]} \mid \overline{X}_{1[\ell]}^{[N]}, \mathcal{G}_{[1:L]}). \tag{4.43}$$

Step (4.35) is obtained by Fano's inequality. Step (4.36) follows from the Markov chain in (4.34) and the data processing inequality. Step (4.37) uses the chain rule of mutual information and the fact that the messages are independent of the channels. Step (4.38) is obtained by reasoning similar to Lemma 4.1. Step (4.39) uses the property that $I(A; B \mid C) \leq I(A; B, D \mid C)$. Step (4.40) is because $(\overline{Y}_{1[\ell]}^{[N]}, \overline{Y}_{2[\ell]}^{[N]})$ is determined by $(\overline{X}_{1[\ell]}^{[N]}, \overline{X}_{2[\ell]}^{[N]}, \mathcal{G}_{[\ell]})$.[5] Step (4.41) uses the definition of mutual information $I(A; B \mid C) = H(B \mid C) - H(B \mid A, C)$ and the non-negativity of entropy in dropping the negative term. Including $\mathcal{G}_{[\ell+1:L]}$ in the conditioning in (4.42) is justified because these channels are independent of all the other terms that appear in the entropy expression. Step (4.43) is simply the chain rule of entropy. Next,

$$I(W_1; \overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) = H(\overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) - H(\overline{Y}_{1[\ell]}^{[N]} \mid W_1, \mathcal{G}_{[1:L]}) \tag{4.44}$$

$$\leq H(\overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) - H(\overline{X}_{1[\ell]}^{[N]} \mid W_1, \mathcal{G}_{[1:L]}) \tag{4.45}$$

$$= H(\overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(\overline{X}_{1[\ell]}^{[N]}; W_1 \mid \mathcal{G}_{[1:L]}) - H(\overline{X}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \tag{4.46}$$

$$\leq H(\overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(Y_{1[\ell-1]}^{[N]}; W_1 \mid \mathcal{G}_{[1:L]}) - H(\overline{X}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \tag{4.47}$$

$$\leq N \log(P) + I(W_1; Y_{1[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) - H(\overline{X}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}). \tag{4.48}$$

[5]Note that $\overline{X}_{i[\ell]}^{[N]}$ has power level $\max(1, \alpha) = 1$ in the weak interference regime ($\alpha \leq 1$). In other words, $\overline{X}_{i[\ell]}^{[N]}$ is equivalent to $(\overline{X}_{i[\ell]}^{[N]})^1$ in the weak interference regime.

Notably in (4.45) we used Sum-set Inequality 1 from Lemma 4.2 as follows. Since $H(\overline{Y}_{1[\ell]}^{[N]} \mid W_1, \mathcal{G}_{[1:L]}) = H(\lfloor G_{11[\ell]}^{[N]}(\overline{X}_{1[\ell]}^{[N]})^1 \rfloor + \lfloor G_{12[\ell]}^{[N]}(\overline{X}_{2[\ell]}^{[N]})^\alpha \rfloor \mid W_1, \mathcal{G}_{[1:L]})$ if we set $\mathcal{W}_S = (W_1, \mathcal{G}_{[1:L]\setminus[\ell]})$, then from Lemma 4.2 we obtain,

$$
\begin{aligned}
No(\log(P)) &= \max(1-1, 0-\alpha)^+ N\log(P) + No(\log(P)) && (4.49)\\
&\geq H(\lfloor G_{21[\ell]}^{[N]}(\overline{X}_{1[\ell]}^{[N]})^1 \rfloor + \lfloor G_{22[\ell]}^{[N]}(\overline{X}_{2[\ell]}^{[N]})^0 \rfloor \mid W_1, \mathcal{G}_{[1:L]}) \\
&\quad - H(\lfloor G_{11[\ell]}^{[N]}(\overline{X}_{1[\ell]}^{[N]})^1 \rfloor + \lfloor G_{12[\ell]}^{[N]}(\overline{X}_{2[\ell]}^{[N]})^\alpha \rfloor \mid W_1, \mathcal{G}_{[1:L]}) && (4.50)\\
&\geq H(\overline{X}_{1[\ell]}^{[N]} \mid W_1, \mathcal{G}_{[1:L]}) - H(\overline{Y}_{1[\ell]}^{[N]} \mid W_1, \mathcal{G}_{[1:L]}) + No(\log(P)). && (4.51)
\end{aligned}
$$

Step (4.47) holds because the output of the relay node is a function of its input signal and the channels of the preceding hops. Specifically, $\overline{X}_{1[\ell]}^{[N]}$ is a function of $X_{1[\ell]}^{[N]}$ according to (4.18); and $X_{1[\ell]}^{[N]}$ is a function of $(Y_{1[\ell-1]}^{[N]}, \mathcal{G}_{[1:\ell-1]})$ according to the relay mapping function, so it is also a function of $(Y_{1[\ell-1]}^{[N]}, \mathcal{G}_{[1:L]})$. Step (4.48) follows from (4.21) which uses the fact that a uniform distribution maximizes entropy over a discrete alphabet of bounded cardinality. By symmetry, it follows from (4.48), that we must also have,

$$
I(W_2; \overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \leq N\log(P) + I(W_2; Y_{2[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) - H(\overline{X}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \tag{4.52}
$$

Adding (4.43),(4.48),(4.52) together, we have

$$
\begin{aligned}
&NR_1 + NR_2 + I(W_1; \overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; \overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \\
&\quad \leq 2N\log(P) + I(W_1; Y_{1[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) \\
&\qquad\qquad + H(\overline{X}_{2[\ell]}^{[N]} \mid \overline{X}_{1[\ell]}^{[N]}, \mathcal{G}_{[1:L]}) - H(\overline{X}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) && (4.53)\\
&\quad = 2N\log(P) + I(W_1; Y_{1[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) - I(\overline{X}_{1[\ell]}^{[N]}; \overline{X}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) && (4.54)\\
&\quad \leq 2N\log(P) + I(W_1; Y_{1[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) \\
&\qquad\qquad - I((\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)}; (\overline{X}_{2[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L]}). && (4.55)
\end{aligned}
$$

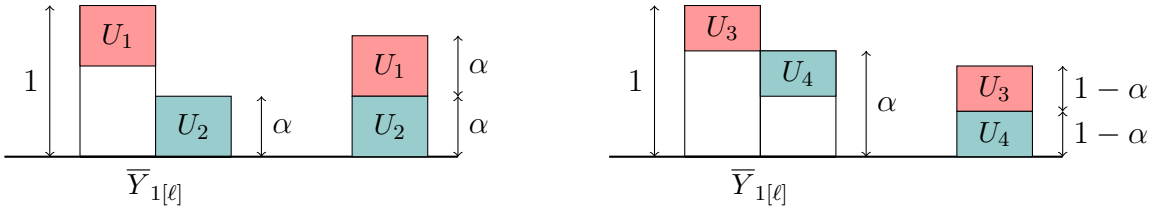

Figure 4.7: *In the left figure,* $\alpha \leq \frac{1}{2}, U_1 = (\overline{X}_{1[\ell]})^{\alpha}, U_2 = (\overline{X}_{2[\ell]})^{\alpha}$*, evidently* U_1, U_2 *can be stacked without elevating either one of them above the level at which it appears in* $\overline{Y}_{1[\ell]}$*. On the right,* $\alpha \geq \frac{1}{2}$*,* $U_3 = (\overline{X}_{1[\ell]})^{1-\alpha}$*,* $U_4 = (\overline{X}_{2[\ell]})^{1-\alpha}$*;* U_3, U_4 *can also be stacked without elevating either of them.*

The definition of mutual information was used to obtain (4.54), and in (4.55) we used the property that $I(A;B) \geq I(f(A);g(B))$ for any functions f, g. Next, we bound $I(W_1; \overline{Y}^{[N]}_{1[\ell]} \mid \mathcal{G}_{[1:L]})$ as follows.

$$I(W_1; \overline{Y}^{[N]}_{1[\ell]} \mid \mathcal{G}_{[1:L]}) = H(\overline{Y}^{[N]}_{1[\ell]} \mid \mathcal{G}_{[1:L]}) - H(\overline{Y}^{[N]}_{1[\ell]} \mid W_1, \mathcal{G}_{[1:L]}) \tag{4.56}$$

$$\leq H(\overline{Y}^{[N]}_{1[\ell]} \mid \mathcal{G}_{[1:L]}) - H((\overline{X}^{[N]}_{1[\ell]})^{\min(\alpha,1-\alpha)}, (\overline{X}^{[N]}_{2[\ell]})^{\min(\alpha,1-\alpha)} \mid W_1, \mathcal{G}_{[1:L]}). \tag{4.57}$$

This step is significant, because it invokes Sum-set Inequality 2 from Lemma 4.3, noting that $(\overline{X}^{[N]}_{1[\ell]})^{\min(1-\alpha,\alpha)}, (\overline{X}^{[N]}_{2[\ell]})^{\min(1-\alpha,\alpha)}$ can be stacked vertically without elevating either of them above their original height in $\overline{Y}^{[N]}_{1[\ell]}$. See Figure 4.7 for an illustration of the stacking.

Similarly,

$$I(W_2; \overline{Y}^{[N]}_{2[\ell]} \mid \mathcal{G}_{[1:L]}) \leq H(\overline{Y}^{[N]}_{2[\ell]} \mid \mathcal{G}_{[1:L]}) - H((\overline{X}^{[N]}_{1[\ell]})^{\min(\alpha,1-\alpha)}, (\overline{X}^{[N]}_{2[\ell]})^{\min(\alpha,1-\alpha)} \mid W_2, \mathcal{G}_{[1:L]}). \tag{4.58}$$

Adding (4.57) and (4.58), we get

$$\begin{aligned} &I(W_1; \overline{Y}^{[N]}_{1[\ell]} \mid \mathcal{G}_{[1:L]}) + I(W_2; \overline{Y}^{[N]}_{2[\ell]} \mid \mathcal{G}_{[1:L]}) \\ &\leq H(\overline{Y}^{[N]}_{1[\ell]} \mid \mathcal{G}_{[1:L]}) + H(\overline{Y}^{[N]}_{2[\ell]} \mid \mathcal{G}_{[1:L]}) - H((\overline{X}^{[N]}_{1[\ell]})^{\min(\alpha,1-\alpha)}, (\overline{X}^{[N]}_{2[\ell]})^{\min(\alpha,1-\alpha)} \mid W_1, \mathcal{G}_{[1:L]}) \\ &\qquad - H((\overline{X}^{[N]}_{1[\ell]})^{\min(\alpha,1-\alpha)}, (\overline{X}^{[N]}_{2[\ell]})^{\min(\alpha,1-\alpha)} \mid W_2, \mathcal{G}_{[1:L]}) \end{aligned} \tag{4.59}$$

$$\leq H(\overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + H(\overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) - H((\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)}, (\overline{X}_{2[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L]})$$
$$- H((\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)}, (\overline{X}_{2[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid W_1, W_2, \mathcal{G}_{[1:L]}) \tag{4.60}$$
$$\leq H(\overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + H(\overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) - H((\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)}, (\overline{X}_{2[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L]}) \tag{4.61}$$

where (4.60) follows from the property that for any three random variables A, B, C, if B, C are independent, then

$$H(A \mid B) + H(A \mid C) \geq H(A) + H(A \mid B, C) \tag{4.62}$$

and (4.61) simply uses the fact that entropy is non-negative. Adding $2 \times (4.32) + 2 \times (4.33) + (4.55) + (4.61)$, we obtain,

$$NR_1 + NR_2 + 2I(W_1; Y_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + 2I(W_2; Y_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]})$$
$$\leq 2N \log(P) + I(W_1; Y_{1[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + H(\overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]})$$
$$+ H(\overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) - I((\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)}; (\overline{X}_{2[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L]})$$
$$- H((\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)}, (\overline{X}_{2[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L]}) \tag{4.63}$$
$$= 2N \log(P) + I(W_1; Y_{1[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + H(\overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]})$$
$$+ H(\overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) - H((\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L]})$$
$$- H((\overline{X}_{2[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L]}) \tag{4.64}$$
$$= 2N \log(P) + I(W_1; Y_{1[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]})$$
$$+ \{H(\overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) - H((\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L]})\}$$
$$+ \{H(\overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) - H((\overline{X}_{2[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L)})\} \tag{4.65}$$
$$\leq 2N \log(P) + I(W_1; Y_{1[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]})$$
$$+ 2 \max(\alpha, 1-\alpha) N \log(P) \tag{4.66}$$

$$= (2 + 2\max(\alpha, 1-\alpha))N\log(P) + I(W_1; Y_{1[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[\ell-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) \tag{4.67}$$

where (4.64) follows from the definition of mutual information, and (4.65) is simply a rearrangement of terms. Step (4.66) is significant because it invokes Sum-set Inequality 1 from Lemma 4.2 as follows.

$$\begin{aligned}
&H(\overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) - H((\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L]}) && (4.68)\\
&\quad = H\left(\left\lfloor G_{11[\ell]}^{[N]}(\overline{X}_{1[\ell]}^{[N]})^1 \right\rfloor + \left\lfloor G_{12[\ell]}^{[N]}(\overline{X}_{2[\ell]}^{[N]})^\alpha \right\rfloor \middle| \mathcal{G}_{[1:L]}\right) - H((\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \mid \mathcal{G}_{[1:L]})\\
&\quad \le H\left(\left\lfloor G_{11[\ell]}^{[N]}(\overline{X}_{1[\ell]}^{[N]})^1 \right\rfloor + \left\lfloor G_{12[\ell]}^{[N]}(\overline{X}_{2[\ell]}^{[N]})^\alpha \right\rfloor \middle| \mathcal{G}_{[1:L]}\right)\\
&\qquad\quad - H\left(\left\lfloor G_{21[\ell]}^{[N]}(\overline{X}_{1[\ell]}^{[N]})^{\min(\alpha,1-\alpha)} \right\rfloor + \left\lfloor G_{22[\ell]}^{[N]}(\overline{X}_{2[\ell]}^{[N]})^0 \right\rfloor \middle| \mathcal{G}_{[1:L]}\right) && (4.69)\\
&\quad \le \max\left(1 - \min(\alpha, 1-\alpha), \alpha - 0\right)^+ N\log(P) && (4.70)\\
&\quad = \max\left(\alpha, 1-\alpha\right) N\log(P) && (4.71)
\end{aligned}$$

Note that (4.67) matches the RHS of (4.31), so that the proof of Lemma 4.4 is complete. With the help of these lemmas, we are now ready to present the converse proof of Theorem 4.1.

4.4.3 Converse Proof for Theorem 4.1

As noted previously, the upper bound $\mathcal{D}_\Sigma \le 2 - \alpha$ in the regime $\frac{4}{7} \le \alpha \le 1$ is immediate, because it corresponds to the sum-GDoF of the MISO broadcast channel [18] formed at the second hop by allowing full cooperation among the relays, which cannot decrease the GDoF. Therefore, we will assume $\alpha \le \frac{4}{7}$ in the following proof. As usual, we will sometimes suppress the $o(\log(P))$ terms for simplicity as they are inconsequential for GDoF studies. Starting

with Fano's inequality, we have

$$
\begin{aligned}
&3NR_1 + 3NR_2 \\
&\quad \le 2I(W_1; Y_{1[2]}^{[N]} \mid \mathcal{G}_{[1:2]}) + 2I(W_2; Y_{2[2]}^{[N]} \mid \mathcal{G}_{[1:2]}) + NR_1 + NR_2 && (4.72)\\
&\quad \le \big(2 + 2\max(1-\alpha, \alpha)\big) N \log(P) + I(W_1; Y_{1[1]}^{[N]} \mid \mathcal{G}_{[1:2]}) + I(W_2; Y_{2[1]}^{[N]} \mid \mathcal{G}_{[1:2]}) && (4.73)\\
&\quad \le \big(2 + 2\max(1-\alpha, \alpha) + 2\max(1-\alpha, \alpha)\big) N \log(P) && (4.74)\\
&\quad = \big(2 + 4\max(1-\alpha, \alpha)\big) N \log(P). && (4.75)
\end{aligned}
$$

where (4.73) is obtained from Lemma 4.4 by setting $\ell = 2$, and (4.74) is essentially the well-known sum-GDoF bound for the single-hop interference channel (corresponding to the first hop) that is obtained in [24, Section III. D] when $\alpha \le \frac{2}{3}$ by using a genie-aided approach. Since $\frac{4}{7} \le \frac{2}{3}$, the bound holds for $\alpha \le \frac{4}{7}$. Normalizing both sides by $3N\log(P)$ and applying the GDoF limit ($P \to \infty$), we obtain the sum-GDoF bound $\mathcal{D}_{\Sigma}^{f.p.} \le \frac{2+4\max(1-\alpha,\alpha)}{3}$. Therefore, when $\alpha \le \frac{1}{2}$, we get the bound $\mathcal{D}_{\Sigma}^{f.p.} \le \frac{6-4\alpha}{3}$, and when $\frac{1}{2} \le \alpha \le \frac{4}{7}$, we obtain the bound $\mathcal{D}_{\Sigma}^{f.p.} \le \frac{2+4\alpha}{3}$.

4.4.4 Converse Proof for Theorem 4.2

In the symmetric multi-hop channel, the bound $2 - \alpha$ for the regime $\frac{2^L}{2^{L+1}-1} \le \alpha \le 1$ is also trivial, because it is the sum-GDoF of the broadcast channel [18] formed in the last hop under finite precision CSIT by allowing full cooperation among the relays which cannot reduce the GDoF. Next, to derive the upper bounds for $\alpha \le \frac{2^L}{2^{L+1}-1}$, we need to recursively apply Lemma 4.4. Starting with Fano's inequality, we have

$$
N(R_1 + R_2)
$$

$$\leq \frac{2I(W_1; Y_{1[L]}^{[N]} \mid \mathcal{G}_{[1:L]}) + 2I(W_2; Y_{2[L]}^{[N]} \mid \mathcal{G}_{[1:L]}) + NR_1 + NR_2}{3} \tag{4.76}$$

$$\leq \frac{(2 + 2\max(1-\alpha, \alpha))N\log(P)}{3} + \frac{I(W_1; Y_{1[L-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[L-1]}^{[N]} \mid \mathcal{G}_{[1:L]})}{3} \tag{4.77}$$

$$= \frac{(2 + 2\max(1-\alpha, \alpha))N\log(P)}{3} + \frac{2I(W_1; Y_{1[L-1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + 2I(W_2; Y_{2[L-1]}^{[N]} \mid \mathcal{G}_{[1:L]})}{6} \tag{4.78}$$

$$\leq \frac{(2 + 2\max(1-\alpha, \alpha))N\log(P)}{3} + \frac{(2 + 2\max(1-\alpha, \alpha))N\log(P)}{6} + \frac{I(W_1; Y_{1[L-2]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[L-2]}^{[N]} \mid \mathcal{G}_{[1:L]})}{6} - \frac{N(R_1+R_2)}{6} \tag{4.79}$$

$$\leq \cdots$$

$$\leq (2 + 2\max(1-\alpha, \alpha)) \left(\sum_{m=0}^{M-1} \frac{1}{3 \times 2^m} \right) N\log(P) + \frac{I(W_1; Y_{1[L-M]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[L-M]}^{[N]} \mid \mathcal{G}_{[1:L]})}{3 \times 2^{M-1}} - \left(\sum_{k=1}^{M-1} \frac{1}{3 \times 2^k} \right) N(R_1+R_2) \tag{4.80}$$

$$\leq \cdots$$

$$\leq (2 + 2\max(1-\alpha, \alpha)) \left(\sum_{m=0}^{L-3} \frac{1}{3 \times 2^m} \right) N\log(P) + \frac{I(W_1; Y_{1[2]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[2]}^{[N]} \mid \mathcal{G}_{[1:L]})}{3 \times 2^{L-3}} - \left(\sum_{k=1}^{L-3} \frac{1}{3 \times 2^k} \right) N(R_1+R_2) \tag{4.81}$$

$$\leq (2 + 2\max(1-\alpha, \alpha)) \left(\sum_{m=0}^{L-2} \frac{1}{3 \times 2^m} \right) N\log(P) + \frac{I(W_1; Y_{1[1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_2; Y_{2[1]}^{[N]} \mid \mathcal{G}_{[1:L]})}{3 \times 2^{L-2}} - \left(\sum_{k=1}^{L-2} \frac{1}{3 \times 2^k} \right) N(R_1+R_2) \tag{4.82}$$

$$\leq (2+2\max(1-\alpha,\alpha))\left(\sum_{m=0}^{L-2}\frac{1}{3\times 2^m}\right)N\log(P)$$
$$+\frac{I(X_{1[1]}^{[N]};Y_{1[1]}^{[N]}\mid\mathcal{G}_{[1:L]})+I(X_{2[1]}^{[N]};Y_{2[1]}^{[N]}\mid\mathcal{G}_{[1:L]})}{3\times 2^{L-2}}-\left(\sum_{k=1}^{L-2}\frac{1}{3\times 2^k}\right)N(R_1+R_2) \tag{4.83}$$

where (4.77), (4.79), (4.80), (4.82) are obtained by setting $\ell = L, L-1, 3, 2$ in Lemma 4.4, respectively. For (4.83) we used the data processing inequality, the Markov chains $W_1 \leftrightarrow X_{1[1]}^{[N]} \leftrightarrow (Y_{1[1]}^{[N]}, \mathcal{G}_{[1:L]})$, $W_2 \leftrightarrow X_{2[1]}^{[N]} \leftrightarrow (Y_{2[1]}^{[N]}, \mathcal{G}_{[1:L]})$ and the independence of channels $\mathcal{G}_{[1:L]}$ with the messages and codewords sent over the first hop. Rearranging (4.83), we obtain

$$\left(1+\frac{2^{L-2}-1}{3\times 2^{L-2}}\right)N(R_1+R_2)\leq(2+2\max(1-\alpha,\alpha))\left(\frac{2^{L-1}-1}{3\times 2^{L-2}}\right)N\log(P)$$
$$+\frac{I(X_{1[1]}^{[N]};Y_{1[1]}^{[N]}\mid\mathcal{G}_{[1:L]})+I(X_{2[1]}^{[N]};Y_{2[1]}^{[N]}\mid\mathcal{G}_{[1:L]})}{3\times 2^{L-2}}. \tag{4.84}$$

With the inequality $I(X_{1[1]}^{[N]};Y_{1[1]}^{[N]}\mid\mathcal{G}_{[1:L]})+I(X_{2[1]}^{[N]};Y_{2[1]}^{[N]}\mid\mathcal{G}_{[1:L]})\leq(2\max(1-\alpha,\alpha))N\log(P)$ obtained in [24, Section III. D][6] when $\alpha\leq\frac{2}{3}$, we have,

$$\left(1+\frac{2^{L-2}-1}{3\times 2^{L-2}}\right)N(R_1+R_2)\leq(2+2\max(1-\alpha,\alpha))\left(\frac{2^{L-1}-1}{3\times 2^{L-2}}\right)N\log(P)$$
$$+\frac{2\max(1-\alpha,\alpha)}{3\times 2^{L-2}}N\log(P). \tag{4.85}$$

Dividing by $(1+\frac{2^{L-2}-1}{3\times 2^{L-2}})$, we get

$$N(R_1+R_2)\leq\frac{(1+\max(\alpha,1-\alpha))\times(2^L-2)+2\max(\alpha,1-\alpha)}{2^L-1}N\log(P) \tag{4.86}$$
$$=\left(1+\max(\alpha,1-\alpha)-\frac{1-\max(\alpha,1-\alpha)}{2^L-1}\right)N\log(P). \tag{4.87}$$

[6]The original proof assumes perfect CSIT, but since the availability of CSIT cannot hurt, such bounds also hold with finite precision CSIT.

Normalizing both sides by $N\log(P)$ and applying the GDoF limit, we obtain the bound,

$$\mathcal{D}_{\Sigma}^{f.p.} \leq 1 + \max(\alpha, 1-\alpha) - \frac{1-\max(\alpha,1-\alpha)}{2^L - 1}. \tag{4.88}$$

Thus, when $\alpha \leq \frac{1}{2}$, we have $\mathcal{D}_{\Sigma}^{f.p.} \leq 2 - \alpha - \frac{\alpha}{2^L-1}$, when $\frac{1}{2} \leq \alpha \leq \frac{2^L}{2^{L+1}-1}$, we have $\mathcal{D}_{\Sigma}^{f.p.} \leq 1 + \alpha - \frac{1-\alpha}{2^L-1}$.

4.4.5 Converse Proof for Theorem 4.3

The upper bound is straightforward by using the cut-set bound and considering the multi-hop channel as a two-unicast channel.

$$\begin{aligned}
NR_1 &\leq I(X_{1[1]}^{[N]}; Y_{1[L]}^{[N]} \mid \mathcal{G}_{[1:L]}) && (4.89)\\
&\leq I(X_{1[1]}^{[N]}; Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, Y_{1[3]}^{[N]}, Y_{2[4]}^{[N]}, \cdots, Y_{1[L]}^{[N]} \mid \mathcal{G}_{[1:L]}) && (4.90)\\
&= I(X_{1[1]}^{[N]}; Y_{1[1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(X_{1[1]}^{[N]}; Y_{2[2]}^{[N]} \mid Y_{1[1]}^{[N]}, \mathcal{G}_{[1:L]})\\
&\qquad + I(X_{1[1]}^{[N]}; Y_{1[3]}^{[N]} \mid Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \mathcal{G}_{[1:L]}) + \cdots\\
&\qquad + I(X_{1[1]}^{[N]}; Y_{1[L]}^{[N]} \mid Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \cdots, Y_{2[L-1]}^{[N]}, \mathcal{G}_{[1:L]}) && (4.91)\\
&\leq I(X_{1[1]}^{[N]}; Y_{1[1]}^{[N]} \mid \mathcal{G}_{[1:L]}) + h(Y_{2[2]}^{[N]} \mid Y_{1[1]}^{[N]}, \mathcal{G}_{[1:L]}) + h(Y_{1[3]}^{[N]} \mid Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \mathcal{G}_{[1:L]}) + \cdots\\
&\qquad + h(Y_{1[L]}^{[N]} \mid Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \cdots, Y_{2[L-1]}^{[N]}, \mathcal{G}_{[1:L]}) && (4.92)\\
&= N\log(P) + h(Y_{2[2]}^{[N]} \mid X_{1[2]}^{[N]}, Y_{1[1]}^{[N]}, \mathcal{G}_{[1:L]}) + h(Y_{1[3]}^{[N]} \mid X_{2[3]}^{[N]}, Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \mathcal{G}_{[1:L]}) + \cdots\\
&\qquad + h(Y_{1[L]}^{[N]} \mid X_{2[L]}^{[N]}, Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \cdots, Y_{2[L-1]}^{[N]}, \mathcal{G}_{[1:L]}) && (4.93)\\
&\leq L \times N\log(P). && (4.94)
\end{aligned}$$

The negative differential entropy terms in (4.92) disappear because upon further conditioning (which cannot increase differential entropy) they reduce to noise terms that are inconsequen-

tial in the GDoF sense:

$$
\begin{aligned}
&h(Y_{i[\ell]}^{[N]} | Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \cdots, Y_{\bar{i}[\ell-1]}^{[N]}, \mathcal{G}_{[1:L]}) \\
&\quad \geq h(Y_{i[\ell]}^{[N]} \mid X_{1[\ell]}^{[N]}, X_{2[\ell]}^{[N]}, Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \cdots, Y_{\bar{i}[\ell-1]}^{[N]}, \mathcal{G}_{[1:L]}) && (4.95) \\
&\quad = h(Z_{i[\ell]}^{[N]} \mid X_{1[\ell]}^{[N]}, X_{2[\ell]}^{[N]}, Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \cdots, Y_{\bar{i}[\ell-1]}^{[N]}, \mathcal{G}_{[1:L]}) && (4.96) \\
&\quad = h(Z_{i[\ell]}^{[N]}) && (4.97) \\
&\quad = No(\log(P)) && (4.98)
\end{aligned}
$$

Note that for (4.95) we used the fact that conditioning cannot increase differential entropy, and (4.97) holds because the noise term is independent of all the conditioning terms.

Next, we note that (4.93) holds since $X_{i[\ell]}^{[N]}$ is determined by $(Y_{i[\ell-1]}^{[N]}, \mathcal{G}_{[1:L]})$. The bound in (4.94) holds because at ℓ^{th} hop, we have

$$
\begin{aligned}
&h(Y_{i[\ell]}^{[N]} \mid X_{\bar{i}[\ell]}^{[N]}, Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \cdots, Y_{\bar{i}[\ell-1]}^{[N]}, \mathcal{G}_{[1:L]}) \\
&\quad = h(\sqrt{P^1} G_{ii[\ell]}^{[N]} X_{i[\ell]}^{[N]} + \sqrt{P^\alpha} G_{i\bar{i}[\ell]}^{[N]} X_{\bar{i}[\ell]}^{[N]} + Z_{i[\ell]}^{[N]} \mid X_{\bar{i}[\ell]}^{[N]}, Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \cdots, Y_{\bar{i}[\ell-1]}^{[N]}, \mathcal{G}_{[1:L]}) && (4.99) \\
&\quad = h(\sqrt{P^1} X_{i[\ell]}^{[N]} + Z_{i[\ell]}^{[N]} \mid X_{\bar{i}[\ell]}^{[N]}, Y_{1[1]}^{[N]}, Y_{2[2]}^{[N]}, \cdots, Y_{\bar{i}[\ell-1]}^{[N]}, \mathcal{G}_{[1:L]}) && (4.100) \\
&\quad \leq h(\sqrt{P^1} X_{i[\ell]}^{[N]} + Z_{i[\ell]}^{[N]}) && (4.101) \\
&\quad \leq N \log(P). && (4.102)
\end{aligned}
$$

Symmetrically, we get

$$
NR_2 \leq L \times N \log(P) + No(\log(P)). \tag{4.103}
$$

Combining (4.94)(4.103) and dividing by $N \log(P)$ at both sides, we get $\mathcal{D}_\Sigma^p \leq 2L$. Since perfect CSIT cannot hurt the GDoF value, we also get $\mathcal{D}_\Sigma^{f.p.} \leq 2L$.

4.5 Achievability

4.5.1 Proof of Achievability for Theorem 4.1

There are 4 sub-cases for the achievable scheme.

- $\alpha \leq \frac{1}{2}$

  The achievable scheme is as follows: message W_i is split into 4 sub-messages $W_i = (W_{i1}, W_{i2}, W_{i3}, W_{i4})$. They carry $d_{i1} = \frac{2\alpha}{3}, d_{i2} = \frac{\alpha}{3}, d_{i3} = 1 - 2\alpha, d_{i4} = \frac{\alpha}{3}$ GDoF respectively. They are encoded into independent Gaussian codebooks producing codewords $X_{i1}, X_{i2}, X_{i3}, X_{i4}$ with powers $\mathrm{E}|X_{i1}|^2 = 1 - P^{-d_{i1}}, \mathrm{E}|X_{i2}|^2 = P^{-d_{i1}} - P^{-d_{i1}-d_{i2}}, \mathrm{E}|X_{i3}|^2 = P^{-d_{i1}-d_{i2}} - P^{-1+d_{i4}}, \mathrm{E}|X_{i4}|^2 = P^{-1+d_{i4}}$. The transmitted signals at the sources are $X_{1[1]} = X_{11} + X_{12} + X_{13} + X_{14}, X_{2[1]} = X_{21} + X_{22} + X_{23} + X_{24}$. Relay $\mathrm{Rx}_{1[1]}$ is able to decode $W_{11}, W_{12}, W_{13}, W_{21}$ successively with the corresponding SINR values $\sim P^{d_{11}}, \sim P^{d_{12}}, \sim P^{d_{13}}, \sim P^{d_{21}}$. Relay $\mathrm{Rx}_{1[1]}$ then reconstructs the codewords $X_{11}, X_{12}, X_{13}, X_{21}$ and removes their contribution from its received signal. The remaining signal above the noise floor is a linear combination of X_{14} and X_{22}, which is denoted as $\mathcal{L}_1(X_{14}, X_{22})$. Relay $\mathrm{Rx}_{1[1]}$ amplifies the remaining signals by power P^{-1}, such that the power is $\mathrm{E}|\mathcal{L}_1|^2 = P^{-1+d_{14}}$. Then, Relay $\mathrm{Rx}_{1[1]}$ splits the messages W_{11}, W_{21} into $W_{11} = (W_{11}^1, W_{11}^2), W_{21} = (W_{21}^1, W_{21}^2)$, with the corresponding GDoF value $d_{11}^1 = d_{11}^2 = d_{21}^1 = d_{21}^2 = \frac{d_{11}}{2} = \frac{\alpha}{3}$. After that, Relay $\mathrm{Tx}_{1[2]} \equiv \mathrm{Rx}_{1[1]}$ re-encodes $W_{11}^1, W_{12}^1, W_{12}, W_{13}, W_{21}^1$ into codewords $X_{11}^1, X_{12}^1, X_{12}, X_{13}, X_{21}^1$ by assigning power $\mathrm{E}|X_{12}|^2 = 1 - P^{-d_{12}}, \mathrm{E}|X_{11}^1|^2 = P^{-d_{12}} - P^{-d_{12}-d_{11}^1}, \mathrm{E}|X_{21}^1|^2 = P^{-d_{12}-d_{11}^1} - P^{-d_{12}-d_{11}^1-d_{21}^1}, \mathrm{E}|X_{21}|^2 = P^{-d_{12}-d_{11}^1-d_{21}^1} - P^{-d_{12}-d_{11}^1-d_{21}^1-d_{21}}, \mathrm{E}|X_{13}|^2 = P^{-d_{12}-d_{11}^1-d_{21}^1-d_{21}} - P^{-1+d_{14}+d_{11}^2}, \mathrm{E}|X_{11}^2|^2 = P^{-1+d_{14}+d_{11}^2} - P^{-1+d_{14}}$. Relay $\mathrm{Tx}_{2[2]} \equiv \mathrm{Rx}_{2[1]}$ proceeds similarly. The transmitted signals at relays are $X_{1[2]} = X_{12} + X_{11}^1 + X_{21}^1 + X_{13} + X_{11}^2 + \mathcal{L}_1, X_{2[2]} = X_{22} + X_{11}^1 + X_{21}^1 + X_{23} + X_{21}^2 + \mathcal{L}_2$. Destination $\mathrm{Rx}_{1[2]}$ is able to decode $W_{12}, W_{11}^1, W_{21}^1, W_{13}, W_{22}, W_{11}^2, W_{14}$ successively, with the corresponding SINR values $\sim$

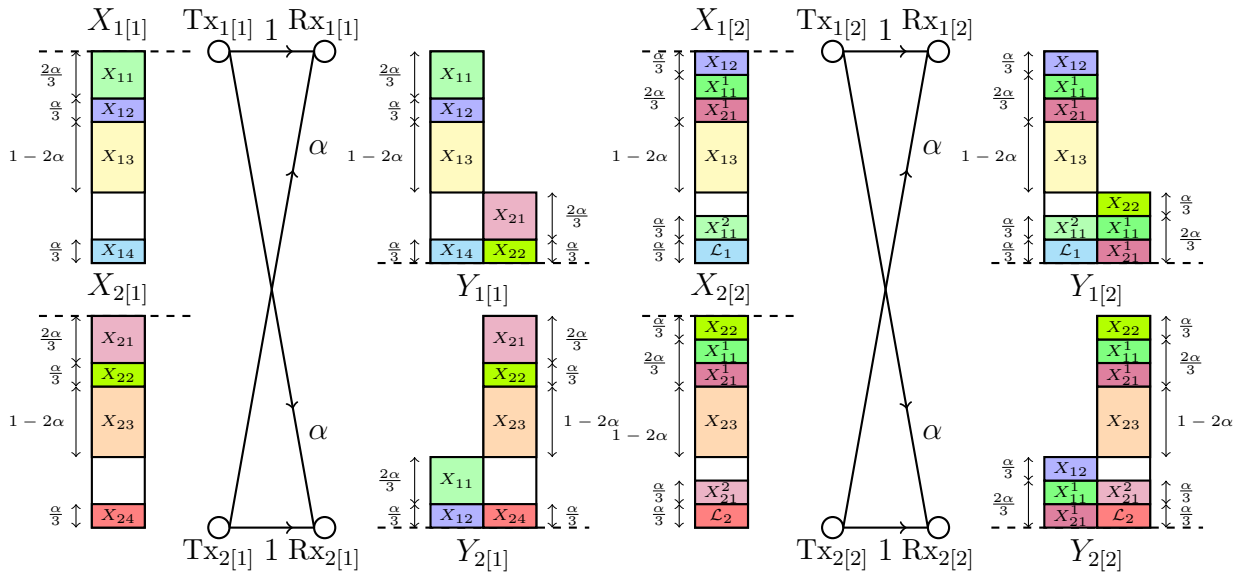

Figure 4.8: *Achievable scheme for $\alpha \leq \frac{1}{2}$. The dashed line at the receivers represents the noise floor, at the transmitters it represents unit power. $\mathcal{L}_1, \mathcal{L}_2$ are short for $\mathcal{L}_1(X_{14}, X_{22}), \mathcal{L}_2(X_{24}, X_{12})$, respectively. The left figure is the first hop while the right figure represents the second hop.*

$P^{d_{12}}, \sim P^{d^1_{11}}, \sim P^{d^1_{21}}, \sim P^{d_{13}}, \sim P^{d_{22}}, \sim P^{d^2_{11}}, \sim P^{d_{14}}$. Destination $\mathrm{Rx}_{2[2]}$ proceeds similarly to decode $W_{22}, W^1_{11}, W^1_{21}, W_{23}, W_{12}, W^2_{21}, W_{24}$ successively. See Figure 4.8 for an illustration.

- $\frac{1}{2} \leq \alpha \leq \frac{4}{7}$

  The achievable scheme is as follows: W_i is split into 3 sub-messages, i.e., $W_i = (W_{i1}, W_{i2}, W_{i3})$, whose corresponding GDoF are $d_{i1} = \frac{2-2\alpha}{3}, d_{i2} = \frac{1-\alpha}{3}, d_{i3} = \frac{5\alpha-2}{3}$. They are encoded into independent Gaussian codebooks producing codewords X_{i1}, X_{i2}, X_{i3} with assigned power $\mathrm{E}|X_{i1}|^2 = 1 - P^{-d_{i1}}, \mathrm{E}|X_{i2}|^2 = P^{-d_{i1}} - P^{-1+d_{i3}}, \mathrm{E}|X_{i3}|^2 = P^{-1+d_{i3}}$. The transmitted signals at the sources are $X_{1[1]} = X_{11} + X_{12} + X_{13}, X_{2[1]} = X_{21} + X_{22} + X_{23}$. Relay $\mathrm{Rx}_{1[1]}$ is able to decode W_{11}, W_{12}, W_{21} successively by treating everything else as noise. Then Relay $\mathrm{Rx}_{1[1]}$ reconstructs and subtracts the codewords X_{11}, X_{12}, X_{21} from its received signal. The remaining signal above the noise floor at Relay $\mathrm{Rx}_{1[1]}$ is a linear combination of X_{13}, X_{22}, which we denote as $\mathcal{L}_1$. Relay $\mathrm{Rx}_{1[1]}$ amplifies the remaining signals such that they carry power $\mathrm{E}|\mathcal{L}_1|^2 = P^{-1+d_{13}}$. Then, the relay $\mathrm{Rx}_{1[1]}$ splits W_{11}, W_{21} into $W_{11} = (W^1_{11}, W^2_{11}), W_{21} = (W^1_{21}, W^2_{21})$, with the corresponding GDoF $d^1_{11} = d^2_{11} = d^1_{22} = d^2_{22} = \frac{d_{11}}{2} = \frac{1-\alpha}{3}$. Re-

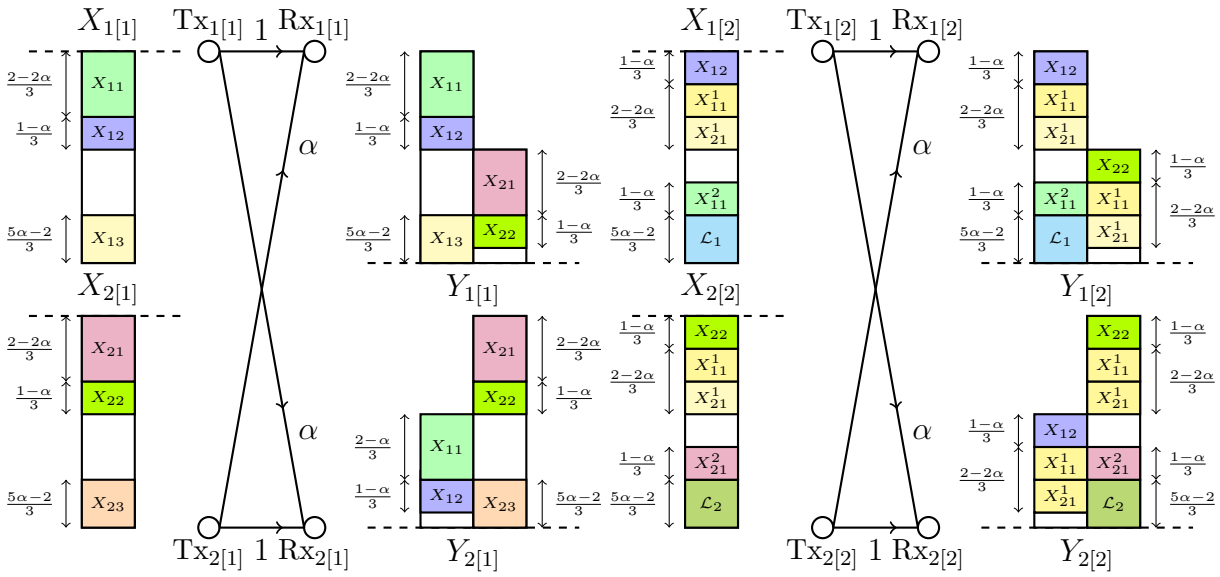

Figure 4.9: *Achievable scheme for* $\frac{1}{2} \leq \alpha \leq \frac{4}{7}$. $\mathcal{L}_1 = \mathcal{L}(X_{13}, X_{22}), \mathcal{L}_2 = \mathcal{L}(X_{23}, X_{12})$. *The left figure is the first hop while the right figure represents the second hop.*

lay $\mathrm{Tx}_{1[2]} \equiv \mathrm{Rx}_{1[1]}$ then re-encodes $W_{12}, W_{11}^1, W_{11}^2, W_{22}^1$ into independent Gaussian codebooks producing codewords $X_{12}, X_{11}^1, X_{11}^2, X_{22}^1$ with power $\mathrm{E}|X_{12}|^2 = 1 - P^{-d_{12}}, \mathrm{E}|X_{11}^1|^2 = P^{-d_{12}} - P^{-d_{12}-d_{11}^1}, \mathrm{E}|X_{21}^1|^2 = P^{-d_{12}-d_{11}^1} - P^{-1+d_{11}^2+d_{13}}, \mathrm{E}|X_{11}^2|^2 = P^{-1+d_{11}^2+d_{13}} - P^{-1+d_{13}}$. Relay $\mathrm{Tx}_{2[2]} \equiv \mathrm{Rx}_{2[1]}$ proceeds similarly. The transmitted signals at the relays are $X_{1[2]} = X_{12}+X_{11}^1+X_{21}^1+X_{11}^2+\mathcal{L}_1, X_{2[2]} = X_{22}+X_{11}^1+X_{21}^1+X_{21}^2+\mathcal{L}_2$. Then destination node $\mathrm{Rx}_{1[2]}$ is able to decode $W_{12}, W_{11}^1, W_{21}^1, W_{22}, W_{11}^2, W_{13}$ successively, while treating the other signals as noise, with the corresponding SINR values $\sim P^{d_{12}}, \sim P^{d_{11}^1}, \sim P^{d_{21}^1}, \sim P^{d_{22}}, \sim P^{d_{11}^2}, \sim P^{d_{13}}$. Destination $\mathrm{Rx}_{2[2]}$ proceeds similarly by decoding $W_{22}, W_{11}^1, W_{21}^1, W_{12}, W_{21}^2, W_{23}$ successively. See Figure 4.9 for an illustration.

- $\alpha \in [4/7, 2/3]$

  In the regime $\alpha \in [4/7, 2/3]$, we have the bound $\mathcal{D}_{\Sigma}^{f.p.} = 2 - \alpha$. The achievable scheme is similar to the case $\alpha \in [1/2, 4/7]$ except that $d_{i1} = \frac{2-2\alpha}{3}, d_{i2} = \frac{1-\alpha}{3}, d_{i3} = \frac{\alpha}{2}$.

- $\frac{2}{3} \leq \alpha \leq 1$

  As noted previously, the achievable scheme in this regime is quite simple, the 2-hop network simply operates as a concatenation of two interference channels using decode-and-forward.

4.5.2 Proof of Achievability for Theorem 4.2

Building on the insights from the 2-hop solution, the achievable scheme for L-hop setting makes use of rate-splitting at the sources and partial decode-and-forward combined with amplify and forward at the relays. It can be visualized as incremental peeling off of interfered layers, such that the relays of the next layer can decode one more interfered layer compared to the relays in the previous layer. The main ideas are illustrated through the following example of a 3-hop network.

Example 4.1. *Suppose* $L = 3, \alpha = \frac{1}{2}$. *Then we have the sum-GDoF value* $\mathcal{D}_{\Sigma}^{f.p.} = 10/7$. *The achievable scheme is illustrated in Figure 4.10.*

In general, the achievable schemes have the following four sub-cases when $\alpha \leq 1$.

- $\alpha \leq \frac{1}{2}$

  In this regime, message W_i is split into $2L$ sub-messages, i.e., $W_i = (W_{i1}, W_{i2}, \cdots, W_{i(2L)})$, which carry $d_{i1} = \frac{\alpha \times 2^L}{2(2^L-1)}$, $d_{i2} = \frac{\alpha \times 2^{L-1}}{2(2^L-1)}$, $\cdots$, $d_{iL} = \frac{\alpha \times 2^1}{2(2^L-1)}$, $d_{iL+1} = 1-2\alpha$, $d_{iL+2} = \frac{\alpha \times 2^{L-1}}{2(2^L-1)}$, $\cdots$, $d_{i(2L)} = \frac{\alpha \times 2^1}{2(2^L-1)}$ GDoF respectively. These sub-messages are encoded into independent Gaussian codebooks with codewords $X_{i1}, X_{i2}, \cdots, X_{i(2L)}$ with the corresponding powers $\mathrm{E}|X_{i1}|^2 = 1$, $\mathrm{E}|X_{i2}|^2 = P^{-d_{i1}}, \mathrm{E}|X_{i3}|^2 = P^{-d_{i1}-d_{i2}}, \cdots, \mathrm{E}|X_{iL}|^2 = P^{-d_{i1}-d_{i2}-\cdots-d_{i(L-1)}}$, $\mathrm{E}|X_{i(L+1)}|^2 = P^{-d_{i1}-d_{i2}-\cdots-d_{i(L)}}, \mathrm{E}|X_{i(L+2)}|^2 = P^{d_{i(L+1)}+d_{i(L+2)}+\cdots+d_{i(2L)}-1}, \cdots, \mathrm{E}|X_{i(2L-1)}|^2 = P^{d_{i(2L-1)}+d_{i(2L)}-1}$, $E|X_{i(2L)}|^2 = P^{d_{i(2L)}-1}$ respectively, up to scaling by an $O(1)$ constant to ensure a sum-power of unity. The relay node $\mathrm{Rx}_{1[1]}$ is able to decode X_{11}, X_{12}, $\cdots$, X_{1L}, $X_{1(L+1)}$, X_{21} successively. The remaining signal above the noise floor of relay $\mathrm{Rx}_{1[1]}$ is the combination of codewords: $(X_{1(L+2)}, \cdots, X_{1(2L)}, X_{22}, \cdots, X_{2(L)})$, and we denote the combination as $\mathcal{L}_{1[1]}$. Relay $\mathrm{Rx}_{1[1]}$ scales this combination (amplify and forward) such that the power of this combination is $P^{d_{1(L+2)}+\cdots+d_{1(2L)}-1} = P^{\frac{\alpha(2^L-2)}{2(2^L-1)}-1}$. Since both relays know messages W_{11}, W_{21}, they split these two messages into two sub-messages: $W_{11} = (W_{11}^1, W_{11}^2), W_{21} = (W_{21}^1, W_{21}^2)$, where $d_{i1}^1 = d_{i1}^2 = \frac{d_{i1}}{2}$. Relay $\mathrm{Tx}_{1[2]} \equiv$

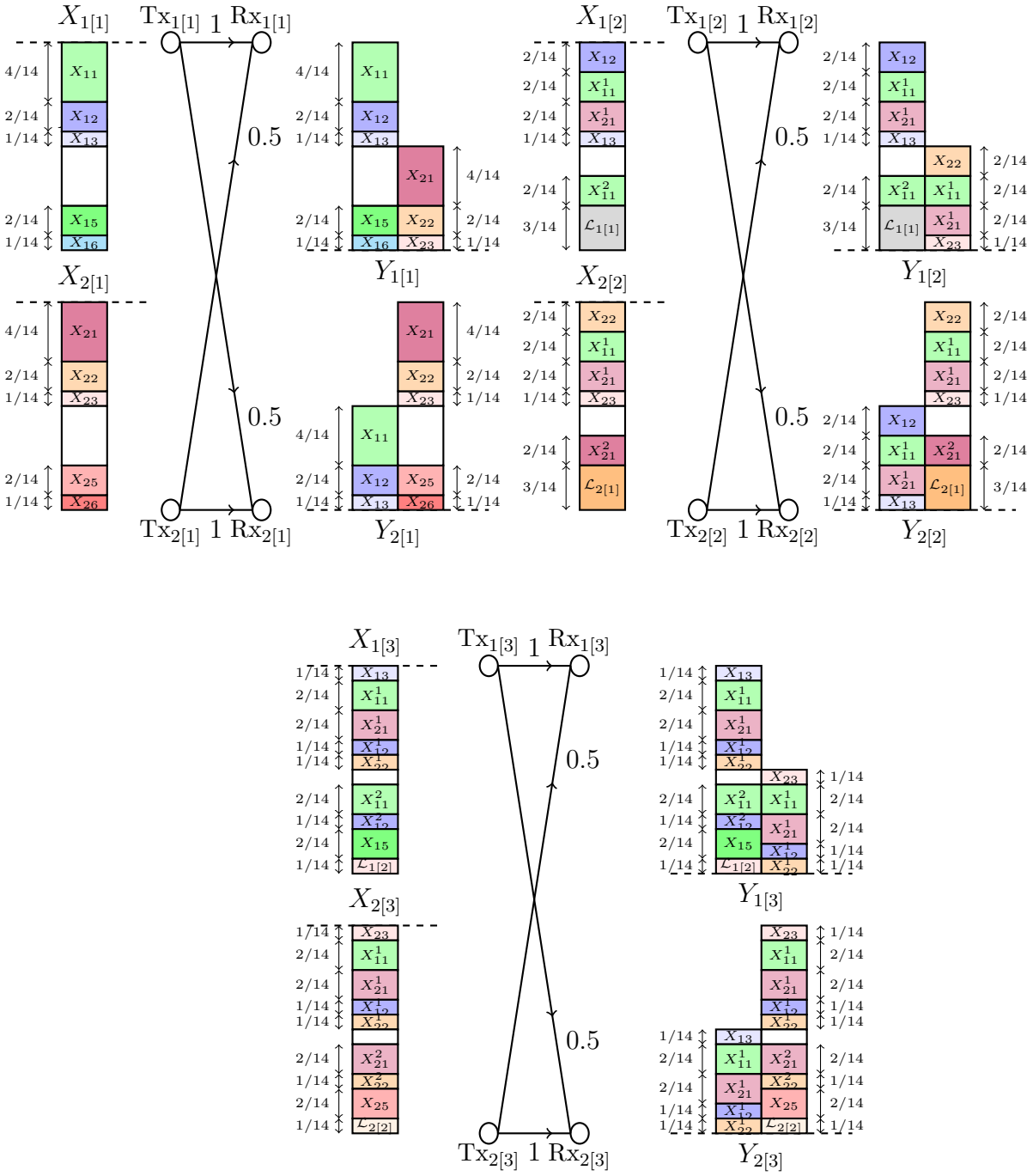

Figure 4.10: *Achievable scheme for* $L = 3, \alpha = 1/2$. X_{14}, X_{24} *have* 0 *GDoF so they are not shown in the figure.* $\mathcal{L}_{1[1]} = \mathcal{L}(X_{15}, X_{22}, X_{16}, X_{23}), \mathcal{L}_{2[1]} = \mathcal{L}(X_{25}, X_{12}, X_{26}, X_{13}), \mathcal{L}_{1[2]} = \mathcal{L}(X_{16}, X_{23}), \mathcal{L}_{2[2]} = \mathcal{L}(X_{26}, X_{13})$. *The top left, top right, bottom figures are the* $1^{st}, 2^{nd}, 3^{rd}$ *hop respectively. The interfered layer at* $Rx_{i[1]}$ *are* $X_{i5}, X_{\bar{i}2}, X_{i6}, X_{\bar{i}3}$. *It can do nothing but amplify and forward this layer. Then, the relay at the next hop,* $Rx_{i[2]}$ *is able to decode* $W_{i5}, W_{\bar{i}2}$, *such that the interfered layer becomes* $X_{i6}, X_{\bar{i}3}$. *After that,* $Rx_{i[3]}$ *is able to decode* $W_{i6}, W_{\bar{i}3}$.

$\mathrm{Rx}_{1[1]}$ re-encodes message W_{12}, W_{11}^1, W_{21}^1, W_{13}, W_{14}, $\cdots$, W_{1L}, $W_{1(L+1)}$, W_{11}^2 into independent Gaussian codebooks producing codewords X_{12}, X_{11}^1, X_{21}^1, X_{13}, X_{14},$\cdots$, X_{1L}, $X_{1(L+1)}$, X_{11}^2 with the corresponding powers $\mathrm{E}|X_{12}|^2 = 1$, $\mathrm{E}|X_{11}^1|^2 = P^{-d_{12}}$, $\mathrm{E}|X_{21}^1|^2 = P^{-d_{12}-d_{11}^1}$, $\mathrm{E}|X_{13}|^2 = P^{-d_{12}-d_{11}^1-d_{21}^1}$, $\cdots$, $\mathrm{E}|X_{1(L+1)}|^2 = P^{-d_{12}-d_{11}^1-d_{21}^1-d_{13}-\cdots-d_{1L}}$, $\mathrm{E}|X_{11}^2|^2 = P^{d_{11}^2+d_{1(L+2)}+\cdots+d_{1(2L)}-1} - P^{d_{1(L+2)}+\cdots+d_{1(2L)}-1}$. Relay $\mathrm{Tx}_{2[2]} \equiv \mathrm{Rx}_{2[1]}$ proceeds similarly. The transmitted signals are $X_{1[2]} = X_{12} + X_{11}^1 + X_{21}^1 + X_{13} + X_{14} + \cdots + X_{1L} + X_{1(L+1)} + X_{11}^2 + \mathcal{L}_{1[1]}$, $X_{1[2]} = X_{22} + X_{11}^1 + X_{21}^1 + X_{23} + X_{24} + \cdots + X_{2L} + X_{2(L+1)} + X_{21}^2 + \mathcal{L}_{2[1]}$. Then, the relay at the next layer: Relay $\mathrm{Rx}_{1[2]}$ is able to decode $W_{12}, W_{11}^1, W_{21}^1, W_{13}, W_{14}, \cdots, W_{1L}, W_{1(L+1)}$, $W_{22}, W_{11}^2, W_{1(L+2)}$ successively. Compared to Relay $\mathrm{Rx}_{1[1]}$, an additional sub-message $W_{1(L+2)}$ can also be decoded at Relay $\mathrm{Rx}_{1[2]}$. The remaining signal above the noise floor of relay $\mathrm{Rx}_{1[2]}$ is the combination of codewords: $(X_{1(L+3)}, \cdots, X_{1(2L)}, X_{23}, \cdots, X_{2(L)})$, which is denoted as $\mathcal{L}_{1[2]}$ and amplified with power P^{-1}. Then relays $\mathrm{Rx}_{1[2]}, \mathrm{Rx}_{2[2]}$ split W_{12}, W_{22}, i.e., $W_{12} = (W_{12}^1, W_{12}^2), W_{22} = (W_{12}^1, W_{12}^2)$, with the corresponding GDoF $d_{i2}^1 = d_{i2}^2 = \frac{d_{i2}}{2}$. Relay $\mathrm{Rx}_{1[2]} \equiv \mathrm{Tx}_{1[3]}$ re-encodes message W_{13}, W_{11}^1, W_{21}^1, W_{12}^1, W_{22}^1, W_{13}, W_{14}, $\cdots, W_{1L}$, $W_{1(L+1)}$, W_{11}^2, W_{12}^2 into independent Gaussian codebooks producing codewords X_{12}, X_{11}^1, X_{21}^1, X_{12}^1, X_{22}^1, X_{13}, X_{14}, $\cdots, X_{1L}, X_{1(L+1)}, X_{11}^2, X_{12}^2$. The transmitted signals at Relay $\mathrm{Rx}_{1[2]} \equiv \mathrm{Tx}_{1[3]}$ is $X_{1[3]} = X_{13} + X_{11}^1 + X_{21}^1 + X_{12}^1 + X_{22}^1 + X_{14} + X_{15} + \cdots + X_{1L} + X_{1(L+1)} + X_{11}^2 + X_{12}^2 + X_{1(L+2)} + \mathcal{L}_{1[2]}$. This idea applies to all the subsequent layers, such that Relay $\mathrm{Rx}_{1[l]}$ is able to decode one more sub-message $W_{i(L+l)}$ compared to the Relay $\mathrm{Rx}_{1[l-1]}$. In other words, the relays at the next layer can decode one more sub-message compared to the relays in the preceding layer. In the ℓ^{th} hop, Relay $\mathrm{Tx}_{1[l+1]} \equiv \mathrm{Rx}_{1[l]}$ splits the message $W_{i(i-1)}$, i.e., $W_{i(i-1)} = (W_{i(i-1)}^1, W_{i(i-1)}^2)$. The transmitted signal at Relay $\mathrm{Tx}_{1[l+1]} \equiv \mathrm{Rx}_{1[l]}$ is $X_{1[l+1]} = X_{1(l+1)} + X_{11}^1 + X_{21}^1 + X_{12}^1 + X_{22}^1 + \cdots + X_{1(l)}^1 + X_{2(l)}^1 + X_{1(l+2)} + \cdots + X_{1L} + X_{2(L+1)} + X_{11}^2 + X_{12}^2 + \cdots X_{1(l)}^2 + \mathcal{L}_{1[l]}$. Therefore, in the last hop, Destination $\mathrm{Rx}_{1[L]}$ can decode $W_{1(L)}, W_{11}^1, W_{21}^1, W_{12}^1, W_{22}^1, \cdots, W_{1(L-1)}^1, W_{2(L-1)}^1, W_{1(L+1)}, W_{11}^2, \cdots, W_{1(L-1)}^2, W_{1(2L)}$ successively.

- $\frac{1}{2} \leq \alpha \leq \frac{2^L}{2^{L+1}-1}$

In the achievable scheme, W_i is split into $(2L-1)$ sub-messages, i.e., $W_i = (W_{i1}, W_{i2}, \cdots, W_{i(2L-1)})$, which carry $\frac{(1-\alpha)\times 2^L}{2(2^L-1)}, \frac{(1-\alpha)\times 2^{L-1}}{2(2^L-1)}, \cdots, \frac{(1-\alpha)\times 2^1}{2(2^L-1)}, \frac{\alpha\times 2^{L-1}}{2(2^L-1)}, \cdots, \frac{\alpha\times 2^1}{2(2^L-1)}, 2\alpha - 1 + \frac{1-\alpha}{2^L-1}$ GDoF respectively. The idea of message splitting, successive decoding, partial decode-and-forward and amplify-and-forward at the relays is similar to the case $\alpha \leq \frac{1}{2}$.

- $\frac{2^L}{2^{L+1}-1} \leq \alpha \leq \frac{2}{3}$

  In the achievable scheme, W_i is split into $(2L-1)$ sub-messages, i.e., $W_i = (W_{i1}, W_{i2}, \cdots, W_{i(2L-1)})$, which carry $\frac{(1-\alpha)\times 2^L}{2(2^L-1)}, \frac{(1-\alpha)\times 2^{L-1}}{2(2^L-1)}, \cdots, \frac{(1-\alpha)\times 2^1}{2(2^L-1)}, \frac{\alpha\times 2^{L-1}}{2(2^L-1)}, \cdots, \frac{\alpha\times 2^1}{2(2^L-1)}, \frac{\alpha}{2} - \frac{(1-\alpha)(2^{L-1}-2)}{2^L-1}$ GDoF respectively. The idea of message splitting, successive decoding, partial decode-and-forward and amplify-and-forward at the relays is similar to the case $\alpha \leq \frac{1}{2}$.

- $\frac{2}{3} \leq \alpha \leq 1$ The achievable scheme for this case is simple as the bound equals to $2-\alpha$, so each hop acts as the interference channel, and a simple decode and forward strategy suffices.

4.5.3 Proof of Achievability for Theorem 4.3

The lower bound is achieved as follows: W_1, W_2 are split into L sub-messages: $W_1 = (W_{11}, W_{12}, \cdots, W_{1L})$, $W_2 = (W_{21}, W_{22}, \cdots, W_{2L})$, and each sub-message carries 1 GDoF. At Source $\mathrm{Tx}_{i[1]}$, the sub-messages $W_{i1}, W_{i2}, \cdots, W_{iL}$ are encoded into independent Gaussian codebooks producing codewords $X_{i1}, X_{i2}, \cdots, X_{iL}$ with powers $\mathrm{E}|X_{i1}|^2 = 1$, $\mathrm{E}|X_{i2}|^2 = P^{-1}$, $\cdots$, $\mathrm{E}|X_{iL}| = P^{-L+1}$, up to an $O(1)$ scaling factor to ensure the sum power of unity. Then the relay node $\mathrm{Rx}_{i[1]}$ is able to decode $W_{\bar{i}1}, W_{\bar{i}2}, \cdots, W_{\bar{i}L}, W_{i1}$ by successive decoding. Next, the relay acts as transmitter $\mathrm{Tx}_{i[2]}$ and re-encodes $W_{\bar{i}2}, \cdots, W_{\bar{i}L}, W_{i1}$ into independent Gaussian codebooks producing codewords $X_{\bar{i}2}, \cdots, X_{\bar{i}L}, X_{i1}$ with power $\mathrm{E}|X_{\bar{i}2}|^2 = 1$, $\mathrm{E}|X_{\bar{i}3}|^2 = P^{-1}$, $\cdots$, $\mathrm{E}|X_{\bar{i}L}|^2 = P^{-L+2}$, $\mathrm{E}|X_{i1}|^2 = P^{-L+1}$, respectively, up to an $O(1)$ normalizing factor. There is a simple interpretation for this scheme: the relays remove (decode and subtract) the sub-message that has the maximum power (topmost layer) and re-transmit the remain-

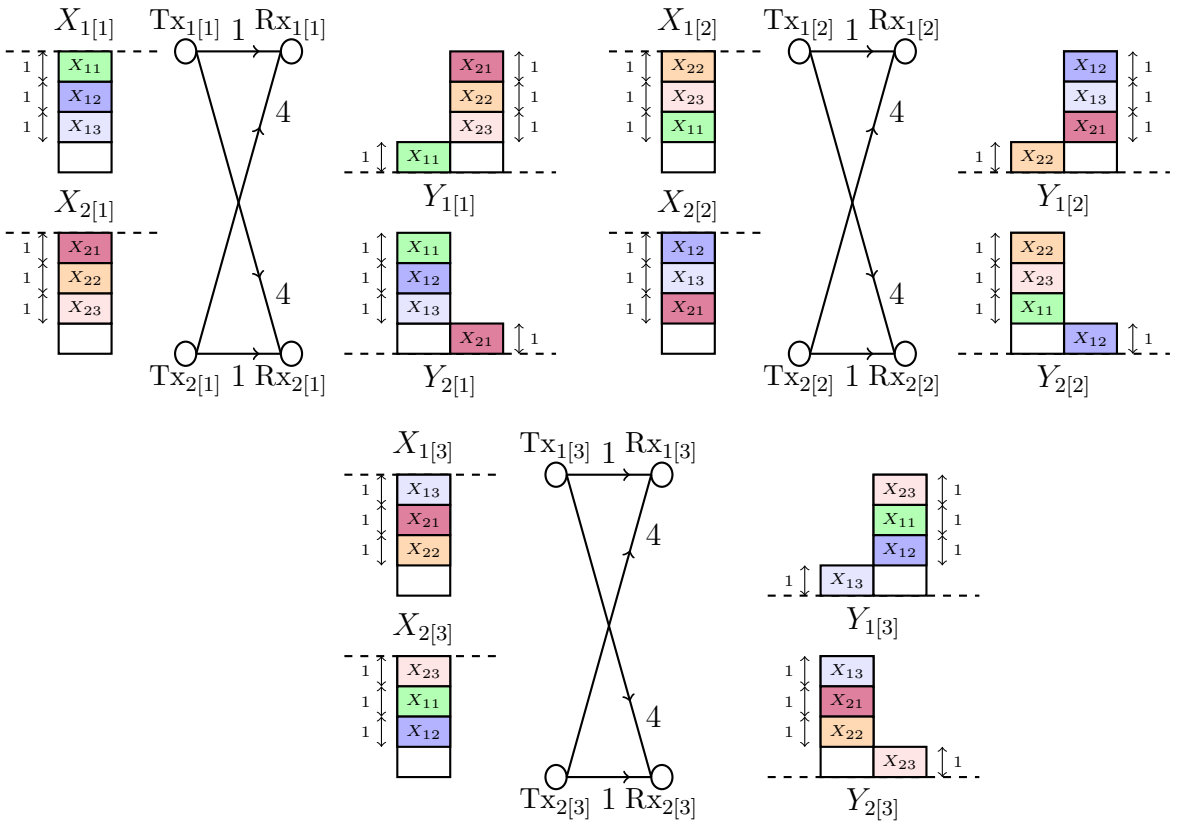

Figure 4.11: *Achievable scheme for* $L = 3, \alpha = 4$.

ing sub-messages by keeping their original layers and scaling power to meet the maximum allowed power level. Hence, if ℓ is even, $\mathrm{Rx}_{i[\ell]}$ is able to decode $W_{i\ell}, \cdots, W_{iL}, W_{\bar{i}1}, \cdots, W_{\bar{i}L}$ successively. Then as a transmitter, $\mathrm{Tx}_{i[\ell+1]}$ transmits $W_{i(\ell+1)}, \cdots, W_{iL}, W_{\bar{i}1}, \cdots, W_{\bar{i}L}$. If ℓ is odd, $\mathrm{Rx}_{i[\ell]}$ is able to decode $W_{\bar{i}\ell}, \cdots, W_{\bar{i}L}, W_{i1}, \cdots, W_{i\ell}$ successively, and then as a transmitter, $\mathrm{Tx}_{i[\ell+1]}$ transmits $W_{\bar{i}(\ell+1)}, \cdots, W_{\bar{i}L}, W_{i1}, \cdots, W_{i\ell}$. Hence, at the last hop, where L is odd, destination $\mathrm{Rx}_{i[L]}$ is able to decode $W_{\bar{i}L}, W_{i1}, \cdots, W_{iL}$ successively. An example of $L = 3, \alpha = 4$ is illustrated in Figure 4.11.

4.6 Extension to the Asymmetric Setting

In this section, we explore the $2\times2\times2$ asymmetric interference channel where the direct link strength values at both hops are 1 and cross link strength values at the first hop and second hop are denoted as $\alpha_{[1]}, \alpha_{[2]}$ respectively. See Fig. 4.12 for an illustration. We will show an achievable scheme that is a rate-splitting between decode-and-forward and quantize-and-forward via an example. Such an scheme is GDoF optimal by extending the converse proof in this chapter from symmetric setting to the asymmetric setting straightforwardly.

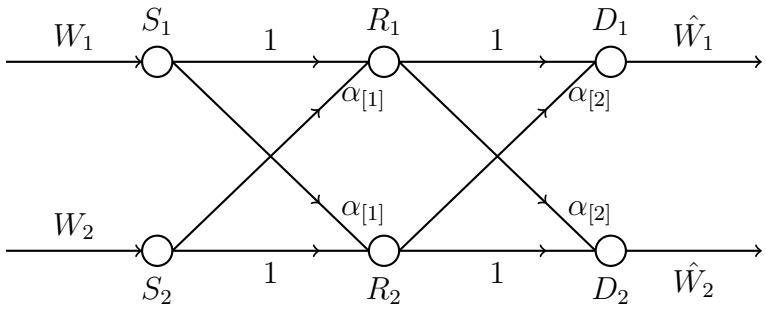

Figure 4.12: *Layered* $2\times2\times2$ *interference channel model.*

Example 4.2. *Let us consider the setting* $\alpha_{[1]}=\frac{1}{2}, \alpha_{[2]}=\frac{1}{4}$*, through which we are able to see the idea of combination of Rate Splitting, Decode-and-forward and Quantize-and-forward schemes. The sum-GDoF value is* $\mathcal{D}_\Sigma = 1.5$*. The achievable scheme is as follows:* W_i *is split into* 3 *sub-messages, i.e.,* $W_i=(W_{i1}, W_{i2}, W_{i3})$*, each carry* $\frac{1}{4}$ *GDoF, they are encoded into independent Gaussian codebooks producing codewords* X_{i1}, X_{i2}, X_{i3} *with the corresponding power* $\mathrm{E}|X_{i1}|^2 = 1-P^{-1/4}, \mathrm{E}|X_{i2}|^2 = P^{-1/4}-P^{-1/2}, \mathrm{E}|X_{i3}|^2 = P^{-1/2}$ *respectively. The transmitted codewords at Source* i *is* $X_{i[1]} = X_{i1}+X_{i2}+X_{i3}$*. At Relay* 1*, it use successive decoding to decode* X*, the remaining signal is the noisy linear combination of* X_{13}, X_{22}*, let us denote it as* $\mathcal{L}_1$*, i.e.,* $\mathcal{L}_1 = G_{11[1]}\sqrt{P}X_{13} + G_{12[1]}\sqrt{P^{1/2}}X_{22} + Z_{1[1]}$*, which contains power* $\delta_{\mathcal{L}_1}, \mathrm{E}|\delta_{\mathcal{L}_1}| = P^{1/4}$*. Vector quantization is used to quantize the remaining signal over* N *channel uses by setting the quadratic distortion at the noise level in order not to incur the GDoF loss, i.e.,* $\mathcal{L}_1 = \hat{\mathcal{L}}_1 + \Delta_1$ *with* $\mathrm{E}|\Delta_1|^2 = 1$*. This in effect extracts the information of the*

remaining signals above the noise level. According to the rate-distortion theorem, $\frac{1}{4}N\log(P)$ *bits are needed to quantize the interfered signal over* N *channel uses. Since* $\hat{\mathcal{L}}_1$ *has* $\frac{1}{4}N\log(P)$ *bits, it is equivalently to say the it has* $1/4$ *GDoF. The transmission at the Relay* 1 *is as follows: It re-encodes* $W_{12}, W_{11}, \hat{\mathcal{L}}_1$ *into independent Gaussian codebooks producing codewords* $X_{12}, X_{11}, X_{\hat{\mathcal{L}}_1}$*. The transmitted signals at Relay* 1 *is* $X_{1[2]} = X_{12}+X_{11}+X_{\hat{\mathcal{L}}_1}$*. Relay* 2 *proceeds similarly. Note that Source* i *and Relay* i *have different power assignment for* X_{i2}, X_{i1}*. They are significant as* X_{i2} *appears above the noise floor and* X_{i1} *appears below at the Destination* $\bar{i}$*, such that Destination* $\bar{i}$ *is able to decode* X_{i2}*. Destination* 1 *is able to use successive decoding to decode* $W_{12}, W_{11}, \hat{\mathcal{L}}_1, W_{22}$ *sequentially. Thus, it is able to reconstruct* X_{22} *from* W_{22}*. Then, since*

$$\hat{\mathcal{L}}_1 = \mathcal{L}_1 - \Delta_1 \tag{4.104}$$

$$= G_{11[1]}\sqrt{P}X_{13} + G_{12[1]}\sqrt{P^{1/2}}X_{22} + Z_{1[1]} - \Delta_1 \tag{4.105}$$

$$= G_{11[1]}\sqrt{P}X_{13} + G_{12[1]}\sqrt{P^{1/2}}X_{22} + \hat{Z}_{1[1]}. \tag{4.106}$$

where $\hat{Z}_{1[1]} \triangleq Z_{1[1]} - \Delta_1$*. Destination* 1 *can subtract the term* $G_{12[1]}\sqrt{P^{1/2}}X_{22}$ *from* $\hat{\mathcal{L}}_1$ *as it already decode* W_{22}*. Since the aggregated noise term* $\hat{Z}_{1[1]}$ *has power* $\mathcal{O}(1)$[7]*, it is able to decode* X_{13} *for* W_{13}*, which contains* $1/4$ *GDoF.*

4.7 Summary

Motivated by the need to understand the robust information-theoretic limits of multi-hop communication networks, in this chapter we initiated the study of the sum-GDoF of layered symmetric L-hop $2 \times 2 \times \cdots \times 2$ networks comprised of 2 nodes in each layer, under finite precision CSIT. As our main contribution, recently introduced sum-set inequalities [20] that

[7]We can observe the noise accumulation due to the quantization, but this is inconsequential in the GDoF sense.

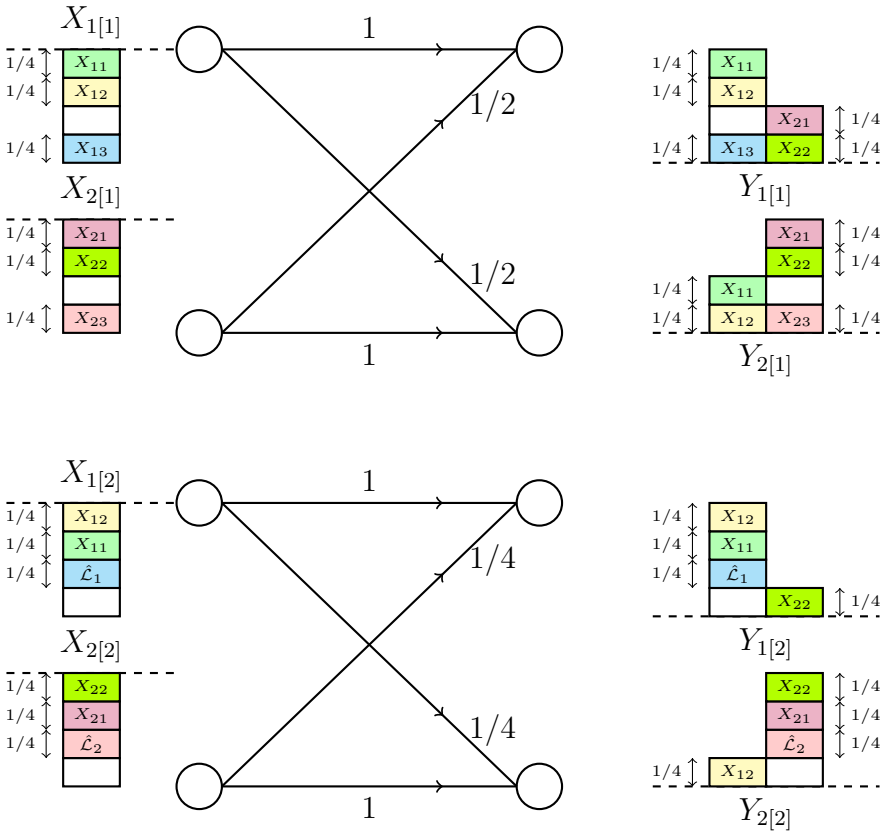

Figure 4.13: *Achievable scheme for* $\alpha_{[1]} = \frac{1}{2}, \alpha_{[2]} = \frac{1}{4}$*, where the dashed line represents the noise floor.* $\mathcal{L}_1 = \mathcal{L}(X_{13}, X_{22}), \mathcal{L}_1 = \mathcal{L}(X_{12}, X_{23})$*. The top figure is the first hop while the bottom figure represents the second hop.*

build upon Aligned Images bounds of [16] were shown to be sufficient to settle the sum-GDoF of this symmetric setting. Notable technical issues that surfaced in our study include the challenge of applying deterministic transformations that were developed for one-hop communication to multi-hop settings, as well as the dependence of coding functions on CSIT that may be available for previous hops and need not satisfy the finite precision assumption. These challenges were overcome through recursive reasoning that applies the deterministic transformation to only one-hop at a time. In terms of optimal solutions, under finite precision CSIT we found that ideas such as Interference Neutralization [11, 61, 47, 38, 59], Aligned Interference Neutralization [34, 64, 33] and Network Diagonalization [58] are too fragile to retain their GDoF benefits, and instead rate-splitting solutions that combine amplify-and-forward and decode-and-forward principles, along with careful layering (superposition) of

messages that allows each successive stage of relays to acquire more common information, are sum-GDoF optimal. The compact expressions obtained from converse bounds prove insightful in designing the optimal achievable schemes. From the big-picture perspective, a takeaway message from the sum-GDoF characterizations is that, on one hand, optimal robust solutions tend to not improve much upon basic alternatives (in this case, the trivial decode-and-forward solution) when all channels are of comparable strength, but on the other hand, when the channel strengths are sufficiently different then significant gains over basic alternatives are possible by optimizing robust solutions. The latter is particularly important for the high-frequency communication networks that motivated this chapter, where due to high path loss, blockages, and often due to directional transmission, there tends to be a much higher spatial variance in channel strengths than in conventional cellular networks. Thus, the results of this chapter, while clearly limited by the simplifying assumptions of layered structure and symmetric gains, nevertheless indicate that significant robust gains are possible by optimizing multi-hop communication for the types of richly diverse topologies that would be typical in high-frequency directed communication networks.

Chapter 5

Conclusion

In this dissertation, by utilizing the Aligned Images set and its extension called Sum-set inequalities, we characterize the GDoF for three channel settings under finite precision CSIT. The corresponding achievable schemes are found to match the upper bounds. The finite precision CSIT assumption blurs the channel coefficients, thus, the structured coding schemes such as zero forcing, interference alignment are filtered out. Random coding schemes that are simpler and more robust become optimal under the finite precision CSIT assumption. The contributions are summarized as follows:

In chapter 2, we explored the optimality condition of Treating Interference as noise under finite precision CSIT. TIN is shown to be GDoF optimal in the CTIN regime, which is a larger regime than the TIN regime under perfect CSIT setting.

In chapter 3, the sum-GDoF characterization for the 2-user interference channel with limited transmitter cooperation under finite precision CSIT is considered. We explored the cooperation benefit for arbitrary channel strength parameters. The cooperation gain is found to be either $0, 1, 1/2$ or $1/3$, i.e., the number of over-the-air bits that each cooperation bit buys is shown to be equal to either $0, 1, 1/2$ or $1/3$. The most interesting and challenging result

is the 1/3 slope that appears only in the strong interference regime under finite precision CSIT. The achievable scheme utilizes rate splitting scheme and the converse proof relies on non-trivial application of Aligned Images set.

In chapter 4, we explore the multi-hop symmetric interference channel under finite precision CSIT assumption. A recursive approach combined with the aligned images sum-set inequalities is utilized for the multi-hop interference channel under finite precision CSIT to characterize the sum-GDoF upper bound. A combination of classical random coding schemes: Rate Splitting, Decode-and-forward and Amplify-and-forward is shown to be GDoF optimal. An interesting observation is that the sum-GDoF value approaches that of one-hop broadcast channel with full transmitter cooperation under finite precision CSIT as the number of hops approaches infinity.

An interesting future direction is to find the sum-GDoF of general asymmetric multi-hop interference channel. The asymmetric multi-hop interference channel is challenging because there are more channel strength parameters. It is an interesting open problem to see whether the current aligned images sum-set inequalities are enough to obtain a tight upper bound. Moreover, we give an example in chapter 4 that shows the quantize-and-forward scheme is necessary to match the upper bound. In general, it is non-trivial to find the optimal random coding schemes for the asymmetric multi-hop networks.

Another research direction under channel uncertainty is to study the computation problems over wireless networks. Computation over multiple access channel [52, 10] attracts researchers' attention in recent years as it is able to reduce the communication overhead for the sensor network and federated learning. Since channel uncertainty is unavoidable in practice, the fundamental limits of the over-the-air computation problem is an interesting research direction when channel uncertainty is involved.

Bibliography

[1] V. Annapureddy and V. Veeravalli. Gaussian interference networks: Sum capacity in the low interference regime and new outer bounds on the capacity region. *IEEE Trans. on Information Theory*, pages 3032–3050, July 2009.

[2] V. S. Annapureddy and V. V. Veeravalli. Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region. *IEEE Transactions on Information Theory*, 55(7):3032–3050, 2009.

[3] V. S. Annapureddy and V. V. Veeravalli. Sum capacity of mimo interference channels in the low interference regime. *IEEE Transactions on Information Theory*, 57(5):2565–2581, 2011.

[4] A. Avestimehr, S. Diggavi, C. Tian, and D. Tse. An approximation approach to network information theory. In *Foundations and Trends in Communication and Information Theory*, volume 12, pages 1–183, 2015.

[5] A. Avestimehr, S. Diggavi, and D. Tse. Wireless network information flow: A deterministic approach. *IEEE Trans. on Inf. Theory*, 57:1872–1905, 2011.

[6] S. Avestimehr, A. Sezgin, and D. Tse. Capacity of the Two Way Relay Channel within a Constant Gap. *European Transactions on Telecommunications*, 21:363 – 374, April 2010.

[7] V. Cadambe and S. Jafar. Interference alignment and the degrees of freedom of wireless X networks. *IEEE Trans. on Information Theory*, (9):3893–3908, Sep 2009.

[8] V. R. Cadambe and S. A. Jafar. Degrees of freedom of wireless networks with relays, feedback, cooperation and full duplex operation. *IEEE Transactions on Information Theory*, 55:2334–2344, May 2009.

[9] G. Caire and S. Shamai. On the achievable throughput of a multiantenna Gaussian broadcast channel. *IEEE Trans. Inform. Theory*, 49(7):1691–1706, July 2003.

[10] X. Cao, G. Zhu, J. Xu, and K. Huang. Optimized power control for over-the-air computation in fading channels. *IEEE Transactions on Wireless Communications*, 19(11):7498–7513, 2020.

[11] Y. Cao and B. Chen. Capacity bounds for two-hop interference networks. *47th Annual Allerton Conference on Communication, Control, and Computing*, abs/0910.1532, 2009.

[12] Y. Chan, J. Wang, and S. A. Jafar. Toward an extremal network theory – robust GDoF gain of transmitter cooperation over TIN. *IEEE Transactions on Information Theory*, 66(6):3827–3845, 2020.

[13] Y.-C. Chan and S. A. Jafar. Exploring Aligned-Images Bounds: Robust Secure GDoF of 3-to-1 Interference Channel. *Technical Report, https://escholarship.org/uc/item/8nh0m0qm*, October 2020.

[14] B. Clerckx, H. Joudeh, C. Hao, M. Dai, and B. Rassouli. Rate splitting for MIMO wireless networks: A promising PHY-layer strategy for LTE evolution. *IEEE Communications Magazine*, pages 98–105, May 2016.

[15] T. M. Cover. Broadcast channels. *IEEE Transactions on Information Theory*, 18(1):2–14, Jan. 1972.

[16] A. G. Davoodi and S. A. Jafar. Aligned image sets under channel uncertainty: Settling conjectures on the collapse of degrees of freedom under finite precision CSIT. *IEEE Transactions on Information Theory*, 62(10):5603–5618, 2016.

[17] A. G. Davoodi and S. A. Jafar. Generalized Degrees of Freedom of the Symmetric K-User Interference Channel under Finite Precision CSIT. *IEEE Transactions on Information Theory*, 63(10):6561–6572, 2017.

[18] A. G. Davoodi and S. A. Jafar. Transmitter cooperation under finite precision CSIT: A GDoF perspective. *IEEE Transactions on Information Theory*, 63(9):6020–6030, 2017.

[19] A. G. Davoodi and S. A. Jafar. Aligned image sets and the generalized degrees of freedom of symmetric MIMO interference channel with partial CSIT. *IEEE Transactions on Information Theory*, 65(1):406–417, Jan. 2019.

[20] A. G. Davoodi and S. A. Jafar. Sum-set inequalities from aligned image sets: Instruments for robust gdof bounds. *IEEE Transactions on Information Theory*, 66(10):6458–6487, 2020.

[21] A. G. Davoodi, B. Yuan, and S. A. Jafar. GDoF of the MISO BC: Bridging the gap between finite precision and perfect CSIT. *IEEE Transactions on Information Theory*, 64(11):7208–7217, Nov 2018.

[22] A. El Gamal and Y.-H. Kim. *Network information theory*. Cambridge University Press, 2011.

[23] R. Etkin and E. Ordentlich. The degrees-of-freedom of the K-User Gaussian interference channel is discontinuous at rational channel coefficients. *IEEE Trans. on Information Theory*, 55:4932–4946, Nov. 2009.

[24] R. Etkin, D. Tse, and H. Wang. Gaussian interference channel capacity to within one bit. *IEEE Transactions on Information Theory*, 54(12):5534–5562, 2008.

[25] C. Geng and S. Jafar. Secure GDoF of $K-$user Gaussian interference channels: When secrecy incurs no penalty. *IEEE Communications Letters*, 19(8):1287–1290, Aug. 2015.

[26] C. Geng and S. Jafar. On the Optimality of Treating Interference as Noise: Compound Interference Networks. *IEEE Transactions on Information Theory*, 62(8):4630–4653, Aug. 2016.

[27] C. Geng and S. Jafar. Power control by GDoF duality of treating interference as noise. *IEEE communications letters*, 22(2):244–247, 2018.

[28] C. Geng, N. Naderializadeh, S. Avestimehr, and S. Jafar. On the Optimality of Treating Interference as Noise. *IEEE Transactions on Information Theory*, 61(4):1753 – 1767, April 2015.

[29] C. Geng, H. Sun, and S. Jafar. On the optimality of treating interference as noise: General message sets. *IEEE Transactions on Information Theory*, 61(7):3722–3736, July 2015.

[30] A. Gholami Davoodi and S. Jafar. Optimality of simple layered superposition coding in the 3 user MISO BC with finite precision CSIT. *IEEE Transactions on Information Theory*, 65(11):7181–7207, Nov 2019.

[31] A. Gholami Davoodi and S. Jafar. Degrees of freedom region of the (M, N_1, N_2) MIMO broadcast channel with partial CSIT: An application of sum-set inequalities based on aligned image sets. *IEEE Transactions on Information Theory*, 66(10):6256–6279, 2020.

[32] T. Gou and S. A. Jafar. Capacity of a class of symmetric SIMO Gaussian interference channels within O (1). *IEEE Trans. on Information Theory*, 57(4):1932–1958, April 2011.

[33] T. Gou, C. Wang, and S. Jafar. Toward full-duplex multihop multiflow – a study of non-layered two unicast wireless networks. *IEEE Journal on Selected Areas in Communications*, 32(9):1738–1751, 2014.

[34] T. Gou, C. Wang, S. Jafar, S. Jeon, and S. Chung. Aligned interference neutralization and the degrees of freedom of the $2 \times 2 \times 2$ interference channel. *IEEE Trans. on Information Theory*, 58(7):4381–4395, July 2012.

[35] T. Han and K. Kobayashi. A new achievable rate region for the interference channel. *IEEE Trans. Inform. Theory*, 27(1):49–60, Jan 1981.

[36] C. Huang, V. Cadambe, and S. Jafar. Interference alignment and the generalized degrees of freedom of the X channel. *IEEE Transactions on Information Theory*, 58(8):5130–5150, August 2012.

[37] W. Huleihel and Y. Steinberg. Channels with cooperation links that may be absent. *IEEE Transactions on Information Theory*, 63(9):5886–5906, Sep. 2017.

[38] I. Issa, S. L. Fong, and A. S. Avestimehr. Two-hop interference channels: Impact of linear schemes. *IEEE Transactions on Information Theory*, 61(10):5463–5489, 2015.

[39] S. Jafar and S. Shamai. Degrees of freedom region for the MIMO X channel. *IEEE Trans. on Information Theory*, 54(1):151–170, Jan. 2008.

[40] S. A. Jafar. Topological Interference Management through Index Coding. *IEEE Trans. on Inf. Theory*, 60(1):"529–568", Jan. 2014.

[41] S. Karmakar and M. K. Varanasi. The capacity region of the mimo interference channel and its reciprocity to within a constant gap. *IEEE Transactions on Information Theory*, 59(8):4781–4797, 2013.

[42] M. Kiamari, C. Wang, and A. S. Avestimehr. On heterogeneous coded distributed computing. In *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pages 1–7, Dec 2017.

[43] A. Lapidoth and S. Shamai. Collapse of degrees of freedom in MIMO broadcast with finite precision CSI. In *Proceedings of 43rd Annual Allerton Conference on Communications, Control and Computing*, Sep 2005.

[44] 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.

[45] S. Li, M. A. Maddah-Ali, Q. Yu, and A. S. Avestimehr. A fundamental tradeoff between computation and communication in distributed computing. *IEEE Transactions on Information Theory*, 64(1):109–128, Jan 2018.

[46] M. Maddah-Ali, A. Motahari, and A. Khandani. Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis. In *IEEE Trans. on Information Theory*, pages 3457–3470, August 2008.

[47] S. Mohajer, S. Diggavi, C. Fragouli, and D. Tse. Approximate Capacity of a Class of Gaussian Interference-Relay Networks. *IEEE Trans. on Information Theory*, 57:2837–2864, May 2011.

[48] A. Motahari, S. Gharan, M. Maddah-Ali, and A. Khandani. Real interference alignment: Exploiting the potential of single antenna systems. *IEEE Transactions on Information Theory*, 60(8):4799–4810, Aug 2014.

[49] A. Motahari and A. Khandani. Capacity bounds for the Gaussian interference channel. *IEEE Transactions on Information Theory*, 55(2):620–643, Feb. 2009.

[50] N. Naderializadeh and A. S. Avestimehr. Itlinq: A new approach for spectrum sharing in device-to-device communication systems. In *2014 IEEE International Symposium on Information Theory*, pages 1573–1577, 2014.

[51] K. Narra, Z. Lin, M. Kiamari, S. Avestimehr, and M. Annavaram. Slack squeeze coded computing for adaptive straggler mitigation. *CoRR*, abs/1904.07098, 2019.

[52] B. Nazer and M. Gastpar. Computation over multiple-access channels. *IEEE Transactions on Information Theory*, 53(10):3498–3516, 2007.

[53] U. Niesen and M. A. Maddah-Ali. Interference alignment: From degrees-of-freedom to constant-gap capacity approximations. *IEEE Trans. on Information Theory*, 59(8):4855–4888, Aug. 2013.

[54] H. Nikbakht, M. Wigger, and S. S. Shitz. Mixed delay constraints in wyner's soft-handoff network. In *2018 IEEE International Symposium on Information Theory (ISIT)*, pages 1171–1175, June 2018.

[55] S. Rini, D. Tuninetti, and N. Devroye. State of the cognitive interference channel: a new unified inner bound, and capacity to within 1.87 bits. *CoRR*, abs/0910.3028, 2009.

[56] A. Sanderovich, O. Somekh, H. V. Poor, and S. Shamai. Uplink macro diversity of limited backhaul cellular network. *IEEE Transactions on Information Theory*, 55(8):3457–3478, Aug 2009.

[57] X. Shang, G. Kramer, and B. Chen. A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels. *IEEE Transactions on Information Theory*, 55(2):689–699, Feb. 2009.

[58] I. Shomorony and A. S. Avestimehr. Degrees of freedom of two-hop wireless networks: Everyone gets the entire cake. *IEEE Transactions on Information Theory*, 60(5):2417–2431, 2014.

[59] I. Shomorony and S. Avestimehr. Two-unicast wireless networks: Characterizing the degrees of freedom. *IEEE Transactions on Information Theory*, 59(1):353–383, 2013.

[60] O. Simeone, N. Levy, A. Sanderovich, O. Somekh, B. Zaidel, H. Poor, and S. Shamai. Cooperative wireless cellular systems: an information theoretic view. In *Foundations and Trends in Communication and Information Theory*, volume 8, pages 1–177, 2012.

[61] O. Simeone, O. Somekh, Y. Bar-Ness, H. V. Poor, and S. Shamai. Capacity of linear two-hop mesh networks with rate splitting, decode-and-forward relaying and cooperation. *roceedings of the 45th Annual Allerton Conference on Communication, Control and Computing, Monticello, IL*, abs/0710.2553, Sep. 2007.

[62] C. Suh and D. Tse. Feedback capacity of the Gaussian interference channel to within 1.7075 bits: the symmetric case. *ArXiv:0901.3580*, abs/0901.3580, 2009.

[63] 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*, Submitted August 2012. Available online: http://arxiv.org/abs/1208.5071.

[64] C. Wang, T. Gou, and S. Jafar. Multiple unicast capacity of 2-source 2-sink networks. *CoRR*, abs/1104.0954, 2011.

[65] I. Wang and D. N. C. Tse. Interference mitigation through limited transmitter cooperation. *IEEE Transactions on Information Theory*, 57(5):2941–2965, May 2011.

[66] J. Wang, B. Yuan, L. Huang, and S. A. Jafar. Sum-gdof of 2-user interference channel with limited cooperation under finite precision CSIT. *IEEE Transactions on Information Theory*, 66(11):6999–7021, 2020.

[67] H. Weingarten, Y. Steinberg, and S. Shamai. The capacity region of the Gaussian MIMO broadcast channel. *IEEE Trans. on Information Theory*, 52:3936–3964, Sep. 2006.

[68] X. Yi and G. Caire. Itlinq+: An improved spectrum sharing mechanism for device-to-device communications. In *2015 49th Asilomar Conference on Signals, Systems and Computers*, pages 1310–1314, 2015.

[69] X. Yi and G. Caire. Optimality of Treating Interference as Noise: A Combinatorial Perspective. *IEEE Transactions on Information Theory*, 62(8):4654–4673, 2016.

[70] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr. Straggler mitigation in distributed matrix multiplication: Fundamental limits and optimal coding. *arXiv preprint arXiv:1801.07487*, 2018.

[71] B. Yuan and S. Jafar. Elevated multiplexing and signal space partitioning in the 2 user MIMO IC with partial CSIT. *IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2016.

Appendix A

Appendix of Chapter 3

A.1 Proof of Lemma 3.1 as a special case of Theorem 3 of [20]

For ease of reference, let us copy here Theorem 3 of [20], but since we only need a special case, let us make a few substitutions into the general statement of Theorem 3 of [20] before we copy it here. Let l, M be arbitrary positive integers, and let $m(1), m(2), \cdots, m(l) \in [M]$ be arbitrary elements of $[M]$ that are in decreasing order, i.e., $m(1) \geq m(2) \geq \cdots \geq m(l)$. Let us substitute $I_j = \{m(j)\}$ in the statement of Theorem 3 of [20]. Furthermore, we specialize $L_j^{\vec{\gamma}\vec{\delta}}(x_1, x_2, \cdots, x_N) = x_j$ for some $j \in [N]$. With these substitutions, the following special case of Theorem 3 of [20] is obtained.

Theorem A.1. *[Special case of Theorem 3 of [20]] Consider M non-negative real numbers $\lambda_1, \cdots, \lambda_M$ and random variables $X_k(t) \in \mathcal{X}_{\lambda_1+\lambda_2+\cdots+\lambda_M}$, $k \in [K]$, $t \in \mathbb{N}$, and another*

random variable W, all independent of $\mathcal{G}$, and define

$$\begin{aligned}
Z(t) &= \sum_{k=1}^{K} \lfloor G_k(t) X_k(t) \rfloor, \\
Z_1(t) &= (X_{k_1}(t))_{\lambda_1+\lambda_2+\cdots+\lambda_{m(1)-1}}^{\lambda_1+\lambda_2+\cdots+\lambda_{m(1)}} \text{ for some } k_1 \in [K], \\
Z_2(t) &= (X_{k_2}(t))_{\lambda_1+\lambda_2+\cdots+\lambda_{m(2)-1}}^{\lambda_1+\lambda_2+\cdots+\lambda_{m(2)}} \text{ for some } k_2 \in [K], \\
&\vdots \\
Z_l(t) &= (X_{k_l}(t))_{\lambda_1+\lambda_2+\cdots+\lambda_{m(l)-1}}^{\lambda_1+\lambda_2+\cdots+\lambda_{m(l)}} \text{ for some } k_l \in [K].
\end{aligned}$$

Then

$$H(Z^{[n]} \mid W, \mathcal{G}) \geq H(Z_1^{[n]}, Z_2^{[n]}, \cdots, Z_l^{[n]} \mid W) + n \times o(\log P), \tag{A.1}$$

if for each $s \in \{1, 2, \cdots, l-1\}$,

$$\lambda_{m(s+1)} + \lambda_{m(s+2)} + \cdots + \lambda_{m(l)} \leq \lambda_1 + \lambda_2 + \cdots + \lambda_{(m(s)-1)}. \tag{A.2}$$

Condition (A.2) of Theorem A.1 has a similar interpretation as conditions (3.27)-(3.30) of Lemma 3.1. Note that sub-section Z_s has level $\ell(Z_s) = \lambda_1 + \lambda_2 + \cdots + \lambda_{m(s)-1}$, and size $\mathcal{T}(Z_s) = \lambda_{m(s)}$. If we vertically stack $Z_1, Z_2, \cdots, Z_l$ in that order from top to bottom, then the LHS of (A.2) is the height of the block Z_s above the ground, i.e., its new height. The RHS of (A.2) is the height of Z_s as it appears in Z, i.e., its original height. Thus, Condition (A.2) simply checks that if $Z_1, \cdots, Z_l$ are stacked vertically in that order from top to bottom, then no sub-section is elevated above its original height in Z.

Next, to see how Lemma 3.1 is implied by Theorem A.1, let us start with sub-sections $U_1, U_2, \cdots, U_m$ that satisfy conditions (3.27)-(3.30), and map the parameters of Lemma 3.1 to the parameters of Theorem A.1. Figure A.1 presents an example that will be illustrative

for this purpose.

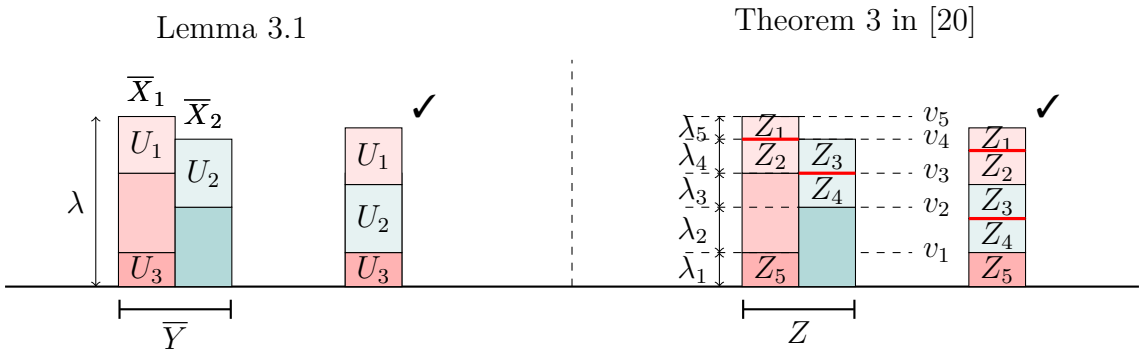

Figure A.1: *Starting from U_1, U_2, U_3 that satisfy the conditions for Lemma 3.1, the figure shows how the parameters are mapped to Theorem 3 of [20] so that the condition* (A.2) *is satisfied. Here, $K = 2, l = 5, M = 5, (m(1), m(2), m(3), m(4), m(5)) = (5, 4, 4, 3, 1), I_1 = \{5\}, I_2 = \{4\}, I_3 = \{4\}, I_4 = \{3\}, I_5 = \{1\}$. Note that in the vertical stacking each Z_i sub-section is not elevated above its original position in Z because in the corresponding vertical stacking each U_j sub-section is not elevated above its original position in $\overline{Y}$.*

Let $\lambda \triangleq \max_{k\in[K]} \mu_k$. Let the interval (a_j, b_j) correspond to the sub-section U_j, and let

$$V = \{a_j : j \in [m]\} \cup \{b_j : j \in [m]\} \cup \{0, \lambda\} \tag{A.3}$$

$$\triangleq \{v_0, v_1, v_2, \cdots, v_M\} \tag{A.4}$$

be the set containing all interval boundaries with $0 = v_0 < v_1 < v_2 < \cdots < v_M = \lambda$. Now we can define $\lambda_1 = v_1 - v_0, \lambda_2 = v_2 - v_1, \cdots, \lambda_M = v_M - v_{M-1}$.

Note that the intervals (v_{i-1}, v_i) partition the interval $(0, \lambda)$ into M contiguous sub-intervals, let us call them v-intervals, so that we have a total of M such v-intervals. If any sub-section U_j spans more than one of these M intervals, let us split it further into smaller sub-sections according to to the v-intervals. These splits are shown with red lines in Fig. A.1 that split each of U_1, U_2. These partitions are reflected in the vertical stack, and then the sub-sections in the vertical stack are labeled as $Z_1, Z_2, \cdots, Z_l$ from top to bottom. This completes the mapping of parameters. The construction trivially guarantees that the condition (A.2) is satisfied. Therefore, according to Theorem A.1, the sum-set inequality (A.1) must hold,

which in turn implies the sum-set inequality (3.25) of Lemma 3.1 holds as well.

A.2 Appendix: Basic Lemmas

The first lemma shows that in the GDoF sense, entropy (even with arbitrary conditioning) does not change if an integer valued random variable is replaced with another integer valued random variable as long as they are always within a constant distance of each other. Note that while $S(i), T(i)$ may depend on P, the RHS of (A.5) is bounded (does not scale with P) so it has 0 GDoF.

Lemma A.1. *For integer valued random variables $S(1), S(2), \cdots, S(n), T(1), T(2), \cdots, T(n)$ and any random variable Z, if $|S(i) - T(i)| \leq c$ for all $i \in [n]$, where c is a finite constant, then*

$$|H(S^{[n]} \mid Z) - H(T^{[n]} \mid Z)| \leq n \log(2c+1). \tag{A.5}$$

Proof. Since $H(T^{[n]} \mid Z) - H(T^{[n]} \mid S^{[n]}, Z) \leq H(S^{[n]} \mid Z) \leq H(T^{[n]} \mid Z) + H(S^{[n]} \mid T^{[n]}, Z)$ we have

$$\begin{aligned}
&|H(S^{[n]} \mid Z) - H(T^{[n]} \mid Z)| \\
&\quad \leq \max(H(S^{[n]} \mid T^{[n]}, Z), H(T^{[n]} \mid S^{[n]}, Z)) && \text{(A.6)} \\
&\quad = \max(H(S^{[n]} - T^{[n]} \mid T^{[n]}, Z), H(T^{[n]} - S^{[n]} \mid S^{[n]}, Z)) && \text{(A.7)} \\
&\quad \leq \max(H(S^{[n]} - T^{[n]}), H(T^{[n]} - S^{[n]})) && \text{(A.8)} \\
&\quad \leq n \log(2c+1). && \text{(A.9)}
\end{aligned}$$

The last step follows because the support of integer valued $S(i) - T(i)$ has cardinality no more than $2c+1$, and entropy of a discrete random variable is bounded by the logarithm of

the cardinality of its support. □

The next lemma shows that additional floor operations can only have a bounded effect, so for example, $\lfloor\sqrt{P^\alpha}c\overline{X}\rfloor$ is within bounded distance from $\lfloor c\lfloor\sqrt{P^\alpha}\overline{X}\rfloor\rfloor$ if c is bounded.

Lemma A.2. *For any $S, c \in \mathbb{R}$,*

$$\Big|\lfloor cS\rfloor - \lfloor c\lfloor S\rfloor\rfloor\Big| \leq (2 + |c|).$$

Proof. Note[1] that $\lfloor c\lfloor S\rfloor\rfloor \in \lfloor cS \pm |c|\rfloor \subset cS \pm (1 + |c|)$ and $\lfloor cS\rfloor \in cS \pm 1$. Therefore, $|\lfloor cS\rfloor - \lfloor c\lfloor S\rfloor\rfloor| \leq (2 + |c|)$. □

The next lemma shows that scaling by a real-valued constant bounded away from zero and infinity followed by a subsequent rounding operation does not change the entropy in the GDoF sense. Thus, for example the entropy $H(\overline{X})$ is equal to the entropy $H(\lfloor G\overline{X}\rfloor \mid \mathcal{G})$ because conditioned on $\mathcal{G}$, the value G is a bounded constant scaling factor.

Lemma A.3. *For any integer valued random variables $S(1), S(2), \cdots, S(n)$, any random variable Z, and constants $c(1), c(2), \cdots, c(n)$ such that $|c(i)| \in [1/\Delta, \Delta]$ for all $i \in [n]$,*

$$\Big|H(S^{[n]} \mid Z) - H(\big(\lfloor c^{[n]}S^{[n]}\rfloor\big) \mid Z)\Big| \leq n \times \log(5 + 2\Delta). \tag{A.10}$$

Proof.

$$\begin{aligned}
&H(S^{[n]} \mid Z) - H(\big(\lfloor c^{[n]}S^{[n]}\rfloor\big) \mid Z) \\
&\leq H(S^{[n]} \mid \big(\lfloor c^{[n]}S^{[n]}\rfloor\big), Z) && \text{(A.11)} \\
&= H\left(S^{[n]} - \left(\lfloor\frac{\lfloor c^{[n]}S^{[n]}\rfloor}{c^{[n]}}\rfloor\right) \mid \big(\lfloor c^{[n]}S^{[n]}\rfloor\big), Z\right) && \text{(A.12)} \\
&\leq n \times \log(5 + 2\Delta). && \text{(A.13)}
\end{aligned}$$

[1] We use the notation $y \pm z$ to denote the interval $[y - z, y + z]$.

In the last step we used Lemma A.1 and the fact that $|S(i) - \lfloor \frac{\lfloor c(i)S(i) \rfloor}{c(i)} \rfloor| \leq (2 + \frac{1}{|c(i)|}) \leq (2+\Delta)$. On the other hand,

$$H(\lfloor c^{[n]} S^{[n]} \rfloor \mid Z) - H(S^{[n]} \mid Z) \leq H(\lfloor c^{[n]} S^{[n]} \rfloor \mid S^{[n]}, Z) \tag{A.14}$$

$$= 0. \tag{A.15}$$

□

Appendix B

Appendix of Chapter 4

B.1 Proof for Lemma 4.1

Let us present the proof of the bound $I(W_1; Y_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \leq I(W_1; \overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]})$, which converts from the original canonical channel to the deterministic channel with conditioning on $\mathcal{G}_{[1:L]}$ instead of $\mathcal{G}_{[1:\ell]}$.[1] Following the proof that appears in the Appendix section of [16], the first step is to limit the input and output to integers. Let an intermediate deterministic channel model of the ℓ^{th} hop have integer inputs $\lfloor \overline{P}^{\max(1,\alpha)} X_{1[\ell]}^{[N]} \rfloor$, $\lfloor \overline{P}^{\max(1,\alpha)} X_{2[\ell]}^{[N]} \rfloor$, and integer outputs[2]

$$\begin{aligned}
\bar{\bar{Y}}_{1[\ell]}(n) &= \lfloor \overline{P}^{1-\max(1,\alpha)} G_{11[\ell]}(n) \lfloor \overline{P}^{\max(1,\alpha)} X_{1[\ell]}(n) \rfloor \rfloor \\
&\quad + \lfloor \overline{P}^{\alpha-\max(1,\alpha)} G_{12[\ell]}(n) \lfloor \overline{P}^{\max(1,\alpha)} X_{2[\ell]}(n) \rfloor \rfloor \qquad \text{(B.1)} \\
\bar{\bar{Y}}_{2[\ell]}(n) &= \lfloor \overline{P}^{\alpha-\max(1,\alpha)} G_{21[\ell]}(n) \lfloor \overline{P}^{\max(1,\alpha)} X_{1[\ell]}(n) \rfloor \rfloor \\
&\quad + \lfloor \overline{P}^{1-\max(1,\alpha)} G_{22[\ell]}(n) \lfloor \overline{P}^{\max(1,\alpha)} X_{2[\ell]}(n) \rfloor \rfloor \qquad \text{(B.2)}
\end{aligned}$$

[1] Since the channels $\mathcal{G}_{[\ell+1:L]}$ do not appear until after the ℓ^{th} hop, the conditioning can be reduced to $\mathcal{G}_{[1:\ell]}$ trivially.

[2] By integer-valued inputs and outputs, we mean that the real and imaginary parts of these inputs and outputs are integer valued.

while the original canonical channel is

$$
\begin{aligned}
Y_{1[\ell]}(n) &= \overline{P}^{1-\max(1,\alpha)} G_{11[\ell]}(n)(\overline{P}^{\max(1,\alpha)} X_{1[\ell]}(n)) \\
&\quad + \overline{P}^{\alpha-\max(1,\alpha)} G_{12[\ell]}(n)(\overline{P}^{\max(1,\alpha)} X_{2[\ell]}(n)) + Z_{1[\ell]}(n) \qquad \text{(B.3)} \\
Y_{2[\ell]}(n) &= \overline{P}^{\alpha-\max(1,\alpha)} G_{21[\ell]}(n)(\overline{P}^{\max(1,\alpha)} X_{1[\ell]}(n)) \\
&\quad + \overline{P}^{1-\max(1,\alpha)} G_{22[\ell]}(n)(\overline{P}^{\max(1,\alpha)} X_{2[\ell]}(n)) + Z_{2[\ell]}(n) \qquad \text{(B.4)}
\end{aligned}
$$

Define

$$
E_{i[\ell]}^{[N]} \triangleq Y_{i[\ell]}^{[N]} - \bar{\bar{Y}}_{i[\ell]}^{[N]}. \qquad \text{(B.5)}
$$

Then we have,

$$
\begin{aligned}
& I(W_1; Y_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \\
&= I(W_1; \bar{\bar{Y}}_{1[\ell]}^{[N]} + E_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \qquad \text{(B.6)} \\
&\le I(W_1; \bar{\bar{Y}}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + I(W_1; E_{1[\ell]}^{[N]} \mid \bar{\bar{Y}}_{1[\ell]}^{[N]}, \mathcal{G}_{[1:L]}) \qquad \text{(B.7)} \\
&\le I(W_1; \bar{\bar{Y}}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + h(E_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \\
&\qquad - h(E_{1[\ell]}^{[N]} \mid W_1, X_{1[\ell]}^{[N]}, X_{2[\ell]}^{[N]}, \bar{\bar{Y}}_{1[\ell]}^{[N]}, \mathcal{G}_{[1:L]}) \qquad \text{(B.8)} \\
&= I(W_1; \bar{\bar{Y}}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + h(E_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \\
&\qquad - h(Z_{1[\ell]}^{[N]} \mid W_1, X_{1[\ell]}^{[N]}, X_{2[\ell]}^{[N]}, \bar{\bar{Y}}_{1[\ell]}^{[N]}, \mathcal{G}_{[1:L]}) \qquad \text{(B.9)} \\
&= I(W_1; \bar{\bar{Y}}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + h(E_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) - h(Z_{1[\ell]}^{[N]}) \qquad \text{(B.10)} \\
&\le I(W_1; \bar{\bar{Y}}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + \sum_{n=1}^{N} h(E_{1[\ell]}(n) \mid \mathcal{G}_{[1:L]}) - h(Z_{1[\ell]}(n)) \qquad \text{(B.11)} \\
&= I(W_1; \bar{\bar{Y}}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + No(\log(P)) \qquad \text{(B.12)}
\end{aligned}
$$

where (B.8) is because conditioning cannot increase differential entropy, (B.10) holds because

the the noise is independent of all messages and channel coefficients, i.e.,

$$I(Z_{1[\ell]}^{[N]}; W_1, X_{1[\ell]}^{[N]}, X_{2[\ell]}^{[N]}, \mathcal{G}_{[1:L]}) = 0. \tag{B.13}$$

For (B.12) note that according to its definition in (B.5), $E_{1[\ell]}(n)$ is the sum of an AWGN term ($Z_{1[\ell]}(n)$) of unit variance, and an independent term whose magnitude is at most $2 + 2\Delta$. Therefore, the variance of $E_{i[\ell]}(n)$ is not more than $(2 + 2\Delta)^2 + 1$. For a given variance constraint Gaussians maximize differential entropy, so the differential entropy $h(E_{i[\ell]}(n) \mid \mathcal{G}_{[1:L]})$ is not more than $\log(2\pi e(1 + (2 + 2\Delta)^2))$, which is $o(\log(P))$. Similarly, $h(Z_{1[\ell]}(n))$ is equal to $\log(2\pi e)$ which is also $o(\log(P))$. So the difference between $I(W_1; Y_{1[N]}^{[N]} \mid \mathcal{G}_{[1:L]})$ and $I(W_1; \bar{\bar{Y}}_{1[N]}^{[N]} \mid \mathcal{G}_{[1:L]})$ approaches 0 when normalized by $N \log(P)$.

Thus, the integer input and output channel with the per-codeword power constraints,

$$\sum_{n=1}^{N} \left((\lfloor \overline{P}^{\max(1,\alpha)} X_{1[\ell]}(n) \rfloor)^2 \right) \leq N P^{\max(1,\alpha)} \tag{B.14}$$

$$\sum_{n=1}^{N} \left((\lfloor \overline{P}^{\max(1,\alpha)} X_{2[\ell]}(n) \rfloor)^2 \right) \leq N P^{\max(1,\alpha)} \tag{B.15}$$

achieves at least the same GDoF as the original canonical channel model.

The next step is to convert the per-codeword power constraints into per-symbol power constraints. Let us define

$$\overline{X}_{1[\ell]}(n) \triangleq \lfloor \overline{P}^{\max(1,\alpha)} X_{1[\ell]}(n) \rfloor \mod \lceil \overline{P}^{\max(1,\alpha)} \rceil \tag{B.16}$$

$$\overline{X}_{2[\ell]}(n) \triangleq \lfloor \overline{P}^{\max(1,\alpha)} X_{2[\ell]}(n) \rfloor \mod \lceil \overline{P}^{\max(1,\alpha)} \rceil \tag{B.17}$$

$$\begin{aligned} \overline{Y}_{1[\ell]}(n) \triangleq & \lfloor \overline{P}^{1-\max(1,\alpha)} G_{11[\ell]}(n) \overline{X}_{1[\ell]}(n) \rfloor \\ & + \lfloor \overline{P}^{\alpha-\max(1,\alpha)} G_{12[\ell]}(n) \overline{X}_{2[\ell]} \rfloor \end{aligned} \tag{B.18}$$

$$\begin{aligned} \overline{Y}_{2[\ell]}(n) = & \lfloor \overline{P}^{\alpha-\max(1,\alpha)} G_{21[\ell]}(n) \overline{X}_{1[\ell]}(n) \rfloor \\ & + \lfloor \overline{P}^{1-\max(1,\alpha)} G_{22[\ell]}(n) \overline{X}_{2[\ell]} \rfloor \end{aligned} \tag{B.19}$$

$$\hat{Y}_{1[\ell]}(n) = \bar{\bar{Y}}_{1[\ell]}(n) - \overline{Y}_{1[\ell]}(n) \tag{B.20}$$

$$\hat{Y}_{2[\ell]}(n) = \bar{\bar{Y}}_{2[\ell]}(n) - \overline{Y}_{2[\ell]}(n) \tag{B.21}$$

Now we have,

$$I(W_1; \bar{\bar{Y}}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}^{[N]}) = I(W_1; \overline{Y}_{1[\ell]}^{[N]} + \hat{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \tag{B.22}$$

$$\leq I(W_1; \overline{Y}_{1[\ell]}^{[N]}, \hat{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \tag{B.23}$$

$$\leq I(W_1; \overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + H(\hat{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \tag{B.24}$$

$$\leq I(W_1; \overline{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) + H(\hat{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[\ell]}). \tag{B.25}$$

It can now be shown that the term $H(\hat{Y}_{1[\ell]}^{[N]} \mid \mathcal{G}_{[\ell]})$ is negligible in the GDoF sense following the same proof as in [16, eq. (124)-(149)]. Therefore, at the ℓ^{th} hop, replacing the long-term (per-codeword) power constraint with short-term (per-symbol) power constraint will not reduce the GDoF value.

The other bound $I(W_2; Y_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \leq I(W_2, \overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]})$ follows by symmetry. The bound $I(W_1, W_2; Y_{1[\ell]}^{[N]}, Y_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]}) \leq I(W_1, W_2; \overline{Y}_{1[\ell]}^{[N]}, \overline{Y}_{2[\ell]}^{[N]} \mid \mathcal{G}_{[1:L]})$ that is needed in (4.38) is obtained similarly.

□