Interference Management in MIMO Wireless Networks

by

Akbar Ghasemi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The scarce and overpopulated radio spectrum is going to present a major barrier to the growth and development of future wireless networks. As such, spectrum sharing seems to be inevitable to accommodate the exploding demand for high data rate applications. A major challenge to realizing the potential advantages of spectrum sharing is interference management. This thesis deals with interference management techniques in non-cooperative networks. In specific, interference alignment is used as a powerful technique for interference management. We use the degrees of freedom (DoF) as the figure of merit to evaluate the performance improvement due to the interference management schemes.

This dissertation is organized in two parts. In the first part, we consider the K-user multiple input multiple output (MIMO) Gaussian interference channel (IC) with M antennas at each transmitter and N antennas at each receiver. This channel models the interaction between K transmitter-receiver pairs sharing the same spectrum for data communication. It is assumed that the channel coefficients are constant and are available at all nodes prior to data transmission. A new cooperative upper-bound on the DoF of this channel is developed which outperforms the known bounds. Also, a new achievable transmission scheme is provided based on the idea of interference alignment. It is shown that the achievable DoF meets the upper-bound when the number of users is greater than a certain threshold, and thus it reveals the channel DoF.

In the second part, we consider communication over MIMO interference and X channels in a fast fading environment. It is assumed that the transmitters obtain the channel state information (CSI) after a finite delay which is greater than the coherence time of the channel. In other words, the CSI at the transmitters becomes outdated prior to being exploited for the current transmission. New transmission schemes are proposed which exploit the knowledge of the past CSI at the transmitters to retrospectively align interference in the subsequent channel uses. The proposed transmission schemes offer DoF gain compared to having no CSI at transmitters. The achievable DoF results are the best known results for these channels. Simple cooperative upper-bounds are developed to prove the tightness of our achievable results for some network configurations.

Acknowledgements

I would like to express my sincere gratitude to my advisor, Professor Amir K. Khandani for his continuous guidance and encouragement throughout the course of my graduate studies. I have had a great opportunity to work with him and have learned a lot from him throughout the course of my PhD research. I wish to thank him for his support which was not limited to research work.

I would like to thank the members of my dissertation committee, Professors Saeed Gazor, Cameron Stewart, Patrick Mitran, and Oleg Michailovich, for taking the time out of their schedules to review my thesis and for providing me with their questions and comments. I would like to especially thank Professor Cameron Stewart for helpful discussion concerning the number theoretic tools used in this work.

I would like to specially acknowledge my friends, Dr. Seyed Abolfazl Motahari and Dr. Mohammad Javad Abdoli for their collaboration and fruitful discussions we had on this research. I also would like to thank Dr. Kamyar Moshksar and Javad Behrouzi Moghaddam for the joint work on other interesting research topics. I would like to extend my gratitude to all members of coding and signal transmission lab, who helped create a very pleasant and friendly learning environment. My warm wishes and thanks go as well to my dear friends Siamak, Hossein, Ghasem, Babak, Bahador, Mohsen, Hamidreza, Pooya, Mahmoud, Shahab, Alireza, Hadi, Mohammad, Mehdi, Ali, Payam, Moein, Ramin, Ehsan, Ershad and many other friends who have enriched my life beyond my studies.

I would like to thank my parents as well my brothers and sisters for their support and encouragement throughout my life. Last but not least, I would like to express my warmest gratitude to my wife, Mitra, for her endless support, patience, and encouragements, without which I would not have finished this long journey of PhD studies.

To my beloved wife, Mitra,
and to my family.

Table of Contents

List of Tables **xv**

List of Figures **xvii**

List of Acronyms **xix**

List of Notations **xxii**

1 Introduction **1**

1.1 Channel state information (CSI) 3

1.2 Interference alignment with full CSI 3

1.3 Interference alignment with no/partial CSI 5

1.4 Dissertation Outline and Main Contributions 6

2 Interference management for constant MIMO Interference channels **9**

2.1 System Model 10

2.2 Main Results and Discussion 12

2.2.1 Upper-bound 12

2.2.2 Asymptotic interference alignment for the IC using rational dimensions 13
2.2.3 Comparing achievable and upper-bound results 18
2.2.4 Joint processing in collocated antennas 22
2.3 DoF upper-bound for the K-user MIMO interference channel 22
2.4 Interference alignment and DoF lower-bound for MIMO IC 25
2.4.1 Preliminaries on Real Interference Alignment 26
2.4.2 Proof of Theorem 2 29
2.5 Conclusions 38

3 Interference alignment with Delayed CSIT 39

3.1 System Model 40
3.2 Main Results and Discussion 43
3.2.1 Main Results 43
3.2.2 Discussion 46
3.3 Proof of Theorem 4 52
3.3.1 An Illustrative Example 52
3.3.2 Proof of Achievability for General Setting 54
3.4 Proof of Theorem 5 62
3.5 Proof of Theorem 6 65
3.6 Proof of Theorem 7 70
3.7 Proof of Theorem 8 73
3.8 Conclusion 74

4 Conclusion **75**

4.1 Future research directions 77

4.1.1 Combining Real Alignment with Signal Space Alignment 77

4.1.2 Real Interference Alignment with finite precision 78

4.1.3 Developing an upper-bound for the networks with distributed transmitters and Delayed CSIT 78

4.1.4 Characterizing the trade-off between the DoF gain and feedback overhead 79

APPENDICES **81**

A Appendices for Chapter 2 **83**

A.1 Proof of (2.19) 83

A.2 The closest rational neighbors of a real number with denominator at most K 85

A.3 Calculating the normalizing constant γ in (2.30) 87

B Appendices for Chapter 3 **89**

B.1 Derivation of rank of matrix $\mathbf{P}^{[k]}$ 89

B.2 Proof of achievability of T_3 for class $\mathcal{C}_8$ 94

B.3 Proof of achievability of T_4 for classes $\mathcal{C}_6$ and $\mathcal{C}_7$ 96

B.4 Proof of achievability of T_5 for classes $\mathcal{C}_8$ and $\mathcal{C}_9$ 97

B.5 Proof of achievability of T_7 for class $\mathcal{C}_9$ 98

B.6 Proof of achievability of T_8 for class $\mathcal{C}_{10}$ 99

References **101**

List of Tables

3.1 Achievable DoF region of the (M_1, M_2, N_1, N_2) MIMO IC with delayed CSIT when $N_2 \geq N_1$. 49

3.2 DoF of the MIMO X Channel with M antennas at each transmitter and N antennas at each receiver 52

3.3 Appropriate parameters for our transmission scheme for each corner point in class $\mathcal{C}_3 - \mathcal{C}_{10}$ 61

List of Figures

2.1 Our achievable and upper-bound on normalized DoF of a $(K, M \times N)$ IC for $K = 5$ and $K = 10$. 14

2.2 Splitting each transmitter and receiver of a MIMO IC into multiple single antenna nodes. Note that in the resulting network there is an independent massage from a transmitter to a receiver if and only if there was desired communication between them in the original network. 15

2.3 Real Interference alignment for the $(K, M \times N)$ IC: the transmit signal of each user is composed of N independent parts which are depicted here by adjacent rectangles. By the real interference alignment, the squares in each column at the receiver side are approximately aligned. 19

2.4 Comparing different results on the DoF of the $(3, M \times N)$ MIMO interference channel where $\beta = \frac{\max(M,N)}{\min(M,N)}$ 21

2.5 Typical shape of function $G(\rho)$ in (2.21). 24

3.1 The two-user MIMO IC and X channel with delayed CSIT 41

3.2 The achievable DoF region for the two-user MIMO IC with $N_2 \geq N_1$ and delayed CSIT (solid line). The DoF region of the same channel with no CSIT (dash-dot line) and full CSIT (dashed line) are also presented for comparison. 48

3.3 The DoF region of the two-user MIMO IC with $M_1 = M_2 = 2$ and $N_1 = N_2 = 1$ and under different assumption on CSIT information: full CSIT (dashed line), delayed CSIT (solid line), no CSIT (dash-dot line) 54

3.4 Outer-bounds on the DoF region of the two-user MIMO IC with $N_2 \geq N_1$ and delayed CSIT for different classes: BC outer-bound and the full CSIT outer-bound are respectively represented by dash-dot and dashed lines. Our achievable DoF region (solid line) is also presented for comparison. 64

A.1 Constructing a Farey sequence using the Stern-Brocot tree. The fractions in layer k form the Farey sequence $\mathcal{F}_k$. 86

List of Acronyms

AWGN	Additive White Gaussian Noise
BC	Broadcast Channel
CSI	Channel State Information
CSIT	Channel State Information at Transmitter
DoF	Degrees of Freedom
HK	Han and Kobayashi
IA	Interference Alignment
IC	Interference Channel
i.i.d.	Independent and Identically Distributed
MAC	Multiple Access Channel
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
SISO	Single-Input Single-Output
SNR	Signal to Noise Ratio
ZF	Zero-Forcing

List of Notations

Boldface Upper-Case Letters	Matrices (or Vectors)
Caligraphic letters	Sets
$\mathbf{A}^T$	Transpose of $\mathbf{A}$
$\mathbf{A}^\dagger$	Conjugate transpose of $\mathbf{A}$
$\mathcal{A} \cup \mathcal{B}$	Union of the sets $\mathcal{A}$ and $\mathcal{B}$
$\mathcal{A} \cap \mathcal{B}$	Intersection of the sets $\mathcal{A}$ and $\mathcal{B}$
$\mathcal{A} \setminus \mathcal{B}$	Difference of the sets $\mathcal{A}$ and $\mathcal{B}$
$\mathbb{C}$	The set of complex numbers
$\mathcal{C}$	The capacity region
$\mathcal{CN}(0, \sigma^2)$	Complex Gaussian distribution with mean 0 and variance σ^2
$\mathcal{D}$	DoF region
DoF	DoF (Sum-DoF)
$\underline{\mathsf{DoF}}$	Achievable DoF
$\mathbb{E}[X]$	The expected value of the random variable X
$f(n) = o(g(n))$	$\lim\limits_{n \to \infty} \frac{f(n)}{g(n)} = 0$
gcd	The greatest common divisor
K	Number of users
$\mathcal{K}$	$\{1, 2, \cdots, K\}$
M	Number of transmit antennas
$\mathcal{M}$	$\{1, 2, \cdots, M\}$

N	Number of receive antennas
$\mathcal{N}$	$\{1, 2, \cdots, N\}$
$\mathbb{R}$	The set of real numbers
$\mathbf{R}$	Rate vector
t	Time index
$\lfloor x \rfloor$	The largest integer not greater than x
$\lceil x \rceil$	The smallest integer not less than x
$\mathbb{Z}$	The set of integers

Chapter 1

Introduction

The ever increasing demand for high data rate transmission has stimulated extensive research in the past few decades to push the spectral efficiency of the point to point systems closer and closer to the celebrated Shannon limit [1]. Fifty years of effort and invention have finally led to the transmission schemes that closely approach this limit at the cost of increasing processing power per information bit [2]. Establishing similar performance limits for multi-user communication networks turns out to be challenging. One of the most fundamental, and yet so far elusive, channels in multi-user information theory is the interference channel (IC). IC models a communication system with several transmitter-receiver pairs, in which each transmitter wishes to communicate with its corresponding receiver while generating interference to all other receivers. Characterizing the capacity region of the IC is one of the long-standing open problems in information theory. Even for the simplest case of two user Gaussian IC, which was first considered in [3], the full characterization of the capacity region is still unknown. In fact, the capacity region of this channel has been characterized only for some ranges of channel coefficients [4–10]. For the general two-user case, a characterization of the capacity region within one bit has been presented in [11].

To increase the bit rate in wireless systems without increasing the bandwidth or power budget, the use of multiple input multiple output (MIMO) systems is a common practice. It is well known that MIMO systems can provide substantial improvement in terms of diversity and/or multiplexing gains. The capacity region of the two-user MIMO IC has been characterized in [12] and [13] to within a constant gap.

By moving from the two-user case to *more than two users*, the capacity characterization of the IC becomes more challenging. To reduce the severe effect of the interference for $K > 2$ users, the use of a new technique known as *interference alignment* is essential [14–16]. Interference alignment is an elegant technique that reduces the effect of the aggregated interference from several users to that of a single user. This is accomplished by assigning a portion of the available time/frequency/space at each receiver to the interference and enforcing all the interfering terms to be received in that portion. There are two versions of interference alignment in the literature: *signal space alignment* and *signal scale alignment*. In signal space alignment, the transmit signal of each user is a linear combination of some vectors where data determines the coefficients of this linear combination. In this approach, interference alignment involves the design of the appropriate vectors for different users such that: i) the interfering terms at each receiver are squeezed into a subspace of the available signal space at that receiver, and ii) the interference subspace can be separated from the desired signal subspace. Signal space alignment is applicable to ICs with multiple antennas or ICs with time varying/frequency selective channel coefficients. Signal scale alignment, on the other hand, uses structured coding, e.g., lattice codes, to align interference at the signal level and is particularly useful for the case of single antenna constant IC (not varying with time/frequency). For the fully connected K-user Gaussian IC ($K > 2$), most of the effort has focused on the characterization of the degrees of freedom (DoF). The DoF for a Gaussian network shows the pre-log factor of the sum-capacity in the limit of increasing Signal to Noise Ratio (SNR). More precisely, the channel sum-capacity (C_Σ) and the channel DoF (DoF) are related to each other by the following relationship:

$$C_\Sigma(\mathrm{SNR}) = \mathsf{DoF}\log_2(\mathrm{SNR}) + o(\log_2(\mathrm{SNR})). \tag{1.1}$$

1.1 Channel state information (CSI)

In its original form [14, 16], interference alignment requires the perfect and instantaneous CSI (simply referred to as full CSI) at all nodes to reveal its full potential. In specific, each transmitter needs to adjust its transmitted signal according to the current CSI to align interference in unintended receivers. It is commonly assumed that the receivers can obtain the CSI through channel estimation phase. The access of transmitters to CSI (CSIT) is generally through the feedback links and is subject to delay and quantization error. It is well known that the quantization error of the CSI due to the finite rate of the feedback links has negligible effect on the performance as long as the feedback bit rate scales sufficiently fast with SNR [17, 18]. The impact of CSIT delay on the performance is more substantial especially in fast fading environment. Specifically, if the feedback delay exceeds the channel coherence time, the CSIT expires prior to the beginning of each channel use and therefore it will be outdated. In the following, we consider the possibility of interference alignment under different assumptions about the CSIT knowledge.

1.2 Interference alignment with full CSI

In [40], Host-Madsen and Nosratinia showed that the DoF of the fully connected K-user Gaussian IC with full CSI is less than or equal to $\frac{K}{2}$. They also conjectured that the DoF of this channel is less than or equal to unity regardless of the number of users when the channel coefficients are constant.

For the case of varying channel coefficients, Cadambe and Jafar in [16] showed that a fully connected K-user Gaussian IC has $\frac{K}{2}$ degrees of freedom, i.e., each user can enjoy half of its available DoF in spite of interfering signals from other users. They also showed in [19] that the DoF of the $M \times N$ X channel with single-antenna nodes and varying channel coefficients is given by $\frac{MN}{M+N-1}$. The achievability scheme of these works is based on the signal space interference alignment.

For the constant channel coefficients, Bresler et al. in [20] computed the capacity region of the many-to-one and one-to-many Gaussian ICs within constant number of bits. In their achievability scheme for the many-to-one Gaussian IC, they introduced the signal scale interference alignment technique. In [21], using the signal scale interference alignment, the authors reported a class of fully connected real constant K-user Gaussian ICs with DoF arbitrarily close to $\frac{K}{2}$. Etkin and Ordentlich in [22] used some results of additive combinatorics to show that for a constant fully connected real Gaussian IC, the DoF is very sensitive to the rationality/irrationality of channel coefficients. They showed that for a fully connected constant real Gaussian IC with rational channel coefficients, the DoF is strictly less than $\frac{K}{2}$. Moreover, they showed that for a class of measure zero of channel coefficients, the DoF is equal to $\frac{K}{2}$. Independently, Motahari *et al.* showed in [23] that for a three-user constant symmetric real Gaussian IC with irrational channel coefficients, the DoF is equal to $\frac{3}{2}$. However, their assumption regarding the channel symmetry restricted its scope to a subset of measure zero of all possible channel coefficients. For a constant Gaussian IC with complex channel coefficients, Cadambe *et al.* in [24] showed that the Host-Madsen and Nosratinia conjecture is not true. By introducing asymmetric complex signaling, they proved that the K-user complex Gaussian IC with constant coefficients has at least 1.2 DoF for almost all values of channel coefficients. Recently, Motahari *et al.* settled the problem in general case by proposing a new type of signal scale interference alignment that can achieve $\frac{K}{2}$ DoF for almost all K-user real Gaussian ICs with constant coefficients [25, 26]. The essence of this new method, called *real alignment*, is to align discrete points along a real axis based on some number-theoretic properties of rational and irrational numbers [26].

Using the results of [16] and [26], one can observe that in a K-user $M \times N$ MIMO interference channel, everyone gets half the cake (the cake being the DoF of a user in the absence of interference) when antenna configuration is symmetric, i.e. $M = N$. For the general K-user $M \times N$ MIMO IC, [27] proved that the DoF per user are even better, i.e., everyone gets $\frac{\beta}{\beta+1}$ (which is greater than or equal to half) of the cake when $\beta = \frac{\max(M,N)}{\min(M,N)}$

is an integer and channels are time-varying. In this study, it is proved that the conclusion of [27] is still applicable even with constant channels and non-integer values of β. To this end, new achievable and upper-bound results are developed.

1.3 Interference alignment with no/partial CSI

The availability of perfect and instantaneous CSI at the receivers can be realized in practice by accurate channel estimation techniques. The full CSI at the transmitters, however, is practically hard to obtain. To overcome this problem, one needs to consider the possibility of IA with no/partial CSIT. Considering DoF as the performance measure, it has been approved that with independent and identically distributed (i.i.d.) Rayleigh fading channel coefficients across time and space, interference alignment is not possible with no CSIT for multi-user channels such as MIMO broadcast channel (BC) [28, 29] , IC, and X channel [28–30]. On the other hand, when the channel coefficients are correlated, the possibility of interference alignment with no CSIT has been demonstrated in [31].

Recently, in [32], Maddah-Ali and Tse introduced a new model for the availability of CSI in the context of multiple input single output (MISO) BC which is interesting from both theoretical and practical standpoints. In this model, which is commonly referred to as *delayed CSIT* model, channel coefficients experience i.i.d. fading across antennas and channel uses. Moreover, each receiver knows its own channel matrices perfectly and instantaneously while all other nodes know it with a finite delay. The remarkable finding of [32] is that the DoF of the MISO BC channel with delayed CSIT can be strictly greater than one, which is the DoF with no CSIT. In other words, even completely outdated CSIT can be exploited to attain DoF gain. Unlike the BC, in networks with distributed transmitters and receivers such as interference and X channels, a fundamental constraint is that each transmitter has only access to its own information symbols. This constraint turns out to be a major bottleneck in exploiting the knowledge of the past CSI at transmitters to achieve a DoF gain. The two-user and three-user MIMO BC with delayed CSIT have been investigated in [33] and [34], respectively. For the two-user single-input single-output (SISO) X channel and three-user SISO IC with delayed CSIT, DoF improvements over no CSIT case were first reported in [35]. In [36], the K-user SISO IC and X channel have been studied under the delayed CSIT assumption wherein new DoF results have been reported. In [37], the DoF of IC and X channel are investigated under the full-duplex transmitter cooperation and delayed CSIT.

In this work, new achievable results for the DoF region of the two-user MIMO IC with delayed CSIT are provided. We then consider the two-user SISO and MIMO X channel and obtain new achievable sum-DoF results under the delayed CSIT assumption. Finally, the K-user MISO Gaussian IC with M antennas at each transmitter is investigated under the delayed CSIT assumption wherein new achievable DoF result is provided.

1.4 Dissertation Outline and Main Contributions

In this dissertation, we address communication over the Gaussian interference and X networks. The following summarizes the main contributions of this dissertation:

In Chapter 2, we consider a K-user MIMO IC with M antennas at each transmitter and N antennas at each receiver. It is assumed that the channel coefficients are constant and known perfectly at all nodes prior to the transmission. First, a new upper-bound for the DoF of this channel is developed. In specific, it is shown that the DoF of this channel is upper-bounded by $K\frac{MN}{M+N}$ when $K \geq K_u = \frac{M+N}{\text{gcd}(M,N)}$. We then show that one can achieve this DoF using real interference alignment technique. This gives an exact characterization of DoF for $K \geq K_u$.

In Chapter 3, we consider the communication over the following channels:

- Two-user MIMO IC with Delayed CSIT
- Two-user MIMO X channel with Delayed CSIT

- K-user X networks with Delayed CSIT
- K-user MISO IC with Delayed CSIT

For the two-user MIMO interference channel, new achievable results on the DoF region are provided and shown to be tight for some antenna configurations. It is observed that, depending on the antenna configuration, the DoF region with delayed CSIT can collapse to the DoF region with no CSIT, strictly lie between DoF regions with no CSIT and full CSIT, or coincide with the DoF region with full CSIT. For the two-user MIMO X channel with M antennas at each transmitter and N antennas at each receiver, new achievable sum-DoFs are obtained which turn out to be tight for all cases except for $1/2 < N/M < 4/3$. In specific, it is shown that the two-user SISO X channel can achieve 6/5 DoF which is better than the previously reported result of 8/7 in [35]. We then extend our analysis to the K-user X networks and show that a DoF of $\frac{4}{3} - \frac{2}{3(3K-1)}$ is achievable for this channel. Finally, the K-user MISO interference channel with $M \geq K$ antennas at each transmitter is investigated under the delayed CSIT assumption wherein a DoF of $\frac{2K}{K+1}$ is achieved. Except for the K-user X networks, all of our achievable DoF results in this chapter are the best known results.

Chapter 2

Interference management for constant MIMO Interference channels

In this chapter*, we consider the K-user MIMO Gaussian interference channel with M antennas at each transmitter and N antennas at each receiver. It is assumed that channel coefficients are constant and are available at all transmitters and receivers. The main objective of this chapter is to characterize the DoF for this channel. Using a new interference alignment technique which has been recently introduced in [26], we show that $\frac{MN}{M+N}K$ degrees of freedom can be achieved for almost all channel realizations. Also, a new upper-bound on the DoF of this channel is provided. This upper-bound coincides with our achievable DoF for $K \geq K_u \triangleq \frac{M+N}{\gcd(M,N)}$, where $\gcd(M,N)$ denotes the greatest common divisor of M and N. This gives an exact characterization of DoF for $M \times N$ MIMO Gaussian interference channel in the case of $K \geq K_u$.

*Part of the work in this chapter has been presented in [38].

2.1 System Model

We consider a constant fully connected K-user MIMO Gaussian IC. This channel is used to model a communication network with K transmitter-receiver pairs. Each transmitter is equipped with M antennas and wishes to communicate with its corresponding receiver, which is equipped with N antennas. All transmitters share a common bandwidth and want to have reliable communication at maximum possible rates. The channel output at the k^{th} receiver is characterized by the following input-output relationship:

$$\mathbf{Y}^{[k]}(t) = \mathbf{H}^{[k1]}\mathbf{X}^{[1]}(t) + \mathbf{H}^{[k2]}\mathbf{X}^{[2]}(t) + \cdots + \mathbf{H}^{[kK]}\mathbf{X}^{[K]}(t) + \mathbf{Z}^{[k]}(t), \tag{2.1}$$

where t is the time index, $k \in \mathcal{K} = \{1, 2, \cdots, K\}$ is the user index, $\mathbf{Y}^{[k]} = (Y_1^{[k]}, \cdots, Y_N^{[k]})^T$ is the $N \times 1$ output signal vector of the k^{th} receiver, $\mathbf{X}^{[j]} = (X_1^{[j]}, \cdots, X_M^{[j]})^T$ is the $M \times 1$ input signal vector of the j^{th} transmitter, $\mathbf{H}^{[kj]} = [h_{nm}^{[kj]}]$ is the $N \times M$ channel matrix between transmitter j and receiver k with the $(n,m)^{th}$ entry specifying the channel gain from the m^{th} antenna of transmitter j to the n^{th} antenna of receiver k, and $\mathbf{Z}^{[k]} = (Z_1^{[k]}, \cdots, Z_N^{[k]})^T$ is $N \times 1$ additive white Gaussian noise (AWGN) vector at the k^{th} receiver. We assume all noise terms are i.i.d. zero mean unit variance Gaussian random variables. It is assumed that each transmitter is subject to an average power constraint P:

$$\mathbb{E}[(\mathbf{X}^{[k]}(t))^{\dagger}(\mathbf{X}^{[k]}(t))] \leq P, \quad k \in \mathcal{K}.$$

Also, let $\mathcal{H}$ denote the set of all channel coefficients, i.e.

$$\mathcal{H} = \{\mathbf{H}^{[k1]}, \mathbf{H}^{[k2]}, \cdots, \mathbf{H}^{[kK]}\}_{k=1}^{K}$$

Transmitter k wishes to communicate a message $W^{[k]} \in \mathcal{W}^{[k]} = \{1, 2, \cdots, 2^{\tau R^{[k]}}\}$ of rate $R^{[k]}$ to receiver k over a block of τ channel uses using a block code of length τ, which is defined as follows:

Definition 1. *A $(2^{\tau\mathbf{R}}, \tau)$ code of block length τ and rate $\mathbf{R} = (R^{[1]}, R^{[2]}, \cdots, R^{[K]})$ for the K-user MIMO Gaussian IC with channel knowledge $\mathcal{H}$ at all nodes is defined as K sets of* *encoding functions* $\{\varphi_{t,\tau}^{[k]}\}_{t=1}^{\tau}$, $1 \leq k \leq K$, *such that*

$$\mathbf{X}^{[k]}(t) = \varphi_{t,\tau}^{[k]}(W^{[k]}, \mathcal{H}) \tag{2.2}$$

together with K *decoding functions* $\psi_{\tau}^{[k]}, 1 \leq k \leq K$, *such that*

$$\hat{W}_{\tau}^{[k]} = \psi_{\tau}^{[k]}(\{\mathbf{Y}^{[k]}(t)\}_{t=1}^{\tau}, \mathcal{H}) \tag{2.3}$$

Let $P_{e,\tau}^{[k]}$ denote the probability of error for receiver k, i.e.

$$P_{e,\tau}^{[k]} = Pr\{W^{[k]} \neq \hat{W}_{\tau}^{[k]}\} \tag{2.4}$$

Then the probability of error for the block code $(2^{\tau\mathbf{R}}, \tau)$ is defined as

$$P_{e,\tau} = \max_{k \in \mathcal{K}} P_{e,\tau}^{[k]} \tag{2.5}$$

The notions of achievable rate and the capacity region for the K-user MIMO Gaussian IC are defined as follows:

Definition 2. *For a given power constraint* P*, a rate tuple* $\mathbf{R}(P)$ *is said to be achievable for the* K*-user MIMO Gaussian IC if there exists a sequence* $\{(2^{\tau\mathbf{R}(P)}, \tau)\}_{\tau=1}^{\infty}$ *of codes such that their probability of error goes to zero as* $\tau \to \infty$*. The closure of the set of all achievable rate tuples is called the capacity region of the channel with power constraint* P *and is denoted by* $\mathcal{C}(P)$*.*

The notion of DoF is defined next.

Definition 3. *To an achievable rate tuple* $\mathbf{R}(P) = (R^{[1]}(P), \cdots, R^{[K]}(P)) \in \mathcal{C}(P)$*, one can correspond an achievable DoF tuple* $(d^{[1]}, \cdots, d^{[K]})$ *provided that:*

$$R^{[k]}(P) = d^{[k]}. \log_2(P) + o(\log_2(P)), \quad k = 1, 2, \cdots, K. \tag{2.6}$$

The set of all achievable DoF tuples is called the DoF region and is denoted by $\mathcal{D}$*.*

Definition 4. *The maximum sum-rate or sum-capacity of the K-user MIMO Gaussian IC is defined as*

$$C_\Sigma(P) = \max_{\mathbf{R}(P)\in\mathcal{C}(P)} \sum_{k=1}^{K} R^{[k]}(P) \tag{2.7}$$

The maximum achievable sum-DoF (or simply channel DoF) is defined as

$$\mathsf{DoF} = \max_{(d_1,\cdots,d_K)\in\mathcal{D}} \sum_{k=1}^{K} d^{[k]} \tag{2.8}$$

For notational consistency, lower and upper bounds on DoF will be denoted by $\underline{\mathsf{DoF}}$ and $\overline{\mathsf{DoF}}$, respectively.

In the sequel, a $(K, M \times N)$ IC refers to a constant fully connected K-user MIMO Gaussian IC with M antennas at each transmitter and N antennas at each receiver. Our primary objective in this chapter is to characterize the DoF of this channel.

2.2 Main Results and Discussion

2.2.1 Upper-bound

The first result of this study, presented in Section 2.3, is an upper-bound for the DoF of the $(K, M \times N)$ MIMO IC. The upper-bound is obtained by allowing full cooperation among groups of users and applying the two-user MIMO IC DoF result of [39] in conjunction with the averaging argument of [40]. The novelty here is in the application of this bound exhaustively to all possible cooperative combinations and choosing the tightest of the resulting bounds. In particular, the DoF of $(K, M \times N)$ MIMO IC is shown to be upper-bounded by $K\frac{MN}{M+N}$ for $K \geq K_u = \frac{M+N}{\gcd(M,N)}$. The upper-bound can be pictorially presented in a more elegant way by defining the *normalized degrees of freedom*. The normalized DoF of a $(K, M \times N)$ IC is defined as:

$$\mathsf{DoF}_{\text{norm}} \triangleq \frac{\mathsf{DoF}}{K \min(M, N)}. \tag{2.9}$$

Note that $K \min(M, N)$ is the DoF of a system consisting of K non-interfering $M \times N$ MIMO channels. Therefore, $\mathsf{DoF}_{\text{norm}}$ is always less than or equal to unity. The normalized upper-bound $\overline{\mathsf{DoF}}_{\text{norm}}$ presented in Section 2.3 is a function of only two parameters K and $\beta \triangleq \frac{\max(M,N)}{\min(M,N)}$ as follows:

$$\overline{\mathsf{DoF}}_{\text{norm}} = \min\{\beta\rho^+, 1 - \rho^-\}, \tag{2.10}$$

where ρ^- and ρ^+ are given by:

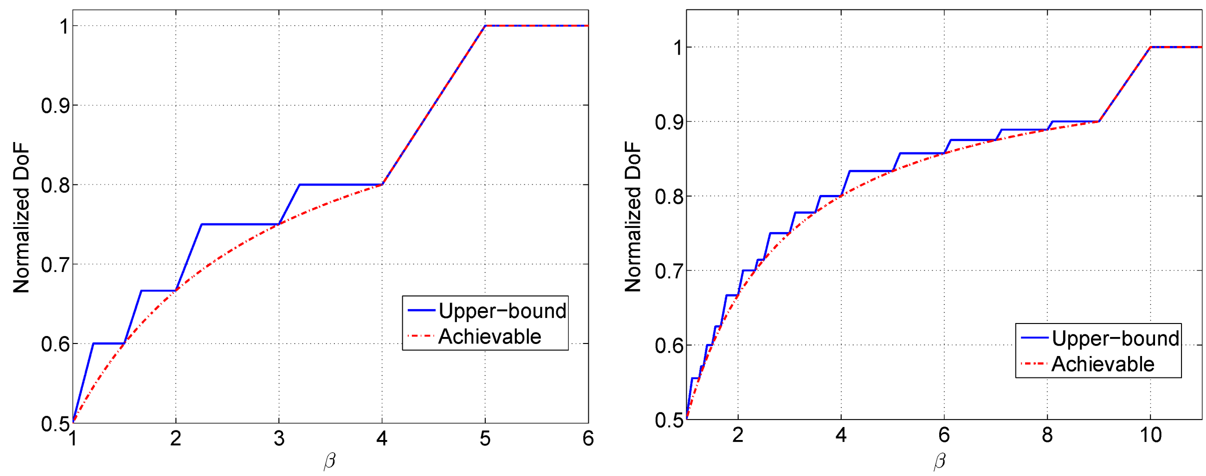
$$\rho^- = \max_{n \in \mathcal{K}} \frac{1}{n} \lfloor \frac{n}{1+\beta} \rfloor, \qquad \rho^+ = \min_{n \in \mathcal{K}} \frac{1}{n} \lceil \frac{n}{1+\beta} \rceil. \tag{2.11}$$

As shown in Fig. 2.1, for each K, the normalized DoF upper-bound is a piecewise linear function of β. Furthermore, as K increases the number of piecewise linear sections in the curve of $\overline{\mathsf{DoF}}_{\text{norm}}$ also increases.

2.2.2 Asymptotic interference alignment for the IC using rational dimensions

To highlight the novel aspects of our transmission scheme, we start with the following observation:

Consider a $(K, M \times N)$ MIMO IC. Now split each transmitter and receiver into multiple single antenna nodes as shown in Fig. 2.2. In this new $KM \times KN$ network, there is an independent message from a transmitter to a receiver if and only if there was desired communication between them in the original network. This lead to an X network setting between the M transmitters and N receivers that correspond to the same original user, and an interference network setting across transmitters and receivers corresponding to different



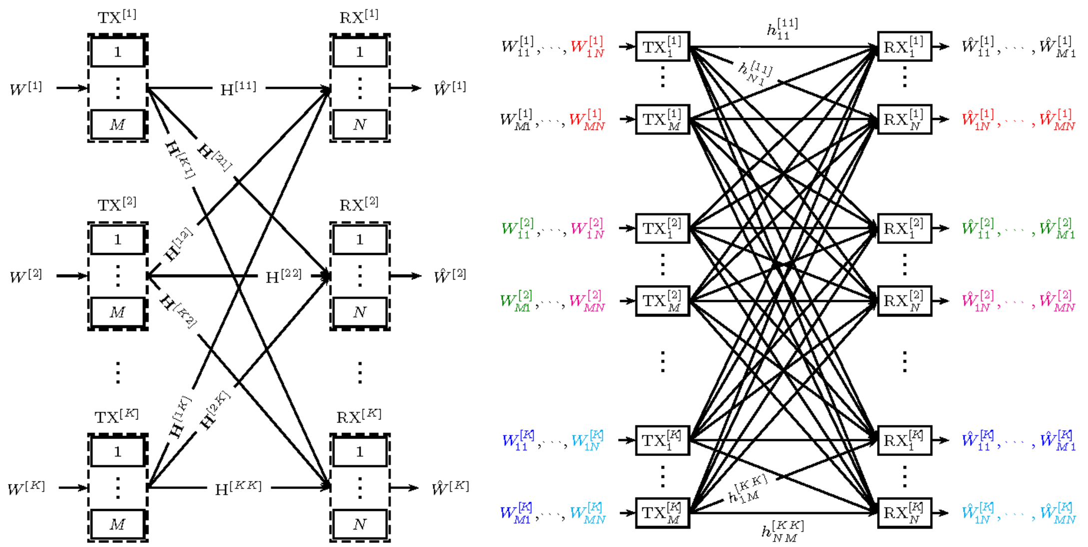
(a) $K = 5$ (b) $K = 10$

Figure 2.1: Our achievable and upper-bound on normalized DoF of a $(K, M \times N)$ IC for $K = 5$ and $K = 10$.

users in the original network (see Fig. 2.2). In this new network, suppose the channels are time-varying/frequency-selective. Then, using both the upper-bound and achievability scheme of [19], it is not difficult to see that the DoF of this new network is $K\frac{MN}{M+N}$. In light of this observation, the main contribution of the current work is two-fold:

- Prove that this DoF value is optimal even in the original $(K, M \times N)$ MIMO IC when $K \geq K_u$.
- Prove that the $K\frac{MN}{M+N}$ DoF value for the network described above, is also achievable when the channel coefficients are constant.

Our transmission scheme builds on the machinery developed in [26] for signal scale alignment over real numbers and also the interference alignment construction proposed in [19] for time varying X channels. Specifically, there are three elements in the achievability proof:



(a) Original $(K, M \times N)$ MIMO IC (b) $KM \times KN$ SISO X network setting with partial message sets

Figure 2.2: Splitting each transmitter and receiver of a MIMO IC into multiple single antenna nodes. Note that in the resulting network there is an independent massage from a transmitter to a receiver if and only if there was desired communication between them in the original network.

- **Rational dimensions framework:** A new framework for the interference alignment is recently suggested in [26] which relies on the theory of Diophantine approximation on manifolds [41], [42]. A new theme in this framework is the notion of rational independence which lets several integer data streams to be multiplexed using rationally independent coefficients. As shown in [26], the celebrated Khintchine-Groshev Theorem guarantees the separability of these data streams almost surely provided that certain conditions are satisfied. The rational independent coefficients in the alignment construction of [26] act like linearly independent vectors in signal space alignment. We exploit this close analogy between rational dimension framework and signal space interference alignment to use the alignment construction of [19] within the rational dimension framework.

- **Alignment construction:** The alignment construction used here is similar to the construction that is introduced by Cadambe and Jafar in [19] for time-varying X channels. This construction is simply a way to construct a set that is almost-invariant to an arbitrarily large number of given linear transformations, whose only requirement is commutativity. As shown in [19], such a set is composed of elements that are simply products of powers of the specified linear transformations, and an initial seed. In this paper, the linear transformations are scalars (channel coefficients), so commutativity is trivially satisfied and the elements of the set are monomials in the channel coefficients.

- **Resolvability:** While the notion of independence in our achievable scheme is rational rather than linear, the argument that establishes this independence is simply follows from the same argument made in [19]. In fact, the main argument for establishing resolvability (that the monomials are distinct) is the same as the proof in [19], where the full rank property of the matrix is proved by showing that each column corresponds to a distinct monomial.

We now provide an intuitive overview of our achievable scheme. A more detailed proof is presented in the next section.

For any arbitrary $\Gamma > 0$, $\Gamma \in \mathbb{N}$, our transmission scheme is constructed in a one dimensional signal space with rational dimension of $(M+N)(f(\Gamma)+o(f(\Gamma)))$, where

$$f(\Gamma) = \Gamma^{KM(KN-1)}.$$

Over this space, each desired sub-message (i.e. $W_{mn}^{[k]}$ in the corresponding X network) in our achievable scheme achieves $f(\Gamma)+o(f(\Gamma))$ DoF. By choosing a large Γ, arbitrarily close to $\frac{1}{M+N}$ DoF can be achieved for each sub-messages. Therefore, as Γ grows to infinity, the achievable DoF for all sub-messages corresponding to a specific user in the original IC is arbitrarily close to $\frac{MN}{M+N}$. The achievable scheme uses real interference alignment over a space of rational dimension $(M+N)(f(\Gamma)+o(f(\Gamma)))$. In specific, sub-message $W_{mn}^{[k]}$ is transmitted over $f(\Gamma)+o(f(\Gamma))$ rationally independent complex numbers at the m^{th} transmit antenna of user k. The decoding at receiver j is based on the Khintchine-Groshev Theorem which guarantees the separability of the rationally independent complex numbers, almost surly. As shown in Fig. 2.3, at the n^{th} antenna of receiver k:

- the $f(\Gamma)+o(f(\Gamma))$ rational dimensions corresponding to $W_{mn}^{[k']}$, $k' \neq k$, align with the $f(\Gamma)+o(f(\Gamma))$ rational dimensions corresponding to $W_{m'n}^{[k']}$, $m' \neq m$. In other words, the rational dimension of the space occupied by the union of the signals corresponding to sub-messages $\{W_{mn}^{[k']} : k' \neq k, m \in \mathcal{M}\}$ is $f(\Gamma)+o(f(\Gamma))$. All of these interfering terms occupy a space with rational dimension $f(\Gamma)+o(f(\Gamma))$.

- the $f(\Gamma)+o(f(\Gamma))$ rational dimensions corresponding to $W_{mn'}^{[k]}$, $n' \neq n$, align with the $f(\Gamma)+o(f(\Gamma))$ rational dimensions corresponding to $W_{m'n''}^{[k']}$, $n'' \neq n$. In other words, the rational dimension of the space occupied by the union of the signals corresponding to sub-messages $\{W_{mn'}^{[k]} : k \in \mathcal{K}, m \in \mathcal{M}, n' \neq n\}$ is $f(\Gamma)+o(f(\Gamma))$. Since the cardinality of the set $\{n' \neq n, n \in \{1,2,\cdots,N\}\}$ is $N-1$, all of these interfering terms will occupy a space with rational dimension $(N-1)(f(\Gamma)+o(f(\Gamma)))$.

It should be emphasize that the overlap between these signals is partial for a fixed value of Γ and the number of rational dimensions that are not align captured by the $o(f(\Gamma))$ factor. That is the perfect alignment happens asymptotically as Γ grows to infinity. Therefore, the rational dimension of all interfering terms at each receive antenna of receiver k is $N(f(\Gamma)+o(f(\Gamma)))$. The rational dimension of signals corresponding to desired sub-messages is $M(f(\Gamma)+o(f(\Gamma)))$. If the set of complex numbers corresponding to desired sub-messages are rationally independent of the complex numbers corresponding to interfering signals, then in a space of rational dimension $(M+N)(f(\Gamma)+o(f(\Gamma)))$, the $M(f(\Gamma)+o(f(\Gamma)))$ desired dimensions can be successfully decoded according to the Khintchin-Groshev Theorem. The precise construction of our achievable scheme can be found in next section.

2.2.3 Comparing achievable and upper-bound results

It is easy to show that in a $(K, M \times N)$ IC, one can always achieve

$$\min\{\max(M,N), K\min(M,N)\}$$

DoF by zero-forcing. Combining this result with $K\frac{MN}{M+N}$ DoF which can be achieved through interference alignment, we obtain

$$\mathsf{DoF} \geq \begin{cases} K\min(M,N)\,\min(1,\frac{\beta}{K}), & K < \beta+1 \\ K\frac{MN}{M+N}, & K \geq \beta+1 \end{cases}, \tag{2.12}$$

or equivalently,

$$\underline{\mathsf{DoF}}_{\text{norm}} = \begin{cases} \min(1,\frac{\beta}{K}), & K < \beta+1 \\ \frac{\beta}{\beta+1}, & K \geq \beta+1 \end{cases}. \tag{2.13}$$

Two examples comparing our achievable result and upper-bound on $\mathsf{DoF}_{\text{norm}}$ are depicted in Fig. 2.1.

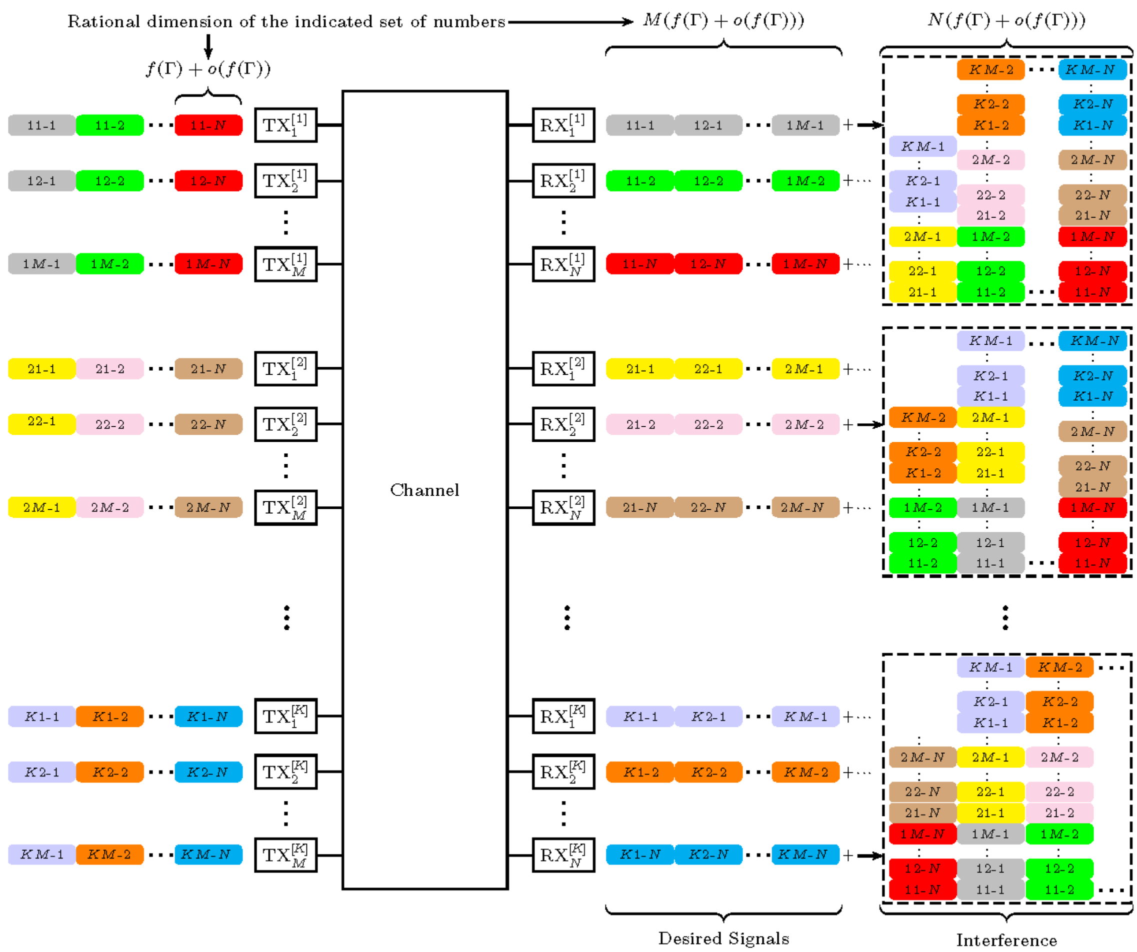


Figure 2.3: Real Interference alignment for the $(K, M \times N)$ IC: the transmit signal of each user is composed of N independent parts which are depicted here by adjacent rectangles. By the real interference alignment, the squares in each column at the receiver side are approximately aligned.

By comparing the achievable and upper-bound results, one can observe that the channel DoF is completely characterized in the following cases:

- β is an integer

  In this case, $\gcd(M, N) = \min(M, N)$, and hence, $K_u = \beta + 1$. Therefore, for $K \geq \beta + 1$, the DoF is equal to $K \frac{MN}{M+N}$, and for $K \leq \beta$ the DoF is equal to $K \min(M, N)$. This is the setting where [27] also has a tight DoF characterization.

- $K \leq \beta + 1$

  In this case, one can easily verify that $\rho^- = 0$ and $\rho^+ = \frac{1}{K}$ in the upper-bound, and therefore, the DoF is equal to $K \min(M, N) \min(1, \beta/K)$. Note that this DoF can be achieved by simple zero-forcing.

- $K \geq K_u$

  In this case, $\rho^- = \rho^+ = \frac{1}{1+\beta}$ in the upper-bound and hence the channel DoF is equal to $K \frac{MN}{M+N}$.

While our results provide a complete characterization of DoF for $K \geq K_u$ and $K \leq 1+\beta$, this characterization for the case of $1 + \beta < K < K_u$ with non-integer values of β seems to be challenging. Recently, it was shown in [43] that for the special case of $K = 3$, one can achieve higher DoF values in this range using signal space interference alignment. The case of $K = 3$ has also been extensively investigated in [44] wherein new achievable and upper-bound results are developed. As shown in Fig. 2.4, the information theoretic upper-bound of [44] is tighter than the cooperative upper-bound developed here for the case of $K = 3$. Also, the upper-bound of [44] reveals that when $\beta = \frac{p+1}{p}$ for $p \in \mathbb{Z}^+$, the $K \frac{MN}{M+N}$ DoF value is tight for any $K > 2$ (see Fig. 2.4 for $K = 3$).

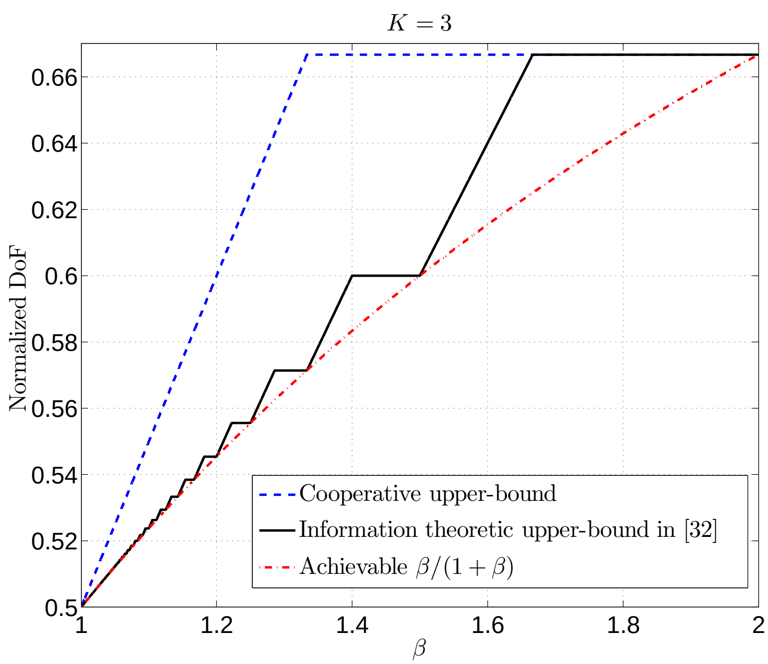


Figure 2.4: Comparing different results on the DoF of the $(3, M \times N)$ MIMO interference channel where $\beta = \frac{\max(M,N)}{\min(M,N)}$

.

2.2.4 Joint processing in collocated antennas

These is no cooperation among the transmit or receive antennas of each user in our transmission scheme. Since our transmission scheme is DoF optimal for $K \geq K_u$, it follows that the DoF advantage of joint processing in collocated antennas vanishes for $K \geq K_u$. In fact, the $(K, M \times N)$ MIMO IC is treated as a $KM \times KN$ X channel with partial message sets in our transmission scheme.

2.3 DoF upper-bound for the K-user MIMO interference channel

The following theorem is the main result of this section.

Theorem 1. *The DoF of a $(K, M \times N)$ IC is upper-bounded by:*

$$\overline{\mathsf{DoF}} \triangleq K \min\left\{\max(M,N)\rho^{+}, \min(M,N)(1-\rho^{-})\right\}, \tag{2.14}$$

where ρ^{+} and ρ^{-} are given by:

$$\rho^{-} = \max_{n \in \mathcal{K}} \frac{\lfloor n\rho_0 \rfloor}{n}, \qquad \rho^{+} = \min_{n \in \mathcal{K}} \frac{\lceil n\rho_0 \rceil}{n}, \tag{2.15}$$

and where $\rho_0 \triangleq \frac{\min(M,N)}{M+N}$ and $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are respectively the floor and the ceiling functions.

Remark 1. *The upper-bound in Theorem 1 is valid for both constant and varying channels.*

Proof. Consider a $(W, M \times N)$ Gaussian IC where $W \leq K$ is a constant. We divide these W users into two disjoint sets of size W_1 and W_2, where $W = W_1 + W_2$. Let us assume that the transmitters in each set are cooperating, and the receivers in each set are cooperating as well. This results in a two-user MIMO Gaussian IC with W_1M, W_2M antennas at transmitters and W_1N, W_2N antennas at their corresponding receivers. It is proved in [39] that for a two-user MIMO Gaussian IC with M_1, M_2 antennas at transmitter 1, 2 and N_1, N_2 antennas at their corresponding receivers, the DoF is equal to:

$$J(M_1, M_2, N_1, N_2) = \min\{M_1 + M_2, N_1 + N_2, \max(M_1, N_2), \max(M_2, N_1)\}. \tag{2.16}$$

Since cooperation does not reduce the capacity, the DoF of the original W-user interference channel does not exceed $J(W_1M, W_2M, W_1N, W_2N)$. Thus, for any $i_1, i_2, \cdots, i_W \in \mathcal{K}$, $i_1 \neq i_2 \neq \cdots \neq i_W$, we have:

$$d^{[i_1]} + d^{[i_2]} + \cdots + d^{[i_W]} \leq J(W_1M, W_2M, W_1N, W_2N), \tag{2.17}$$

where $d^{[k]}$ denotes the DoF of user k. Adding up all inequalities similar to (2.17), the DoF of the K-user Gaussian IC is upper-bounded as:

$$\mathsf{DoF} \leq \frac{K}{W} J(W_1M, W_2M, W_1N, W_2N). \tag{2.18}$$

It is proved in Appendix A.1 that the function $J(W_1M, W_2M, W_1N, W_2N)$ can be upper-bounded as:

$$J(W_1M, W_2M, W_1N, W_2N) \leq \max\{\max(M, N)W_{\min}, \min(M, N)W_{\max}\}, \tag{2.19}$$

where $W_{\max} = \max(W_1, W_2)$ and $W_{\min} = \min(W_1, W_2)$. Combining (2.19) and (2.18), we have:

$$\mathsf{DoF} \leq K\, G(\rho), \tag{2.20}$$

where $\rho \triangleq \frac{W_{\min}}{W}$ and

$$G(\rho) \triangleq \max\{\max(M, N)\rho, \min(M, N)(1 - \rho)\}. \tag{2.21}$$

A typical plot of $G(\rho)$ is depicted in Fig. 2.5. To obtain the tightest upper-bound, we need to minimize $G(\rho)$ over the rational number ρ. However, there are two constraints on ρ:
C1) $0 \leq \rho \leq \frac{1}{2}$,
C2) the denominator of ρ as a rational number in lowest terms can not exceed K. Thus,

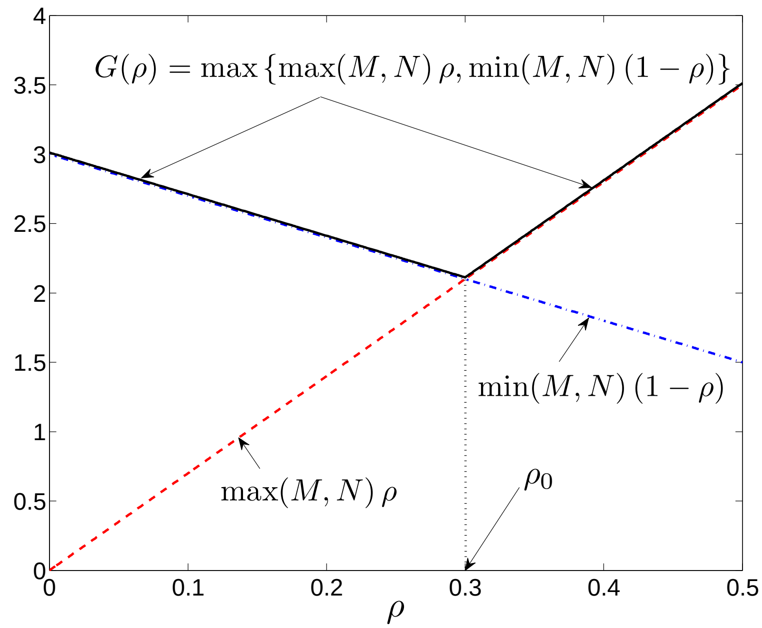


Figure 2.5: Typical shape of function $G(\rho)$ in (2.21).

the goal is to minimize $G(\rho)$ subject to the constraints C1 and C2. It is straightforward to show that (see also Fig. 2.5) without any constraint on ρ, the function $G(\rho)$ is minimized when:

$$\max(M,N)\rho = \min(M,N)(1-\rho). \tag{2.22}$$

Equivalently, $G(\rho)$ is minimized at $\rho = \rho_0$, where ρ_0 was defined in Theorem 1. Although $\rho = \rho_0$ satisfies constraint C1, it does not generally satisfy constraint C2 because the denominator of ρ_0 in the simplest form can exceed K. Therefore, to find the optimal ρ that minimizes $G(\rho)$ subject to the constraints C1 and C2, we need to find the closest rational neighbors of ρ_0 with denominator not exceeding K. Let ρ^- and ρ^+ denote the closest rational neighbors of ρ_0 with denominator not exceeding K such that $0 \le \rho^- \le \rho \le \rho^+$.

From (2.20), for such ρ^+ and ρ^-, we have:

$$\begin{aligned} \mathsf{DoF} &\leq K \max\{\max(M,N)\rho^+, \min(M,N)(1-\rho^+)\} = K \max(M,N)\rho^+ \\ \mathsf{DoF} &\leq K \max\{\max(M,N)\rho^-, \min(M,N)(1-\rho^-)\} = K \min(M,N)(1-\rho^-) \end{aligned}. \tag{2.23}$$

Therefore, the final upper-bound can be expressed as:

$$\mathsf{DoF} \leq K \min\left\{\max(M,N)\rho^+, \min(M,N)(1-\rho^-)\right\}. \tag{2.24}$$

The problem of finding the closest rational neighbors of a real number with denominator less than or equal to K is addressed in the following lemma whose proof can be found in Appendix A.2:

Lemma 1. *Let $\alpha \in (0,1)$ be a real number. Given a positive integer K, the closest rational neighbors of α ($\alpha^- \leq \alpha \leq \alpha^+$) with denominator not exceeding K are given by:*

$$\alpha^- = \max_{n\in\{1,2,\cdots,K\}} \frac{\lfloor n\alpha \rfloor}{n}, \tag{2.25}$$

$$\alpha^+ = \min_{n\in\{1,2,\cdots,K\}} \frac{\lceil n\alpha \rceil}{n}. \tag{2.26}$$

Now, (2.15) easily follows from the above lemma and the proof is complete. □

2.4 Interference alignment and DoF lower-bound for MIMO IC

The following is the main result of this section.

Theorem 2. *For the $(K, M \times N)$ IC, we can achieve $K\frac{MN}{M+N}$ degrees of freedom for* almost all *channel realizations.*

A new method for interference alignment has been recently introduced by Motahari *et al.* in [26]. By applying arguments from the field of Diophantine approximation in Number

Theory, they showed that interference alignment can be performed based on the properties of rational and irrational numbers. Using this new type of alignment, which the authors called *real interference alignment*, the DoF of the K-user *constant* Gaussian IC with single antenna can be achieved for almost all channel realizations. Since our achievability scheme is based on an extension of real interference alignment, we first review the basic ideas behind this technique. In our discussions, we follow the footsteps of [26] and [22].

2.4.1 Preliminaries on Real Interference Alignment

Real interference alignment essentially mimics, in one dimension, the basic rules of signal-space interference alignment. In signal space interference alignment, the transmit signal of each user is a linear combination of some constant vectors in Euclidean space, which hereafter will be called *transmit directions*, where data determines the coefficients of this linear combination. In this setup, interference alignment is realized by simultaneous design of appropriate transmit directions for different users such that:
i) Interfering signals from other users are received aligned at the intended receiver. In other words, all interfering terms at each receiver fall into a subspace of the available signal space at that receiver. This condition will be referred to as *alignment condition.*
ii) The interference subspace can be separated from the desired signal subspace at each receiver. This condition will be referred to as *separability condition.*

Note that transmit directions are selected according to the channel coefficients. In signal space alignment, when both alignment and separability conditions are satisfied, we can separate the desired signal from aligned interfering signals by zero-forcing.

Consider a K-user Gaussian IC with a single antenna at all nodes where channel coefficients are all constant. To introduce the counterparts of separability and alignment conditions in real interference alignment, we need the notion of *rational independence.*

Definition 5 (rational independence)**.** *The complex numbers* $\omega_1, \omega_2, \cdots, \omega_m$ *are said to be* *rationally independent if whenever integers* $k_1, k_2, ..., k_m$ *satisfy*

$$k_1\omega_1 + k_2\omega_2 + \cdots + k_m\omega_m = 0,$$

we should have $k_i = 0$ *for* $i = 1, \cdots, m$, *i.e., the only representation of zero as a linear combination of* ω_i, $i = 1, \cdots, m$ *is the trivial solution.*

Next, the notion of rational dimension is defined:

Definition 6 (rational dimension). *The rational dimension of complex numbers* ω_1, ω_2, $\cdots$, ω_m *is defined as the smallest natural number* n *such that all numbers* ω_i, $i = 1, \cdots, m$ *can be represented as rational linear combinations of* n *fixed rationally independent complex numbers. The rational dimension of a set* $\mathcal{A}$ *of numbers will be denoted by* $\dim(\mathcal{A})$.

Suppose that $\omega_1, \omega_2, \cdots, \omega_m$ are rationally independent numbers. Therefore, for arbitrary integers $k_1, k_2, ..., k_m$, not all of them equal to zero, we have $|k_1\omega_1 + k_2\omega_2 + \cdots + k_m\omega_m| > 0$. The problem of finding a non-zero lower-bound on the absolute value of an integer linear combination of rationally independent numbers is closely related to metric Diophantine approximation in Number Theory [41]. The following theorem which is an extension of the Khintchine-Groshev Theorem in metric Diophantine approximation [45] provides a quantitative lower-bound on the absolute value of an integer linear combination of complex numbers.

Theorem 3 (Khintchine-Groshev for complex numbers). *Assume* $\epsilon > 0$ *is an arbitrary positive constant. For* almost all ℓ*-tuples* $\boldsymbol{\omega} = (\omega_1, \omega_2, \cdots, \omega_\ell)$ *of complex numbers, one can find a constant* c *such that the inequality*

$$|p + q_1\omega_1 + q_2\omega_2 + \cdots + q_\ell\omega_\ell| > \frac{c}{(\max_i q_i)^{(\ell-1)/2+\epsilon}} \tag{2.27}$$

holds for all $p \in \mathbb{Z}$ *and all* $q = (q_1, q_2, \cdots, q_\ell) \in \mathbb{Z}^\ell \setminus \mathbf{0}$.

It is important to note that the Khintchine-Groshev Theorem is valid for "almost all" complex numbers. That is the Lebesgue measure of those numbers satisfying the

Khintchine-Groshev Theorem is one. It should be pointed out here that the Khintchine-Groshev Theorem is not valid even for all rationally independent complex numbers.

The complex numbers ω_i, $i = 1, \cdots, \ell$, in the Khintchine-Groshev Theorem could be independent quantities or they can lie on some well-behaved manifold. Specifically, the Khintchine-Groshev Theorem is still valid when all the complex numbers ω_i, $i = 1, \cdots, \ell$ are different monomials in $m < l$ independent variables [45].

Consider two sets $\mathcal{A}$ and $\mathcal{B}$ of complex numbers with rational dimensions $\dim(\mathcal{A})$ and $\dim(\mathcal{B})$, respectively. We define the *alignment index* of $\mathcal{A}$ and $\mathcal{B}$, which is denoted by $\chi(\mathcal{A}, \mathcal{B})$, as:

$$\chi(\mathcal{A}, \mathcal{B}) \triangleq \frac{\dim(\mathcal{A} \bigcup \mathcal{B})}{\max(\dim(\mathcal{A}), \dim(\mathcal{B}))}.$$

It is easy to see that $\chi(\mathcal{A}, \mathcal{A}) = 1$ for any non-empty set $\mathcal{A}$. Furthermore, one can readily see that $\chi(\mathcal{A}, \mathcal{B}) \geq 1$ for any two non-empty sets $\mathcal{A}$ and $\mathcal{B}$. The alignment index of more than two sets is similarly defined as the ratio of the rational dimension of their union to the maximum of the individual rational dimensions.

Now, consider two sequences $\mathcal{A}_n$ and $\mathcal{B}_n$ of sets where the cardinalities of $\mathcal{A}_n$ and $\mathcal{B}_n$ grows to infinity as $n \to \infty$. We define the notion of *asymptotic alignment* as follows:

Definition 7 (Asymptotic alignment)**.** *Two sequences $\mathcal{A}_n$ and $\mathcal{B}_n$ of sets are called asymptotically aligned if* $\limsup_{n\to\infty} \chi(\mathcal{A}_n, \mathcal{B}_n) = 1$.

The above definition can be generalized to more than two sequences of sets. In other words, S sequences of sets $\mathcal{A}_n^{[1]}, \cdots, \mathcal{A}_n^{[S]}$ are call asymptotically aligned if the $\limsup$ of their alignment index goes to unity as $n \to \infty$.

Consider two sequences of discrete random variables X_n and Y_n that are uniformly distributed over $\mathcal{A}_n$ and $\mathcal{B}_n$, respectively. If $\mathcal{A}_n$ and $\mathcal{B}_n$ are asymptotically aligned, the random sequences X_n and Y_n will be called asymptotically aligned.

Example 1. *Consider the following sequences of sets:*

$$\mathcal{A}_n = \{a_1^{n_1} a_2^{n_2} a_3^{n_3} : \; n_i \in \{0, 1, \cdots, n\}\}, \quad n = 1, 2, \cdots$$

where a_1, a_2, and a_3 are selected as three rationally independent real numbers such that for every n all the elements of $\mathcal{A}_n$ are rationally independent. According to the Khintchine-Groshev theorem, almost all triples of real numbers satisfy this condition. One can easily confirm that $\dim(\mathcal{A}_n) = (n+1)^3$. *Under this condition, the two sequences $a_1 \cdot \mathcal{A}_n$ and $a_2 \cdot \mathcal{A}_n$ of sets are asymptotically aligned. The reason is that $[a_1 \cdot \mathcal{A}_n \bigcup a_2 \cdot \mathcal{A}_n] \subset A_{n+1}$ and hence $\chi(a_1 \cdot \mathcal{A}_n, a_2 \cdot \mathcal{A}_n) \leq \frac{(n+2)^3}{(n+1)^3}$ which tends to one as $n \to \infty$.*

2.4.2 Proof of Theorem 2

Consider a $(K, M \times N)$ IC where each user satisfies a power constraint P. For any $\epsilon > 0$, we will provide a transmission scheme that achieves $\sum_{k=1}^{K} R_k = \frac{KMN}{M+N}(1-\epsilon)\log_2 P - o(\log_2 P)$, showing that $\mathsf{DoF} \geq \frac{KMN}{M+N}$.

In our achievable scheme, each transmitter uses its antennas separately, i.e., there is no cooperation among transmit antennas of each user. In fact, user k relies on M independent codebooks $\mathcal{C}_m^{[k]}(P, \epsilon, \tau)$, $m = 1, \cdots, M$, of block length τ where $\mathcal{C}_m^{[k]}(P, \epsilon, \tau)$ is associated with its m^{th} transmit antenna. Each codebook $\mathcal{C}_m^{[k]}(P, \epsilon, \tau)$, $m \in \mathcal{M}$, is obtained by a linear combination of N independent sub-codebooks $\mathcal{C}_{mn}^{[k]}(P, \epsilon, \tau)$, $n = 1, \cdots, N$. More precisely, the transmit symbol from the m^{th} antenna of user k at time index t can be expressed as:

$$X_m^{[k]}(t) = \sum_{n=1}^{N} h_{nm}^{[kk]} X_{mn}^{[k]}(t), \quad t = 1, \cdots, \tau, \tag{2.28}$$

where $\left(X_m^{[k]}(1), \cdots, X_m^{[k]}(\tau)\right) \in \mathcal{C}_m^{[k]}(P, \epsilon, \tau)$ and $\left(X_{mn}^{[k]}(1), \cdots, X_{mn}^{[k]}(\tau)\right) \in \mathcal{C}_{mn}^{[k]}(P, \epsilon, \tau)$. The sub-codebook $\mathcal{C}_{mn}^{[k]}(P, \epsilon, \tau)$ is intended to be decoded at the n^{th} receive antenna of user k. Each sub-codebook $\mathcal{C}_{mn}^{[k]}(P, \tau)$ is in turn obtained by adding L independent sub-sub-codebooks $\mathcal{C}_{mnl}^{[k]}(P, \epsilon, \tau)$, $\ell = 1, \cdots, L$, i.e.,

$$X_{mn}^{[k]}(t) = \sum_{\ell=1}^{L} X_{mn\ell}^{[k]}(t), \quad t = 1, \cdots, \tau, \tag{2.29}$$

where $\left(X_{mn\ell}^{[k]}(1), \cdots, X_{mn\ell}^{[k]}(\tau)\right) \in \mathcal{C}_{mn\ell}^{[k]}(P, \epsilon, \tau)$ and $L \in \mathbb{N}$ is a design parameter which will be determined later. Each sub-sub-codebook $\mathcal{C}_{mn\ell}^{[k]}(P, \epsilon, \tau)$ is generated i.i.d. according to a uniform distribution over $\Lambda_{mn\ell}^{[k]}(P, \epsilon)$, where:

$$\Lambda_{mn\ell}^{[k]}(P, \epsilon) \triangleq \gamma P^{\frac{\nu-2+4\epsilon}{2(\nu+2\epsilon)}} \omega_{mn\ell}^{[k]} \cdot \{-Q, -Q+1, \cdots, Q\}, \tag{2.30}$$

in which:

- $Q \triangleq \lfloor P^{\frac{1-\epsilon}{\nu+2\epsilon)}} \rfloor$.
- γ is a normalizing constant selected such that the average transmit power of each user does not exceed P. In Appendix A.3, we calculate the normalizing constant γ and show that it is independent of ν and P.
- $\nu \in \mathbb{N}$ is an important design parameter which controls the cardinality of $\Lambda_{mn\ell}^{[k]}(P, \epsilon)$ as well as the magnitude of its elements. Since $|\Lambda_{mn\ell}^{[k]}(P, \epsilon)| = 2Q + 1 \leq 2P^{\frac{1-\epsilon}{\nu+2\epsilon}} + 1$, we refer to ν as the *rate control parameter*.
- $\omega_{mn\ell}^{[k]}$ is a real number which should be properly selected according to the channel coefficients for the purpose of interference alignment.

Since $\gamma P^{\frac{\nu-2+4\epsilon}{2(\nu+2\epsilon)}}$ does not depend on m, n, and ℓ, the symbol $X_{mn}^{[k]}(t)$ can be considered as a random integer linear combination of L real numbers $\omega_{mn1}^{[k]}, \cdots, \omega_{mnL}^{[k]}$ multiplied by $\gamma P^{\frac{\nu-2+4\epsilon}{2(\nu+2\epsilon)}}$, i.e.,

$$X_{mn}^{[k]}(t) = \gamma P^{\frac{\nu-2+4\epsilon}{2(\nu+2\epsilon)}} \sum_{\ell=1}^{L} B_{mn\ell}^{[k]}(t) \omega_{mn\ell}^{[k]}, \tag{2.31}$$

where $B_{mn\ell}^{[k]}(t)$'s are independently and uniformly distributed over $\{-Q, -Q+1, \cdots, Q\}$. Each $B_{mn\ell}^{[k]}(t)$ will be referred to as *a data stream*. By substituting (2.31) in (2.28), the transmit symbol of user k on its m^{th} antenna can be reformulated as:

$$X_{m}^{[k]}(t) = \gamma P^{\frac{\nu-2+4\epsilon}{2(\nu+2\epsilon)}} \sum_{n=1}^{N} \sum_{\ell=1}^{L} B_{mn\ell}^{[k]}(t) h_{nm}^{[kk]} \omega_{mn\ell}^{[k]}. \tag{2.32}$$

We observe that $X_m^{[k]}(t)$ is a random integer linear combination of NL complex numbers $h_{nm}^{[kk]}\omega_{mn\ell}^{[k]}$, $n \in \mathcal{N}$, $\ell \in \mathcal{L}$. The complex numbers $h_{nm}^{[kk]}\omega_{mn\ell}^{[k]}$, $k \in \mathcal{K}$, $m \in \mathcal{M}$, $n \in \mathcal{N}$, $\ell \in \mathcal{L}$ act like beamforming vectors in signal space alignment and will be referred to as *modulation pseudo-vectors*. Let us define $\Omega_{mn}^{[k]}$ as:

$$\Omega_{mn}^{[k]} \triangleq \left\{\omega_{mn1}^{[k]}, \cdots, \omega_{mnL}^{[k]}\right\}. \tag{2.33}$$

Since the NL pseudo-vectors $h_{nm}^{[kk]} \cdot \Omega_{mn}^{[k]}$, $n \in \mathcal{N}$ carry independent data streams, they are required to be rationally independent, i.e.,

$$\dim\left(\bigcup_{n=1}^{N}\left[h_{nm}^{[kk]} \cdot \Omega_{mn}^{[k]}\right]\right) = NL, \quad \forall k \in \mathcal{K} \text{ and } \forall m \in \mathcal{M}. \tag{2.34}$$

Using the above signalling scheme, the received signal at the n^{th} antenna of receiver k at time index t can be expressed as:

$$Y_n^{[k]}(t)=\sum_{k'=1}^{K}\sum_{m=1}^{M} h_{nm}^{[kk']}X_m^{[k']}+Z_n^{[k]}(t)=\gamma P^{\frac{\nu-2+4\epsilon}{2(\nu+2\epsilon)}} \sum_{k'=1}^{K}\sum_{m=1}^{M}\sum_{n'=1}^{N}\sum_{\ell=1}^{L} B_{mn'\ell}^{[k']}(t)h_{nm}^{[kk']}h_{n'm}^{[k'k']}\omega_{mn'\ell}^{[k']}+Z_n^{[k]}(t) \tag{2.35}$$

$$\begin{aligned}= \gamma P^{\frac{\nu-2+4\epsilon}{2(\nu+2\epsilon)}} \Bigg[& \underbrace{\sum_{m=1}^{M}\sum_{\ell=1}^{L} B_{mn\ell}^{[k]}(t)\, (h_{nm}^{[kk]})^2\, \omega_{mn\ell}^{[k]}}_{\text{desired}} + \underbrace{\sum_{m=1}^{M}\sum_{\substack{n'=1\\n'\neq n}}^{N}\sum_{\ell=1}^{L} B_{mn'\ell}^{[k]}(t)\, h_{nm}^{[kk]}h_{n'm}^{[kk]}\, \omega_{mn'\ell}^{[k]}}_{\text{self-interference}} \\ & + \underbrace{\sum_{\substack{k'=1\\k'\neq k}}^{K}\sum_{m=1}^{M}\sum_{n'=1}^{N}\sum_{\ell=1}^{L} B_{mn'\ell}^{[k']}(t)\, h_{nm}^{[kk']}h_{n'm}^{[k'k']}\, \omega_{mn'\ell}^{[k']}}_{\text{multi-user interference}}\Bigg] + Z_n^{[k]}(t).\end{aligned} \tag{2.36}$$

As we see from (2.35), the modulation pseudo-vectors from different transmit antennas of different users appear in $Y_n^{[k]}(t)$ after multiplication with the corresponding channel coefficients. For example, the modulation pseudo-vector $h_{n'm}^{[k'k']}\omega_{mn'\ell}^{[k']}$ which is originated from the

m^{th} antenna of user k' appears in $Y_n^{[k]}$ as $h_{nm}^{[kk']} h_{n'm}^{[k'k']} \omega_{mn'\ell}^{[k']}$. We refer to $h_{nm}^{[kk']} h_{n'm}^{[k'k']} \omega_{mn'\ell}^{[k']}$ as a *received pseudo-vector* in $Y_n^{[k]}(t)$. According to this terminology, $Y_n^{[k]}(t)$ is a noisy version of an integer linear combination of $LMNK$ received pseudo-vectors. Each received pseudo-vector has a data stream as its coefficient. We observe from (2.36) that three different components appear in $Y_n^{[k]}(t)$:

- The desired component which contains LM data streams. Each desired data stream in $Y_n^{[k]}(t)$ (i.e., $B_{mn\ell}^{[k]}(t)$) can be represented by an ordered pair (m, ℓ), $m \in \mathcal{M}$, $\ell \in \mathcal{L}$.
- The self-interference component which contains $LM(N-1)$ data streams. All data streams in this component are originated from transmitter k.
- The multi-user interference component which contains $LMN(K-1)$ data streams. All the data streams in this component are originated from interfering users.

Let us define $\tilde{Y}_n^{[k]}(t)$ as the noise-free part of $Y_n^{[k]}(t)$. The received pseudo-vectors in $\tilde{Y}_n^{[k]}(t)$ are not necessarily rationally independent and therefore some of them may be expressed as rational linear combinations of the rest. Let us momentarily assume that $\tilde{Y}_n^{[k]}(t)$ is known at the n^{th} antenna of receiver k. We then can recover a data stream from $\tilde{Y}_n^{[k]}(t)$ provided that its corresponding received pseudo-vector can not be represented as a rational linear combination of the other received pseudo-vectors in $\tilde{Y}_n^{[k]}(t)$. Accordingly, all the desired data streams at the n^{th} antenna of receiver k can be obtained from $\tilde{Y}_n^{[k]}(t)$ if the received pseudo-vectors $(h_{nm}^{[kk]})^2 \omega_{mn\ell}^{[k]}$, $m \in \mathcal{M}$, $\ell \in \mathcal{L}$ can not be expressed as rational linear combinations of $h_{nm}^{[kk']} h_{n'm}^{[k'k']} \omega_{mn'\ell}^{[k']}$, $k' \in \mathcal{K}$, $m \in \mathcal{M}$, $n' \in \mathcal{N}$, $\ell \in \mathcal{L}$, $(k', n') \neq (k, n)$. This condition will be referred to as *the separability condition for the n^{th} antenna of receiver k*, parallel to the separability condition for signal space alignment. According to this terminology, if the separability condition holds at the n^{th} antenna of receiver k, all the desired data streams at the n^{th} antenna of receiver k can be uniquely determined from $\tilde{Y}_n^{[k]}$. However, what we have received in the n^{th} antenna of receiver k is $Y_n^{[k]}$ which is a noisy version of $\tilde{Y}_n^{[k]}$. Therefore, to recover the desired data streams at the n^{th} antenna of receiver k, we further require to accurately estimate $\tilde{Y}_n^{[k]}$ from $Y_n^{[k]}$. To this aim, let $\mu_n^{[k]}$ denote the rational dimension of the received pseudo-vectors at the n^{th} antenna of receiver k. Apparently, $\mu_n^{[k]} \leq LMNK$. As we shall see shortly, if the rate control parameter ν in (2.30) is selected as:

$$\nu = \max_{k \in \mathcal{K}, n \in \mathcal{N}} \mu_n^{[k]}, \tag{2.37}$$

then we would be able to identify $\tilde{Y}_n^{[k]}$ in $Y_n^{[k]}$ with high probability for all $k \in \mathcal{K}$ and all $n \in \mathcal{N}$.

Each user decodes its data on different receive antennas separately. In other words, there is no cooperation among receive antennas of each user. There are ML desired data streams at the signal received by each antenna of every user. To decode each part, we treat the other parts as well as the interfering signals as i.i.d. noise and therefore as $\tau \to \infty$ the following rate is achievable for data stream (m, ℓ) of the signal received on the n^{th} antenna of receiver k:

$$R_{mn\ell}^{[k]} = I(X_{mn\ell}^{[k]}; Y_n^{[k]}) = H(X_{mn\ell}^{[k]}) - H(X_{mn\ell}^{[k]} | Y_n^{[k]}), \ \ m \in \mathcal{M}, \, l \in \mathcal{L}, \tag{2.38}$$

where for the notational simplicity, we omitted the time index t. It is obvious that:

$$H(X_{mn\ell}^{[k]}) = \log_2 |\Lambda_{mn\ell}^{[k]}(P, \epsilon)| \approx \frac{(1-\epsilon)}{\nu + 2\epsilon} \log_2 P + 1. \tag{2.39}$$

In the following, we prove that if the modulation pseudo-vectors at all transmitters are selected such that the separability condition holds at all receive antennas of all receivers, then we *almost always* have:

$$\limsup_{P \to \infty} H(X_{mn\ell}^{[k]} | Y_n^{[k]}) \leq c_0, \ \ \forall k \in \mathcal{K}, \, \forall m \in \mathcal{M}, \, \forall n \in \mathcal{N}, \, \forall \ell \in \mathcal{L}, \tag{2.40}$$

where c_0 is some constant independent of P. Consequently, user k can almost always achieve $R_{mn\ell}^{[k]} = \frac{(1-\epsilon)}{\nu+2\epsilon} \log_2 P + o(\log_2 P)$ by decoding the (m, ℓ) data stream of its desired signal component on the n^{th} receive antenna. Since there are ML desired data streams in the signal received by the n^{th} antenna of user k and since ϵ can be made arbitrarily small,

it follows that $\mathsf{DoF} \geq \frac{LMNK}{\nu}$.
Next, we show that (2.40) is valid under the above-mentioned conditions. Let

$$\Theta_n^{[k]}(P,\epsilon) \triangleq \left\{ \sum_{k'=1}^{K} \sum_{m=1}^{M} \sum_{n'=1}^{N} \sum_{\ell=1}^{L} h_{nm}^{[kk']} h_{n'm}^{[k'k']} \lambda_{mn'\ell}^{[k']} : \lambda_{mn'\ell}^{[k']} \in \Lambda_{mn'\ell}^{[k']}(P,\epsilon) \right\}, \ k \in \mathcal{K}, \ n \in \mathcal{N}. \tag{2.41}$$

Note that $\Theta_n^{[k]}(P,\epsilon)$ is the support set of the random variable $\tilde{Y}_n^{[k]}$ which is the noise-free part of $Y_n^{[k]}$. We can estimate $\tilde{Y}_n^{[k]}$ from $Y_n^{[k]}$ using the following estimator:

$$\widehat{\tilde{Y}_n^{[k]}} = \underset{\theta \in \Theta_n^{[k]}(P,\epsilon)}{\operatorname{argmin}} |Y_n^{[k]} - \theta|. \tag{2.42}$$

An error may occur using this estimation whenever the absolute value of the additive Gaussian noise $Z_n^{[k]}$ is greater than half of the minimum distance of the set $\Theta_n^{[k]}(P,\epsilon)$. That is

$$\Pr\{\widehat{\tilde{Y}_n^{[k]}} \neq \tilde{Y}_n^{[k]}\} \leq \Pr\left\{ |Z_n^{[k]}| \geq \frac{d_{\min}(\Theta_n^{[k]}(P,\epsilon))}{2} \right\} \leq 2\exp\left(-\frac{d_{\min}^2(\Theta_n^{[k]}(P,\epsilon))}{8} \right), \tag{2.43}$$

where the last inequality follows from the properties of Gaussian distribution. As we discussed earlier, if the separability condition holds at all antennas of all receivers, we can uniquely determine $X_{mn\ell}^{[k]}$ from $\tilde{Y}_n^{[k]}, \forall m \in \mathcal{M}$ and $\forall l \in \mathcal{L}$. Hence, $\Pr\{\widehat{X}_{mn\ell}^{[k]} \neq X_{mn\ell}^{[k]}\} \leq \Pr\{\widehat{\tilde{Y}_n^{[k]}} \neq \tilde{Y}_n^{[k]}\}$. Therefore, we can upper-bound $H(X_{mn\ell}^{[k]}|Y_n^{[k]})$ using the data processing and Fano's inequalities [22]:

$$\begin{aligned} H(X_{mn\ell}^{[k]}|Y_n^{[k]}) &\leq H(X_{mn\ell}^{[k]}|\hat{X}_{mnl}^{[k]}) \leq 1 + \Pr\{\hat{X}_{mn\ell}^{[k]} \neq X_{mn\ell}^{[k]}\} \log_2(|\Lambda_{mn\ell}^{[k]}(P,\epsilon)|) \\ &\leq 1 + 2\exp\left(-\frac{d_{\min}^2(\Theta_n^{[k]}(P,\epsilon))}{8} \right) \times \left[\frac{(1-\epsilon)}{\nu + 2\epsilon} \log_2 P + 1 + o(1) \right] \end{aligned} \tag{2.44}$$

Finally, we show that if ν is selected according to (2.37), then we almost always have $d_{\min}(\Theta_n^{[k]}(P,\epsilon)) \geq \varrho P^{\frac{\epsilon}{2}}$ for some constant ϱ. Accordingly, (2.40) follows from (2.44). If we select ν as in (2.37), then each $\theta_n^{[k]} \in \Theta_n^{[k]}(P,\epsilon)$ is a rational linear combination of at most

ν rationally independent complex numbers and therefore it can be expressed as:

$$\theta_n^{[k]} = \gamma P^{\frac{\nu-2+4\epsilon}{2(\nu+2\epsilon)}} \sum_{i=1}^{\nu} \delta_{ni}^{[k]} T_{ni}^{[k]}, \tag{2.45}$$

where $T_{ni}^{[k]}$'s, $i = 1, \cdots, \nu$, represent ν rationally independent received pseudo-vectors* at the n^{th} antenna of receiver k and $\delta_{ni}^{[k]}$'s , $i = 1, \cdots, \nu$ are the corresponding integer coefficients. Since at most KM independent data streams may arrive along the same received pseudo-vector $T_{ni}^{[k]}$, it follows that $|\delta_{ni}^{[k]}| \leq KMQ$. The minimum distance $d_{\min}(\Theta_n^{[k]}(P, \epsilon))$ is the minimum value of $|\theta_n^{[k]} - {\theta'}_n^{[k]}|$, $\forall \theta_n^{[k]} \in \Theta_n^{[k]}(P, \epsilon)$, $\forall {\theta'}_n^{[k]} \in \Theta_n^{[k]}(P, \epsilon) \setminus \theta_n^{[k]}$. The quantity $|\theta_n^{[k]} - {\theta'}_n^{[k]}|$ can be expressed as:

$$|\theta_n^{[k]} - {\theta'}_n^{[k]}| = \gamma P^{\frac{\nu-2+4\epsilon}{2(\nu+2\epsilon)}} \left| \sum_{i=1}^{\nu} T_{ni}^{[k]} (\delta_{ni}^{[k]} - {\delta'}_{ni}^{[k]}) \right|. \tag{2.46}$$

According to the Khintchine-Groshev Theorem, for every $\epsilon > 0$ there exists some constant c_1 such that:

$$\left| \sum_{i=1}^{\nu} T_{ni}^{[k]} (\delta_{ni}^{[k]} - {\delta'}_{ni}^{[k]}) \right| \geq \frac{c_1}{(2KMQ)^{(\nu-2)/2+\epsilon}} \tag{2.47}$$

for almost all received pseudo-vectors $T_{ni}^{[k]}$'s, $i = 1, \cdots, \nu$. Therefore, the minimum distance $d_{\min}(\Theta_n^{[k]}(P, \epsilon))$ is lower-bounded by:

$$d_{\min}(\Theta_n^{[k]}(P, \epsilon)) \geq \varrho P^{\frac{\nu-2+4\epsilon}{2(\nu+2\epsilon)}} P^{-\frac{(1-\epsilon)(\nu-2+2\epsilon)}{2(\nu+2\epsilon)}} = \varrho' P^{\frac{\epsilon}{2}} \tag{2.48}$$

for almost all received pseudo-vectors $T_{ni}^{[k]}$'s, $i = 1, \cdots, \nu$, where $\varrho' = c_1 \gamma (2KM)^{-((\nu-2)/2+\epsilon)}$ is a constant independent of P. Since the lower-bound on the minimum distance is obtained using the Khintchine-Groshev Theorem, we use the term "almost always" in statements concerning our achievability result.

So far, we established that for almost all modulation pseudo-vectors $h_{mn}^{[kk]} \omega_{mn\ell}^{[k]}$, $k \in \mathcal{K}$, $m \in \mathcal{M}$, $n \in \mathcal{N}$, $\ell \in \mathcal{L}$ satisfying the separability condition at all antennas of all

*Note that according to the separability condition, out of these ν rationally independent received pseudo-vectors, ML ones are $(h_{nm}^{[kk]})^2 \omega_{mn\ell}^{[k]}$, $m \in \mathcal{M}$, $\ell \in \mathcal{L}$.

receivers, the proposed scheme can achieve $\frac{LMNK}{\nu}$ degrees of freedom where ν represents the maximum number of rationally independent received pseudo-vectors across all receive antennas of all users. In general, ν can be as large as $LMNK$ and therefore DoF strongly depends on the value of ν. In the sequel, we show that if the modulation pseudo-vectors are properly selected according to the channel coefficients, the value of ν can approach $(M+N)L$, and consequently, $K\frac{MN}{M+N}$ degrees of freedom is almost always achievable. As mentioned earlier, reducing ν by an appropriate selection of modulation pseudo-vectors is counterpart to the alignment condition in signal space alignment. We define $\mathcal{H}_m^{[k]}$ as the set of channel coefficients from the m^{th} antenna of user k to all receive antennas of different users. That is:

$$\mathcal{H}_m^{[k]} \triangleq \{h_{1m}^{[1k]}, h_{2m}^{[1k]}, \cdots, h_{Nm}^{[1k]}, h_{1m}^{[2k]}, h_{2m}^{[2k]}, \cdots, h_{Nm}^{[2k]}, \cdots, h_{1m}^{[Kk]}, h_{2m}^{[Kk]}, \cdots, h_{Nm}^{[Kk]}\}.$$

Note that $|\mathcal{H}_m^{[k]}| = KN$, $\forall k \in \mathcal{K}$, $\forall m \in \mathcal{M}$. For each $n \in \mathcal{N}$, we define E_n as:

$$E_n \triangleq \bigcup_{k=1}^{K} \bigcup_{m=1}^{M} \left[h_{nm}^{[kk]}.(\mathcal{H}_m^{[k]} \setminus h_{nm}^{[kk]}) \right]. \tag{2.49}$$

Note that each element of E_n is the product of two channel coefficients. That is if $e \in E_n$, then e can be represented as $h_{nm}^{[kk]} h_{n'm}^{[k'k]}$ for some $k \in \mathcal{K}$, $k' \in \mathcal{K}$, $m \in \mathcal{M}$, $n' \in \mathcal{N}$ where $(k, n) \neq (k', n')$. One can verify that $|E_n| = KM(KN-1)$, $\forall n \in \mathcal{N}$. For a positive integer Γ and for each $m \in \mathcal{M}$, $n \in \mathcal{N}$, $k \in \mathcal{K}$, we select $\Omega_{mn}^{[k]}$ as:

$$\Omega_{mn}^{[k]} = \left\{ \prod_{i=1}^{|E_n|} e_i^{s_i} \ : \ e_i \in E_n, \ s_i \in \{0, 1, \cdots, \psi_{mn}^{[k]}(e_i)\} \right\}, \tag{2.50}$$

where $\psi_{mn}^{[k]}(\cdot)$ are functions described by:

$$\psi_{mn}^{[k]}(e) = \begin{cases} \Gamma - 1, & \text{if } e \in h_{nm}^{[kk]}.(\mathcal{H}_m^{[k]} \setminus h_{nm}^{[kk]}) \\ \Gamma, & \text{Otherwise} \end{cases}. \tag{2.51}$$

We claim that if the real numbers $\omega_{mn\ell}^{[k]}$ are selected from $\Omega_{mn}^{[k]}$ in (2.50), then the separability condition holds at all antennas of all receivers and moreover ν can approach $(M+N)L$.

First, we notice that elements of $\Omega_{mn}^{[k]}$ are different monomials in the variables e_i's and therefore they are almost always linearly independent. From (2.49), (2.50), and (2.51), one can verify that the number of modulation pseudo-vectors, L, which is equal to the cardinality of $\Omega_{mn}^{[k]}$, is given by

$$L = \Gamma^{KN-1}(\Gamma+1)^{(KM-1)(KN-1)}. \tag{2.52}$$

Next, consider the received signal at the n^{th} antenna of receiver k at time index t. From (2.36), we see that:

- Received pseudo-vectors corresponding to the desired component of $Y_n^{[k]}(t)$ are the elements of $\bigcup_{m=1}^{M}(h_{nm}^{[kk]})^2 \cdot \Omega_{mn}^{[k]}$.
- Received pseudo-vectors corresponding to the self-interference component of $Y_n^{[k]}(t)$ are the elements of $\mathcal{B}_n^{[k]} \triangleq \bigcup_{m=1}^{M}\bigcup_{\substack{n'=1\\n'\neq n}}^{N}\left[h_{nm}^{[kk]}h_{n'm}^{[kk]}\cdot\Omega_{mn'}^{[k]}\right]$.
- Received pseudo-vectors corresponding to the multi-user interference component of $Y_n^{[k]}(t)$ are the elements of $\mathcal{G}_n^{[k]} \triangleq \bigcup_{\substack{k'=1\\k'\neq k}}^{K}\bigcup_{m=1}^{M}\bigcup_{n'=1}^{N}\left[h_{nm}^{[kk']}h_{n'm}^{[k'k']}\cdot\Omega_{mn'}^{[k']}\right]$.

Since $(h_{nm}^{[kk]})^2 \notin E_n$, $\forall k \in \mathcal{K}$, $\forall m \in \mathcal{M}$, $\forall n \in \mathcal{N}$, it follows that the received pseudo-vectors corresponding to the desired component can not be expressed as rational linear combinations of the other received pseudo-vectors and therefore the separability condition holds at all antennas of all receivers. We then notice that:

$$\begin{array}{ll} h_{nm}^{[kk]}h_{n'm}^{[kk]} \in E_{n'}, & \forall m \in \mathcal{M},\ n' \neq n \\ h_{nm}^{[kk']}h_{n'm}^{[k'k']} \in E_{n'}, & \forall m \in \mathcal{M},\ k' \neq k \end{array}. \tag{2.53}$$

Since each element of $\Omega_{mn'}^{[k]}$, $n' \neq n$, is a monomial in the variables e_i''s where $e_i' \in E_{n'}$, and because of (2.53), each element of $\bigcup_{m=1}^{M}\left[h_{nm}^{[kk]}h_{n'm}^{[kk]}\cdot\Omega_{mn'}^{[k]}\right]$ is again a monomial in e_i''s with a degree at most Γ for each variable. Similarly, since each element of $\Omega_{mn'}^{[k']}$, $k' \neq k$ is a monomial in e_i''s where $e_i' \in E_{n'}$, and because of (2.53), each element of

$\bigcup_{\substack{k'=1\\k'\neq k}}^{K} \bigcup_{m=1}^{M} \left[h_{nm}^{[kk']} h_{n'm}^{[k'k']} \cdot \Omega_{mn'}^{[k']}\right]$ is again a monomial in e_i''s with a degree at most Γ for each variable. Hence,

$$\dim(\mathcal{B}_n^{[k]} \bigcup \mathcal{G}_n^{[k]}) \leq N(\Gamma+1)^{KM(KN-1)}. \tag{2.54}$$

Therefore,

$$\mu_n^{[k]} \leq ML + N(\Gamma+1)^{KM(KN-1)}. \tag{2.55}$$

Recall that $\mu_n^{[k]}$ is the rational dimension of the received pseudo-vectors at the n^{th} antenna of receiver k. We then have:

$$\nu \leq ML + N(\Gamma+1)^{KM(KN-1)}. \tag{2.56}$$

Therefore, from (2.52) and (2.56) the achievable DoF is given by:

$$\underline{\mathsf{DoF}} = \frac{KMN\Gamma^{KN-1}(\Gamma+1)^{(KM-1)(KN-1)}}{M\Gamma^{KN-1}(\Gamma+1)^{(KM-1)(KN-1)} + N(\Gamma+1)^{KM(KN-1)}}.$$

Noting that Γ is an arbitrary integer, as $\Gamma \to \infty$, the achievable DoF tends to $K\frac{MN}{M+N}$.

2.5 Conclusions

We studied the fully conected K-user MIMO Gaussian IC with constant channel coefficients. New results on the DoF of channel are obtained. Using real interference alignment technique, we developed a transmission scheme which can achieve a DoF value which is higher than all previously known results. We also introduced a new upper-bound on the DoF of this, which coincides with our achievable DoF when the number of users is larger than some threshold, which depends on the number of transmit and receive antennas. Our results reveal a complete characterization of DoF for a wide range of M, N, and K values.

Chapter 3

Interference alignment with Delayed CSIT

In this chapter*, we consider Gaussian interference and X channels in i.i.d. fading environment. It is assumed that transmitters have access to the CSI after a finite delay, a model which is referred to as delayed CSIT model. In this model, the transmitters knowledge of CSI becomes outdated prior to being used for the current transmission. We first study the two-user MIMO interference channel. New achievable results on the degrees of freedom (DoF) region of this channel are provided and shown to be tight for some antenna configurations. It is observed that, depending on the antenna configuration, the DoF region with delayed CSIT can collapse to the DoF region with no CSIT, strictly lie between DoF regions with no CSIT and full CSIT, or coincide with the DoF region with full CSIT. Next, we consider the two-user MIMO X channel. This is a generalization of the IC in which there is an independent message from each transmitter to each receiver. Using a new coding scheme, new achievable sum-DoFs are obtained which turn out to be tight for all cases except possibly for $1/2 < N/M < 4/3$. In specific, we show that the two-user SISO X channel can achieve a DoF of $6/5$ which is better than the previous result of $8/7$.

*Part of the work in this chapter has been presented in [46] and [47].

We then further generalize our analysis to the K-user SISO X network, a network with K transmitter-receiver pairs in which each transmitter has an independent message for each receiver. We show that one can achieve $\frac{4}{3} - \frac{2}{3(3K-1)}$ DoF over this channel. Finally, the K-user multiple-input single-output (MISO) interference channel with $M \geq K$ antennas at each transmitter is investigated under the delayed CSIT assumption wherein a new lower bound of $\frac{2K}{K+1}$ on the sum-DoF is presented. Interference alignment is the main ingredient of our transmission schemes to obtain DoF improvements over the no CSIT case. It is realized retrospectively through a multi-phase transmission scheme in which each transmitter uses its knowledge of past CSI to regulate its subsequent transmissions such that the interference subspace at each receiver is not expanded.

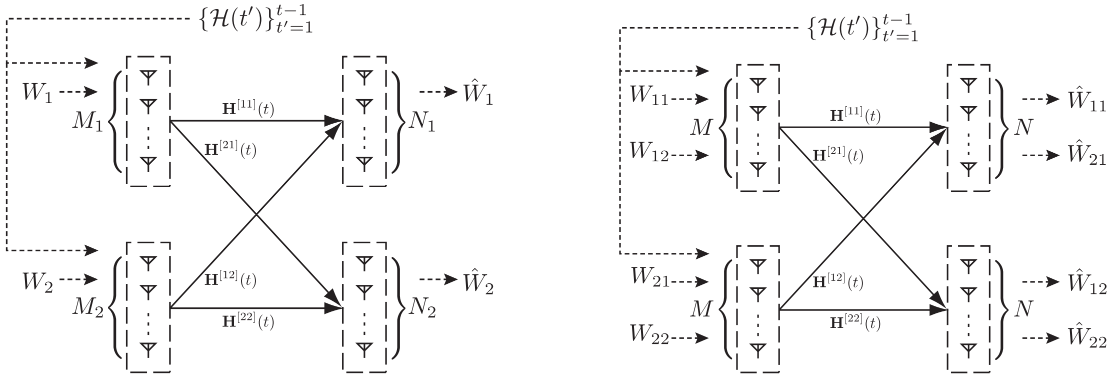
3.1 System Model

An (M_1, M_2, N_1, N_2) MIMO Gaussian IC, i.e., a MIMO interference channel with M_i antennas at transmitter i (TX_i) and N_j antennas at receiver j (RX_j), $i, j \in \{1, 2\}$, as shown in Fig. 3.1a, is described by the following input-output relationship:

$$\mathbf{Y}^{[j]}(t) = \mathbf{H}^{[j1]}(t)\mathbf{X}^{[1]}(t) + \mathbf{H}^{[j2]}(t)\mathbf{X}^{[2]}(t) + \mathbf{Z}^{[j]}(t), \quad j = 1, 2, \tag{3.1}$$

where t, $t = 1, 2, \cdots$, is the time index, $\mathbf{X}^{[i]}(t) \in \mathbb{C}^{M_i}$ is the transmitted vector of TX_i, $\mathbf{Y}^{[j]}(t) \in \mathbb{C}^{N_j}$ is the received vector at RX_j, $\mathbf{H}^{[ji]}(t) \in \mathbb{C}^{N_j \times M_i}$ is the channel matrix between TX_i and RX_j, and $\mathbf{Z}^{[j]}(t) \in \mathbb{C}^{N_j}$ is the complex AWGN vector at RX_j. Each transmitter is required to satisfy the power constraint P. TX_i wishes to transmit message $W_i \in \mathcal{W}_i \in \{1, 2, \cdots, 2^{\tau R_i(P)}\}$ with rate $R_i(P)$ to RX_i, $i = 1, 2$ over a block of τ channel uses. We further assume that the channel coefficients are i.i.d. standard complex Gaussian random variables across time and space and are independent of receivers' noise.

An (M_1, M_2, N_1, N_2) MIMO X channel, as depicted in Fig. 3.1b, is defined by the same input-output relationship (3.1) and under the power constraint P for each transmitter . In



(a) The (M_1, M_2, N_1, N_2) MIMO IC

(b) The (M, M, N, N) MIMO X channel

Figure 3.1: The two-user MIMO IC and X channel with delayed CSIT

the MIMO X channel, however, there are four independent messages: $W_{11}, W_{12}, W_{21}, W_{22}$, where W_{ij} denotes a message from TX_i to RX_j.

The X network is a generalization of X channel to the cases with more than two transmitters or receivers. A K user X network models a communication system with K transmitters and K receivers in which each transmitter has an independent message for every receiver.

The K-user MISO Gaussian IC with M antennas at each transmitter consists of K transmitter-receiver pairs in which each transmitter wishes to communicate with its intended receiver.

The knowledge of CSI at the transmitters and receivers are summarized in the following definition:

Definition 8 (Delayed CSIT for IC and X Channel)**.** *Each receiver knows all its incoming channel coefficients in time slot t, perfectly and instantaneously, while having access to the channel coefficients of the other receivers with one time slot delay. Each transmitter has access to all channel coefficients after one time slot delay via noiseless feedback links.*

Let $\mathcal{H}(t)$ denote the set of all channel coefficients at time slot t. A block code for interference channel with delayed CSIT is defined next.

Definition 9. *A $(2^{\tau\mathbf{R}}, \tau)$ code of block length τ and rate $\mathbf{R} = (R^{[1]}, R^{[2]})$ for the two-user MIMO Gaussian IC with delayed CSIT is defined as two sets of encoding functions $\{\varphi_{t,\tau}^{[k]}\}_{t=1}^{\tau}$, $1 \leq k \leq 2$, such that*

$$\mathbf{X}^{[k]}(t) = \varphi_{t,\tau}^{[k]}(W^{[k]}, \{\mathcal{H}(t')\}_{t'=1}^{t-1})$$

together with two decoding functions $\psi_{\tau}^{[k]}, 1 \leq k \leq 2$, such that

$$\hat{W}^{[k]} = \psi_{\tau}^{[k]}(\{\mathbf{Y}^{[k]}(t)\}_{t=1}^{\tau}, \{\mathcal{H}(t')\}_{t'=1}^{\tau-1}, \{\mathbf{H}^{[kj]}(\tau)\}_{j=1}^{2}) \tag{3.2}$$

The notions of achievable rate, capacity region, DoF region, and channel DoF are defined exactly as in section 2.1. In the following, the DoF region of the two-user MIMO IC with delayed CSIT will be denoted by $\mathcal{D}_{\text{IC}}^{\text{d-CSI}}$. Also, the DoF regions of this channel with full and no CSIT will be denoted by $\mathcal{D}_{\text{IC}}^{\text{f-CSI}}$ and $\mathcal{D}_{\text{IC}}^{\text{n-CSI}}$, respectively.

The definition of a block code for the MIMO X channel, the K-user X network, and the K-user MISO IC with delayed CSIT is similar to the two-user IC and will not be repeated here. Also, the notions of an achievable rate and the capacity region are defined similarly. The DoF region $\mathcal{D}_{\text{X}}$ of the two-user X channel is also defined as in the IC case for the DoF tuple $(d_{11}, d_{12}, d_{21}, d_{22})$. The sum-DoF (or simply DoF) of this channel is defined as:

$$\mathsf{DoF}_{\text{X}} \triangleq \max_{\mathcal{D}_{\text{X}}}(d_{11} + d_{12} + d_{21} + d_{22}). \tag{3.3}$$

The DoF of the X channel with delayed CSIT is denoted by $\mathsf{DoF}_{\text{X}}^{\text{d-CSI}}$. In this work, we study the MIMO X channel under the delayed CSIT assumption and with $M_1 = M_2 = M$, $N_1 = N_2 = N$, and denote its DoF by $\mathsf{DoF}_{\text{X}}^{\text{d-CSI}}(M, N)$. The DoFs region of this channel with full and no CSIT are respectively denoted by $\mathsf{DoF}_{\text{X}}^{\text{f-CSI}}(M, N)$ and $\mathsf{DoF}_{\text{X}}^{\text{n-CSI}}(M, N)$. The DoF of K-user X network is defined as

$$\mathsf{DoF}_{K\text{-X}} \triangleq \max_{\mathcal{D}_{K\text{-X}}} \sum_{i=1}^{K} \sum_{j=1}^{K} d_{ij}, \tag{3.4}$$

where $\mathcal{D}_{K\text{-X}}$ denotes the DoF region of the K-user X network. The DoF of this channel with delayed CSIT is denoted by $\mathsf{DoF}^{\text{d-CSI}}_{K\text{-X}}$. Finally, the DoF of K-user MISO IC is defined as

$$\mathsf{DoF}_{K\text{-IC}}(M) \triangleq \max_{\mathcal{D}_{K\text{-IC}}(M)} (d_1 + d_2 + \cdots + d_K), \tag{3.5}$$

where $\mathcal{D}_{K\text{-IC}}(M)$ denotes the DoF region of the K-user MISO IC with M antennas at each transmitter. The DoF of this channel with delayed CSIT is denoted by $\mathsf{DoF}^{\text{d-CSI}}_{K\text{-IC}}(M)$. In this work, we study the DoF of the K-user MISO IC under the delayed CSIT assumption and with $M \geq K$ antennas at each transmitter.

3.2 Main Results and Discussion

3.2.1 Main Results

Consider the (M_1, M_2, N_1, N_2) MIMO Gaussian IC. Without loss of generality, we assign index 2 to the user with more receive antennas, i.e., $N_2 \geq N_1$. In the case of $N_1 = N_2$, we assign index 2 to the user with less transmit antennas.

The following theorem provides an inner-bound on the DoF region of the two-user MIMO IC with delayed CSIT:

Theorem 4. *$\mathcal{D}^{d\text{-}CSI}_{IC} \supseteq \mathcal{D}^{d\text{-}CSI}_{IC,in}$, where the inner-bound on the DoF region is defined as*

$$\begin{aligned}
\mathcal{D}^{d\text{-}CSI}_{IC,in} \triangleq \Bigg\{(d_1, d_2) \in \mathbb{R}^2_+ \Bigg|\; & \mathcal{I}_1 : \quad 0 \leq d_1 \leq M_1, && \mathcal{I}_2 : \quad \frac{d_1}{N_1} + \frac{d_2}{\max(M_2', N_1)} \leq 1, \\
& \mathcal{I}_3 : \quad 0 \leq d_2 \leq M_2, && \mathcal{I}_4 : \quad \frac{d_1}{\max(M_1', N_2)} + \frac{d_2}{N_2} \leq 1, \\
& \mathcal{I}_5 : \quad (1 + \frac{L}{N_1}) d_1 + d_2 \leq N_1 + N_2 \Bigg\},
\end{aligned} \tag{3.6}$$

and

$$M'_k \triangleq \min(M_k, N_1 + N_2), \quad k = 1, 2, \tag{3.7}$$

$$L \triangleq N_1 + N_2 - M'_1. \tag{3.8}$$

The proof is presented in section 3.3.

We should point out here that some of the inequalities in (3.6) may be inactive for some antenna configurations. Moreover, the tightness of the above inner bound for some antenna configurations is summarized in the following theorem:

Theorem 5. *The achievable DoF region described in Theorem 4 is tight in the following cases:*

a) $M_2 \leq N_1$

b) $N_1 < M_1 \leq N_2$ *and* $M_2 \geq N_1 + N_2$

c) $\min(M_1, M_2) \geq N_1 + N_2$

d) $M_1 \leq \Delta < N_1 \leq N_2 < L < M_2$

e) $M_1 \leq \Delta' < N_1 \leq N_2 < M_2 \leq L,$

where

$$\Delta \triangleq \frac{N_1(N_1 - M_1)}{N_2 - M_1}, \tag{3.9}$$

$$\Delta' \triangleq \frac{N_1(M_2 - N_2)}{M_2 - N_1}. \tag{3.10}$$

The proof is presented in section 3.4.

Remark 2. *Theorem 4 has been reported in an independent and concurrent study [48]. They also showed that the achievable DoF region of Theorem 4 is tight for all antenna configurations and gives the DoF region of the Channel.*

The following theorem provides a lower-bound on the sum-DoF of the two user MIMO X channel with delayed CSIT:

Theorem 6.

$$\mathsf{DoF}_X^{d\text{-}CSI}(M,N) \geq \begin{cases} \frac{4}{3}N, & \frac{N}{M} \leq \frac{1}{2} \\ \frac{2N}{M+4N}(M+2N), & \frac{1}{2} < \frac{N}{M} < 1 \\ \frac{6}{5}N, & 1 \leq \frac{N}{M} < \frac{4}{3} \\ \frac{4MN}{2M+N}, & \frac{4}{3} \leq \frac{N}{M} < 2 \\ 2M, & 2 \leq \frac{N}{M} \end{cases}. \tag{3.11}$$

Furthermore, the above lower bound is tight for all values of M and N except possibly for $1/2 < N/M < 4/3$.

The achievability proof is presented in section 3.5. The converse proof is presented in the discussion.

Theorem 7. *For the K-user X network with delayed CSIT, $\mathsf{DoF}_X^{d\text{-}CSI} \geq \frac{4}{3} - \frac{2}{3(3K-1)}$.*

The proof is presented in section 3.6.

Finally, Theorem 8 presents our result on the DoF of the K user MISO IC with delayed CSIT:

Theorem 8. *For $M \geq K$, we have $\mathsf{DoF}_{K\text{-}IC}^{d\text{-}CSI}(M) \geq \frac{2K}{K+1}$.*

The proof is presented in section 3.7.

3.2.2 Discussion

In the two-user MIMO IC with full CSIT, the channel DoF can be achieved using beamforming at transmitters and zero-forcing at receivers. To design appropriate beamforming vectors, the knowledge of the current CSI at transmitters is crucial. In delayed CSIT model, however, the transmitters only have access to past CSI and therefore the achievable scheme of the full CSIT model is not applicable. As we shall see in our achievable scheme, the knowledge of delayed CSIT can be exploited efficiently using the idea of interference alignment. The core idea is that both transmitters first send a certain amount of information intended for their corresponding receivers. As a result, a subspace of the signal space at each receiver is occupied by interference. Then, each transmitter uses its knowledge of past CSI to regulate its subsequent transmissions such that the pre-existing interference subspace at each receiver is not expanded.

The DoF regions of the two-user MIMO IC with full CSIT and no CSIT have been previously established and are restated here:

Theorem 9 (DoF region of two-user MIMO IC with full CSIT [39]). *Let $\mathcal{D}_{IC}^{f\text{-}CSI}$ denote the DoF region of the (M_1, M_2, N_1, N_2) MIMO IC with full CSIT. Then*

$$\mathcal{D}_{IC}^{f\text{-}CSI} = \Big\{(d_1, d_2) \in \mathbb{R}_+^2 \Big|\ 0 \le d_1 \le \min(M_1, N_1), \quad 0 \le d_2 \le \min(M_2, N_2),$$
$$d_1 + d_2 \le \min\Big\{M_1 + M_2, N_1 + N_2, \max(M_1, N_2), \max(M_2, N_1)\Big\}\Big\}.$$

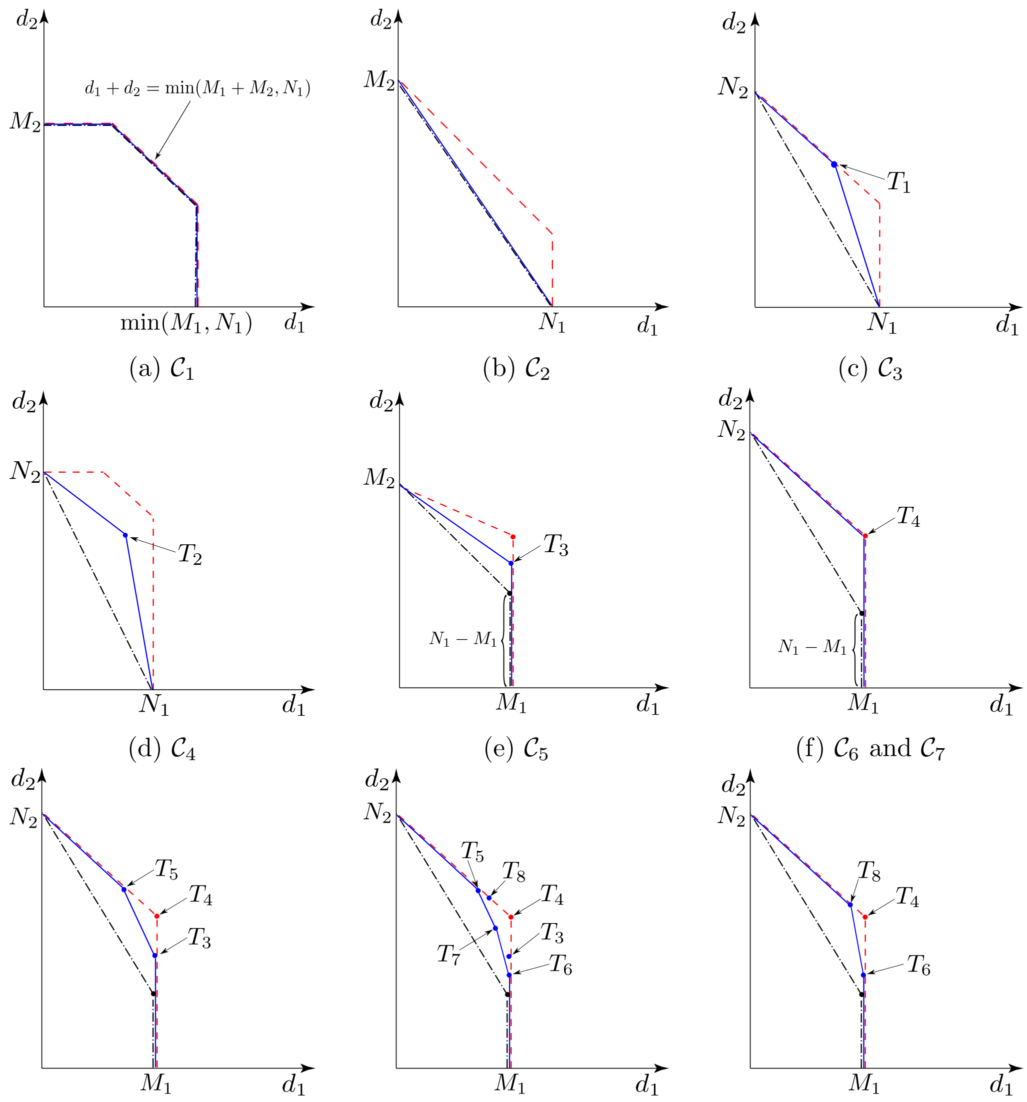
Theorem 10 (DoF region of two-user MIMO IC with no CSIT [30], [28], [49]). *Let $\mathcal{D}_{IC}^{n\text{-}CSI}$ denote the DoF region of the (M_1, M_2, N_1, N_2) MIMO IC with no CSIT. Then for $N_2 \ge N_1$*

$$\mathcal{D}_{IC}^{n\text{-}CSI} = \Big\{(d_1, d_2) \in \mathbb{R}_+^2 \Big|\ 0 \le d_1 \le \min(M_1, N_1), \quad 0 \le d_2 \le \min(M_2, N_2),$$
$$d_1 + \frac{\min(M_2, N_1) - A}{\min(M_2, N_2) - A}(d_2 - A) \le \min(M_1, N_1)\Big\},$$

where $A \triangleq \min(M_1 + M_2, N_1) - \min(M_1, N_1)$.

To examine our achievable DoF region for the two-user MIMO IC with delayed CSIT, we consider 10 possibilities for different values of M_1, M_2, N_1, and N_2, as summarized in Table 3.1. These 10 classes cover all antenna configurations and are mutually exclusive. For each class in Table 3.1, our achievable DoF region under the delayed CSIT assumption is presented and compared with the DoF regions under full CSIT and no CSIT assumptions. Fig. 3.2 also shows our achievable DoF region for each of these 10 classes. For each class the achievable DoF region is a polygon whose corner points are labeled in the figure. Some comments are in order:

- Our achievable DoF region with delayed CSIT is larger that the DoF region with no CSIT except for classes $\mathcal{C}_1$ and $\mathcal{C}_2$.
- Delayed CSIT does not incur any loss in DoF compared to the full CSIT in classes $\mathcal{C}_6$ and $\mathcal{C}_7$.
- In all antenna configurations except for class $\mathcal{C}_4$, the sum-DoF of the two-user MIMO IC is the same regardless of the knowledge of CSI at transmitters. For class $\mathcal{C}_4$, however, our achievable sum-DoF with delayed CSIT is strictly greater than the channel sum-DoF with no CSIT. Moreover, based on the upper bound developed in the proof of Theorem 5, one can conclude that, for this class, the channel sum-DoF with delayed CSIT is also strictly less than that with full CSIT. For example, when $\min(M_1, M_2) \geq N_1 + N_2$, the sum-DoF with no CSIT is equal to $\max(N_1, N_2)$ while with delayed CSIT it is equal to $\frac{(N_1^2+N_2^2)(N_1+N_2)}{N_1^2+N_2^2+N_1N_2}$, and with full CSIT it is equal to $N_1 + N_2$.
- In delayed CSIT model, each transmitter is assumed to have all the channel matrices with a unit delay. However, in our achievable scheme, each transmitter only requires to know a delayed version of its own channel matrices to all receivers (delayed local CSIT).



(a) $\mathcal{C}_1$ (b) $\mathcal{C}_2$ (c) $\mathcal{C}_3$

(d) $\mathcal{C}_4$ (e) $\mathcal{C}_5$ (f) $\mathcal{C}_6$ and $\mathcal{C}_7$

(g) $\mathcal{C}_8$ (h) $\mathcal{C}_9$ (i) $\mathcal{C}_{10}$

Corner Point (d_1, d_2)	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8
d_1	$\frac{N_1(M_2'-N_2)}{M_2'-N_1}$	$\frac{M_1'N_1(M_2'-N_2)}{M_1'M_2'-N_1N_2}$	M_1	M_1	$\frac{N_1(M_2-N_2)}{M_2-N_1}$	M_1	$\frac{N_1(N_1+N_2-M_2)}{N_1+L-M_2}$	$\frac{N_1^2}{L}$
d_2	$\frac{M_2'(N_2-N_1)}{M_2'-N_1}$	$\frac{M_2'N_2(M_1'-N_1)}{M_1'M_2'-N_1N_2}$	$\frac{M_2(N_1-M_1)}{N_1}$	$N_2 - M_1$	$\frac{M_2(N_2-N_1)}{M_2-N_1}$	$\frac{L(N_1-M_1)}{N_1}$	$\frac{M_2(N_1-M_1)}{N_1+L-M_2}$	$N_2 - \frac{N_1^2}{L}$

Figure 3.2: The achievable DoF region for the two-user MIMO IC with $N_2 \geq N_1$ and delayed CSIT (solid line). The DoF region of the same channel with no CSIT (dash-dot line) and full CSIT (dashed line) are also presented for comparison.

Table 3.1: Achievable DoF region of the (M_1, M_2, N_1, N_2) MIMO IC with delayed CSIT when $N_2 \geq N_1$.

Class	Class Definition	Achievable region $\mathcal{D}_{\text{IC,in}}^{\text{d-CSI}}$	Comparison
$\mathcal{C}_1$	$M_2 \leq N_1 \leq N_2$	$d_1 \leq M_1, d_2 \leq M_2$ $d_1 + d_2 \leq N_1$	$\mathcal{D}_{\text{IC}}^{\text{n-CSI}} = \mathcal{D}_{\text{IC}}^{\text{d-CSI}} = \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$
$\mathcal{C}_2$	$N_1 < M_1,\ N_1 < M_2 \leq N_2$	$\frac{d_1}{N_1} + \frac{d_2}{M_2} \leq 1$	$\mathcal{D}_{\text{IC}}^{\text{n-CSI}} \subseteq \mathcal{D}_{\text{IC}}^{\text{d-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$
$\mathcal{C}_3$	$N_1 \leq M_1 < N_2 < M_2$	$d_1 + d_2 \leq N_2$ $\frac{d_1}{N_1} + \frac{d_2}{M_2'} \leq 1$	$\mathcal{D}_{\text{IC}}^{\text{n-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{d-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$
$\mathcal{C}_4$	$N_1 \leq N_2 < \min(M_1, M_2)$	$\frac{d_1}{M_1'} + \frac{d_2}{N_2} \leq 1$ $\frac{d_1}{N_1} + \frac{d_2}{M_2'} \leq 1$	$\mathcal{D}_{\text{IC}}^{\text{n-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{d-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$
$\mathcal{C}_5$	$M_1 < N_1 < M_2 < N_2$	$d_1 \leq M_1$ $\frac{d_1}{N_1} + \frac{d_2}{M_2} \leq 1$	$\mathcal{D}_{\text{IC}}^{\text{n-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{d-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$
$\mathcal{C}_6$	$M_1 \leq \Delta \leq N_1 < N_2 < L < M_2$	$d_1 \leq M_1$ $d_1 + d_2 \leq N_2$	$\mathcal{D}_{\text{IC}}^{\text{n-CSI}} \subseteq \mathcal{D}_{\text{IC}}^{\text{d-CSI}} = \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$
$\mathcal{C}_7$	$M_1 \leq \Delta' \leq N_1 < N_2 < M_2 \leq L$	$d_1 \leq M_1$ $d_1 + d_2 \leq N_2$	$\mathcal{D}_{\text{IC}}^{\text{n-CSI}} \subseteq \mathcal{D}_{\text{IC}}^{\text{d-CSI}} = \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$
$\mathcal{C}_8$	$\Delta' < M_1 < N_1 < N_2 \leq M_2 \leq L$	$d_1 \leq M_1,\ d_1 + d_2 \leq N_2$ $\frac{d_1}{N_1} + \frac{d_2}{M_2} \leq 1$	$\mathcal{D}_{\text{IC}}^{\text{n-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{d-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$
$\mathcal{C}_9$	$\Delta < M_1 \leq N_1 < N_2 \leq L < M_2 \leq N_1 + N_2 - \Delta$	$d_1 \leq M_1,\ d_1 + d_2 \leq N_2$ $\frac{d_1}{N_1} + \frac{d_2}{M_2} \leq 1$ $(1+\frac{L}{N_1})d_1 + d_2 \leq N_1 + N_2$	$\mathcal{D}_{\text{IC}}^{\text{n-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{d-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$
$\mathcal{C}_{10}$	$\Delta < M_1 < N_1 \leq N_2 < L < N_1 + N_2 - \Delta < M_2$	$d_1 \leq M_1$ $d_1 + d_2 \leq N_2$ $(1+\frac{L}{N_1})d_1 + d_2 \leq N_1 + N_2$	$\mathcal{D}_{\text{IC}}^{\text{n-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{d-CSI}} \subset \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$

From [14] and [50], the DoF of the two-user MIMO X channel with full CSIT is given by:

$$\mathsf{DoF}_{\mathrm{X}}^{\text{f-CSI}}(M,N) = \min\left\{2\min(M,N), \frac{4}{3}\max(M,N)\right\}. \tag{3.12}$$

On the other hand, it was shown in [49] that with no CSIT, the DoF of this channel collapses to:

$$\mathsf{DoF}_{\mathrm{X}}^{\text{n-CSI}}(M,N) = \min(N, 2M). \tag{3.13}$$

If we allow the transmitters in the (M, M, N, N) MIMO X channel to cooperate, we reach to a two-user MIMO broadcast channel with $2M$ antennas at the transmitter and N antennas at each receiver whose DoF region with delayed CSIT was characterized in [33] as:

$$\frac{d_1}{\min(2M, 2N)} + \frac{d_2}{\min(2M, N)} \leq 1, \tag{3.14}$$

$$\frac{d_1}{\min(2M, N)} + \frac{d_2}{\min(2M, 2N)} \leq 1. \tag{3.15}$$

Since cooperation cannot shrink the capacity region, the above DoF region can serve as an outer-bound for the DoF region of the MIMO X channel with delayed CSIT by just replacing d_1 by $d_{11} + d_{12}$ and replacing d_2 by $d_{21} + d_{22}$. From (3.14) and (3.15), one can obtain the sum-DoF of the MIMO BC channel with delayed CSIT which is an upper-bound for the sum-DoF of the MIMO X channel:

$$\mathsf{DoF}_{\mathrm{BC}}^{\text{d-CSI}}(2M,N) = \frac{4\min(2M,N)\min(M,N)}{\min(2M,N) + 2\min(M,N)}. \tag{3.16}$$

Our achievable DoF results for the MIMO X channel with delayed CSIT are summarized in Table 3.2 along with the channel DoFs with full and no CSIT and also the broadcast upper-bound (3.16). From the table, one can make the following observations:

- For $\frac{N}{M} < 2$, delayed CSIT improves the channel sum-DoF compared to the no CSIT case.

- For $\frac{N}{M} \leq \frac{1}{2}$ or $\frac{N}{M} \geq \frac{4}{3}$, our achievable sum-DoF is tight and characterize the channel sum-DoF. It is interesting to note that the channel sum-DoF under the delayed CSIT assumption lies strictly between channel sum-DoF with full CSIT and channel sum-DoF with no CSIT. Also, for these cases the sum-DoF of the X channel with delayed CSIT coincides with sum-DoF of the BC obtained by allowing cooperation between transmitters.

- When $M = N$, our achievable DoF is equal to $\frac{6}{5}N$. Therefore, the DoF of SISO X channel is lower-bound by $\frac{6}{5}$. This is strictly better than the previously reported DoF of $\frac{8}{7}$ in [35].

- The DoF of the MIMO X channel with feedback and delayed CSIT has been characterized in [51]. Comparing their results with Theorem 6 reveals that in the presence of delayed CSIT, feedback can only increase DoF for $1/2 < N/M < 4/3$. We do not know at this stage whether the gap between the achieved DoF and the outer bound is due to the weakness of the coding scheme or a new outer bound is expected.

In [19], it has been proved that the K-user X network with a single antenna at each node and with full CSIT has $\frac{K^2}{2K-1}$ degrees of freedom. Without CSIT, however, the DoF of this channel collapses to one [49]. From Theorem 7, we can conclude that the DoF of the single-antenna K-user X network with delayed CSIT is strictly greater than that with no CSIT. Better achievable results on the DoF of X networks with $K \geq 3$ users have been recently presented in [36]..

The K-user MISO IC with $M \geq K$ antennas at each transmitter is known to have K degrees of freedom with full CSIT. With no CSIT, however, the DoF of this channel collapses to one. From Theorem 8, one can see that the DoF of this channel with delayed CSIT is strictly greater than one. In our achievable scheme, each transmitter only needs to know its channel matrices to the other receivers with a unit delay and global delayed CSIT is *not* required at transmitters. To obtain a simple upper-bound for the DoF of this channel, we allow transmitters to cooperate. Since cooperation does not reduce the

Table 3.2: DoF of the MIMO X Channel with M antennas at each transmitter and N antennas at each receiver

Case No.	$\mathsf{DoF}_{\mathrm{X}}^{\text{d-CSI},ach}(M,N)$	$\mathsf{DoF}_{\mathrm{X}}^{\text{f-CSI}}(M,N)$	$\mathsf{DoF}_{\mathrm{X}}^{\text{n-CSI}}(M,N)$	$\mathsf{DoF}_{\mathrm{BC}}^{\text{d-CSI}}(2M,N)$
$\frac{N}{M} \leq \frac{1}{2}$	$\frac{4}{3}N$	$2N$	N	$\frac{4}{3}N$
$\frac{1}{2} < \frac{N}{M} < 1$	$\frac{2N}{M+4N}(M+2N)$	$\min(2N, \frac{4M}{3})$	N	$\frac{4}{3}N$
$1 \leq \frac{N}{M} < \frac{4}{3}$	$\frac{6}{5}N$	$\min(2M, \frac{4N}{3})$	N	$\frac{4MN}{2M+N}$
$\frac{4}{3} \leq \frac{N}{M} < 2$	$\frac{4MN}{2M+N}$	$\min(2M, \frac{4N}{3})$	N	$\frac{4MN}{2M+N}$
$2 \leq \frac{N}{M}$	$2M$	$2M$	$2M$	$2M$

capacity, the DoF of the K-user MISO IC with M antennas at each transmitter and delayed CSIT is upper bounded by the DoF of a MK-user MISO broadcast channel with delayed CSIT, which from [52], is equal to $\frac{K}{1+\frac{1}{2}+\cdots+\frac{1}{K}}$.

3.3 Proof of Theorem 4

In this section, we prove the DoF region stated in Theorem 1 is achievable. For illustration purpose, we first elaborate on our achievable scheme for a two-user MIMO IC with two antennas at each transmitter and a single antenna at each receiver. We then present our achievable scheme for general setting.

3.3.1 An Illustrative Example

Consider a two-user MIMO IC with $M_1 = M_2 = 2$ and $N_1 = N_2 = 1$. We first notice that the DoF region of this channel with perfect CSIT is the unit square characterized by $d_i \leq 1$, $i = 1, 2$. Also, the DoF region with no CSIT is the time division region described by $d_1 + d_2 \leq 1$. These regions are depicted in Fig. 3.3. In the following, we show that the DoF region of this channel with delayed CSIT is a quadrilateral whose corner points are $(0,0)$, $(0,1)$, $(1,0)$, and $(2/3, 2/3)$ as depicted in Fig. 3.3. Since the corner points $(1,0)$ and $(0,1)$ are trivially achievable, we only need to show the achievability of $(d_1, d_2) = (2/3, 2/3)$. To this end, we propose a transmission scheme which operates in two distinct phases over three consecutive channel uses:

Phase I: This phase takes one channel use in which each transmitter sends two independent coded symbols for its intended receiver. Specifically, let us assume that TX_1 sends the symbol $u_r^{[1]}, r = 1,2$, over its r^{th} transmit antenna while TX_2 sends $u_s^{[2]}, s = 1,2$, over its s^{th} transmit antenna. By neglecting the noise terms at the receiver side, the following signals are observed by the receivers:

$$\begin{aligned} y^{[1]}(1) &= h_{11}^{[11]}(1)u_1^{[1]} + h_{12}^{[11]}(1)u_2^{[1]} + \underbrace{h_{11}^{[12]}(1)u_1^{[2]} + h_{12}^{[12]}(1)u_2^{[2]}}_{I^{[1]}(1)}, \\ y^{[2]}(1) &= \underbrace{h_{11}^{[21]}(1)u_1^{[1]} + h_{12}^{[21]}(1)u_2^{[1]}}_{I^{[2]}(1)} + h_{11}^{[22]}(1)u_1^{[2]} + h_{12}^{[22]}(1)u_2^{[2]}, \end{aligned} \tag{3.17}$$

where $I^{[1]}(1)$ and $I^{[2]}(1)$ are defined as:

$$\begin{aligned} I^{[1]}(1) &\triangleq h_{11}^{[12]}(1)u_1^{[2]} + h_{12}^{[12]}(1)u_2^{[2]}, \\ I^{[2]}(1) &\triangleq h_{11}^{[21]}(1)u_1^{[1]} + h_{12}^{[21]}(1)u_2^{[1]}. \end{aligned}$$

According to the delayed CSIT assumption, each transmitter has access to the channel coefficients by a unit delay. Therefore, TX_1 and TX_2 have respectively access to $I^{[2]}(1)$ and $I^{[1]}(1)$ by the end of this phase. From (3.17), one can observe that if we deliver both $I^{[1]}(1)$ and $I^{[2]}(1)$ to RX_1 then it will be able to resolve its desired information symbols ($u_1^{[1]}$ and $u_2^{[1]}$). A similar observation can be made for RX_2. Hence, our goal in Phase II boils down to delivering $I^{[1]}(1)$ and $I^{[2]}(1)$ to both receivers.

Phase II: In this phase, we deliver $I^{[1]}(1)$ and $I^{[2]}(1)$ to both receivers. This can be simply accomplished in two channel uses by time division.

Since each transmitter has sent two independent information symbols for its intended receiver in three channel uses, the DoF pair $(2/3, 2/3)$ has been achieved. In Section 3.4,

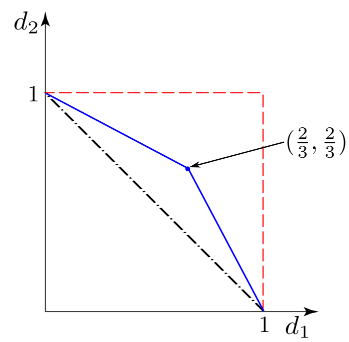


Figure 3.3: The DoF region of the two-user MIMO IC with $M_1 = M_2 = 2$ and $N_1 = N_2 = 1$ and under different assumption on CSIT information: full CSIT (dashed line), delayed CSIT (solid line), no CSIT (dash-dot line)

we will prove that the above region is indeed the DoF region of the $(2, 2, 1, 1)$ MIMO IC with delayed CSIT.

3.3.2 Proof of Achievability for General Setting

We now proceed to prove our achievability result for the general (M_1, M_2, N_1, N_2) MIMO IC. Our transmission scheme consists of W consecutive channel uses during which each transmitter operates in two distinct phases. Specifically, TX_k, $k = 1, 2$, has two phases of transmission: Phase I-k which spans the first W_k channel uses and Phase II-k which takes the remaining $W - W_k$ channel uses. In the following, we describe the transmission phases for TX_k, $k = 1, 2$, in detail:

Phase I-k: TX_k sends random linear combinations of μ_k information symbols for its intended receiver at each channel use of this phase, where $\mu_k \leq M'_k W_k$ will be determined later and M'_k is defined by (3.7). Due to the interference, it is not generally possible for RX_k to solve the equations received in Phase I-k for its desired information symbols. In Phase II-k, TX_k exploits its knowledge of the past CSI to provide the receivers with useful equations that can eventually help them to resolve their intended symbols.

Phase II-k: No fresh information symbol is sent over the channel during this phase. According to the delayed CSIT assumption, TX_k has access to all channel coefficients of Phase I-k, and thereby, is aware of the interference terms observed by its non-intended receiver ($\mathrm{RX}_{\bar{k}}$) during that phase. By retransmission of these interference terms, TX_k enables its non-intended receiver to cancel the effect of interference observed in Phase I-k. Specifically, let $I_n^{[k]}(j)$, $1 \leq n \leq N_k$, denote the interference observed by the n^{th} receive antenna of RX_k at the j^{th} channel use. At each channel use of Phase II-k, TX_k transmits M_k random linear combinations of the $N_{\bar{k}}W_k$ interference terms $\{I_n^{[k]}(j) : 1 \leq n \leq N_{\bar{k}},\ 1 \leq j \leq W_k\}$ over its transmit antennas.

Each term $I_n^{[1]}(j)$, $1 \leq j \leq W_2$, contains μ_2 independent symbols transmitted by TX_2 in Phase I-2. Similarly, $I_n^{[2]}(j)$, $1 \leq j \leq W_1$, contains μ_1 independent symbols transmitted by TX_1 in Phase I-1. Hence, during Phase I-2, RX_1 observes $\min(\mu_2, N_1W_2)$ independent interference terms across all its receive antennas. Likewise, RX_2 observes $\min(\mu_1, N_2W_1)$ independent interference terms across all its receive antennas during Phase I-1.

Before we go into the details, we show that inequalities $\mathcal{I}_1 - \mathcal{I}_4$ in (3.6) can be easily obtained using an equation counting argument in our achievable scheme: RX_1 needs to decode its μ_1 desired information symbols transmitted by TX_1 during Phase I-1 along with the $\min(\mu_2, N_1W_2)$ independent interference terms caused by TX_2 during Phase I-2. Since there are N_1 antennas at RX_1, it obtains N_1W equations over the entire transmission scheme. Therefore, a necessary condition for RX_1 to resolve the desired variables and interference terms is given by:

$$\mu_1 + \min(\mu_2, N_1W_2) \leq N_1W. \tag{3.18}$$

Similarly, the following condition is necessary at RX_2:

$$\mu_2 + \min(\mu_1, N_2W_1) \leq N_2W. \tag{3.19}$$

By selecting $\mu_k = M'_kW_k$, $k = 1, 2$, in our transmission scheme and in view of the fact that $d_1 = \frac{\mu_1}{W}$ and $d_2 = \frac{\mu_2}{W}$, (3.18) and (3.19) yield the inequalities $\mathcal{I}_2$ and $\mathcal{I}_4$ in (3.6).

From $\frac{W_i}{W} \leq 1$, $k = 1, 2$, it follows that $d_1 \leq M_1$ and $d_2 \leq M_2$, which are the same as the inequalities $\mathcal{I}_1$ and $\mathcal{I}_3$ in (3.6). Since for classes $\mathcal{C}_3 - \mathcal{C}_8$, the inequality $\mathcal{I}_5$ in (3.6) is not active, the proof of Theorem 4 for these classes boils down to the sufficiency of the above equation counting argument. For classes $\mathcal{C}_9$ and $\mathcal{C}_{10}$, however, the above equation counting argument is not sufficient and we need to consider the additional inequality $\mathcal{I}_5$.

Next, we derive sufficient conditions on μ_1, μ_2, W_1, W_2, and W that guarantee the achievability of DoF pair $(d_1, d_2) = (\frac{\mu_1}{W}, \frac{\mu_2}{W})$ by our transmission scheme. To this end, we need to ensure the rank of $N_k W$ equations available at RX_k, $k = 1, 2$, is not less than the total number of desired quantities in that receiver. Let $\mathbf{U}^{[k]}$ represent the vector containing all the information symbols of TX_k, i.e.,

$$\mathbf{U}^{[1]} \triangleq \left[u_1^{[1]}, u_2^{[1]}, \cdots, u_{\mu_1}^{[1]}\right]^T, \tag{3.20}$$

$$\mathbf{U}^{[2]} \triangleq \left[u_1^{[2]}, u_2^{[2]}, \cdots, u_{\mu_2}^{[2]}\right]^T. \tag{3.21}$$

Also, let $\mathbf{X}^{[k]}_{[m:n]}$, $n \geq m$, denote the vector containing all the transmitted signals by TX_k during the interval $t = m, m+1, \cdots, n$. That is,

$$\mathbf{X}^{[k]}_{[m:n]} \triangleq [(\mathbf{X}^{[k]}(m))^T, (\mathbf{X}^{[k]}(m+1))^T, \cdots, (\mathbf{X}^{[k]}(n))^T]^T. \tag{3.22}$$

According to our transmission scheme, at each channel use, the information symbols of each transmitter are multiplied by some precoding matrix before transmission. Let $\mathbf{F}^{[k]}_{[m:n]}$ represent the precoding matrices used by TX_k during the interval $t = m, m+1, \cdots, n$, i.e.,

$$\mathbf{X}^{[k]}_{[m:n]} = \mathbf{F}^{[k]}_{[m:n]} \mathbf{U}^{[k]}. \tag{3.23}$$

Specifically, $\mathbf{F}^{[1]}_{[1:W_1]}$ and $\mathbf{F}^{[1]}_{[W_1+1:W]}$ are respectively the precoding matrices of TX_1 in the first and second phase of transmission with respective sizes $M_1 W_1 \times \mu_1$ and $M_1(W - W_1) \times \mu_1$. Also, $\mathbf{F}^{[2]}_{[1:W_2]}$ and $\mathbf{F}^{[2]}_{[W_2+1:W]}$ are respectively the precoder matrices of TX_2 in the first and second phase of transmission with respective sizes $M_2 W_2 \times \mu_2$ and $M_2(W - W_2) \times \mu_2$. It is important to mention that the precoding matrices $\mathbf{F}^{[1]}_{[1:W_1]}$, $\mathbf{F}^{[2]}_{[1:W_2]}$ are randomly selected and reveal to all nodes ahead of transmission. The elements of $\mathbf{F}^{[1]}_{[W_1+1:W]}$ and $\mathbf{F}^{[2]}_{[W_2+1:W]}$, however, are selected according to the delayed CSIT as we will see in the following. We define $\mathbf{H}^{[ji]}_{[m:n]}$, $n \geq m$, as a block diagonal matrix containing the channel matrices from TX_i to RX_j during the interval $t = m, m+1, \cdots, n$, i.e.,

$$\mathbf{H}^{[ji]}_{[m:n]} \triangleq \mathbf{diag}\Big(\mathbf{H}^{[ji]}(m), \mathbf{H}^{[ji]}(m+1), \cdots, \mathbf{H}^{[ji]}(n)\Big). \tag{3.24}$$

Since at phase two, each transmitter sends random linear combinations of interference terms observed by its non-intended receiver during phase one, the precoding matrices $\mathbf{F}^{[k]}_{[W_k+1:W]}$, $k = 1, 2$, can be expressed as:

$$\begin{aligned} \mathbf{F}^{[1]}_{[W_1+1:W]} &= \mathbf{G}^{[1]}_{[W_1+1:W]}\mathbf{H}^{[21]}_{[1:W_1]}\mathbf{F}^{[1]}_{[1:W_1]}, \\ \mathbf{F}^{[2]}_{[W_2+1:W]} &= \mathbf{G}^{[2]}_{[W_2+1:W]}\mathbf{H}^{[12]}_{[1:W_2]}\mathbf{F}^{[2]}_{[1:W_2]}, \end{aligned} \tag{3.25}$$

where $\mathbf{G}^{[1]}_{[W_1+1:W]}$ and $\mathbf{G}^{[2]}_{[W_2+1:W]}$ are random matrices with respective sizes $M_1(W - W_1) \times N_2W_1$ and $M_2(W - W_2) \times N_1W_2$.

The vector $\mathbf{Y}^{[k]}_{[1:W]}$ which contains all the N_kW received signals of RX_k can be expressed as:

$$\begin{aligned} \mathbf{Y}^{[1]}_{[1:W]} &= \mathbf{H}^{[11]}_{[1:W]}\mathbf{X}^{[1]}_{[1:W]} + \mathbf{H}^{[12]}_{[1:W]}\mathbf{X}^{[2]}_{[1:W]}, \\ \mathbf{Y}^{[2]}_{[1:W]} &= \mathbf{H}^{[21]}_{[1:W]}\mathbf{X}^{[1]}_{[1:W]} + \mathbf{H}^{[22]}_{[1:W]}\mathbf{X}^{[2]}_{[1:W]}. \end{aligned} \tag{3.26}$$

The interference terms observed by RX_1 during Phase I-2 are given by $\mathbf{H}^{[12]}_{[1:W_2]}\mathbf{F}^{[2]}_{[1:W_2]}\mathbf{U}^{[2]}$. Since the rank of $\mathbf{H}^{[12]}_{[1:W_2]}\mathbf{F}^{[2]}_{[1:W_2]}$ is equal to $\min(N_1W_2, \mu_2)$[§], all these terms can be expressed in terms of $\min(N_1W_2, \mu_2)$ independent interference terms. More precisely, the matrix $\mathbf{H}^{[12]}_{[1:W_2]}\mathbf{F}^{[2]}_{[1:W_2]}$ can be decomposed as:

$$\mathbf{H}^{[12]}_{[1:W_2]}\mathbf{F}^{[2]}_{[1:W_2]} = \mathbf{L}^{[12]}_{[1:W_2]}\mathbf{R}^{[12]}_{[1:W_2]}, \tag{3.27}$$

where $\mathbf{L}^{[12]}_{[1:W_2]}$ and $\mathbf{R}^{[12]}_{[1:W_2]}$ are of size $N_1W_2 \times \min(N_1W_2, \mu_2)$ and $\min(N_1W_2, \mu_2) \times \mu_2$ respectively. Such a decomposition is trivial because one of $\mathbf{L}^{[12]}_{[1:W_2]}$ and $\mathbf{R}^{[12]}_{[1:W_2]}$ is indeed the

[§]Note that $\mu_2 \leq M_2W_2$.

identity matrix and the other one is $\mathbf{H}^{[12]}_{[1:W_2]}\mathbf{F}^{[2]}_{[1:W_2]}$. The vector of independent interference terms at RX_1 can then be defined as:

$$\mathbf{I}^{[1]} \triangleq \mathbf{R}^{[12]}_{[1:W_2]}\mathbf{U}^{[2]} \tag{3.28}$$

By repeating the above argument for RX_2, the vector of independent interference terms at RX_2 is defined as:

$$\mathbf{I}^{[2]} \triangleq \mathbf{R}^{[21]}_{[1:W_1]}\mathbf{U}^{[1]}, \tag{3.29}$$

where $\mathbf{R}^{[21]}_{[1:W_1]}$ is obtained using a decomposition similar to (3.27):

$$\mathbf{H}^{[21]}_{[1:W_1]}\mathbf{F}^{[1]}_{[1:W_1]} = \mathbf{L}^{[21]}_{[1:W_1]}\mathbf{R}^{[21]}_{[1:W_1]}. \tag{3.30}$$

RX_k wishes to decode the vector of information symbols $\mathbf{U}^{[k]}$; however, to this end, it requires to decode the vector $\mathbf{I}^{[k]}$ of independent interference terms as well. Let us define the vector $\mathbf{E}^{[k]}$ as the vector containing all the variables that should be decoded at RX_k, i.e.,

$$\mathbf{E}^{[k]} \triangleq [(\mathbf{U}^{[k]})^T, (\mathbf{I}^{[k]})^T]^T. \tag{3.31}$$

Note that $\mathbf{E}^{[1]}$ and $\mathbf{E}^{[2]}$ are of sizes $\mu_1+\min(N_1W_2, \mu_2)$ and $\mu_2+\min(N_2W_1, \mu_1)$ respectively. Using (3.23), (3.24) and (3.25), we can re-express (3.26) in the following form:

$$\mathbf{Y}^{[k]}_{[1:W]} = \mathbf{P}^{[k]}\mathbf{E}^{[k]}, \quad k = 1, 2, \tag{3.32}$$

where the coefficient matrices $\mathbf{P}^{[1]}$ and $\mathbf{P}^{[2]}$ can be represented in the following forms:

$$\begin{aligned}
\mathbf{P}^{[1]} &= \left(\begin{array}{c|c} \mathbf{H}^{[11]}_{[1:W_1]}\mathbf{F}^{[1]}_{[1:W_1]} & \mathbf{L}^{[12]}_{[1:W_2]} \\ \hline \mathbf{H}^{[11]}_{[W_1+1:W]}\mathbf{F}^{[1]}_{[W_1+1:W]} & \mathbf{H}^{[12]}_{[W_2+1:W]}\mathbf{G}^{[2]}_{[W_2+1:W]}\mathbf{L}^{[12]}_{[1:W_2]} \end{array}\right), \\
\mathbf{P}^{[2]} &= \left(\begin{array}{c|c} \mathbf{H}^{[22]}_{[1:W_2]}\mathbf{F}^{[2]}_{[1:W_2]} & \mathbf{L}^{[21]}_{[1:W_1]} \\ \hline \mathbf{H}^{[22]}_{[W_2+1:W]}\mathbf{F}^{[2]}_{[W_2+1:W]} & \mathbf{H}^{[21]}_{[W_1+1:W]}\mathbf{G}^{[1]}_{[W_1+1:W]}\mathbf{L}^{[21]}_{[1:W_1]} \end{array}\right),
\end{aligned} \tag{3.33}$$

where it is assumed that $W_1 < W_2$[§].

To ensure RX_k can obtain $\mathbf{E}^{[k]}$ from (3.32), we need to show that the rank of matrix $\mathbf{P}^{[k]}$ is not less than the size of $\mathbf{E}^{[k]}$. That is,

$$\mathrm{rank}(\mathbf{P}^{[k]}) \geq \mu_k + \min(N_k W_{\bar{k}}, \mu_{\bar{k}}), \quad k = 1, 2. \tag{3.34}$$

We define:

$$\begin{aligned} u &\triangleq \arg\max_k \{W_k\}, \\ \ell &\triangleq \{1, 2\} \backslash \{u\}. \end{aligned} \tag{3.35}$$

In Appendix B.1, it is proved that the rank of $\mathbf{P}^{[k]}$ is given by:

$$\mathrm{rank}(\mathbf{P}^{[k]}) = \min \left\{ \mu_k + \min(\mu_{\bar{k}}, N_k W_{\bar{k}}) \, , \, r_1^{[k]} + r_2^{[k]} + r_3^{[k]} \right\}, \tag{3.36}$$

where

$$\begin{aligned} r_1^{[k]} &\triangleq \min \left\{ N_k W_\ell, \mu_\ell + M_u W_\ell, \mu_1 + \mu_2 \right\} \\ r_2^{[k]} &\triangleq \min \left\{ N_k (W_u - W_\ell), \min \left\{ M_\ell (W_u - W_\ell), N_u W_\ell, \mu_\ell \right\} + \min \left\{ M_u (W_u - W_\ell), \mu_u \right\} \right\} \\ r_3^{[k]} &\triangleq \min \left\{ N_k (W - W_u), \min \left\{ M_k (W - W_u), N_{\bar{k}} W_k, \mu_k \right\} + \min \left\{ M_{\bar{k}} (W - W_u), N_k W_{\bar{k}}, \mu_{\bar{k}} \right\} \right\}. \end{aligned} \tag{3.37}$$

From (3.36), one can infer that the rank condition (3.34) is equivalent to the following condition:

$$\mu_k + \min(\mu_{\bar{k}}, N_k W_{\bar{k}}) \leq r_1^{[k]} + r_2^{[k]} + r_3^{[k]}, \quad k = 1, 2. \tag{3.38}$$

To show that $\mathcal{D}_{\mathrm{IC,in}}^{\mathrm{d\text{-}CSI}}$ is indeed achievable by our transmission scheme, we need to prove that all its corner points are achievable. In order to show that a given corner point (d_1^*, d_2^*) is

[§]When $W_1 \geq W_2$, $\mathbf{P}^{[1]}$ and $\mathbf{P}^{[2]}$ are again given by (3.33) but with different sizes of the constituting sub-matrices.

achievable, it suffices to find positive integers W^*, W_1^*, W_2^*, μ_1^*, and μ_2^* which simultaneously satisfy $d_1^* = \frac{\mu_1^*}{W^*}$, $d_2^* = \frac{\mu_2^*}{W^*}$, and the rank conditions (3.38). In Table 3.3, for each nontrivial corner point in class $\mathcal{C}_3 - \mathcal{C}_{10}$, the appropriate values of W^*, W_1^*, W_2^*, μ_1^*, and μ_2^* which satisfy $d_1^* = \frac{\mu_1^*}{W^*}$ and $d_2^* = \frac{\mu_2^*}{W^*}$ are presented. Thus, the proof of achievability for each corner point boils down to verification of the rank conditions (3.38) with corresponding values of W^*, W_1^*, W_2^*, μ_1^*, and μ_2^* in Table 3.3. Also, we can further simplify the proof using the following observations:

- The achievability of T_1 for class $\mathcal{C}_3$ simply follows from the achievability of T_5 for class $\mathcal{C}_9$. The reason is that if $(M_1, M_2, N_1, N_2) \in \mathcal{C}_3$, then $(N_1, M_2', N_1, N_2) \in \mathcal{C}_9$. Therefore, if in class $\mathcal{C}_3$, TX_1 uses only N_1 out of its M_1 antennas[§] and TX_2 uses only M_2' out of its M_2 antennas, the achievability of T_1 for class $\mathcal{C}_3$ will result from achievability of T_5 for class $\mathcal{C}_9$.

- The achievability of T_3 for class $\mathcal{C}_5$ simply follows from the achievability of T_3 for class $\mathcal{C}_8$. The reason is that if $(M_1, M_2, N_1, N_2) \in \mathcal{C}_5$, then $(M_1, M_2, N_1, M_2) \in \mathcal{C}_8$. Therefore, if RX_2 in class $\mathcal{C}_5$ uses only M_2 out of its N_2 antennas[†], the achievability of T_3 for class $\mathcal{C}_5$ will result from achievability of T_3 for class $\mathcal{C}_8$.

- The achievability of T_6 for classes $\mathcal{C}_9$ and $\mathcal{C}_{10}$ follows from the achievability of T_3 for class $\mathcal{C}_8$. In fact, one can readily show that if $(M_1, M_2, N_1, N_2) \in \{\mathcal{C}_9, \mathcal{C}_{10}\}$, then $(M_1, N_1 + N_2 - M_1, N_1, N_2) \in \mathcal{C}_8$. Since T_3 is achievable for class $\mathcal{C}_8$, if TX_2 uses only $N_1 + N_2 - M_1$ out of its M_2 transmit antennas[‡], we can achieve T_3 for classes $\mathcal{C}_9$ and $\mathcal{C}_{10}$ with M_2 replaced by $N_1 + N_2 - M_1$. Moreover, it is easy to see that the corner point T_6 in $(M_1, M_2, N_1, N_2) \in \{\mathcal{C}_9, \mathcal{C}_{10}\}$ MIMO IC is exactly equal to the corner point T_3 in $(M_1, N_1 + N_2 - M_1, N_1, N_2) \in \mathcal{C}_8$.

Therefore, to complete the proof of Theorem 4, we only need to prove the following cases:

[§]Note that $M_1 \geq N_1$ for class $\mathcal{C}_3$.

[†]Note that $N_2 > M_2$ for class $\mathcal{C}_5$.

[‡]Recall that for class $\mathcal{C}_9$ and $\mathcal{C}_{10}$, we have: $\Delta < M_1 \leq N_1 \leq N_2 < N_1 + N_2 - M_1 < M_2$.

Table 3.3: Appropriate parameters for our transmission scheme for each corner point in class $\mathcal{C}_3 - \mathcal{C}_{10}$

Point	Class	W^*	W_1^*	W_2^*	μ_1^*	μ_2^*
T_1	$\mathcal{C}_3$	$M_2' - N_1$	W^*	$N_2 - N_1$	$N_1(M_2' - N_2)$	$M_2'W_2^*$
T_2	$\mathcal{C}_4$	$M_1'M_2' - N_1N_2$	$N_1(M_2' - N_2)$	$N_2(M_1' - N_1)$	$M_1'W_1^*$	$M_2'W_2^*$
T_3	$\mathcal{C}_5$, $\mathcal{C}_8$	N_1	W^*	$N_1 - M_1$	$M_1W_1^*$	$M_2W_2^*$
T_4	$\mathcal{C}_6$, $\mathcal{C}_7$	$\min(M_2, L)$	W^*	$N_2 - M_1$	$M_1W_1^*$	$\min(M_2, L)W_2^*$
T_5	$\mathcal{C}_8$, $\mathcal{C}_9$	$M_2 - N_1$	W^*	$N_2 - N_1$	$N_1(M_2 - N_2)$	$M_2W_2^*$
T_6	$\mathcal{C}_9$, $\mathcal{C}_{10}$	N_1	W^*	$N_1 - M_1$	$M_1W_1^*$	AW_2^*
T_7	$\mathcal{C}_9$	$N_1 + L - M_2$	W^*	$N_1 - M_1$	$N_1(N_1 + N_2 - M_2)$	$M_2W_2^*$
T_8	$\mathcal{C}_{10}$	L	W^*	$N_2 - M_1$	N_1^2	$N_2L - N_1^2$

- The achievability of T_2 for class $\mathcal{C}_4$ which is explained in the following.
- The achievability of T_3 for class $\mathcal{C}_8$ (see Appendix B.2)
- The achievability of T_4 for classes $\mathcal{C}_6$ and $\mathcal{C}_7$ (see Appendix B.3)
- The achievability of T_5 for classes $\mathcal{C}_8$ and $\mathcal{C}_9$ (see Appendix B.4)
- The achievability of T_7 for class $\mathcal{C}_9$ (see Appendix B.5)
- The achievability of T_8 for class $\mathcal{C}_{10}$ (see Appendix B.6)

As a showcase, we prove the achievability of corner point T_2 of class $\mathcal{C}_4$ in the following. According to Table 3.3, we only need to show that $W^* = M_1'M_2' - N_1N_2$, $W_1^* = N_1(M_2' - N_2)$, $W_2^* = N_2(M_1' - N_1)$, $\mu_1^* = M_1'W_1^*$, and $\mu_2^* = M_2'W_2^*$ satisfy the rank conditions in (3.38).

From (3.37), for $k = 1, 2$, we have

$$
\begin{aligned}
r_1^{[k]} &\overset{\text{(a)}}{=} N_k W_\ell^* \,, \\
r_2^{[k]} &\overset{\text{(b)}}{=} N_k (W_u^* - W_\ell^*), \\
r_3^{[k]} &\overset{\text{(b)}}{=} \min\left\{ N_k (W^* - W_u^*), \min\left(M_k (W^* - W_u^*), N_{\bar{k}} W_k^* \right) + \min\left(M'_{\bar{k}} (W^* - W_u^*), N_k W_{\bar{k}}^* \right) \right\} \\
&\overset{\text{(c)}}{=} N_k (W^* - W_u^*),
\end{aligned}
$$

where (a) and (b) are true since $\min(M'_1, M'_2) > N_2 > N_1$ in class $\mathcal{C}_4$. To prove (c), we need to prove $N_1 W_2^* + N_2 W_1^* > N_2 (W^* - W_u^*)$. In fact, we prove the following stronger inequality:

$$N_1 W_2^* + N_2 W_1^* \geq N_2 (W^* - W_2^*). \tag{3.39}$$

To do so, we note that $W^* = W_1^* + \frac{M'_2}{N_2} W_2^*$, and therefore, (3.39) is reduced to $M'_2 \leq N_1 + N_2$ which is obviously true due to (3.7).

From the above expressions, it simply follows that $r_1^{[k]} + r_2^{[k]} + r_3^{[k]} = N_k W^*$, $k = 1, 2$. On the other hand, one can easily check that

$$
\begin{aligned}
\mu_1^* + \min(\mu_2^*, N_1 W_2^*) &= M'_1 W_1^* + N_1 W_2^* = N_1 W^* \\
\mu_2^* + \min(\mu_1^*, N_2 W_1^*) &= M'_2 W_2^* + N_2 W_1^* = N_2 W^*,
\end{aligned}
$$

and therefore, the rank conditions (3.38) are met with equality and the proof is complete.

3.4 Proof of Theorem 5

In this section, we prove the tightness of our achievable DoF region for the antenna configurations stated in Theorem 5. The DoF region of the two-user MIMO IC with full CSIT is obviously an outer-bound for the DoF region of the same channel with delayed CSIT. Also, by allowing transmitters to cooperate, the two-user MIMO IC is converted into the two-user MIMO broadcast channel (BC) whose DoF region with delayed CSIT has been characterized [52], [33]. Since cooperation does not shrink the capacity region, the DoF region of the resulting MIMO BC is also an outer-bound for DoF region of the original MIMO IC. The above arguments are summarized in the following lemma:

Lemma 2.

$$\mathcal{D}_{IC,out}^{d\text{-}CSI} = \mathcal{D}_{IC}^{f\text{-}CSI} \bigcap \mathcal{D}_{BC}^{d\text{-}CSI} \tag{3.40}$$

where $\mathcal{D}_{BC}^{d\text{-}CSI}$ is the union of all $(d_1, d_2) \in \mathbb{R}_+^2$ which satisfy the following two inequalities:

$$\begin{aligned} \frac{d_1}{\min(M_1+M_2, N_1+N_2)} + \frac{d_2}{\min(M_1+M_2, N_2)} &\leq 1 \\ \frac{d_1}{\min(M_1+M_2, N_1)} + \frac{d_2}{\min(M_1+M_2, N_1+N_2)} &\leq 1. \end{aligned} \tag{3.41}$$

In Fig. 3.4, for each of the classes $\mathcal{C}_1 - \mathcal{C}_{10}$, the regions $\mathcal{D}_{\text{IC}}^{\text{f-CSI}}$ and $\mathcal{D}_{\text{BC}}^{\text{d-CSI}}$ are depicted together with the achievable region $\mathcal{D}_{\text{IC,in}}^{\text{d-CSI}}$. The following conclusions can be inferred from this figure:

- For class $\mathcal{C}_1$, the DoF regions with full CSIT and no CSIT coincide and are equal to $\mathcal{D}_{\text{IC,in}}^{\text{d-CSI}}$. Therefore, any kind of CSI at transmitters provides no benefit in terms of DoF for this class. This corresponds to case a) in Theorem 5.

- For class $\mathcal{C}_3$, the region $\mathcal{D}_{\text{IC,out}}^{\text{d-CSI}}$ is described by $\frac{d_1}{N_1} + \frac{d_2}{N_1+N_2} \leq 1$ and $d_1 + d_2 \leq N_2$. On the other hand, the region $\mathcal{D}_{\text{IC,in}}^{\text{d-CSI}}$ is described by $d_1 + d_2 \leq N_2$ and $\frac{d_1}{N_1} + \frac{d_2}{M_2'} \leq 1$ for this class. Therefore, the achievable DoF region is tight provided that $M_2 \geq N_1 + N_2$. This corresponds to case b) in Theorem 5.

- For class $\mathcal{C}_4$, $\mathcal{D}_{\text{IC,out}}^{\text{d-CSI}} = \mathcal{D}_{\text{BC}}^{\text{d-CSI}}$ and is described by $\frac{d_1}{N_1+N_2} + \frac{d_2}{N_2} \leq 1$ and $\frac{d_1}{N_1} + \frac{d_2}{N_1+N_2} \leq 1$. On the other hand, the region $\mathcal{D}_{\text{IC,in}}^{\text{d-CSI}}$ is described by $\frac{d_1}{M_1'} + \frac{d_2}{N_2} \leq 1$ and $\frac{d_1}{N_1} + \frac{d_2}{M_2'} \leq 1$. Therefore, the achievable DoF region is tight when $\min(M_1, M_2) \geq N_1 + N_2$. This corresponds to case c) in Theorem 5.

- For classes $\mathcal{C}_6$ and $\mathcal{C}_7$, $\mathcal{D}_{\text{IC,in}}^{\text{d-CSI}} = \mathcal{D}_{\text{IC,out}}^{\text{d-CSI}} = \mathcal{D}_{\text{IC}}^{\text{f-CSI}}$ which is described by $d_1 \leq M_1$ and $d_1 + d_2 \leq N_2$. This corresponds to cases d) and e) in Theorem 5.

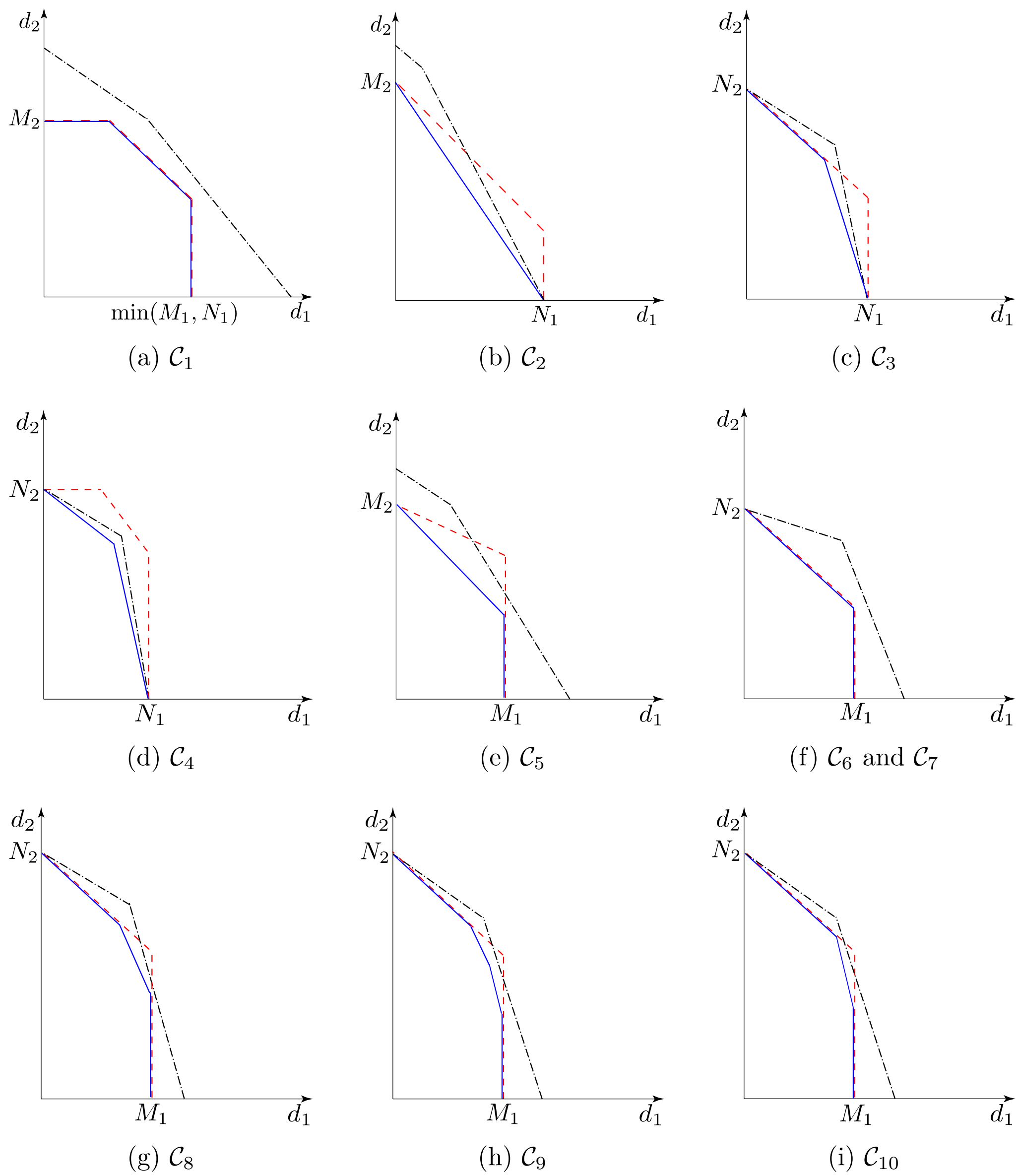


Figure 3.4: Outer-bounds on the DoF region of the two-user MIMO IC with $N_2 \geq N_1$ and delayed CSIT for different classes: BC outer-bound and the full CSIT outer-bound are respectively represented by dash-dot and dashed lines. Our achievable DoF region (solid line) is also presented for comparison.

3.5 Proof of Theorem 6

We consider each of the five cases in Theorem 6 separately:

a) $2 \leq \frac{N}{M}$

In this case, the channel DoFs with full and no CSIT coincide and are equal to $2M$. Therefore, the channel DoF with delayed CSIT is also equal to $2M$.

b) $\frac{4}{3} \leq \frac{N}{M} < 2$

Our transmission scheme consists of three phases: Phase I is dedicated to RX_1, i.e., in this phase, each transmitter sends some information symbols desired by RX_1. In Phase II, which is assigned to RX_2, each transmitter sends some information symbols for RX_2. Finally, in Phase III, each transmitter sends some redundant information to help receivers to resolve their desired symbols. In the following, we assume that a variables are desired by RX_1 and b variables are desired by RX_2. The details of our transmission scheme are as follows:

Phase I: This phase takes one channel use in which each of TX_1 and TX_2 sends M independent information symbols for RX_1. Let $a_1^{[k]}, \cdots, a_M^{[k]}$, $k \in \{1, 2\}$, denote the TX_k transmitted symbols during Phase I. By the end of this phase, RX_1 has N equations in terms of $2M$ desired unknowns. Since $2M > N$, RX_1 needs $2M - N$ extra equations to be able to resolve its desired information symbols. Now, let us look at the second receiver: RX_2 has also N equations which contain no information for RX_2 and can serve as the extra equations RX_1 needs in order to resolve its intended symbols. Each of these N equations, however, is a linear combination of information symbols of both transmitters, and therefore, can not be locally generated at one transmitter. To overcome this problem, RX_2 eliminates M variables $a_1^{[2]}, \cdots, a_M^{[2]}$ from its received equations[§] to obtain $N - M$

[§]This can be simply accomplished by solving $a_1^{[2]}, \cdots, a_M^{[2]}$ in terms of other unknowns and then substituting them in the remaining equations.

linearly independent equations which are solely in terms of $a_1^{[1]}, \cdots, a_M^{[1]}$. Let $J_1^{[2]}, \cdots, J_{N-M}^{[2]}$ denote these equations. RX_2 can also eliminate M variables $a_1^{[1]}, \cdots, a_M^{[1]}$ to obtain $N-M$ linearly independent equations which are solely in terms of $a_1^{[2]}, \cdots, a_M^{[2]}$. These equations are denoted by $V_1^{[2]}, \cdots, V_{N-M}^{[2]}$. Therefore, from the N received equations at RX_2, $2(N-M)$ equations are formed† which are solely in terms of information symbols of one transmitter. According to the delayed CSIT assumption, TX_1 has access to $J_1^{[2]}, \cdots, J_{N-M}^{[2]}$ at the end of Phase I. Similarly, TX_2 has access to $V_1^{[2]}, \cdots, V_{N-M}^{[2]}$ at the end of Phase I. Since $2(N-M) \geq 2M-N$, if we somehow deliver any $2M-N$ of $J_1^{[2]}, \cdots, J_{N-M}^{[2]}, V_1^{[2]}, \cdots, V_{N-M}^{[2]}$ to RX_1, it has enough equations to resolve its desired information symbols. This goal will be achieved in Phase III.

Phase II: This phase is similar to Phase I by exchanging the role of receivers. During this phase, each of TX_1 and TX_2 sends M information symbols for RX_2. Let $b_1^{[1]}, \cdots, b_M^{[1]}$ denote the information symbols transmitted by TX_1 and $b_1^{[2]}, \cdots, b_M^{[2]}$ denote the information symbols transmitted by TX_2. RX_1 can eliminate the variables $b_1^{[1]}, \cdots, b_M^{[1]}$ from its received equations to obtain $N-M$ equations $V_1^{[1]}, \cdots, V_{N-M}^{[1]}$ which are solely in terms of $b_1^{[2]}, \cdots, b_{2M}^{[2]}$. RX_1 can also eliminate the variables $b_1^{[2]}, \cdots, b_M^{[2]}$ from its received equations to obtain $N-M$ equations $J_1^{[1]}, \cdots, J_{N-M}^{[1]}$ which are solely in terms of $b_1^{[1]}, \cdots, b_M^{[1]}$. Since $2(N-M) \geq 2M-N$, if we somehow deliver any $2M-N$ of $J_1^{[1]}, \cdots, J_{N-M}^{[1]}, V_1^{[1]}, \cdots, V_{N-M}^{[1]}$ to RX_2, it has enough equations to resolve its desired information symbols. This goal will be achieved in Phase III.

Phase III: The linear combination $J_1^{[1]}+J_1^{[2]}$ is solely in terms of information symbols of TX_1 and the channel coefficients in the first and second channel uses. Therefore, according to the delayed CSIT assumption, it is available at TX_1 at the and of Phase II. Moreover, if we deliver $J_1^{[1]}+J_1^{[2]}$ to RX_1, it can eliminate the effect of $J_1^{[1]}$ to obtain $J_1^{[2]}$ which is a useful equation in terms of information symbols desired by RX_1. Similarly, by delivering $J_1^{[1]}+J_1^{[2]}$ to RX_2, it can eliminate the effect of $J_1^{[2]}$ to obtain $J_1^{[1]}$ which is a useful equation in terms of information symbols desired by RX_2. Therefore, $J_1^{[1]} + J_1^{[2]}$ is a useful equation for both

†Note that $2(N-M) \leq N$.

receivers. Since each receiver only needs $2M-N$ extra equations to resolve its desired symbols, our goal in Phase III it to deliver the $2M-N$ linear combinations $J_1^{[1]}+J_1^{[2]}$, $\cdots$, $J_{N-M}^{[1]}+J_{N-M}^{[2]}$, $V_1^{[1]}+V_1^{[2]}$, $\cdots$, $V_{3M-2N}^{[1]}+V_{3M-2N}^{[2]}$ to both receivers. Since there are N antennas at each receiver, this can be simply accomplished in $\frac{2M-N}{N}$ channel uses[§].

Since $2(M+M)=4M$ information symbols were transmitted in $1+1+\frac{2M-N}{N}=\frac{2M+N}{N}$ channel uses, we have achieved a DoF of $\frac{4MN}{2M+N}$.

c) $1 \leq \frac{N}{M} < \frac{4}{3}$ Our transmission scheme for this case consists of three phases as follows:

Phase I: This phase takes two channel uses. At each channel use, TX_1 sends M independent information symbols for RX_1. Let $a_1^{[1]},\cdots,a_{2M}^{[1]}$ denote the TX_1 transmitted symbols during Phase I. TX_2 sends $3N-2M$ information symbols intended for RX_1 in this phase by transmitting random linear combinations of these symbols over its transmit antennas. Notice that since $1<\frac{N}{M}<\frac{4}{3}$ we have $0<3N-2M<2M$. Let $a_1^{[2]},\cdots,a_{3N-2M}^{[2]}$ denote TX_2 transmitted symbols during Phase I. By the end of Phase I, RX_2 has $2N$ equations in terms of $2M+3N-2M=3N$ unknowns. From this system of equations, RX_2 can eliminate $3N-2M$ variables $a_1^{[2]},\cdots,a_{3N-2M}^{[2]}$ to obtain $2N-(3N-2M)=2M-N$ independent equations which are solely in terms of $a_1^{[1]},\cdots,a_{2M}^{[1]}$. Let $J_1^{[2]},\cdots,J_{2M-N}^{[2]}$ denote these equations. RX_2 can also eliminate $2M$ variables $a_1^{[1]},\cdots,a_{2M}^{[1]}$ to obtain $2N-2M=2(N-M)$ independent equations which are solely in terms of $a_1^{[2]},\cdots,a_{3N-2M}^{[2]}$. Let $V_1^{[2]},\cdots,V_{2(N-M)}^{[2]}$ denote these equations. Thus, from the $2N$ received equations at RX_2 during Phase I, $2M-N+2(N-M)=N$ independent equations are formed which are solely in terms of information symbols of one transmitter. Since RX_1 needs $3N-2N=N$ extra equations to resolve its received symbols during Phase I, we should deliver all the linear combinations $J_1^{[2]},\cdots,J_{2M-N}^{[2]},V_1^{[2]},\cdots,V_{2(N-M)}^{[2]}$ to RX_1. This goal is accomplished in Phase III.

Phase II: This phase takes two channel uses. At each channel use, TX_1 sends M independent information symbols for RX_2. Let $b_1^{[1]},\cdots,b_{2M}^{[1]}$ denote the TX_1 transmitted

[§]Note that the fractional channel uses can be easily handled by sufficient repetition of Phase I and Phase II.

symbols during Phase II. TX_2 sends $3N-2M$ information symbols intended for RX_2 in this phase by transmitting random linear combinations of these symbols over its transmit antennas. Let $b_1^{[2]}, \cdots, b_{3N-2M}^{[2]}$ denote TX_2 transmitted symbols during Phase II. From its received equations during Phase II, RX_1 can eliminate $3N-2M$ variables $b_1^{[2]}, \cdots, b_{3N-2M}^{[2]}$ to obtain $2M-N$ equations which are solely in terms of $b_1^{[1]}, \cdots, b_{2M}^{[1]}$. Let $J_1^{[1]}, \cdots, J_{2M-N}^{[1]}$ denote these equations. RX_1 can also eliminate $2M$ variables $b_1^{[1]}, \cdots, b_{2M}^{[1]}$ to obtain $2(N-M)$ equations which are solely in terms of $b_1^{[2]}, \cdots, b_{3N-2M}^{[2]}$. Let $V_1^{[1]}, \cdots, V_{2(N-M)}^{[1]}$ denote these equations. RX_2 requires all the linear combinations $J_1^{[1]}, \cdots, J_{2M-N}^{[1]}, V_1^{[1]}, \cdots, V_{2(N-M)}^{[1]}$ to resolve its desired symbols. This goal is achieved in Phase III.

Phase III: The linear combination $J_1^{[1]} + J_1^{[2]}$ is solely in terms of information symbols of TX_1 and the channel coefficients in the first and second channel uses. Therefore, according to the delayed CSIT assumption, it is available at TX_1 at the and of Phase II. Moreover, if we deliver $J_1^{[1]} + J_1^{[2]}$ to both receivers, each of them will obtain a useful equation in terms of its desired information symbols. Since each receiver needs N extra equations to resolve its desired symbols, our goal in Phase III is to deliver the N linear combinations $J_1^{[1]} + J_1^{[2]}$, $\cdots$, $J_{2M-N}^{[1]} + J_{2M-N}^{[2]}$, $V_1^{[1]} + V_1^{[2]}$, $\cdots$, $V_{2(N-M)}^{[1]} + V_{2(N-M)}^{[2]}$ to both receivers. This can be simply accomplished in one channel use[§].

Since $6N$ information symbols were transmitted in $2+2+1=5$ channel uses, we have achieved a DoF of $\frac{6}{5}N$.

d) $\frac{1}{2} < \frac{N}{M} < 1$

Our transmission scheme consists of three phases as follows:

Phase I: This phase takes two channel uses. At each channel use, TX_1 sends M independent information symbols for RX_1. Let $a_1^{[1]}, \cdots, a_{2M}^{[1]}$ denote the TX_1 transmitted symbols during Phase I. TX_2 sends $2N-M$ information symbols intended for RX_1 in the first channel use and retransmits them in the second channel use. Notice that since

[§]Note that $2M-N < M$ and $2(N-M) < M$.

$\frac{1}{2} < \frac{N}{M} < 1$ we have $0 < 2N - M < M$. Let $a_1^{[2]}, \cdots, a_{2N-M}^{[2]}$ denote TX_2 transmitted symbols during Phase I. By the end of Phase I, RX_2 has $2N$ equations in terms of $2M + 2N - M = M + 2N$ unknowns. From this system of equations, RX_2 can eliminate $2N - M$ variables $a_1^{[2]}, \cdots, a_{2N-M}^{[2]}$ to obtain M equations which are solely in terms of $a_1^{[1]}, \cdots, a_{2M}^{[1]}$. Let $J_1^{[2]}, \cdots, J_M^{[2]}$ denote these equations. Since the linear combinations $J_1^{[2]}, \cdots, J_M^{[2]}$ are in terms of TX_1 symbols and according to the delayed CSIT assumption, TX_1 has access to these linear combinations at the end of Phase I. If we somehow deliver $J_1^{[2]}, \cdots, J_M^{[2]}$ to RX_1, it will have enough equations to resolve its desired information symbols. This goal will be achieved at Phase III.

Phase II: This phase is similar to Phase I by exchanging the role of receivers. By the end of this phase, TX_1 and TX_2 have respectively sent $2N - M$ and $2M$ fresh information symbols intended for RX_2. Let $b_1^{[1]}, \cdots, b_{2N-M}^{[1]}$ denote the information symbols transmitted by TX_1 and $b_1^{[2]}, \cdots, b_{2M}^{[2]}$ denote the information symbols transmitted by TX_2. Similar to Phase I, RX_1 can eliminate the variables $b_1^{[1]}, \cdots, b_{2N-M}^{[1]}$ from the $2N$ equations available at this receiver to obtain M equations in terms of $b_1^{[2]}, \cdots, b_{2M}^{[2]}$. Let $V_1^{[1]}, \cdots, V_M^{[1]}$ denote these equations. Similar to Phase I, TX_2 has access to these linear combinations at the end of Phase II. If we somehow deliver $V_1^{[1]}, \cdots, V_M^{[1]}$ to RX_2, it will have enough equations to resolve its desired information symbols. This goal will be achieved at Phase III.

Phase III: Our objective in this phase is to deliver $J_1^{[2]}, \cdots, J_M^{[2]}$ to RX_1 and to deliver $V_1^{[1]}, \cdots, V_M^{[1]}$ to RX_2. Since $V_1^{[2]}, \cdots, V_M^{[2]}$ are available at both TX_1 and RX_2 and since $J_1^{[1]}, \cdots, J_M^{[1]}$ are available at both TX_2 and RX_1, Phase III will be accomplished in $\frac{M}{N}$ channel uses†.

Since $2(2N + M)$ information symbols were transmitted in $2 + 2 + \frac{M}{N} = \frac{4N+M}{N}$ channel uses, we have achieved a DoF of $\frac{2N}{M+4N}(M + 2N)$.

e) $\frac{N}{M} \leq \frac{1}{2}$

†Note that the fractional channel uses can be easily handled by sufficient repetition of Phase I and Phase II.

Since the MIMO interference channel is contained in the MIMO X channel and the DoF of the two-user MIMO IC with $M \geq 2N$ is equal to $\frac{4}{3}N$ (see class $\mathcal{C}_4$), we can conclude that $\frac{4}{3}N$ is also an achievable DoF for the two-user MIMO X channel.

3.6 Proof of Theorem 7

In this section, we prove Theorem 7 for the case of $K = 3$. The proof for the general case immediately follows from the solution to the 3-user network .

We wish to prove that $\frac{5}{4}$ DoF are achievable on the 3-user X network with delayed CSIT. To this aim, we consider 12 channel uses over which a total of 15 independent information symbols will be transmitted. More precisely, during 12 channel uses, each transmitter can send two independent information symbols to one receiver and a single information symbol to the other two receivers. In the following, we are assumed that u variables are intended for receiver one, v variables are intended for receiver two, and w variables are intended for receiver three. Our transmission scheme consists of four phases:

Phase 1 consists of the three channel uses and is dedicated to receiver one. In each channel use of this phase, transmitter one sends a new information symbol for receiver one. Transmitter two sends an information symbol for receiver one in its first channel use and retransmits it in the next two channel uses of this phase. Transmitter 3 acts like transmitter two. Let $u_1^{[1]}, u_2^{[1]}, u_3^{[1]}$ be the information symbols sent by transmitter one and $u_1^{[2]}$ and $u_1^{[3]}$ be the information symbols sent by transmitter two and three, respectively. The following signals are received at receiver k, $k = 1, 2, 3$:

$$\begin{aligned} y^{[k]}(1) &= h^{[k1]}(1)u_1^{[1]} + h^{[k2]}(1)u_1^{[2]} + h^{[k3]}(1)u_1^{[3]} \\ y^{[k]}(2) &= h^{[k1]}(2)u_2^{[1]} + h^{[k2]}(2)u_1^{[2]} + h^{[k3]}(2)u_1^{[3]} \\ y^{[k]}(3) &= h^{[k1]}(3)u_3^{[1]} + h^{[k2]}(3)u_1^{[2]} + h^{[k3]}(3)u_1^{[3]} \end{aligned} \tag{3.42}$$

From the above system of linear equations for $k = 2$, receiver two can eliminate the variables $u_1^{[2]}$ and $u_1^{[3]}$ to obtain a linear combination of $u_1^{[1]}, u_2^{[1]}, u_3^{[1]}$. Let $L^{[21]}$ denote this linear combination. Receiver three can do the same thing to obtain a different linear combination of $u_1^{[1]}, u_2^{[1]}, u_3^{[1]}$. Let $L^{[31]}$ denote this linear combination. Note that $L^{[21]}$ and $L^{[31]}$ are almost surely linearly independent.

Phase 2 consists of the three channel uses and is dedicated to receiver two. In each channel use of this phase, transmitter two sends a new information symbol for receiver two. Transmitter one sends an information symbol for receiver two in its first channel use and retransmits it in the next two channel uses of this phase. Transmitter 3 acts like transmitter one. Let $v_1^{[2]}, v_2^{[2]}, v_3^{[2]}$ be the information symbols sent by transmitter two and $v_1^{[1]}$ and $v_1^{[3]}$ be the information symbols sent by transmitter one and three, respectively. The following signals are received at receiver k, $k = 1, 2, 3$:

$$\begin{aligned}
y^{[k]}(4) &= h^{[k1]}(4)v_1^{[1]} + h^{[k2]}(4)v_1^{[2]} + h^{[k3]}(4)v_1^{[3]} \\
y^{[k]}(5) &= h^{[k1]}(5)v_1^{[1]} + h^{[k2]}(5)v_2^{[2]} + h^{[k3]}(5)v_1^{[3]} \\
y^{[k]}(6) &= h^{[k1]}(6)v_1^{[1]} + h^{[k2]}(6)v_3^{[2]} + h^{[k3]}(6)v_1^{[3]}
\end{aligned} \tag{3.43}$$

From the above system of linear equations for $k = 1$, receiver one can eliminate the variables $v_1^{[1]}$ and $v_1^{[3]}$ to obtain a linear combination of $v_1^{[2]}, v_2^{[2]}, v_3^{[2]}$. Let $L^{[12]}$ denote this linear combination. Receiver three can do the same thing to obtain a different linear combination of $v_1^{[2]}, v_2^{[2]}, v_3^{[2]}$. Let $L^{[32]}$ denote this linear combination. Note that $L^{[12]}$ and $L^{[32]}$ are almost surely linearly independent.

Phase 3 consists of the three channel uses and is dedicated to receiver three. In each channel use of this phase, transmitter three sends a new information symbol for receiver three. Transmitter one sends an information symbol for receiver three in its first channel use and retransmits it in the next two channel uses of this phase. Transmitter 2 acts like transmitter one. Let $w_1^{[3]}, w_2^{[3]}, w_3^{[3]}$ be the information symbols sent by transmitter 3 and $w_1^{[1]}$ and $w_1^{[2]}$ be the information symbols sent by transmitter one and two, respectively.

The following signals are received at receiver k, $k = 1, 2, 3$:

$$\begin{aligned} y^{[k]}(7) &= h^{[k1]}(7)w_1^{[1]} + h^{[k2]}(7)w_1^{[2]} + h^{[k3]}(7)w_1^{[3]} \\ y^{[k]}(8) &= h^{[k1]}(8)w_1^{[1]} + h^{[k2]}(8)w_1^{[2]} + h^{[k3]}(8)w_2^{[3]} \\ y^{[k]}(9) &= h^{[k1]}(9)w_1^{[1]} + h^{[k2]}(9)w_1^{[2]} + h^{[k3]}(9)w_3^{[3]} \end{aligned} \tag{3.44}$$

From the above system of linear equations for $k = 1$, receiver one can eliminate the variables $w_1^{[1]}$ and $w_1^{[2]}$ to obtain a linear combination of $w_1^{[3]}, w_2^{[3]}, w_3^{[3]}$. Let $L^{[13]}$ denote this linear combination. Receiver two can do the same thing to obtain a different linear combination of $w_1^{[3]}, w_2^{[3]}, w_3^{[3]}$. Let $L^{[23]}$ denote this linear combination. Note that $L^{[13]}$ and $L^{[23]}$ are almost surely linearly independent.

Phase 4 consists of the three channel uses. No new information symbol is transmitted during this phase. In the first channel use of this phase, transmitter one sends $L^{[21]}$, transmitter two sends $L^{[12]}$, and transmitter three sends nothing. In the second channel use of this phase, transmitter one sends $L^{[31]}$, transmitter two sends nothing, and transmitter three sends $L^{[13]}$. Finally, in the last channel use of this phase, transmitter one sends nothing, transmitter two sends $L^{[32]}$, and transmitter three sends $L^{[23]}$.

Now, we argue that each receiver has enough number of equations to resolve its intended information symbols. First, we note that by the end of phase 3, receiver one knows $L^{[12]}$ and $L^{[13]}$, receiver two knows $L^{[21]}$ and $L^{[23]}$, and receiver three knows $L^{[31]}$ and $L^{[32]}$. We now consider each receiver separately:

Receiver 1: By the knowledge of $L^{[12]}$ and $L^{[13]}$, this receiver can resolve $L^{[21]}$ and $L^{[31]}$ during the phase 4. Moreover, $y^{[1]}(1), y^{[1]}(2), y^{[1]}(3), L^{[21]}$, and $L^{[31]}$ form a system of linear equations in the five variables $u_1^{[1]}, u_2^{[1]}, u_3^{[1]}, u_1^{[2]}$, and $u_1^{[3]}$. One can readily check that these equations are almost surely linearly independent and hence receiver one can extract its desired symbols by solving this system of linear equations.

Receiver 2: By the knowledge of $L^{[21]}$ and $L^{[23]}$, this receiver can resolve $L^{[12]}$ and $L^{[32]}$ during the phase 4. Moreover, $y^{[2]}(4), y^{[2]}(5), y^{[2]}(6), L^{[12]}$, and $L^{[32]}$ form a system of linear equations in the five variables $v_1^{[1]}, v_1^{[2]}, v_2^{[2]}, v_3^{[2]}$, and $v_1^{[3]}$. One can readily check that these equations are almost surely linearly independent and hence receiver two can extract its desired symbols by solving this system of linear equations.

Receiver 3: By the knowledge of $L^{[31]}$ and $L^{[32]}$, this receiver can resolve $L^{[13]}$ and $L^{[23]}$ during the phase 4. Moreover, $y^{[3]}(7), y^{[3]}(8), y^{[3]}(9), L^{[13]}$, and $L^{[23]}$ form a system of linear equations in the five variables $w_1^{[1]}, w_1^{[2]}, w_1^{[3]}, w_2^{[3]}$, and $w_3^{[3]}$. One can readily check that these equations are almost surely linearly independent and hence receiver one can extract its desired symbols by solving this system of linear equations.

3.7 Proof of Theorem 8

In this section, we prove that a sum-DoF of $\frac{2K}{K+1}$ is achievable for the MISO IC with $M \geq K$ antennas at each transmitter and with delayed CSIT. Our transmission scheme consists of two phases:

Phase I: This phase takes K channel uses. Only TX_k is active at channel use k, $1 \leq k \leq K$. Let us consider channel use k: At this channel use, TX_k sends K independent information symbols over its first K antennas[§]. Each receiver then observes a linear combination of these symbols. RX_k can decode its desired information symbols if it is provided with all of the linear combinations observed in other receivers. This is the objective of Phase II. Note that by the end of Phase I, RX_k has received a linear combination in terms of information symbols desired by RX_1, a linear combination in terms of information symbols desired by RX_2, $\cdots$, and a linear combination in terms of information symbols desired by RX_K.

Phase II: This phase takes $\binom{K}{2}$ channel uses. No new information symbol is transmitted during this phase. At each channel use, only two transmitters are active and all other transmitters are silent. We first select two users. There are $\binom{K}{2} = K(K-1)/2$ possibilities for such a selection. By transmitting their respective linear combinations received in Phase I, these two users can exchange the linear combinations they know about each other in one channel use. After $K(K-1)/2$ channel uses, each receiver has enough linear combinations to decode its own information symbols.

[§]The remaining $M - K$ antennas are inactive.

Since K^2 information symbols have been transmitted in $K+K(K-1)/2$ channel uses, we achieve a sum-DoF of $\frac{K^2}{K+K(K-1)/2} = \frac{2K}{K+1}$ and the proof is complete.

3.8 Conclusion

We obtained new results on the DoF region of the two-user MIMO Gaussian IC with delayed CSIT. The tightness of our achievable scheme was proved for some antenna configurations. The two-user MIMO X channel, K-user X network, and K-user MISO IC were also considered wherein new achievable DoF results were obtained and shown to be tight in some cases.

Interference alignment was the main ingredient of our transmission schemes to obtain DoF improvements over the no CSIT case. It was realized by a multi-phase transmission in which each transmitter uses its knowledge of past CSI to regulate its subsequent transmissions such that the interference subspace at each receiver is not expanded. Even though we showed DoF improvement over the no CSIT case, the problem of DoF characterization for the two-user MIMO X channel and K-user MISO IC with delayed CSIT remains open due to lack of tight upper-bounds.

Chapter 4

Conclusion

In this dissertation, we studied interference alignment as a powerful technique for interference mitigation in wireless networks. Our main focus was on MIMO wireless networks with distributed transmitters and receivers, namely IC and X channel. To make progress, channel DoF, which provides a first order approximation of the channel capacity, was used as the figure of merit. The slightest improvement in the channel DoF is translated to an unbounded gap in the channel capacity in high-SNR regime. Therefore, DoF investigation of channels whose capacity region are unknown have a profound impact on our understanding of the behavior of these channels in practical ranges of SNR.

In Chapter 2, we studied the K-user MIMO IC with constant channel coefficients. Constant channels with distributed transmitters and receivers (like IC and X channel) do not lend themselves to solutions based on the signal space interference alignment and in most cases there is a gap between the lower and upper bound results obtained using this approach. We started by extending a recently introduced interference alignment technique, known as real alignment, to the more general case of constant MIMO ICs. Using this method, we obtained a new achievable DoF for the K-user constant MIMO Gaussian interference channel. Our achievable DoF result outperforms the existing results which are based on the signal space interference alignment. To evaluate our achievability scheme, we developed a new upper-bound on the DoF of the K-user MIMO Gaussian interference channel. Our upper-bound, which is valid for both constant and time-varying channels, is the tightest known bound for the K-user MIMO IC. By comparing this new upper-bound with our achievable DoF, the optimality (in the sense of DoF) of real alignment technique was established when the number of users is larger than a certain threshold. In specific, it was shown than the K-user constant MIMO IC with M antennas ta each transmitter and N antennas ta each receiver is equal to $K\frac{MN}{M+N}$ when $K \geq \frac{M+N}{\gcd(M,N)}$. Our results suggest that, from a degrees of freedom point of view, the advantage of the joint processing in co-located antennas at the transmitters and receivers vanishes as the number of users exceed a certain threshold. This is in sharp contrast to the signal space interference alignment which relies on the MIMO benefits.

In Chapter 3, we considered the possibility of interference alignment for the fast fading channels with partial CSI at the transmitters. We considered the delayed CSIT model in which transmitters have access to the CSI after a finite delay which is greater than the channel coherence time. We considered several channels including the two-user MIMO IC, the two-user MIMO X channel, the K-user X network, and the K-user MISO IC. For each of these channels, we proposed new transmission schemes under the delayed CSIT assumption which provide DoF advantage compared to the no CSIT case. The main ingredient of our transmission schemes was interference alignment. It was realized by a multi-phase transmission in which each transmitter uses its knowledge of past CSI to regulate its subsequent transmissions such that the interference subspace at each receiver is not expanded. We first considered the two-user MIMO IC where an achievable DoF region for this channel was obtained. We also developed a simple outer-bound on the DoF region of this channel which meets our achievable DoF region in some cases, and thus, characterized the channel DoF region with delayed CSIT for certain classes of antenna configurations. We then studied the two-user MIMO X channel under the delayed CSIT assumption. For the two-user MIMO X channel with M antennas at each transmitter and N antennas at each receiver, the channel DoF was characterized for all values of

M and N except possibly for $1/2 < N/M < 4/3$. It was proved that the DoF of this channel coincides with the DoF of a two-user MIMO BC which is obtained by allowing full cooperation between transmitters for $N/M > 4/3$. The K-user SISO X network was considered next, and shown to achieve $\frac{4}{3} - \frac{2}{3(3K-1)}$ DoF. Finally, the K-user MISO IC was investigated under the delayed CSIT assumption and it was proved that this channel can achieve $\frac{2K}{K+1}$ DoF.

4.1 Future research directions

This dissertation can be followed in different directions, some of which are highlighted as follows:

4.1.1 Combining Real Alignment with Signal Space Alignment

Although the proposed achievable scheme for the K-user MIMO IC in Chapter 2 is optimum* when the number of users is above a certain threshold, it is not generally optimum for the small number of users. This is because we do not take advantage of the potential cooperation among the transmit and receive antennas of each user. In fact, the extension of the real interference alignment performs optimally for $K \geq K_u$ according to our results. On the other hand, we know that the DoF upper-bound for $K \leq K_l$ is achieved using the advantages of MIMO processing. Therefore, it seems that for $K_l < K < K_u$, a combination of real alignment and signal space alignment is required to approach the the channel DoF. To develop such a hybrid approach, it is required to extend the Khintchine-Groshev theorem to the case of several linear forms with dependent integer coefficients. To our best knowledge, there is no such extension in literature. The case of $K = 3$ users has been extensively investigated and to a great extent solved recently†. However, in the studying of

*in the sense of DoF

†see [44] and references therein

K-user IC, the 3-user case is somehow special in the sense that many of the transmission schemes proposed for this channel are not easily extended to the case of $K \geq 4$ users. Therefore, it seems that a separate treatment for $K \geq 4$ users is required.

4.1.2 Real Interference Alignment with finite precision

Real interference alignment is known to be a powerful technique to establish asymptotic results like DoF characterization. However, it is not clear whether this technique can predict the channel capacity in finite SNR regime. In fact, as it was mentioned in Chapter 2, the DoF of the K-user constant IC is a discontinuous function of channel coefficients and is very sensitive to the rationality/irrationality of channel coefficients. In specific, one might argue that the irrationally of the channel coefficients is fundamental in real interference alignment and hence the scheme might not work at the presence of unavoidable quantization errors. Very recently, it was shown in [53] that the real interference alignment can be used to obtain constant gap capacity results for the two-user X channel. This important study proved that, at least for the two-user X channel, the everywhere discontinuity of the DoF in the channel coefficients is indeed a consequence of the definition of DoF as a limiting expression and not fundamental to the real interference alignment. An interesting future direction is to combine our extended version of real interference alignment with the method developed in [53] to obtain constant gap capacity characterization for the K-user MIMO IC.

4.1.3 Developing an upper-bound for the networks with distributed transmitters and Delayed CSIT

Most of our results in Chapter 3 are achievable DoF results. Even though cooperative outer-bounds were shown to be tight in some cases, they are not generally sufficient to characterize the DoF of channels like two-user X channel. It is worth mentioning that the only known upper bounds in the literature on the DoF of channels with delayed CSIT are for the K-user MISO BC, and the two-user MIMO IC. While those upper bounds proved to be tight for their corresponding channels, their extension to other channels seems to be not a straightforward task. An important future direction of this work is to develop new upper-bounds on the DoF of channels studied here.

4.1.4 Characterizing the trade-off between the DoF gain and feedback overhead

The study of communication channels under the delayed CSIT model reveals that the channel DoF are not entirely lost even with the completely outdated CSI at the transmitters. In other words, the CSI obtained through feedback links with some finite delay can be efficiently exploited to attain capacity gains. On the other hand, the overhead of providing delayed CSI to the transmitters in the delay CSIT model is substantial and may overwhelm the capacity gains. An interesting future direction is to characterize the trade off between the capacity gain and the feedback overhead. This is especially useful in fading channels with correlated fading across several channel uses.

APPENDICES

Appendix A

Appendices for Chapter 2

A.1 Proof of (2.19)

In this appendix, we prove that

$$J(W_1M, W_2M, W_1N, W_2N) \leq \max\{\max(M,N)W_{\min}, \min(M,N)W_{\max}\}, \tag{A.1}$$

where $W_{\min} = \min(W_1, W_2)$ and $W_{\max} = \max(W_1, W_2)$. First, note that

$$\begin{aligned} J(W_1M, W_2M, W_1N, W_2N) &= \min\{WM, WN, \max(W_1M, W_2N), \max(W_2M, W_1N)\} \\ &\leq \min\{\max(W_1M, W_2N), \max(W_2M, W_1N)\}. \end{aligned} \tag{A.2}$$

Due to the symmetry, without loss of generality, we prove (A.1) for the case of $M \geq N$. We consider two cases:

1. $W_1 \geq W_2$
   In this case, $\max(W_1M, W_2N) = W_1M$. To evaluate (A.2), we differentiate between two cases:

- $W_1 N \geq W_2 M$
  In this case, $\max(W_2 M, W_1 N) = W_1 N$. Therefore, (A.2) reduces to:
$$\begin{aligned} J(W_1 M, W_2 M, W_1 N, W_2 N) &\leq \min\{W_1 M, W_1 N\} = W_1 N \\ &= \max\{W_1 N, W_2 M\} = \max\{W_{\max} N, W_{\min} M\}. \end{aligned}$$

- $W_1 N < W_2 M$
  In this case $\max(W_2 M, W_1 N) = W_2 M$. Therefore, (A.2) reduces to:
$$\begin{aligned} J(W_1 M, W_2 M, W_1 N, W_2 N) &\leq \min\{W_1 M, W_2 M\} = W_2 M \\ &= \max\{W_1 N, W_2 M\} = \max\{W_{\max} N, W_{\min} M\}. \end{aligned}$$

2. $W_1 < W_2$
   In this case, $\max(W_2 M, W_1 N) = W_2 M$. To evaluate (A.2), we again differentiate between two cases:

- $W_1 M \geq W_2 N$
  In this case, $\max(W_1 M, W_2 N) = W_1 M$. Therefore, (A.2) reduces to:
$$\begin{aligned} J(W_1 M, W_2 M, W_1 N, W_2 N) &\leq \min\{W_1 M, W_2 M\} = W_1 M \\ &= \max\{W_1 M, W_2 N\} = \max\{W_{\max} N, W_{\min} M\}. \end{aligned}$$

- $W_1 M < W_2 N$
  In this case, $\max(W_1 M, W_2 N) = W_2 N$. Therefore, (A.2) reduces to:
$$\begin{aligned} J(W_1 M, W_2 M, W_1 N, W_2 N) &\leq \min\{W_2 N, W_2 M\} = W_2 N \\ &= \max\{W_1 M, W_2 N\} = \max\{W_{\max} N, W_{\min} M\}. \end{aligned}$$

This completes the proof.

A.2 The closest rational neighbors of a real number with denominator at most K

In this appendix, we study how closely a real number can be approximated by rational numbers that have a given bound on the size of their denominators. Specifically, for a real number α and a positive integer K, we are looking for two rational numbers α^- and α^+ such that $\alpha^- \leq \alpha \leq \alpha^+$ and moreover α^- and α^+ are closer to α than any other rational number with denominator at most K. Given α and K, there is an elegant method to find the rationals α^- and α^+ using the so called Farey sequence [54]. A Farey sequence of order N consists of all irreducible fractions from $[0, 1]$ with denominator not exceeding N, arranged in order of increasing magnitude. The Farey sequence of order N will be denoted by $\mathcal{F}_N$. For example $\mathcal{F}_5 = \{\frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1}\}$. Farey sequences of any order can be obtained using Stern-Brocot tree [54]. As it is depicted in Fig. A.1 , at the first layer of this tree we have two fractions $\frac{0}{1}$ and $\frac{1}{1}$. Layer i is obtained from layer $i-1$ by keeping all the fractions from layer $i-1$ and by inserting $\frac{m+m'}{n+n'}$ between two adjacent fractions $\frac{m}{n}$ and $\frac{m'}{n'}$ in layer $i-1$ whenever it is possible to do so without getting a denominator that is exceeding i. Using this procedure, fractions at the i-th layer of the tree constitute the Farey sequence of order i.

Suppose that $\alpha \in [0, 1)$ is a given real number, and the goal is to calculate the closest rational neighbors of α with denominator not exceeding a given positive integer K. To do this, we need to find the place of α in the sequence $\mathcal{F}_K$. If $\alpha \in \mathcal{F}_k$, then $\alpha^- = \alpha^+ = \alpha$. If $\alpha \notin \mathcal{F}_k$, then we can find its closest rationals α^- and α^+ by:

$$\alpha^- = \max_{\substack{q \in \mathcal{F}_K \\ q < \alpha}} q, \qquad \alpha^+ = \min_{\substack{q \in \mathcal{F}_K \\ q > \alpha}} q. \tag{A.3}$$

For example, the closest rational neighbors of $\alpha = \sqrt{2} - 1$ with denominator not exceeding 5 are $\alpha^- = \frac{2}{5}$ and $\alpha^+ = \frac{1}{2}$. In this method, for a given K, we first need to construct the sequence $\mathcal{F}_K$ and then solve the optimization problem in (A.3). Lemma 1 provides an

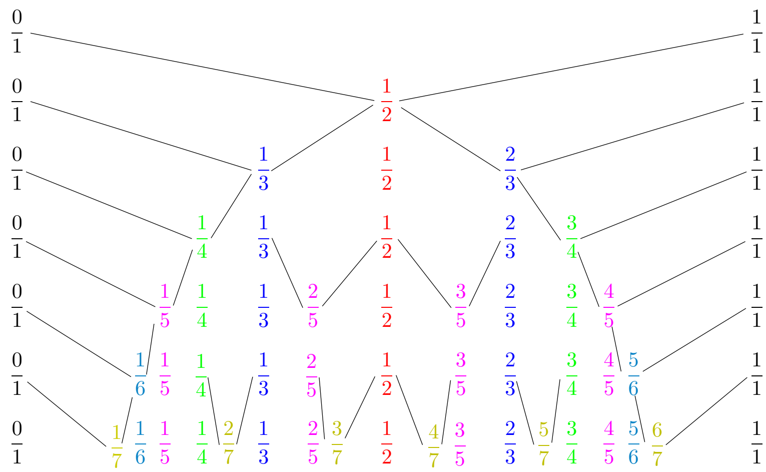


Figure A.1: Constructing a Farey sequence using the Stern-Brocot tree. The fractions in layer k form the Farey sequence $\mathcal{F}_k$.

alternative approach to find the closest rational neighbors of a given real number α with denominator at most K without the help of Farey sequence.

Proof of Lemma 1. To prove (2.25), let us assume that $\max_{n\in\{1,\cdots,K\}} \frac{\lfloor n\alpha \rfloor}{n} = \frac{\lfloor n_0\alpha \rfloor}{n_0}$ for some $n_0 \in \{1,\cdots,K\}$. Note that $\frac{\lfloor n_0\alpha \rfloor}{n_0} \leq \alpha$ and $\frac{\lfloor n_0\alpha \rfloor}{n_0} \in \mathcal{F}_K$. We claim that among all fractions in $\mathcal{F}_K$ that are less than α, the fraction $\frac{\lfloor n_0\alpha \rfloor}{n_0}$ is the closest to α. We prove our claim by contradiction. Assume we can find a fraction $\frac{p}{q}$, $(p,q) = 1$ such that $\frac{p}{q} \in \mathcal{F}_K$ and $\frac{\lfloor n_0\alpha \rfloor}{n_0} < \frac{p}{q} \leq \alpha$. It then follows that:

$$p \leq q\alpha. \tag{A.4}$$

On the other hand, since $q \leq K$, it follows that $\frac{\lfloor q\alpha \rfloor}{q} \leq \frac{\lfloor n_0\alpha \rfloor}{n_0}$ and since $\frac{\lfloor n_0\alpha \rfloor}{n_0} < \frac{p}{q}$ it follows that

$$p > \lfloor q\alpha \rfloor. \tag{A.5}$$

Combining (A.4) and (A.5), we have $\lfloor q\alpha \rfloor < p \leq q\alpha$ which is a contradiction because p is an integer. We can prove (2.26) by a similar argument. $\square$

A.3 Calculating the normalizing constant γ in (2.30)

The average transmit power of user k can be calculated as follows:

$$\sum_{m=1}^{M}\mathbb{E}\left[(X_m^{[k]})^2\right]=\sum_{m=1}^{M}\sum_{n=1}^{N}(h_{nm}^{[kk]})^2\mathbb{E}\left[(X_{mn}^{[k]})^2\right]=\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{\ell=1}^{L}(h_{nm}^{[kk]})^2\mathbb{E}\left[(X_{mn\ell}^{[k]})^2\right]. \tag{A.6}$$

On the other hand, since $X_{mn\ell}^{[k]}$ is uniformly distributed over $\Lambda_{mn\ell}^{[k]}(P,\epsilon)$, it follows that

$$\mathbb{E}\left[(X_{mn\ell}^{[k]})^2\right]=\frac{1}{|\Lambda_{mn\ell}^{[k]}(P,\epsilon)|}\sum_{x\in\Lambda_{mn\ell}^{[k]}(P,\epsilon)}x^2, \tag{A.7}$$

where $|\Lambda_{mn\ell}^{[k]}(P,\epsilon)|$ denotes the size of the set $\Lambda_{mn\ell}^{[k]}(P,\epsilon)$ which is equal to $2Q+1$. Therefore,

$$\mathbb{E}\left[(X_{mn\ell}^{[k]})^2\right]=\frac{\gamma^2P^{\frac{\nu-2+4\epsilon}{\nu+2\epsilon}}\left(\omega_{mn\ell}^{[k]}\right)^2}{2Q+1}\sum_{q=-Q}^{Q}q^2=\gamma^2P^{\frac{\nu-2+4\epsilon}{\nu+2\epsilon}}\left(\omega_{mn\ell}^{[k]}\right)^2\frac{Q(Q+1)}{3}. \tag{A.8}$$

Substituting (A.8) in (A.6) and noting that $Q(Q+1)\approx P^{\frac{2(1-\epsilon)}{\nu+2\epsilon}}$ for large values of P, we obtain:

$$\sum_{m=1}^{M}\mathbb{E}\left[(X_m^{[k]})^2\right]\approx\frac{1}{3}\gamma^2P\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{\ell=1}^{L}(h_{nm}^{[kk]}\omega_{mn\ell}^{[k]})^2. \tag{A.9}$$

Therefore, the power constraint P at all transmitters is satisfied if

$$\gamma^2=\min_{k\in\mathcal{K}}\frac{3}{\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{\ell=1}^{L}(h_{nm}^{[kk]}\omega_{mn\ell}^{[k]})^2}.$$

Appendix B

Appendices for Chapter 3

B.1 Derivation of rank of matrix $\mathbf{P}^{[k]}$

We use the following simple lemmas in derivation of rank$(\mathbf{P}^{[k]})$ without any proof:

Lemma 3. *Let $\mathbf{A}_m$, $m = 1, \cdots, r$, be random matrices of size $u \times v_m$ generated independently according to continuous distributions. Then,*

$$\text{rank}([\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_r]) = \min\Big\{u, \sum_{m=1}^{r} \text{rank}(A_m)\Big\}, \qquad \textit{almost surely.}$$

Lemma 4. *Let $\mathbf{A}_m$, $m = 1, \cdots, r$, be random matrices which are generated independently according to continuous distributions and are such that the matrix multiplication $\mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_r$ is defined. Then,*

$$\text{rank}(\mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_r) = \min\Big\{\text{rank}(\mathbf{A}_1), \text{rank}(\mathbf{A}_2) \cdots, \text{rank}(\mathbf{A}_r)\Big\}, \qquad \textit{almost surely.}$$

The matrix $\mathbf{P}^{[k]}$ in (3.33) can be partitioned into six sub-matrices $\mathbf{P}_{ij}^{[k]}$, $1 \leq i \leq 3, 1 \leq$

$j \leq 2$, as follows:

$$\mathbf{P}^{[k]} = \left(\begin{array}{c|c} \mathbf{P}_{11}^{[k]} & \mathbf{P}_{12}^{[k]} \\ \hline \mathbf{P}_{21}^{[k]} & \mathbf{P}_{22}^{[k]} \\ \hline \mathbf{P}_{31}^{[k]} & \mathbf{P}_{32}^{[k]} \end{array}\right) = \left(\begin{array}{c|c} N_k W_\ell \times \mu_k & N_k W_\ell \times \min\{\mu_{\bar{k}}, N_k W_{\bar{k}}\} \\ \hline N_k(W_u - W_\ell) \times \mu_k & N_k(W_u - W_\ell) \times \min\{\mu_{\bar{k}}, N_k W_{\bar{k}}\} \\ \hline N_k(W - W_u) \times \mu_k & N_k(W - W_u) \times \min\{\mu_{\bar{k}}, N_k W_{\bar{k}}\} \end{array}\right). \tag{B.1}$$

To calculate the rank of $\mathbf{P}^{[k]}$, let us define:

$$\mathbf{Q}_i^{[k]} \triangleq [\mathbf{P}_{i1}^{[k]}, \mathbf{P}_{i2}^{[k]}], \quad r_i^{[k]} \triangleq \operatorname{rank}(\mathbf{Q}_i^{[k]}), \quad i = 1, 2, 3. \tag{B.2}$$

Since the sub-matrices $\mathbf{P}_{i1}^{[k]}$ and $\mathbf{P}_{i2}^{[k]}$ are independent for each $i \in \{1, 2, 3\}$, using Lemma 3, one can write:

$$\begin{aligned} r_1^{[k]} &= \min\left\{N_k W_\ell, \operatorname{rank}(\mathbf{P}_{11}^{[k]}) + \operatorname{rank}(\mathbf{P}_{12}^{[k]})\right\}, \\ r_2^{[k]} &= \min\left\{N_k(W_u - W_\ell), \operatorname{rank}(\mathbf{P}_{21}^{[k]}) + \operatorname{rank}(\mathbf{P}_{22}^{[k]})\right\}, \\ r_3^{[k]} &= \min\left\{N_k(W - W_u), \operatorname{rank}(\mathbf{P}_{31}^{[k]}) + \operatorname{rank}(\mathbf{P}_{32}^{[k]})\right\}, \\ \operatorname{rank}(\mathbf{P}^{[k]}) &= \min\left\{\mu_k + \min\{\mu_{\bar{k}}, N_k W_{\bar{k}}\}, r_1^{[k]} + r_2^{[k]} + r_3^{[k]}\right\}. \end{aligned}$$

Based on the values of W_k and $W_{\bar{k}}$, we consider two different cases:

(a) $W_k \leq W_{\bar{k}}$

- $\mathbf{P}_{11}^{[k]} = \mathbf{H}_{[1:W_k]}^{[kk]} \mathbf{F}_{[1:W_k]}^{[k]}$, and thus, $\operatorname{rank}(\mathbf{P}_{11}^{[k]}) = \min\{\mu_k, N_k W_k\}$.
- $\mathbf{P}_{21}^{[k]}$ and $\mathbf{P}_{31}^{[k]}$ are given by:

$$\mathbf{P}_{21}^{[k]} = \mathbf{H}_{[W_k+1:W_{\bar{k}}]}^{[kk]} \mathbf{G}_{[W_k+1:W_{\bar{k}}]}^{[k]} \mathbf{H}_{[1:W_k]}^{[\bar{k}k]} \mathbf{F}_{[1:W_k]}^{[k]}, \tag{B.3}$$

$$\mathbf{P}_{31}^{[k]} = \mathbf{H}_{[W_{\bar{k}}+1:W]}^{[kk]} \mathbf{G}_{[W_{\bar{k}}+1:W]}^{[k]} \mathbf{H}_{[1:W_k]}^{[\bar{k}k]} \mathbf{F}_{[1:W_k]}^{[k]}, \tag{B.4}$$

where $\mathbf{G}^{[k]}_{[W_k+1:W_{\bar{k}}]}$ and $\mathbf{G}^{[k]}_{[W_{\bar{k}}+1:W]}$ contain respectively the first $M_k(W_{\bar{k}} - W_k)$ rows and the last $M_k(W - W_{\bar{k}})$ rows of the precoding matrix $\mathbf{G}^{[k]}_{[W_k+1:W]}$, i.e.,

$$\mathbf{G}^{[k]}_{[W_k+1:W]} = \left(\begin{array}{c} \mathbf{G}^{[k]}_{[W_k+1:W_{\bar{k}}]} \\ \hline \mathbf{G}^{[k]}_{[W_{\bar{k}}+1:W]} \end{array}\right). \tag{B.5}$$

Therefore, $\mathbf{G}^{[k]}_{[W_k+1:W_{\bar{k}}]}$ (resp. $\mathbf{G}^{[k]}_{[W_{\bar{k}}+1:W]}$) is of rank $\min\left\{M_k(W_{\bar{k}} - W_k), N_{\bar{k}}W_k\right\}$ (resp. $\min\left\{M_k(W - W_{\bar{k}}), N_{\bar{k}}W_k\right\}$) almost surely. Since the above matrices are randomly generated independent of each other, by Lemma 4, we have

$$\begin{aligned}
\operatorname{rank}(\mathbf{P}^{[k]}_{21}) &= \min\Bigg\{ \operatorname{rank}(\mathbf{H}^{[kk]}_{[W_k+1:W_{\bar{k}}]}), \operatorname{rank}(\mathbf{G}^{[k]}_{[W_k+1:W_{\bar{k}}]}), \\
&\qquad \operatorname{rank}(\mathbf{H}^{[\bar{k}k]}_{[1:W_k]}), \operatorname{rank}(\mathbf{F}^{[k]}_{[1:W_k]}) \Bigg\} \\
&= \min\Bigg\{ \min\{M_k, N_k\}(W_{\bar{k}} - W_k), \min\left\{M_k(W_{\bar{k}} - W_k), N_{\bar{k}}W_k\right\}, \\
&\qquad \min\{M_k, N_{\bar{k}}\}W_k, \mu_k \Bigg\} \\
&= \min\Bigg\{ \min\{M_k, N_k\}(W_{\bar{k}} - W_k), N_{\bar{k}}W_k, \mu_k \Bigg\},
\end{aligned} \tag{B.6}$$

and Similarly,

$$\operatorname{rank}(\mathbf{P}^{[k]}_{31}) = \min\Bigg\{ \min\{M_k, N_k\}(W - W_{\bar{k}}), N_{\bar{k}}W_k, \mu_k \Bigg\}. \tag{B.7}$$

– $\mathbf{P}^{[k]}_{12}$ and $\mathbf{P}^{[k]}_{22}$ are such that

$$\mathbf{L}^{[k\bar{k}]}_{[1:W_{\bar{k}}]} = \left(\begin{array}{c} \mathbf{P}^{[k]}_{12} \\ \hline \mathbf{P}^{[k]}_{22} \end{array}\right). \tag{B.8}$$

As such,

$$\operatorname{rank}(\mathbf{P}^{[k]}_{12}) = \min\left\{\mu_{\bar{k}}, \min\{M_{\bar{k}}, N_k\}W_k\right\}, \tag{B.9}$$

$$\operatorname{rank}(\mathbf{P}^{[k]}_{22}) = \min\left\{\mu_{\bar{k}}, \min\{M_{\bar{k}}, N_k\}(W_{\bar{k}} - W_k)\right\}, \tag{B.10}$$

almost surely.

– $\mathbf{P}_{32}^{[k]}$ is given by

$$\mathbf{P}_{32}^{[k]} = \mathbf{H}_{[W_{\bar{k}}+1:W]}^{[k\bar{k}]} \mathbf{G}_{[W_{\bar{k}}+1:W]}^{[\bar{k}]} \mathbf{L}_{[1:W_{\bar{k}}]}^{[k\bar{k}]}. \tag{B.11}$$

The three matrices on the right hand side of (B.11) are independent of each other, and thus, using Lemma 4,

$$\begin{aligned}
\mathrm{rank}(\mathbf{P}_{32}^{[k]}) &= \min\left\{ \mathrm{rank}(\mathbf{H}_{[W_{\bar{k}}+1:W]}^{[k\bar{k}]}), \mathrm{rank}(\mathbf{G}_{[W_{\bar{k}}+1:W]}^{[\bar{k}]}), \mathrm{rank}(\mathbf{L}_{[1:W_{\bar{k}}]}^{[k\bar{k}]}) \right\} \\
&= \min\left\{ \min\{M_{\bar{k}}, N_k\}(W-W_{\bar{k}}), \min\left\{ M_{\bar{k}}(W-W_{\bar{k}}), N_k W_{\bar{k}} \right\}, \min\left\{ \mu_{\bar{k}}, N_k W_{\bar{k}} \right\} \right\} \\
&= \min\left\{ \min\{M_{\bar{k}}, N_k\}(W-W_{\bar{k}}), N_k W_{\bar{k}}, \mu_{\bar{k}} \right\}.
\end{aligned} \tag{B.12}$$

Combining the above results, we have

$$\begin{aligned}
r_1^{[k]} &= \min\left\{ N_k W_k, \min\left\{ \mu_k, N_k W_k \right\} + \min\left\{ \mu_{\bar{k}}, \min\{M_{\bar{k}}, N_k\} W_k \right\} \right\} \\
&= \min\left\{ N_k W_k, \mu_k + M_{\bar{k}} W_k, \mu_1 + \mu_2 \right\}, \\
r_2^{[k]} &= \min\left\{ N_k(W_{\bar{k}} - W_k), \min\left\{ \min\{M_k, N_k\}(W_{\bar{k}} - W_k), N_{\bar{k}} W_k, \mu_k \right\} + \right. \\
&\qquad\qquad \left. \min\left\{ \min\{M_{\bar{k}}, N_k\}(W_{\bar{k}} - W_k), \mu_{\bar{k}} \right\} \right\} \\
&= \min\left\{ N_k(W_{\bar{k}} - W_k), \min\left\{ M_k(W_{\bar{k}} - W_k), N_{\bar{k}} W_k, \mu_k \right\} + \min\left\{ M_{\bar{k}}(W_{\bar{k}} - W_k), \mu_{\bar{k}} \right\} \right\}, \\
r_3^{[k]} &= \min\left\{ N_k(W - W_{\bar{k}}), \min\left\{ \min\{M_k, N_k\}(W - W_{\bar{k}}), N_{\bar{k}} W_k, \mu_k \right\} \right. \\
&\qquad\qquad \left. + \min\left\{ \min\{M_{\bar{k}}, N_k\}(W - W_{\bar{k}}), N_k W_{\bar{k}}, \mu_{\bar{k}} \right\} \right\} \\
&= \min\left\{ N_k(W - W_{\bar{k}}), \min\left\{ M_k(W - W_{\bar{k}}), N_{\bar{k}} W_k, \mu_k \right\} + \min\left\{ M_{\bar{k}}(W - W_{\bar{k}}), N_k W_{\bar{k}}, \mu_{\bar{k}} \right\} \right\}.
\end{aligned}$$

(b) $W_k > W_{\bar{k}}$

Parallel to the arguments for the case of $W_k \le W_{\bar{k}}$, we have the following results for this case:

$$\mathrm{rank}(\mathbf{P}_{11}^{[k]}) = \min\Big\{\mu_k, \min\{M_k, N_k\}W_{\bar{k}}\Big\},$$
$$\mathrm{rank}(\mathbf{P}_{21}^{[k]}) = \min\Big\{\mu_k, \min\{M_k, N_k\}(W_k - W_{\bar{k}})\Big\}$$

$$\mathrm{rank}(\mathbf{P}_{31}^{[k]}) = \min\Big\{\mathrm{rank}(\mathbf{H}_{[W_k+1:W]}^{[kk]}), \mathrm{rank}(\mathbf{G}_{[W_k+1:W]}^{[k]}), \mathrm{rank}(\mathbf{H}_{[1:W_k]}^{[\bar{k}k]}), \mathrm{rank}(\mathbf{F}_{[1:W_k]}^{[k]})\Big\}$$
$$= \min\Big\{\min\{M_k, N_k\}(W - W_k), \min\Big\{M_k(W - W_k), N_{\bar{k}}W_k\Big\}, \min\{M_k, N_{\bar{k}}\}W_k, \mu_k\Big\}$$
$$= \min\Big\{\min\{M_k, N_k\}(W - W_k), \min\{M_k, N_{\bar{k}}\}W_k, \mu_k\Big\},$$

$$\mathrm{rank}(\mathbf{P}_{12}^{[k]}) = \mathrm{rank}(\mathbf{L}_{[1:W_{\bar{k}}]}^{[k\bar{k}]}) = \min\Big\{\mu_{\bar{k}}, N_k W_{\bar{k}}\Big\},$$

$$\mathrm{rank}(\mathbf{P}_{22}^{[k]}) = \min\Big\{\mathrm{rank}(\mathbf{H}_{[W_{\bar{k}}+1:W_k]}^{[k\bar{k}]}), \mathrm{rank}(\mathbf{G}_{[W_{\bar{k}}+1:W_k]}^{[\bar{k}]}), \mathrm{rank}(\mathbf{L}_{[1:W_{\bar{k}}]}^{[k\bar{k}]})\Big\}$$
$$= \min\Big\{\min\{M_{\bar{k}}, N_k\}(W_k - W_{\bar{k}}), \min\Big\{M_{\bar{k}}(W_k - W_{\bar{k}}), N_k W_{\bar{k}}\Big\}, \min\Big\{\mu_{\bar{k}}, N_k W_{\bar{k}}\Big\}\Big\}$$
$$= \min\Big\{\min\Big\{M_{\bar{k}}, N_k)(W_k - W_{\bar{k}}\Big\}, N_k W_{\bar{k}}, \mu_{\bar{k}}\Big\},$$

$$\mathrm{rank}(\mathbf{P}_{32}^{[k]}) = \min\Big\{\min\{M_{\bar{k}}, N_k\}(W - W_k), N_k W_{\bar{k}}, \mu_{\bar{k}}\Big\}.$$

$$r_1^{[k]} = \min\Big\{N_k W_{\bar{k}}, \min\Big\{\mu_k, \min\{M_k, N_k\}W_{\bar{k}}\Big\} + \min\Big\{\mu_{\bar{k}}, N_k W_{\bar{k}}\Big\}\Big\}$$
$$= \min\Big\{N_k W_{\bar{k}}, \mu_{\bar{k}} + M_k W_{\bar{k}}, \mu_1 + \mu_2\Big\},$$

$$
\begin{aligned}
r_2^{[k]} &= \min\Bigg\{N_k(W_k - W_{\bar{k}}), \min\Big\{\min\{M_k, N_k\}(W_k - W_{\bar{k}}), \mu_k\Big\} \\
&\qquad\qquad + \min\Big\{\min\{M_{\bar{k}}, N_k\}(W_k - W_{\bar{k}}), N_k W_{\bar{k}}, \mu_{\bar{k}}\Big\}\Bigg\} \\
&= \min\Big\{N_k(W_k - W_{\bar{k}}), \min\Big\{M_{\bar{k}}(W_k - W_{\bar{k}}), N_k W_{\bar{k}}, \mu_{\bar{k}}\Big\} + \min\Big\{M_k(W_k - W_{\bar{k}}), \mu_k\Big\}\Big\},
\end{aligned}
$$

$$
\begin{aligned}
r_3^{[k]} &= \min\Big\{N_k(W - W_k), \min\Big\{\min\{M_k, N_k\}(W - W_k), \min\{M_k, N_{\bar{k}}\}W_k, \mu_k\Big\} \\
&\qquad\qquad + \min\Big\{\min\{M_{\bar{k}}, N_k\}(W - W_k), N_k W_{\bar{k}}, \mu_{\bar{k}}\Big\}\Big\} \\
&= \min\Big\{N_k(W - W_k), \min\Big\{M_k(W - W_k), N_{\bar{k}} W_k, \mu_k\Big\} + \min\Big\{M_{\bar{k}}(W - W_k), N_k W_{\bar{k}}, \mu_{\bar{k}}\Big\}\Big\}.
\end{aligned}
$$

The above results can be summarized as follows:

$$
\begin{aligned}
\operatorname{rank}(\mathbf{P}^{[k]}) &= \min\Big\{\mu_k + \min\{\mu_{\bar{k}}, N_k W_{\bar{k}}\}, r_1^{[k]} + r_2^{[k]} + r_3^{[k]}\Big\}, \\
r_1^{[k]} &= \min\Big\{N_k W_\ell, \mu_\ell + M_u W_\ell, \mu_1 + \mu_2\Big\}, \\
r_2^{[k]} &= \min\Big\{N_k(W_u - W_\ell), \min\Big\{M_\ell(W_u - W_\ell), N_u W_\ell, \mu_\ell\Big\} + \min\Big\{M_u(W_u - W_\ell), \mu_u\Big\}\Big\}, \\
r_3^{[k]} &= \min\Big\{N_k(W - W_u), \min\Big\{M_k(W - W_u), N_{\bar{k}} W_k, \mu_k\Big\} + \min\Big\{M_{\bar{k}}(W - W_u), N_k W_{\bar{k}}, \mu_{\bar{k}}\Big\}\Big\},
\end{aligned}
$$

where $u = \arg\max_k\{W_k\}$ and $\ell = \{1, 2\}\backslash\{u\}$.

B.2 Proof of achievability of T_3 for class $\mathcal{C}_8$

To show point $T_3 = (M_1, \frac{M_2(N_1 - M_1)}{N_1})$ is achievable for class $\mathcal{C}_8$, we show that $W^* = W_1^* = N_1$, $W_2^* = N_1 - M_1$, $\mu_1^* = M_1 W_1^*$, and $\mu_2^* = M_2 W_2^*$ satisfy the rank conditions in (3.38).

Since $W_1^* > W_2^*$, we have $u = 1$ and $\ell = 2$. Substituting in (3.37), we have

$$\begin{aligned}
r_1^{[1]} &= \min\Big\{N_1 W_2^*, \mu_2^* + M_1 W_2^*, \mu_1^* + \mu_2^*\Big\} \overset{\text{(a)}}{=} N_1 W_2^*, \\
r_2^{[1]} &= \min\Big\{N_1(W_1^* - W_2^*), \min\Big\{M_2(W_1^* - W_2^*), N_1 W_2^*, \mu_2^*\Big\} + \min\Big\{M_1(W_1^* - W_2^*), \mu_1^*\Big\}\Big\} \\
&\overset{\text{(b)}}{=} \min\Big\{N_1(W_1^* - W_2^*), N_1 W_2^* + M_1(W_1^* - W_2^*)\Big\} \overset{\text{(c)}}{=} N_1(W_1^* - W_2^*), \\
r_3^{[1]} &= 0,
\end{aligned}$$

where (a) and (b) follow from $M_1 < N_1 < M_2$ and (c) follows from $N_1 M_1 \leq N_1^2 - N_1 M_1 + M_1^2$ which is obviously true. Hence, $r_1^{[1]} + r_2^{[1]} + r_3^{[1]} = N_1 W_1^* = N_1^2$. On the other hand, $\mu_1^* + \min\{\mu_2^*, N_1 W_2^*\} = M_1 N_1 + N_1(N_1 - M_1) = N_1^2$, and therefore, the rank condition is verified for RX_1. For RX_2, we have

$$\begin{aligned}
r_1^{[2]} &= \min\Big\{N_2 W_2^*, \mu_2^* + M_1 W_2^*, \mu_1^* + \mu_2^*\Big\} \overset{\text{(a)}}{=} N_2 W_2^* = N_2(N_1 - M_1), \\
r_2^{[2]} &= \min\Big\{N_2(W_1^* - W_2^*), \min\Big\{M_2(W_1^* - W_2^*), N_1 W_2^*, \mu_2^*\Big\} + \min\Big\{M_1(W_1^* - W_2^*), \mu_1^*\Big\}\Big\} \\
&\overset{\text{(b)}}{=} \min\Big\{N_2(W_1^* - W_2^*), N_1 W_2^* + M_1(W_1^* - W_2^*)\Big\} = \min\Big\{N_2 M_1, N_1(N_1 - M_1) + M_1^2\Big\}, \\
r_3^{[2]} &= 0,
\end{aligned}$$

where (a) and (b) follow from $N_1 < N_2 < M_2$. Hence,

$$r_1^{[2]} + r_2^{[2]} + r_2^{[3]} = \min\Big\{N_1 N_2, (N_1 + N_2)(N_1 - M_1) + M_1^2\Big\}. \tag{B.13}$$

On the other hand,

$$\mu_2^* + \min\{\mu_1^*, N_2 W_1^*\} = M_2(N_1 - M_1) + M_1 N_1. \tag{B.14}$$

Recall that for class $\mathcal{C}_8$, $M_1 > \Delta'$ and $M_2 \leq A$. From $M_1 > \Delta'$, it follows that $N_1 N_2 > M_2(N_1 - M_1) + M_1 N_1$. From $M_2 \leq A$, it follows that $(N_1 + N_2)(N_1 - M_1) + M_1^2 > M_2(N_1 - M_1) + M_1 N_1$. Thus, the right hand side of (B.13) is greater than or equal to (B.14) and the proof is complete.

B.3 Proof of achievability of T_4 for classes $\mathcal{C}_6$ and $\mathcal{C}_7$

To show point $T_4 = (M_1, N_2 - M_1)$ is achievable for classes $\mathcal{C}_6$ and $\mathcal{C}_7$, we show that $W^* = W_1^* = \min\{M_2, L\}$, $W_2^* = N_2 - M_1$, $\mu_1^* = M_1 W_1^*$, and $\mu_2^* = \min\{M_2, L\}W_2^*$ satisfy the rank conditions in (3.38) for these classes. Since $W_1^* > W_2^*$, we have $u = 1$ and $\ell = 2$. Substituting in (3.37), we have

$$\begin{aligned}
r_1^{[k]} &= \min\left\{N_k W_2^*, \mu_2^* + M_1 W_2^*, \mu_1^* + \mu_2^*\right\} \overset{(a)}{=} N_k W_2^*,\\
r_2^{[k]} &= \min\left\{N_k(W_1^* - W_2^*), \min\left\{M_2(W_1^* - W_2^*), N_1 W_2^*, \mu_2^*\right\} + \min\left\{M_1(W_1^* - W_2^*), \mu_1^*\right\}\right\}\\
&\overset{(b)}{=} \min\left\{N_k(W_1^* - W_2^*), N_1 W_2^* + M_1(W_1^* - W_2^*)\right\} \overset{(c)}{=} N_k(W_1^* - W_2^*),\\
r_3^{[k]} &= 0,
\end{aligned}$$

where (a) and (b) follow from the assumption $M_1 \leq N_1 < N_2 < \min\{M_2, L\}$ and (c) follows from $N_2(W_1^* - W_2^*) \leq N_1 W_2^* + M_1(W_1^* - W_2^*)$ which is a straightforward result of the fact that $\min\{M_2, L\} - N_2 + M_1 \leq N_1$ for classes $\mathcal{C}_6$ and $\mathcal{C}_7$. Hence,

$$\begin{aligned}
r_1^{[1]} + r_2^{[1]} + r_3^{[1]} &= N_1 W_1^*,\\
r_1^{[2]} + r_2^{[2]} + r_3^{[2]} &= N_2 W_1^*.
\end{aligned} \tag{B.15}$$

On the other hand,

$$\begin{aligned}
\mu_1^* + \min\{\mu_2^*, N_1 W_2^*\} &= M_1 W_1^* + N_1 W_2^*,\\
\mu_2^* + \min\{\mu_1^*, N_2 W_1^*\} &= \min\{M_2, L\}W_2^* + M_1 W_1^*.
\end{aligned} \tag{B.16}$$

Therefore, the rank conditions are met if:

$$M_1 W_1^* + N_1 W_2^* \leq N_1 W_1^*, \tag{B.17}$$

$$\min\{M_2, L\}W_2^* + M_1 W_1^* \leq N_2 W_1^*. \tag{B.18}$$

It is easy to see that (B.18) holds with equality. For class $\mathcal{C}_6$, (B.17) is equivalent to $M_1 \leq \Delta$ which is valid for class $\mathcal{C}_6$. For class $\mathcal{C}_7$, (B.17) is equivalent to $M_1 \leq \Delta'$ which is valid for class $\mathcal{C}_7$.

B.4 Proof of achievability of T_5 for classes $\mathcal{C}_8$ and $\mathcal{C}_9$

To show $T_5 = (\frac{N_1(M_2-N_2)}{M_2-N_1}, \frac{M_2(N_2-N_1)}{M_2-N_1})$ is achievable for classes $\mathcal{C}_8$ and $\mathcal{C}_9$, we show that $W^* = W_1^* = M_2 - N_1$, $W_2^* = N_2 - N_1$, $\mu_1^* = N_1(M_2 - N_2)$, and $\mu_2^* = M_2 W_2^*$ satisfy the rank conditions in (3.38). Since $W_1^* > W_2^*$, we have $u = 1$ and $\ell = 2$. Substituting in (3.37), we have

$$r_1^{[k]} = \min\left\{N_k W_2^*, \mu_2^* + M_1 W_2^*, \mu_1^* + \mu_2^*\right\} \overset{\text{(a)}}{=} N_k W_2^*,$$

$$r_2^{[k]} = \min\left\{N_k(W_1^* - W_2^*), \min\left\{M_2(W_1^* - W_2^*), N_1 W_2^*, \mu_2^*\right\} + \min\left\{M_1(W_1^* - W_2^*), \mu_1^*\right\}\right\}$$
$$\overset{\text{(b)}}{=} \min\left\{N_k(W_1^* - W_2^*), N_1 W_2^* + M_1(W_1^* - W_2^*)\right\} \overset{\text{(c)}}{=} N_k(W_1^* - W_2^*),$$

$$r_3^{[k]} = 0,$$

where (a) and (b) follow from $M_1 < N_1 < M_2$. To prove (c), we need to show that $N_1 W_2^* + M_1(W_1^* - W_2^*) \geq N_2(W_1^* - W_2^*)$, or equivalently,

$$\frac{M_2 - N_2}{N_2 - N_1} \leq \frac{N_1}{N_2 - M_1}. \tag{B.19}$$

We prove (B.19) for each class separately:

- For class $\mathcal{C}_9$, (B.19) follows from the following chain of relations:
$$\frac{M_2 - N_2}{N_2 - N_1} = \frac{N_1(M_2 - N_2)}{(N_1 - \Delta)(N_2 - M_1)} \overset{\text{(a)}}{\leq} \frac{N_1}{N_2 - M_1}, \tag{B.20}$$
where (a) follows from the assumption $M_2 < N_1 + N_2 - \Delta$ for class $\mathcal{C}_9$.

- For class $\mathcal{C}_8$, we have
$$\frac{M_2 - N_2}{N_2 - N_1} = \frac{\Delta'}{N_1 - \Delta'}.$$
On the other hand, from $M_2 \leq L$, it follows that $\Delta' \leq \Delta = \frac{N_1(N_1 - M_1)}{N_2 - M_1}$. Since $M_1 > \Delta'$ and Δ is a decreasing function of M_1, we have
$$\Delta' < \frac{N_1(N_1 - M_1)}{N_2 - M_1} < \frac{N_1(N_1 - \Delta')}{N_2 - \Delta'} \;\Rightarrow\; \frac{\Delta'}{N_1 - \Delta'} < \frac{N_1}{N_2 - \Delta'} < \frac{N_1}{N_2 - M_1}, \tag{B.21}$$
which is the desired result.

Hence,

$$\begin{aligned} r_1^{[1]} + r_2^{[1]} + r_3^{[1]} &= N_1 W_1^* = N_1(M_2 - N_1), \\ r_1^{[2]} + r_2^{[2]} + r_3^{[2]} &= N_2 W_1^* = N_2(M_2 - N_1). \end{aligned} \tag{B.22}$$

On the other hand,

$$\begin{aligned} \mu_1^* + \min\{\mu_2^*, N_1 W_2^*\} &= \mu_1^* + N_1 W_2^* = N_1(M_2 - N_1), \\ \mu_2^* + \min\{\mu_1^*, N_2 W_1^*\} &= \mu_2^* + \mu_1^* = N_2(M_2 - N_1). \end{aligned} \tag{B.23}$$

Therefore, the rank conditions are met and the proof is complete.

B.5 Proof of achievability of T_7 for class $\mathcal{C}_9$

To show $T_7 = (\frac{N_1(N_1+N_2-M_2)}{N_1+L-M_2}, \frac{M_2(N_1-M_1)}{N_1+L-M_2})$ is an achievable point for class $\mathcal{C}_9$, we show that $W^* = W_1^* = N_1+L-M_2$, $W_2^* = N_1 - M_1$, $\mu_1^* = N_1(N_1+N_2-M_2)$, and $\mu_2^* = M_2 W_2^*$ satisfy the rank conditions in (3.38). Since $W_1^* > W_2^*$, we have $u = 1$ and $\ell = 2$. Substituting in (3.37), we have

$$\begin{aligned} r_1^{[k]} &= \min\left\{N_k W_2^*, \mu_2^* + M_1 W_2^*, \mu_1^* + \mu_2^*\right\} \overset{\text{(a)}}{=} N_k W_2^*, \\ r_2^{[k]} &= \min\left\{N_k(W_1^* - W_2^*), \min\left\{M_2(W_1^* - W_2^*), N_1 W_2^*, \mu_2^*\right\} + \min\left\{M_1(W_1^* - W_2^*), \mu_1^*\right\}\right\} \\ &\overset{\text{(b)}}{=} \min\left\{N_k(W_1^* - W_2^*), N_1 W_2^* + M_1(W_1^* - W_2^*)\right\} \overset{\text{(c)}}{=} N_k(W_1^* - W_2^*), \\ r_3^{[k]} &= 0, \end{aligned}$$

where (a) and (b) follow from $M_1 < N_1 < M_2$ and the fact that $\mu_1^* = N_1(W_1^* - W_2^*)$. To prove (c), we need to show that $N_2(W_1^* - W_2^*) \le N_1 W_2^* + M_1(W_1^* - W_2^*)$, or equivalently,

$$\frac{M_2 - N_2}{N_2 - N_1} \le \frac{N_1}{N_2 - M_1},$$

which was proved in (B.20). Hence, $r_1^{[k]} + r_2^{[k]} + r_3^{[k]} = N_k W_1^*$, $k = 1, 2$. On the other hand,

$$\mu_1^* + \min\{\mu_2^*, N_1 W_2^*\} = \mu_1^* + N_1 W_2^* = N_1 W_1^*,$$
$$\mu_2^* + \min\{\mu_1^*, N_2 W_1^*\} = \mu_2^* + \mu_1^* = M_2 W_2^* + N_1(W_1^* - W_2^*).$$

and therefore, the rank condition holds at RX_1. The rank condition holds at RX_2 provided that $M_2 W_2^* + N_1(W_1^* - W_2^*) \le N_2 W_1^*$, or equivalently,

$$\frac{M_2 - N_2}{N_2 - N_1} \le \frac{N_1}{N_2 - M_1},$$

which is again true according to (B.20).

B.6 Proof of achievability of T_8 for class $\mathcal{C}_{10}$

To show $T_8 = (\frac{N_1^2}{L}, N_2 - \frac{N_1^2}{L})$ is an achievable point for class $\mathcal{C}_{10}$, we show that $W^* = W_1^* = L$, $W_2^* = N_2 - M_1$, $\mu_1^* = N_1^2$, and $\mu_2^* = N_2 L - N_1^2$ satisfy the rank conditions in (3.38). Since $W_1^* > W_2^*$, we have $u = 1$ and $\ell = 2$. Substituting in (3.37), we have

$$\begin{aligned}
r_1^{[k]} &= \min\left\{N_k W_2^*, \mu_2^* + M_1 W_2^*, \mu_1^* + \mu_2^*\right\} \overset{(a)}{=} N_k W_2^*,\\
r_2^{[k]} &= \min\left\{N_k(W_1^* - W_2^*), \min\left\{M_2(W_1^* - W_2^*), N_1 W_2^*, \mu_2^*\right\} + \min\left\{M_1(W_1^* - W_2^*), \mu_1^*\right\}\right\}\\
&\overset{(b)}{=} \min\left\{N_k(W_1^* - W_2^*), N_1 W_2^* + M_1(W_1^* - W_2^*)\right\} \overset{(c)}{=} N_k(W_1^* - W_2^*),\\
r_3^{[k]} &= 0,
\end{aligned}$$

where (a) and (b) follow from $\mu_2^* > N_2 W_2^*$ and $W_1^* - W_2^* = N_1$, and (c) follows from the fact that $N_2(W_1^* - W_2^*) = N_1 W_2^* + M_1(W_1^* - W_2^*) = N_1 N_2$. Hence, $r_1^{[k]} + r_2^{[k]} + r_3^{[k]} = N_k W_1^*$, $k = 1, 2$. On the other hand,

$$\mu_1^* + \min\{\mu_2^*, N_1 W_2^*\} = \mu_1^* + N_1 W_2^* = N_1 W_1^*,$$
$$\mu_2^* + \min\{\mu_1^*, N_2 W_1^*\} = \mu_2^* + \mu_1^* = N_2 W_1^*,$$

and therefore, the rank conditions hold.

References

[1] C. E. Shannon and W. Weaver, "The mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[2] G. Forney and D. Costello, "Channel coding: The road to channel capacity," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1150–1177, 2007.

[3] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *The Annals of Probability*, vol. 2, no. 5, pp. 805–814, 1974.

[4] A. Carleial, "A case where interference does not reduce capacity (corresp.)," *Information Theory, IEEE Transactions on*, vol. 21, no. 5, pp. 569–570, 1975.

[5] H. Sato, "The capacity of the Gaussian interference channel under strong interference (corresp.)," *Information Theory, IEEE Transactions on*, vol. 27, no. 6, pp. 786–788, 1981.

[6] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *Information Theory, IEEE Transactions on*, vol. 55, no. 2, pp. 620–643, 2009.

[7] 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," *Information Theory, IEEE Transactions on*, vol. 55, no. 7, pp. 3032–3050, 2009.

References

[8] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1345–1356, 2004.

[9] H. Sato, "On degraded Gaussian two-user channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 5, pp. 637–640, 1978.

[10] M. Costa, "On the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 31, no. 5, pp. 607–615, 1985.

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

[12] E. Telatar and D. Tse, "Bounds on the capacity region of a class of interference channels," in *IEEE Int. Symp. Information Theory*, pp. 2871–2874, 2007.

[13] S. Karmakar and M. Varanasi, "Capacity of the mimo interference channel to within a constant gap," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pp. 2193–2197, IEEE, 2011.

[14] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3457–3470, 2008.

[15] M. Maddah-Ali, A. Motahari, and A. Khandani, "Signaling over mimo multi-base systems: Combination of multi-access and broadcast schemes," in *Information Theory, 2006 IEEE International Symposium on*, pp. 2104 –2108, july 2006.

[16] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3425–3441, 2008.

[17] N. Jindal, "Mimo broadcast channels with finite-rate feedback," *Information Theory, IEEE Transactions on*, vol. 52, no. 11, pp. 5045–5060, 2006.

[18] R. Krishnamachari and M. Varanasi, "Interference alignment under limited feedback for mimo interference channels," in *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pp. 619–623, IEEE, 2010.

[19] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom of wireless X networks," *Information Theory, IEEE Transactions on*, vol. 55, no. 9, pp. 3893–3908, 2009.

[20] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many gaussian interference channels," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4566–4592, 2010.

[21] V. Cadambe, S. Jafar, and S. Shamai, "Interference alignment on the deterministic channel and application to fully connected awgn interference networks," in *Information Theory Workshop, 2008. ITW'08. IEEE*, pp. 41–45, IEEE, 2008.

[22] R. H. Etkin and E. Ordentlich, "The degrees-of-freedom of the K-user Gaussian interference channel is discontinuous at rational channel coefficients," *Information Theory, IEEE Transactions on*, vol. 55, no. 11, pp. 4932–4946, 2009.

[23] A. Motahari, A. Khandani, and S. Gharan, "On the degrees of freedom of the 3-user gaussian interference channel: The symmetric case," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, pp. 1914–1918, IEEE, 2009.

[24] V. Cadambe, S. Jafar, and C. Wang, "Interference alignment with asymmetric complex signalingsettling the høst-madsen–nosratinia conjecture," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4552–4565, 2010.

[25] A. Motahari, S. Gharan, and A. Khandani, "Real interference alignment with real numbers," *arXiv preprint arXiv:0908.1208*, 2009.

[26] A. Motahari, S. Gharan, M. Maddah-Ali, and A. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *arXiv preprint arXiv:0908.2282*, 2009.

[27] T. Gou and S. Jafar, "Degrees of freedom of the K-user $M \times N$ MIMO interference channel," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6040–6057, 2010.

[28] C. Huang, S. A. Jafar, S. Shamai, and S. Vishwanath, "On degrees of freedom region of MIMO networks without CSIT," *Arxiv preprint arXiv:0909.4017*, 2009.

[29] C. S. Vaze and M. K. Varanasi, "The degrees of freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *Arxiv preprint arXiv:0909.5424*, 2009.

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

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

[32] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pp. 1188–1195, 2010.

[33] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region of the two-user MIMO broadcast channel with delayed CSIT," in *Information Theory, 2011. ISIT 2011. IEEE International Symposium on*, pp. 331–335, 2011.

[34] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the degrees of freedom of three-user MIMO broadcast channel with delayed CSIT," in *Information Theory, 2011. ISIT 2011. IEEE International Symposium on*, pp. 341–345, 2011.

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

[36] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the degrees of freedom of K-user SISO interference and X channels with delayed CSIT," *Submitted to Information Theory, IEEE Transactions on, Available online at: Arxiv preprint arXiv:1109.4314*, 2011.

[37] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "Interference and X networks with noisy cooperation and feedback," *Submitted to Information Theory, IEEE Transactions on, Available online at: Arxiv preprint arXiv:1201.5805*, 2012.

[38] A. Ghasemi, A. Motahari, and A. Khandani, "Interference alignment for the K-user MIMO interference channel," in *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pp. 360–364, IEEE, 2010.

[39] S. A. Jafar and M. Fakhereddin, "Degrees of freedom for the MIMO interference channel," in *Information Theory, 2006. ISIT 2006. IEEE International Symposium on*, 2006.

[40] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Information Theory, 2005. ISIT 2005. IEEE International Symposium on*, pp. 2065–2069, 2005.

[41] V. Bernik, D. Kleinbock, and G. Margulis, "Khintchine-type theorems on manifolds: the convergence case for standard and multiplicative versions," *International Mathematics Research Notices*, no. 9, pp. 453–486, 2001.

[42] V. Beresnevich, "A Groshev type theorem for convergence on manifolds," *Acta Mathematica Hungarica 94*, no. 1-2, pp. 99–130, 2002.

References

[43] M. Amir, A. El-Keyi, and M. Nafie, "A new achievable dof region for the 3-user MxN symmetric interference channel," *arXiv preprint arXiv:1105.4026*, 2011.

[44] C. Wang, T. Gou, and S. Jafar, "Subspace alignment chains and the degrees of freedom of the three-user mimo interference channel," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, pp. 2471–2475, IEEE, 2012.

[45] D. Kleinbock, "Baker-sprindžuk conjectures for complex analytic manifolds," in *Algebraic groups and Arithmetic, Tata Inst Fund Res*, Citeseer, 2004.

[46] A. Ghasemi, A. Motahari, and A. Khandani, "On the degrees of freedom of X channel with delayed csit," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pp. 767–770, IEEE, 2011.

[47] A. Ghasemi, M. Abdoli, and A. Khandani, "On the degrees of freedom of MIMO x channel with delayed csit," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, pp. 1892–1896, IEEE, 2012.

[48] C. Vaze and M. Varanasi, "The degrees of freedom region of the mimo interference channel with delayed csit," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pp. 2203–2207, IEEE, 2011.

[49] C. S. Vaze and M. K. Varanasi, "The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *Information Theory, IEEE Transactions on*, vol. 58, no. 8, pp. 5354–5374, 2012.

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

[51] R. Tandon, S. Mohajer, H. V. Poor, and S. Shamai, "On X-channels with feedback and delayed CSI," *Arxiv preprint arXiv:1201.6313*, 2012.

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

[53] U. Niesen and M. ali M.A., “Interference alignment: From degrees-of-freedom to constant-gap capacity approximations,” in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, pp. 2077–2081, IEEE, 2012.

[54] R. Graham, D. Knuth, and O. Patashnik, *Concrete mathematics: a foundation for computer science*. Addison-Wesley Reading, MA, 1994.