ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2014-ENST-0066

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Communication et Electronique »

présentée et soutenue publiquement par

Xinping YI

lle 21 Octobre 2014

Contrôle des interférences dans les réseaux sans-fils en présence d'informations limitées sur le canal

Directeur de thèse : **David GESBERT**

Jury

M. Dirk SLOCK, Professeur, EURECOM	Président du jury
M. Eduard JORSWIECK, Professeur, Dresden University of Technology (TU Dresden)	Rapporteur
M. Abdellatif ZAIDI, Professeur Associé, Université Paris-Est Marne la Vallée	Rapporteur
M. Bruno CLERCKX, Lecturer, Imperial College	Examinateur
M. Sheng YANG, Maître de Conférences, SUPELEC	Examinateur

TELECOM ParisTech
école de l'Institut Télécom - membre de ParisTech

THÈSE

DISSERTATION

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
from TELECOM ParisTech

Specialization: Communication and Electronics

Xinping YI

Interference Management in Wireless Networks with Channel Uncertainty

Will be defended on 21st October, 2014 before a committee composed of:

President of the Jury		
Professor	Dirk SLOCK	EURECOM
Reviewers		
Professor	Eduard JORSWIECK	TU Dresden
Professor	Abdellatif ZAIDI	Université Paris-Est Marne la Vallée
Examiners		
Professor	Bruno CLERCKX	Imperial College
Professor	Sheng YANG	SUPELEC
Thesis Supervisor		
Professor	David GESBERT	EURECOM

THÈSE

présentée pour l'obtention du grade de

Docteur de TELECOM ParisTech

Spécialité: Communication et Electronique

Xinping YI

Contrôle des interférences dans les réseaux sans-fils en présence d'informations limitées sur le canal

Soutenance de thèse prévue le 21 Octobre 2014 devant le jury composé de :

Président du jury		
Professeur	Dirk SLOCK	EURECOM
Rapporteurs		
Professeur	Eduard JORSWIECK	TU Dresden
Professeur	Abdellatif ZAIDI	Université Paris-Est Marne la Vallée
Examinateurs		
Professeur	Bruno CLERCKX	Imperial College
Professeur adjoint	Sheng YANG	SUPELEC
Directeur de thèse		
Professeur	David GESBERT	EURECOM

Abstract

Having multiple terminals access simultaneously the wireless medium enhances the capacity of cellular networks, whereas inevitably introduces multiuser interference, which in turn limits spectral efficiency. Hence, interference management constitutes a major means to improve spectral efficiency. Crucially, the spectral efficiency improvement by most interference management techniques usually stems from the assumption of the availability of channel state information at the transmitters (CSIT) obtained through feedback. The inaccuracy and/or latency of channel knowledge via feedback are major sources of channel uncertainty that affect network performance to a great extent. This is particularly true of methods relying on multiple-antenna precoding.

This thesis focuses on interference management with channel uncertainty in multiuser multiple-input multiple-output (MIMO) networks, where the channel uncertainty comes from feedback *delays* as well as the strictly *limited capacity* of feedback links. As such, in the extreme cases, the transmitters may either possess *sufficiently precise* CSI but with *large latency* or have access *instantaneously* to a *coarse* channel information (e.g., topological feedback, with one bit indicating whether the channel is strong or weak). The former renders the available CSI feedback obsolete under a fast fading channel, and the latter makes the transmitters *almost blind* to the exception of the binary indicator of channel strength. This thesis focuses on different regimes of CSIT availability, trying to address two fundamental problems: (1) How to best exploit delayed feedback? (2) How to best exploit drastically reduced feedback (e.g., such a topology-related feedback)?

Regarding delayed feedback, a recent breakthrough has shown that even completely outdated channel feedback is still useful. This surprising finding results from the idea of retrospective interference alignment, which achieves substantial degrees of freedom (DoF) in the infinite signal-to-noise-ratio (SNR) regime, even if the channel is independent and identically distributed (i.i.d.) across time. Although inspiring and fascinating from a conceptual point of

view, this result is subject to improvement. It can be seen as optimistic in that it intrinsically focuses on the asymptotic SNR behavior, leaving aside in particular the question of how shall precoding be done practically using stale CSIT at finite SNR. But it can also be seen as pessimistic in that it assumes the channel is i.i.d. across time, where delayed CSIT bears no correlation with current channel realizations. In this regard, we will address two problems under delayed feedback settings: (1) How does delayed CSIT improve sum rate performance at *finite* SNR? (2) Can we do better if the channel exists *correlation across time* in multiuser MIMO networks such that an estimate of current CSIT is also available?

When it comes to topological feedback, at first glance, a limitation to sole topological information is like a drop in a CSIT ocean, making it difficult for the transmitters to extract substantial DoF gain, seemingly useless in the sense of DoF. Recently, interference networks with no CSIT except for the network connectivity graph have been studied under the so-called topological interference management (TIM) framework. A surprising fact was revealed that the network performance can be significantly improved with this sole topological information, *provided that the network is partially connected*. Remarkably, interference alignment was shown to be beneficial over orthogonal access schemes under TIM settings. Nevertheless, there are still many interesting open problems: (1) When orthogonal access is optimal under TIM settings? (2) Is topological feedback beneficial in the context of an interference network where a message exchange mechanism between transmitters pre-exists (e.g., like in cellular networks enabled with a coordinated multipoint mechanism)?

In this thesis, above questions are addressed mainly from the information theoretic perspective, but also from signal processing and communication theoretic ones, from which some new interference management techniques are proposed to combat channel uncertainty and improve spectral efficiency, whereas the fundamental limitations of some existing interference management techniques are also revealed.

Abrégé

L' accès simultané de multiples terminaux au réseau sans fil améliore la capacité des réseaux cellulaires, tandis qu'il introduit inévitablement des interférences entre utilisateurs, ce qui peut à son tour limiter l'efficacité spectrale. Par conséquent, la gestion des interférences constitue un des principaux moyens d'améliorer l'efficacité spectrale. Fondamentalement, l'amélioration de l'efficacité spectrale par la plupart des techniques de gestion d'interférence repose généralement sur la disponibilité de l'informations de canal aux émetteurs obtenues grâce à un « feedback » de la part des utilisateurs. L'inexactitude et/ou le délai dans l'acquisition de la connaissance du canal sont les principales sources d'incertitude sur l'état du canal qui affectent les performances du réseau. Cela est particulièrement vrai pour des méthodes à entrées multiples et sorties multiples (MIMO) reposant sur des méthodes de précodage au niveau des antennes.

Cette thèse porte sur la contrôlé des interférences en présence d'incertitude de canal dans les réseaux MIMO multi-utilisateurs, dans le cas où l'incertitude de canal provient de retards de rétroaction ainsi que de la capacité limitée des systèmes de feedback. Ainsi, dans les cas extrêmes, les émetteurs peuvent posséder soit des estimées suffisamment précises mais avec une grande latence, soit avoir accès instantanément à une information de canal grossière (par exemple, le feedback topologique, avec un bit indiquant si le canal est fort ou faible). Le premier rend les estimées disponibles obsolètes dans un canal changeant rapidement, et le second revient à considérer des émetteurs presque aveugle à l'exception de l'indicateur binaire sur l'amplitude du canal. Cette thèse porte sur les différents régimes de la disponibilité de l'information de canal, en essayant de répondre à deux problèmes fondamentaux: (1) Comment exploiter d'une manière optimale une information de canal retardée? (2) Comment exploiter au mieux une information fortement réduite (comme par exemple un feedback topologique)?

Dans cette thèse, les questions ci-dessus sont adressées principalement avec la perspective de la théorie de l'information ainsi que de la théorie

du traitement du signal et des communications. À partir de celles-ci, de nouvelles techniques de gestion des interférences sont proposées pour lutter contre l'incertitude de canal et améliorer l'efficacité spectrale, tandis que les limites fondamentales de certaines techniques existantes de gestion des interférences sont également révélés.

Acknowledgements

During my Ph.D. study at EURECOM, I have benefited tremendously from the interaction with many extraordinary people here in research and personal life.

First of all, special thanks and appreciation should be given to my Ph.D. advisor Prof. David Gesbert, for his continual inspiration, guidance, and support during my time at EURECOM. It was he who offered me an opportunity to focus on research as a doctoral student, and introduced me to the exciting research field of interference management with delayed channel feedback. Thanks to him, I was in a flexible research environment with some freedom to choose research topics in my last year. I have been also taught how to act professionally to the external industry projects. It is also due to his support so that I could make a three-month academic visit at The University of California, Irvine.

In addition to my advisor, I would like to thank all the committee members for agreeing to serve as the jury of my defense. Thanks to Prof. Eduard Jorswieck and Prof. Abdellatif Zaidi for reading my thesis. Special thanks to Prof. Sheng Yang, who led me to information theory community and showed me the way out when I was stuck in challenging research problems, and to Prof. Bruno Clerckx and Prof. Dirk Slock for having interesting discussion with me. Special thanks are also to Prof. Mari Kobayashi at Supelec for the fruitful collaboration in the first two years of my Ph.D. as well as to Prof. Syed Ali Jafar for hosting me at The University of California at Irvine and having discussion with me about some interesting problems.

I would like to express my thanks to the colleagues and friends at EURECOM, including Haifan Yin, Qianrui Li, Xiwen Jiang, Jingjing Zhang, Dr. Paul de Kerret, Martina Cardone, Rajeev Gangula, Milltiades Filippou, Manijeh Bashar, Konstantinos Alexandris, George Arvanitakis, Sandeep Kottath, and many others, with whom I had the most important and rememberable three years in my life so far. Thanks also to some friends at UC Irvine, including Hua Sun, Chunhua Geng, Feng Jiang, Sundar Rajan

Krishnamurthy, and many others, with whom I spent a great time at Irvine.

Last, but not the least, I would like to thank my wife for her support, patience and encouragement during my Ph.D. study, and my parents for their unconditional love and care throughout my life.

Contents

List of Figures

List of Tables

Acronyms

Here are the main acronyms used in this thesis. The meaning of an acronym is also indicated when it is first used.

AWGN	Additive White Gaussian Noise
BC	Broadcast Channel
CoMP	Coordinated Multiple Points
CSI	Channel State Information
CSIT	Channel State Information at the Transmitter
CSIR	Channel State Information at the Receiver
DoF	Degrees of Freedom
FDMA	Frequency Division Multiple Access
IA	Interference Alignment
IC	Interference Channel
i.i.d.	Independent and Identically Distributed
LTE	Long-Term Evolution
MAC	Multiple Access Channel
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
MU-MIMO	Multiple-User Multiple-Input Multiple-Output
SIMO	Single-Input Multiple-Output
SINR	Signal-to-Interference-plus-Noise-Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise-Ratio
SU-MIMO	Single-User Multiple-Input Multiple-Output
TDMA	Time Division Multiple Access
TIM	Topological Interference Management
TIN	Treating Interference as Noise
ZF	Zero-Forcing

Notations

Common notations are list as below and they will be recalled in each chapter.
Some specific notations will be defined when used.

a, A	Variable
$\mathcal{A}$	Set
$\boldsymbol{a}$	Vector
$\boldsymbol{A}$	Matrix
a^*	Conjugate of a variable a
$\boldsymbol{A}^\mathsf{T}$	Transpose of a matrix $\boldsymbol{A}$
$\boldsymbol{A}^\mathsf{H}$	Hermitian transpose of a matrix $\boldsymbol{A}$
$\bar{\mathcal{A}}$, $\mathcal{A}^c$	The complementary set of $\mathcal{A}$
$\lvert\mathcal{A}\rvert$	The cardinality of the set $\mathcal{A}$
$A_\mathcal{S}$	$\{A_i, i \in \mathcal{S}\}$
$\mathcal{A}_\mathcal{S}$	$\cup_{i\in\mathcal{S}}\mathcal{A}_i$
$\boldsymbol{A}_\mathcal{S}$	The submatrix of $\boldsymbol{A}$ with the rows out of $\mathcal{S}$ removed
$\mathcal{A}\backslash a$	$\{x\|x \in \mathcal{A}, x \neq a\}$
$\mathcal{A}_1\backslash\mathcal{A}_2$	$\{x\|x \in \mathcal{A}_1, x \notin \mathcal{A}_2\}$
$\cap, \cup$	Intersection, union of two sets
$\boldsymbol{A}_{ij}$, $[\boldsymbol{A}]_{ij}$, or $[\boldsymbol{A}]_{i,j}$	The ij-th entry of the matrix $\boldsymbol{A}$
$\boldsymbol{A}_i$ or $[\boldsymbol{A}]_i$	The i-th row of $\boldsymbol{A}$
$\boldsymbol{A}_{[k_1:k_2]}$	The submatrix of $\boldsymbol{A}$ from k_1-th row to k_2-th row when $k_1 \leq k_2$
$\boldsymbol{A}^{-1}$	Inverse of a matrix $\boldsymbol{A}$
$\mathrm{Tr}(\boldsymbol{A})$	Trace of a matrix $\boldsymbol{A}$
$\lVert\boldsymbol{a}\rVert^2$	Norm of a vector $\boldsymbol{a}$
$\lVert\boldsymbol{A}\rVert_\mathrm{F}$	Frobenius norm of a matrix $\boldsymbol{A}$
$\mathrm{rank}(\boldsymbol{A})$	Rank of a matrix $\boldsymbol{A}$
$\det(\boldsymbol{A})$	Determinant of a matrix $\boldsymbol{A}$
$\boldsymbol{a}^\perp$	Normalized orthogonal component of a non-zero vector $\boldsymbol{a}$
$\mathbf{0}_K$	$K \times K$ zero matrix

$\mathbf{0}_{M\times N}$	$M \times N$ zero matrix
$\mathbf{I}_K$	$K \times K$ identity matrix
$\boldsymbol{e}_i$	The i-th column of identity matrix $\mathbf{I}_K$
$f(P) \sim g(P)$	$\lim_{P\to\infty} \frac{f(P)}{g(P)} = C$, where $C > 0$ is a constant that does not scale as P
$\boldsymbol{A} \preceq \boldsymbol{B}$	$\boldsymbol{B} - \boldsymbol{A}$ is positive semidefinite
$(x)^+$,$[x]^+$	$\max\{x, 0\}$
$\mathbb{R}_+^n$	Set of n-tuples of non-negative real numbers
$f = O(g)$	Standard Landau notation, i.e., $\lim \frac{f}{g} \leq C$ where the limit depends on the context
$\mathcal{K}$	$\{1, 2, \ldots, K\}$
$[n]$	$\{1, 2, \ldots, n\}$
$\mathbf{1}(\cdot)$	Indicator function with values 1 when the parameter is true and 0 otherwise
$\min, \max$	Minimum, maximum
$\inf, \sup$	Infimum, supremum
$\mathbb{E}_A B$	Expectation of B over random variable A
$\log(a)$	Logarithm base 2 of a positive real number a
$\mathcal{N}(0, \sigma^2)$ or $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$	Complex circularly symmetric Gaussian distribution with zero mean and variance σ^2
$\mathrm{diag}(\mathcal{A})$	Diagonal matrix with entries being the elements of $\mathcal{A}$

Chapter 1

Résumé [Français]

Les demandes croissantes récent à haut débit de streaming multimédia et des services de réseautage social sur les applications mobiles (tels que les smartphones et les tablettes omniprésents) nécessitent des techniques de débits plus élevés pour les réseaux cellulaires sans fil. La rareté du spectre de fréquences radio dans moins de 10 GHz, qui sont le plus approprié pour les communications sans fil, entrave les progrès. Avec rares spectre, une utilisation plus efficace du spectre devient de plus en plus importante.

Le principal défi pour améliorer l'efficacité spectrale est de compenser les dégradations de signaux qui sont causées par la nature de la propagation sans fil. Le signal transmis dans les réseaux cellulaires sans fil souffre d'une variété de dégradations de canal radio, tels que la propagation perte de trajet, de retarder la propagation, l'étalement Doppler, observation locale, décoloration macroscopique/microscopique, et surtout les interférences [1]. En raison de la propriété de radiodiffusion de support sans fil, des signaux à un utilisateur interfèrent tous les voisins qui fonctionnent à la même fréquence. Pareille ingérence limite la réutilisation des ressources spectrales (temps, fréquence, code, etc.) et est considéré comme l'un des principaux goulets d'étranglement qui limitent le débit global dans les réseaux sans fil.

Dans les réseaux cellulaires, il y a deux principales sources d'interférences: intracellulaire et inter-cellule. L'interférence intracellulaire provient de transmission simultanée de plusieurs utilisateurs qui partagent la même bande de fréquence dans la même cellule. L'interférence inter-cellule au bord de la cellule est causée par la coexistence de non-coopérative ou imparfaitement coopéré plusieurs cellules fonctionnant à la même bande de fréquence. L'existence de l'interférences d'intracellulaire et inter-cellule dégrade la performance globale du réseau. Gestion globalement des interférences est devenue

l'une des principales méthodes pour améliorer l'efficacité spectrale dans les réseaux cellulaires sans fil.

1.1 Gestion des interférences dans les réseaux sans fil

Avec l'évolution des systèmes de communication sans fil, la complexité des méthodes de gestion des interférences est passée d'élimination des interférences (par exemple, la réutilisation des fréquences, l'accès orthogonale), le rejet et l'annulation d'interférence multi-utilisateur multi-cellules [par exemple, de forçage à zéro (ZF) précodage/décodage, multiprogrammation], à la coordination d'interférence et de l'exploitation (par exemple, multipoint coordonnée (CoMP) de transmission, l'alignement d'interférence (IA)).

Stratégies de réutilisation de fréquence sont déployées dans les réseaux cellulaires classiques où la même bande de fréquences est réutilisée par les cellules non-adjacentes seulement. En tant que tel, l'interférence intercellulaire entre deux cellules adjacentes est totalement évitée. Dans chaque cellule, un seul utilisateur est autorisé à accéder au support sans fil par répartition dans le temps ou par répartition en fréquence orthogonale accès (TDMA/FDMA), et ainsi l'interférence intracellulaire est aussi évitée. Néanmoins, ces techniques d'accès orthogonales conduisent généralement à la performance globale du réseau sous-optimale, malgré qu'ils puissent réduire la complexité de la conception du système.

L'utilisation d'antennes multiples permet à multiple-input multiple-output (MIMO) à la transmission, offrant la possibilité d'amplifier la capacité du réseau en exploitant la diversité spatiale et les gains de multiplexage [2]. La clé de transmission MIMO dans le cas où l'utilisateur unique (SU-MIMO) se trouve dans l'espace-temps de traitement du signal, dans lequel à la fois du temps et de dimensions spatiales sont explorées dans l'utilisation de plusieurs antennes réparties dans l'espace [3]. Une famille de techniques espace-temps de traitement du signal des deux côtés émetteur et récepteur ont été développés, comme forçage à zéro beamforming/détection et interférences annulation, par lequel l'interférence est pré-annulé à l'émetteur et/ou rejeté au niveau du récepteur. Communications multi-utilisateurs permettent à plusieurs utilisateurs accèdent au même support sans fil en même temps et la fréquence, agissant comme un rappel physique de la performance de la couche [4]. En raison de la transformation/décodage conjoint à la station de base et la liaison montante multi-utilisateur (généralement modélisé comme de multiples canaux d'accès, MAC) permet la transmission simultanée de

plusieurs utilisateurs sans interférer entre eux. La liaison descendante multi-utilisateur est généralement modélisée comme canaux de diffusion (BC), dans laquelle la station de base envoie différents messages à plusieurs utilisateurs. Les informations conclusions théoriques suggèrent que la stratégie de transmission optimale pour multi-utilisateurs MIMO (MU-MIMO) BC se compose d'une pré-annulation technique d'interférence (dite codage sale papier) combinée avec un ordonnancement d'utilisateur et charge de puissance algorithme explicite [5]. À son tour, certaines techniques de gestion d'interférence pratiques ont été proposées impliquant des concepts tels que linéaire/précodage non linéaire, la sélection de l'utilisateur, et le chargement de puissance de remplissage à l'eau, par lequel le compromis performances et la complexité a été striken. Comparé à SU-MIMO, MU-MIMO est plus à l'abri des limitations sans fil de propagation de canal, comme le déficit canal de rang, le canal de corrélation, et la propagation de ligne de vue.

Quand il se agit de multicellulaires réseaux, la coopération de station de base, aussi connu comme MIMO de réseau et de transmission CoMP, stimule théoriquement la performance du réseau par un traitement commun dans les stations de base avec les messages utilisateurs éventuellement partagés ainsi que des informations d'état de canal [6]. Si les messages des utilisateurs sont partagés par toutes les stations de base via des liens de transport de retour, le réseau cellulaire ensemble forme une grande BC virtuel. Plus précisément, la coordination conjointe entre toutes les stations de base à travers une liaison de raccordement idéal, tous les signaux peuvent être traités conjointement, de sorte que l'interférence inter-cellule est exploitée pour transmettre des messages utiles. Techniques interférences d'annulation et de rejet à deux émetteurs et récepteurs côtés ont été intensivement étudiés dans multiutilisateur MIMO BC, où, par exemple les symboles sont précodés conjointement à la station de base en les faisant coucher dans l'espace nul engendré par les canaux des autres utilisateurs afin de ne pas gêner la utilisateurs non intéressés. En revanche, se il n'y a pas d'échange de messages, le traitement conjoint ne est pas possible et les réseaux cellulaires sont traitées comme des canaux d'interférence (IC). Il est difficile de faire les signaux souhaités se trouvent dans l'espace nul des canaux de tous les autres utilisateurs que le nombre d'utilisateurs augmente, compte tenu d'un nombre limité d'antennes à chaque station de base. Il semble que le réseau est limité par les interférences et la performance globale du réseau sera délimitée par une valeur constante comme le nombre d'émetteurs/récepteurs augmente.

Ce point de vue communément admise a été contestée par le travail [7] en 2008 par Cadambe et Jafar, qui a montré que le réseau ne est pas fondamentalement limité par les interférences et de la performance du taux

de somme de canaux d'interférence peut être réduit que le nombre de paires d'émetteur/récepteur. Plus précisément, dans un canal d'interférence entrée unique sortie unique (SISO), chaque utilisateur peut réaliser la moitié de sa capacité de canal sans interférences, qui est quel que soit le nombre d'utilisateurs. Ce résultat surprenant est dû à l'idée d'alignement d'interférence, par lequel les signaux transmis sont coordonnés par l'intermédiaire de précodage linéaire de telle sorte que les signaux parasites se trouvent dans un sous-espace de dimension réduite et sont séparables de l'une souhaitée au niveau de chaque récepteur. L'alignement d'interférence a attiré beaucoup d'attentions de la théorie de l'information, la communication, et les communautés de traitement du signal au cours des six dernières années [8]. D'autres preuves montrant la force de l'alignement d'interférence ont été trouvés dans les réseaux sans fil d'interférence [9-11], les réseaux de X [12,13], les réseaux cellulaires [14 à 16], et les réseaux multi-flux multi-hop [17, 18], pour ne en nommer que quelques-uns.

Du côté du récepteur, le traitement de interférence comme bruit (TIN) est une technique de gestion d'interférence populaire en raison de sa faible complexité et la robustesse de canaliser l'incertitude. Le régime de TIN a été montré dans [19-21] pour avoir une capacité de somme optimale dans les canaux d'interférence lorsque l'interférence est assez faible. Plus récemment, la condition suffisante de l'optimalité du régime de TIN dans les canaux d'interférence K-utilisateur a été créé en [22]. Remarquablement, si la puissance du signal désiré (en échelle dB) de chaque utilisateur n'est pas inférieure à la somme de la force (en échelle dB) de l'interférence la plus forte depuis et vers cet utilisateur, système de TIN atteint la région des capacités au sein d'un intervalle constant.

1.2 Gestion des interférences avec l'incertitude de canal

1.2.1 L'incertitude de canal

Fondamentalement, les avantages de la plupart des techniques de gestion des interférences viennent généralement de l'hypothèse de la disponibilité des connaissances de canal aux émetteurs. Avec la connaissance de canal, les émetteurs sont en mesure d'allouer adaptative puissance selon la force de canal, d'effectuer une pré-annuler interférences multi-utilisateur selon les instructions de canal, pour planifier les utilis/d'exploiter la diversité multi-utilisateur, et/ou de coordonner les signaux transmis de manière à aligner

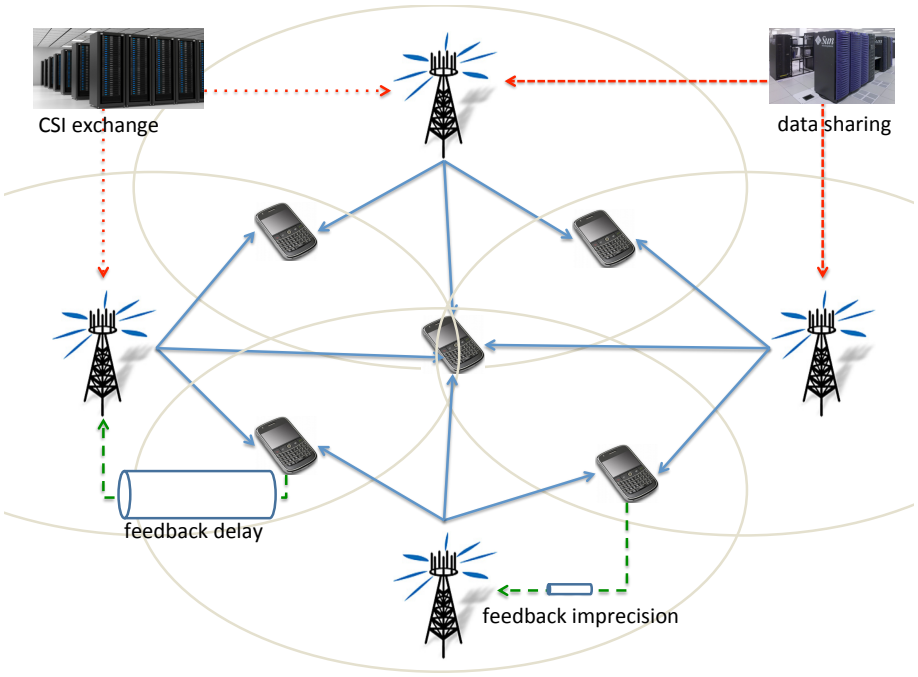


Figure 1.1: réseau cellulaire sans fil avec l'incertitude de canal causé par des retours d'imprécision et de retard.

les interférences au récepteurs. Surtout, la disponibilité des connaissances de canal global à chaque émetteur/récepteur est habituellement exigée par les techniques d'alignement d'interférence et MIMO réseau classiques. Néanmoins, ce gain substantiel sur le cas lorsque l'émetteur n'a pas connaissance de canal est au détriment des énormes ressources de rétroaction sur la liaison montante. En effet, la connaissance de canal à des émetteurs est généralement obtenue grâce à la rétroaction. Les utilisateurs premiers estimation des états de canaux dans la phase de formation, puis nourrissent l'estimation retournée à sa station de base de desserte via les liens capacité de rétroaction limitée. L'inexactitude et/ou la latence de la connaissance de canal (c'est à dire, le canal incertitude) affecte largement les performances du réseau. L'impact de l'incertitude de canal à la performance du système grâce à la rétroaction à taux limité a été intensivement étudiée dans la dernière décennie (voir [23, 24] et les références incluses). Dans les systèmes de communication pratiques, l'incertitude de canal provient principalement de deux contraintes de liens de rétroaction capacité limitée.

- Commentaires imprécision: En raison de la capacité limitée des liens de rétroaction, les coefficients de canal/vecteurs doivent être quantifiées au récepteurs avant d'être renvoyée à des émetteurs selon la capacité de la liaison. L'erreur de quantification est la principale source d'imprécision de la rétroaction.

- Retard de la rétroaction: Si le retour arrive à émetteurs dans le temps de cohérence, puis les rétroactions de canal offre une certaine connaissance liée à l'état actuel du canal. Le niveau de corrélation est directement lié au retard exprimé en fraction du temps de cohérence de canal. Toutefois, si la latence de rétroaction dépasse le temps de cohérence du canal, les rétroactions de canal devient indépendante de l'état de canal actuel.

Sans doute, la rétroaction de canal avec une formation de pilote limité au niveau des récepteurs va encore dégrader la précision de rétroaction et à son tour la performance du réseau. Ce point n'est pas abordé dans cette thèse où la rétroaction de canal parfaite à récepteurs est assumée. Un réseau cellulaire sans fil illustrative est illustré à la figure 1.1, où la rétroaction imprécision et le retard sont deux principales sources d'incertitude de canal, et le partage de données et l'échange CSI sont également des principaux moyens de gestion des interférences à surmonter l'incertitude de canal.

1.2.2 Figures de mérite: Réalisable Taux et degrés de liberté

Parce que nous concentrons principalement sur les performances du réseau dans le régime rapport signal sur bruit (SNR), degrés de liberté (DOF, également connu comme un gain de multiplexage) métrique seront employés dans cette thèse que le principal critère de mérite. Par exception, dans le chapitre 3, nous considérons la performance de SNR fini et prenons taux réalisables comme mesure de performance.

1.2.3 Taux réalisables

Le taux réalisables sur un canal point-à-point peut être généralement calculée, en supposant que la signalisation gaussien sur l'émetteur et le traitement des interférences que le bruit au niveau du récepteur, par

$$R = \log(1 + \mathrm{SINR}) \tag{1.1}$$

où SINR est le rapport entre la puissance du signal désiré et la puissance de la somme des interférences et du bruit.

1.2.4 Degrés de Liberté

Comme la mesure de la performance des chapitres suivants, DoF caractérisation sert de la capacité de premier ordre rapprochement des réseaux sans fil, à partir de laquelle émergent des techniques et des idées de gestion d'interférence nouveaux. Le nombre de DoF représente la pente avec laquelle les taux augmentent avec le logarithme du SNR. A noter que lors de la prise limitations du système supplémentaires en compte comme matériel imparfait, niveaux de modulation finis, et le coût de formation de canal dans un environnement variant dans le temps, le taux de somme sature inévitablement dans la très grande limite de SNR [25]. Cependant, la DoF peut être démontré significative dans un intervalle raisonnable de SNR pratiques pour les systèmes bien conçus, et il se est avéré utile pour comprendre les limites fondamentales de plusieurs protocoles de communication, telles que l'alignement d'interférence (IA) [8] et multicellulaire MIMO [6], entre autres. Mathématiquement, la valeur réalisable DoF est définie comme

$$d = \lim_{P\to\infty} \frac{R}{\log P} \tag{1.2}$$

où P est la puissance d'émission et R est le taux réalisables. Notez que les définitions plus précises du DoF seront données plus tard en cas de besoin.

1.2.5 Network Performance avec l'incertitude de canal

Pour le canal MIMO point à point, des informations d'état de canal (CSI) à l'émetteur (CSIT) est habituellement utilisé pour allouer la puissance à l'émetteur, et donc l'incertitude canal ne est pas important en termes de DoF pour SU-MIMO. Bien que la somme DoF de liaison montante multi-utilisateurs MAC n'est pas affectée par l'absence partielle ou de la CSIT, l'incertitude de canal pourrait sérieusement dégrader les performances de liaison descendante BC et IC dans le sens de DoF. Dans suit, les performances en termes de DoF est décrite à l'égard de diverses disponibilité de rétroaction de canal.

La rétroaction parfaite

Lorsque la connaissance de canal avec une précision infinie et sans latence est disponible à l'émetteur (appelé «parfaite CSIT"), la valeur DoF optimale de K-utilisateur M émission antenne MISO BC est $\min\{M, K\}$ réalisé par des stratégies linéaires comme ZF formation de faisceau. Quand il s'agit de K-utilisateur MIMO BC avec M antennes à l'émetteur et les antennes Nk au

récepteur ke, la somme optimale DoF $\min\{M, \sum_{k=1}^{K} N_k\}$ peut également être obtenue par ZF précodage aux deux émetteurs et récepteurs [26]. Bien que la somme optimale DoF de K-utilisateur canaux d'interférence SISO ont été montrés pour être $\frac{K}{2}$ [7], la caractérisation DoF de l'affaire MIMO générale avec des configurations d'antenne arbitraires est ouvert et encore attirer l'attention [11, 27]. En particulier, pour le mode MIMO à deux IC, où les émetteurs et les récepteurs sont équipés de M_1, M_2, N_1, N_2 antennes, respectivement, la somme est DoF $\min\{M_1 + M_2, N_1 + N_2, \max\{M_1, N_2\}, \max\{M_1, N_2\}\}$ avec une parfaite CSIT [28]. Néanmoins, la performance promise par CSIT parfaite est déprimé par les dures contraintes de précision infinie et la rétroaction latence zéro, ce qui n'est pas pratique.

Pas de rétroaction

D'autre part, en l'absence de connaissance des réalisations de canal à l'émetteur (dénommé "pas CSIT"), un effondrement du DoF a été prédit par l'information des études théoriques. Plus précisément, certains travaux antérieurs avec aucun réglage de la CSIT ont observé cet effondrement du DoF, sous l'hypothèse supplémentaire d'homogénéité, comme i.i.d décoloration isotrope, canal dégradation ou l'équivalence statistique des récepteurs [29-32]. Par exemple, dans un deux utilisateurs MISO BC avec générique (c'est à dire, des coefficients de canal sont tirées d'une distribution continue) et constante ou variable dans le temps canaux, si l'émetteur est aveugle aux Etats canal deux utilisateurs de, la somme la plus connue DoF externe lié est $\frac{4}{3}$ alors la limite intérieure plus connu est toujours 1. Remarquable, un résultat par Jafar [33] étonnamment montré que, sous un modèle de bloc fondu hétérogène, la connaissance des intervalles de cohérence de canal peut améliorer somme DoF. Jetez un deux utilisateurs MISO BC à titre d'exemple, supposons qu'un utilisateur éprouve temps évanouissement sélectif tandis que l'autre éprouve évanouissements sélectifs en fréquence, le DoF optimal $\frac{4}{3}$ peut être atteint. Néanmoins, si la borne extérieure ou intérieure est serrée en général aucun réglage de la CSIT n'est encore inconnu. Il a en outre revendiqué dans [34] que l'effondrement DoF avec imprécision fixe de la CSIT dont l'erreur ne se adaptent pas avec SNR, et la valeur d'un DoF deux utilisateurs MISO BC est supérieure délimitée par $\frac{4}{3}$. Ce DoF limite supérieure a été prouvé à être serré par Gou, Jafar et Wang dans [35], et par Maddah-Ali dans [36] sous le paramètre composé d'état fini, dans lequel DoF sont présentés pour être robuste pour canaliser l'incertitude et l'optimale $\frac{4}{3}$ DoF sont réalisables pour le composé à états finis MISO BC, quel que soit le nombre d'états. D'autre part, il a été conjecturé que la somme d'un DoF deux utilisateurs MISO BC

doit se effondrer à 1 sous le n° réglage général de la CSIT sans hypothèses supplémentaires. Pour prouver ou réfuter cette conjecture a été un problème de longue date ouverte, jusqu'à ce qu'une récente percée par Davoodi et Jafar [37], qui a montré que la limite intérieure est serré tel que DoF effondrement de l'unité que conjecture. Remarquablement, c'est le premier résultat de montrer l'effondrement total de DoF dans l'incertitude de canal.

La rétroaction avec précision finie

Il est bien connu que la pleine DoF en BC peut être maintenu sous CSIT imparfaite si l'erreur dans CSIT diminue à $O(P^{-1})$ as P grandit [38, 39]. L'interprétation est qu'une telle stratégie de quantification maintient le bruit de quantification à un niveau ne dépassant pas le bruit thermique, évitant ainsi de faire quantification un goulot d'étranglement de la transmission que le SNR se développe. En outre, pour le cas de la corrélation dans le temps de telle sorte que le canal émetteur peut prédire l'état actuel de décomposition erreur comme $O(P^{-\alpha})$ pour une constante d'évanouissement $\alpha \in [0, 1]$, ZF ne peut atteindre une fraction optimale de la α DoF par utilisateur [38, 39]. Ce résultat révèle le goulot d'étranglement d'une famille de précodage régimes s'appuyant uniquement sur la précision de rétroaction instantanée CSIT que la corrélation temporelle diminue.

Commentaires avec précision finie a également été considéré dans les canaux d'interférence en liaison avec l'alignement d'interférence. Parmi beaucoup d'autres, dans [40], le nombre de bits de rétroaction nécessaires via diffusion liens de rétroaction a été caractérisée de maintenir le plein DoF avec une parfaite CSIT mondiale. Le nombre minimum de radiodiffusion bits de rétroaction devrait échelle que le nombre d'utilisateurs et $\log P$, ce qui signifie que l'incertitude du canal diminue à $O(P^{-1})$ en tant que P se développe, ce qui concorde avec le cas de la BC.

La rétroaction avec délai

Quand il s'agit du retard de rétroaction, la sagesse conventionnelle suggère de prévoir le canal actuel de la rétroaction retardée en exploitant le temps de canal corrélation. L'état de canal prédit est ensuite utilisé pour précodage comme si c'est l'état du canal vrai. Il fonctionne (mais pas nécessairement optimale) lorsque le temps de cohérence est plus grande que le délai global de rétroaction. Sinon, le retour retardé ne porte aucune information sur le canal vrai courant et le précodage construit sur cette prédiction n'offre aucun gain de multiplexage du tout. Lorsque les évaluations CSI est entièrement

obsolète (c'est à dire, pas corrélés avec le canal vrai actuelle, dénommée pleinement "retardé CSIT"), il semblerait telles informations de canal est non exploitable en vue d'améliorer les gains de multiplexage. Récemment, ce point de vue a été contestée par une information intéressante travail théorique [41], dans lequel Maddah- Ali et Tse ont montré un résultat surprenant que même complètement dépassée CSIT peut être très utile en termes de degrés de liberté, tant qu'il décrit exactement passé réalisations de canal, ce est à dire, l'erreur dans la description des états de canaux dernières doivent se dégrader au moins aussi vite que $O(P^{-1})$. Pour les deux utilisateurs MISO BC, le schéma proposé dans [41] (dénommé «MAT») réalise la DoF de $\frac{2}{3}$ par utilisateur - indépendamment de la corrélation temporelle - la réalisation de meilleures DoF strictement que ce qui est obtenu sans CSIT, même dans des situations extrêmes, lorsque les évaluations de canal retardée est fait totalement obsolète par un retard de rétroaction dépassement du temps de cohérence du canal. Le rôle de la parfaite CSIT retardé peut être réinterprété comme un retour d'informations caractérisant le passé de signal/interférences entendu par les récepteurs. Cette information côté permet au transmetteur de procéder à l'alignement d'interférence «rétrospective» dans le domaine de l'espace et du temps, comme l'a démontré dans différents systèmes de réseau multi-utilisateur [42-47]. Malgré son optimalité DoF, ces régimes MAT inspiration sont conçus en supposant le pire des cas où les évaluations de canal retardée ne fournissent aucune information sur l'état actuel du canal. Le complément d'enquête sur le canal incertitude dans cette voie comprend la soi-disant «alternatif CSIT» [48], dans lequel les utilisateurs rencontrent des temps variables disponibilité de CSI, par exemple, l'incertitude de canal varie selon les réglages parfaits, retardée et ne CSIT, et beaucoup d'autres. Cette rétroaction retardée est la forme de l'incertitude de canal considéré dans les chapitres 3 et 4.

La rétroaction topologique

Dans les réseaux partiellement connectés, la connaissance de la connectivité réseau peut être assez facile à acquérir par les émetteurs via une rétroaction (dénommé «rétroaction topologique»), que le long terme connectivité (statistique) de canal varie souvent plus lent que les réalisations de canal, et réaction de tête de cette information topologique (avec une valeur binaire indiquant si le canal est forte ou faible) est négligeable. Plus précisément, la seule connaissance de la connectivité réseau aux émetteurs via une rétroaction topologique se est avérée bénéfique pour améliorer les performances du réseau dans les réseaux d'interférence partiellement connectés, si le canal est lent

fondu [49] ou évanouissement rapide [50, 51]. Remarquablement, les canaux d'interférence avec aucune CSIT delà de l'information topologique ont été formulés dans le cadre de la gestion des interférences topologique, dans lequel ce problème de rétroaction topologique a été montré pour être équivalent à l'index de codage problème bien défini [52] sous solutions linéaires. Cette rétroaction topologique est la configuration envisagée dans les chapitres 5 et 6.

1.3 Objectifs et méthodologie

1.3.1 Objectifs

En général, cette thèse se concentre sur l'étude de la gestion des interférences avec le canal d'incertitude dans les réseaux multi-MIMO. Comme indiqué précédemment, une caractéristique commune derrière une grande partie de l'analyse des techniques de gestion d'interférence a été la disponibilité de la CSIT parfaite instantanée, avec des exceptions face à soi-disant régimes de rétroaction limitées [23, 24, 34, 39]. En pratique, toutefois, l'acquisition de parfaite (ou même suffisamment précise) CSI aux émetteurs est difficile, si possible, notamment pour les chaînes évanouissement rapide. Les commentaires de canal souffrent de retards ainsi que la capacité strictement limité de liens de rétroaction. Ainsi, dans les cas extrêmes, on peut soit considérer une CSI suffisamment précise mais avec une grande latence ou accéder instantanément à une information de canal grossier (par exemple, un bit indiquant si le canal est forte ou faible), dans laquelle le premier rend les disponibles évaluations CSI totalement obsolètes sous le canal évanouissement rapide, et celui-ci fait l'émetteur presque aveugle, sauf l'indicateur binaire de la force du canal. Cette thèse porte sur différents régimes à l'égard de la disponibilité de la CSIT, en essayant de répondre aux deux problèmes fondamentaux suivants:

- Comment exploiter au mieux la rétroaction retardée?
- Comment exploiter au mieux la rétroaction topologique?

1.3.2 Méthodologie

Pour ces deux problèmes, nous regardons essentiellement sur eux à partir d'un point de vue théorique informations, dont nous tirons l'information limites extérieures théoriques et les systèmes de conception de faisabilité d'approcher les limites extérieures. L'optimalité s'affiche lorsque la faisabilité

coïncide avec les limites extérieures. L'exception est le chapitre 3, dans lequel nous nous concentrons sur l'aspect de traitement du signal telles que la conception précodeur en minimisant l'erreur quadratique moyenne et de maximiser l'information mutuelle.

Les techniques utilisées dans le chapitre 3 comprennent l'optimisation, algorithme itératif, et les équations différentielles de la matrice. Dans le chapitre 4, les limites extérieures sont obtenus en utilisant les techniques de génie-aide de la limite, l'inégalité extrémal, et la capacité ergodique des canaux MIMO avec l'incertitude, alors que la faisabilité est construit sur le codage bloc-Markov et de décodage arrière, ainsi que le concept de roman quantification de l'interférence. Le chapitre 5 présente quelques outils en théorie des graphes et combinatoire, et le problème de la gestion des interférences avec retour topologique sont connecté à certains problèmes bien définis dans la théorie des graphes. Outre les informations techniques de la limite théorique, chapitre 6 se appuie également lien entre les problèmes de gestion des interférences et des problèmes de coloration de graphe ainsi que l'indice de codage problèmes.

1.3.3 Hypothèses

Afin de rendre plus docile problèmes concernés, nous faisons les hypothèses suivantes dans cette thèse.

- **Haute SNR**: En prenant DoF comme la principale mesure de la performance, les problèmes difficiles tels que la conception de précodeur optimale et la répartition de puissance sont sensiblement simplifiées, SNR tend vers l'infini. Comme telle analyse, plus docile peut être fait pour avoir un aperçu sur la façon de concevoir des schémas. Cette hypothèse s'applique aux chapitres 4,5 et 6, tandis que dans le chapitre 3, nous prenons la performance de vitesse finie SNR réalisables en compte.

- **Parfait CSI au niveau des récepteurs (CSIR)**: Tout au long de cette thèse, nous supposons estimation de canal est parfaitement effectuée au niveau des récepteurs au cours de la phase de formation, et CSI est parfaitement connue par les récepteurs.

- **la mobilité de l'utilisateur**: En raison de la mobilité de l'utilisateur, des coefficients de canal varient au fil du temps. Dans les chapitres 3, le canal d'intérêt varie de l'emplacement à la fente, comme rétroaction retard dépasse canal temps de cohérence tels que la rétroaction de canal

est vulnérable à être dépassée. En revanche, le chapitre 4 prend canal corrélation temporelle en compte, où la rétroaction retard est dans le temps de cohérence du canal. Dans les chapitres 5 et 6, les canaux garder constante pendant canal temps de cohérence et varient lorsqu'il dépasse.

- **signalisation Gaussien**: Lorsque nous calculons le taux réalisables (au chapitre 3) et les systèmes de conception de faisabilité (chapitre 4), la signalisation Gaussienne est généralement supposée.
- **filtres linéaires**: Quand nous nous concentrons sur les performances de haute SNR, l'utilisation de filtre linéaire ou non linéaire ne fait aucune différence à DoF métrique. Pour simplifier, nous employons habituellement filtres linéaires à deux côtés émetteur et récepteur.
- **Traiter les interférences que le bruit**: Mis à part le décodage conjoint dans le chapitre 4, nous traitons habituellement interférence bruit à récepteurs.

1.4 Contributions de cette thèse

Cette thèse est composée d'un membre des contributions sur la façon de gérer l'interférence avec la rétroaction retardée ou topologique. Dans ce qui suit, nous présentons toutes les contributions de cette thèse avec des publications.

1.4.1 Gestion des interférences avec rétroaction différé

Lorsque les évaluations CSI est entièrement pas à jour (c'est à dire, retardé CSIT), une percée récente [41] a montré que même les évaluations de canal complètement dépassée est toujours utile. Ce résultat surprenant est basé sur l'idée de l'alignement d'interférence rétrospective (aussi connu comme"l'alignement de MAT") qui permet à la reconstruire de l'émetteur et retransmettre l'interférence entendu au récepteurs dans le passé par rétrospection la CSIT retardé, de sorte que les récepteurs peuvent aligner l'ingérence parfaitement et récupérer souhaité symboles avec succès, ce qui en fait un régime optimal en termes de DoF dans le régime de SNR infinie. Bien inspirant et fascinant d'un point de vue conceptuel, ce résultat est sujet à amélioration. Il peut être considéré comme optimiste en ce qu'il met l'accent sur le comportement intrinsèque de SNR asymptotique, en laissant de côté en particulier la question de savoir comment serait précodage être fait en utilisant pratiquement CSIT vicié au SNR finie. Mais il peut aussi être

considéré qu'il assume le canal est indépendant et identiquement distribuées (i.i.d) dans le temps, où retardée CSIT porte aucune corrélation avec les réalisations de canaux actuels. La thèse étudiera les deux problèmes suivants basés sur un tel cadre de rétroaction retardé:

- Pouvons-nous faire mieux au SNR finie? Comment retardé CSIT peux améliorer la performance de taux de somme?
- Pouvons-nous faire mieux si le canal présente une corrélation dans le temps? Comment ne corrélation temporelle affecte la région DoF du cas général MIMO?

Précodage avec CSIT différé: le cas fini SNR

Dans le chapitre 3, nous formulerons un problème similaire avec le réglage de la CSIT retardé, mais au SNR finie. Nous proposons une première construction pour le précodeur qui correspond aux résultats précédents au infinie SNR encore atteint un bon compromis entre l'alignement d'interférence et l'amélioration du signal au finie SNR, permettant l'amélioration significative des performances dans des contextes pratiques. Nous allons présenter deux méthodes de précodage générales avec nombre arbitraire d'utilisateurs au moyen de MMSE virtuelle et l'optimisation de l'information mutuelle, parvenir à un bon compromis entre l'amélioration du signal et l'alignement d'interférence. En particulier, en optimisant le rapprochement de l'information mutuelle, nous arrivons à une solution de forme fermée pratique, qui offre un remarquable bon compromis entre l'alignement d'interférence et l'amélioration du signal. Ces résultats présentés dans le chapitre 3 ont été publiés dans

- Xinping Yi and David Gesbert, "Precoding methods for the MISO broadcast channel with delayed CSIT," IEEE Transactions on Wireless Communications, vol. 12, no. 5, pp. 2344–2354, May 2013.
- Xinping Yi and David Gesbert, "Precoding on the broadcast MIMO channel with delayed CSIT: The finite SNR case," in Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'12), Kyoto, Japan, March 2012.

Les réseaux MIMO avec CSIT différé: Le cas corrélé Temps

En raison du comportement d'étalement Doppler fini de canaux à évanouissement, c'est le cas dans de nombreuses situations de la vie réelle que les

réalisations de canal dernières peuvent fournir des informations sur les actuels. Avec retard CSIT, le bénéfice de courant imparfaite tels CSIT a été exploitée d'abord dans [53] pour la MISO BC lequel un schéma de transmission roman a été proposé qui améliore l'alignement sur MAT pur dans la construction précodeurs basé sur retardée et estimée CSIT actuelle. La caractérisation complète de la DoF optimale pour cette CSIT mixte a été rapporté plus tard dans [54, 55] pour MISO avant BC sous ce paramètre.

Malheureusement, l'extension au cas MIMO avec des configurations d'antenne arbitraires n'est pas une étape triviale, même avec la courante hypothèse de la qualité de la CSIT symétrique. Les principaux défis sont de deux ordres: (a) la dimension spatiale supplémentaire au côté récepteur introduit un compromis non négligeable entre le signal utile et les interférences mutuelles, et (b) l'asymétrie de recevoir des configurations d'antenne résultats de l'écart de commune message capacité de décodage à différents récepteurs. En particulier, le nombre total de flux qui peuvent être livrés sous forme de messages communs aux deux récepteurs est inévitablement limitée par la faible une (c'est à dire, avec moins d'antennes).

Dans le chapitre 4, nous allons considérer les réseaux MIMO corrélées dans le temps (la BC et IC) où l'émetteur (s) a/ont une) retardé CSI obtenu à partir d'un canal de rétroaction de latence sujettes ainsi que 2) imparfaite actuelle CSIT, obtenu, par exemple, à partir de prédiction sur la base de ces derniers échantillons de canaux selon l'une corrélation temporelle. Les régions DoF pour la diffusion et l'interférence des réseaux MIMO-deux utilisateurs avec des configurations d'antenne générale dans de telles conditions sont entièrement caractérisées, en fonction de l'indicateur de qualité de prédiction. Plus précisément, un cadre unifié simple est proposée, ce qui nous permet d'atteindre la région optimale DoF pour les configurations d'antenne générales et qualités actuelles de la CSIT. Ce cadre se appuie sur le codage bloc-Markov avec une interférence quantification, combinant de manière optimale l'utilisation des deux CSIT jour et instantanée. Une caractéristique frappante de notre travail est que, en faisant varier la répartition du pouvoir, chaque point dans la région DoF peut être réalisé avec un seul système.

Nous établissons aussi des limites extérieures de la région DoF MIMO BC et IC en fonction de l'exposant courant de la qualité de la CSIT. En introduisant un signal virtuel reçu pour l'IC, nous relions le bien externe lié à celui de la BC, arrivé à un résultat consolidé externe similaire pour les deux cas. En plus des techniques de délimitation génie-assistée et l'application de l'inégalité extrémal, nous développons un ensemble de limites supérieure et inférieure de la capacité ergodique pour les canaux MIMO, qui est essentiel pour le cas MIMO mais pas extensible de MISO. Dans la suite, nous proposons

une nouvelle façon systématique pour prouver la faisabilité, au lieu de vérifier la faisabilité de chaque point de la région borne extérieure de coin, comme cela se fait habituellement dans la littérature. Cela contraste avec la plupart des preuves existantes dans la littérature où la faisabilité de chaque point d'angle est cochée. Ces résultats présentés au chapitre 4 ont été publiés dans

- Xinping Yi, Sheng Yang, David Gesbert, and Mari Kobayashi, "The degrees of freedom region of temporally-correlated MIMO networks with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 594-614, January 2014.

- Xinping Yi, David Gesbert, Sheng Yang, and Mari Kobayashi, "Degrees of freedom of time-correlated broadcast channels with delayed CSIT: The MIMO case," in Proc. of *IEEE International Symposium on Information Theory (ISIT'13)*, Istanbul, Turkey, July 2013.

- Xinping Yi, David Gesbert, Sheng Yang, and Mari Kobayashi, "On the DoF of the multiple-antenna time correlated interference channel with delayed CSIT," in Proc. *Asilomar Conference on Signals and Systems (Invited Paper)*, Pacific Grove, CA, USA, November 2012. (long version arXiv: 1204.3046)

Les articles ci-dessus constituent des généralisations de certains de nos travaux précédents portant sur MISO seulement les résultats, qui ne sont pas présentés dans cette thèse.

- Sheng Yang, Mari Kobayashi, David Gesbert, and Xinping Yi, "Degrees of freedom of time correlated MISO broadcast channels with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 315-328, January 2013.

- Sheng Yang, Mari Kobayashi, David Gesbert, and Xinping Yi, "Degrees of Freedom of MISO broadcast channel with perfect delayed and imperfect current CSIT," in Proc. *IEEE Information Theory Workshop (ITW'12)*, Lausanne, Switzerland, September 2012.

- Mari Kobayashi, Sheng Yang, David Gesbert, and Xinping Yi, "On the Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," in Proc. of *IEEE Intern. Symposium on Information Theory (ISIT'12)*, Cambridge, MA, USA, July 2012.

1.4.2 Gestion des interférences avec rétroaction topologique

À première vue, une limitation à un bit évaluations de la CSIT est comme une goutte dans un océan, ce qui rend difficile pour les émetteurs d'extraire le gain DoF substantielle de la coopération, semble inutile dans le sens de DoF. Il a été rapporté dans [29, 30] que DoF substantielle ne peut être réalisé dans l'IC ou le scénario de la BC sans CSIT. Un examen approfondi de ces résultats pessimistes révèle que la plupart des réseaux sont entièrement connecté, dans lequel ne importe quel émetteur peut interférer avec ne importe quel récepteur non conforme dans le réseau.

Grâce à placement aléatoire des nœuds, le fait que le pouvoir se désintègre rapidement avec la distance, l'existence d'obstacles, et les effets d'ombrage locales, nous pouvons soutenons que certains liens d'interférence sont inévitablement beaucoup plus faible que les autres, suggérant l'utilisation d'un graphe partiellement connecté à modéliser, au moins approximativement, la topologie du réseau. Une question intéressante a ensuite soulevé, si la connectivité partielle pourrait être mise à profit pour permettre l'utilisation d'une certaine forme détendue de la CSIT tout en réalisant une performance DoF substantielle. En particulier, l'exploitation de l'information topologique, indiquant simplement que des liens interférents sont suffisamment faibles pour être approchée par zéro interférence et qui les liens sont trop fort, est d'un grand intérêt pratique.

Récemment, réseaux d'ingérence sans CSIT sauf pour le graphe de connectivité réseau ont été étudiés dans le cadre de la gestion dite d'interférence topologique (TIM) [49]. Un effet surprenant a révélé que la performance du réseau peut être sensiblement améliorée sous les seules informations de topologie, à condition que le réseau soit partiellement connecté. Remarquablement, l'alignement d'interférence a été prouvé pour être bénéfique sur les schémas d'accès orthogonaux, même à l'insu de la voie à la réalisation émetteurs.

Bien que le problème TIM soit avérée équivalent au problème de codage de l'indice sous solutions linéaires [49], il y a encore de nombreux problèmes ouverts intéressants dans paramètre TIM. Par exemple, deux problèmes sont énumérés comme suit:

- En l'absence de la CSIT, l'accès orthogonal est optimal dans les réseaux d'interférence entièrement connectés, tandis que l'alignement d'interférence offre plus de gains sous Paramètres TIM. Une question intéressante se pose alors quant à l'accès orthogonale est optimale dans les paramètres TIM.

- Compte tenu de l'avantage de la rétroaction topologique dans les réseaux d'interférence, une question logique est de savoir si les évaluations topologique est bénéfique dans le contexte d'un réseau d'interférence où un mécanisme d'échange de messages entre les émetteurs pré-existe. C'est, quel est l'avantage de la coopération de l'émetteur sous Paramètres TIM?

Topologique gestion des interférences: l'optimalité de l'accès Orthogonal

Dans le chapitre 5, la DoF optimale des multiples problèmes TIM unicast sont entièrement caractérisés par l'intermédiaire de systèmes d'accès orthogonaux simples pour une sous-classe de topologies de réseaux cellulaires. En particulier, il sera montré que l'accès orthogonale atteint le optimale symétrique DoF, résumer DoF, et la région DoF du unidimensionnel (où les deux les émetteurs et les récepteurs sont placés le long d'une ligne droite) réseaux cellulaires convexes [où la couverture cellulaire est convexe telle que chaque émetteur (ou récepteur) se connecte à récepteurs consécutifs (ou émetteurs)] avec des ensembles de messages arbitraires (c'est à dire, le réglage unicast multiples général). Remarquablement, la convexité physique se avère être la limitation fondamentale dans tous les réseaux cellulaires convexes unidimensionnelles. Ces résultats présentés au chapitre 5 ont été soumis pour publication.

- Xinping Yi, Hua Sun, Syed Ali Jafar, and David Gesbert, "Topological interference management: The optimality of orthogonal access," to be submitted, 2014.

Gestion des interférences topologique avec la Coopération émetteur

Dans futurs LTE-A réseaux cellulaires par exemple, un mécanisme transport de retour de routage assure que les stations de base, qui a été sélectionné à coopérer dans le cadre de CoMP, de recevoir une copie des messages à transmettre. Pourtant, l'échange de CSI en temps opportun est difficile en raison de l'obsolescence rapide des instantanée CSI et la latence de signalisation transport de retour liens. Dans ce cas, un canal de diffusion sur les émetteurs distribués s'ensuit avec le manque de CSIT instantanée. Dans le chapitre 6, nous considérons la mise TIM où la coopération de l'émetteur typique est activée. Nous allons montrer que la seule information topologique peut être exploitée dans ce cas d'améliorer strictement DoF

lorsque le réseau ne est pas entièrement connecté. C'est une hypothèse raisonnable dans la pratique. Plusieurs limites internes basés sur la graphique coloration, revêtement de hypergraphe, et l'alignement d'interférence de sous-espace seront proposées, avec quelques limites extérieures construites sur la séquence du générateur, les paramètres composés et codage de l'indice, pour caractériser la DoF symétrique pour les réseaux réguliers ainsi que pour identifier le des conditions pour parvenir à une certaine quantité de DoF pour les topologies de réseau arbitraires. Ces résultats présentés au chapitre 6 ont été publiés dans

- Xinping Yi and David Gesbert, "Topological interference management with transmitter cooperation," in Proc. of *IEEE International Symposium on Information Theory (ISIT'14)*, Hawaii, USA, July 2014.

- Xinping Yi and David Gesbert, "Topological interference management with transmitter cooperation," *IEEE Transactions on Information Theory*, July 2014, under revision.

1.5 Autres contributions

Autres contributions généralisant également certains des résultats ci-dessus au cours de thèse de doctorat ne sont pas présentés dans cette thèse.

Chaînes MISO diffusion avec CSIT différé: Le K-utilisateur de cas

L'extension à un K-utilisateur MISO BC a également été étudiée lorsque l'émetteur a accès à CSI retardé en plus imparfaite CSIT actuelle. Nous avons contribué à la caractérisation de la région de DoF dans un tel cadre en dérivant une borne extérieure et en fournissant un système réalisable. Remarquablement, la DoF réalisables est obtenu par l'élaboration d'un nouveau système d'alignement d'interférence rétrospective, qui se appuie à la fois sur le principe de l'alignement de MAT et ZF précodage pour atteindre un plus grand DoF. Cette contribution a été publiée dans:

- Paul de Kerret, Xinping Yi, and David Gesbert, "On the degrees of freedom of the K-User time-correlated broadcast channel with delayed CSIT", in Proc. of *IEEE International Symposium on Information Theory (ISIT'13)*, Istanbul, Turkey, July 2013. (long version arXiv: 1301.3138)

MIMO réseau avec retards transport de retour

Une application intéressante de la CSIT retardée dans les réseaux cellulaires a également été étudiée. Nous avons considéré le problème de la liaison descendante précodage pour les réseaux MIMO multi-cellulaires où les émetteurs sont fournis avec CSI imparfaite. Plus précisément, chaque émetteur reçoit une estimation de canal retardé avec le retard étant spécifique à chaque composante de canal. Ce modèle est particulièrement adapté pour les scénarios où un utilisateur Feeds Retour son CSI à sa base de desserte seulement comme il est prévu dans les futurs réseaux LTE. Nous avons analysé l'impact du retard lors de l'échange de CSI basée transport de retour sur la performance de taux par MIMO réseau. Nous avons souligné combien retard peut considérablement dégrader les performances du système si les méthodes de précodage existants sont utilisés. Une stratégie de formation de faisceau robuste solution de rechange a été proposée, atteindre la performance maximale, au sens DoF. Cette contribution a été publiée dans:

- Xinping Yi, Paul de Kerret, and David Gesbert, "The DoF of network MIMO with backhaul delays," in Proc. of *IEEE International Conference on Communications (ICC'13)*, Budapest, Hungry, June 2013. (long version, arXiv: 1210.5470.)

Chapter 2

Introduction

The recent growing demands for high-rate multimedia streaming and social networking services on mobile applications (such as the omnipresent smartphones and tablets) require the development of dramatically higher throughput techniques for wireless cellular networks. Radio spectrum scarcity in frequencies below 10 GHz, which are most suitable to wireless communications, impedes the progress in this regard. With scarce spectrum, a more efficient utilization of spectrum becomes more and more crucial.

The main challenge to improve spectral efficiency is to compensate for signal degradations caused by the nature of wireless propagation. The signal transmitted in wireless cellular networks suffers from a variety of radio channel degradations, such as propagation path loss, delay spread, doppler spread, local shadowing, and macroscopic/microscopic fading, and most importantly interference [1]. Due to the broadcasting nature of wireless medium, signals to one user will interfere all neighboring users that are operating at the same frequency. As such, interference restricts the reusability of spectral resources (time, frequency, code, etc.) and is regarded as one of the major bottlenecks limiting the overall throughput in wireless networks.

In cellular networks, there exist two major sources of interference: intracell and intercell interference. The intracell interference comes from simultaneous transmission of multiple users that share the same frequency band in the same cell. The intercell interference at the cell edge is due to the coexistence of noncooperative or imperfectly cooperated multiple cells operating at the same frequency band. The existence of intracell and intercell interference degrades the overall network performance. Interference management overall has become one of the main methods to improve spectral efficiency in wireless cellular networks.

2.1 Interference Management in Wireless Networks

Along with the evolution of wireless communication systems, the complexity of methods for interference management has been growing from interference avoidance (e.g., frequency reuse, orthogonal access), to multiuser multicell interference rejection and cancelation [e.g., zero-forcing (ZF) precoding/decoding, multiuser scheduling], to interference coordination and exploitation (e.g., coordinated multipoint (CoMP) transmission, interference alignment (IA)).

Frequency reuse strategies are deployed in conventional cellular networks where the same frequency band is reused by non-adjacent cells only. As such, the intercell interference between two adjacent cells is totally avoided. In each cell, only one user is allowed to access the wireless medium by time-division or frequency-division orthogonal access (TDMA/FDMA), and thus intracell interference is avoided as well. Nevertheless, these orthogonal access techniques usually lead to suboptimal overall network performance, although they can reduce the complexity of system design.

The use of multiple antennas enables multiple-input multiple-output (MIMO) transmission, offering the opportunity to amplify the network capacity by exploiting spatial diversity and multiplexing gains [2]. The key of MIMO transmission in single user case (SU-MIMO) lies in space-time signal processing, in which both time and spatial dimensions are explored in the use of multiple spatially distributed antennas [3]. A family of space-time signal processing techniques at both transmitter and receiver sides were developed, such as zero forcing beamforming/detection and interference cancelation, by which the interference is precanceled at the transmitter and/or rejected at the receiver.

Multiuser communications enable multiple users access the same wireless medium at the same time and frequency, acting mainly as a physical layer performance booster [4]. Thanks to the joint processing/decoding at the base station, the multiuser uplink (usually modeled as multiple access channels, MAC) allows for the simultaneous transmission of multiple users without interfering one another. The multiuser downlink is usually modeled as broadcast channels (BC), in which the base station sends different messages to multiple users. The information theoretic findings suggest that the optimal transmission strategy for multiuser MIMO (MU-MIMO) BC consists of an interference pre-cancelation technique (so-called dirty-paper coding) combined with an explicit user scheduling and power loading algorithm [5]. In turn, some practical interference management techniques have been proposed involving concepts such as linear/nonlinear precoding, user selection, and

water-filling power loading, by which the performance-complexity tradeoff was striken. Compared to SU-MIMO, MU-MIMO appears more immune to wireless channel propagation limitations, such as channel rank deficiency, channel correlation, and line-of-sight propagation.

When it comes to multicell networks, base station cooperation, also known as network MIMO and CoMP transmission, theoretically boosts the network performance by joint processing at base stations with possibly shared users' messages as well as channel state information [6]. If users' messages are shared by all base stations via backhaul links, the overall cellular network forms a large virtual BC. Specifically, with joint coordination among all base stations through an ideal backhaul link, all the signals can be jointly processed, such that intercell interference is exploited to transmit useful messages. Interference cancelation and rejection techniques at both transmitter and receiver sides were intensively studied in multiuser MIMO BC, where for instance the symbols are jointly precoded at the base station making them lie in the null space spanned by other users' channels so as not to interfere the non-interested users. In contrast, if there is no message sharing, joint processing is not possible and cellular networks are treated as interference channels (IC). It is challenging to make the desired signals lie in the null space of all other users' channels as the number of users increases, given a limited number of antennas at each base station. It would seem that the network is interference-limited and the overall network performance will be bounded by a constant as the number of transmitters/receivers increases.

This commonly accepted viewpoint was challenged by the work [7] in 2008 by Cadambe and Jafar, who showed that the network is not fundamentally interference-limited and the sum rate performance of interference channels can be scaled as the number of transmitter/receiver pairs. Specifically, in a single-input single-output (SISO) interference channel, each user can achieve half of his interference-free channel capacity, which is irrespective of the number of users. This surprising result is due to the idea of interference alignment, by which the transmitted signals are coordinated via linear precoding such that the interfering signals lie in a reduced dimensional subspace and are separable from the desired one at each receiver. Interference alignment has attracted plenty of attentions from the information theory, communication, and signal processing communities over the past six years [8]. Further evidences showing the strength of interference alignment were found in wireless interference networks [9–11], X networks [12, 13], cellular networks [14–16], and multi-hop multi-flow networks [17, 18], to name a few.

At the receiver side, treating interference as noise (TIN) is a popular interference management technique, thanks to its low complexity and robust-

ness to channel uncertainty. The TIN scheme was shown in [19–21] to be sum capacity optimal in the interference channels when the interference is weak enough. Most recently, the sufficient condition of the optimality of TIN scheme in K-user interference channels was established in [22]. Remarkably, if the desired signal strength (in dB scale) of each user is no less than the sum of the strength (in dB scale) of the strongest interference *from* and *to* this user, TIN scheme achieves the capacity region within a constant gap.

2.2 Interference Management with Channel Uncertainty

2.2.1 Channel Uncertainty

Crucially, the benefits of most interference management techniques usually stem from the assumption of the availability of channel knowledge at the transmitters. With channel knowledge, the transmitters are able to adaptively allocate power according to the channel strength, to precancel multiuser interference according to the channel directions, to schedule users to exploit multiuser diversity, and/or to coordinate transmitted signals so as to align interference at receivers. Especially, the availability of global channel knowledge at each transmitter/receiver is usually required by the conventional interference alignment and network MIMO techniques. Nevertheless, such substantial gain over the case when the transmitter totally lacks channel knowledge is at the expense of huge feedback resources on the uplink.

Indeed, the channel knowledge at transmitters is usually obtained through feedback, in which the users first estimate channel states in the training phase and then feed the estimation back to its serving base station via the capacity-limited feedback links. The inaccuracy and/or latency of channel knowledge (i.e., channel uncertainty) affects network performance to a great extent. The impact of channel uncertainty to system performance owing to limited-rate feedback was intensively investigated in the past decade (see [23, 24] and the references therein). In practical communication systems, the channel uncertainty comes mainly from two constraints of capacity-limited feedback links.

- **Feedback imprecision**: Due to the limited capacity of feedback links, the channel coefficients/vectors should be quantized at receivers before being fed back to transmitters according to link capacity. The quantization error is the main source of feedback imprecision.

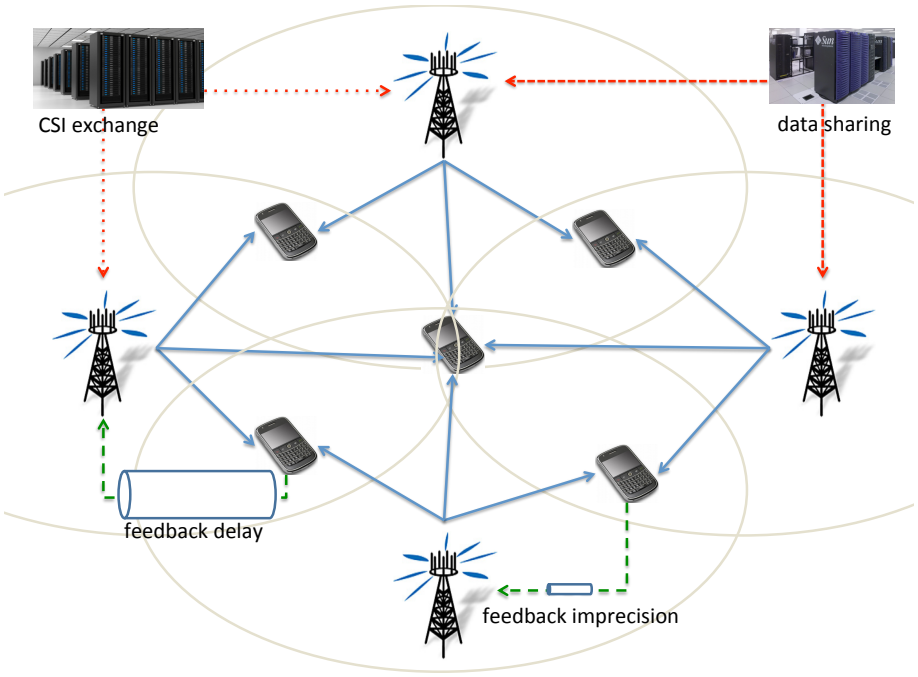


Figure 2.1: Wireless cellular network with channel uncertainty caused by feedback imprecision and delay.

- **Feedback delay**: If the feedback arrives at transmitters within the coherence time, then the channel feedback offers some knowledge, correlated with the current channel state. The correlation level is directly related to the delay expressed as a fraction of the channel coherence time. However, if the feedback latency exceeds the channel coherence time, the channel feedback becomes independent of current channel state.

Admittedly, the channel estimation with limited pilot training at the receivers will further degrade feedback precision and in turn the network performance. This point is not addressed in the thesis, and perfect channel estimation at receivers is assumed.

An illustrative wireless cellular network is shown in Fig. 2.1, where feedback imprecision and delay are two main sources of channel uncertainty, and data sharing and CSI exchange are also major means of interference management to overcome channel uncertainty.

2.2.2 Figures of Merit: Achievable Rate and Degrees of Freedom

Since we mainly focus on network performance in the high signal-to-noise ratio (SNR) regime, degrees of freedom (DoF, also known as multiplexing gain) metric will be employed in this thesis as the main figure of merit. As an exception, in Chapter 3, we consider finite SNR performance and take achievable rate as the performance metric.

Achievable Rate

The achievable rate over a point-to-point channel can be usually computed, assuming Gaussian signaling at the transmitter and treating interference as noise at the receiver, by

$$R = \log(1 + \text{SINR}) \tag{2.1}$$

where SINR is the ratio between the desired signal power and the sum power of interference and noise.

Degrees of Freedom

As the performance metric of later chapters, DoF characterization serves as the first-order capacity approximation for wireless networks, from which novel interference management techniques and insights emerge. The number of DoF represents the slope with which the rate increases with the logarithm of SNR. Note that when taking additional system limitations into account such as imperfect hardware, finite modulation levels, and cost of channel training in a time-varying environment, the sum rate inevitably saturates in the very large SNR limit [25]. However, the DoF can be shown to be meaningful within a reasonable interval of practical SNRs for properly designed systems, and it has proved useful in understanding the fundamental limits of several communication protocols, such as interference alignment (IA) [8] and multi-cell MIMO [6] among many others. Mathematically, the achievable DoF value is defined as

$$d = \lim_{P \to \infty} \frac{R}{\log P} \tag{2.2}$$

where P is the transmit power and R is the achievable rate. Note that the more specific definitions of DoF will be given later when needed.

2.2.3 Network Performance with Channel Uncertainty

For the point-to-point MIMO channel, channel state information (CSI) at transmitter (CSIT) is usually utilized to allocate power at the transmitter, and thus channel uncertainty is not important in terms of DoF for SU-MIMO. While the sum DoF of multiuser uplink MAC is not affected by partial or lack of CSIT, channel uncertainty could severely degrade the performance of downlink BC and IC in the sense of DoF. In what follows, the performance in terms of DoF is described with regard to various channel feedback availability.

Perfect Feedback

When channel knowledge with infinite precision and zero-latency is available at the transmitter (referred to as "perfect CSIT"), the optimal DoF value of K-user M transmit-antenna MISO BC is $\min\{M, K\}$ achieved by linear strategies such as ZF beamforming. When it comes to K-user MIMO BC with M antennas at the transmitter and N_k antennas at kth receiver, the optimal sum DoF $\min\{M, \sum_{k=1}^{K} N_k\}$ can also be achieved by ZF precoding at both transmitters and receivers [26]. While the optimal sum DoF of K-user SISO interference channels were shown to be $\frac{K}{2}$ [7], the DoF characterization of the general MIMO case with arbitrary antenna configurations is open and still attracting more attention [11, 27]. Particularly, for the two-user MIMO IC, where the transmitters and receivers are equipped with M_1, M_2, N_1, N_2 antennas, respectively, the sum DoF is $\min\{M_1+M_2, N_1+N_2, \max\{M_1, N_2\}, \max\{M_1, N_2\}\}$ with perfect CSIT [28]. Nevertheless, the performance promised by perfect CSIT is depressed by the harsh constraints of infinite precision and zero-latency feedback, which is impractical.

No Feedback

On the other hand, in the absence of knowledge of channel realizations at the transmitter (referred to as "no CSIT"), a collapse of DoF was predicted by information theoretic studies. Specifically, some previous works with no CSIT settings have observed this collapse of DoF, under the additional assumption of homogeneity, such as i.i.d. isotropic fading, channel degradedness or statistical equivalence of receivers [29–32]. For instance, in a two-user MISO BC with generic (i.e., channel coefficients are drawn from a continuous distribution) and constant or time-varying channels, if the transmitter is blind to both users' channel states, then the best known sum DoF outer bound is $\frac{4}{3}$ whereas the best known inner bound is still 1. Remarkably,

a surprising result by Jafar [33] showed that, under some heterogeneous block fading model, the sole knowledge of channel coherence intervals can improve sum DoF. For example, for a two-user MISO BC, given that one user experiences time-selective fading whereas the other one experiences frequency-selective fading, the optimal DoF $\frac{4}{3}$ can be achieved. Nevertheless, whether the outer or inner bound is tight in general no CSIT setting is still unknown.

It was further claimed in [34] that DoF collapse with fixed imprecision of CSIT whose error does not scale with SNR, and the DoF value of a two-user MISO BC is upper bounded by $\frac{4}{3}$. This DoF upper bound was proved to be tight by Gou, Jafar and Wang in [35], and by Maddah-Ali in [36] under the finite state compound setting, in which DoF are shown to be robust to channel uncertainty and the optimal $\frac{4}{3}$ DoF are achievable for the finite state compound MISO BC, regardless of the number of states. On the other hand, it was conjectured that the sum DoF of a two-user MISO BC must collapse to 1 under the general no CSIT setting without any additional assumptions. To prove or disprove this conjecture has been a longstanding open problem, until a recent breakthrough by Davoodi and Jafar [37], who showed that the inner bound is tight such that DoF collapse to unity as conjectured. Remarkably, this is the first result to show the total collapse of DoF under channel uncertainty.

Feedback with Finite Precision

It is also well known that the full DoF in BC can be maintained under imperfect CSIT if the error in CSIT decreases as $O(P^{-1})$ as P grows [38, 39]. The interpretation is that such a quantization strategy maintains the quantization noise at a level no greater than the thermal noise, hence avoiding to make quantization a bottleneck of transmission as the SNR grows. Moreover, for the case of the temporally correlated fading channel such that the transmitter can predict the current state with error decaying as $O(P^{-\alpha})$ for some constant $\alpha \in [0, 1]$, ZF can only achieve a fraction α of the optimal DoF per user [38, 39]. This result somehow reveals the bottleneck of a family of precoding schemes relying only on the feedback precision of instantaneous CSIT as the temporal correlation decreases.

Feedback with finite precision was also considered in interference channels in conjunction with interference alignment. Among many others, in [40], the number of required feedback bits via broadcast feedback links was characterized to maintain full DoF with perfect global CSIT. The minimum number of broadcasting feedback bits should scale as the number of users

and $\log P$, meaning that the channel uncertainty decreases as $O(P^{-1})$ as P grows, which agrees with the BC case.

Feedback with Delay

When it comes to the feedback delay, the conventional wisdom suggests to predict the current channel from the delayed feedback by exploiting the channel time correlation. The predicted channel state is then used for precoding as if it is the true channel state. It works (although not necessarily optimal) when the coherence time is larger than the overall feedback delay. Otherwise, the delayed feedback bears no information on the current true channel, and the precoding built upon this prediction offers no multiplexing gain at all.

When the CSI feedback is fully obsolete (i.e., uncorrelated with the current true channel, referred to as fully "delayed CSIT"), it would seem such channel information is non-exploitable in view of improving multiplexing gains. Recently, this viewpoint was challenged by an interesting information theoretic work [41], in which Maddah- Ali and Tse showed a surprising result that even completely outdated CSIT can be very useful in terms of degrees of freedom, as long as it accurately describes past channel realizations, i.e., the error in describing past channel states should decay at least as fast as $O(P^{-1})$. For the two-user MISO BC, the proposed scheme in [41] (referred to as "MAT") achieves the DoF of $\frac{2}{3}$ per user – irrespectively of the temporal correlation – achieving strictly better DoF than what is obtained without any CSIT, even in extreme situations when the delayed channel feedback is made totally obsolete by a feedback delay exceeding the channel coherence time. The role of perfect delayed CSIT can be re-interpreted as a feedback of information characterizing the past signal/interference heard by the receivers. This side information enables the transmitter to perform "retrospective" interference alignment in the space and time domain, as demonstrated in different multiuser network systems [42–47]. Despite its DoF optimality, these MAT-inspired schemes are designed assuming the worst case scenario where the delayed channel feedback provides no information about the current channel state.

The further investigation on channel uncertainty in this avenue includes the so-called "alternating CSIT" [48], in which the users experience time-varying CSI availability, e.g., the channel uncertainty varies among perfect, delayed and no CSIT settings, and many others.

This delayed feedback is the form of channel uncertainty considered in Chapters 3 and 4.

Topological Feedback

In partially connected networks, the knowledge of network connectivity can be fairly easy to be acquired by the transmitters via feedback (referred to as "topological feedback"), as the long term (statistical) channel connectivity often varies slower than channel realizations, and the feedback overhead of this topological information (with binary value indicating whether the channel is strong or weak) is negligible. Specifically, the sole knowledge of network connectivity at the transmitters via topological feedback was shown to be beneficial to improve network performance in partially connected interference networks, whether the channel is slow fading [49] or fast fading [50, 51]. Remarkably, the interference channels with no CSIT beyond topological information were formulated under the topological interference management framework, in which this topological feedback problem was shown to be equivalent to the well-defined index coding problem [52] under linear solutions.

This topological feedback is the setup envisioned in Chapters 5 and 6.

2.3 Objectives and Methodology

2.3.1 Objectives

In general, this thesis focuses on the study of interference management with channel uncertainty in multiuser MIMO networks. As stated earlier, a common feature behind much of the analysis of interference management techniques has been the availability of perfect instantaneous CSIT, with exceptions dealing with so-called limited feedback schemes [23, 24, 34, 39]. In practice, however, the acquisition of perfect (or even sufficiently precise) CSI at the transmitters is difficult, if not impossible, especially for fast fading channels. The channel feedback suffers from *delays* as well as the strictly *limited capacity* of feedback links. As such, in the extreme cases, we may either consider a *sufficiently precise* CSI but with *large latency* or have access *instantaneously* to a *coarse* channel information (e.g., one bit indicating whether the channel is strong or weak), in which the former renders the available CSI feedback fully obsolete under the fast fading channel, and the latter makes the transmitter *almost blind* except the binary indicator of channel strength. This thesis focuses on different regimes with respect of the CSIT availability, trying to address the following two fundamental problems:

- How to best exploit delayed feedback?
- How to best exploit topological feedback?

2.3.2 Methodology

For these two problems, we mainly look at them from an information theoretic perspective, in which we derive information theoretic outer bounds and design achievability schemes to approach the outer bounds. The optimality will be shown when the achievability coincides with the outer bounds. The exception is Chapter 3, in which we focus on the signal processing aspect such as precoder design by minimizing mean square error and maximizing mutual information.

The techniques used in Chapter 3 include optimization, iterative algorithm, and matrix differential equations. In Chapter 4, the outer bounds are derived by utilizing the genie-aid bounding techniques, extremal inequality, and ergodic capacity of MIMO channels with uncertainty, whereas the achievability is built upon block-Markov encoding and backward decoding, as well as the novel concept of interference quantization. Chapter 5 introduces some tools in graph theory and combinatorics, and the interference management problem with topological feedback is connected to some well-defined problems in graph theory. Apart from the information theoretic bounding techniques, Chapter 6 also builds connection between interference management problems and graph coloring problems as well as index coding problems.

2.3.3 Assumptions

In order to make involved problems more tractable, we make the following assumptions in this thesis.

- **High SNR**: By taking DoF as the main performance metric, the challenging problems such as optimal precoder design and power allocation are substantially simplified, as SNR tends to infinity. As such, more tractable analysis can be made to gain insight on how to design schemes. This assumption applies to Chapters 4,5, and 6, while in Chapter 3 we take finite SNR achievable rate performance into account.

- **Perfect CSI at the receivers (CSIR)**: Throughout this thesis, we assume channel estimation is perfectly performed at the receivers during the training phase, and CSI is perfectly known by receivers.

- **User mobility**: Due to user's mobility, channel coefficients vary over time. In Chapters 3, the channel of interest varies from slot to slot, as feedback delay exceeds channel coherence time such that channel feedback is vulnerable to be outdated. In contrast, Chapter 4 takes channel time correlation into account, where feedback delay is within

channel coherence time. In Chapters 5 and 6, the channels keep constant during channel coherence time and vary when exceeds.

- **Gaussian signaling**: When we calculate achievable rate (in Chapter 3) and design achievability schemes (in Chapter 4), Gaussian signaling is usually assumed.

- **Linear filters**: When we focus on high SNR performance, the use of linear or nonlinear filter makes no difference to DoF metric. For simplicity, we usually employ linear filters at both transmitter and receiver sides.

- **Treat interference as noise**: Apart from the joint decoding in Chapter 4, we usually treat interference as noise at receivers.

2.4 Contributions of This Thesis

This thesis is made up of a member of contributions on how to manage interference with delayed or topological feedback. In what follows, we outline all contributions in this thesis and subsequent publications.

2.4.1 Interference Management with Delayed Feedback

When the CSI feedback is fully outdated (i.e., delayed CSIT), a recent breakthrough [41] has shown that even the completely outdated channel feedback is still useful. This surprising result is based on the idea of retrospective interference alignment (also known as "MAT alignment") which allows the transmitter reconstruct and retransmit the overheard interference at receivers in the past by retrospecting the delayed CSIT, so that the receivers can align the interference perfectly and recover desired symbols successfully, making it an optimal scheme in terms of DoF in the infinite SNR regime.

Although inspiring and fascinating from a conceptual point of view, this result is subject to improvement. It can be seen as optimistic in that it intrinsically focuses on the asymptotic SNR behavior, leaving aside in particular the question of how shall precoding be done practically using stale CSIT at finite SNR. But it can also be seen as pessimistic in that it assumes the channel is independent and identically distributed (i.i.d.) across time, where delayed CSIT bears no correlation with current channel realizations.

The thesis in this regard will investigate the following two problems based on such a delayed feedback setting:

- Can we do better at finite SNR? How does delayed CSIT improve sum rate performance?
- Can we do better if the channel exhibits correlation across time? How does time correlation affect the DoF region of the general MIMO case?

Precoding with Delayed CSIT: The Finite SNR Case

In Chapter 3, we will formulate a similar problem under the delayed CSIT setting, yet *at finite SNR*. We propose a first construction for the precoder which matches the previous results at infinite SNR yet reaches a useful trade-off between interference alignment and signal enhancement at finite SNR, allowing for significant performance improvement in practical settings. We will present two general precoding methods with arbitrary number of users by means of virtual MMSE and mutual information optimization, achieving good compromise between signal enhancement and interference alignment. In particular, by optimizing the approximation of mutual information, we arrive at a convenient closed-form solution, which offers a remarkably good compromise between interference alignment and signal enhancement.

These results presented in Chapter 3 were published in

- Xinping Yi and David Gesbert, "Precoding methods for the MISO broadcast channel with delayed CSIT," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2344–2354, May 2013.
- Xinping Yi and David Gesbert, "Precoding on the broadcast MIMO channel with delayed CSIT: The finite SNR case," in Proc. of *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'12)*, Kyoto, Japan, March 2012.

MIMO Networks with Delayed CSIT: The Time Correlated Case

Owing to the finite Doppler spread behavior of fading channels, it is the case in many real life situations that past channel realizations can provide *some* information about current ones. Together with delayed CSIT, the benefit of such imperfect current CSIT was first exploited in [53] for the MISO BC whereby a novel transmission scheme was proposed which improves over pure MAT alignment in constructing precoders based on delayed *and* estimated current CSIT. The full characterization of the optimal DoF for this mixed CSIT was later reported in [54, 55] for MISO BC under this setting.

Unfortunately, the extension to the MIMO case with arbitrary antenna configurations is not a trivial step, even with the symmetric current CSIT

quality assumption. The main challenges are two-fold: (a) the extra spatial dimension at the receiver side introduces a non-trivial tradeoff between useful signal and mutual interference, and (b) the asymmetry of receive antenna configurations results in the discrepancy of common-message-decoding capability at different receivers. In particular, the total number of streams that can be delivered as common messages to both receivers is inevitably limited by the weaker one (i.e., with fewer antennas).

In Chapter 4, we will consider the time-correlated MIMO networks (BC and IC) where the transmitter(s) has/have 1) delayed CSI obtained from a latency-prone feedback channel as well as 2) imperfect current CSIT, obtained, e.g., from prediction on the basis of these past channel samples according to temporal correlation. The DoF regions for the two-user broadcast and interference MIMO networks with general antenna configurations under such conditions are fully characterized, as a function of the prediction quality indicator. Specifically, a simple unified framework is proposed, allowing us to attain optimal DoF region for the general antenna configurations and current CSIT qualities. Such a framework builds upon *block-Markov encoding* with *interference quantization*, optimally combining the use of both outdated and instantaneous CSIT. A striking feature of our work is that, by varying the power allocation, every point in the DoF region can be achieved with one single scheme.

We also establish outer bounds on the DoF region for MIMO BC and IC as a function of the current CSIT quality exponent. By introducing a virtual received signal for the IC, we nicely link the outer bound to that of the BC, arriving at a similar outer bound result for both cases. In addition to the genie-aided bounding techniques and the application of the extremal inequality, we develop a set of upper and lower bounds of ergodic capacity for MIMO channels, which is essential for the MIMO case but not extendible from MISO. In the sequel, we propose a new systematic way to prove the achievability, instead of checking the achievability of every corner point of the outer bound region, as typically done in the literature. This contrasts with most existing proofs in the literature where the achievability of each corner point is checked.

These results presented in Chapter 4 were published in

- Xinping Yi, Sheng Yang, David Gesbert, and Mari Kobayashi, "The degrees of freedom region of temporally-correlated MIMO networks with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 594-614, January 2014.

- Xinping Yi, David Gesbert, Sheng Yang, and Mari Kobayashi, "Degrees

of freedom of time-correlated broadcast channels with delayed CSIT: The MIMO case," in Proc. of *IEEE International Symposium on Information Theory (ISIT'13)*, Istanbul, Turkey, July 2013.

- Xinping Yi, David Gesbert, Sheng Yang, and Mari Kobayashi, "On the DoF of the multiple-antenna time correlated interference channel with delayed CSIT," in Proc. *Asilomar Conference on Signals and Systems (Invited Paper)*, Pacific Grove, CA, USA, November 2012. (long version arXiv: 1204.3046)

The above papers constitute generalizations of some of our previous works addressing MISO only results, which are not presented in this thesis.

- Sheng Yang, Mari Kobayashi, David Gesbert, and Xinping Yi, "Degrees of freedom of time correlated MISO broadcast channels with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 315-328, January 2013.

- Sheng Yang, Mari Kobayashi, David Gesbert, and Xinping Yi, "Degrees of Freedom of MISO broadcast channel with perfect delayed and imperfect current CSIT," in Proc. *IEEE Information Theory Workshop (ITW'12)*, Lausanne, Switzerland, September 2012.

- Mari Kobayashi, Sheng Yang, David Gesbert, and Xinping Yi, "On the Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," in Proc. of *IEEE Intern. Symposium on Information Theory (ISIT'12)*, Cambridge, MA, USA, July 2012.

2.4.2 Interference Management with Topological Feedback

At first glance, a limitation to one-bit feedback is like a drop in a CSIT ocean, making it difficult for the transmitters to extract substantial DoF gain from cooperation, seemingly useless in the sense of DoF. It has been reported in [29, 30] that a substantial DoF cannot be realized in IC or BC scenario without CSIT. A closer examination of these pessimistic results however reveals that many of the considered networks are fully connected, in which any transmitter interferes with any non-intended receiver in the network.

Owing to the nodes' random placement, the fact that power decays fast with distance, the existence of obstacles, and local shadowing effects, we may argue that certain interference links are unavoidably much weaker than others, suggesting the use of a partially-connected graph to model, at least

approximately, the network topology. An interesting question then arises as to whether the partial connectivity could be leveraged to allow the use of some relaxed form of CSIT while still achieving a substantial DoF performance. In particular the exploitation of topological information, simply indicating which of the interfering links are weak enough to be approximated by zero interference and which links are too strong to do so, is of great practical interest.

Recently, interference networks with no CSIT except for the network connectivity graph have been studied under the so-called topological interference management (TIM) framework [49]. A surprising fact revealed that the network performance can be significantly improved under the sole topology information, *provided that the network is partially connected*. Remarkably, interference alignment was proven to be beneficial over orthogonal access schemes, even without the knowledge of channel realization at transmitters.

While the TIM problem is shown to be equivalent to the index coding problem under linear solutions [49], there are still many interesting open problems in TIM setting. For instance, two problems are listed as follows:

- With no CSIT, orthogonal access is optimal in fully connected interference networks, while interference alignment offers more gains under TIM settings. One interesting question then arises as to when orthogonal access is optimal under TIM settings.

- Given the benefit of topological feedback in interference networks, a logical question is whether topological feedback is beneficial in the context of an interference network where a message exchange mechanism between transmitters pre-exists. That is, what is the benefit of transmitter cooperation under TIM settings?

Topological Interference Management: The Optimality of Orthogonal Access

In Chapter 5, the optimal DoF of multiple unicast TIM problems are fully characterized via simple orthogonal access schemes for a subclass of cellular network topologies. In particular, it will be shown that the orthogonal access achieves the optimal symmetric DoF, sum DoF, and DoF region of the one-dimensional (where both the transmitters and the receivers are placed along a straight line) convex cellular networks [where cell coverage is convex such that every transmitter (or receiver) connects to consecutive receivers (or transmitters)] with arbitrary message sets (i.e., the general multiple unicast

setting). Remarkably, the physical convexity turns out to be the fundamental limitation in all one-dimensional convex cellular networks.

These results presented in Chapter 5 were submitted for publication.

- Xinping Yi, Hua Sun, Syed Ali Jafar, and David Gesbert, "Topological interference management: The optimality of orthogonal access in convex cellular networks," to be submitted, 2014.

Topological Interference Management with Transmitter Cooperation

In future LTE-A cellular networks for instance, a backhaul routing mechanism ensures that base stations selected to cooperate under the CoMP framework receive a copy of the messages to be transmitted. Still, the exchange of timely CSI is challenging due to the rapid obsolescence of instantaneous CSI and the latency of backhaul signaling links. In this case, a broadcast channel over distributed transmitters ensues, with lack of instantaneous CSIT.

In Chapter 6, we consider the TIM setting where a typical transmitter cooperation is enabled. We will show that the sole topological information can also be exploited in this case to strictly improve DoF when the network is not fully connected, which is a reasonable assumption in practice. Several inner bounds based on graph coloring, hypergraph covering, and subspace interference alignment will be proposed, together with some outer bounds built upon generator sequence, compound settings and index coding, to characterize the symmetric DoF for the regular networks as well as to identify the conditions to achieve a certain amount of DoF for the arbitrary network topologies.

These results presented in Chapter 6 were published in

- Xinping Yi and David Gesbert, "Topological interference management with transmitter cooperation," in Proc. of *IEEE International Symposium on Information Theory (ISIT'14)*, Hawaii, USA, July 2014.

- Xinping Yi and David Gesbert, "Topological interference management with transmitter cooperation," *IEEE Transactions on Information Theory*, July 2014, under revision.

2.5 Other Contributions

Other contributions also generalizing some of above results during the course of PhD thesis are not presented in this thesis.

MISO Broadcast Channels with Delayed CSIT: The K-user Case

The extension to a K-user MISO BC was also studied when the transmitter has access to delayed CSI in addition to imperfect current CSIT. We contributed to the characterization of DoF region in such a setting by deriving an outer bound and providing an achievable scheme. Remarkably, the achievable DoF is obtained by developing a new retrospective interference alignment scheme, which builds upon both the principle of the MAT alignment and ZF precoding to achieve a larger DoF. This contribution was published in

- Paul de Kerret, Xinping Yi, and David Gesbert, "On the degrees of freedom of the K-User time-correlated broadcast channel with delayed CSIT", in Proc. of *IEEE International Symposium on Information Theory (ISIT'13)*, Istanbul, Turkey, July 2013. (long version arXiv: 1301.3138)

Network MIMO with Backhaul Delays

An interesting application of delayed CSIT in cellular networks was also studied. We considered the problem of downlink precoding for multi-cell MIMO networks where transmitters are provided with imperfect CSI. Specifically, each transmitter receives a delayed channel estimate with the delay being specific to each channel component. This model is particularly adapted to the scenarios where a user feeds back its CSI to its serving base only as it is envisioned in future LTE networks. We analyzed the impact of the delay during the backhaul-based CSI exchange on the rate performance achieved by Network MIMO. We highlighted how delay can dramatically degrade system performance if existing precoding methods are used. An alternative robust beamforming strategy was proposed, achieving the maximal performance, in DoF sense. This contribution was published in

- Xinping Yi, Paul de Kerret, and David Gesbert, "The DoF of network MIMO with backhaul delays," in Proc. of *IEEE International Conference on Communications (ICC'13)*, Budapest, Hungry, June 2013. (long version, arXiv: 1210.5470.)

Part I

Interference Management with Delayed Feedback

Chapter 3

Precoding Methods with Delayed CSIT: The Finite SNR Case

The impact of channel uncertainty caused by feedback delay and how to exploit delayed feedback will be considered in this chapter from a signal processing perspective, with emphasis on sum-rate performance at finite SNR.

Recent results in [41] have shown that precoding in multiuser downlink channel with outdated channel feedback due to feedback delay could lead to data rates much beyond the ones obtained without any CSIT, even in extreme situations when such a channel feedback is made totally obsolete by the feedback delay exceeding the channel coherence time. This surprising result is based on the idea of retrospective interference alignment. Thanks to delayed channel feedback, the transmitter is allowed to retrospect the past channel realizations, reconstruct, and retransmit overheard interference in the past at receivers. As such, the receivers are able to cancel out the overheard interference completely (as if the interferences are aligned perfectly in time domain) and recover desired symbols successfully.

In this chapter, we formulate a similar problem under such a delayed CSIT setting, yet at *finite* SNR. We propose a first construction for the precoder which matches the previous results at infinite SNR yet reaches a useful trade-off between interference alignment and signal enhancement at finite SNR, allowing for significant performance improvements in practical settings.

3.1 Introduction

In multi-user downlink MIMO systems, also known as MIMO BC, the ability to beamform (i.e. linearly precode) multiple data streams simultaneously to several users (up to N) comes nevertheless at a price in terms of requiring the base station transmitter to be informed of the channel coefficients of all served users [4]. In frequency division duplex scenarios (the bulk of available wireless standards today), this implies establishing a feedback link from the mobiles to the base station which can carry CSI related information, in quantized format. A common limitation of such an approach, perceived by many to be a key hurdle toward a more widespread use of multiuser MIMO methods in real-life networks, lies in the fact that the feedback information typically arrives back to the transmitter with a delay which may cause a severe degradation when comparing the obtained feedback CSIT with the actual current CSI. Pushed to the extreme, and considering a feedback delay with the same order of magnitude as the channel coherence time, the available CSIT feedback becomes completely obsolete (i.e., uncorrelated with the current true channel) and, seemingly non-exploitable in view of designing the precoding coefficients.

Recently, this commonly accepted viewpoint was challenged by an interesting information-theoretic work [56] which suggest that precoding on the multi-user MIMO channel with delayed CSIT could substantially increase DoF beyond the ones obtained without any CSIT, even in extreme situations when the delayed channel feedback is made totally obsolete by a feedback delay exceeding the channel coherence time. This surprising result is built upon a novel idea of retrospective interference alignment which allows the transmitter reconstruct and retransmit the overheard interference at receivers in the past by retrospecting the past channel realizations (i.e., outdated channel feedback), so that the receivers can align the interference perfectly, achieving the optimal DoF at the *infinite* SNR.

The premise in [56] is a time-slotted MISO BC with a common transmitter serving multiple users and having a delayed version of CSIT, where the delay causes the CSIT to be fully uncorrelated with the current channel. In this situation, it has been shown that the transmitter can still exploit the stale channel information: The transmitter tries to reproduce the interference generated to the users in the past time slots, while at the same time making sure the forwarded interference occupies a subspace of reduced dimension, compatible with its cancelation at the user's side. Building on such ideas, [56] constructs a transmission protocol which was shown to achieve the maximum DoF for the delayed CSIT broadcast MIMO channel. Precoding

on delayed CSIT MIMO channels has recently attracted more interesting work, dealing with DoF analysis on extended channels, like the X channel and interference channels [42, 43, 57], but also performance analysis including effects of feedback [58] and training [59]. In the example of the two-antenna transmitter, two-user MISO BC, the maximum DoF were shown in [56] to be $\frac{4}{3}$, less than the value of 2 which would be obtained with perfect CSIT, but strictly larger than the single DoF obtained in the absence of any CSIT. This demonstrated that completely obsolete channel feedback is actually useful.

Although fascinating from a conceptual point of view, these results are intrinsically focussed on the asymptotic SNR behavior (i.e., DoF), leaving aside in particular the question of how shall precoding be done practically using stale CSIT at finite SNR (e.g., sum rate). This chapter precisely tackles this question. In what follows we obtain the following key results:

- We show finite SNR precoding using delayed CSIT can be achieved using a combination of retrospective interference alignment together with a signal enhancement strategy.

- We propose a first construction for the precoder generalizing the ideas of [56], namely Generalized MAT (GMAT), which matches the previous results at infinite SNR yet reaches a useful trade-off between interference alignment and signal enhancement at finite SNR, allowing for significant sum-rate performance improvement in practical settings.

- We present two general precoding methods with arbitrary number of users by means of virtual minimum mean square error (MMSE) and mutual information optimization, achieving good compromise between signal enhancement and interference alignment. The precoder coefficients are interpreted as beamforming vector coefficients in an equivalent interference channel scenario, which can be optimized in a number of ways.

Numerical evaluation reveals a substantial performance benefit in terms of data rate in the low to moderate SNR region, but coinciding with the DoF results in [56] when SNR grows to infinity.

The rest of the chapter is organized as follows. In Section 3.2, the channel model of interest is described and the proposed GMAT protocol is detailed first in the two-user case then is generalized to the K-user case. Section 3.3 focuses on the precoder optimization methods based on MMSE and mutual information criteria. Discussion on the multiplexing gain and an interesting interpretation from an equivalent MIMO interference channel is given in

Section 3.4. Numerical examples showing the advantages of the new methods are discussed in Section 3.5. Finally, Section 3.5 concludes the chapter.

Notation: Matrices and vectors are represented as uppercase and lowercase letters, and transpose and conjugate transpose of a matrix are denoted as $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{\mathsf{H}}$, respectively. Further, $\mathrm{Tr}(\cdot)$, $\|\cdot\|$ and $\|\cdot\|_{\mathrm{F}}$ represent respectively the trace of a matrix, the norm of a vector and a Frobenius norm of a matrix. We reserve $[\boldsymbol{A}]_{m,n}$ to denote the element at the m-th row and n-th column of matrix $\boldsymbol{A}$, and $|\mathcal{S}|$ to the cardinality of the set $\mathcal{S}$. Finally, an order-k message denoted by $u_{\mathcal{S}}$ ($|\mathcal{S}| = k$) refers to a linear combination of k distinct symbol vectors intended to k different users in set $\mathcal{S}$.

3.2 System Model

Consider a K-user MISO downlink system with a base station transmitter equipped with K antennas and K single-antenna users, as shown in Fig. 3.1. The received signal in the t-th time slot at j-th user ($j \in \{1, \ldots, K\}$) can be represented as

$$y_j(t) = \sqrt{\frac{P}{K}} \boldsymbol{h}_j^{\mathsf{T}}(t)\boldsymbol{x}(t) + z_j \tag{3.1}$$

where $\boldsymbol{h}_j^{\mathsf{T}}(t) = [h_{j1}(t) \ \ldots \ h_{jK}(t)]$ is the multi-antenna channel vector from the transmitter to j-th user in the t-th time slot, with h_{jk} being the channel coefficient from k-th transmit antenna to j-th user, $\boldsymbol{x}(t)$ is the $K \times 1$ vector of signals sent from the array of K transmit antennas satisfying $\mathbb{E}(\boldsymbol{x}(t)\boldsymbol{x}^{\mathsf{H}}(t)) = \mathbf{I}$, P denotes transmit power, and z_j is the additive Gaussian noise with zero mean and unit variance. A time slotted transmission protocol in the downlink direction is considered.

As in [56], the point made in this chapter is that delayed feedback can be of use to the transmitter including the extreme situation where a feedback delay of one unit of time creates a full decorrelation with the current downlink channel. For this reason, we base ourselves on the framework of so-called delayed CSIT [42, 43, 56–59] by which at time t, it is assumed that user-j has perfect knowledge of $\{\boldsymbol{h}_j(s)\}_{s=1}^{t}$ and of the delayed CSIT of other users $\{\boldsymbol{h}_k(s)\}_{s=1}^{t-1}, k \neq j$, while the transmitter is informed perfectly $\{\boldsymbol{h}_j(s)\}_{s=1}^{t-1}, \forall j$. The accessibility of such delayed CSI at other terminals has been justified in previous work such as [58] by the following model. The users feed back their CSI to the transmitter with delays, then the transmitter broadcasts all the CSI to all the users such that all users have access to

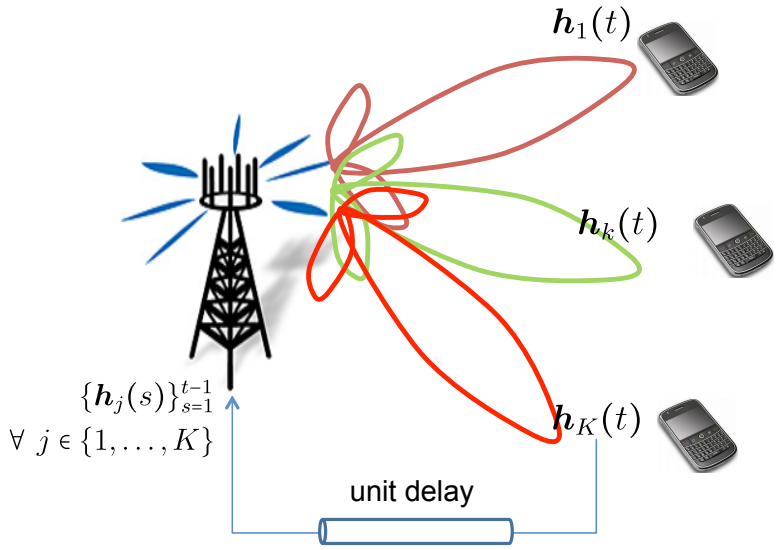


Figure 3.1: Multiuser MISO downlink broadcast channel with delayed feedback.

other users' delayed CSI[1]. Nevertheless, there exists another more efficient scenario for sharing the delayed CSI across users. It is based on the notion of "broadcast uplink feedback", i.e., the terminals broadcast their CSI which is then captured by any overhearing device, which includes both the transmitter and the other terminals. Furthermore, we make no assumption about any correlation between the channel vectors across multiple time slots (could be fully uncorrelated), making it is impossible for the transmitter to use classical MU-MIMO precoding to serve the users, since the transmitter possesses some CSIT possibly independent from the actual channel.

Recently, Maddah-Ali and Tse [56] proposed an algorithm (referred to in this chapter as "MAT" algorithm) under such a delayed CSIT setting obtaining DoF strictly beyond those obtained without any CSIT, even in extreme situations when the delayed CSIT is made totally obsolete. In general, for the K-user case, a K-phase transmission protocol can achieve the maximum DoF $\frac{K}{1+\frac{1}{2}+\cdots+\frac{1}{K}}$. Although such rates are inferior to the ones obtained under the full CSIT setting (cf. K symbols/channel use for K transmit antenna case), they are substantially higher than what was

[1]Clearly, the broadcast phase may introduce some additional delays. The transmitter then exploits the largest delayed version of the CSI, which is common with the one received by the users.

previously reported for the no CSIT case (cf. 1 symbol/channel use regardless of K).

Although optimal in terms of DoF at infinite SNR, we point out that the MAT algorithm can be substantially improved at finite SNR. The key reason is that, at finite SNR, a good scheme will not attempt to use all DoF to eliminate the interference but will try to strike a compromise between interference canceling and enhancing the detectability of the desired signal in the presence of noise. Taking into account this property of basic receivers leads us to revisit the design of the protocol and in particular the design of the precoding coefficients as functions of the knowledge of past channel vectors under the name of GMAT (i.e., generalized MAT algorithm).

First, we proceed by reviewing the proposed protocol in the two-user case, highlighting the connections with the original MAT algorithm. We then generalize the protocol to respectively the three- and K-user cases. In the next section, we then turn to the problem of the optimization of the precoders.

3.2.1 GMAT for the Two-User Case

The transmission of GMAT in the first two time slots is identical to the MAT algorithm, with[2]

$$\boldsymbol{x}(1) = \boldsymbol{s}_A, \quad \boldsymbol{x}(2) = \boldsymbol{s}_B \tag{3.2}$$

where $\boldsymbol{x}(t)$ $(t = 1, 2)$ is the 2×1 signal vector sent from the transmitter at time slot t, $\boldsymbol{s}_A$ and $\boldsymbol{s}_B$ are 2×1 symbol vectors intended to user A and B, respectively, satisfying $\mathbb{E}\{\boldsymbol{s}_i \boldsymbol{s}_i^{\mathsf{H}}\} = \mathbf{I}$. In the third time slot, the transmitter now sends

$$\boldsymbol{x}(3) = \begin{bmatrix} u_{AB} \\ 0 \end{bmatrix} \tag{3.3}$$

where u_{AB} corresponds to an order-2 message (i.e., a combination of two individual user messages) in the following form

$$u_{AB} = \boldsymbol{w}_1^{\mathsf{T}} \boldsymbol{s}_A + \boldsymbol{w}_2^{\mathsf{T}} \boldsymbol{s}_B \tag{3.4}$$

where $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are precoding vectors satisfying the power constraint $\|\boldsymbol{w}_1\|^2 + \|\boldsymbol{w}_2\|^2 \leq 2$ and can be a function of $\boldsymbol{h}_i(1)$ and $\boldsymbol{h}_i(2)$ $(i = A, B)$

[2]For the notational simplicity, we use the index exchangeably, where both A and 1 correspond to the first user/component, and so forth.

according to the delayed CSIT model. Note that this power constraint balances the transmit power used over three time slots. The signal vector received over the three time slots at user A can be rewritten as

$$\bar{\boldsymbol{y}}_A = \sqrt{\frac{P}{2}}\bar{\boldsymbol{H}}_{A1}\boldsymbol{s}_A + \sqrt{\frac{P}{2}}\bar{\boldsymbol{H}}_{A2}\boldsymbol{s}_B + \boldsymbol{z}_A, \tag{3.5}$$

where $\bar{\boldsymbol{y}}_A = [y_A(1)\ y_A(2)\ y_A(3)]^\mathsf{T}$ is the concatenated received signal vector at user A in overall three time slots, $\boldsymbol{z}_A = [z_A(1)\ z_A(2)\ z_A(3)]^\mathsf{T}$ is the Gaussian noise vector, and the effective signal and interference channel matrices are

$$\bar{\boldsymbol{H}}_{A1} = \begin{bmatrix} \boldsymbol{h}_A^\mathsf{T}(1) \\ \boldsymbol{0} \\ h_{A1}(3)\boldsymbol{w}_1^\mathsf{T} \end{bmatrix},\ \bar{\boldsymbol{H}}_{A2} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{h}_A^\mathsf{T}(2) \\ h_{A1}(3)\boldsymbol{w}_2^\mathsf{T} \end{bmatrix}, \tag{3.6}$$

and, by analogy, for user B, we get the interference and signal matrices:

$$\bar{\boldsymbol{H}}_{B1} = \begin{bmatrix} \boldsymbol{h}_B^\mathsf{T}(1) \\ \boldsymbol{0} \\ h_{B1}(3)\boldsymbol{w}_1^\mathsf{T} \end{bmatrix},\ \bar{\boldsymbol{H}}_{B2} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{h}_B^\mathsf{T}(2) \\ h_{B1}(3)\boldsymbol{w}_2^\mathsf{T} \end{bmatrix}. \tag{3.7}$$

A Particular Case (MAT Algorithm)

We point out that the MAT algorithm [56] can be derived as a particular case of the above method, with $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ specified as

$$\boldsymbol{w}_1 = \boldsymbol{h}_B(1), \quad \boldsymbol{w}_2 = \boldsymbol{h}_A(2). \tag{3.8}$$

The key idea behind the original MAT solution in (3.8) is that the interference $\boldsymbol{s}_B$ seen by user A arrives with an effective channel matrix $\bar{\boldsymbol{H}}_{A2}$ which is of rank one, making it possible for user A to combine the three received signals in order to retrieve $\boldsymbol{s}_A$ while canceling out $\boldsymbol{s}_B$ completely. This process is referred to as alignment of interference signal $\boldsymbol{s}_B$, as it mimics the approach taken in interference channels in e.g., [7]. A similar property is exploited in (3.8) at user B as well by making $\bar{\boldsymbol{H}}_{B1}$ be rank 1.

Interpretation of GMAT v.s. MAT

A drawback of the original MAT solution in (3.8) is to optimize the precoders from the point of view of interference alone while the signal matrices $\bar{\boldsymbol{H}}_{A1}$ and $\bar{\boldsymbol{H}}_{B2}$ are ignored. Although this approach is optimal from an information theoretic (i.e., DoF) point of view, it is suboptimal at finite SNR.

In contrast, here, the role of introduced beamformer $\boldsymbol{w}_1$ is to strike a balance between aligning the interference channel of $\boldsymbol{s}_A$ at user B and enhancing the detectability of $\boldsymbol{s}_A$ at user A. In algebraic terms, this can be interpreted as having a compromise between obtaining a rank deficient $\bar{\boldsymbol{H}}_{B1}$ and an orthogonal matrix for $\bar{\boldsymbol{H}}_{A1}$. When it comes to $\boldsymbol{w}_2$, the compromise is between obtaining a rank deficient $\bar{\boldsymbol{H}}_{A2}$ and an orthogonal matrix for $\bar{\boldsymbol{H}}_{B2}$. How to achieve this trade-off in practice is addressed in Section 3.3.

It is also important to note that, there might be alternative fashions of constructing finite SNR precoders based on delayed CSIT. For instance, an interesting question is: Can delayed feedback be exploited already in the second time slot with gains on the finite SNR performance? The intuitive answer to this question is yes. However, the use of precoders in the last time slot only generates a strong symmetry and handling of the users, which in turn allows for closed-form and insightful solutions. This symmetric property is also maintained in the MAT algorithm.

3.2.2 GMAT for the Three-User Case

Similarly to the MAT algorithm, the proposed GMAT sends 18 symbols in a total of three phases, which include 6, 3, and 2 time slots, respectively, offering the effective DoF of $\frac{18}{11}$ symbols/slot. In the first phase, 6 symbol vectors carrying all 18 symbols are sent in 6 consecutive time slots in a way identical to the initial MAT

$$\boldsymbol{x}(1) = \boldsymbol{s}_A^1,\ \boldsymbol{x}(2) = \boldsymbol{s}_B^1,\ \boldsymbol{x}(3) = \boldsymbol{s}_C^1, \tag{3.9}$$

$$\boldsymbol{x}(4) = \boldsymbol{s}_A^2,\ \boldsymbol{x}(5) = \boldsymbol{s}_B^2,\ \boldsymbol{x}(6) = \boldsymbol{s}_C^2 \tag{3.10}$$

where $\boldsymbol{s}_i^1$ and $\boldsymbol{s}_i^2$ $(i = A, B, C)$ are 3×1 symbol vectors (i.e., order-1 messages) intended to user-i. As in the two-user case, we do not introduce channel dependent precoding in the first phase in order to preserve symmetry across the users. Instead, feedback based precoding is introduced in the second phase.

Phase-2 involves 3 time slots, in each of which two order-2 messages (i.e., a linear combination of two order-1 messages) are sent from the first two transmit antennas:

$$\boldsymbol{x}(7) = \begin{bmatrix} u_{AB}^1 \\ u_{AB}^2 \\ 0 \end{bmatrix},\ \boldsymbol{x}(8) = \begin{bmatrix} u_{AC}^1 \\ u_{AC}^2 \\ 0 \end{bmatrix},\ \boldsymbol{x}(9) = \begin{bmatrix} u_{BC}^1 \\ u_{BC}^1 \\ 0 \end{bmatrix} \tag{3.11}$$

where the order-2 messages are constructed by

$$u_{AB}^1 = \boldsymbol{w}_{12}^{1\,\intercal} \boldsymbol{s}_A^1 + \boldsymbol{w}_{21}^{1\,\intercal} \boldsymbol{s}_B^1,\ u_{AB}^2 = \boldsymbol{w}_{12}^{2\,\intercal} \boldsymbol{s}_A^2 + \boldsymbol{w}_{21}^{2\,\intercal} \boldsymbol{s}_B^2 \tag{3.12}$$

$$u_{AC}^1 = \boldsymbol{w}_{13}^{1\top}\boldsymbol{s}_A^1 + \boldsymbol{w}_{31}^{1\top}\boldsymbol{s}_C^1,\ u_{AC}^2 = \boldsymbol{w}_{13}^{2\top}\boldsymbol{s}_A^2 + \boldsymbol{w}_{31}^{2\top}\boldsymbol{s}_C^2 \tag{3.13}$$

$$u_{BC}^1 = \boldsymbol{w}_{23}^{1\top}\boldsymbol{s}_B^1 + \boldsymbol{w}_{32}^{1\top}\boldsymbol{s}_C^1,\ u_{BC}^2 = \boldsymbol{w}_{23}^{2\top}\boldsymbol{s}_B^2 + \boldsymbol{x}_{32}^{2\top}\boldsymbol{s}_C^2 \tag{3.14}$$

where u_{ij}^1 and u_{ij}^2 $(i \neq j)$ are two realizations of the order-2 message dedicated to both user-i and user-j, and $\boldsymbol{w}_{ji}^1 \in \mathbb{C}^{3\times 1}, \boldsymbol{w}_{ji}^2 \in \mathbb{C}^{3\times 1}, 1 \leq i, j \leq 3$ can be arbitrary vector functions of $\boldsymbol{h}_i(t), i = A, B, C,\ t = 1, \cdots, 6$. The responsibility of Phase-2 is to provide independent equations with regard to $\boldsymbol{s}_i^1$ (or $\boldsymbol{s}_i^2$) by utilizing the overheard interferences in the previous phase.

Finally, in the last phase, channel dependent precoding is not introduced as this allows to obtain decoupled optimization problems for each of the $\boldsymbol{w}_{ji}^l$ $(l = 1, 2)$ as will be made in Section 3.3. In this phase, two order-3 messages (i.e., linear combinations of order-2 messages) are sent at the first transmit antenna within two consecutive time slots, i.e.,

$$\boldsymbol{D}(10) = \begin{bmatrix} u_{ABC}^1 \\ 0 \\ 0 \end{bmatrix},\ \boldsymbol{x}(11) = \begin{bmatrix} u_{ABC}^2 \\ 0 \\ 0 \end{bmatrix} \tag{3.15}$$

where u_{ABC}^l $(l = 1, 2)$ is the order-3 message which is identical to the original MAT algorithm

$$\begin{aligned} u_{ABC}^l = a_1^l(h_{C1}(7)u_{AB}^1 + h_{C2}(7)u_{AB}^2) + a_2^l(h_{B1}(8)u_{AC}^1 \\ + h_{B2}(8)u_{AC}^2) + a_3^l(h_{A1}(9)u_{BC}^1 + h_{A2}(9)u_{BC}^2) \end{aligned}$$

where $\{a_j^l\}$ $(j = 1, 2, 3)$ are chosen in a way similar to the original MAT, i.e., arbitrary yet linearly independent sets of coefficients and known by both transmitter and receivers.

Without loss of generality, we treat user A as the target user, and the compact received signal model in matrix format over the 11 time slots can be given by

$$\bar{\boldsymbol{y}}_A = \sqrt{\frac{P}{3}}\sum_{l=1}^{2}\bar{\boldsymbol{H}}_{A1}^l\boldsymbol{s}_A^l + \sqrt{\frac{P}{3}}\sum_{l=1}^{2}\bar{\boldsymbol{H}}_{A2}^l\boldsymbol{s}_B^l + \sqrt{\frac{P}{3}}\sum_{l=1}^{2}\bar{\boldsymbol{H}}_{A3}^l\boldsymbol{s}_C^l + \boldsymbol{z}_A$$

where the equivalent channel matrices can be formulated as

$$\bar{\boldsymbol{H}}_{A1}^l = \begin{bmatrix} \tilde{\boldsymbol{H}}_{A1}^l \\ \boldsymbol{D}_A^l(2)\boldsymbol{W}_1^l(2) \\ \boldsymbol{D}_A^l(3)\boldsymbol{W}_1^l(3) \end{bmatrix}, \bar{\boldsymbol{H}}_{A2}^l = \begin{bmatrix} \tilde{\boldsymbol{H}}_{A2}^l \\ \boldsymbol{D}_A^l(2)\boldsymbol{W}_2^l(2) \\ \boldsymbol{D}_A^l(3)\boldsymbol{W}_2^l(3) \end{bmatrix}, \tag{3.16}$$

$$\bar{\boldsymbol{H}}_{A3}^l = \begin{bmatrix} \tilde{\boldsymbol{H}}_{A3}^l \\ \boldsymbol{D}_A^l(2)\boldsymbol{W}_3^l(2) \\ \boldsymbol{D}_A^l(3)\boldsymbol{W}_3^l(3) \end{bmatrix} \in \mathbb{C}^{11\times 3} \tag{3.17}$$

in which

$$\tilde{\boldsymbol{H}}_{Aj}^{l} = \begin{bmatrix} \mathbf{0}_{m_1^l \times 3} \\ \boldsymbol{h}_A(m_1^l + 1) \\ \mathbf{0}_{n_1^l \times 3} \end{bmatrix} \in \mathbb{C}^{6\times 3} \tag{3.18}$$

with $m_1^l = (3(l-1)+j-1)$, $n_1^l = 6-3(l-1)-j$, $\boldsymbol{D}_A^l(2) = \mathrm{diag}\{h_{Al}(7), h_{Al}(8), h_{Al}(9)\}$, $\boldsymbol{D}_A^l(3) = \mathrm{diag}\{h_{A1}(10), h_{A1}(11)\}$, and

$$\boldsymbol{W}^l(2) = \Bigg[\underbrace{\begin{bmatrix} \boldsymbol{w}_{12}^{l\,\mathsf{T}} \\ \boldsymbol{w}_{13}^{l\,\mathsf{T}} \\ \mathbf{0}_{1\times 3} \end{bmatrix}}_{\boldsymbol{W}_1^l(2)} \quad \underbrace{\begin{bmatrix} \boldsymbol{w}_{21}^{l\,\mathsf{T}} \\ \mathbf{0}_{1\times 3} \\ \boldsymbol{w}_{23}^{l\,\mathsf{T}} \end{bmatrix}}_{\boldsymbol{W}_2^l(2)} \quad \underbrace{\begin{bmatrix} \mathbf{0}_{1\times 3} \\ \boldsymbol{w}_{31}^{l\,\mathsf{T}} \\ \boldsymbol{w}_{32}^{l\,\mathsf{T}} \end{bmatrix}}_{\boldsymbol{W}_3^l(2)} \Bigg] \in \mathbb{C}^{3\times 9} \tag{3.19}$$

is the global precoding matrix (referred to as the order-2 message generation matrix) in which $\boldsymbol{W}_j^l(2)$ is corresponding to user-j.

Given the order-2 message generation matrix $\boldsymbol{W}_j^l(2) \in \mathbb{C}^{3\times 3}$, the precoding matrix for the third phase (referred to as order-3 message generation matrix) can be recursively obtained by

$$\boldsymbol{W}_j^l(3) = \boldsymbol{C}^l(2)\boldsymbol{\Lambda}^l(2)\boldsymbol{W}_j^l(2) \in \mathbb{C}^{2\times 3}, \quad j = 1, 2, 3 \tag{3.20}$$

where $\boldsymbol{\Lambda}^l(2) = \mathrm{diag}\{h_{Cl}(7), h_{Bl}(8), h_{Al}(9)\}$ is set identically to MAT for simplicity, and

$$\boldsymbol{C}^l(2) = \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \end{pmatrix} \tag{3.21}$$

is a constant matrix known by both transmitter and receivers.

A Particular Case (MAT Algorithm)

The original MAT algorithm can be deduced from the proposed method by selecting

$$\boldsymbol{W}^1(2) = \begin{bmatrix} \boldsymbol{h}_B^{\mathsf{T}}(1) & \boldsymbol{h}_A^{\mathsf{T}}(2) & \mathbf{0}_{1\times 3} \\ \boldsymbol{h}_C^{\mathsf{T}}(1) & \mathbf{0}_{1\times 3} & \boldsymbol{h}_A^{\mathsf{T}}(3) \\ \mathbf{0}_{1\times 3} & \boldsymbol{h}_C^{\mathsf{T}}(2) & \boldsymbol{h}_B^{\mathsf{T}}(3) \end{bmatrix} \tag{3.22}$$

and $\boldsymbol{W}^2(2)$ in an analogous way.

Similarly to the two-user case, interferences carrying unintended symbols $\boldsymbol{s}_B^l$ and $\boldsymbol{s}_C^l$ are aligned perfectly at user A, and hence matrices $\bar{\boldsymbol{H}}_{A2}^l$ and

$\bar{\boldsymbol{H}}_{A3}^{l}$ are rank deficient with total rank of 5, making the useful symbol $\boldsymbol{s}_A^l$ retrievable from the left 6-dimensional interference-free subspace. For the proposed GMAT algorithm, we seek to balance signal orthogonality (conditioning of $\bar{\boldsymbol{H}}_{A1}^{l}$) and perfect interference alignment by a careful design of $\boldsymbol{W}^l(2)$.

3.2.3 GMAT for the General K-User Case

In the K-user case, the maximum achievable DoF are $d = \frac{K}{\sum_{k=1}^{K} \frac{1}{k}}$ [56]. Let $d = \frac{K^2 L}{T}$, where T is an integer representing the overall required time slots and L is the number of repeated transmission to guarantee T to be an integer. Without loss of generality, we assume $L = (K-1)!$. The total T times slots can be divided into K phases. In Phase-1, there consist of LK time slots. As the same way to the MAT algorithm, an order-1 message $\boldsymbol{x}(t)$ is sent in t-th time slot, i.e.,

$$\boldsymbol{x}(t) = \boldsymbol{s}_i^l,\ l = 1, \cdots, L \tag{3.23}$$

satisfying $t = L(l-1) + i$, where $\boldsymbol{s}_i^l$ is the $K \times 1$ symbol vector intended to user-i.

From Phase-2 to Phase-K, the transmission of GMAT is similar to MAT algorithm. Phase-k ($2 \le k \le K$) requires $T_k \triangleq \frac{LK}{k}$ time slots, with each time slot transmitting $K-k+1$ order-k messages from $K-k+1$ transmit antennas, i.e.,

$$\boldsymbol{x}(t) = \begin{bmatrix} u_{\mathcal{S}_k}^1 & \cdots & u_{\mathcal{S}_k}^{K-k+1} & 0 & \cdots & 0 \end{bmatrix}^{\mathsf{T}} \tag{3.24}$$

where $u_{\mathcal{S}_k}^j$ ($1 \le j \le K-k+1$) is the j-th message realization of the order-k message that can be generated by

$$\boldsymbol{u}_{\mathcal{S}_k}^l = \boldsymbol{W}^l(k)\boldsymbol{s}^l \tag{3.25}$$

where $\boldsymbol{u}_{\mathcal{S}_k}^l$ is the $Q_k \times 1$ vector ($Q_k \triangleq \binom{K}{k}$) with each element being order-k message that can be interpreted as the combination of any k symbol vectors from $\{\boldsymbol{s}_i^l\}$ ($1 \le l \le L$); $\mathcal{S}_k$ is the set of dedicated users and satisfies $|\mathcal{S}_k| = k$; $\boldsymbol{s}^l = [\boldsymbol{s}_1^{l\,\mathsf{T}} \cdots \boldsymbol{s}_K^{l\,\mathsf{T}}]^{\mathsf{T}} \in \mathcal{C}^{K^2 \times 1}$ is the concatenated symbol vector, and $\boldsymbol{W}^l(k) \in \mathcal{C}^{Q_k \times K^2}$ is the order-k message generation matrix, whose definition is as follows:

Definition 3.1 (**Order-k Message Generation Matrix**). *The order-k message generation matrix $\boldsymbol{W}^l(k) = \begin{bmatrix} \boldsymbol{W}_1^l(k) & \cdots & \boldsymbol{W}_K^l(k) \end{bmatrix}$ ($2 \le k \le K$) is a $Q_k \times K^2$ matrix which satisfies:*

1. *it contains k nonzero and $K-k$ zero blocks in each row, where each block is $1 \times K$ row vector;*
2. *the positions of nonzero blocks of any two rows are not identical; and*
3. *it contains all possibilities of k nonzero positions out of total K positions in each row.*

We point out that the order-k message is desired by those k users whose symbols are contained, and will be overheard by other $K-k$ users as an interference.

Based on the above definition, the signal model of K-user GMAT protocol can be extended as

$$\bar{\boldsymbol{y}}_i = \sqrt{\frac{P}{K}} \sum_{l=1}^{L} \bar{\boldsymbol{H}}_{ii}^{l} \boldsymbol{s}_i^l + \sqrt{\frac{P}{K}} \sum_{l=1}^{L} \sum_{j=1, j\neq i}^{K} \bar{\boldsymbol{H}}_{ij}^{l} \boldsymbol{s}_j^l + \boldsymbol{z}_i \tag{3.26}$$

where

$$\bar{\boldsymbol{H}}_{ij}^{l} = \begin{bmatrix} \tilde{\boldsymbol{H}}_{ij}^{l}(1) \\ \vdots \\ \tilde{\boldsymbol{H}}_{ij}^{l}(k) \\ \vdots \\ \tilde{\boldsymbol{H}}_{ij}^{l}(K) \end{bmatrix} \in \mathbb{C}^{T \times K} \tag{3.27}$$

with $T = \sum_{i=1}^{K} T_k$, is defined as follows:

- The first submatrix corresponds to the effective channel matrix in Phase-1, which can be given by

$$\tilde{\boldsymbol{H}}_{ij}^{l}(1) = \begin{bmatrix} \boldsymbol{0}_{m_1^l \times K} \\ \boldsymbol{h}_i(t) \\ \boldsymbol{0}_{n_1^l \times K} \end{bmatrix} \in \mathbb{C}^{T_1 \times K} \tag{3.28}$$

  where $j = 1, \ldots, K$, $l = 1, \ldots, L$, $m_1^l = (K(l-1) + j - 1)$, $n_1^l = KL - K(l-1) - j$, and $t = m_1^l + 1$;

- The k-th submatrix ($2 \leq k \leq K-1$) which corresponds to Phase-k can be formulated as

$$\tilde{\boldsymbol{H}}_{ij}^{l}(k) = \begin{bmatrix} \boldsymbol{0}_{m_k^l \times K} \\ \boldsymbol{D}_i^l(k) \boldsymbol{W}_j^l(k) \\ \boldsymbol{0}_{n_k^l \times K} \end{bmatrix} \in \mathbb{C}^{T_k \times K} \tag{3.29}$$

where $m_k^l = \left(\lceil \frac{l \cdot l_k}{L} \rceil - 1\right) Q_k$, $n_k^l = T_k - \lceil \frac{l \cdot l_k}{L} \rceil Q_k$ with $l_k = \frac{T_k}{Q_k}$, and $\boldsymbol{D}_i^l(k) = \text{diag}\{h_{is}(t)\} \in \mathbb{C}^{Q_k \times Q_k}$ corresponds to the present channel over which the order-k message is sent in Phase-k with $s = ((l \cdot l_k) \mod L) \mod k$ and t being the index of time slot. In general, $\boldsymbol{W}_j^l(k)$ ($k \geq 2$) is the order-k message generation matrix specified to user-j, which is recursively defined according to

$$\boldsymbol{W}_j^l(k+1) = \boldsymbol{C}^l(k)\boldsymbol{\Lambda}^l(k)\boldsymbol{W}_j^l(k) \tag{3.30}$$

where $\boldsymbol{C}^l(k) \in \mathbb{C}^{Q_{k+1} \times Q_k}$ is a constant matrix known by transmitter and all users, satisfying: (1) each row contains $k+1$ nonzero elements, and (2) the positions of nonzero elements of any two rows are different one another; and $\boldsymbol{\Lambda}^l(k) \in \mathbb{C}^{Q_k \times Q_k}$ is a diagonal matrix whose elements are chosen to be functions of the channel coefficients in Phase-k, so that the interference overheard can be aligned within a limited dimensional subspace. For simplicity, we place emphasis on $\boldsymbol{W}_j^l(k)$, letting $\boldsymbol{\Lambda}^l(k)$ be predetermined as the channel coefficients in Phase-k, as did in the original MAT algorithm.

- The last submatrix is corresponding to the last phase, i.e.,

$$\tilde{\boldsymbol{H}}_{ij}^l(K) = \boldsymbol{D}_i^l(K)\boldsymbol{W}_j^l(K) \in \mathbb{C}^{T_K \times K} \tag{3.31}$$

where $\boldsymbol{W}_j^l(K)$ is defined similarly to (3.30), in which $\boldsymbol{C}^l(K-1) \in \mathbb{C}^{T_K \times Q_{K-1}}$ is a full rank constant matrix without zero elements, and $\boldsymbol{D}_i^l(K) = \text{diag}\{h_{i1}(t)\} \in \mathbb{C}^{T_K \times T_K}$ ($t \in [T - T_K + 1, T]$) contains channel coefficients during Phase-K.

For further illustration, we take the four-user case for example to show its order-2 message generation matrix, i.e.,

$$\boldsymbol{W}^l(2) = \begin{bmatrix} \boldsymbol{w}_{12}^{l\,\top} & \boldsymbol{w}_{21}^{l\,\top} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{w}_{13}^{l\,\top} & \boldsymbol{0} & \boldsymbol{w}_{31}^{l\,\top} & \boldsymbol{0} \\ \boldsymbol{w}_{14}^{l\,\top} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{w}_{41}^{l\,\top} \\ \boldsymbol{0} & \boldsymbol{w}_{23}^{l\,\top} & \boldsymbol{w}_{32}^{l\,\top} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{w}_{24}^{l\,\top} & \boldsymbol{0} & \boldsymbol{w}_{42}^{l\,\top} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{w}_{34}^{l\,\top} & \boldsymbol{w}_{43}^{l\,\top} \end{bmatrix} \tag{3.32}$$

where $\boldsymbol{w}_{ji}^l \in \mathbb{C}^{K \times 1}$ is the beamforming vector aiming at the compromise between user-i and user-j. This formulation collapses to (3.12)-(3.14) for the three-user case and to (3.4) for the two-user case.

A Particular Case (MAT Algorithm)

Particularly for the four-user case, the original MAT algorithm is a specialized GMAT algorithm by setting order-2 message generation matrix as

$$\boldsymbol{W}^1(2) = \begin{bmatrix} \boldsymbol{h}_B^{\mathsf{T}}(1) & \boldsymbol{h}_A^{\mathsf{T}}(2) & \mathbf{0} & \mathbf{0} \\ \boldsymbol{h}_C^{\mathsf{T}}(1) & \mathbf{0} & \boldsymbol{h}_A^{\mathsf{T}}(3) & \mathbf{0} \\ \boldsymbol{h}_D^{\mathsf{T}}(1) & \mathbf{0} & \mathbf{0} & \boldsymbol{h}_A^{\mathsf{T}}(4) \\ \mathbf{0} & \boldsymbol{h}_C^{\mathsf{T}}(2) & \boldsymbol{h}_B^{\mathsf{T}}(3) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{h}_D^{\mathsf{T}}(2) & \mathbf{0} & \boldsymbol{h}_B^{\mathsf{T}}(4) \\ \mathbf{0} & \mathbf{0} & \boldsymbol{h}_D^{\mathsf{T}}(3) & \boldsymbol{h}_C^{\mathsf{T}}(4) \end{bmatrix} \tag{3.33}$$

for $l = 1$ and similarly for other l. For example, for user A, the interference channels $\bar{\boldsymbol{H}}_{Aj}^l$ $(j \neq 1)$ are perfectly aligned, leaving $K = 4$ interference-free dimensions for desired signal, and therefore making the intended symbols retrievable at user A. Similarly for other users, all symbols can be recovered. Hence, 96 symbols are delivered within 50 time slots, yielding the sum DoF of $\frac{48}{25}$.

It is worth noting that the higher order messages can be delivered by the combination of lower order messages. For example, from Phase-k to K, the messages delivered to the receivers aim at completely decoding the order-k message. To avoid too many parameters being optimized which requires huge complexity, we will focus merely on the design of the order-2 message generation matrices $\{\boldsymbol{W}_j^l(2)\}$.

3.3 GMAT Optimization Design

There exist several distinct avenues for computing the delayed CSIT based precoders (i.e., matrices $\{\boldsymbol{W}_j^l(2)\}$). Two of them are briefly described in the following subsections. The first is based on the optimization of a virtual MMSE metric, yielding an iterative solution, while the second one considers the maximization of an approximation of the mutual information, yielding suboptimal yet closed-form solutions. Note that none of these approaches have anything in common with finite SNR interference alignment methods with non-delayed CSIT, such as, e.g., [60,61], since the nature of our problem is fully conditioned by the delayed CSIT scenario. In all cases below, the design of the precoders obeys two principles: (i) the precoders are functions of delayed channel feedback, and (ii) the design is based on the exploitation of alignment-orthogonality trade-off that is underpinned by eq-(3.5), (3.16), and (3.26).

3.3.1 Virtual MMSE Metric

In the following, we describe an approach based on a virtual MMSE metric (referred to later as "GMAT-MMSE") for the two-user case, and subsequently generalize it to the K-user case.

Special $K = 2$ Case

One difficulty in the precoder design lies in the fact that, since the transmitter does not know $\boldsymbol{h}_i(3)$ at Slot-3, the optimization of the precoder in (3.6) and (3.7) cannot involve such information even though the channel realizations on the third time slot clearly affect the overall rate performance. The question therefore is whether a meaningful criterion can be formulated for the optimization of the precoder that IS NOT a function of the non-delayed CSIT. The answer is positive. In what follows, we first offer an intuitive treatment of this problem. Then, a rigorous mathematical argument is offered for it in the next subsection based on mutual information bounds.

In order to derive an optimization model that does no longer depend on the non-delayed CSIT coefficients $h_{A1}(3)$ and $h_{B1}(3)$, we observe that the key trade-off between alignment of interference and desired signal orthogonality is in fact independent from the realizations of $h_{A1}(3)$ and $h_{B1}(3)$, since such coefficients impact on the amplitudes of the precoders but not on their directions. Hence, it is natural to formulate a virtual signal model that skips dependency on the unknown CSIT yet preserves the above mentioned trade-off:

$$\boldsymbol{y}_i = \sqrt{\frac{P}{2}}\boldsymbol{H}_{i1}\boldsymbol{s}_A + \sqrt{\frac{P}{2}}\boldsymbol{H}_{i2}\boldsymbol{s}_B + \boldsymbol{z}_i, i = A, B \tag{3.34}$$

where the virtual channel matrices are now modified from (3.6) and (3.7) according to:

$$\boldsymbol{H}_{i1} = \begin{bmatrix} \boldsymbol{h}_i^{\mathsf{T}}(1) \\ \boldsymbol{0} \\ \boldsymbol{w}_1^{\mathsf{T}} \end{bmatrix}, \ \boldsymbol{H}_{i2} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{h}_i^{\mathsf{T}}(2) \\ \boldsymbol{w}_2^{\mathsf{T}} \end{bmatrix}, \ i = A, B. \tag{3.35}$$

Given $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$, the optimum RX MMSE filter at user-i over this virtual channel is given by

$$\boldsymbol{V}_i = \sqrt{\rho}\left(\rho\boldsymbol{H}_{i1}\boldsymbol{H}_{i1}^{\mathsf{H}} + \rho\boldsymbol{H}_{i2}\boldsymbol{H}_{i2}^{\mathsf{H}} + \mathbf{I}\right)^{-1}\boldsymbol{H}_{ii} \tag{3.36}$$

where $\rho = \frac{P}{K}$ (here $K = 2$), and the corresponding minimal MSE is

$$J_i(\boldsymbol{w}_1, \boldsymbol{w}_2) = \mathrm{Tr}\left(\mathbf{I} - \rho\boldsymbol{H}_{ii}^{\mathsf{H}}(\rho\boldsymbol{H}_{i1}\boldsymbol{H}_{i1}^{\mathsf{H}} + \rho\boldsymbol{H}_{i2}\boldsymbol{H}_{i2}^{\mathsf{H}} + \mathbf{I})^{-1}\boldsymbol{H}_{ii}\right).$$

Note that we exchangeably use A and 1 to represent the first user, and so forth.

Hence, the optimal $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ can be obtained from the following optimization problem, i.e.,

$$\min_{\boldsymbol{w}_1,\boldsymbol{w}_2:\|\boldsymbol{w}_1\|^2+\|\boldsymbol{w}_2\|^2\leq 2} J = J_A(\boldsymbol{w}_1,\boldsymbol{w}_2) + J_B(\boldsymbol{w}_1,\boldsymbol{w}_2).$$

In practice, the gradient based approaches can be used to perform optimization although the convexity of the problem is not guaranteed [62, 63].

General K-User Case

In Phase-k, the transmitter does not know $\boldsymbol{h}_i(t)$ in Slot-t, where $t = \sum_{l=1}^{k-1} T_l + 1, \cdots, \sum_{l=1}^{k} T_l$. Similarly to the two-user case, the virtual received signal can be generalized as $(i = 1, \cdots, K)$

$$\boldsymbol{y}_i = \sqrt{\frac{P}{K}} \sum_{l=1}^{L} \boldsymbol{H}_{ii}^l \boldsymbol{s}_i^l + \sqrt{\frac{P}{K}} \sum_{l=1}^{L} \sum_{j=1,j\neq i}^{K} \boldsymbol{H}_{ij}^l \boldsymbol{s}_j^l + \boldsymbol{z}_i, \tag{3.37}$$

where

$$\boldsymbol{H}_{ij}^l = \left[\tilde{\boldsymbol{H}}_{ij}^{l\,\mathsf{T}} \ \cdots \ \mathbf{0}_{K\times m_k^l} \ \boldsymbol{W}_j^{l\,\mathsf{T}}(k) \ \mathbf{0}_{K\times n_k^l} \ \cdots \ \boldsymbol{W}_j^{l\,\mathsf{T}}(K)\right]^{\mathsf{T}}$$

whose elements are defined in Section 3.2.

Similarly, given $\boldsymbol{W}_j^l(2)$, the optimum MMSE filter for $\boldsymbol{s}_i^l$ at user-i becomes

$$\boldsymbol{V}_i^l = \sqrt{\rho} \left(\rho \sum_{l=1}^{L} \sum_{j=1}^{K} \boldsymbol{H}_{ij}^l \boldsymbol{H}_{ij}^{l\,\mathsf{H}} + \mathbf{I} \right)^{-1} \boldsymbol{H}_{ii}^l \tag{3.38}$$

where $\rho = \frac{P}{K}$ is the normalized transmit power, and the corresponding minimal MSE is

$$\begin{aligned} &J_i^l(\boldsymbol{W}_j^l(2), j = 1, \cdots, K) \\ &= \mathrm{Tr}\left(\mathbf{I} - \rho \boldsymbol{H}_{ii}^{l\,\mathsf{H}} \left(\rho \sum_{l=1}^{L} \sum_{j=1}^{K} \boldsymbol{H}_{ij}^l \boldsymbol{H}_{ij}^{l\,\mathsf{H}} + \mathbf{I} \right)^{-1} \boldsymbol{G}_{ii}^l \right). \end{aligned}$$

The optimal solutions of $\{\boldsymbol{W}_j^l(2), j = 1, \cdots, K\}$ in the sense of virtual MMSE at the receiver side are now given by:

$$\min_{\boldsymbol{W}_j^l(2), j=1,\cdots,K} \quad J = \sum_{l=1}^{L} \sum_{i=1}^{K} J_i^l(\boldsymbol{W}_j^l(2)) \tag{3.39}$$

$$s.t. \quad \sum_{l=1}^{L}\sum_{j=1}^{K}\|\boldsymbol{W}_j^l(2)\|_{\mathrm{F}}^2 \leq KT_2 \tag{3.40}$$

where T_2 was defined in Section 3.2.3.

As the above optimization does not lend itself easily to a closed-form solution, we propose an iterative procedure, based on the gradient descent of the cost function J, where $\boldsymbol{W}_j^l(2)$ is iteratively updated according to

$$\hat{\boldsymbol{W}}_j^l(2)[n+1] = \hat{\boldsymbol{W}}_j^l(2)[n] - \beta\frac{\partial(J)}{\partial \boldsymbol{W}_j^l(2)} \tag{3.41}$$

where n is the iteration index and β is a small step size. The partial derivation is given in the Appendix. Nevertheless, to circumvent non-convexity issue, we explore an alternative optimization method below.

3.3.2 Mutual Information Metric

Here, we propose an approach based on maximizing an approximation of the mutual information, yielding a convenient closed-form solution for $\{\boldsymbol{W}_j^l(2)\}$. In the following, we will start with the two-user case to gain insight, and then generalize it to the K-user case.

Special $K = 2$ Case

Recall that

$$\bar{\boldsymbol{y}}_A = \sqrt{\rho}\bar{\boldsymbol{H}}_{A1}\boldsymbol{s}_A + \sqrt{\rho}\bar{\boldsymbol{H}}_{A2}\boldsymbol{s}_B + \boldsymbol{z}_A \tag{3.42}$$

where $\rho = \frac{P}{K}$ (here $K = 2$), $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are functions of $\boldsymbol{h}_i(j), i = A, B, j = 1, 2$ and satisfy power constraint $\|\boldsymbol{w}_1\|^2 + \|\boldsymbol{w}_2\|^2 \leq 2$. Consequently, the exact mutual information of user A can be calculated by

$$I(\boldsymbol{s}_A; \bar{\boldsymbol{y}}_A) = \log\det\left(\mathbf{I} + \left(\mathbf{I} + \rho\bar{\boldsymbol{H}}_{A2}\bar{\boldsymbol{H}}_{A2}^{\mathsf{H}}\right)^{-1}\rho\bar{\boldsymbol{H}}_{A1}\bar{\boldsymbol{H}}_{A1}^{\mathsf{H}}\right) \tag{3.43}$$

$$= \log\det\left(\mathbf{I} + \rho\begin{bmatrix}1 & 0\\ 0 & \frac{1+\|\boldsymbol{h}_A^{\mathsf{H}}(2)\|^2}{\Delta_1(\boldsymbol{w}_2)}\end{bmatrix}\begin{bmatrix}\|\boldsymbol{h}_A^{\mathsf{H}}(1)\|^2 & h_{A1}^*(3)\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{h}_A(1)\\ h_{A1}(3)\boldsymbol{h}_A^{\mathsf{H}}(1)\boldsymbol{w}_1 & |h_{A1}(3)|^2\|\boldsymbol{w}_1\|^2\end{bmatrix}\right) \tag{3.44}$$

$$= \log\left(1 + \rho\|\boldsymbol{h}_A(1)\|^2 + \frac{\Theta_1(\boldsymbol{w}_1)}{\Delta_1(\boldsymbol{w}_2)}\right) \tag{3.45}$$

where the second line can be easily obtained by permuting rows 2 and 3 in $\bar{\boldsymbol{H}}_{A1}$ and $\bar{\boldsymbol{H}}_{A2}$, and the third line is due to the characteristic polynomial equality [64]

$$\det(\mathbf{I}+\rho\boldsymbol{M}) = 1+\rho\ \mathrm{Tr}(\boldsymbol{M})+\rho^2\det(\boldsymbol{M}), \tag{3.46}$$

where $\boldsymbol{M}$ is a 2×2 Hermitian matrix. By analogy, the mutual information of user B can be given by

$$I(\boldsymbol{s}_B;\bar{\boldsymbol{y}}_B) = \log\left(1+\rho\|\boldsymbol{h}_B(2)\|^2+\frac{\Theta_2(\boldsymbol{w}_2)}{\Delta_2(\boldsymbol{w}_1)}\right) \tag{3.47}$$

where $\Theta_i(\boldsymbol{w}_j)$ and $\Delta_i(\boldsymbol{w}_j)$ are defined as

$$\Theta_1(\boldsymbol{w}_1) = \rho C_2|h_{A1}(3)|^2(C_1\|\boldsymbol{w}_1\|^2-\rho\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{h}_A(1)\boldsymbol{h}_A(1)^{\mathsf{H}}\boldsymbol{w}_1) \tag{3.48}$$

$$\Delta_1(\boldsymbol{w}_2) = C_2(1+\rho|h_{A1}(3)|^2\|\boldsymbol{w}_2\|^2)-\rho^2|h_{A1}(3)|^2\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{h}_A(2)\boldsymbol{h}_A(2)^{\mathsf{H}}\boldsymbol{w}_2 \tag{3.49}$$

$$\Theta_2(\boldsymbol{w}_2) = \rho C_3|h_{B1}(3)|^2(C_4\|\boldsymbol{w}_2\|^2-\rho\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{h}_B(2)\boldsymbol{h}_B(2)^{\mathsf{H}}\boldsymbol{w}_2) \tag{3.50}$$

$$\Delta_2(\boldsymbol{w}_1) = C_3(1+\rho|h_{B1}(3)|^2\|\boldsymbol{w}_1\|^2)-\rho^2|h_{B1}(3)|^2\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{h}_B(1)\boldsymbol{h}_B(1)^{\mathsf{H}}\boldsymbol{w}_1 \tag{3.51}$$

with

$$C_1 = 1+\rho\|\boldsymbol{h}_A(1)\|^2,\quad C_2 = 1+\rho\|\boldsymbol{h}_A(2)\|^2 \tag{3.52}$$

$$C_3 = 1+\rho\|\boldsymbol{h}_B(1)\|^2,\quad C_4 = 1+\rho\|\boldsymbol{h}_B(2)\|^2. \tag{3.53}$$

By imposing a symmetric constraint for power allocation between $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$, e.g., $\|\boldsymbol{w}_1\|^2=\|\boldsymbol{w}_2\|^2=1$ for simplicity, the sum mutual information can be deduced to

$$I(\boldsymbol{s}_A;\bar{\boldsymbol{y}}_A)+I(\boldsymbol{s}_B;\bar{\boldsymbol{y}}_B) \approx \log\left(1+\frac{\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{R}_1\boldsymbol{w}_1}{\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{R}_2\boldsymbol{w}_2}\right) + \log\left(1+\frac{\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{Q}_2\boldsymbol{w}_2}{\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{Q}_1\boldsymbol{w}_1}\right)+\log(C_1C_4) \tag{3.54}$$

where

$$\boldsymbol{R}_1 = C_2\left(\mathbf{I}+\rho\boldsymbol{h}_A^{\perp}(1)\boldsymbol{h}_A^{\perp\mathsf{H}}(1)\right) \tag{3.55}$$

$$\boldsymbol{R}_2 = C_1\left(\gamma_1\mathbf{I}+\rho\boldsymbol{h}_A^{\perp}(2)\boldsymbol{h}_A^{\perp\mathsf{H}}(2)\right) \tag{3.56}$$

$$\boldsymbol{Q}_1 = C_4\left(\gamma_2\mathbf{I}+\rho\boldsymbol{h}_B^{\perp}(1)\boldsymbol{h}_B^{\perp\mathsf{H}}(1)\right) \tag{3.57}$$

$$\boldsymbol{Q}_2 = C_3 \left(\mathbf{I} + \rho \boldsymbol{h}_B^{\perp}(2) \boldsymbol{h}_B^{\perp \mathsf{H}}(2) \right) \tag{3.58}$$

with

$$\gamma_1 = \frac{1 + \rho \|\boldsymbol{h}_A(2)\|^2}{\rho |h_{A1}(3)|^2} + 1, \quad \gamma_2 = \frac{1 + \rho \|\boldsymbol{h}_B(1)\|^2}{\rho |h_{B1}(3)|^2} + 1 \tag{3.59}$$

and $\boldsymbol{h}_i^{\perp}(j) \in \mathbb{C}^{2\times 1}$ is the orthogonal component of $\boldsymbol{h}_i(j)$ $(i = A, B, j = 1, 2)$ satisfying

$$\boldsymbol{h}_i(j) \boldsymbol{h}_i^{\mathsf{H}}(j) + \boldsymbol{h}_i^{\perp}(j) \boldsymbol{h}_i^{\perp \mathsf{H}}(j) = \|\boldsymbol{h}_i(j)\|^2 \mathbf{I}. \tag{3.60}$$

The maximization of the mutual information in closed-form is very challenging in the arbitrary SNR regime. In this chapter, we investigate the possibility of studying the high-enough SNR regime (i.e., high enough to produce tractable solutions) while maintaining an SNR regime that is finite-enough so as to preserve the key notion of alignment-orthogonality trade-off exposed earlier in Section 3.2. Thus, we further approximate the mutual information as

$$I(\boldsymbol{s}_A; \bar{\boldsymbol{y}}_A) + I(\boldsymbol{s}_B; \bar{\boldsymbol{y}}_B) \approx \log \left(\frac{\boldsymbol{w}_1^{\mathsf{H}} \boldsymbol{R}_1 \boldsymbol{w}_1}{\boldsymbol{w}_2^{\mathsf{H}} \boldsymbol{R}_2 \boldsymbol{w}_2} \frac{\boldsymbol{w}_2^{\mathsf{H}} \boldsymbol{Q}_2 \boldsymbol{w}_2}{\boldsymbol{w}_1^{\mathsf{H}} \boldsymbol{Q}_1 \boldsymbol{w}_1} \right) + \log(C_1 C_4) \tag{3.61}$$

which can be optimized by separately maximizing two Rayleigh Quotients, i.e.,

$$\max_{\|\boldsymbol{w}_1\|^2=1} \log \left(\frac{\boldsymbol{w}_1^{\mathsf{H}} \boldsymbol{R}_1 \boldsymbol{w}_1}{\boldsymbol{w}_1^{\mathsf{H}} \boldsymbol{Q}_1 \boldsymbol{w}_1} \right) = \max_{\|\boldsymbol{w}_1\|^2=1} \frac{\boldsymbol{w}_1^{\mathsf{H}} \left(\mathbf{I} + \rho \boldsymbol{h}_A^{\perp}(1) \boldsymbol{h}_A^{\perp \mathsf{H}}(1) \right) \boldsymbol{w}_1}{\boldsymbol{w}_1^{\mathsf{H}} \left(\gamma_2 \mathbf{I} + \rho \boldsymbol{h}_B^{\perp}(1) \boldsymbol{h}_B^{\perp \mathsf{H}}(1) \right) \boldsymbol{w}_1} \tag{3.62}$$

$$\max_{\|\boldsymbol{w}_2\|^2=1} \log \left(\frac{\boldsymbol{w}_2^{\mathsf{H}} \boldsymbol{Q}_2 \boldsymbol{w}_2}{\boldsymbol{w}_2^{\mathsf{H}} \boldsymbol{R}_2 \boldsymbol{w}_2} \right) = \max_{\|\boldsymbol{w}_2\|^2=1} \frac{\boldsymbol{w}_2^{\mathsf{H}} \left(\mathbf{I} + \rho \boldsymbol{h}_B^{\perp}(2) \boldsymbol{h}_B^{\perp \mathsf{H}}(2) \right) \boldsymbol{w}_2}{\boldsymbol{w}_2^{\mathsf{H}} \left(\gamma_1 \mathbf{I} + \rho \boldsymbol{h}_A^{\perp}(2) \boldsymbol{h}_A^{\perp \mathsf{H}}(2) \right) \boldsymbol{w}_2}. \tag{3.63}$$

Hence, we can obtain the optimal solutions $\boldsymbol{w}_1^{opt}$ and $\boldsymbol{w}_2^{opt}$, which are given by the dominant generalized eigenvectors of the pairs $(\boldsymbol{R}_1, \boldsymbol{Q}_1)$ and $(\boldsymbol{Q}_2, \boldsymbol{R}_2)$, respectively.

Nevertheless, these solutions can be found to be dependent on parameters γ_1 and γ_2 which in turn depend on the unknown channel coefficients in Slot-3. Fortunately, it is possible to average their impact and obtain a lower bound on mutual information that no longer depends on such coefficients. Aware of the convexity of mutual information approximation in eq-(3.61) with regard

to $|h_{A1}(3)|^2$ and $|h_{B1}(3)|^2$, we further lower-bound it by applying Jensen's inequality, i.e.,

$$\mathbb{E}_{|h_{B1}(3)|^2} \log\left(\frac{\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{R}_1\boldsymbol{w}_1}{\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{Q}_1\boldsymbol{w}_1}\right) \geq \log\left(\frac{\boldsymbol{w}_1^{\mathsf{H}}\left(\mathbf{I}+\rho\boldsymbol{h}_A^{\perp}(1)\boldsymbol{h}_A^{\perp\mathsf{H}}(1)\right)\boldsymbol{w}_1}{\boldsymbol{w}_1^{\mathsf{H}}\left(\bar{\gamma}_2\mathbf{I}+\rho\boldsymbol{h}_B^{\perp}(1)\boldsymbol{h}_B^{\perp\mathsf{H}}(1)\right)\boldsymbol{w}_1}\right)+\log\left(\frac{C_2}{C_4}\right)$$

$$\mathbb{E}_{|h_{A1}(3)|^2} \log\left(\frac{\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{Q}_2\boldsymbol{w}_2}{\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{R}_2\boldsymbol{w}_2}\right) \geq \log\left(\frac{\boldsymbol{w}_2^{\mathsf{H}}\left(\mathbf{I}+\rho\boldsymbol{h}_B^{\perp}(2)\boldsymbol{h}_B^{\perp\mathsf{H}}(2)\right)\boldsymbol{w}_2}{\boldsymbol{w}_2^{\mathsf{H}}\left(\bar{\gamma}_1\mathbf{I}+\rho\boldsymbol{h}_A^{\perp}(2)\boldsymbol{h}_A^{\perp\mathsf{H}}(2)\right)\boldsymbol{w}_2}\right)+\log\left(\frac{C_3}{C_1}\right)$$

with

$$\bar{\gamma}_1 = 1+\|\boldsymbol{h}_A(2)\|^2+1/\rho, \quad \bar{\gamma}_2 = 1+\|\boldsymbol{h}_B(1)\|^2+1/\rho$$

being independent of the unknown channel coefficients $h_{A1}(3)$ and $h_{B1}(3)$ where $\mathbb{E}[|h_{A1}(3)|^2] = \mathbb{E}[|h_{B1}(3)|^2] = 1$, such that the original optimization problem can be alternatively done by

$$\max_{\|\boldsymbol{w}_i\|^2=1} \frac{\boldsymbol{w}_i^{\mathsf{H}}\left(\mathbf{I}+\rho\boldsymbol{h}_{\bar{i}}^{\perp}(i)\boldsymbol{h}_{\bar{i}}^{\perp\mathsf{H}}(i)\right)\boldsymbol{w}_i}{\boldsymbol{w}_i^{\mathsf{H}}\left(\bar{\gamma}_{\bar{i}}\mathbf{I}+\rho\boldsymbol{h}_{\bar{i}}^{\perp}(i)\boldsymbol{h}_{\bar{i}}^{\perp\mathsf{H}}(i)\right)\boldsymbol{w}_i} \tag{3.64}$$

where $i, \bar{i} = 1, 2$ and $i \neq \bar{i}$. As mentioned earlier, we exchangeably use A and 1 to represent the first user, and so forth.

Interestingly, the above objective function can be interpreted as dual SINR in a two-user interference channel. Define

$$\mathrm{DSINR}_i \triangleq \frac{\boldsymbol{w}_i^{\mathsf{H}}\left(\mathbf{I}+\rho\boldsymbol{h}_{\bar{i}}^{\perp}(i)\boldsymbol{h}_{\bar{i}}^{\perp\mathsf{H}}(i)\right)\boldsymbol{w}_i}{\boldsymbol{w}_i^{\mathsf{H}}\left(\bar{\gamma}_{\bar{i}}\mathbf{I}+\rho\boldsymbol{h}_{\bar{i}}^{\perp}(i)\boldsymbol{h}_{\bar{i}}^{\perp\mathsf{H}}(i)\right)\boldsymbol{w}_i} \tag{3.65}$$

which is referred to as a *regularized* SINR in a dual two-user interference channel with a desired channel $\boldsymbol{h}_{\bar{i}}^{\perp}$ and interference channel $\boldsymbol{h}_{\bar{i}}^{\perp}$, where $\boldsymbol{w}_i$ is interpreted as a receive filter. Thus, the optimization problem in eq-(3.64) can be equivalently done by maximizing the regularized SINR in the dual MISO interference channels. Note that the regularization lies in not only the interference channels but also the desired channels. This solution is referred to later as "GMAT-DSINR".

General K-User Case

Recall that the definition of DSINR in eq-(3.65) for the two-user case, where $\boldsymbol{w}_i$ is determined by the orthogonal channels of itself and also its peers. According to the structure of $\boldsymbol{W}^l(2)$ for the K-user case, we can follow this approach and design each nonzero submatrices $\boldsymbol{w}_{ji}^l$ separately. For each $\boldsymbol{w}_{ji}^l$, the dual interference channel can be constructed by the orthogonal channels. Thus, the regularized dual SINR can be formulated as (e.g., $l = 1$)

$$\mathrm{DSINR}_{ji}^l = \frac{\boldsymbol{w}_{ji}^{l\,\mathsf{H}} \left(\mathbf{I} + \rho \sum_{k \neq i} \boldsymbol{h}_k^{\perp}(j) \boldsymbol{h}_k^{\perp \mathsf{H}}(j) \right) \boldsymbol{w}_{ji}^l}{\boldsymbol{w}_{ji}^{l\,\mathsf{H}} \left(\bar{\gamma}_{ji} \mathbf{I} + \rho \boldsymbol{h}_i^{\perp}(j) \boldsymbol{h}_i^{\perp \mathsf{H}}(j) \right) \boldsymbol{w}_{ji}^l}, \tag{3.66}$$

where $j \neq i$, $\boldsymbol{w}_{ji}^l \in \mathbb{C}^{K \times 1}$ is the i-th (when $i < j$) or $(i-1)$-th (when $i > j$) nonzero block of $\boldsymbol{W}_j^l(2)$, $\boldsymbol{h}_i^{\perp}(j) \in \mathbb{C}^{K \times K}$ is one representation of the null space of $\boldsymbol{h}_i(j)$ with the same norm[3], and

$$\bar{\gamma}_{ji} = \|\boldsymbol{w}_{ji}^l\|^2 + \|\boldsymbol{h}_i(j)\|^2 + 1/\rho \tag{3.67}$$

where $\|\boldsymbol{w}_{ji}^l\|^2$ can be chosen to satisfy the overall transmit power constraint, and T_2 was defined earlier in Section 3.2.3. Note that the numerator and denominator of DSINR_{ji}^l represent the requirements of signal orthogonality and interference alignment, respectively. While the latter aims at aligning $\boldsymbol{w}_{ji}^l$ as close as possible to the interference component $\boldsymbol{h}_i^{\perp}(j)$, the former tries to make $\boldsymbol{w}_{ji}^l$ as orthogonal as possible to the spanned subspace by all the channel vectors $\boldsymbol{h}_k^{\perp}(j)$ with $k \neq i$.

Accordingly, the optimal $\boldsymbol{w}_{ji}^l$ can be obtained by separately optimizing $(\forall\ i, j,\ i \neq j)$

$$\max_{\boldsymbol{w}_{ji}^l} \quad \mathrm{DSINR}_{ji}^l,\ j \neq i \tag{3.68}$$

$$s.t. \quad \sum_{l=1}^{L} \sum_{j=1}^{K} \|\boldsymbol{W}_j^l(2)\|_{\mathrm{F}}^2 \leq K T_2 \tag{3.69}$$

where the corresponding solution can be simply obtained by generalized eigenvalue decomposition. By maximizing the dual SINR, $\boldsymbol{w}_{ji}^l$ is preferred to keep aligned along with $\boldsymbol{h}_j(j)$ while to be as orthogonal to $\boldsymbol{h}_k(j)$ $(\forall\ k \neq i)$ as possible. Consequently, the optimal solution of $\boldsymbol{w}_{ji}^l$ balances signal orthogonality with interference alignment between user-j's and other users' dual orthogonal channels at j-th time slot.

[3]We abuse here the vector notation to represent the corresponding orthogonal channel matrix for the sake of consistence.

3.4 Discussion

3.4.1 Multiplexing Gain of GMAT

In the following, we show the GMAT algorithm possesses the same multiplexing gain as original MAT. We consider the two-user case for example. According to equations from (3.54) to (3.64), we have

$$\lim_{\rho\to\infty}\frac{\mathbb{E}\log\left(\max_{\|\boldsymbol{w}_1\|^2=1}\frac{\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{Q}_1\boldsymbol{w}_1}{\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{Q}_1\boldsymbol{w}_1}\right)}{\log\rho}=\lim_{\rho\to\infty}\frac{\mathbb{E}\log\left(\frac{\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{R}_1\boldsymbol{w}_1}{\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{Q}_1\boldsymbol{w}_1}\right)\Big|_{\boldsymbol{w}_1=\frac{\boldsymbol{h}_B(1)}{\|\boldsymbol{h}_B(1)\|}}}{\log\rho}=1 \tag{3.70}$$

$$\lim_{\rho\to\infty}\frac{\mathbb{E}\log\left(\max_{\|\boldsymbol{w}_2\|^2=1}\frac{\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{Q}_2\boldsymbol{w}_2}{\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{R}_2\boldsymbol{w}_2}\right)}{\log\rho}=\lim_{\rho\to\infty}\frac{\mathbb{E}\log\left(\frac{\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{Q}_2\boldsymbol{w}_2}{\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{R}_2\boldsymbol{w}_2}\right)\Big|_{\boldsymbol{w}_2=\frac{\boldsymbol{h}_A(2)}{\|\boldsymbol{h}_A(2)\|}}}{\log\rho}=1. \tag{3.71}$$

Thus, together with the fact that $\lim_{\rho\to\infty}\frac{\mathbb{E}\log(C_1C_4)}{\log\rho}=2$, the multiplexing gain can be achieved with

$$\mathrm{MG}=\lim_{\rho\to\infty}\frac{\mathbb{E}\max_{\|\boldsymbol{w}_i\|^2=1}(I(\boldsymbol{s}_A;\bar{\boldsymbol{y}}_A)+I(\boldsymbol{s}_B;\bar{\boldsymbol{y}}_B))}{3\log\rho}=\frac{4}{3}$$

which is identical to the original MAT algorithm. Intuitively, at high SNR, the signal orthogonality becomes no relevance, thus our solution naturally seeks perfect interference alignment as in MAT.

3.4.2 Single-Beam MIMO Interference Channel Interpretation

To understand more clearly the roles of desired signal orthogonality and interference alignment, we transform the mutual information equality (3.54) into another form, and further interpret their relationship from the point of view of a two-user single-beam MIMO interference channel. The strong benefit of this interpretation is that the problem of computing the precoders lends itself to classical precoding techniques in the MIMO interference channel. Based on eq-(3.54), the sum mutual information equation can be further transformed to the form

$$I(\boldsymbol{s}_A;\bar{\boldsymbol{y}}_A)+I(\boldsymbol{s}_B;\bar{\boldsymbol{y}}_B) \tag{3.72}$$

$$=\log\left(1+\frac{\alpha_1\rho\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{h}_A(1)\boldsymbol{h}_A^{\mathsf{H}}(1)\boldsymbol{w}_1+\alpha_2\rho\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{h}_A^{\perp}(1)\boldsymbol{h}_A^{\perp\mathsf{H}}(1)\boldsymbol{w}_1}{\sigma_1^2+\beta_3\rho\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{h}_A(2)\boldsymbol{h}_A^{\mathsf{H}}(2)\boldsymbol{w}_2+\beta_4\rho\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{h}_A^{\perp}(2)\boldsymbol{h}_A^{\perp\mathsf{H}}(2)\boldsymbol{w}_2}\right) \tag{3.73}$$

$$+\log\left(1+\frac{\beta_1\rho\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{h}_B(2)\boldsymbol{h}_B^{\mathsf{H}}(2)\boldsymbol{w}_2+\beta_2\rho\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{h}_B^{\perp}(2)\boldsymbol{h}_B^{\perp\mathsf{H}}(2)\boldsymbol{w}_2}{\sigma_2^2+\alpha_3\rho\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{h}_B(1)\boldsymbol{h}_B^{\mathsf{H}}(1)\boldsymbol{w}_1+\alpha_4\rho\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{h}_B^{\perp}(1)\boldsymbol{h}_B^{\perp\mathsf{H}}(1)\boldsymbol{w}_1}\right)+\log(C_1C_4) \tag{3.74}$$

where

$$\alpha_1=\frac{\alpha_2}{1+\rho\|\boldsymbol{h}_A(1)\|^2},\quad \alpha_2=\frac{1+\rho\|\boldsymbol{h}_A(2)\|^2}{\rho\|\boldsymbol{h}_A(1)\|^2}, \tag{3.75}$$

$$\alpha_3=\frac{1}{\rho|h_{B1}(3)|^2\|\boldsymbol{w}_1\|^2},\quad \alpha_4=\alpha_3+1, \tag{3.76}$$

$$\beta_1=\frac{\beta_2}{1+\rho\|\boldsymbol{h}_B(2)\|^2},\quad \beta_2=\frac{1+\rho\|\boldsymbol{h}_B(1)\|^2}{\rho\|\boldsymbol{h}_B(2)\|^2}, \tag{3.77}$$

$$\beta_3=\frac{1}{\rho|h_{A1}(3)|^2\|\boldsymbol{w}_2\|^2},\quad \beta_4=\beta_3+1, \tag{3.78}$$

$$\sigma_1^2=\frac{1}{\rho|h_{A1}(3)|^2}+\|\boldsymbol{w}_2\|^2,\sigma_2^2=\frac{1}{\rho|h_{B1}(3)|^2}+\|\boldsymbol{w}_1\|^2. \tag{3.79}$$

According to eq-(3.73) and eq-(3.74), the sum mutual information can be treated as that of two-user MIMO interference channels with 2 antennas at each transmitter and receiver, as shown in Fig. 3.2. Note that $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ act as the transmit beamformers, where the single beam is transmitted from each transmitter.

Accordingly, the received signals at two receivers can be equivalently expressed as

$$\boldsymbol{y}_1=\sqrt{\rho}\boldsymbol{H}_1\boldsymbol{w}_1s_1+\sqrt{\rho}\boldsymbol{H}_2\boldsymbol{w}_2s_2+\boldsymbol{n}_1 \tag{3.80}$$

$$\boldsymbol{y}_2=\sqrt{\rho}\boldsymbol{G}_2\boldsymbol{w}_2s_2+\sqrt{\rho}\boldsymbol{G}_1\boldsymbol{w}_1s_1+\boldsymbol{n}_2 \tag{3.81}$$

where

$$\boldsymbol{H}_1=\begin{bmatrix}\sqrt{\alpha_1}\boldsymbol{h}_A^{\mathsf{H}}(1)\\ \sqrt{\alpha_2}\boldsymbol{h}_A^{\perp\mathsf{H}}(1)\end{bmatrix},\quad \boldsymbol{H}_2=\begin{bmatrix}\sqrt{\beta_3}\boldsymbol{h}_A^{\mathsf{H}}(2)\\ \sqrt{\beta_4}\boldsymbol{h}_A^{\perp\mathsf{H}}(2)\end{bmatrix}, \tag{3.82}$$

$$\boldsymbol{G}_1=\begin{bmatrix}\sqrt{\alpha_3}\boldsymbol{h}_B^{\mathsf{H}}(1)\\ \sqrt{\alpha_4}\boldsymbol{h}_B^{\perp\mathsf{H}}(1)\end{bmatrix},\quad \boldsymbol{G}_2=\begin{bmatrix}\sqrt{\beta_1}\boldsymbol{h}_B^{\mathsf{H}}(2)\\ \sqrt{\beta_2}\boldsymbol{w}_B^{\perp\mathsf{H}}(2)\end{bmatrix} \tag{3.83}$$

and the additive noises follow the distribution $\boldsymbol{n}_i\sim\mathcal{CN}(0,\frac{\sigma_i^2}{2}\mathbf{I})$.

Consequently, the received SINR for both users can be written, respectively, as

$$\mathrm{SINR}_1=\frac{\rho\|\boldsymbol{H}_1\boldsymbol{w}_1\|^2}{\sigma_1^2+\rho\|\boldsymbol{H}_2\boldsymbol{w}_2\|^2}=\frac{\rho\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{H}_1^{\mathsf{H}}\boldsymbol{H}_1\boldsymbol{w}_1}{\sigma_1^2+\rho\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{H}_2^{\mathsf{H}}\boldsymbol{H}_2\boldsymbol{w}_2} \tag{3.84}$$

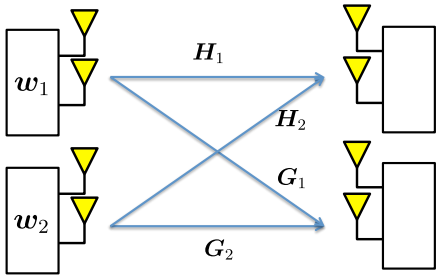


Figure 3.2: Interpretation as an MIMO interference channel.

$$\mathrm{SINR}_2 = \frac{\rho\|\boldsymbol{G}_2\boldsymbol{w}_2\|^2}{\sigma_2^2 + \rho\|\boldsymbol{G}_1\boldsymbol{w}_1\|^2} = \frac{\rho\boldsymbol{w}_2^{\mathsf{H}}\boldsymbol{G}_2^{\mathsf{H}}\boldsymbol{G}_2\boldsymbol{w}_2}{\sigma_2^2 + \rho\boldsymbol{w}_1^{\mathsf{H}}\boldsymbol{G}_1^{\mathsf{H}}\boldsymbol{G}_1\boldsymbol{w}_1} \tag{3.85}$$

which are identical to those in eq-(3.73-3.74). Note that the approach in the previous section that skips the dependency on the unknown channel coefficients $h_{A1}(3)$, $h_{B1}(3)$ can also be applied. We omit the details here to avoid redundancy. Hence, existing precoder design methods in the two-user single-beam MIMO interference channels with perfect CSIT, e.g., [65–68], can be used here in the context of delayed CSIT precoding. Instead of going into details about those solutions, we take the classic MRT and ZF precoders here for example,

$$\boldsymbol{w}_1^{MRT} = \boldsymbol{U}_{\max}(\boldsymbol{H}_1^{\mathsf{H}}\boldsymbol{H}_1), \qquad \boldsymbol{w}_2^{MRT} = \boldsymbol{U}_{\max}(\boldsymbol{G}_2^{\mathsf{H}}\boldsymbol{G}_2) \tag{3.86}$$

$$\boldsymbol{w}_1^{ZF} = \boldsymbol{U}_{\min}(\boldsymbol{G}_1^{\mathsf{H}}\boldsymbol{G}_1), \qquad \boldsymbol{w}_2^{ZF} = \boldsymbol{U}_{\min}(\boldsymbol{H}_2^{\mathsf{H}}\boldsymbol{H}_2) \tag{3.87}$$

where $\boldsymbol{U}_{\max}(\cdot)$ and $\boldsymbol{U}_{\min}(\cdot)$ are the generalized eigenvectors corresponding to the largest and smallest eigenvalues, respectively. Interestingly, for the first user, it is worth noting that $\alpha_1 < \alpha_2$ and therefore $\boldsymbol{w}_1^{MRT} \to \boldsymbol{h}_A^{\perp}(1)$, means perfect orthogonality of desired signal is preferred. On the other hand, $\alpha_3 < \alpha_4$, which denotes $\boldsymbol{w}_1^{ZF} \to \boldsymbol{h}_B(1)$, corresponds to the preference of perfect interference alignment. Our proposed GMAT-MMSE and GMAT-DSINR solutions offer a trade-off between them, yielding a better performance at finite SNR.

3.5 Numerical Results

The effectiveness of the proposed solutions is evaluated in terms of the sum rate per time slot in bps/Hz over a correlated rayleigh fading channel, where

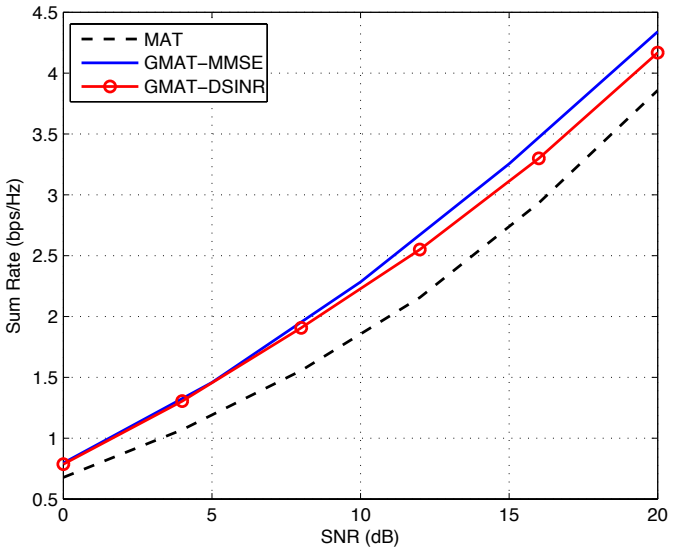


Figure 3.3: Sum rate vs. SNR for the two-user case.

the concatenated channel matrix in slot-t can be formulated as

$$\boldsymbol{H}(t) = \boldsymbol{R}_r^{1/2} \boldsymbol{H}_w(t) \boldsymbol{R}_t^{1/2} \tag{3.88}$$

where $\boldsymbol{H}_w(t)$ is normalized i.i.d. rayleigh fading channel matrix, and $\boldsymbol{R}_t$, $\boldsymbol{R}_r$ are transmit and receive correlation matrices with (i,j)-th entry being $\tau_t^{|i-j|}$ and $\tau_r^{|i-j|}$ [69, 70], respectively, where τ_t and τ_r are randomly chosen within $[0, 1)$. Note that the users' channel vectors are the rows of $\boldsymbol{H}(t)$.

The parameters in the simulation are set as follows. Maximum 500 gradient-descent iterations for the GMAT-MMSE with $\beta = 0.01$. The performance is averaged over 1000 channel realizations. Recall that the present channel coefficients (cf. $\boldsymbol{D}_i(k)$, e.g., $h_{A1}(3)$ and $h_{B1}(3)$ for the two-user case) are unknown to the transmitter and therefore are circumvented for transmit precoder design, while they should be taken into account at the receiver. Naturally, such a mismatch would result in performance degradation, but our proposed precoding methods are verified to be always effective thanks to the efficient trade-off between interference alignment and signal enhancement.

We show in Fig. 3.3 for the two-user case the sum rate comparison among GMAT-MMSE with the iteratively updated $\boldsymbol{w}_1$, $\boldsymbol{w}_2$, GMAT-DSINR with closed-form solutions in eq-(3.64), and the original MAT algorithm with

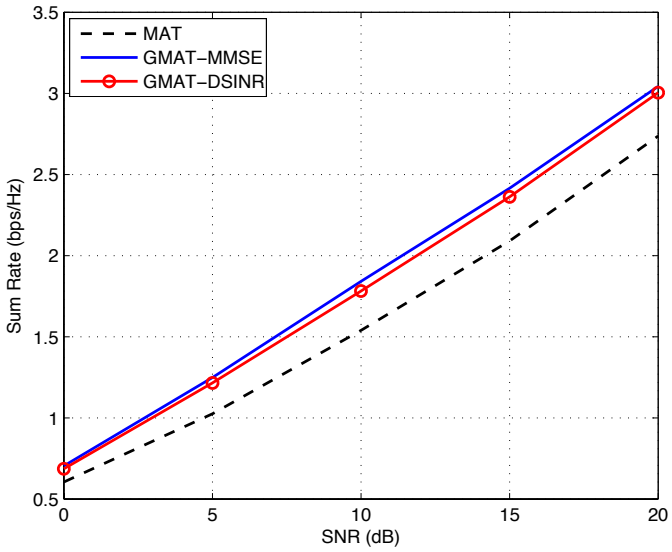


Figure 3.4: Sum rate vs. SNR for the three-user case.

$\boldsymbol{w}_1 = \boldsymbol{h}_B(1)$, $\boldsymbol{w}_2 = \boldsymbol{h}_A(2)$, with the same power constraint $\|\boldsymbol{w}_1\|^2 + \|\boldsymbol{w}_2\|^2 \leq 2$ for all. Note that MMSE receiver is used here to set up a reference for comparison together with GMAT-MMSE solution, such that we are able to show how good the closed-form solution can achieve compared to the iterative one at finite SNR. In Fig. 3.3, the gap of sum rate between GMAT and MAT illustrates improvement of GMAT-MMSE and GMAT-DSINR algorithms over the initial MAT concept, demonstrating the benefit of the trade-off between interference alignment and desired signal orthogonality enhancement. Compared with the original MAT algorithm, both GMAT approaches have gained substantial improvement at finite SNR and possessed the same slope, which implies the same multiplexing gain, at high SNR. Interestingly, the closed-form solution performs as well as the iterative one, indicating the effectiveness of the mutual information approximation.

In Fig. 3.4, we present the similar performance comparison for the three-user case with *MMSE receiver*. The GMAT-MMSE solution updates order-2 message generation matrix $\{\boldsymbol{W}_j^l(2)\}$ iteratively, while the original MAT algorithm set it according to eq-(3.22) and the GMAT-DSINR solution is obtained by optimizing eq-(3.68) and eq-(3.69). All these methods hold the same power allocation. With more transmit antennas and users, the same insights regarding the trade-off between signal orthogonality and in-

terference alignment can be always obtained. It is interesting to note that, GMAT-DSINR performs as well as GMAT-MMSE, despite the distributed optimization.

3.6 Summary

We generalize the concept of precoding over a multi-user MISO broadcast channel with delayed CSIT for arbitrary number of users case, by proposing a precoder construction algorithm, which achieves the same DoF at infinite SNR yet reaches a useful trade-off between interference alignment and signal enhancement at finite SNR. Our proposed precoding concept lends itself to a variety of optimization methods, e.g., virtual MMSE and mutual information solutions, achieving good compromise between signal orthogonality and interference alignment. An interesting question is also the diversity gain performance of schemes combining current and delayed CSIT. Clearly our scheme will achieve the same diversity performance as the original MAT since it converges to it in the high SNR regime. The question of whether modified schemes can be devised to address the DoF-diversity trade-off is an interesting open problem.

Remarkably, our work has triggered a line of interesting works since it published in [71, 72], including [73], among many others, in which the precoding methods were considered in the delayed CSIT setting at finite SNR, together with the additional statistical CSIT.

3.7 Appendix

3.7.1 Gradient Descent Parameter for GMAT-MMSE

Let $[\boldsymbol{H}_{ij}^l]_{m,n} = \boldsymbol{e}_m^{\mathsf{H}} \boldsymbol{H}_{ij}^l \boldsymbol{e}_n$ be the m-th row and n-th column element of $\boldsymbol{H}_{ij}^l$. Particularly,

$$[\boldsymbol{H}_{ij}^l]_{m,n} = \boldsymbol{e}_{m'}^{\mathsf{H}} \boldsymbol{W}_j^l(k) \boldsymbol{e}_n \tag{3.89}$$

when $m = \sum_{s=1}^{k-1} T_s + m'$ where $1 \leq m' \leq T_k$ and $1 \leq n \leq K$. Here, $\boldsymbol{e}_m$ is defined as a binary vector with only one '1' at m-th row. By differentiating over $\boldsymbol{W}_j^l(2)$, we have the differentiation

$$\frac{\partial [\boldsymbol{H}_{ij}^l]_{m,n}}{\partial \boldsymbol{W}_j^{l\ \mathsf{T}}(2)} = \left(\frac{\partial [\boldsymbol{H}_{ij}^l]_{m,n}}{\partial \boldsymbol{W}_j^l(2)} \right)^{\mathsf{T}} \tag{3.90}$$

$$= \begin{cases} \mathbf{0} & if\ m \le T_1 \\ \boldsymbol{e}_n \boldsymbol{e}_{m'}^{\mathsf{H}} & if\ T_1 + 1 \le m \le T_1 + T_2 \\ \boldsymbol{e}_n \boldsymbol{e}_{m'}^{\mathsf{H}} \prod_{t=2}^{k-1} \boldsymbol{C}^l(t) \boldsymbol{\Lambda}^l(t) & if\ \sum_{s=1}^{k-1} T_s + 1 \le m \le \sum_{s=1}^{k} T_s\ when\ k \ge 3 \end{cases} \tag{3.91}$$

$$= \boldsymbol{e}_n \boldsymbol{e}_m^{\mathsf{H}} \boldsymbol{Q}^l \tag{3.92}$$

where

$$\boldsymbol{Q}^l = \begin{bmatrix} \mathbf{0}_{T_1 \times K} \\ \mathbf{0}_{m_2^l \times K} \\ \mathbf{I} \\ \mathbf{0}_{n_2^l \times K} \\ \vdots \\ \prod_{t=2}^{K-1} \boldsymbol{C}^l(t) \boldsymbol{\Lambda}^l(t) \end{bmatrix}. \tag{3.93}$$

Note that we abuse vector $\boldsymbol{e}_{m'}$ with various dimensions T_k according to the corresponding matrices $\boldsymbol{W}_j^l(k)$ for the sake of notational simplicity. Then, it follows that

$$\frac{\partial [\boldsymbol{H}_{ij}^l]_{m,n}}{\partial [\boldsymbol{W}_j^{l\ \mathsf{T}}(2)]_{p,q}} = \boldsymbol{e}_m^{\mathsf{H}} \boldsymbol{Q}^l \boldsymbol{e}_p \boldsymbol{e}_q^{\mathsf{H}} \boldsymbol{e}_n \tag{3.94}$$

where $1 \le p \le T_k$, $1 \le q \le K$, and we have

$$\frac{\partial \boldsymbol{H}_{ij}^l}{\partial [\boldsymbol{W}_j^{l\ \mathsf{T}}(2)]_{p,q}} = \boldsymbol{Q}^l \boldsymbol{e}_p \boldsymbol{e}_q^{\mathsf{H}} \tag{3.95}$$

Finally, according to the chain rule of matrix differentiation [63, 74], we have

$$\frac{\partial \left(J_i^l\right)}{\partial [\boldsymbol{W}_j^{l\ \mathsf{T}}(2)]_{p,q}} = \mathrm{Tr}\left(\left(\frac{\partial J_i^l}{\partial \boldsymbol{H}_{ij}^l} \right)^{\mathsf{T}} \frac{\partial \boldsymbol{H}_{ij}^l}{\partial [\boldsymbol{W}_j^{l\ \mathsf{T}}(2)]_{p,q}} \right) \tag{3.96}$$

$$= \mathrm{Tr}\left(\boldsymbol{e}_q^{\mathsf{H}} \left(\frac{\partial J_i^l}{\partial \boldsymbol{H}_{ij}^l} \right)^{\mathsf{T}} \boldsymbol{Q}^l \boldsymbol{e}_p \right). \tag{3.97}$$

So, for the K-user case, the Gaussian descent parameter can be calculated by

$$\frac{\partial (J)}{\partial \boldsymbol{W}_j^l(2)} = \sum_{i=1}^{K} \frac{\partial \left(J_i^l\right)}{\partial \boldsymbol{W}_j^l(2)} = \sum_{i=1}^{K} \left(\frac{\partial J_i^l}{\partial \boldsymbol{H}_{ij}^l} \right)^{\mathsf{T}} \boldsymbol{Q}^l \tag{3.98}$$

where

$$\left(\frac{\partial J_i^l}{\partial \boldsymbol{H}_{ii}^l}\right)^{\mathsf{T}} = f\left(\sqrt{\rho}\boldsymbol{H}_{ii}^l, \rho\sum_{l=1}^{L}\sum_{j=1,j\neq i}^{K}\boldsymbol{H}_{ij}^l\boldsymbol{H}_{ij}^{l\,\mathsf{H}} + \mathbf{I}\right)$$
$$\left(\frac{\partial J_i^l}{\partial \boldsymbol{H}_{ij}^l}\right)^{\mathsf{T}} = g\left(\sqrt{\rho}\boldsymbol{H}_{ij}^l, \rho\sum_{l=1}^{L}\sum_{k=1,k\neq j}^{K}\boldsymbol{H}_{ik}^l\boldsymbol{H}_{ik}^{l\,\mathsf{H}} + \mathbf{I}\right)$$

where

$$f(\boldsymbol{A},\boldsymbol{B}) = -\boldsymbol{A}^{\mathsf{H}}\left(\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}}+\boldsymbol{B}\right)^{-1}\boldsymbol{B}\left(\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}}+\boldsymbol{B}\right)^{-1}$$
$$g(\boldsymbol{A},\boldsymbol{B}) = \boldsymbol{A}^{\mathsf{H}}\left(\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}}+\boldsymbol{B}\right)^{-1}\left(\boldsymbol{B}-\mathbf{I}\right)\left(\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}}+\boldsymbol{B}\right)^{-1}.$$

Chapter 4

MIMO Networks with Delayed CSIT: The Time-Correlated Case

While the previous chapter focused on throughput maximization of time i.i.d. MISO BC with delayed CSIT at finite SNR, in this chapter we take into account channel time correlation in MIMO networks. It is quite challenging to optimize throughput at finite SNR with channel time correlation. Instead, we maximize DoF region, from an information theoretic perspective, as a first-order capacity approximation at infinite SNR and leave the exact capacity region characterization to future study.

In particular, we consider the time-correlated MIMO BC and IC where the transmitter(s) has/have (i) delayed CSI obtained from a latency-prone feedback channel as well as (ii) imperfect current CSIT, obtained, e.g., from prediction on the basis of these past channel samples based on the temporal correlation. The DoF regions for the two-user broadcast and interference MIMO networks with general antenna configuration are fully characterized, as a function of the prediction quality indicator. Specifically, a simple unified framework is proposed, allowing to attain optimal DoF region for the general antenna configurations and current CSIT qualities. Such a framework builds upon block-Markov encoding with interference quantization, optimally combining the use of both outdated and instantaneous CSIT. A striking feature is that, by varying the power allocation, every point in the DoF region can be achieved with one single scheme. As a result, instead of checking the achievability of every corner point of the outer bound region, as typically done in the literature, we propose a new systematic way to prove the achievability.

4.1 Introduction

While the capacity region of the MIMO BC was established in [26], the characterization of the capacity of Gaussian IC has been a long-standing open problem, even for the two-user single-antenna case. Recent progress sheds light on this problem from various perspectives, among which the authors in [28] characterized the DoF region for the two-user MIMO IC. In most works, the DoF analysis for multiuser channels involves the full knowledge of CSI at both the transmitter and receiver sides. In practice, however, the acquisition of perfect CSI at the transmitters is difficult especially for fast fading channels. The CSIT obtained via feedback suffers from delays, which renders the available CSIT feedback possibly fully obsolete (i.e., fully "delayed CSIT") under the fast fading channel.

As mentioned in previous chapter, under this delayed CSIT setting, a novel scheme (termed in this chapter as "MAT alignment") was proposed in [41] for the MISO BC to demonstrate that even the completely outdated channel feedback is still useful. The precoders are designed achieving strictly better DoF than what is obtained without any CSIT. The essential ingredient for the proposed scheme in [41] lies in the use of a multi-slot protocol initiating with the transmission of unprecoded information symbols to the user terminals, followed by the *analog* forwarding of the interferences created in the previous time slots. Most recently, generalizations under the similar principle to the MIMO BC [75], MIMO IC [43] settings, the MIMO BC with secrecy constraints [47], among others, were also addressed, where the DoF regions are fully characterized with arbitrary antenna configurations, again establishing DoF strictly beyond the ones obtained without CSIT [29, 30, 32] but below the ones with perfect CSIT [26, 28]. Note that other recent interesting lines of work combining instantaneous and delayed forms of feedback were reported in [48, 76–78].

Albeit inspiring and fascinating from a conceptual point of view, these works made an assumption that the channel is i.i.d. across time, where the delayed CSIT bears no correlation with the current channel realization. Hence, these results pessimistically consider that no estimate for the *current* channel realization is producible to the transmitter. Owing to the finite Doppler spread behavior of fading channels, it is however the case in many real life situations that the past channel realizations can provide *some* information about the current one. Therefore a scenario where the transmitter is endowed with delayed CSI in addition to some (albeit imperfect) estimate of the current channel is of practical relevance. Together with the delayed CSIT, the benefit of such imperfect current CSIT was first exploited in [53] for the MISO

BC whereby a novel transmission scheme was proposed which improves over pure MAT alignment in constructing precoders based on delayed *and* current CSIT estimate. The full characterization of the optimal DoF for this hybrid CSIT was later reported in [54, 55] for MISO BC under this setting. The key idea behind the schemes (termed hereafter as "α-MAT alignment") in [53–55] lies in the modification of the MAT alignment such that i) the initial time slot involves transmission of *precoded* symbols, which enables to reduce the power of mutual interferences and efficiently compress them; ii) the subsequent slots perform a digital transmission of quantized residual interferences together with new private symbols. Most recently, this philosophy was extended to the MIMO networks (BC/IC) but only with symmetric antenna configurations [79], as well as the K-user MISO case [80]. The generalization to the MISO BC with different qualities of imperfect current CSIT was also studied in [81]. Remarkably, the authors of [81] showed that, in order to balance the asymmetry of the CSIT quality, an infinite number of time slots are required. As such, they extended the number of phases of the α-MAT alignment [54] to infinity and varied the length of each phase.

Unfortunately, extending the previous results to the MIMO case with arbitrary antenna configurations is not a trivial step, even with the symmetric current CSIT quality assumption. The main challenges are two-fold: (a) the extra spatial dimension at the receiver side introduces a non-trivial tradeoff between the useful signal and the mutual interference, and (b) the asymmetry of receive antenna configurations results in the discrepancy of common-message-decoding capability at different receivers. In particular, the total number of streams that can be delivered as common messages to both receivers is inevitably limited by the weaker one (i.e., with fewer antennas). Such a constraint prevents the system from achieving the optimal DoF of the symmetric case by simply extending the previous schemes developed in [79].

To counter these new challenges posed by the asymmetry of antenna configurations, we develop a new strategy that balances the discrepancy of common-message-decoding capability at two receivers. This allows us to fully characterize the DoF region of both MIMO BC and MIMO IC, achieved by a unified and simple scheme built upon *block-Markov encoding*. This encoding concept was first introduced in [82] for relay channels and then became a standard tool for communication problems involving interaction between nodes, such as feedback (e.g., [83,84]) or user cooperation (e.g., [85]). It turns out that our problem with both delayed and instantaneous CSIT, closely related to [83], can also be solved with this scheme. As it will become clear later, in each block, the transmitter superimposes the common information

about the interferences created in the past block (due to the imperfect instantaneous CSIT) on the new private information (thus creating new interferences). At the receiver side, backward decoding is employed, i.e., the decoding of each block relies on the common side information from the decoding of future blocks. Due to the repetitive nature in each block, the proposed scheme can be uniquely characterized with the parameters such as the power allocation and rate splitting of the superposition. Surprisingly enough, our block-Markov scheme can also include the asymmetry of current CSIT with a simple parameter change, and thus somehow balance the global asymmetry, i.e., antenna asymmetry and CSIT asymmetry, in the system.

More specifically, in this chapter we obtain the following key results:

- We establish outer bounds on the DoF region for the two-user temporally-correlated MIMO BC and IC with perfect delayed and imperfect current CSIT, as a function of the current CSIT quality *exponent*. By introducing a virtual received signal for the IC, we nicely link the outer bound to that of the BC, arriving at the similar outer bound results for both cases. In addition to the genie-aided bounding techniques and the application of the extremal inequality in [54], we develop a set of upper and lower bounds of ergodic capacity for MIMO channels, which is essential for the MIMO case but not extendible from MISO.

- We propose a unified framework relying on block-Markov encoding uniquely parameterized by the rate splitting and power allocation, by which the optimal DoF regions confined by the outer bounds are achievable with perfect delayed plus imperfect current CSIT. For any antenna and current CSIT settings, every point in the outer bound region can be achieved with one single scheme. For instance, the MIMO BC with $M = 3$, $N_1 = 2$ and $N_2 = 1$ achieves optimal sum DoF $\frac{15+4\alpha_1+2\alpha_2}{7}$ when $3\alpha_1 - 2\alpha_2 \leq 1$ and $\frac{7+2\alpha_2}{3}$ otherwise, where α_1 and α_2 are imperfect current CSIT qualities for both users' channels. This smoothly connects three special cases: the case with pure delayed CSIT [75] ($\alpha_1 = \alpha_2 = 0$), that with perfect current CSIT [26] ($\alpha_1 = \alpha_2 = 1$), and that with perfect CSIT at Receiver 1 and delayed CSIT at Receiver 2 [86] ($\alpha_1 = 1, \alpha_2 = 0$).

- We propose a new systematic way to prove the achievability. In the proposed framework, the achievability region is defined by the decodability conditions in terms of the rate splitting and power allocation. The achievability is proved by mapping the outer bound region into a set of proper rate and power allocation and showing that this set lies

within the decodability region. This contrasts with most existing proofs in literature where the achievability of each corner point is checked.

It is worth noting that our results cover the previously reported particular cases: the perfect CSIT setting [26,28] (i.e., current CSIT of perfect quality), the pure delayed CSIT setting [43] (i.e., current CSIT of zero quality), the partial/hybrid CSIT MIMO BC/IC case [42, 86, 87] (with perfect CSIT at one receiver and delayed CSIT at the other one), and the special MISO case [53–55] with $N_1 = N_2 = 1$, symmetric MIMO case [79], as well as the MISO case with asymmetric current CSIT qualities [81]. In a parallel work [88], a similar scheme was independently revealed, also built on the block-Markov encoding, evolving from the multi-phase scheme initially proposed in [81]. While they focus on the MISO BC in a more general evolving CSIT setting, our work deals with a wider class of channel configurations (both MIMO BC and IC) with static CSIT.

The rest of the chapter is organized as follows. We present the system model and assumptions in the coming section, followed by the main results on DoF region characterization for both MIMO BC and MIMO IC cases in Section 4.3. Some illustrative examples of the achievability schemes are provided in Section 4.4, followed by the general formulation in Section 4.5. In Section 4.6, we present the proofs of outer bounds. Finally, we conclude the chapter in Section 4.7.

Notation: Matrices and vectors are represented as uppercase and lowercase letters, respectively. Matrix transport, Hermitian transport, inverse, rank, determinant and the Frobenius norm of a matrix are denoted by $\boldsymbol{A}^{\mathsf{T}}$, $\boldsymbol{A}^{\mathsf{H}}$, $\boldsymbol{A}^{-1}$, $\mathrm{rank}(\boldsymbol{A})$, $\det(\boldsymbol{A})$ and $\|\boldsymbol{A}\|_{\mathrm{F}}$, respectively. $\boldsymbol{A}_{[k_1:k_2]}$ represents the submatrix of $\boldsymbol{A}$ from k_1-th row to k_2-th row when $k_1 \leq k_2$. $\boldsymbol{h}^{\perp}$ is the normalized orthogonal component of any non-zero vector $\boldsymbol{h}$. We use $\mathbf{I}_M$ to denote an $M \times M$ identity matrix where the dimension is omitted whenever confusion is not probable. The approximation $f(P) \sim g(P)$ is in the sense of $\lim_{P\to\infty} \frac{f(P)}{g(P)} = C$, where $C > 0$ is a constant that does not scale as P. Partial ordering of Hermitian matrices is denoted by $\succeq$ and $\preceq$, i.e., $\boldsymbol{A} \preceq \boldsymbol{B}$ means $\boldsymbol{B} - \boldsymbol{A}$ is positive semidefinite. Logarithms are in base 2. $(x)^+$ means $\max\{x, 0\}$, and $\mathbb{R}_+^n$ represents the set of n-tuples of non-negative real numbers. $f = O(g)$ follows the standard Landau notation, i.e., $\lim \frac{f}{g} \leq C$ where the limit depends on the context. With some abuse of notation, we use $O_X(g)$ to denote any f such that $\mathbb{E}_X(f) = O(\mathbb{E}_X(g))$. Finally, the range or null spaces mentioned in this chapter refer to the column spaces.

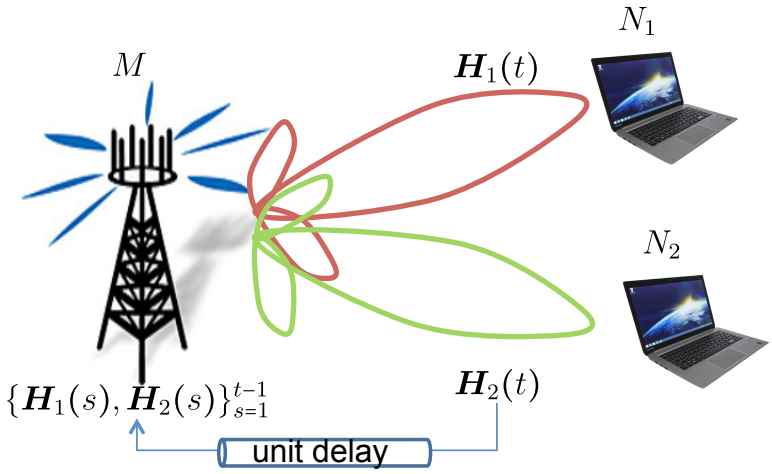


Figure 4.1: MIMO broadcast channel with delayed CSIT.

4.2 System Model

4.2.1 Two-user MIMO Broadcast Channel

For a two-user (M, N_1, N_2) MIMO broadcast channel with M antennas at the transmitter and N_i antennas at Receiver i, as in Fig. 4.1, the discrete time signal model is given by

$$\boldsymbol{y}_i(t) = \boldsymbol{H}_i(t)\boldsymbol{x}(t) + \boldsymbol{z}_i(t) \tag{4.1}$$

for any time instant t, where $\boldsymbol{H}_i(t) \in \mathbb{C}^{N_i \times M}$ is the channel matrix for Receiver i $(i = 1, 2)$; $\boldsymbol{z}_i(t) \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_{N_i})$ is the normalized additive white Gaussian noise (AWGN) vector at Receiver i and is independent of channel matrices and transmitted signals; the coded input signal $\boldsymbol{x}(t) \in \mathbb{C}^{M \times 1}$ is subject to the power constraint $\mathbb{E}\big(\|\boldsymbol{x}(t)\|^2\big) \leq P$, $\forall\, t$.

4.2.2 Two-user MIMO Interference Channel

For a two-user (M_1, M_2, N_1, N_2) MIMO interference channel with M_i antennas at Transmitter i and N_j antennas at Receiver j, for $i, j = 1, 2$, as in Fig. 4.2, the discrete time signal model is given by

$$\boldsymbol{y}_i(t) = \boldsymbol{H}_{i1}(t)\boldsymbol{x}_1(t) + \boldsymbol{H}_{i2}(t)\boldsymbol{x}_2(t) + \boldsymbol{z}_i(t) \tag{4.2}$$

for any time instant t, where $\boldsymbol{H}_{ji}(t) \in \mathbb{C}^{N_j \times M_i}$ $(i, j = 1, 2)$ is the channel matrix between Transmitter i and Receiver j; the coded input signal $\boldsymbol{x}_i(t) \in \mathbb{C}^{M_i \times 1}$ is subject to the power constraint $\mathbb{E}\big(\|\boldsymbol{x}_i(t)\|^2\big) \leq P$ for $i = 1, 2$, $\forall\, t$.

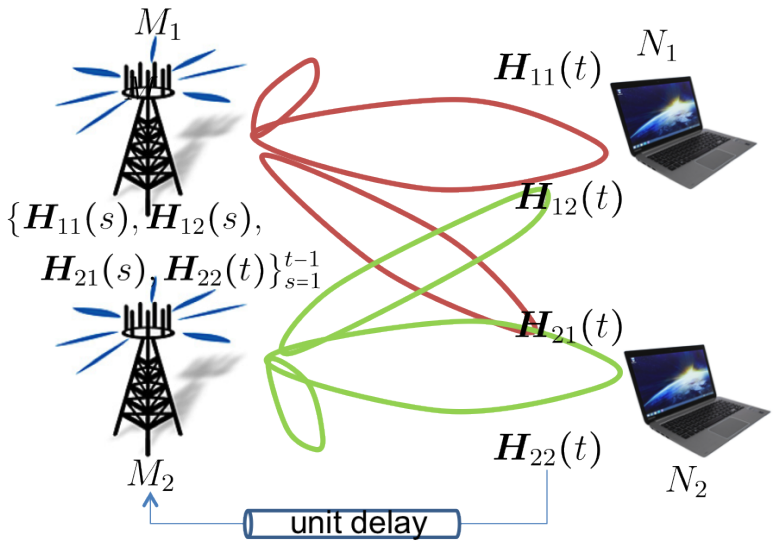


Figure 4.2: MIMO interference channel with delayed CSIT.

In the rest of this chapter, we refer to MIMO BC/IC as MIMO networks. For notational brevity, we define the ensemble of channel matrices, i.e., $\mathcal{H}(t) \triangleq \{\boldsymbol{H}_1(t), \boldsymbol{H}_2(t)\}$ (resp. $\mathcal{H}(t) \triangleq \{\boldsymbol{H}_{11}(t), \boldsymbol{H}_{21}(t), \boldsymbol{H}_{12}(t), \boldsymbol{H}_{22}(t)\}$), as the channel state for BC (resp. IC). We further define $\mathcal{H}^k \triangleq \{\mathcal{H}(t)\}_{t=1}^k$, and $\hat{\mathcal{H}}^k \triangleq \{\hat{\mathcal{H}}(t)\}_{t=1}^k$, where $k = 1, \cdots, n$.

4.2.3 Assumptions and Definitions

Assumption 4.1 (perfect delayed and imperfect current CSIT)**.** *At each time instant t, the transmitters know perfectly the delayed CSI $\mathcal{H}^{t-1}$, and obtain an imperfect estimate of the current CSI $\hat{\mathcal{H}}(t)$, which could, for instance, be produced by standard prediction based on past samples. The current CSIT estimate is modeled by*

$$\boldsymbol{H}_i(t) = \hat{\boldsymbol{H}}_i(t) + \tilde{\boldsymbol{H}}_i(t) \tag{4.3}$$

$$\boldsymbol{H}_{ij}(t) = \hat{\boldsymbol{H}}_{ij}(t) + \tilde{\boldsymbol{H}}_{ij}(t) \tag{4.4}$$

for BC and IC, respectively, where estimation error $\tilde{\boldsymbol{H}}_i(t)$ (resp. $\tilde{\boldsymbol{H}}_{ij}(t)$) and the estimate $\hat{\boldsymbol{H}}_i(t)$ (resp. $\hat{\boldsymbol{H}}_{ij}(t)$) are mutually independent, and each entry is assumed[1] *to be $\mathcal{N}_{\mathbb{C}}\left(0, \sigma_i^2\right)$ and $\mathcal{N}_{\mathbb{C}}\left(0, 1-\sigma_i^2\right)$. Further, we assume the*

[1]We make the above assumption on the fading distribution to simplify the presentation, although the results can be applied to a broader class of distributions.

following Markov chain

$$(\mathcal{H}^{t-1}, \hat{\mathcal{H}}^{t-1}) \to \hat{\mathcal{H}}(t) \to \mathcal{H}(t), \quad \forall t, \tag{4.5}$$

which means $\mathcal{H}(t)$ is independent of $(\mathcal{H}^{t-1}, \hat{\mathcal{H}}^{t-1})$ conditional on $\hat{\mathcal{H}}(t)$. Furthermore, at the end of the transmission, i.e., at time instant n, the receivers know perfectly $\mathcal{H}^n$ and $\hat{\mathcal{H}}^n$.

It readily follows that, for any fat submatrix $\boldsymbol{H}$ of $\boldsymbol{H}_i$ or $\boldsymbol{H}_{ij}$,

$$\mathbb{E}(\log\det(\boldsymbol{H}\boldsymbol{H}^{\mathsf{H}})) > -\infty \text{ and } \mathbb{E}(\log\det(\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^{\mathsf{H}})) = O(1)$$

when σ_i^2 goes to 0.

The assumption on CSIR is in accordance with previous works with delayed CSIT, and does not add any limitation over the assumption made in [41, 43, 75]. We point out that only local CSIT/CSIR (the channel links with which the node is connected) is really helpful and leads to the same result. Nevertheless, we assume the CSIT/CSIR to be available in a global fashion for simplicity of presentation.

We are interested in characterizing the DoF of the above system as functions of the quality of current CSIT, thus bridging between the two previously investigated extremes which are the perfect instantaneous CSIT and the fully outdated (non-instantaneous) CSIT cases. As it was established in previous works [53, 54], the imperfect current CSIT has beneficial value (in terms of improving the DoF) only if the CSIT estimation error decays at least exponentially with the SNR or faster. Thus it is reasonable to study the regime by which the CSIT quality can be parameterized by an indicator $\alpha_i \geq 0$ such that:

$$\alpha_i \triangleq -\lim_{P\to\infty} \frac{\log \sigma_i^2}{\log P} \tag{4.6}$$

if the limit exists. This α_i indicates the quality of current CSIT corresponding to Receiver i at high SNR. While $\alpha_i = 0$ reflects the case with no current CSIT, $\alpha_i \to \infty$ corresponds to that with perfect instantaneous CSIT. As a matter of fact, when $\alpha_i \geq 1$, the quality of the imperfect current CSIT is sufficient to avoid the DoF loss, and ZF precoding with this imperfect CSIT is able to achieve the maximum DoF [39]. Therefore, we focus on the case $\alpha_i \in [0, 1]$ henceforth. The connections between the above model and the linear prediction over existing time-correlated channel models with prescribed user mobility are highlighted in [53, 54]. According to the definition of the estimated current CSIT, we have $\mathbb{E}\big(|\boldsymbol{h}_k^{\mathsf{H}}(t)\hat{\boldsymbol{h}}_k^{\perp}(t)|^2\big) = \sigma_i^2 \sim P^{-\alpha_i}$, with $\boldsymbol{h}_k^{\mathsf{H}}$

representing any row of channel matrices $\boldsymbol{H}_i(t)$ (resp. $\boldsymbol{H}_{ij}(t)$), and $\hat{\boldsymbol{h}}_k^{\mathsf{H}}$ being its corresponding estimate.

A rate pair (R_1, R_2) is said to be *achievable* for the two-user MIMO networks with perfect delayed and imperfect current CSIT if there exists a $\big(2^{nR_1}, 2^{nR_2}, n\big)$ code scheme with:

- two message sets $\mathcal{W}_1 \triangleq [1 : 2^{nR_1}]$ and $\mathcal{W}_2 \triangleq [1 : 2^{nR_2}]$, from which two independent messages W_1 and W_2 intended respectively to Receiver 1 and Receiver 2 are uniformly chosen;

- one encoding function for (each) transmitter:

$$\begin{array}{ll} \text{BC:} & \boldsymbol{x}(t) = f_t\big(W_1, W_2, \mathcal{H}^{t-1}, \hat{\mathcal{H}}^t\big) \\ \text{IC:} & \boldsymbol{x}_i(t) = f_{i,t}\big(W_i, \mathcal{H}^{t-1}, \hat{\mathcal{H}}^t\big),\ i = 1, 2; \end{array} \tag{4.7}$$

- one decoding function at the corresponding receiver,

$$\hat{W}_j = g_j\big(\boldsymbol{Y}_j^n, \mathcal{H}^n, \hat{\mathcal{H}}^n\big),\ j = 1, 2 \tag{4.8}$$

  for Receiver j, where $\boldsymbol{Y}_j^n \triangleq \{\boldsymbol{y}_j(t)\}_{t=1}^n$,

such that the average decoding error probability $P_e^{(n)}$, defined as $P_e^{(n)} \triangleq \mathbb{P}\big((W_1, W_2) \neq (\hat{W}_1, \hat{W}_2)\big)$, vanishes as the code length n tends to infinity. The capacity region $\mathcal{C}$ is defined as the set of all achievable rate pairs. Accordingly, the DoF region can be defined as follows:

Definition 4.1 (DoF region)**.** *The DoF region for the two-user MIMO network is defined as*

$$\begin{aligned} \mathcal{D} = \Bigg\{ (d_1, d_2) \in \mathbb{R}_+^2 |\ \forall (w_1, w_2) \in \mathbb{R}_+^2, w_1 d_1 + w_2 d_2 \\ \leq \limsup_{P \to \infty} \left(\sup_{(R_1, R_2) \in \mathcal{C}} \frac{w_1 R_1 + w_2 R_2}{\log P} \right) \Bigg\}. \end{aligned} \tag{4.9}$$

4.3 Main Results

According to the assumptions and definitions in the previous section, the main results of this chapter are stated as the following two theorems:

Theorem 4.1. *For the two-user (M, N_1, N_2) MIMO BC with delayed and imperfect current CSIT, the optimal DoF region $\{(d_1, d_2)|(d_1, d_2) \in \mathbb{R}_+^2\}$ is characterized by*

$$d_1 \leq \min\{M, N_1\}, \tag{4.10a}$$

$$d_2 \leq \min\{M, N_2\}, \tag{4.10b}$$

$$d_1 + d_2 \leq \min\{M, N_1 + N_2\}, \tag{4.10c}$$

$$\frac{d_1}{\min\{M, N_1\}} + \frac{d_2}{\min\{M, N_1 + N_2\}} \leq 1 + \frac{\min\{M, N_1 + N_2\} - \min\{M, N_1\}}{\min\{M, N_1 + N_2\}}\alpha_1, \tag{4.10d}$$

$$\frac{d_1}{\min\{M, N_1 + N_2\}} + \frac{d_2}{\min\{M, N_2\}} \leq 1 + \frac{\min\{M, N_1 + N_2\} - \min\{M, N_2\}}{\min\{M, N_1 + N_2\}}\alpha_2, \tag{4.10e}$$

where $\alpha_i \in [0, 1]$ $(i = 1, 2)$ indicates the current CSIT quality exponent of Receiver i's channel.

Proof. The proof of achievability will be presented in Section 4.4 showing some insights with toy examples, and in Section 4.5 for the general formulation. The converse proof will be given in Section 4.6 focusing on (4.10d) and (4.10e), because the first three bounds correspond to the upper bounds under perfect CSIT settings and thus hold trivially under delayed and imperfect current CSIT settings. □

Remark 4.1. *This result yields a number of previous results as special cases: the delayed CSIT case [75] for $\alpha_1 = \alpha_2 = 0$, where the sum DoF bound* (4.10c) *is inactive; perfect CSIT case [26] for $\alpha_1 = \alpha_2 = 1$, where the weighted sum DoF bounds* (4.10d) *and* (4.10e) *are inactive; partial CSIT (i.e., perfect CSIT for one channel and delayed CSIT for the other one) case [86] for $\alpha_1 = 1, \alpha_2 = 0$, where only* (4.10b) *and* (4.10e) *are active; delayed CSIT in MISO BC for $N_1 = N_2 = 1$ [54, 55, 81].*

Before presenting the optimal DoF region for MIMO IC, we specify two conditions.

Definition 4.2 (Condition C_k)**.** *Given $k \in \{1, 2\}$, the condition C_k holds, indicating the following inequalities*

$$M_k \geq N_j, \quad M_j < N_k, \quad M_1 + M_2 > N_1 + N_2 \tag{4.11}$$

are true, $\forall\ j \in \{1, 2\},\ j \neq k$.

Remark 4.2. *This definition that points out the existence of the corresponding outer bound, is different from that in [43], in which the condition implies the activation of the outer bounds.*

Theorem 4.2. *For the two-user (M_1, M_2, N_1, N_2) MIMO IC with delayed and imperfect current CSIT, the optimal DoF region $\{(d_1, d_2)|(d_1, d_2) \in \mathbb{R}_+^2\}$ is characterized by*

$$d_1 \leq \min\{M_1, N_1\}, \tag{4.12a}$$

$$d_2 \leq \min\{M_2, N_2\}, \tag{4.12b}$$

$$\begin{aligned} d_1 + d_2 \leq \min\{&M_1 + M_2, N_1 + N_2, \\ &\max\{M_1, N_2\}, \max\{M_2, N_1\}\}, \end{aligned} \tag{4.12c}$$

$$\begin{aligned} \frac{d_1}{\min\{M_2, N_1\}} + \frac{d_2}{\min\{M_2, N_1 + N_2\}} \leq & \frac{\min\{N_1, M_1 + M_2\}}{\min\{M_2, N_1\}} \\ & + \frac{\min\{M_2, N_1 + N_2\} - \min\{M_2, N_1\}}{\min\{M_2, N_1 + N_2\}}\alpha_1, \end{aligned} \tag{4.12d}$$

$$\begin{aligned} \frac{d_1}{\min\{M_1, N_1 + N_2\}} + \frac{d_2}{\min\{M_1, N_2\}} \leq & \frac{\min\{N_2, M_1 + M_2\}}{\min\{M_1, N_2\}} \\ & + \frac{\min\{M_1, N_1 + N_2\} - \min\{M_1, N_2\}}{\min\{M_1, N_1 + N_2\}}\alpha_2, \end{aligned} \tag{4.12e}$$

$$d_1 + \frac{N_1 + 2N_2 - M_2}{N_2} d_2 \leq N_1 + N_2 + (N_1 - M_2)\alpha_2, \quad \textit{if } C_1 \textit{ holds} \tag{4.12f}$$

$$d_2 + \frac{N_2 + 2N_1 - M_1}{N_1} d_1 \leq N_1 + N_2 + (N_2 - M_1)\alpha_1, \quad \textit{if } C_2 \textit{ holds} \tag{4.12g}$$

where $\alpha_i \in [0, 1]$ $(i = 1, 2)$ indicates the current CSIT quality exponent corresponds to Receiver i.

Proof. The general formulation of achievability will be presented in Section 4.5, and the converse will be given in Section 4.6. For the converse, the first three inequalities correspond to the outer bounds for the case of perfect CSIT, which should also hold for our setting. Hence, it is sufficient to prove the last four bounds. Due to the symmetry property of the bounds (4.12d) and (4.12e), (4.12f) and (4.12g), it is sufficient to prove the bounds (4.12d) and (4.12f). □

Remark 4.3. *Some previous reported results can be regarded as special cases of our results: delayed CSIT case [43] for $\alpha_1 = \alpha_2 = 0$; perfect CSIT case [28] for $\alpha_1 = \alpha_2 = 1$, where the weighted sum DoF bounds* (4.12d)-(4.12g) *are inactive; hybrid CSIT (i.e., perfect CSIT for one channel and delayed CSIT for the other one) case [87] for $\alpha_1 = 1, \alpha_2 = 0$, where the bounds* (4.12e) *and* (4.12f) *are active.*

4.4 Achievability: Toy Examples

To introduce the main idea of our achievability scheme, we revisit MAT [41] and α-MAT alignment [53–55] for the case of MISO BC in Section 4.4.1, followed by an alternative way built on *block-Markov encoding and backward decoding* in Section 4.4.2, as well as some examples in Section 4.4.3 and 4.4.4 showing that block-Markov encoding allows us to balance the asymmetry both in current CSIT qualities and antenna configurations. Although MAT [41] and α-MAT alignment [53–55] appear to be conceptually different, these schemes boil down into a *single* block-Markov encoding scheme (of an infinite number of *constant-length* blocks). In fact, both schemes can be represented exactly in the same manner with different parameters.

4.4.1 MAT v.s. α-MAT Revisit

Let us take the simplest antenna configuration, i.e., $(2,1,1)$ BC, as an example. Recall that both MAT and α-MAT deliver symbol under the same structure. Specifically, in the first phase (Phase I), two independent messages w_1 and w_2 are encoded into two independent vectors $\boldsymbol{u}_1(w_1)$ and $\boldsymbol{u}_2(w_2)$ with different covariance matrices $\boldsymbol{Q}_1 \triangleq \mathbb{E}\left(\boldsymbol{u}_1\boldsymbol{u}_1^{\mathsf{H}}\right)$ and $\boldsymbol{Q}_2 \triangleq \mathbb{E}\left(\boldsymbol{u}_2\boldsymbol{u}_2^{\mathsf{H}}\right)$. The sum of these vectors are sent out, i.e.,

$$\begin{aligned} &\boldsymbol{x}[1] = \boldsymbol{u}_1 + \boldsymbol{u}_2, \\ &s.t. \begin{cases} \text{MAT:} & \boldsymbol{Q}_1 = \boldsymbol{Q}_2 = P\mathbf{I}, \\ \alpha\text{-MAT:} & \begin{cases} \boldsymbol{Q}_1 = P_1\boldsymbol{\Phi}_{\hat{h}_2} + P_2\boldsymbol{\Phi}_{\hat{h}_2^{\perp}} \\ \boldsymbol{Q}_2 = P_1\boldsymbol{\Phi}_{\hat{h}_1} + P_2\boldsymbol{\Phi}_{\hat{h}_1^{\perp}} \end{cases} \end{cases} \end{aligned} \tag{4.13}$$

where $P_1 \sim P^{1-\alpha}$, $P_2 = P - P_1 \sim P$, $\forall\, \alpha \in [0,1]$, and $\boldsymbol{\Phi}_h \triangleq \frac{\boldsymbol{h}\boldsymbol{h}^{\mathsf{H}}}{\|\boldsymbol{h}\|^2}$. Each receiver experiences some interferences caused by the symbols dedicated to the other receiver

$$\begin{cases} \eta_1 \triangleq \boldsymbol{h}_1^{\mathsf{H}}\boldsymbol{u}_2 \\ \eta_2 \triangleq \boldsymbol{h}_2^{\mathsf{H}}\boldsymbol{u}_1 \end{cases} s.t. \begin{cases} \text{MAT:} & \mathbb{E}\left(|\eta_i|^2\right) \sim P \\ \alpha\text{-MAT:} & \mathbb{E}\left(|\eta_i|^2\right) \sim P^{1-\alpha} \end{cases} \tag{4.14}$$

Then, the task of the second phase is to multicast the interferences (η_1, η_2) to *both* receivers. The main difference between the MAT and α-MAT lies in the way in which the interferences are sent. While the analog version of η_k is sent in two slots with MAT, the digitized version is sent with α-MAT instead. Note that the covariance matrices $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$, or equivalently, the spatial precoding and power allocation, of α-MAT are such that the mutual interferences (η_1, η_2) have a reduced power level $P^{1-\alpha}$. According to the rate-distortion theorem [89], each interference η_k, $k = 1, 2$, can be compressed with a source codebook of size $P^{1-\alpha}$ or $(1-\alpha)\log P$ bits into an index l_k, in such a way that the average distortion between η_k and the source codeword $\hat{\eta}_k(l_k)$ is comparable to the AWGN level [54]. Then, the index l_k is encoded with a channel codebook into a codeword $\boldsymbol{x}_c(l_k) \sim P\mathbf{I}_2$ and sent as the common message to both receivers. Thanks to the reduced range of l_k, there is still room to transmit private messages. The structure of the two slots in the second phase (Phase II) is

$$\begin{cases} \text{MAT:} & \boldsymbol{x}[2] = \boldsymbol{v}_k \eta_k, \\ \alpha\text{-MAT:} & \boldsymbol{x}[2] = \boldsymbol{x}_c(l_k) + \boldsymbol{u}_{p1} + \boldsymbol{u}_{p2} \end{cases} \tag{4.15}$$

where $k = 1, 2$, $\boldsymbol{v}_k$ is a randomly chosen vector; the covariances of the private signals $\boldsymbol{u}_{p1}$ and $\boldsymbol{u}_{p2}$ are respectively $\boldsymbol{Q}_{u_{p1}} = P^{\alpha}\boldsymbol{\Phi}_{\hat{h}_2^{\perp}}$ and $\boldsymbol{Q}_{u_{p2}} = P^{\alpha}\boldsymbol{\Phi}_{\hat{h}_1^{\perp}}$ in such a way that they are drown in the AWGN at the unintended receivers without creating noticeable interferences (at high SNR). At Receiver k, the common messages l_1 and l_2 are first decoded from the two slots in Phase II, by treating the private signal $\boldsymbol{u}_{p1}$ or $\boldsymbol{u}_{p2}$ as noise. The common messages are then used to 1) reconstruct η_1 and η_2 that will be used with the received signal in Phase I to decode w_k and recover $2-\alpha$ DoF, and 2) to reconstruct $\boldsymbol{x}_c(l_k)$ and remove it from the received signals in Phase II so as to decode the private messages and recover 2α DoF (in two slots). In the end, $2-\alpha+2\alpha = 2+\alpha$ DoF per user are achievable in three slots, yielding average DoF of $\frac{2+\alpha}{3}$ per user. The interested readers may refer to [54] for more details of α-MAT alignment.

4.4.2 An Alternative: Block-Markov Implementation

In fact, both MAT and α-MAT can be implemented in a block-Markov fashion, the concept of which is shown in Fig. 4.3 for $\alpha = 0$. The common message $\boldsymbol{x}_c(l_{b-1})$ comes from the previous block $b-1$, and $\boldsymbol{u}_k(w_{kb})$ is the new private message dedicated to Receiver k ($k = 1, 2$). Essentially, we "squeeze" the Phase II of block $b-1$ and the Phase I of block b into one single block, with proper power and rate scaling.

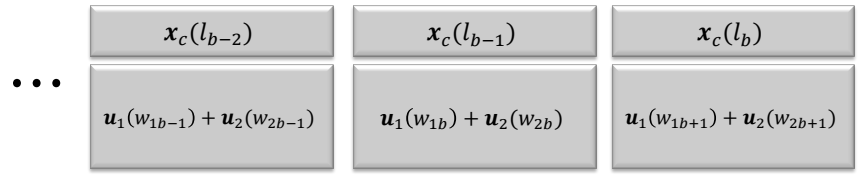


Figure 4.3: Block-Markov Encoding.

The transmission consists of B blocks of length n. For simplicity of demonstration, we set $n = 1$. In block b, the transmitter sends a mixture of two new private messages w_{1b} and w_{2b} together with one common message l_{b-1}, for $b = 1, \ldots, B$. As it will become clear, the message l_{b-1} is the compression index of the mutual interferences experienced by the receivers in the previous block $b-1$. By encoding w_{1b}, w_{2b}, and l_{b-1} into $\boldsymbol{u}_1(w_{1b})$, $\boldsymbol{u}_2(w_{2b})$, and $\boldsymbol{x}_c(l_{b-1})$, respectively, with independent channel codebooks, the transmitted signal is written as

$$\boldsymbol{x}[b] = \boldsymbol{x}_c(l_{b-1}) + \boldsymbol{u}_1(w_{1b}) + \boldsymbol{u}_2(w_{2b}), \quad b = 1, \ldots, B \tag{4.16}$$

where we set $l_0 = 1$ to initiate the transmission and $w_{1B} = w_{2B} = 1$ to end it. As before, the common message $\boldsymbol{x}_c(l_{b-1})$ is with power P, whereas the precoding in $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ is with a reduced power, parameterized by A, A', with $0 \leq A, A' \leq 1$, such that

$$\boldsymbol{Q}_1 = P^A \boldsymbol{\Phi}_{\hat{h}_2} + P^{A'} \boldsymbol{\Phi}_{\hat{h}_2^\perp}, \quad \boldsymbol{Q}_2 = P^A \boldsymbol{\Phi}_{\hat{h}_1} + P^{A'} \boldsymbol{\Phi}_{\hat{h}_1^\perp} \tag{4.17}$$

where $A \triangleq (A' - \alpha)^+$. The mutual interferences are defined similarly and their powers are now reduced

$$y_1[b] = \underbrace{\boldsymbol{h}_1^{\mathsf{H}} \boldsymbol{x}_c(l_{b-1})}_{P} + \underbrace{\boldsymbol{h}_1^{\mathsf{H}} \boldsymbol{u}_1(w_{1b})}_{P^{A'}} + \underbrace{\boldsymbol{h}_1^{\mathsf{H}} \boldsymbol{u}_2(w_{2b})}_{\eta_{1b} \sim P^A} \tag{4.18}$$

$$y_2[b] = \underbrace{\boldsymbol{h}_2^{\mathsf{H}} \boldsymbol{x}_c(l_{b-1})}_{P} + \underbrace{\boldsymbol{h}_2^{\mathsf{H}} \boldsymbol{u}_2(w_{2b})}_{P^{A'}} + \underbrace{\boldsymbol{h}_2^{\mathsf{H}} \boldsymbol{u}_1(w_{1b})}_{\eta_{2b} \sim P^A} \tag{4.19}$$

where we omit the block indices for the channel coefficients as well as the AWGN for brevity. At the end of block b, (η_{1b}, η_{2b}) are compressed with a codebook of size P^{2A} into an index $l_b \in \left\{1, \ldots, P^{2A}\right\}$. The distortion between (η_{1b}, η_{2b}) and $(\hat{\eta}_1(l_b), \hat{\eta}_2(l_b))$ is at the noise level.

At the end of B blocks, Receiver k would like to retrieve $w_{k1}, \ldots, w_{k,B-1}$. Let us focus on Receiver 1, without loss of generality. In this particular case,

l_{b-1} can be decoded at the end of block b, by treating the private signals as noise, i.e., with signal-to-interference-and-noise-ratio (SINR) level $P^{1-A'}$, for $b = 2, \ldots, B$. The correct decoding of l_{b-1} is guaranteed if the SINR can support the DoF of $2A$ for the common message, i.e.,

$$2A \leq 1 - A'. \tag{4.20}$$

Given that this condition is satisfied, $l_0, l_1, \ldots, l_{B-1}$ are available to both receivers. Therefore, η_{1b}, η_{2b}, $b = 1, \ldots, B-1$, are known, up to the noise level. To decode w_{1b}, Receiver 1 uses η_{1b}, η_{2b}, and l_{b-1} to form the following 2×2 MIMO system

$$\begin{bmatrix} y_1[b] - \boldsymbol{h}_1^{\mathsf{H}} \boldsymbol{x}_c(l_{b-1}) - \eta_{1b} \\ \eta_{2b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_1^{\mathsf{H}} \\ \boldsymbol{h}_2^{\mathsf{H}} \end{bmatrix} \boldsymbol{u}_1(w_{1b}) \tag{4.21}$$

where the equivalent channel matrix has rank 2 almost surely. This decoding strategy for the private message boils down to the backward decoding, where the mutual interferences (η_{1b}, η_{2b}) decoded in the future block are utilized in current block as side information. From the covariance matrix $\boldsymbol{Q}_1$ of $\boldsymbol{u}_1$ from (4.17), we deduce that the correct decoding of w_{1b} is guaranteed if the DoF d_{1b} of w_{1b} satisfy

$$d_{1b} \leq A + A'. \tag{4.22}$$

Combining (4.20) and (4.22), it readily follows that the optimal A' should equalize (4.20), i.e., $A'^* = \frac{1+2\alpha}{3}$. Thus, we achieve $d_{1b} = \frac{2+\alpha}{3}$. Due to the symmetry, d_{2b} has the same value. Finally, we have

$$d_k = \frac{1}{B} \sum_{b=1}^{B-1} d_{kb} = \frac{B-1}{B} \frac{2+\alpha}{3}, \quad k = 1, 2 \tag{4.23}$$

which goes to $\frac{2+\alpha}{3}$ when $B \to \infty$.

By now, we have shown that both MAT and α-MAT schemes can be interpreted under a common framework of block-Markov encoding with power allocation parameters (A, A') and that they only differ from the choice of these parameters. As we will show in the following subsections, the strength (or benefit) of the block-Markov encoding framework becomes evident in the asymmetric system setting, where the original α-MAT alignment fails to achieve the optimal DoF in general.

Table 4.1: Parameter Setting for the $(2,1,1)$ BC Case ($\alpha_1 \geq \alpha_2$)

Condition	A_1'	A_2'	Corner Point (d_1, d_2)
$2\alpha_1 - \alpha_2 \leq 1$	$A_1' = \frac{1+\alpha_1+\alpha_2}{3}$	$A_2' = \frac{1+\alpha_1+\alpha_2}{3}$	$\left(\frac{2+2\alpha_1-\alpha_2}{3}, \frac{2-\alpha_1+2\alpha_2}{3}\right)$
	$A_1' = \frac{1+\alpha_2}{2}$	$A_2' = \alpha_1$	$(1, \alpha_1)$
$2\alpha_1 - \alpha_2 > 1$	$A_1' = \frac{1+\alpha_2}{2}$	$A_2' = \frac{1+\alpha_2}{2}$	$(1, \frac{1+\alpha_2}{2})$
-	$A_1' = \alpha_2$	$A_2' = \frac{1+\alpha_1}{2}$	$(\alpha_2, 1)$

4.4.3 Asymmetry in Current CSIT Qualities

Let us consider again the MISO BC case but assume now that the CSIT qualities of two channels are different, i.e., $\alpha_1 \neq \alpha_2$, where α_k $(k = 1, 2)$ is for Receiver k. The signal model is in the exact same form as in (4.16) with a more general precoding, parameterized by A_k, A_k', with $0 \leq A_k, A_k' \leq 1$, such that

$$\boldsymbol{Q}_1 = P^{A_1}\boldsymbol{\Phi}_{\hat{h}_2} + P^{A_1'}\boldsymbol{\Phi}_{\hat{h}_2^{\perp}}, \quad \boldsymbol{Q}_2 = P^{A_2}\boldsymbol{\Phi}_{\hat{h}_1} + P^{A_2'}\boldsymbol{\Phi}_{\hat{h}_1^{\perp}} \tag{4.24}$$

where $A_k \triangleq (A_k' - \alpha_j)^+$, $j \neq k \in \{1, 2\}$. Following the same footsteps as in the symmetric case, it is readily shown that $\eta_{1b} \sim P^{A_2}$ and $\eta_{2b} \sim P^{A_1}$ and that (η_{1b}, η_{2b}) can be compressed up to the noise level with a source codebook of size $P^{A_1+A_2}$. The decoding at both receivers is the same as before. To decode the common message l_{b-1} by treating the private signals as noise, since the SINR is $P^{1-A_1'}$ at Receiver 1 and $P^{1-A_2'}$ at Receiver 2, the DoF of the common message should satisfy

$$A_1 + A_2 \leq \min\{1 - A_1', 1 - A_2'\}. \tag{4.25}$$

Using the common messages l_b and l_{b-1} as side information, w_{1b} and w_{2b} can be decoded at the respective receivers if

$$d_{1b} \leq A_1 + A_1' \quad \text{and} \quad d_{2b} \leq A_2 + A_2'. \tag{4.26}$$

By carefully selecting the parameters A_1' and A_2', all corner points of the DoF outer bound can be achieved, as shown in Table 4.1 on the top of this page where the condition is to make sure the corner points exist. Note that the corner point $(\alpha_2, 1)$ always exists as long as $\alpha_1 \geq \alpha_2$.

4.4.4 Asymmetry in Antenna Configurations

We use the $(4, 3, 2)$ MIMO BC case to show that the block-Markov encoding can achieve the optimal performance in asymmetric antenna settings. Recall

that, in the previous subsections, the backward decoding is performed to decode the private messages, and that the common messages can be decoded *block by block*. In this case, however, we also need backward decoding to decode the common messages as well.

The same transmission signal model (4.16) is used here, with the following precoding, parameterized by A_k and A'_k, $k = 1, 2$, $0 \leq A_k \leq A'_k \leq 1$:

$$\boldsymbol{Q}_1 = P^{A_1}\boldsymbol{\Phi}_{\hat{H}_2} + P^{A'_1}\boldsymbol{\Phi}_{\hat{H}_2^{\perp}}, \quad \boldsymbol{Q}_2 = P^{A_2}\boldsymbol{\Phi}_{\hat{H}_1} + P^{A'_2}\boldsymbol{\Phi}_{\hat{H}_1^{\perp}} \tag{4.27}$$

where A_k, $k \neq j \in \{1, 2\}$, is defined as

$$A_k \triangleq \begin{cases} (A'_k - \alpha_j)^+, & d_k \leq 4 - N_j\alpha_j, \\ \frac{d_k - (4 - N_j)}{N_j}, & d_k > 4 - N_j\alpha_j \end{cases} \tag{4.28}$$

with $d_k \in \mathbb{R}_+$ being the achievable DoF associated with Receiver k. It is readily verified that $A'_k - \alpha_j \leq A_k \leq A'_k$ is always true, such that the created interference at intended Receiver j is of power level A_k, and the desired signal at Receiver k is of level A'_k.

We recall that the common message $\boldsymbol{x}_c(l_{b-1})$ is transmitted with power P and that the ranks of $\boldsymbol{\Phi}_{\hat{H}_2}$, $\boldsymbol{\Phi}_{\hat{H}_2^{\perp}}$, $\boldsymbol{\Phi}_{\hat{H}_1}$, and $\boldsymbol{\Phi}_{\hat{H}_1^{\perp}}$ are respectively 2, 2, 3, and 1, almost surely. The received signals are now vectors given by

$$\boldsymbol{y}_1[b] = \underbrace{\boldsymbol{H}_1\boldsymbol{x}_c(l_{b-1})}_{P\mathbf{I}_3} + \boldsymbol{H}_1\boldsymbol{u}_1(w_{1b}) + \underbrace{\boldsymbol{H}_1\boldsymbol{u}_2(w_{2b})}_{\boldsymbol{\eta}_{1b} \sim P^{A_2}\mathbf{I}_3}, \tag{4.29}$$

$$\boldsymbol{y}_2[b] = \underbrace{\boldsymbol{H}_2\boldsymbol{x}_c(l_{b-1})}_{P\mathbf{I}_2} + \boldsymbol{H}_2\boldsymbol{u}_2(w_{2b}) + \underbrace{\boldsymbol{H}_2\boldsymbol{u}_1(w_{1b})}_{\boldsymbol{\eta}_{2b} \sim P^{A_1}\mathbf{I}_2}. \tag{4.30}$$

Following the same footsteps as in the single receive antenna case, it is readily shown that $(\boldsymbol{\eta}_{1b}, \boldsymbol{\eta}_{2b})$ can be compressed up to the noise level with a source codebook of size $P^{2A_1+3A_2}$. For convenience, let us define

$$d_\eta \triangleq 2A_1 + 3A_2. \tag{4.31}$$

Unlike the MISO case where the common messages can be decoded independently in each block without loss of optimality, backward decoding is required to *jointly* decode the common and private messages in the general MIMO case, in order to achieve the optimal DoF. As we will see later on, the common rate can be improved with backward decoding in general. The decoding starts at block B. Since w_{1B} and w_{2B} are both known, the private signals can be removed from the received signals $\boldsymbol{y}_1[B]$ and $\boldsymbol{y}_2[B]$. The

common message l_{B-1} can be decoded at both receivers if $d_\eta \leq 2$. At block b, for $b = B-1, \ldots, 2$, assuming l_b is known perfectly from the decoding of block $b+1$, $\boldsymbol{\eta}_{1b}$ and $\boldsymbol{\eta}_{2b}$ can be reconstructed up to the noise level. The following MIMO system can be obtained

$$\begin{bmatrix} \boldsymbol{y}_1[b] - \boldsymbol{\eta}_{1b} \\ \boldsymbol{\eta}_{2b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_1 \\ 0 \end{bmatrix} \boldsymbol{x}_c(l_{b-1}) + \begin{bmatrix} \boldsymbol{H}_1 \\ \boldsymbol{H}_2 \end{bmatrix} \boldsymbol{u}_1(w_{1b}). \tag{4.32}$$

Note that this is a multiple-access channel (MAC) from which l_{b-1} and w_{1b} can be correctly decoded if the rate pair lies within the following region

$$d_\eta \leq 3 \tag{4.33}$$
$$d_{1b} \leq 2A_1 + 2A_1' \tag{4.34}$$
$$d_\eta + d_{1b} \leq 3 + 2A_1, \tag{4.35}$$

whose general proof is provided in Appendix 4.8.1. Let us set d_{1b} to equalize (4.34). Then, (4.33) and (4.35) imply $d_\eta \leq 3 - 2A_1'$. Similar analysis on Receiver 2 will lead to $d_\eta \leq 2 - A_2'$, by setting $d_{2b} = A_2' + 3A_2$. Therefore, from (4.31), we obtain the following constraint

$$2A_1 + 3A_2 \leq \min\left\{3 - 2A_1', 2 - A_2'\right\} \tag{4.36}$$

to achieve any (d_{1b}, d_{2b}) such that

$$d_{1b} \leq 2A_1 + 2A_1' \quad \text{and} \quad d_{2b} \leq A_2' + 3A_2. \tag{4.37}$$

By letting $B \to \infty$, $d_1 = 2A_1 + 2A_1'$ and $d_2 = A_2' + 3A_2$ can be achieved for any $A_1', A_2' \leq 1$ given the definition of (A_1, A_2) in (4.28), as long as (4.36) is satisfied. We can show that, by properly choosing (A_1', A_2'), all the corner points given by the outer bound can be achieved. For example, by setting $\alpha_1 = \alpha_2 = \alpha$, the values (A_1', A_2') that achieve the corner points are illustrated in Table 4.2 on the top of the next page. Note that $(\frac{12}{5}, \frac{4}{5} + \alpha)$ exists only when $\alpha \leq \frac{4}{5}$, whereas $(3\alpha, 4 - 3\alpha)$ and $(4 - 2\alpha, 2\alpha)$ exist only when $\alpha > \frac{4}{5}$.

4.5 Achievability: the General Formulation

As aforementioned, the key ingredients of the achievability scheme consist of:

- *block-Markov encoding* with a *constant* block length: the fresh messages in the current block and the interferences created in the past blocks are encoded together with the proper rate splitting and power scaling;

Table 4.2: Parameter Setting for the $(4,3,2)$ BC Case with $\alpha_1 = \alpha_2 = \alpha$.

Corner Point (d_1, d_2)	Cond.	(A_1', A_2')	(A_1, A_2)	d_η
$(3, \alpha)$	$\alpha \leq \frac{1}{2}$	$(\frac{3+2\alpha}{4}, \alpha)$	$(\frac{3-2\alpha}{4}, 0)$	$\frac{3-2\alpha}{2}$
	$\alpha > \frac{1}{2}$	$(1, \alpha)$	$(\frac{1}{2}, 0)$	1
$(2\alpha, 2)$	$\alpha \leq \frac{2}{3}$	$(\alpha, \frac{2+3\alpha}{4})$	$(0, \frac{2-\alpha}{4})$	$\frac{6-3\alpha}{4}$
	$\alpha > \frac{2}{3}$	$(\alpha, 1)$	$(0, \frac{1}{3})$	1
$(\frac{12}{5}, \frac{4}{5} + \alpha)$	$\alpha \leq \frac{4}{5}$	$(\frac{3}{5} + \frac{1}{2}\alpha, \frac{1}{5} + \alpha)$	$(\frac{3}{5} - \frac{1}{2}\alpha, \frac{1}{5})$	$\frac{9}{5} - \alpha$
$(3\alpha, 4 - 3\alpha)$	$\alpha > \frac{4}{5}$	$(1, 1)$	$(\frac{3\alpha-2}{2}, 1 - \alpha)$	1
$(4 - 2\alpha, 2\alpha)$	$\alpha > \frac{4}{5}$	$(1, 1)$	$(1 - \alpha, \frac{2\alpha-1}{3})$	1

- *spatial precoding* with imperfect current CSIT: with proper power allocation over the range and null spaces of the inaccurate current channel, the interference power at unintended receiver can be reduced as compared to that without any CSIT;

- *interference quantization*: instead of forwarding the overheard interference directly in an analog way as done in pure delayed CSIT scenario, the reduced-power interferences are compressed first with a reduced number of bits, and forwarded in a digital fashion with lower rate;

- *backward decoding*: the messages are decoded from the last block to the first one, where in each block the messages are decoded with the aid of side information provided by the blocks in the future.

In the following, the general achievability scheme will be described in detail for BC and IC respectively.

4.5.1 Broadcast Channels

First of all, we notice that the region (4.10) given in Theorem 4.1 does not depend on M (resp. N_k) when $M > N_1 + N_2$ (resp. $N_k > M$). Therefore, it is sufficient to prove the achievability for the case $M \leq N_1 + N_2$ and $N_k \leq M$. And the achievability for the other cases can be inferred by simply switching off the additional transmit/receive antennas. Thus, it yields

$$\begin{aligned} M &= \min\{M, N_1 + N_2\}, \\ N_k &= \min\{M, N_k\}, \quad k = 1, 2. \end{aligned} \tag{4.38}$$

Block-Markov encoding

The block-Markov encoding has the same structure as before, namely,

$$\boldsymbol{x}[b] = \boldsymbol{x}_c(l_{b-1}) + \boldsymbol{u}_1(w_{1b}) + \boldsymbol{u}_2(w_{2b}), \quad b = 1, \dots, B \tag{4.39}$$

where we recall that we set $l_0 = 1$ to initiate the transmission and $w_{1B} = w_{2B} = 1$ to end it.

Spatial precoding

Both $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{C}^{M\times 1}$ are precoded signals of M streams, such that

$$\boldsymbol{Q}_1 = P^{A_1}\boldsymbol{\Phi}_{\hat{H}_2} + P^{A_1'}\boldsymbol{\Phi}_{\hat{H}_2^\perp}, \quad \boldsymbol{Q}_2 = P^{A_2}\boldsymbol{\Phi}_{\hat{H}_1} + P^{A_2'}\boldsymbol{\Phi}_{\hat{H}_1^\perp} \tag{4.40}$$

where the rank of $\boldsymbol{\Phi}_{\hat{H}_k}$ is N_k whereas the rank of $\boldsymbol{\Phi}_{\hat{H}_k^\perp}$ is $M - N_k$, $k = 1, 2$. In other words, for Receiver k, N_j streams are sent in the subspace of the unintended Receiver j with power level A_k and the other $M - N_j$ streams are sent in the null space of Receiver j with power level A_k', where (A_k, A_k') satisfies

$$0 \le A_k \le A_k' \le 1 \quad \text{and} \quad A_k \ge A_k' - \alpha_j \tag{4.41}$$

for $j \neq k \in \{1, 2\}$. Note that the above condition guarantees that the interference at Receiver j has power level A_k and the desired signal at Receiver k is of power level A_k'.

Interference quantization

Recall that the common message $\boldsymbol{x}_c(l_{b-1})$ is sent with power P. The received signals in block b are given by

$$\boldsymbol{y}_1[b] = \underbrace{\boldsymbol{H}_1\boldsymbol{x}_c(l_{b-1})}_{P\mathbf{I}_{N_1}} + \boldsymbol{H}_1\boldsymbol{u}_1(w_{1b}) + \underbrace{\boldsymbol{H}_1\boldsymbol{u}_2(w_{2b})}_{\boldsymbol{\eta}_{1b}\sim P^{A_2}\mathbf{I}_{N_1}}, \tag{4.42}$$

$$\boldsymbol{y}_2[b] = \underbrace{\boldsymbol{H}_2\boldsymbol{x}_c(l_{b-1})}_{P\mathbf{I}_{N_2}} + \boldsymbol{H}_2\boldsymbol{u}_2(w_{2b}) + \underbrace{\boldsymbol{H}_2\boldsymbol{u}_1(w_{1b})}_{\boldsymbol{\eta}_{2b}\sim P^{A_1}\mathbf{I}_{N_2}}. \tag{4.43}$$

It is readily shown that $(\boldsymbol{\eta}_{1b}, \boldsymbol{\eta}_{2b})$ can be compressed up to the noise level with a source codebook of size $P^{N_2A_1+N_1A_2}$ into an index l_b. For convenience, let us define

$$d_{\eta_1} \triangleq N_1A_2, \; d_{\eta_2} \triangleq N_2A_1, \text{ and } d_\eta \triangleq d_{\eta_1} + d_{\eta_2}. \tag{4.44}$$

Backward decoding

The decoding starts at block B. Since w_{1B} and w_{2B} are both known, the private signals can be removed from the received signals $\boldsymbol{y}_1[B]$ and $\boldsymbol{y}_2[B]$. The common message l_{B-1} can be decoded at both receivers if $d_\eta \leq \min\{N_1, N_2\}$. At block b, assuming l_b is known perfectly from the decoding of block $b+1$, $\boldsymbol{\eta}_{1b}$ and $\boldsymbol{\eta}_{2b}$ can be reconstructed up to the noise level, for $b = B-1, \ldots, 2$. The following MIMO system can be obtained at Receiver k, $k = 1, 2$

$$\begin{bmatrix} \boldsymbol{y}_k[b] - \boldsymbol{\eta}_{kb} \\ \boldsymbol{\eta}_{jb} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_k \\ 0 \end{bmatrix} \boldsymbol{x}_c(l_{b-1}) + \begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{H}_j \end{bmatrix} \boldsymbol{u}_k(w_{kb}) \tag{4.45}$$

for $j \neq k \in \{1, 2\}$. Since the common message l_{b-1} and the private message w_{kb} are both desired by Receiver k, this system corresponds to a multiple-access channel (MAC). As formally proved in Appendix 4.8.1, Receiver k can decode correctly both messages if the following conditions are satisfied.

$$d_\eta \leq N_k \tag{4.46}$$

$$d_{kb} \leq N_j A_k + (M - N_j) A'_k \tag{4.47}$$

$$d_\eta + d_{kb} \leq N_k + N_j A_k. \tag{4.48}$$

Let us choose d_{kb} to be equal to the right hand side of (4.47) for $k = 1, 2$ and $b = 1, .., B-1$. Then, the equality in (4.47) together with (4.44), (4.46), (4.48) implies, when letting $B \to \infty$, the following lemma.

Lemma 4.1 (decodability condition for BC)**.** *Let us define*

$$\begin{aligned} \mathcal{A}_{\text{BC}} \triangleq \big\{ &(A_1, A'_1, A_2, A'_2) \mid A_k, A'_k \in [0, 1], \\ &A'_k - \alpha_j \leq A_k \leq A'_k, \forall\, k \neq j \in \{1, 2\} \big\} \end{aligned} \tag{4.49}$$

$$\mathcal{D}_{\text{BC}} \triangleq \{(d_1, d_2) \mid d_k \in [0, N_k], \quad \forall\, k \in \{1, 2\}\} \tag{4.50}$$

and

$$f_{A\text{-}d} \ : \mathcal{A}_{\text{BC}} \to \mathcal{D}_{\text{BC}} \tag{4.51}$$

$$(A_k, A'_k) \mapsto d_k \triangleq N_j A_k + (M - N_j) A'_k, \forall k \neq j \in \{1, 2\}. \tag{4.52}$$

Then $(d_1, d_2) = f_{A\text{-}d}(\boldsymbol{A})$, for some $\boldsymbol{A} \in \mathcal{A}_{\text{BC}}$, is achievable with the proposed scheme, if

$$d_{\eta_1} + d_1 \leq N_1, \tag{4.53}$$

$$d_{\eta_2} + d_2 \leq N_2. \tag{4.54}$$

where we recall $d_{\eta_1} \triangleq N_1 A_2$ and $d_{\eta_2} \triangleq N_2 A_1$.

Remark 4.4. *In the above lemma, d_{η_k} can be interpreted as the degrees of freedom occupied by the interference at Receiver k. Therefore,* (4.53) *and* (4.54) *are clearly outer bounds for any transmission strategies, i.e., the sum of the dimension of the useful signal and the dimension of the interference signal at the receiver side cannot exceed the total dimension of the signal space. These bounds are in general not tight except for special cases such as the "strong interference" regime where interference can be decoded completely and removed or the "weak interference" regime where the interference can be treated as noise while the useful signal power dominates the received power. Remarkably, the proposed scheme achieves these outer bounds. This is due to two of the main ingredients of our scheme, namely, the block-Markov encoding and interference quantization. The block-Markov encoding places the digitized interference in the "upper level" of the signal space (with full power) and thus "pushes" the channel into the "strong interference" regime in which the digitized interference can be decoded thanks to the structure brought by the interference quantization.*

Definition 4.3 (achievable region for BC)**.** *Let us define*

$$\mathcal{I}_A^{\mathrm{BC}} \triangleq \left\{ (A_1, A_1', A_2, A_2') \in \mathcal{A}_{\mathrm{BC}} \;\middle|\; \begin{array}{l} (d_1, d_2) = f_{A\text{-}d}(A_1, A_1', A_2, A_2'), \\ \frac{d_k}{N_k} \leq 1 - A_j, \quad k \neq j \in \{1, 2\} \end{array} \right\} \tag{4.55}$$

and the achievable DoF region of the proposed scheme

$$\mathcal{I}_d^{\mathrm{BC}} \triangleq f_{A\text{-}d}(\mathcal{I}_A^{\mathrm{BC}}) \triangleq \left\{ (d_1, d_2) \;\middle|\; \begin{array}{l} (d_1, d_2) = f_{A\text{-}d}(A_1, A_1', A_2, A_2'), \\ (A_1, A_1', A_2, A_2') \in \mathcal{A}_{\mathrm{BC}}, \\ \frac{d_k}{N_k} \leq 1 - A_j, \quad k \neq j \in \{1, 2\} \end{array} \right\}. \tag{4.56}$$

Achievability analysis

In the following, we would like to show that any pair (d_1, d_2) in the outer bound region defined by (4.10), hereafter referred to as $\mathcal{O}_d^{\mathrm{BC}}$, can be achieved by the proposed strategy. Therefore, it is sufficient to show that $\mathcal{O}_d^{\mathrm{BC}} \subseteq \mathcal{I}_d^{\mathrm{BC}}$. The main idea is as follows. If there exists a function

$$f_{d\text{-}A} \;:\; \mathcal{O}_d^{\mathrm{BC}} \to \mathcal{A}_{\mathrm{BC}} \tag{4.57}$$

such that

$$(d_1, d_2) = f_{A\text{-}d}(f_{d\text{-}A}(d_1, d_2)), \quad \text{and} \tag{4.58}$$

$$f_{d\text{-}A}(d_1, d_2) \in \mathcal{I}_A^{\mathrm{BC}}, \tag{4.59}$$

then for every $(d_1, d_2) \in \mathcal{O}_d^{\mathrm{BC}}$ we can use the power allocation $(A_1, A_1', A_2, A_2') = f_{d\text{-}A}(d_1, d_2)$ on the proposed scheme to achieve it, i.e.,

$$\mathcal{O}_d^{\mathrm{BC}} = f_{A\text{-}d}(f_{d\text{-}A}(\mathcal{O}_d^{\mathrm{BC}})) \subseteq f_{A\text{-}d}(\mathcal{I}_A^{\mathrm{BC}}) = \mathcal{I}_d^{\mathrm{BC}} \tag{4.60}$$

from which the achievability is proved. Now, we define formally the power allocation function.

Definition 4.4 (power allocation for BC)**.** *Let us define* $f_{d\text{-}A} : \mathcal{O}_d^{\mathrm{BC}} \to \mathcal{A}_{\mathrm{BC}}$*:*

$$(d_1, d_2) \mapsto (A_1, A_1') \triangleq f_1(d_1),\ (A_2, A_2') \triangleq f_2(d_2) \tag{4.61}$$

where f_k*,* $j \neq k \in \{1, 2\}$*, is specified as below.*

- *When* $M = N_j$*:* $A_k' = A_k = \frac{d_k}{M}$*;*
- *When* $M > N_j$ *and* $d_k < M - N_j\alpha_j$*:* $A_k = (A_k' - \alpha_j)^+$*, and thus*

$$A_k' = \begin{cases} \frac{d_k}{M - N_j}, & \text{if } d_k < (M - N_j)\alpha_j; \\ \frac{d_k + N_j\alpha_j}{M}, & \text{otherwise}; \end{cases} \tag{4.62}$$

- *When* $M > N_j$ *and* $d_k \geq M - N_j\alpha_j$*:* $A_k' = 1$*, and thus* $A_k = \frac{d_k - (M - N_j)}{N_j}$*.*

It is readily shown that, for any $(d_1, d_2) \in \mathcal{O}_d^{\mathrm{BC}}$, the resulting power allocation always lies in $\mathcal{A}_{\mathrm{BC}}$ as defined in (4.49) and that (4.58) is always satisfied. It remains to show that (4.59) holds as well, i.e., the decodability condition in (4.55) is satisfied. To that end, for any $(d_1, d_2) \in \mathcal{O}_d^{\mathrm{BC}}$, we first define $(A_1, A_1', A_2, A_2') \triangleq f_{d\text{-}A}(d_1, d_2)$ which implies $d_j = N_k A_j + (M - N_k)A_j'$, $j \neq k \in \{1, 2\}$. Applying this equality on the constraints in the outer bound $\mathcal{O}_d^{\mathrm{BC}}$ in (4.10), we have

$$\frac{d_k}{N_k} \leq \frac{M - (M - N_k)A_j'}{N_k} - A_j, \tag{4.63}$$

$$\frac{d_k}{N_k} \leq 1 - \left[\frac{(M - N_k)(A_j' - \alpha_k) + N_k A_j}{M}\right]^+ \tag{4.64}$$

for $k \neq j \in \{1, 2\}$, where the first one is from the sum rate constraint (4.10c) whereas the second one is from the rest of the constraints in (4.10). The final step is to show that either of (4.63) and (4.64) implies the last constraint in (4.55):

- When $M = N_k$, (4.64) is identical to the last constraint in (4.55);
- When $M > N_k$ and $d_j \geq M - N_k\alpha_k$, we have $A'_j = 1$ according to the mapping $f_{d\text{-}A}$ defined in Definition 4.4. Hence, (4.63) is identical to the last constraint in (4.55);
- When $M > N_k$ and $d_j < M - N_k\alpha_k$, we have $A_j = (A'_j - \alpha_k)^+$ according to Definition 4.4. Hence,

$$\left[\frac{(M-N_k)(A'_j-\alpha_k)+N_kA_j}{M}\right]^+ \geq A_j \tag{4.65}$$

  with which (4.64) implies the last constraint in (4.55).

By now, we have proved the achievability through the existence of a proper power allocation function such that (4.58) and (4.59) are satisfied for every pair (d_1, d_2) in the outer bound.

4.5.2 Interference Channels

The proposed scheme for MIMO IC is similar to that for BC, with the differences that (a) the interferences can only be reconstructed at the transmitter from which the symbols are sent, and (b) antenna configuration does matter at both transmitters and receivers. Further, as with the broadcast channel, we notice that the region (4.12) given in Theorem 4.2 does not depend on M_k (resp. N_k) when $M_k > N_1 + N_2$ (resp. $N_k > M_1 + M_2$), $k = 1, 2$. Therefore, it is sufficient to prove the achievability for the case $M_k \leq N_1 + N_2$ and $N_k \leq M_1 + M_2$, $k = 1, 2$, since the achievability for the other cases can be inferred by simply switching off the additional transmit/receive antennas. Thus, it yields

$$\begin{aligned} M_k &= \min\{M_k, N_1 + N_2\}, \\ N_k &= \min\{N_k, M_1 + M_2\}, \quad k = 1, 2. \end{aligned} \tag{4.66}$$

We also define for notational convenience

$$N'_1 \triangleq \min\{N_1, M_2\}, \quad N'_2 \triangleq \min\{N_2, M_1\}. \tag{4.67}$$

Block-Markov encoding

The block-Markov encoding is done independently at both transmitters

$$\boldsymbol{x}_1[b] = \boldsymbol{x}_{1c}(l_{1,b-1}) + \boldsymbol{u}_1(w_{1b}), \tag{4.68}$$

$$\boldsymbol{x}_2[b] = \boldsymbol{x}_{2c}(l_{2,b-1}) + \boldsymbol{u}_2(w_{2b}), \quad b = 1, \dots, B \tag{4.69}$$

where we set $l_{1,0} = l_{2,0} = 1$ to initiate the transmission and $w_{1B} = w_{2B} = 1$ to end it.

Spatial precoding

The signal $\boldsymbol{u}_k \in \mathbb{C}^{M_k \times 1}$, $k = 1, 2$, is precoded signal of M_k streams, such that

$$\boldsymbol{Q}_1 = P^{A_1} \boldsymbol{\Phi}_{\hat{H}_{21}} + P^{A_1'} \boldsymbol{\Phi}_{\hat{H}_{21}^{\perp 1}} + P^{A_1''} \boldsymbol{\Phi}_{\hat{H}_{21}^{\perp 2}}, \tag{4.70}$$

$$\boldsymbol{Q}_2 = P^{A_2} \boldsymbol{\Phi}_{\hat{H}_{12}} + P^{A_2'} \boldsymbol{\Phi}_{\hat{H}_{12}^{\perp 1}} + P^{A_2''} \boldsymbol{\Phi}_{\hat{H}_{12}^{\perp 2}} \tag{4.71}$$

where we use $\hat{\boldsymbol{H}}_{jk}^{\perp 1}$ (resp. $\hat{\boldsymbol{H}}_{jk}^{\perp 2}$) to denote any matrix spanning the $(M_k - N_j' - \xi_k)$-dimensional (resp. ξ_k-dimensional) subspace of the null space of $\hat{\boldsymbol{H}}_{jk}$ where ξ_k will be specified later on. Therefore, the rank of $\boldsymbol{\Phi}_{\hat{H}_{jk}}$ is N_j' whereas the rank of $\boldsymbol{\Phi}_{\hat{H}_{jk}^{\perp 1}}$ and $\boldsymbol{\Phi}_{\hat{H}_{jk}^{\perp 2}}$ are respectively $M_k - N_j' - \xi_k$ and ξ_k, $k = 1, 2$. The power levels (A_k, A_k', A_k'') satisfy

$$\begin{gathered} A_k, A_k', A_k'' \in [0, 1], \\ A_k \leq A_k', \ A_k'' \leq A_k', \quad \text{and} \quad A_k \geq A_k' - \alpha_j \end{gathered} \tag{4.72}$$

for $j \neq k \in \{1, 2\}$. Note that the above condition guarantees that the interference at Receiver j has power level A_k and the desired signal at Receiver k at power level A_k'.

Interference quantization

Recall that the common messages $\boldsymbol{x}_{1c}(l_{1,b-1})$ and $\boldsymbol{x}_{2c}(l_{2,b-1})$ are sent with power P. The received signals in block b are given by

$$\begin{aligned} \boldsymbol{y}_1[b] = & \underbrace{\boldsymbol{H}_{11}\boldsymbol{x}_{1c}(l_{1,b-1}) + \boldsymbol{H}_{12}\boldsymbol{x}_{2c}(l_{2,b-1})}_{P\mathbf{I}_{N_1}} \\ & + \boldsymbol{H}_{11}\boldsymbol{u}_1(w_{1b}) + \underbrace{\boldsymbol{H}_{12}\boldsymbol{u}_2(w_{2b})}_{\boldsymbol{\eta}_{1b} \sim P^{A_2}\mathbf{I}_{N_1'}}, \end{aligned} \tag{4.73}$$

$$\begin{aligned} \boldsymbol{y}_2[b] = & \underbrace{\boldsymbol{H}_{22}\boldsymbol{x}_{2c}(l_{2,b-1}) + \boldsymbol{H}_{21}\boldsymbol{x}_{1c}(l_{1,b-1})}_{P\mathbf{I}_{N_2}} \\ & + \boldsymbol{H}_{22}\boldsymbol{u}_2(w_{2b}) + \underbrace{\boldsymbol{H}_{21}\boldsymbol{u}_1(w_{1b})}_{\boldsymbol{\eta}_{2b} \sim P^{A_1}\mathbf{I}_{N_2'}}. \end{aligned} \tag{4.74}$$

It is readily shown that $\boldsymbol{\eta}_{1b}$ and $\boldsymbol{\eta}_{2b}$ can be compressed *separately* up to the noise level with two independent source codebooks of size $P^{N_1'A_2}$ and $P^{N_2'A_1}$, into indices $l_{2,b}$ and $l_{1,b}$, respectively. For convenience, let us define

$$d_{\eta_1} \triangleq N_1'A_2,\ d_{\eta_2} \triangleq N_2'A_1,\ \text{and}\ d_{\eta} \triangleq d_{\eta_1} + d_{\eta_2}. \tag{4.75}$$

Backward decoding

The decoding starts at block B. Since w_{1B} and w_{2B} are both known, the private signals can be removed from the received signals $\boldsymbol{y}_1[B]$ and $\boldsymbol{y}_2[B]$. The common messages $l_{1,B-1}$ and $l_{2,B-1}$ can be decoded at both receivers if

$$d_{\eta_k} \leq \min\left\{M_j, N_1, N_2\right\}, \tag{4.76}$$

$$d_{\eta_1} + d_{\eta_2} \leq \min\left\{N_1, N_2\right\}, \tag{4.77}$$

i.e., the common rate pair should lie within the intersection of MAC regions at both receivers for the common messages. At block b, assuming both $l_{1,b}$ and $l_{2,b}$ are known perfectly from the decoding of block $b+1$, $\boldsymbol{\eta}_{1b}$ and $\boldsymbol{\eta}_{2b}$ can be reconstructed up to the noise level, for $b = B-1, \ldots, 2$. The following MIMO system can be obtained at Receiver k

$$\begin{bmatrix}\boldsymbol{y}_k[b] - \boldsymbol{\eta}_{kb}\\ \boldsymbol{\eta}_{jb}\end{bmatrix} = \begin{bmatrix}\boldsymbol{H}_{kk}\\ 0\end{bmatrix}\boldsymbol{x}_{kc}(l_{k,b-1}) + \begin{bmatrix}\boldsymbol{H}_{kj}\\ 0\end{bmatrix}\boldsymbol{x}_{jc}(l_{j,b-1}) + \begin{bmatrix}\boldsymbol{H}_{kk}\\ \boldsymbol{H}_{jk}\end{bmatrix}\boldsymbol{u}_k(w_{kb}) \tag{4.78}$$

for $j \neq k \in \{1,2\}$. Note that this system corresponds to a multiple-access channel from which the three independent messages $l_{1,b-1}$, $l_{2,b-1}$, and w_{kb} are to be decoded. It will be shown in the Appendix 4.8.1 that the three messages can be correctly decoded if the DoF quadruple $(d_{\eta_1}, d_{\eta_2}, d_{1b}, d_{2b})$ lies within the following region

$$d_{kb} \leq N_j'A_k + (M_k - N_j' - \xi_k)A_k' + \xi_k A_k'' \tag{4.79}$$

$$d_{\eta_k} \leq \min\left\{M_j, N_1, N_2\right\} \tag{4.80}$$

$$d_{\eta_1} + d_{\eta_2} \leq \min\left\{N_1, N_2\right\} \tag{4.81}$$

$$d_{\eta_k} + d_{kb} \leq N_k' + \min\left\{M_k - N_j', N_k - N_k'\right\}A_k' + N_j'A_k \tag{4.82}$$

$$d_{\eta_j} + d_{kb} \leq \min\left\{M_k, N_k\right\} + N_j'A_k \tag{4.83}$$

$$d_{\eta_1} + d_{\eta_2} + d_{kb} \leq N_k + N_j'A_k. \tag{4.84}$$

Now, let us fix

$$d_{kb} \triangleq N'_j A_k + (M_k - N'_j - \xi_k)A'_k + \xi_k A''_k \tag{4.85}$$

$$d_{\eta_j} \triangleq N'_j A_k \tag{4.86}$$

from which we can reduce the region defined by (4.79)-(4.84). First, we remove (4.79) that is implied by (4.85). Second, (4.80) is not active as it is implied by (4.86) and (4.81). Third, (4.81) is implied by (4.84) and (4.85). Finally, from (4.86), (4.83) is equivalent to $d_{kb} \leq \min\{M_k, N_k\}$ that is implied by (4.85). Therefore, by letting $B \to \infty$, we have the following counterpart of Lemma 4.1 for interference channels.

Lemma 4.2 (decodability condition for IC)**.** *Let us define*

$$\mathcal{A}_{\text{IC}} \triangleq \left\{ (A_1, A'_1, A''_1, A_2, A'_2, A''_2) \;\middle|\; \begin{array}{l} A_k, A'_k, A''_k \in [0,1] \\ A'_k - \alpha_j \leq A_k \leq A'_k, \;\; A''_k \leq A'_k, \\ \xi_k A''_k \leq N'_k(1 - A_j), \quad k \neq j \in \{1,2\} \end{array} \right\} \tag{4.87}$$

$$\mathcal{D}_{\text{IC}} \triangleq \left\{ (d_1, d_2) \mid d_k \in [0, \min\{M_k, N_k\}], \quad \forall\, k \in \{1,2\} \right\} \tag{4.88}$$

and

$$f_{A\text{-}d} \;:\; \mathcal{A}_{\text{IC}} \to \mathcal{D}_{\text{IC}} \tag{4.89}$$

$$(A_k, A'_k, A''_k) \mapsto d_k \triangleq N'_j A_k + (M_k - N'_j - \xi_k)A'_k + \xi_k A''_k, \quad \forall k \neq j \in \{1,2\} \tag{4.90}$$

where

$$\xi_k \triangleq \begin{cases} (M_k - N'_j)^+ - (N_k - N'_k)^+, & \text{if } C_k \text{ holds} \\ 0, & \text{otherwise.} \end{cases} \tag{4.91}$$

Then $(d_1, d_2) = f_{A\text{-}d}(\boldsymbol{A})$, for some $\boldsymbol{A} \in \mathcal{A}_{\text{IC}}$, is achievable with the proposed scheme, if

$$d_{\eta_1} + d_1 \leq N_1, \tag{4.92}$$

$$d_{\eta_2} + d_2 \leq N_2. \tag{4.93}$$

where we recall $d_{\eta_1} \triangleq N'_1 A_2$ and $d_{\eta_2} \triangleq N'_2 A_1$.

Proof. It has been shown that with (4.86) and (4.90), only (4.82) and (4.84) are active. With ξ_k defined in (4.91), we can verify that $M_k - N'_j - \xi_k = \min\left\{M_k - N'_j, N_k - N'_k\right\}$. Thus, from (4.86), (4.90), (4.91), and the last constraint in (4.87), it follows that (4.82) always holds. Finally, the only constraint that remains is (4.84) that can be equivalently written as (4.92) and (4.93). □

Definition 4.5 (achievable region for IC)**.** *Let us define*

$$\mathcal{I}_A^{\mathrm{IC}} \triangleq \left\{ (A_1, A'_1, A''_1, A_2, A'_2, A''_2) \in \mathcal{A}_{\mathrm{IC}} \;\middle|\; \begin{array}{l} (d_1, d_2) = f_{A\text{-}d}(A_1, A'_1, A''_1, A_2, A'_2, A''_2), \\ \frac{d_k}{N'_k} \leq \frac{N_k}{N'_k} - A_j, \quad k \neq j \in \{1,2\} \end{array} \right\} \tag{4.94}$$

and the achievable DoF region of the proposed scheme

$$\mathcal{I}_d^{\mathrm{IC}} \triangleq f_{A\text{-}d}(\mathcal{I}_A^{\mathrm{IC}}) \triangleq \left\{ (d_1, d_2) \;\middle|\; \begin{array}{l} (d_1, d_2) = f_{A\text{-}d}(A_1, A'_1, A''_1, A_2, A'_2, A''_2), \\ (A_1, A'_1, A''_1, A_2, A'_2, A''_2) \in \mathcal{A}_{\mathrm{IC}}, \\ \frac{d_k}{N'_k} \leq \frac{N_k}{N'_k} - A_j, \quad k \neq j \in \{1,2\} \end{array} \right\}. \tag{4.95}$$

Achievability analysis

The analysis is similar to the BC case, i.e., it is sufficient to find a function $f_{d\text{-}A} : \mathcal{O}_d^{\mathrm{IC}} \to \mathcal{A}_{\mathrm{IC}}$ where $\mathcal{O}_d^{\mathrm{IC}}$ denotes the outer bound region defined by (4.12), such that

$$(d_1, d_2) = f_{A\text{-}d}(f_{d\text{-}A}(d_1, d_2)), \quad \text{and} \tag{4.96}$$

$$f_{d\text{-}A}(d_1, d_2) \in \mathcal{I}_A^{\mathrm{IC}}. \tag{4.97}$$

Now, we define formally the power allocation function.

Definition 4.6 (power allocation for IC)**.** *Let us define γ_k, $k \neq j \in \{1,2\}$, as*

$$\gamma_k \triangleq \min\left\{1, \frac{M_j - d_j}{\xi_k}\right\}. \tag{4.98}$$

Then, we define $f_{d\text{-}A} : \mathcal{O}_d^{\mathrm{IC}} \to \mathcal{A}_{\mathrm{IC}}$*:*

$$(d_1, d_2) \mapsto \begin{array}{l} (A_1, A_1', A_1'') \triangleq f_1(d_1, d_2), \\ (A_2, A_2', A_2'') \triangleq f_2(d_1, d_2) \end{array} \tag{4.99}$$

where f_k*,* $k \neq j \in \{1, 2\}$*, such that* (4.90) *is satisfied, and that*

- *when* $M_k = N_j'$*:* $A'' = A_k' = A_k = \frac{d_k}{M_k}$*;*
- *when* $M_k > N_j'$*,* $d_k < (M_k - N_j')\gamma_k + N_j'(\gamma_k - \alpha_j)^+$*:*

$$A_k = (A_k' - \alpha_j)^+, \quad A_k' = A_k'' < \gamma_k; \tag{4.100}$$

- *when* $M_k > N_j'$*,* $d_k \geq (M_k - N_j')\gamma_k + N_j'(\gamma_k - \alpha_j)^+$*, and* $\gamma_k < 1$*:*

$$A_k = (A_k' - \alpha_j)^+, \quad A_k' > A_k'' = \gamma_k; \tag{4.101}$$

- *when* $M_k > N_j'$*,* $d_k \geq (M_k - N_j')\gamma_k + N_j'(\gamma_k - \alpha_j)^+$*, and* $\gamma_k = 1$*:*

$$A_k' = A_k'' = 1. \tag{4.102}$$

First, one can verify, with some basic manipulations that, $f_{d\text{-}A}(\mathcal{O}_d^{\mathrm{IC}}) \subseteq \mathcal{A}_{\mathrm{IC}}$. Second, (4.96) is satisfied by construction. Finally, it remains to show that (4.97) holds as well, i.e., the decodability condition in (4.94) is satisfied. Since the region $\mathcal{O}_d^{\mathrm{IC}}$ depends on whether the condition C_k holds, we prove the achievability accordingly.

Neither C_1 nor C_2 holds ($\xi_1 = \xi_2 = 0$)

For any $(d_1, d_2) \in \mathcal{O}_d^{\mathrm{IC}}$, we can define $(A_1, A_1', A_1'', A_2, A_2', A_2'') \triangleq f_{d\text{-}A}(d_1, d_2)$ which implies, in this case,

$$d_j = N_k' A_j + (M_j - N_k')A_j', \quad j \neq k \in \{1, 2\}. \tag{4.103}$$

Applying this equality on the constraints in the outer bound $\mathcal{O}_d^{\mathrm{IC}}$ in (4.12), we have

$$\begin{aligned} \frac{d_k}{N_k'} \leq & \frac{\min\left\{\max\left\{M_1, N_2\right\}, \max\left\{M_2, N_1\right\}\right\}}{N_k'} \\ & - \frac{(M_j - N_k')A_j'}{N_k'} - A_j, \end{aligned} \tag{4.104}$$

$$\frac{d_k}{N'_k} \leq \frac{\min\{M_k, N_k\}}{N'_k} - \left[\frac{\min\{M_k, N_k\} - N_k}{N'_k} + \frac{(M_j - N'_k)(A'_j - \alpha_k) + N'_k A_j}{M_j}\right]^+, \tag{4.105}$$

for $k \neq j \in \{1, 2\}$, where the first one is from the sum rate constraint (4.12c) whereas the second one is from the rest of the constraints in (4.12). The final step is to show that either of (4.104) and (4.105) implies the last constraint in (4.94).

- When $M_j = N'_k$, (4.105) implies the last constraint in (4.94) because $\frac{\min\{M_k, N_k\} - N_k}{N_k} \leq 0$;
- When $M_j > N'_k$ and $d_j \geq M_j - N'_k \alpha_k$, we have $A'_j = 1$ according to the mapping $f_{d\text{-}A}$ defined in Definition 4.6, since $\gamma_j = 1$. Hence, the right hand side (RHS) of (4.104) is not greater than $\frac{N_k}{N'_k} - A_j$, which implies the last constraint in (4.94);
- When $M_j > N'_k$ and $d_j < M_j - N'_k \alpha_k$, we have $A_j = (A'_j - \alpha_k)^+$ according to Definition 4.6 with $\gamma_j = 1$. Since $\frac{\min\{M_k, N_k\} - N_k}{N_k} \leq 0$, we can show that

$$\left[\frac{\min\{M_k, N_k\} - N_k}{N_k} + \frac{(M_j - N'_k)(A'_j - \alpha_k) + N'_k A_j}{M_j}\right]^+ \geq \frac{\min\{M_k, N_k\} - N_k}{N_k} + A_j \tag{4.106}$$

  with which (4.105) implies the last constraint in (4.94).

C_k holds ($\xi_k > 0$, $\xi_j = 0$)

In this case, it is readily shown, from (4.90) and (4.91), that

$$d_k = N_j A_k + (N_k - M_j) A'_k + \xi_k A''_k, \tag{4.107}$$
$$d_j = M_j A_j. \tag{4.108}$$

Applying the mapping $d_j = M_j A_j$ on (4.12c) results in

$$\frac{d_k}{N'_k} \leq \frac{\min\{M_k, N_k\}}{N'_k} - A_j \tag{4.109}$$

that always implies $\frac{d_k}{N'_k} \leq \frac{N_k}{N'_k} - A_j$. Due to the asymmetry, we also need to prove that $\frac{d_j}{N_j} \leq 1 - A_k$. Therefore, the final step is to show that it can be implied by at least one of the constraints in (4.12), together with (4.107) and (4.108).

- When $d_k < (M_k - N_j)\gamma_k + N_j(\gamma_k - \alpha_j)^+$, we have $A'_k = A''_k < \gamma_k$ according to (4.100). Therefore, $d_k = N_j A_k + (M_k - N_j)A'_k$, plugging which into (4.12e), we obtain

$$\begin{aligned}\frac{d_j}{N_j} \leq & \frac{\min\{M_j, N_j\}}{N_j} - \left[\frac{\min\{M_j, N_j\} - N_j}{N_j}\right. \\ & \left. + \frac{(M_k - N_j)(A'_k - \alpha_j) + N_j A_k}{M_k}\right]^+ \qquad (4.110) \\ \leq & \frac{\min\{M_j, N_j\}}{N_j} - \left[\frac{\min\{M_j, N_j\} - N_j}{N_j} + A_k\right] \qquad (4.111)\end{aligned}$$

  where the $[\cdot]^+$ in (4.110) is from the single user bound (4.12b); the last inequality is due to $A_k = (A'_k - \alpha_j)^+$ and $\frac{\min\{M_j, N_j\} - N_j}{N_j} \leq 0$.

- When $d_k \geq (M_k - N_j)\gamma_k + N_j(\gamma_k - \alpha_j)^+$, we have $A'_k \geq A''_k = \gamma_k$ according to (4.101) and (4.102).

  - If $\gamma_k < 1$, then $A''_k = \gamma_k = \frac{M_j - d_j}{\xi_k}$ and $d_k = (N_k - M_j)A'_k + M_j - d_j + N_j A_k$. Plugging the latter into (4.12f), we obtain

$$\begin{aligned}\frac{d_j}{N_j} \leq & \frac{\min\{M_j, N_j\}}{N_j} - \left[\frac{\min\{M_j, N_j\} - N_j}{N_j}\right. \\ & \left. + \frac{(N_k - M_j)(A'_k - \alpha_j) + N_j A_k}{N_k + N_j - M_j}\right]^+ \qquad (4.112) \\ \leq & \frac{\min\{M_j, N_j\}}{N_j} - \left[\frac{\min\{M_j, N_j\} - N_j}{N_j} + A_k\right] \qquad (4.113)\end{aligned}$$

    where the $[\cdot]^+$ in (4.112) is from the single user bound (4.12b); the last inequality is due to $A_k = (A'_k - \alpha_j)^+$ and $\frac{\min\{M_j, N_j\} - N_j}{N_j} \leq 0$.

  - If $\gamma_k = 1$, then $A'_k = A''_k = 1$ and $d_k = M_k - N_j + N_j A_k$. Plugging the latter into (4.12c), we obtain

$$\begin{aligned}\frac{d_j}{N_j} & \leq \frac{\min\{M_k, N_k\} - M_k + N_j - N_j A_k}{N_j} \qquad (4.114) \\ & \leq 1 - A_k. \qquad (4.115)\end{aligned}$$

Thus, the last constraint in (4.94) is shown in all cases. By now, we have proved the achievability through the existence of a proper power allocation function such that (4.96) and (4.97) are satisfied for every pair (d_1, d_2) in the outer bound.

4.6 Converse

To obtain the outer bounds, the following ingredients are essential:

- *Genie-aided* bounding techniques by providing side information of one receiver to the other one [43, 75];
- *Extremal inequality* to bound the weighted difference of conditional differential entropies [90, 91];
- *Ergodic capacity upper and lower bounds* for MIMO channels with channel uncertainty.

In the following, we first present the proof of outer bound (4.10d) for MIMO BC and (4.12d) for MIMO IC, referred to in this section as L_4. It should be noticed that both bounds share the same structure. Then, we give the proof of bound (4.12f) for the MIMO IC case, referred to in this section as L_6, when the condition C_1 holds.

4.6.1 Proof of Bound L_4

We first provide the outer bounds by employing the genie-aided techniques for BC and IC, respectively, reaching a similar formulation of the rate bounds. With the help of extremal inequalities, the weighted sum rates are further bounded. Finally, the bounds in terms of (α_1, α_2) are obtained by deriving novel ergodic capacity bounds for MIMO channels with channel uncertainty.

To obtain the outer bounds, we adopt a genie-aided upper bounding technique reminisced in [43, 75], by providing Receiver 2 the side information of Receiver 1's message W_1 as well as received signal $\boldsymbol{Y}_1^n$. For notational brevity, we define a virtual received signal as

$$\bar{\boldsymbol{y}}_i(t) \triangleq \begin{cases} \boldsymbol{H}_i(t)\boldsymbol{x}(t) + \boldsymbol{z}_i(t) & \text{for BC} \\ \boldsymbol{H}_{i2}(t)\boldsymbol{x}_2(t) + \boldsymbol{z}_i(t) & \text{for IC} \end{cases} \tag{4.116}$$

and we also define $\boldsymbol{X}^n \triangleq \{\boldsymbol{x}(t)\}_{t=1}^n$, $\boldsymbol{X}_i^n \triangleq \{\boldsymbol{x}_i(t)\}_{t=1}^n$, $\boldsymbol{Y}_i^k \triangleq \{\boldsymbol{y}_i(t)\}_{t=1}^k$, and $\bar{\boldsymbol{Y}}_i^k \triangleq \{\bar{\boldsymbol{y}}_i(t)\}_{t=1}^k$. Denote also $n\epsilon_n \triangleq 1 + nRP_e^{(n)}$ where ϵ_n tends to zero as $n \to \infty$ by the assumption that $\lim_{n\to\infty} P_e^{(n)} = 0$.

Broadcast Channel

The genie-aided model is a degraded BC $\boldsymbol{X}^n \to (\boldsymbol{Y}_1^n, \boldsymbol{Y}_2^n) \to \boldsymbol{Y}_1^n$, and therefore we bound the achievable rates by applying Fano's inequality as

$$
\begin{aligned}
& n(R_1 - \epsilon_n) \\
& \leq I(W_1; \boldsymbol{Y}_1^n | \mathcal{H}^n, \hat{\mathcal{H}}^n) && (4.117)\\
& = \sum_{t=1}^{n} I(W_1; \boldsymbol{y}_1(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \boldsymbol{Y}_1^{t-1}) && (4.118)\\
& = \sum_{t=1}^{n} \Big(h(\boldsymbol{y}_1(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \boldsymbol{Y}_1^{t-1}) \; - h(\boldsymbol{y}_1(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \boldsymbol{Y}_1^{t-1}, W_1) \Big) && (4.119)\\
& \leq \sum_{t=1}^{n} \left(h(\boldsymbol{y}_1(t) | \mathcal{H}(t)) - h(\boldsymbol{y}_1(t) | \mathcal{U}(t), \mathcal{H}(t)) \right) && (4.120)\\
& \leq n N_1' \log P - \sum_{t=1}^{n} h(\bar{\boldsymbol{y}}_1(t) | \mathcal{U}(t), \mathcal{H}(t)) + n \cdot O(1) && (4.121)\\
& n(R_2 - \epsilon_n) \\
& \leq I(W_2; \boldsymbol{Y}_1^n, \boldsymbol{Y}_2^n, W_1 | \mathcal{H}^n, \hat{\mathcal{H}}^n) && (4.122)\\
& = I(W_2; \boldsymbol{Y}_1^n, \boldsymbol{Y}_2^n | W_1, \mathcal{H}^n, \hat{\mathcal{H}}^n) && (4.123)\\
& = \sum_{t=1}^{n} I(W_2; \boldsymbol{y}_1(t), \boldsymbol{y}_2(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \boldsymbol{Y}_1^{t-1}, \boldsymbol{Y}_2^{t-1}, W_1) && (4.124)\\
& \leq \sum_{t=1}^{n} I(\boldsymbol{x}(t); \boldsymbol{y}_1(t), \boldsymbol{y}_2(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \boldsymbol{Y}_1^{t-1}, \boldsymbol{Y}_2^{t-1}, W_1) && (4.125)\\
& = \sum_{t=1}^{n} \Big(h(\boldsymbol{y}_1(t), \boldsymbol{y}_2(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \boldsymbol{Y}_1^{t-1}, \boldsymbol{Y}_2^{t-1}, W_1) && (4.126)\\
& \qquad - h(\boldsymbol{y}_1(t), \boldsymbol{y}_2(t) | \boldsymbol{x}(t), \mathcal{H}^n, \hat{\mathcal{H}}^n, \boldsymbol{Y}_1^{t-1}, \boldsymbol{Y}_2^{t-1}, W_1) \Big) && (4.127)\\
& \leq \sum_{t=1}^{n} h(\boldsymbol{y}_1(t), \boldsymbol{y}_2(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \boldsymbol{Y}_1^{t-1}, \boldsymbol{Y}_2^{t-1}, W_1) && (4.128)\\
& = \sum_{t=1}^{n} h(\bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t) | \mathcal{U}(t), \mathcal{H}(t)) && (4.129)
\end{aligned}
$$

where $\mathcal{U}(t) \triangleq \left\{ \bar{\boldsymbol{Y}}_1^{t-1}, \bar{\boldsymbol{Y}}_2^{t-1}, \mathcal{H}^{t-1}, \hat{\mathcal{H}}^t, W_1 \right\}$ for BC and $N_1' \triangleq \min\{M, N_1\}$; (4.120) is from (4.116) and because (a) conditioning reduces differential

entropy, and (b) $\{\bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t)\}$ are independent of $\mathcal{H}_{t+1}^n$ and $\hat{\mathcal{H}}_{t+1}^n$, given the past states and channel outputs; (4.121) follows the fact that the rate of the point-to-point $M \times N_1$ MIMO channel (i.e., between the transmitter and Receiver 1) is bounded by $\min\{M, N_1\} \log P + O(1)$; (4.123) is due to the independence between W_1 and W_2; (4.125) follows date processing inequality; (4.128) is obtained by noticing (a) translation does not change differential entropy, (b) Gaussian noise terms are independent from instant to instant, and are also independent of the channel matrices and the transmitted signals, and (c) the differential entropy of Gaussian noise with normalized variance is non-negative and finite.

Interference Channel

Given the message and corresponding channel states, part of the received signal is deterministic and therefore removable without mutual information loss. Hence, similarly to the BC case, we formulate a degraded channel model, i.e., $\boldsymbol{X}_2^n \to \left(\bar{\boldsymbol{Y}}_1^n, \bar{\boldsymbol{Y}}_2^n\right) \to \bar{\boldsymbol{Y}}_1^n$. By applying Fano's inequality, the achievable rate of Receiver 1 and Receiver 2 can be bounded as

$$
\begin{aligned}
& n(R_1 - \epsilon_n) \\
& \le I(W_1; \boldsymbol{Y}_1^n | \mathcal{H}^n, \hat{\mathcal{H}}^n) && (4.130)\\
& = I(W_1, W_2; \boldsymbol{Y}_1^n | \mathcal{H}^n, \hat{\mathcal{H}}^n) - I(W_2; \boldsymbol{Y}_1^n | W_1, \mathcal{H}^n, \hat{\mathcal{H}}^n) && (4.131)\\
& \le n\tilde{N}_1 \log P - I(W_2; \boldsymbol{Y}_1^n | W_1, \mathcal{H}^n, \hat{\mathcal{H}}^n) + n \cdot O(1) && (4.132)\\
& = n\tilde{N}_1 \log P - h(\boldsymbol{Y}_1^n | W_1, \mathcal{H}^n, \hat{\mathcal{H}}^n) \\
& \qquad + h(\boldsymbol{Y}_1^n | W_1, W_2, \mathcal{H}^n, \hat{\mathcal{H}}^n) + n \cdot O(1) && (4.133)\\
& = n\tilde{N}_1 \log P - h(\boldsymbol{Y}_1^n | W_1, \mathcal{H}^n, \hat{\mathcal{H}}^n) + n \cdot O(1) && (4.134)\\
& = n\tilde{N}_1 \log P - h(\bar{\boldsymbol{Y}}_1^n | \mathcal{H}^n, \hat{\mathcal{H}}^n) + n \cdot O(1) && (4.135)\\
& \le n\tilde{N}_1 \log P - \sum_{t=1}^{n} h(\bar{\boldsymbol{y}}_1(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \bar{\boldsymbol{Y}}_1^{t-1}, \bar{\boldsymbol{Y}}_2^{t-1}) + n \cdot O(1) && (4.136)\\
& = n\tilde{N}_1 \log P - \sum_{t=1}^{n} h(\bar{\boldsymbol{y}}_1(t) | \mathcal{U}(t), \mathcal{H}(t)) + n \cdot O(1) && (4.137)\\
& n(R_2 - \epsilon_n) \\
& \le I(W_2; \boldsymbol{Y}_1^n, \boldsymbol{Y}_2^n, W_1 | \mathcal{H}^n, \hat{\mathcal{H}}^n) && (4.138)\\
& = I(W_2; \boldsymbol{Y}_1^n, \boldsymbol{Y}_2^n | W_1, \mathcal{H}^n, \hat{\mathcal{H}}^n) && (4.139)\\
& = I(W_2; \bar{\boldsymbol{Y}}_1^n, \bar{\boldsymbol{Y}}_2^n | \mathcal{H}^n, \hat{\mathcal{H}}^n) && (4.140)
\end{aligned}
$$

$$= \sum_{t=1}^{n} I(W_2; \bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \bar{\boldsymbol{Y}}_1^{t-1}, \bar{\boldsymbol{Y}}_2^{t-1}) \tag{4.141}$$

$$\leq \sum_{t=1}^{n} I(\boldsymbol{x}_2(t); \bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \bar{\boldsymbol{Y}}_1^{t-1}, \bar{\boldsymbol{Y}}_2^{t-1}) \tag{4.142}$$

$$= \sum_{t=1}^{n} \Big(h(\bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \bar{\boldsymbol{Y}}_1^{t-1}, \bar{\boldsymbol{Y}}_2^{t-1}) \tag{4.143}$$

$$- h(\bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t) | \boldsymbol{x}_2(t), \mathcal{H}^n, \hat{\mathcal{H}}^n, \bar{\boldsymbol{Y}}_1^{t-1}, \bar{\boldsymbol{Y}}_2^{t-1}) \Big) \tag{4.144}$$

$$\leq \sum_{t=1}^{n} h(\bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t) | \mathcal{H}^n, \hat{\mathcal{H}}^n, \bar{\boldsymbol{Y}}_1^{t-1}, \bar{\boldsymbol{Y}}_2^{t-1}) \tag{4.145}$$

$$= \sum_{t=1}^{n} h(\bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t) | \mathcal{U}(t), \mathcal{H}(t)) \tag{4.146}$$

where we define $\mathcal{U}(t) \triangleq \left\{ \bar{\boldsymbol{Y}}_1^{t-1}, \bar{\boldsymbol{Y}}_2^{t-1}, \mathcal{H}^{t-1}, \hat{\mathcal{H}}^t \right\}$ for IC and $\tilde{N}_1 \triangleq \min\{M_1 + M_2, N_1\}$; (4.132) follows the fact that the mutual information at hand is upper bounded by the rate of the $(M_1 + M_2) \times N_1$ point-to-point MIMO channel created by letting the two transmitters cooperate perfectly, given by $\min\{M_1 + M_2, N_1\} \log P + O(1)$; (4.134) is due to the fact that (a) transmitted signal $\boldsymbol{X}_i^n$ is a deterministic function of messages W_i, $\mathcal{H}^n$, and $\hat{\mathcal{H}}^{n-1}$ as specified in (4.7) for $i = 1, 2$, (b) translation does change differential entropy, and (c) the differential entropy of Gaussian noise with normalized variance is non-negative and finite; (4.135) and (4.140) are obtained because translation preserves differential entropy; (4.136) is because conditioning reduces differential entropy; (4.145) is because (a) translation does not change differential entropy, (b) Gaussian noise terms are independent from instant to instant, and are also independent of the channel matrices and the transmitted signals, and (c) the differential entropy of Gaussian noise with normalized variance is non-negative and finite; (4.146) is obtained due to the independence $\{\bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t)\}$ of $\mathcal{H}_{t+1}^n$ and $\hat{\mathcal{H}}_{t+1}^n$, given the past state and channel outputs.

It is worth noting that BC and IC share the common structure of the achievable rate bounds, and therefore can be further bounded in a similar way. To avoid redundancy, we give the proof for IC, which can be straightforwardly extended to BC.

Define

$$\boldsymbol{S}(t) \triangleq \begin{bmatrix} \boldsymbol{H}_{12}(t) \\ \boldsymbol{H}_{22}(t) \end{bmatrix}, \quad \hat{\boldsymbol{S}}(t) \triangleq \begin{bmatrix} \hat{\boldsymbol{H}}_{12}(t) \\ \hat{\boldsymbol{H}}_{22}(t) \end{bmatrix}, \\ \boldsymbol{K}(t) \triangleq \mathbb{E}\{\boldsymbol{x}_2(t)\boldsymbol{x}_2^{\mathsf{H}}(t) \,|\, \mathcal{U}(t)\}. \tag{4.147}$$

Let $p = \min\{M_2, N_1 + N_2\}$ and $q = \min\{M_2, N_1\}$. By following the footsteps in [54], we have

$$\frac{1}{p} h(\bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t) | \mathcal{U}(t), \mathcal{H}(t)) - \frac{1}{q} h(\bar{\boldsymbol{y}}_1(t) | \mathcal{U}(t), \mathcal{H}(t)) \tag{4.148}$$

$$\leq \mathbb{E}_{\hat{\boldsymbol{S}}(t)} \max_{\substack{\boldsymbol{K} \succeq 0, \\ \mathrm{tr}(\boldsymbol{K}) \leq P}} \mathbb{E}_{\boldsymbol{S}(t)|\hat{\boldsymbol{S}}(t)} \left(\frac{1}{p} \log\det(\mathbf{I} + \boldsymbol{S}(t)\boldsymbol{K}(t)\boldsymbol{S}^{\mathsf{H}}(t)) \right. \\ \left. - \frac{1}{q} \log\det(\mathbf{I} + \boldsymbol{H}_{12}(t)\boldsymbol{K}(t)\boldsymbol{H}_{12}^{\mathsf{H}}(t)) \right) \tag{4.149}$$

$$\leq -\frac{\min\{M_2, N_1 + N_2\} - \min\{M_2, N_1\}}{\min\{M_2, N_1 + N_2\}} \log \sigma_1^2 + O(1) \tag{4.150}$$

where (4.149) is obtained by applying extremal inequality [90,91] for degraded outputs; the last inequality is obtained from the following lemma:

Lemma 4.3. *For two random matrices $\boldsymbol{S} = \hat{\boldsymbol{S}} + \tilde{\boldsymbol{S}} \in \mathbb{C}^{L \times M}$ and $\boldsymbol{H} = \hat{\boldsymbol{H}} + \tilde{\boldsymbol{H}} \in \mathbb{C}^{N \times M}$ with $L \geq N$, $\tilde{\boldsymbol{S}}$, $\tilde{\boldsymbol{H}}$ are respectively independent of $\hat{\boldsymbol{S}}$, $\hat{\boldsymbol{H}}$, and the entries of $\tilde{\boldsymbol{H}}$ are i.i.d. $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$. Then, given any $\boldsymbol{K} \succeq 0$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_M \geq 0$, we have*

$$\frac{1}{\min\{M, L\}} \mathbb{E}_{\tilde{\boldsymbol{S}}} \log\det(\mathbf{I} + \boldsymbol{S}\boldsymbol{K}\boldsymbol{S}^{\mathsf{H}}) \\ - \frac{1}{\min\{M, N\}} \mathbb{E}_{\tilde{\boldsymbol{H}}} \log\det(\mathbf{I} + \boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^{\mathsf{H}}) \\ \leq -\frac{\min\{M, L\} - \min\{M, N\}}{\min\{M, L\}} \log(\sigma^2) + O_{\hat{S}}(1) + O_{\hat{H}}(1) \tag{4.151}$$

as σ^2 goes to 0.

Proof. See Appendix B. □

Remark:

- This lemma can be regarded as the weighted difference of the ergodic capacity for two MIMO channels with uncertainty, where $\tilde{\boldsymbol{S}}$ and $\tilde{\boldsymbol{H}}$ are channel uncertainties. It can also be interpreted as the ergodic capacity difference of two Ricean MIMO channels with line-of-sight components $\hat{\boldsymbol{S}}$, $\hat{\boldsymbol{H}}$, and fading components $\tilde{\boldsymbol{S}}$, $\tilde{\boldsymbol{H}}$.

- This lemma also shows the change of the ergodic capacity per dimension as the dimensionality decreases. In other words, as the channel dimension decreases, the difference of the ergodic capacity per dimension is bounded by the dimension difference and the channel uncertainty.

According to the Markov chain $\boldsymbol{X}_2^n \to \left(\bar{\boldsymbol{Y}}_1^n, \bar{\boldsymbol{Y}}_2^n\right) \to \bar{\boldsymbol{Y}}_1^n$, we upper-bound the weighted sum rate as

$$
\begin{aligned}
& n\left(\frac{R_1}{\min\{M_2, N_1\}} + \frac{R_2}{\min\{M_2, N_1+N_2\}} - \epsilon_n\right) && (4.152)\\
& \quad \le n \cdot \frac{\min\{M_1+M_2, N_1\}}{\min\{M_2, N_1\}} \log P \\
& \quad + \sum_{t=1}^{n} \left(\frac{1}{\min\{M_2, N_1+N_2\}} h(\bar{\boldsymbol{y}}_1(t), \bar{\boldsymbol{y}}_2(t) | \mathcal{U}(t), \mathcal{H}(t)) \right. && (4.153)\\
& \qquad \left. - \frac{1}{\min\{M_2, N_1\}} h(\bar{\boldsymbol{y}}_1(t) | \mathcal{U}(t), \mathcal{H}(t)) \right) + n \cdot O(1) && (4.154)\\
& \quad \le n \frac{\min\{M_1+M_2, N_1\}}{\min\{M_2, N_1\}} \log P + n \cdot O(1) \\
& \qquad + n \frac{\min\{M_2, N_1+N_2\} - \min\{M_2, N_1\}}{\min\{M_2, N_1+N_2\}} \alpha_1 \log P && (4.155)
\end{aligned}
$$

and another outer bound can be similarly obtained by exchanging the roles of Receiver 1 and Receiver 2. Accordingly, the corresponding outer bound L_4 of the DoF region is obtained by the definition.

4.6.2 Proof of Bound L_6

This bound is active when C_1 holds, i.e., $M_1 \ge N_2$, $N_1 > M_2$, and $M_1 + M_2 > N_1 + N_2$. The proof follows the same lines of thought in [43]. Since $N_1 > M_2$, we formulate a virtual received signal

$$
\begin{aligned}
\tilde{\boldsymbol{y}}_1(t) &\triangleq \boldsymbol{U}\boldsymbol{y}_1(t) \\
&= \boldsymbol{U}\boldsymbol{H}_{11}(t)\boldsymbol{x}_1(t) + \boldsymbol{U}\boldsymbol{H}_{12}(t)\boldsymbol{x}_2(t) + \boldsymbol{U}\boldsymbol{z}_1(t) && (4.156)
\end{aligned}
$$

where $\boldsymbol{U} \in \mathbb{C}^{N_1 \times N_1}$ is any unitary matrix such that the last $N_1 - M_2$ rows of $\boldsymbol{U}(t)\boldsymbol{H}_{12}(t)$ are with all zeros and is independent of the rest of random variables. Therefore, the last $N_1 - M_2$ outputs in $\tilde{\boldsymbol{y}}_1(t)$ are interference free, i.e., $\tilde{\boldsymbol{y}}_{1[M_2+1:N_1]}(t) \sim \boldsymbol{H}_{1[M_2+1:N_1]1}(t)\boldsymbol{x}_1(t) + \boldsymbol{z}_{1[M_2+1:N_1]}(t)$. For convenience, we also define

$$
\tilde{\boldsymbol{y}}_2(t) \triangleq \boldsymbol{H}_{21}(t)\boldsymbol{x}_1(t) + \boldsymbol{z}_2(t). \qquad (4.157)
$$

Starting with Fano's inequality, the achievable rate can be bounded as

$$
\begin{aligned}
&n(R_1-\epsilon_n)\\
&\le I(W_1;\mathbf{Y}_1^n|\mathcal{H}^n,\hat{\mathcal{H}}^n) && (4.158)\\
&= I(W_1;\tilde{\mathbf{Y}}_1^n|\mathcal{H}^n,\hat{\mathcal{H}}^n) && (4.159)\\
&= I(W_1;\tilde{\mathbf{Y}}_{1[1:M_2]}^n|\mathcal{H}^n,\hat{\mathcal{H}}^n,\tilde{\mathbf{Y}}_{1[M_2+1:N_1]}^n)\\
&\qquad + I(W_1;\tilde{\mathbf{Y}}_{1[M_2+1:N_1]}^n|\mathcal{H}^n,\hat{\mathcal{H}}^n) && (4.160)\\
&\le n(M_2-d_2)\log P + n\cdot O(1)\\
&\qquad + I(W_1;\tilde{\mathbf{Y}}_{1[M_2+1:N_1]}^n|\mathcal{H}^n,\hat{\mathcal{H}}^n) && (4.161)\\
&\le n(M_2-d_2)\log P + n\cdot O(1)\\
&\qquad + I(W_1;\tilde{\mathbf{Y}}_{1[M_2+1:N_1]}^n,\tilde{\mathbf{Y}}_2^n|\mathcal{H}^n,\hat{\mathcal{H}}^n) && (4.162)\\
&= n(M_2-d_2)\log P + n\cdot O(1)\\
&\quad + \sum_{t=1}^{n} I\big(W_1;\tilde{\boldsymbol{y}}_{1[M_2+1:N_1]}(t),\tilde{\boldsymbol{y}}_2(t)|\mathcal{H}^n,\hat{\mathcal{H}}^n,\tilde{\mathbf{Y}}_{1[M_2+1:N_1]}^{t-1},\tilde{\mathbf{Y}}_2^{t-1}\big) && (4.163)\\
&\le n(M_2-d_2)\log P + n\cdot O(1)\\
&\quad + \sum_{t=1}^{n} h(\tilde{\boldsymbol{y}}_{1[M_2+1:N_1]}(t),\tilde{\boldsymbol{y}}_2(t)|\mathcal{H}^n,\hat{\mathcal{H}}^n,\tilde{\mathbf{Y}}_{1[M_2+1:N_1]}^{t-1},\tilde{\mathbf{Y}}_2^{t-1}) && (4.164)\\
&= n(M_2-d_2)\log P + n\cdot O(1)\\
&\qquad + \sum_{t=1}^{n} h(\tilde{\boldsymbol{y}}_{1[M_2+1:N_1]}(t),\tilde{\boldsymbol{y}}_2(t)|\mathcal{U}(t),\mathcal{H}(t)) && (4.165)\\
&n(R_2-\epsilon_n)\\
&\le I(W_2;\mathbf{Y}_2^n|\mathcal{H}^n,\hat{\mathcal{H}}^n) && (4.166)\\
&= I(W_1,W_2;\mathbf{Y}_2^n|\mathcal{H}^n,\hat{\mathcal{H}}^n) - I(W_1;\mathbf{Y}_2^n|W_2,\mathcal{H}^n,\hat{\mathcal{H}}^n) && (4.167)\\
&\le nN_2\log P - I(W_1;\tilde{\mathbf{Y}}_2^n|\mathcal{H}^n,\hat{\mathcal{H}}^n) + n\cdot O(1) && (4.168)\\
&\le nN_2\log P - h(\tilde{\mathbf{Y}}_2^n|\mathcal{H}^n,\hat{\mathcal{H}}^n) + h(\tilde{\mathbf{Y}}_2^n|W_1,\mathcal{H}^n,\hat{\mathcal{H}}^n) + n\cdot O(1) && (4.169)\\
&= nN_2\log P - h(\tilde{\mathbf{Y}}_2^n|\mathcal{H}^n,\hat{\mathcal{H}}^n) + n\cdot O(1) && (4.170)\\
&\le nN_2\log P - \sum_{t=1}^{n} h(\tilde{\boldsymbol{y}}_2(t)|\mathcal{H}^n,\hat{\mathcal{H}}^n,\tilde{\mathbf{Y}}_{1[M_2+1:N_1]}^{t-1},\tilde{\mathbf{Y}}_2^{t-1}) + n\cdot O(1) && (4.171)\\
&= nN_2\log P - \sum_{t=1}^{n} h(\tilde{\boldsymbol{y}}_2(t)|\mathcal{U}(t),\mathcal{H}(t)) + n\cdot O(1) && (4.172)
\end{aligned}
$$

where $\mathcal{U}(t) \triangleq \{\mathcal{H}^{t-1}, \hat{\mathcal{H}}^t, \tilde{\boldsymbol{Y}}^{t-1}_{1[M_2+1:N_1]}, \tilde{\boldsymbol{Y}}^{t-1}_2\}$ and $\tilde{\boldsymbol{Y}}^k_i \triangleq \{\tilde{\boldsymbol{y}}_i(t)\}^k_{t=1}$, $i = 1, 2$; (4.159) holds due to the fact that unitary transformation does not change the mutual information; (4.161) comes from Lemma 6 in [43], given by

$$I(W_1; \tilde{\boldsymbol{Y}}^n_{1[1:M_2]} | \mathcal{H}^n, \hat{\mathcal{H}}^n, \tilde{\boldsymbol{Y}}^n_{1[M_2+1:N_1]}) \leq n(M_2 - d_2) \log P + n \cdot O(1) \quad (4.173)$$

where a similar proof can be straightforwardly obtained; (4.164) holds because (a) $\tilde{\boldsymbol{y}}_{1[M_2+1:N_1]}(t)$ and $\tilde{\boldsymbol{y}}_2(t)$ are deterministic functions of W_1, $\mathcal{H}^n$ and $\hat{\mathcal{H}}^n$, (b) translation does not change differential entropy, and (c) the differential entropy of Gaussian noise with normalized variance is non-negative and finite; (4.168) follows that the mutual information at hand is upper bounded by the capacity of an $(M_1 + M_2) \times N_2$ point-to-point MIMO channel, i.e., $N_2 \log P + O(1)$ since $M_1 + M_2 > N_2$ from the condition C_1; (4.170) holds because $\tilde{\boldsymbol{y}}_2(t)$ is a deterministic function of W_1, given channel states, and the differential entropy of the normalized Gaussian noise is finite; (4.171) is due to conditioning reduces the differential entropy; (4.165) and the last equality are due to that the received signals at instant t are independent of $\mathcal{H}^n_{t+1}$ and $\hat{\mathcal{H}}^n_{t+1}$, given the past states and channel outputs.

Next, we define

$$\boldsymbol{S}(t) \triangleq \begin{bmatrix} \boldsymbol{H}_{1[M_2+1:N_1]1}(t) \\ \boldsymbol{H}_{21}(t) \end{bmatrix} \in \mathbb{C}^{(N_1+N_2-M_2)\times M_1}. \quad (4.174)$$

Similarly to the proof for bound L_4, we obtain the weighted difference of two differential entropies by applying the extremal inequality and Lemma 4.3

$$\frac{1}{p} h(\tilde{\boldsymbol{y}}_{1[M_2+1:N_1]}(t), \tilde{\boldsymbol{y}}_2(t) | \mathcal{U}(t), \mathcal{H}(t)) - \frac{1}{q} h(\tilde{\boldsymbol{y}}_2(t) | \mathcal{U}(t), \mathcal{H}(t)) \leq -\frac{N_1 - M_2}{N_1 + N_2 - M_2} \log \sigma_2^2 + O(1) \quad (4.175)$$

where we set $p = \min\{M_1, N_1 + N_2 - M_2\} = N_1 + N_2 - M_2$ and $q = \min\{M_1, N_2\} = N_2$.

Finally, we have

$$n \left(\frac{R_1}{N_1 + N_2 - M_2} + \frac{R_2}{N_2} - \epsilon_n \right) \quad (4.176)$$

$$\leq n \left(1 + \frac{M_2 - d_2}{N_1 + N_2 - M_2} \right) \log P + \sum_{t=1}^{n} \left(\frac{1}{N_1 + N_2 - M_2} h(\tilde{\boldsymbol{y}}_{1[M_2+1:N_1]}(t), \tilde{\boldsymbol{y}}_2(t) | \mathcal{U}(t), \mathcal{H}(t)) \right. \quad (4.177)$$

$$- \frac{1}{N_2} h(\tilde{\boldsymbol{y}}_2(t)|\mathcal{U}(t), \mathcal{H}(t)) \Bigg) + n \cdot O(1) \tag{4.178}$$

$$\leq n \left(1 + \frac{M_2 - d_2}{N_1 + N_2 - M_2}\right) \log P + n \frac{N_1 - M_2}{N_1 + N_2 - M_2} \alpha_2 \log P + n \cdot O(1) \tag{4.179}$$

which leads to

$$d_1 + \frac{N_1 + 2N_2 - M_2}{N_2} d_2 \leq N_1 + N_2 + (N_1 - M_2)\alpha_2. \tag{4.180}$$

By exchanging the roles of Receiver 1 and Receiver 2, the outer bound (4.12g) can be obtained straightforwardly when the condition C_2 holds.

4.7 Summary

In this chapter, we focus on the two-user MIMO broadcast and interference channels where the transmitter(s) has/have access to both delayed CSIT and an estimate of current CSIT. Specifically, the DoF region of MIMO networks (BC/IC) in this setting with general antenna configuration and general current CSIT qualities has been fully characterized, thanks to a simple yet unified framework employing interference quantization, block-Markov encoding and backward decoding techniques. Our DoF regions generalize a number of existing results under more specific CSIT settings, such as delayed CSIT [43, 75], perfect CSIT [26, 28], partial/hybrid/mixed CSIT [42, 86, 87]. The results further shed light on the benefits of the temporally correlated channel, when such correlation can be opportunistically taken into account for system designs. The capacity region characterization in the entire SNR regime is still an interesting open problem and left to future work.

4.8 Appendix

4.8.1 Achievable rate regions for the related MAC channels

Broadcast Channels

Let us focus on Receiver k, $k \neq j \in \{1, 2\}$, without loss of generality. The channel in (4.45) is a MAC, rewritten as

$$\underbrace{\begin{bmatrix} \boldsymbol{y}_k[b] - \boldsymbol{\eta}_{kb} \\ \boldsymbol{\eta}_{jb} \end{bmatrix}}_{Y_k} = \underbrace{\begin{bmatrix} \boldsymbol{H}_k \\ 0 \end{bmatrix}}_{S_1} X_c + \underbrace{\begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{H}_j \end{bmatrix}}_{S_2} X_k + Z_k \tag{4.181}$$

where $X_c \triangleq \boldsymbol{x}_c(l_{b-1})$ and $X_k \triangleq \boldsymbol{u}_k(w_{kb})$ are independent, with rate R_c and R_k, respectively; Z_k is the AWGN. It is well known [89] that a rate pair (R_c, R_k) is achievable in the channel if

$$R_c \leq I(X_c; Y_k \,|\, X_k, S) \tag{4.182}$$
$$R_k \leq I(X_k; Y_k \,|\, X_c, S) \tag{4.183}$$
$$R_c + R_k \leq I(X_c, X_k; Y_k \,|\, S) \tag{4.184}$$

for any input distribution $p_{X_c X_k} = p_{X_c} p_{X_k}$; $S \triangleq \{S_1, S_2\}$ denotes the state of the channel. Let $X_c \sim \mathcal{N}_{\mathbb{C}}(0, \boldsymbol{Q}_c)$ and $X_k \sim \mathcal{N}_{\mathbb{C}}(0, \boldsymbol{Q}_k)$ with $\boldsymbol{Q}_c \triangleq P\mathbf{I}_M$ and $\boldsymbol{Q}_k \triangleq P^{A'_k} \boldsymbol{\Phi}_{\hat{H}_j^{\perp}} + P^{A_k} \boldsymbol{\Phi}_{\hat{H}_j}$. It readily follows that[2]

$$\begin{aligned} I(X_c; Y_k \,|\, X_k) &= \log\det(\mathbf{I} + P\boldsymbol{S}_1 \boldsymbol{S}_1^{\mathsf{H}}) \\ &= N_k \log P + O(1) \end{aligned} \tag{4.185}$$
$$\begin{aligned} I(X_k; Y_k \,|\, X_c) &= \log\det(\mathbf{I} + \boldsymbol{S}_2 \boldsymbol{Q}_k \boldsymbol{S}_2^{\mathsf{H}}) \\ &= ((M - N_j) A'_k + N_j A_k) \log P + O(1) \end{aligned} \tag{4.186}$$

since $\boldsymbol{S}_2 \in \mathbb{C}^{(N_1+N_2)\times M}$ has rank M almost surely, given the assumption $N_1 + N_2 \geq M$. For the sum rate constraint, we have

$$\begin{aligned} &I(X_c, X_k; Y_k) \\ &\quad = h(Y_k) - h(Z_k) && (4.187) \\ &\quad = h(\boldsymbol{H}_j X_k + Z_{k2}) + O(1) \\ &\qquad + h(\boldsymbol{H}_k(X_c + X_k) + Z_{k1} \,|\, \boldsymbol{H}_j X_k + Z_{k2}) && (4.188) \\ &\quad \geq h(\boldsymbol{H}_j X_k + Z_{k2}) + O(1) \\ &\qquad + h(\boldsymbol{H}_k(X_c + X_k) + Z_{k1} \,|\, \boldsymbol{H}_j X_k + Z_{k2}, X_k) && (4.189) \\ &\quad = h(\boldsymbol{H}_j X_k + Z_{k2}) + h(\boldsymbol{H}_k X_c + Z_{k1}) + O(1) && (4.190) \\ &\quad = N_j A_k \log P + N_k \log P + O(1) && (4.191) \end{aligned}$$

where we define Z_{k1} and Z_{k2} the first and second parts of the noise vector Z_k; the second equality is from the chain rule and the fact that the Gaussian noise Z_k is normalized; (4.189) is due to conditioning reduces differential entropy; (4.190) is from the independence between X_c and X_k and between the noises and the rest; the first term in (4.191) is essentially the differential entropy of the interference $\boldsymbol{\eta}_{jb}$. By relating the rate pair (R_c, R_k) to the DoF pair (d_η, d_{kb}), (4.46)-(4.48) is straightforward.

[2]Hereafter, we omit for notational brevity the expectation on the channel states S, whenever possible, which does not change the high SNR behavior in this case. We consider any realization $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ instead.

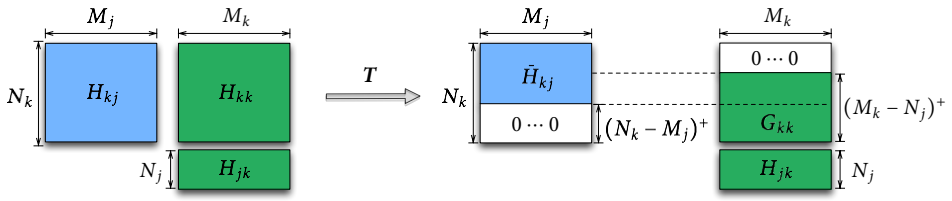


Figure 4.4: Visualization of the interplay between X_{jc} and X_k.

Interference Channels

In (4.78), each receiver sees a MAC with three independent messages. Let us focus on Receiver k, $k \neq j \in \{1,2\}$, without loss of generality. The channel in (4.78) is rewritten as

$$\underbrace{\begin{bmatrix} \boldsymbol{y}_k[b] - \boldsymbol{\eta}_{kb} \\ \boldsymbol{\eta}_{jb} \end{bmatrix}}_{Y_k} = \underbrace{\begin{bmatrix} \boldsymbol{H}_{kk} \\ 0 \end{bmatrix}}_{S_{k1}} X_{kc} + \underbrace{\begin{bmatrix} \boldsymbol{H}_{kj} \\ 0 \end{bmatrix}}_{S_{k2}} X_{jc} + \underbrace{\begin{bmatrix} \boldsymbol{H}_{kk} \\ \boldsymbol{H}_{jk} \end{bmatrix}}_{S_{k3}} X_k + Z_k \tag{4.192}$$

where $X_{kc} \triangleq \boldsymbol{x}_{kc}(l_{k,b-1})$, $X_{jc} \triangleq \boldsymbol{x}_{jc}(l_{j,b-1})$, and $X_k \triangleq \boldsymbol{u}_k(w_{kb})$, $k \neq j \in \{1,2\}$, are three independent signals, with rate R_{kc}, R_{jc}, and R_k, respectively; Z_k is the AWGN. It is well known [89] that a rate triple (R_{kc}, R_{jc}, R_k) is achievable in the channel if

$$R_{kc} \leq I(X_{kc}; Y_k \,|\, X_{jc}, X_k) \tag{4.193}$$
$$R_{jc} \leq I(X_{jc}; Y_k \,|\, X_{kc}, X_k) \tag{4.194}$$
$$R_k \leq I(X_k; Y_k \,|\, X_{kc}, X_{jc}) \tag{4.195}$$
$$R_{kc} + R_{jc} \leq I(X_{kc}, X_{jc}; Y_k \,|\, X_k) \tag{4.196}$$
$$R_{kc} + R_k \leq I(X_{kc}, X_k; Y_k \,|\, X_{jc}) \tag{4.197}$$
$$R_{jc} + R_k \leq I(X_{jc}, X_k; Y_k \,|\, X_{kc}) \tag{4.198}$$
$$R_{kc} + R_{jc} + R_k \leq I(X_{kc}, X_{jc}, X_k; Y_k) \tag{4.199}$$

for any $p_{X_{kc}X_{jc}X_k} = p_{X_{kc}} p_{X_{jc}} p_{X_k}$, where we omit the conditioning on the channel states S as in the BC case for brevity. Let $X_{kc} \sim \mathcal{N}_{\mathbb{C}}(0, \boldsymbol{Q}_{kc})$ and $X_k \sim \mathcal{N}_{\mathbb{C}}(0, \boldsymbol{Q}_k)$ with $\boldsymbol{Q}_{kc} \triangleq P\mathbf{I}_{M_k}$ and $\boldsymbol{Q}_k \triangleq P^{A_k}\boldsymbol{\Phi}_{\hat{H}_{jk}} + P^{A'_k}\boldsymbol{\Phi}_{\hat{H}_{jk}^{\perp 1}} + P^{A''_k}\boldsymbol{\Phi}_{\hat{H}_{jk}^{\perp 2}}$. It is readily shown that

$$\begin{aligned} I(X_{kc}; Y_k \,|\, X_{jc}, X_k) &= \log\det(\mathbf{I} + P\boldsymbol{S}_{k1}\boldsymbol{S}_{k1}^{\mathsf{H}}) \\ &= \min\{M_k, N_k\} \log P + O(1) \end{aligned} \tag{4.200}$$

$$
\begin{aligned}
I(X_{jc}; Y_k \,|\, X_{kc}, X_k) &= \log\det(\mathbf{I} + P\boldsymbol{S}_{k2}\boldsymbol{S}_{k2}^{\mathsf{H}}) \\
&= \min\{M_j, N_k\}\log P + O(1) && (4.201)\\
I(X_k; Y_k \,|\, X_{kc}, X_{jc}) &= \log\det(\mathbf{I} + \boldsymbol{S}_{k3}\boldsymbol{Q}_k\boldsymbol{S}_{k3}^{\mathsf{H}}) \\
&= (N_j' A_k + (M_k - N_j' - \xi_k)A_k' \\
&\qquad + \xi_k A_k'')\log P + O(1) && (4.202)\\
I(X_{jc}, X_{kc}; Y_k \,|\, X_k) &= \log\det(\mathbf{I} + P\boldsymbol{S}_{k1}\boldsymbol{S}_{k1}^{\mathsf{H}} + P\boldsymbol{S}_{k2}\boldsymbol{S}_{k2}^{\mathsf{H}}) \\
&= N_k \log P + O(1) && (4.203)
\end{aligned}
$$

since $\boldsymbol{S}_{k3} \in \mathbb{C}^{(N_1+N_2)\times M_k}$ has rank M_k almost surely, given the assumption $N_1 + N_2 \geq M_k$. Following the same steps as (4.187)-(4.191), we can obtain

$$
I(X_{kc}, X_k; Y_k \,|\, X_{jc}) \;\geq\; N_j' A_k \log P + \min\{M_k, N_k\}\log P + O(1). \quad (4.204)
$$

It remains to bound the RHS of (4.198) and (4.199). First, using the chain rule, we have

$$
I(X_{jc}, X_k; Y_k \,|\, X_{kc}) = I(X_k; Y_k \,|\, X_{kc}) + I(X_{jc}; Y_k \,|\, X_k, X_{kc}) \quad (4.205)
$$

where the scaling of the second term is already shown in (4.201). The first term can be interpreted as the rate of X_k by treating X_{jc} as noise in a two-user MAC with a channel matrix in the block upper triangular form $\begin{bmatrix} \boldsymbol{H}_{kj} & \boldsymbol{H}_{kk} \\ & \boldsymbol{H}_{jk} \end{bmatrix}$. As shown in Fig. 4.4, since $\boldsymbol{H}_{kj}$, $\boldsymbol{H}_{kk}$, and $\boldsymbol{H}_{jk}$ are mutually independent, there exists an invertible row transformation $\boldsymbol{T}$ that can convert the $(N_1 + N_2) \times (M_1 + M_2)$ matrix to the form on the right, almost surely. The interference created by X_{jc} is through the matrix $\bar{\boldsymbol{H}}_{kj}$, only affecting the overlapping part between X_{jc} and X_k, as shown in Fig. 4.4. Note that the dimension of the overlapping is $((M_k - N_j)^+ - (N_k - M_j)^+)^+$ that coincides with the definition of ξ_k in (4.91). From Fig. 4.4, the interference-free received signal for X_k is $\tilde{Y}_k = \begin{bmatrix} \boldsymbol{G}_{kk} \\ \boldsymbol{H}_{jk} \end{bmatrix} X_k + \tilde{Z}_k$. Thus,

$$
\begin{aligned}
I(X_k; Y_k \,|\, X_{kc}) &\geq I(X_k; \tilde{Y}_k) && (4.206)\\
&= \log\det\left(\mathbf{I} + \begin{bmatrix} \boldsymbol{G}_{kk} \\ \boldsymbol{H}_{jk} \end{bmatrix} \boldsymbol{Q}_k \begin{bmatrix} \boldsymbol{G}_{kk} \\ \boldsymbol{H}_{jk} \end{bmatrix}^{\mathsf{H}}\right) + O(1) && (4.207)\\
&\geq \log\det\left(\mathbf{I} + \begin{bmatrix} \tilde{\boldsymbol{G}}_{kk}' & \tilde{\boldsymbol{G}}_{kk} \\ \tilde{\boldsymbol{H}}_{kk}' & \tilde{\boldsymbol{H}}_{kk} \end{bmatrix} \begin{bmatrix} P^{A_k'}\mathbf{I}_{M_k - N_j' - \xi_k} & \\ & P^{A_k}\mathbf{I}_{N_j'} \end{bmatrix}\right. \\
&\qquad \left. \cdot \begin{bmatrix} \tilde{\boldsymbol{G}}_{kk}' & \tilde{\boldsymbol{G}}_{kk} \\ \tilde{\boldsymbol{H}}_{kk}' & \tilde{\boldsymbol{H}}_{kk} \end{bmatrix}^{\mathsf{H}}\right) + O(1) && (4.208)\\
&= ((M_k - N_j' - \xi_k)A_k' + N_j' A_k)\log P + O(1) && (4.209)
\end{aligned}
$$

where the $O(1)$ term in (4.207) is from the fact that the covariance of the noise $\tilde{Z}_k$ depends on $\boldsymbol{T}$ that does not scale with P; $\boldsymbol{G}_{kk}$ and $\boldsymbol{H}_{jk}$ remain independent. Next, let $\boldsymbol{Q}_k = \boldsymbol{U}_{jk}\,\mathrm{diag}(P^{A''_k}\mathbf{I}_{\xi_k}, P^{A'_k}\mathbf{I}_{M_k-N'_j-\xi_k}, P^{A_k}\mathbf{I}_{N'_j})\,\boldsymbol{U}_{jk}^{\mathsf{H}}$ be the eigenvalue decomposition of $\boldsymbol{Q}_k$ and define the column partitions $\left[\tilde{\boldsymbol{G}}''_{kk}\,\tilde{\boldsymbol{G}}'_{kk}\,\tilde{\boldsymbol{G}}_{kk}\right] \triangleq \boldsymbol{G}_{kk}\boldsymbol{U}_{jk}$ and $\left[\tilde{\boldsymbol{H}}''_{jk}\,\tilde{\boldsymbol{H}}'_{jk}\,\tilde{\boldsymbol{H}}_{jk}\right] \triangleq \boldsymbol{H}_{jk}\boldsymbol{U}_{jk}$ where the number of columns of the submatrices is ξ_k, $M_k - N'_j - \xi_k$, and N'_j, respectively; inequality (4.208) is from the fact that removing one column block and the corresponding diagonal block of size ξ_k can only reduce the log-determinant; the last equality is from the fact that the square matrix $\begin{bmatrix}\tilde{\boldsymbol{G}}'_{kk} & \tilde{\boldsymbol{G}}_{kk}\\ \tilde{\boldsymbol{H}}'_{jk} & \tilde{\boldsymbol{H}}_{jk}\end{bmatrix}$ has full rank, almost surely, for the following reasons: 1) the matrices G and H are mutually independent since the column transform $\boldsymbol{U}_{jk}$ is unitary and independent of the G matrices; 2) the rows related to the matrices H are linearly independent, since it can be proved that $\mathrm{rank}(\tilde{\boldsymbol{H}}_{jk}) = \mathrm{rank}(\boldsymbol{H}_{jk}\boldsymbol{\Phi}_{\hat{H}_{jk}}\boldsymbol{H}_{jk}^{\mathsf{H}}) = \min\{M_k, N_j\}$, i.e., $\tilde{\boldsymbol{H}}_{jk}$ has full rank; 3) the rows related to the matrices G are linearly independent as well. Plugging (4.209) and (4.201) into (4.205), we have

$$\begin{aligned} &I(X_{jc}, X_k; Y_k \,|\, X_{kc}) \\ &\qquad \geq (N'_k + (M_k - N'_j - \xi_k)A'_k + N'_j A_k)\log P + O(1). \end{aligned} \tag{4.210}$$

Finally, for the sum rate constraint (4.199), we follow the same steps as (4.187)-(4.191), we can obtain

$$\begin{aligned} &I(X_{kc}, X_{jc}, X_k; Y_k) \\ &\qquad \geq N'_j A_k \log P + \min\{M_k + M_j, N_k\}\log P + O(1) &(4.211)\\ &\qquad = (N_k + N'_j A_k)\log P + O(1) &(4.212) \end{aligned}$$

By relating the rate pair (R_{kc}, R_{jc}, R_k) to the DoF pair $(d_{\eta_1}, d_{\eta_2}, d_{kb})$, (4.79)-(4.84) are straightforward.

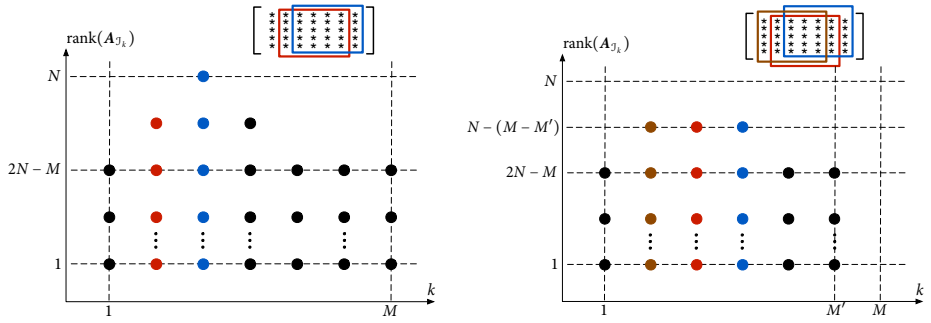
4.8.2 Proof of Lemma 4.3

In order to prove Lemma 4.3, we provide the following preliminary results stated as Lemma 4.4-4.7.

Let $\boldsymbol{A} \in \mathbb{C}^{N\times M}$, $N \leq M$, be a full rank matrix and $\boldsymbol{A'} \in \mathbb{C}^{N\times M'}$, $M' \leq M$, be a submatrix of $\boldsymbol{A}$. We have the following lemmas that will be repeatedly used in the rest of the proof.

Lemma 4.4 (rank of submatrix)**.**

$$\mathrm{rank}(\boldsymbol{A'}) \geq \mathrm{rank}(\boldsymbol{A}) - (M - M'). \tag{4.213}$$



(a) An example for Lemma 4.5 (b) An example for Lemma 4.6

Figure 4.5: Illustrations of the worst-case ranks of the submatrices from a sliding window. For each k, the number of vertical dots represents the rank of the submatrix $\boldsymbol{A}_{\mathcal{I}_k}$. In particular, the number of red (resp. blue) dots is the rank of the submatrix selected by the red (resp. blue) window. The sum of the ranks can be found by counting the number of dots.

Lemma 4.5. *Let $\mathcal{I}_1, \ldots, \mathcal{I}_M$ be a cyclic sliding window of size N on the set of indices $\{1, \ldots, M\}$, i.e.,*

$$\mathcal{I}_k \triangleq \{(k+i)_M + 1 \,:\, i \in [0, N-1]\}, \quad k = 1, \ldots, M. \tag{4.214}$$

If the columns of $\boldsymbol{A}$ are arranged such that $\mathrm{rank}(\boldsymbol{A}_{\mathcal{I}_k}) = N$ for some $k \in [1, M]$, then

$$\sum_{k=1}^{M} \mathrm{rank}(\boldsymbol{A}_{\mathcal{I}_k}) \geq N^2 \tag{4.215}$$

where $\boldsymbol{A}_{\mathcal{I}_k}$ is the matrix composed of N columns of $\boldsymbol{A}$ defined by $\mathcal{I}_k$, i.e., $\boldsymbol{A}_{\mathcal{I}_k} \triangleq [A_{j,i}]_{j\in[1,N], i\in\mathcal{I}_k}$.

Proof. The sketch of the proofs for the above lemma is illustrated in Fig. 4.5a. Given that there exists k such that the submatrix selected by the window is full rank N (the blue window in Fig.4.5a), the rank of the submatrix selected by the window $\mathcal{I}_{k+1}$ or $\mathcal{I}_{k-1}$ (the red window in Fig.4.5b) is lower bounded by $N-1$. By applying the same argument, it is readily shown that the rank of the submatrix selected by the window $\mathcal{I}_{k+2}$ or $\mathcal{I}_{k-2}$ is lower bounded by $N-2$. This lower bound keeps decreasing when the window slides away from the blue one, until it hits another lower bound $N-(M-N) = 2N-M$ given by Lemma 4.4. The submatrices within the sliding windows are of rank $2N-M$, which lasts $M-1-2(M-N) = 2N-M-1$ times. With the help

of Fig.4.5a, a lower bound on the sum of the ranks of all the submatrices visited by the sliding window, can be obtained by counting the dots in the figure, i.e.,

$$N + 2\sum_{i=1}^{M-N}(N-i) + (2N-M)(2N-M+1) = N^2. \tag{4.216}$$

In fact, this can be found easily by "completing the triangle", the number of dots in which is N^2. □

Lemma 4.6. *$\boldsymbol{A}' \in \mathbb{C}^{N\times M'}$, $N \le M' \le M$, is a submatrix of $\boldsymbol{A}$. We define $\mathcal{I}'_1, \ldots, \mathcal{I}'_{M'}$ as a cyclic sliding window of size N on the set of indices $\{1, \ldots, M'\}$, i.e.,*

$$\mathcal{I}'_k \triangleq \{(k+i)_{M'} + 1 \,:\, i \in [0, N-1]\}, \quad k = 1, \ldots, M'. \tag{4.217}$$

If the columns of $\boldsymbol{A}'$ are arranged such that the first $\mathrm{rank}(\boldsymbol{A}')$ columns of $\boldsymbol{A}'_{\mathcal{I}'_k}$ are linear independent for some $k \in [1, M]$, then we have

$$\sum_{k=1}^{M'} \mathrm{rank}(\boldsymbol{A}'_{\mathcal{I}'_k}) \ge N(N-(M-M')) \tag{4.218}$$

where $\boldsymbol{A}'_{\mathcal{I}'_k}$ is the submatrix of $\boldsymbol{A}'$ with N columns defined by $\mathcal{I}'_k$, i.e., $\boldsymbol{A}'_{\mathcal{I}'_k} \triangleq [A'_{j,i}]_{j\in[1,N], i\in\mathcal{I}'_k}$.

Proof. The sketch of the proofs for the above lemma is illustrated in Fig.4.5b. Given that there exists k such that the submatrix selected by the window has rank $r = N-(M-M')$ given by Lemma 4.4 and that the first r columns are linearly independent (the blue window in Fig.4.5b), the rank of the submatrix selected by the windows $\mathcal{I}'_{k-1}, \ldots, \mathcal{I}'_{k-(N-r)}$ (the red and brown windows in Fig.4.5b) is lower bounded by $r-1$. This lower bound keeps decreasing when the window slides go away from these positions, until it hits another lower bound $N-(M-N) = 2N-M$ given by Lemma 4.4. With the help of Fig.4.5b, a lower bound on the sum of the ranks of all the submatrices visited by the sliding window, can be obtained by counting the dots in the Figure. In fact, after some basic computations, it turns out that there are $N(N-(M-M'))$ dots. □

Lemma 4.7. *Given $\boldsymbol{H} = \hat{\boldsymbol{H}} + \tilde{\boldsymbol{H}} \in \mathbb{C}^{N\times M}$, $N \le M$, with the entries of $\tilde{\boldsymbol{H}}$ being i.i.d. $\mathcal{N}_{\mathbb{C}}\left(0, \sigma^2\right)$, $\sigma > 0$, then*

$$\mathbb{E}_{\tilde{\boldsymbol{H}}} \log\det(\boldsymbol{H}\boldsymbol{H}^{\mathsf{H}}) \ge (N - \mathrm{rank}(\hat{\boldsymbol{H}}))\log\sigma^2 + O_{\hat{H}}(1) \tag{4.219}$$

as σ^2 goes to 0.

Proof. According to [92, Lemma 1], for any $\boldsymbol{G} = \hat{\boldsymbol{G}} + \tilde{\boldsymbol{G}} \in \mathbb{C}^{N\times N}$ with the entries in $\tilde{\boldsymbol{G}}$ i.i.d. $\mathcal{N}_{\mathbb{C}}(0,1)$ independent of $\hat{\boldsymbol{G}}$, we have

$$\mathbb{E}_{\tilde{G}} \log\det(\boldsymbol{G}\boldsymbol{G}^{\mathsf{H}}) \geq \sum_{i=1}^{\tau} \log(\lambda_i(\hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^{\mathsf{H}})) + O(1) \tag{4.220}$$

where $\tau \leq \mathrm{rank}(\hat{\boldsymbol{G}})$ is the number of eigenvalues of $\boldsymbol{G}$ that are larger than 1. From here, it follows that

$$\mathbb{E}_{\tilde{G}} \log\det(\boldsymbol{G}\boldsymbol{G}^{\mathsf{H}}) \geq \sum_{i=1}^{\mathrm{rank}(\hat{\boldsymbol{G}})} \log(1+\lambda_i(\hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^{\mathsf{H}})) + O(1) \tag{4.221}$$

since the remaining $\mathrm{rank}(\hat{\boldsymbol{G}}) - \tau$ eigenvalues are smaller than 1 and do not contribute more than $O(1)$ to the expectation. Therefore, for any $\sigma > 0$, we can apply the above inequality to $\sigma^{-1}\boldsymbol{H}$ and have

$$\begin{aligned}
&\mathbb{E}_{\tilde{\boldsymbol{H}}} \log\det((\sigma^{-1}\boldsymbol{H})(\sigma^{-1}\boldsymbol{H})^{\mathsf{H}}) \\
&\quad\geq \sum_{i=1}^{\mathrm{rank}(\hat{\boldsymbol{H}})} \log(\lambda_i(\sigma^{-2}\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^{\mathsf{H}})) + O(1) && (4.222)\\
&\quad= -\mathrm{rank}(\hat{\boldsymbol{H}})\log\sigma^2 + \sum_{i=1}^{\mathrm{rank}(\hat{\boldsymbol{H}})} \log(\lambda_i(\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^{\mathsf{H}})) + O(1) && (4.223)\\
&\quad= -\mathrm{rank}(\hat{\boldsymbol{H}})\log\sigma^2 + O_{\hat{H}}(1) && (4.224)
\end{aligned}$$

where the last equality is from Assumption 4.1 that $\mathbb{E}_{\hat{H}}(\log\det(\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^{\mathsf{H}})) > -\infty$. □

In the following, we prove Lemma 4.3 case by case according to the value of M[3]. First, let us recall that $N \leq L$. Since the case with $M \leq N$ is trivial, we focus on the cases with $N < M < L$ and $M \geq L$.

Case A: $N < M < L$

Let us define M' as the number of eigenvalues of $\boldsymbol{K}$ that are not smaller than 1[4], and let $\boldsymbol{K} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathsf{H}}$ be the eigenvalue decomposition of $\boldsymbol{K}$. We

[3]The technique employed in this proof was first developed in our earlier version of this chapter [79], and later applied and extended to tackle the K-user MISO case in [80, 92].

[4]Or any constant $c > 0$ that is independent of any parameter in the system. Note that M' can depend on $\hat{\boldsymbol{S}}$ and the SNR P.

first establish the following upper bound:

$$\det(\mathbf{I}+\boldsymbol{SKS}^{\mathsf{H}}) = \det(\mathbf{I}+\boldsymbol{\Lambda V}^{\mathsf{H}}\boldsymbol{S}^{\mathsf{H}}\boldsymbol{SV}^{\mathsf{H}}) \tag{4.225}$$

$$\leq \det(\mathbf{I}+\lambda_{\max}(\boldsymbol{V}^{\mathsf{H}}\boldsymbol{S}^{\mathsf{H}}\boldsymbol{SV})\boldsymbol{\Lambda}) \tag{4.226}$$

$$\leq \det(\mathbf{I}+\|\boldsymbol{S}\|_{\mathrm{F}}^2\,\boldsymbol{\Lambda}) \tag{4.227}$$

where the last inequality is due to $\lambda_{\max}(\boldsymbol{V}^{\mathsf{H}}\boldsymbol{S}^{\mathsf{H}}\boldsymbol{SV}) \leq \|\boldsymbol{SV}\|_{\mathrm{F}}^2 = \|\boldsymbol{S}\|_{\mathrm{F}}^2$. Therefore, we have

$$\mathbb{E}_{\tilde{\boldsymbol{S}}}\log\det(\mathbf{I}+\boldsymbol{SKS}^{\mathsf{H}}) \leq \log\det(\mathbf{I}+\mathbb{E}_{\tilde{\boldsymbol{S}}}(\|\boldsymbol{S}\|_{\mathrm{F}}^2)\,\boldsymbol{\Lambda}) \tag{4.228}$$

$$\leq \log\det(\mathbf{I}+\mathbb{E}_{\tilde{\boldsymbol{S}}}(\|\boldsymbol{S}\|_{\mathrm{F}}^2)\,\boldsymbol{\Lambda}') + O_{\hat{S}}(1) \tag{4.229}$$

where the first inequality is from (4.227) on which we apply Jensen's inequality; $\boldsymbol{\Lambda}'$ is a diagonal matrix composed of the M' largest eigenvalues of $\boldsymbol{K}$.

Next, let $\boldsymbol{\Phi} \triangleq \boldsymbol{HV}$, $\boldsymbol{\Phi}' \triangleq \hat{\boldsymbol{H}}\boldsymbol{V}$. Without loss of generality, we assume that the columns of $\boldsymbol{\Phi}$ and $\boldsymbol{\Phi}'$ are arranged such that the conditions in Lemma 4.5 and Lemma 4.6 are satisfied (i.e., $\mathrm{rank}(\boldsymbol{\Phi}_{\mathcal{I}}) = N$, where $\mathcal{I}$ is the cyclic window with size N, and $\boldsymbol{\Phi}_{\mathcal{I}}$ is defined as in Lemma 4.5), respectively. This also implies that the eigenvalues in $\boldsymbol{\Lambda}$ and $\boldsymbol{\Lambda}'$ are arranged accordingly. In the following, given different values of M', we prove that

$$\mathbb{E}_{\tilde{\boldsymbol{H}}}\log\det(\mathbf{I}+\boldsymbol{HKH}^{\mathsf{H}}) \geq \frac{N}{M}\log\det(\boldsymbol{\Lambda}') + \frac{N(M-N)}{M}\log\sigma^2 + O_{\hat{H}}(1). \tag{4.230}$$

Case $M' = M$

In this case, we have

$$\det(\mathbf{I}+\boldsymbol{HKH}^{\mathsf{H}}) = \det(\mathbf{I}+\boldsymbol{\Phi\Lambda\Phi}^{\mathsf{H}}) \tag{4.231}$$

$$= \sum_{\mathcal{I}\subseteq\{1,\ldots,N\}} \det(\boldsymbol{\Lambda}_{\mathcal{I}})\det(\boldsymbol{\Phi}_{\mathcal{I}}^{\mathsf{H}}\boldsymbol{\Phi}_{\mathcal{I}}) \tag{4.232}$$

$$\geq \sum_{k=1}^{M} \det(\boldsymbol{\Phi}_{\mathcal{I}_k}^{\mathsf{H}}\boldsymbol{\Phi}_{\mathcal{I}_k})\det(\boldsymbol{\Lambda}_{\mathcal{I}_k}) \tag{4.233}$$

where (4.232) is an application of the identity $\det(I+\boldsymbol{A}) = \sum_{\mathcal{I}\subseteq\{1,\ldots,M\}}\det(\boldsymbol{A}_{\mathcal{II}})$ for any $\boldsymbol{A}\in\mathbb{C}^{M\times M}$ [93]; the lower bound is obtained by only considering a sliding window of size N for all the possible sub-determinant. Thus,

$$\log\det(\mathbf{I}+\boldsymbol{HKH}^{\mathsf{H}})$$

$$\geq \log\left(\sum_{k=1}^{M}\det(\boldsymbol{\Phi}_{\mathcal{I}_k}^{\mathsf{H}}\boldsymbol{\Phi}_{\mathcal{I}_k})\det(\boldsymbol{\Lambda}_{\mathcal{I}_k})\right) \tag{4.234}$$

$$\geq \log\left(\frac{1}{M}\sum_{k=1}^{M}\det(\boldsymbol{\Phi}_{\mathcal{I}_k}^{\mathsf{H}}\boldsymbol{\Phi}_{\mathcal{I}_k})\det(\boldsymbol{\Lambda}_{\mathcal{I}_k})\right) \tag{4.235}$$

$$\geq \frac{1}{M}\log\left(\prod_{k=1}^{M}\det(\boldsymbol{\Phi}_{\mathcal{I}_k}^{\mathsf{H}}\boldsymbol{\Phi}_{\mathcal{I}_k})\det(\boldsymbol{\Lambda}_{\mathcal{I}_k})\right) \tag{4.236}$$

$$= \frac{1}{M}\left(N\log\det(\boldsymbol{\Lambda}) + \sum_{k=1}^{M}\log\det(\boldsymbol{\Phi}_{\mathcal{I}_k}^{\mathsf{H}}\boldsymbol{\Phi}_{\mathcal{I}_k})\right) \tag{4.237}$$

where (4.236) holds since arithmetic mean is not smaller than geometric mean; the last equality is from the sliding window property $\prod_{k=1}^{M}\det(\boldsymbol{\Lambda}_{\mathcal{I}_k}) = \det(\boldsymbol{\Lambda})^N$. Finally, we have

$$\begin{aligned}
&\mathbb{E}_{\tilde{\boldsymbol{H}}}\log\det(\mathbf{I}+\boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^{\mathsf{H}}) \\
&\quad\geq \frac{1}{M}\left(N\log\det(\boldsymbol{\Lambda}) + \sum_{k=1}^{M}\mathbb{E}_{\tilde{\boldsymbol{H}}}\log\det(\boldsymbol{\Phi}_{\mathcal{I}_k}^{\mathsf{H}}\boldsymbol{\Phi}_{\mathcal{I}_k})\right) && (4.238)\\
&\quad\geq \frac{1}{M}\left(N\log\det(\boldsymbol{\Lambda}) + \log\sigma^2\sum_{k=1}^{M}(N-\operatorname{rank}(\hat{\boldsymbol{\Phi}}_{\mathcal{I}_k}))\right) \\
&\qquad\qquad + O_{\hat{H}}(1) && (4.239)\\
&\quad= \frac{1}{M}\left(N\log\det(\boldsymbol{\Lambda}) + \log\sigma^2\left(MN - \sum_{k=1}^{M}\operatorname{rank}(\hat{\boldsymbol{\Phi}}_{\mathcal{I}_k})\right)\right) \\
&\qquad\qquad + O_{\hat{H}}(1) && (4.240)\\
&\quad\geq \frac{N}{M}\left(\log\det(\boldsymbol{\Lambda}') + (M-N)\log\sigma^2\right) + O_{\hat{H}}(1) && (4.241)
\end{aligned}$$

where $\hat{\boldsymbol{\Phi}} \triangleq \hat{\boldsymbol{H}}\boldsymbol{V}$ and hence $\operatorname{rank}(\hat{\boldsymbol{\Phi}}) = \operatorname{rank}(\hat{\boldsymbol{H}})$; (4.239) is from Lemma 4.7; the last inequality is from Lemma 4.5 and that $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}'$ as $M = M'$.

Case $M > M' \geq N$

For this case, we can first get

$$\begin{aligned}
\det(\mathbf{I}+\boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^{\mathsf{H}}) &= \det(\mathbf{I}+\boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^{\mathsf{H}}) \\
&\geq \det(\mathbf{I}+\boldsymbol{\Phi}'\boldsymbol{\Lambda}'(\boldsymbol{\Phi}')^{\mathsf{H}}). && (4.242)
\end{aligned}$$

Following the same footsteps as in (4.238)-(4.240), we obtain

$$\begin{aligned}
&\mathbb{E}_{\tilde{\boldsymbol{H}}} \log\det(\mathbf{I} + \boldsymbol{HKH}^{\mathsf{H}}) \\
&\quad \geq \frac{1}{M'}\left(N\log\det(\boldsymbol{\Lambda}') + \log\sigma^2\left(M'N - \sum_{k=1}^{M'} \operatorname{rank}(\hat{\boldsymbol{\Phi}}'_{\mathcal{I}'_k})\right)\right) \\
&\qquad\qquad + O_{\hat{H}}(1) && (4.243)\\
&\quad \geq \frac{N}{M'}\left(\log\det(\boldsymbol{\Lambda}') + (M-N)\log\sigma^2\right) + O_{\hat{H}}(1) && (4.244)\\
&\quad \geq \frac{N}{M}\log\det(\boldsymbol{\Lambda}') + \frac{N(M-N)}{M}\log\sigma^2 + O_{\hat{H}}(1) && (4.245)
\end{aligned}$$

where the inequality (4.244) is from Lemma 4.6.

Case $M' < N$

From (4.242) and given that $M' < N$, we have

$$\begin{aligned}
&\mathbb{E}_{\tilde{\boldsymbol{H}}} \log\det(\mathbf{I} + \boldsymbol{HKH}^{\mathsf{H}}) \\
&\quad \geq \log\det(\boldsymbol{\Lambda}') + \log\sigma^2\left(M' - \operatorname{rank}(\hat{\boldsymbol{\Phi}}')\right) + O_{\hat{H}}(1) && (4.246)\\
&\quad \geq \log\det(\boldsymbol{\Lambda}') + \log\sigma^2\left(M' - (N - (M - M'))\right) + O_{\hat{H}}(1) && (4.247)\\
&\quad = \log\det(\boldsymbol{\Lambda}') + (M-N)\log\sigma^2 + O_{\hat{H}}(1) && (4.248)\\
&\quad \geq \frac{N}{M}\log\det(\boldsymbol{\Lambda}') + \frac{N(M-N)}{M}\log\sigma^2 + O_{\hat{H}}(1) && (4.249)
\end{aligned}$$

where (4.247) is from $\log\sigma^2 \leq 0$ and $\operatorname{rank}(\hat{\boldsymbol{\Phi}}') \geq N - (M - M')$.

By now, (4.230) has been proved in all configurations of (M, N, M'). Combining (4.229) and (4.230), we have

$$\begin{aligned}
&N\,\mathbb{E}_{\tilde{\boldsymbol{S}}} \log\det(\mathbf{I} + \boldsymbol{SKS}^{\mathsf{H}}) - M\,\mathbb{E}_{\tilde{\boldsymbol{H}}} \log\det(\mathbf{I} + \boldsymbol{HKH}^{\mathsf{H}}) \\
&\qquad \leq -N(M-N)\log\sigma^2 + O_{\hat{S}}(1) + O_{\hat{H}}(1) \\
&\qquad\qquad + N\,\log\det\left(\mathbb{E}_{\tilde{\boldsymbol{S}}}(\|\boldsymbol{S}\|_{\mathrm{F}}^2)\,\mathbf{I} + (\boldsymbol{\Lambda}')^{-1}\right) && (4.250)\\
&\qquad \leq -N(M-N)\log\sigma^2 + O_{\hat{S}}(1) + O_{\hat{H}}(1) && (4.251)
\end{aligned}$$

where the last inequality is from the fact that $\boldsymbol{\Lambda}' \succeq \mathbf{I}$ by construction and hence $\log\det\left(\mathbb{E}_{\tilde{\boldsymbol{S}}}(\|\boldsymbol{S}\|_{\mathrm{F}}^2)\,\mathbf{I} + (\boldsymbol{\Lambda}')^{-1}\right) \leq M'\log(1 + \mathbb{E}_{\tilde{\boldsymbol{S}}}(\|\boldsymbol{S}\|_{\mathrm{F}}^2)) = O_{\hat{S}}(1)$. This completes the proof of (4.151) for the case $N < M < L$.

Case B: $M \geq L$

For the first term in (4.151), we bound it as follows

$$\begin{aligned}
&\mathbb{E}_{\tilde{\boldsymbol{S}}} \log\det(\mathbf{I} + \boldsymbol{S}\boldsymbol{K}\boldsymbol{S}^{\mathsf{H}}) \\
&\quad = \mathbb{E}_{\tilde{\boldsymbol{S}}} \log\det(\mathbf{I} + \boldsymbol{U}_{\boldsymbol{S}}\boldsymbol{\Sigma}_{\boldsymbol{S}}\boldsymbol{V}_{\boldsymbol{S}}^{\mathsf{H}}\boldsymbol{K}\boldsymbol{V}_{\boldsymbol{S}}\boldsymbol{\Sigma}_{\boldsymbol{S}}\boldsymbol{U}_{\boldsymbol{S}}^{\mathsf{H}}) && (4.252)\\
&\quad = \mathbb{E}_{\tilde{\boldsymbol{S}}} \log\det(\mathbf{I} + \boldsymbol{\Sigma}_{\boldsymbol{S}}^2\boldsymbol{V}_{\boldsymbol{S}}^{\mathsf{H}}\boldsymbol{K}\boldsymbol{V}_{\boldsymbol{S}}) && (4.253)\\
&\quad \leq \mathbb{E}_{\tilde{\boldsymbol{S}}} \log\det(\mathbf{I} + \lambda_{\max}(\boldsymbol{\Sigma}_{\boldsymbol{S}}^2)\boldsymbol{V}_{\boldsymbol{S}}^{\mathsf{H}}\boldsymbol{K}\boldsymbol{V}_{\boldsymbol{S}}) && (4.254)\\
&\quad = \sum_{i=1}^{L} \mathbb{E}_{\tilde{\boldsymbol{S}}} \log(1 + \lambda_{\max}(\boldsymbol{S}\boldsymbol{S}^{\mathsf{H}})\lambda_i(\boldsymbol{V}_{\boldsymbol{S}}^{\mathsf{H}}\boldsymbol{K}\boldsymbol{V}_{\boldsymbol{S}})) && (4.255)\\
&\quad \leq \sum_{i=1}^{L} \mathbb{E}_{\tilde{\boldsymbol{S}}} \log(1 + \lambda_{\max}(\boldsymbol{S}\boldsymbol{S}^{\mathsf{H}})\lambda_i) && (4.256)\\
&\quad \leq \sum_{i=1}^{L} \mathbb{E}_{\tilde{\boldsymbol{S}}} \log(1 + \|\boldsymbol{S}\|_{\mathrm{F}}^2\lambda_i) && (4.257)\\
&\quad \leq \sum_{i=1}^{L} \log(1 + \mathbb{E}_{\tilde{\boldsymbol{S}}}(\|\boldsymbol{S}\|_{\mathrm{F}}^2)\,\lambda_i) && (4.258)\\
&\quad = \log\det(\mathbf{I} + \mathbb{E}_{\tilde{\boldsymbol{S}}}(\|\boldsymbol{S}\|_{\mathrm{F}}^2)\,\boldsymbol{\Lambda}'') && (4.259)\\
&\quad = \log\det(\mathbf{I} + \mathbb{E}_{\tilde{\boldsymbol{S}}}(\|\boldsymbol{S}\|_{\mathrm{F}}^2)\,\boldsymbol{\Lambda}''') + O_{\hat{S}}(1) && (4.260)
\end{aligned}$$

where in (4.252), $\boldsymbol{S} = \boldsymbol{U}_{\boldsymbol{S}}\boldsymbol{\Sigma}_{\boldsymbol{S}}\boldsymbol{V}_{\boldsymbol{S}}^{\mathsf{H}}$ with $\boldsymbol{\Sigma}_{\boldsymbol{S}} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{V}_{\boldsymbol{S}} \in \mathbb{C}^{M\times L}$; (4.253) comes from the equality $\det(\mathbf{I} + \boldsymbol{A}\boldsymbol{B}) = \det(\mathbf{I} + \boldsymbol{B}\boldsymbol{A})$; (4.256) is due to Poincare Separation Theorem [93] that $\lambda_i(\boldsymbol{V}_{\boldsymbol{S}}^{\mathsf{H}}\boldsymbol{K}\boldsymbol{V}_{\boldsymbol{S}}) \leq \lambda_i(\boldsymbol{K})$ for $i = 1, \cdots, N$; (4.257) is from the fact that $\lambda_{\max}(\boldsymbol{S}\boldsymbol{S}^{\mathsf{H}}) \leq \|\boldsymbol{S}\|_{\mathrm{F}}^2$; (4.258) is obtained by applying Jensen's inequality; $\boldsymbol{\Lambda}'' \triangleq \mathrm{diag}(\lambda_1, \cdots, \lambda_L)$ and $\boldsymbol{\Lambda}''' \triangleq \mathrm{diag}(\lambda_1, \cdots, \lambda_{\min\{L,M'\}})$ with M' being the number of eigenvalues that are not smaller than 1, i.e., $\Lambda''' \succeq \mathbf{I}$.

For the second term in (4.151), we use the following lower bound

$$\begin{aligned}
&\mathbb{E}_{\tilde{\boldsymbol{H}}} \log\det(\mathbf{I} + \boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^{\mathsf{H}}) \\
&\quad = \mathbb{E}_{\tilde{\boldsymbol{H}}} \log\det(\mathbf{I} + \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^{\mathsf{H}}) && (4.261)\\
&\quad \geq \mathbb{E}_{\tilde{\boldsymbol{H}}} \log\det(\mathbf{I} + \boldsymbol{\Phi}'\boldsymbol{\Lambda}''\boldsymbol{\Phi}'^{\mathsf{H}}) && (4.262)\\
&\quad \geq \frac{N}{L}\log\det(\boldsymbol{\Lambda}''') + \frac{N(L-N)}{L}\log(\sigma^2) + O_{\hat{H}}(1) && (4.263)
\end{aligned}$$

where $\boldsymbol{\Phi} \triangleq \boldsymbol{H}\boldsymbol{V} \in \mathbb{C}^{N\times M}$ with $\boldsymbol{V}$ being the unitary matrix containing the eigenvectors of $\boldsymbol{K}$, i.e., $\boldsymbol{K} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathsf{H}}$ with $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_M)$; in

(4.262), $\boldsymbol{\Phi}' = \boldsymbol{H}\boldsymbol{V}' \in \mathbb{C}^{N\times L}$ with $\boldsymbol{V}'$ being the first L columns of $\boldsymbol{V}$, and multiplying by the matrix $\boldsymbol{V}'$ does not change the distribution property, and $\boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^{\mathsf{H}} \succeq \boldsymbol{\Phi}'\boldsymbol{\Lambda}''\boldsymbol{\Phi}'^{\mathsf{H}}$; the last inequality is obtained from (4.230) in the previous subsection.

Finally, it is readily shown that, following the same steps as in (4.250) and (4.251),

$$\frac{1}{L}\mathbb{E}_{\tilde{\boldsymbol{S}}} \log\det(\mathbf{I} + \boldsymbol{S}\boldsymbol{K}\boldsymbol{S}^{\mathsf{H}}) - \frac{1}{N}\mathbb{E}_{\tilde{\boldsymbol{H}}} \log\det(\mathbf{I} + \boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^{\mathsf{H}}) \\ \leq -\frac{L-N}{L}\log(\sigma^2) + O_{\hat{S}}(1) + O_{\hat{H}}(1). \quad (4.264)$$

This completes the proof of (4.151) for the case $M \geq L$.

Part II

Interference Management with Topological Feedback

Chapter 5

Topological Interference Management: The Optimality of Orthogonal Access

Whereas interference management with delayed feedback was studied in the first half of this thesis, the focus will be placed on the another source of channel uncertainty, namely the sole topological feedback, in this and next chapters.

Interference networks with no CSIT except for the knowledge of the connectivity graph via topological feedback have been recently studied under the topological interference management (TIM) framework [49]. In this chapter, we investigate the limitation of the sole topological feedback under this TIM framework, particularly focusing on the one-dimensional convex cellular networks, in which both the transmitters and the receivers are placed along a straight line, and the signal coverage of each transmitter is convex such that every transmitter interferes consecutive receivers. The optimal DoF are fully characterized in such a single-antenna cellular network under the TIM setting. Specifically, it is shown that orthogonal access achieves the optimal symmetric DoF, sum DoF, and DoF region of these one-dimensional convex cellular networks with arbitrary message demands (i.e., the general multiple unicast setting). This conclusion can be also extended to a larger class of network topologies when there do not exist long cycles (with length no less than six) in topology graphs.

5.1 Introduction

The benefits of multiuser communication in either IC or BC stem from the availability of CSIT. Most efforts on limited [23, 24, 39], imperfect [34, 39], or delayed feedback settings [41, 43, 54, 55, 94], among others [33, 48, 76, 78], rely on the assumption that the transmitters are endowed with various forms of certain amount channel information, so that a good fraction of the DoF achieved in the perfect CSIT can be obtained. Such an assumption is hard to realize in many practical scenarios, such as cellular networks [95], especially when the feedback resource is extremely scarce. Conversely, it has been reported [29–32] that DoF collapse in IC or BC scenario with no CSIT. A closer examination of these pessimistic results however reveals that many of the considered networks are fully connected, in that any transmitter interferes with any non-intended receiver in the network.

Owing to the nodes' random placement, the fact that power decays fast with distance, the existence of obstacles, and local shadowing effects, we may argue that certain interference links are unavoidably much weaker than others, suggesting the use of a partially-connected graph to model, at least approximately, the network topology. An interesting question then arises as to whether the partial connectivity could be leveraged to allow the use of some relaxed form of CSIT while still achieving a substantial DoF performance. In particular the exploitation of topological information, simply indicating which of the interfering links are weak enough to be approximated by zero interference and which links are too strong to do so, is of great practical interest.

Recently, interference networks with no CSI except for the knowledge of the connectivity graph at the transmitters have been formulated as the topological interference management (TIM) problem [49]. This TIM problem was nicely bridged to the "index coding" problem [52, 96], which attains a lot of attention in the past decade from both information theory and computer science communities. It has been demonstrated to be insightful and convenient to look into the TIM and index coding problems from an interference alignment perspective [8, 97]. The optimal linear solutions to the TIM and index coding problems are usually revealed by the optimal vector subspace assignment under interference alignment principles [98]. Another line of works by [99, 100] reveals the rate of index coding problem from a graph theoretic perspective, in which local chromatic number offers a new upper bound of broadcast rate to index coding problems, correspondingly a new achievability to TIM problems. Remarkably, this local coloring approach can be somehow interpreted as the one-to-one interference alignment, and

thus we refer to these approaches as alignment-based approaches.

As alignment-based approaches stimulate the further advance on TIM and index coding problems, such as TIM with alternating CSIT [101, 102], and TIM with multiple antenna [103], to name a few, some depressed facts were observed that those sophisticated alignment-based approaches offer no gain over the orthogonal schemes (which can be also interpreted from an interference alignment perspective [49]) for some class of networks. For instance, in [104], the one-dimensional connectivity convexity totally prohibits profit from the alignment-based approaches. Orthogonal access was proven to be sum DoF optimal in the one-dimensional convex cellular network with convex message set [104]. The one-dimensional convex network topology refers to the placement of all transmitters and receivers along a straight line, such that the transmitter that interferes a receiver will cause interference to all other closer receivers and the receiver that hears a transmitter will hear signals from all other closer transmitters. This connectivity convexity captures the physical phenomenon that signal strength is relying on the physical locations of transmitters and receivers, where the received signals are stronger with physically closer nodes than those are farther away. In real-life cellular networks, however, this convexity might be not so practical at the base station side, due to the various power allocation at the base stations. For instance, the user may hear the signal originates from a faraway base station with higher transmit power, while it does not hear from nearby ones with lower transmit power. A natural question then arises as to whether orthogonal access is still sum DoF optimal with convexity at only one side.

Up to date, the one-dimensional convex cellular networks [104] might be the only explicit subclass of network topologies where orthogonal access is sum DoF optimal. What else? In addition, while the sum DoF metric offers us the performance of the whole network, it does not capture the individual behavior of distinct users, where the network topology asymmetry does matter. Is orthogonal access still optimal as to the performance metrics, such as the DoF region. As we may be aware of, the understanding of sum DoF is far from the characterization of overall DoF region, such that the extension from the former to the latter is really a huge leap and hence totally nontrivial. Then the consequently arisen question might be whether the sum DoF or the symmetric DoF offer us some insight toward the characterization of DoF region.

Furthermore, the optimality of orthogonal access in [104] is restricted to the convex message sets, where the transmitter has to possess messages to consecutive receivers. This constraint is not practical and does hurt in real-life applications. For instance, in a social-oriented application, a transmitter may

be interested in exchanging messages with any other receivers, not need to the nearby ones. A crucial question then arises as to whether the optimality of orthogonal access still holds when message sets are not convex any more.

In this chapter, we try to address these aforementioned questions, still under the multiple unicast TIM framework mainly in the one-dimensional convex cellular networks. Specifically, we obtain the following results:

- Orthogonal access is DoF optimal in all of one-dimensional convex interference channels even if the convexity in network topology is only at one (either transmitter or receiver) side. The optimal symmetric DoF, sum DoF, and DoF region are characterized with aid of some well-defined graph theoretic parameters.

- Orthogonal access is DoF optimal for a class of cellular networks (where the network topology graph contains no cycles with length greater than four, including one-dimensional convex networks as special cases) with arbitrary message sets. The optimal solutions (i.e., symmetric DoF, sum DoF and DoF region) of multiple unicast TIM problems in such a subclass of connectivity graphs are comprehensively unveiled.

The rest of the chapter is organized as follows. The general system model and some basic definitions are given in the coming section. In Section 5.3, we present the main results of the one-dimensional convex interference channels and chordal cellular networks, followed by the proofs in Section 5.4. We conclude the chapter in Section 5.5, together with some interesting discussions.

Notations: Throughout this chapter, we let A and $\mathcal{A}$ represent a variable and a set, respectively. In addition, $\bar{\mathcal{A}}$ is the complementary set of $\mathcal{A}$, and $|\mathcal{A}|$ is the cardinality of the set $\mathcal{A}$. Denote by $A_{\mathcal{S}} \triangleq \{A_i, i \in \mathcal{S}\}$ and $\mathcal{A}_{\mathcal{S}} \triangleq \cup_{i\in\mathcal{S}}\mathcal{A}_i$. Define $\mathcal{A}\backslash a \triangleq \{x|x \in \mathcal{A}, x \neq a\}$ and $\mathcal{A}_1\backslash\mathcal{A}_2 \triangleq \{x|x \in \mathcal{A}_1, x \notin \mathcal{A}_2\}$.

5.2 System Model

5.2.1 Channel Model

We consider a cellular network with M base stations (a.k.a. sources, transmitters), which are labeled as S_1, S_2, ..., S_M, and N user terminals (a.k.a. destinations, receivers), which are labeled as D_1, D_2, ..., D_N. Both transmitters and receivers are equipped with one single antenna each. The partial connectivity of the network is modeled through the received signal equation for

Receiver j at time instant t by:

$$Y_j(t) = \sum_{i \in \mathcal{T}_j} h_{ji}(t) X_i(t) + Z_j(t) \tag{5.1}$$

where $h_{ji}(t)$ is the channel coefficient between source i and destination j, the transmitted signal $X_i(t)$ is subject to the average power constraint, i.e., $\mathbb{E}\left(|X_i(t)|^2\right) \leq P$, with P being the average transmit power, and $Z_j(t)$ is the Gaussian noise with zero-mean and unit-variance and is independent of transmitted signals and channel coefficients. We denote by $\mathcal{T}_n$ the transmit set containing the indices of sources that are *connected* to destination n, for $n \in \{1, 2, \ldots, N\}$, and by $\mathcal{R}_n$ the receive set consisting of the indices of destinations that are *connected* to source m, for $m \in \{1, 2, \ldots, M\}$. In practice, the partial connectivity may be modeled by taking those interference links that are "weak enough" (due to distance and/or shadowing) to zero. For instance in [49], a reasonable model was suggested whereby a link is disconnected if the received signal power falls below the effective noise floor. An interesting discussion in this regard will be given in Section 5.5, together with some numerical results, showing how the results obtained in TIM settings reflect to the real-life wireless networks. However, other models maybe envisioned and the study of how robust the derived schemes are with respect to modeling errors is an open problem beyond the scope of this thesis.

Conforming with the TIM framework, the actual channel realizations are not available at the sources, yet the network topology (i.e., $\mathcal{T}_m, \mathcal{R}_n, \forall m, n$) is known by all sources and destinations. We follow the similar assumption on channel coefficients as in [104], where the nonzero channel coefficients can keep constant or vary over time, but are statistically indistinguishable one another by transmitters. At the destination side, only desired channel coefficients are required to known by destinations. The network topology is assumed to be fixed throughout the duration of communication.

5.2.2 Message Sets

Regarding the message set, we follow the general multiple unicast model in [49], where each message originates from one unique source and intends for one unique destination. As such, each source may have multiple independent messages that could intend for multiple destinations and each destination may desire multiple independent messages that could originate from multiple sources.

- Message set at source S_i, i.e., $\mathcal{W}(S_i),\ i \in \{1, \ldots, M\}$

- Message set at destination D_j, i.e., $\mathcal{W}(D_j),\ j \in \{1, \ldots, N\}$

In particular, interference channels with each source possessing one message that intends for a unique destination, and X networks with each source having a message to each connected destination are special cases.

According to the message sets at sources and destinations, there exist three mutually exclusive possibilities between source S_i and destination D_j:

- **Weak link** ($D_j \nrightarrow S_i$): if the channel is weak and the corresponding edge is absent in network topology.
- **Desired link** ($S_i \rightarrow D_j$): if there is a message W_{ij} originates from source S_i to destination D_j, which is illustrated as a solid black edge.
- **Interfering link** ($S_i - D_j$): if there is a connection between source S_i and destination D_j, but there are no desired messages between them. These links are referred to as interfering links and shown as dashed red edges.

We denote by $S_i \rightarrowtail D_j$ if either $S_i \rightarrow D_j$ or $S_i - D_j$, to indicate that there is a strong link between S_i and D_j.

5.2.3 Definitions

Throughout this part (this chapter and the next chapter), we treat the network topologies of cellular networks as undirected bipartite graphs, where the sources and destinations are vertices, and the connectivities between them are represented as edges. A few definitions of basic graph parameters pertaining to graph theory [105–107] are now recalled, while some more definitions will be given in later sections when needed. Note that all graphs considered in this thesis are finite, simple and loopless.

Definition 5.1 (Basic Graph Parameters [106])**.**

- *The* **chromatic number** *of $\mathcal{G}$, denoted by $\chi(\mathcal{G})$, is the smallest number of colors that assign to the vertices of $\mathcal{G}$, such that no two adjacent vertices have colors in common.*
- *A* **clique** *is a subgraph of a graph $\mathcal{G}$ where any two vertices in this subgraph are adjacent.*
- *The* **independent set** *of a graph $\mathcal{G}$ is a set of vertices such that any two vertices are not adjacent. The* **independent set number***, denoted by $\alpha(\mathcal{G})$, is the cardinality of the largest independent set.*

In this chapter, we consider the symmetric DoF (i.e., the maximum DoF which can be achieved by all users simultaneously), sum DoF (i.e., the maximum total DoF achievable by all users), and DoF region (i.e., the tuple of all achievable DoF for all users) as our main figures of merit.

Definition 5.2 (Symmetric DoF, Sum DoF, and DoF Region)**.**

$$d_{\text{sym}} = \limsup_{P\to\infty} \sup_{(R_{\text{sym}},\ldots,R_{\text{sym}})\in\mathcal{C}} \frac{R_{\text{sym}}}{\log P} \tag{5.2}$$

$$d_{\text{sum}} = \limsup_{P\to\infty} \sup_{(R_1,\ldots,R_{|\mathcal{W}|})\in\mathcal{C}} \frac{\sum_{i\in\mathcal{W}} R_i}{\log P} \tag{5.3}$$

$$\mathcal{D} = \Big\{(d_1,\ldots,d_{|\mathcal{W}|}) \in \mathbb{R}_+^{|\mathcal{W}|} \Big| \forall (w_1,\ldots,w_{|\mathcal{W}|}) \in \mathbb{R}_+^{|\mathcal{W}|} \tag{5.4}$$

$$\sum_{i\in\mathcal{W}} w_i d_i \le \limsup_{P\to\infty} \sup_{(R_1,\ldots,R_{|\mathcal{W}|})\in\mathcal{C}} \frac{\sum_{i\in\mathcal{W}} w_i R_i}{\log P} \Big\} \tag{5.5}$$

where $\mathcal{W}$ is the message set, the capacity region $\mathcal{C}$ is the set of all achievable rate tuples, and P is the average power constraint.

Definition 5.3 (Orthogonal Access)**.** *Orthogonal access is such that the involved messages $\mathcal{W}_o$ are orthogonal [104], i.e.,*

$$\mathcal{D}(W_j) \nrightarrow \mathcal{S}(W_i), \forall\ W_i, W_j \in \mathcal{W}_o,\ i \neq j \tag{5.6}$$

where $\mathcal{S}(W_i)$ is the collection of sources from which W_i originates, and $\mathcal{D}(W_j)$ is the collection of destinations for which W_j intends.

Definition 5.4 (Message Conflict Graphs)**.** *The* **message conflict graph** *is a graph with vertices being messages and edges between two vertices if the two messages are conflicting with one another. The two messages are conflicting if they originate from the same source, intend for the same destination, or one's source interferes the other's destination.*

Clearly, the orthogonal messages are not conflicting, and the conflicting messages are non-orthogonal. The orthogonal messages belong to an independent set of the message conflict graphs.

5.3 Main Results

In the following, we start with the simple special straight line networks, where the sources and destinations form a one-dimensional convex interference channel with one destination demanding one message from one single

source, followed by the more general cellular networks with arbitrary message demands. For all these networks, we show briefly that orthogonal access achieves optimal DoF. The related definitions and detailed proofs are relegated to the next section.

5.3.1 Straight Line Networks

In straight line networks, the base stations and users are placed in a straight street, with each base station (source) serving one user (destination), which forms a one-dimensional convex interference channel. The message W_i originates from a single source S_i and intends for a single destination D_i, i.e., $M = N$. Each source has only one message, i.e., $\mathcal{W}(S_i) = \{W_i\}$, and each destination demands one message, i.e., $\mathcal{W}(D_j) = \{W_j\}$.

Compared to [104], we relax the network convexity to only one side, which is more practical in the real-life scenarios. It is because the coverage by the base stations is inclined to spread over the users nearby, given power constraints at base stations. On the other hand, as the transmitter power constraints at base stations may differ from one another, even if the user could not hear a base station nearby, it may hear other base station farther away with higher transmit power. As such, the user side does not have to comply with the convexity property. An exemplary network topology is shown in Fig. 5.1.

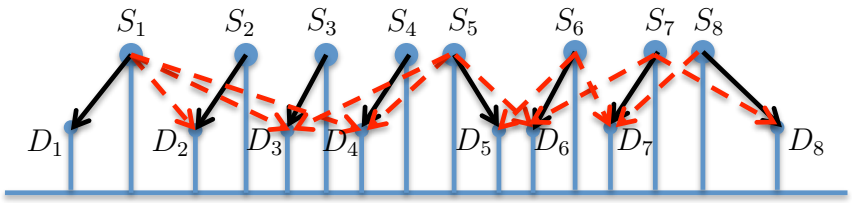


Figure 5.1: A straight line network that can be represented by a one-dimensional convex interference channel.

Let us consider hereafter the convexity only at the sources, and the case with convexity only at the destinations follows similarly. We adopt the relative placement of the destinations, where the placement that the destination node D_{j_1} is "to the left of" (resp. "to the right of") another destination node D_{j_2} is denoted by $D_{j_1} \leq D_{j_2}$ (resp. $D_{j_1} \geq D_{j_2}$). The source convexity property in the straight line networks can be highlighted in the following definition:

Definition 5.5 (Source Convexity Property)**.** *For the source convexity, if $S_i \rightarrowtail D_{j_1}$ and $S_i \rightarrowtail D_{j_2}$ where $D_{j_1} \leq D_{j_2}$, then $S_i \rightarrowtail D_k$ for all k satisfying $D_{j_1} \leq D_k \leq D_{j_2}$.*

Clearly, if a source can interfere a destination on its left (resp. right) side, then it will interfere all other destinations on the left (resp. right) that are closer. On the other hand, if a source cannot interfere a destination on its left (resp. right) side, then it will not interfere all other destinations on the left (resp. right) that are farther away. For cellular networks with this convexity property, we obtain the following result.

Theorem 5.1. *Orthogonal access is DoF optimal for multiple unicast TIM problems in all one-dimensional convex interference channels. The optimal symmetric DoF, sum DoF and DoF region achieved by orthogonal schemes are given by*

$$d_{\text{sym}} = \frac{1}{\chi(\mathcal{G})} \tag{5.7}$$

$$d_{\text{sum}} = \alpha(\mathcal{G}) \tag{5.8}$$

$$\mathcal{D} = \left\{ (d_1, \ldots, d_{|\mathcal{V}|}) \in \mathbb{R}_+^{|\mathcal{V}|} \;\middle|\; \sum_{i \in \mathscr{C}_j} d_i \leq 1, \forall\, \mathscr{C}_j \in \mathfrak{C}(\mathcal{G}) \right\} \tag{5.9}$$

where $\mathcal{G}$ is the message conflict graph, $\mathcal{V}$ is the vertex set of $\mathcal{G}$ corresponding to the message set, $\chi(\mathcal{G})$ and $\alpha(\mathcal{G})$ are chromatic number and independent set number of $\mathcal{G}$ as defined earlier, respectively, and $\mathfrak{C}(\mathcal{G})$ is the collection of all possible cliques in $\mathcal{G}$.

Proof. See Section 5.4.2. □

Remark 5.1. *As a matter of fact, the one-dimensional convex networks can be represented as convex bipartite graphs [106], where for every vertex in one disjoint set (e.g., sources), the vertices adjacent to it in the other disjoint set (e.g., destinations) are consecutive. The graphs are called biconvex graphs, if this property applies to both disjoint sets. The convex cellular networks considered in [104] are representable by biconvex bipartite graphs, while here we consider a superclass: convex bipartite graphs.*

Example 5.1. *Consider a five-user convex interference channel as shown in Fig. 5.2, where the convexity applies to the sources only. According to*

Theorem 5.1, the optimal symmetric DoF, optimal sum DoF, and DoF region are given respectively by $d_{\text{sym}} = \frac{1}{3}$, $d_{\text{sum}} = 2$, *and*

$$\mathcal{D} = \left\{ (d_1, \ldots, d_5) \in \mathbb{R}_+^5 \;\middle|\; \begin{array}{l} d_1 + d_2 + d_3 \leq 1 \\ d_1 + d_4 \leq 1, d_3 + d_5 \leq 1 \\ d_4 + d_5 \leq 1 \end{array} \right\}. \tag{5.10}$$

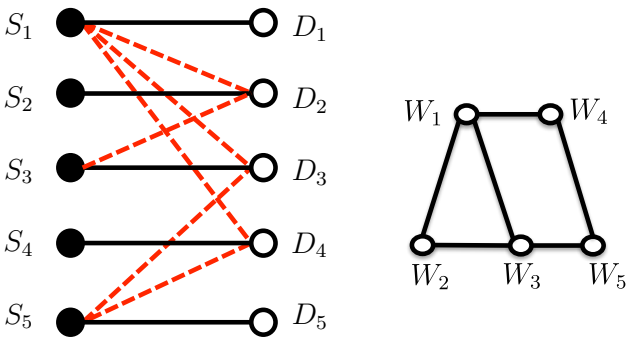


Figure 5.2: A five-user one-dimensional convex interference channel. The left is the network topology graph with convexity applied to the sources only, and the right is its conflict graph.

5.3.2 Chordal Cellular Networks

In what follows, we consider a larger class of cellular networks, whose network topologies are modeled by the following graphs.

Definition 5.6 (Chordal Bipartite Graphs [108])**.** *A chordal bipartite graph is an undirected bipartite graph in which every cycle of length at least six has a chord.*

The Chordal cellular networks are represented by chordal bipartite graphs $\mathcal{G}$. Regarding the characterization of chordal bipartite graphs, one approach is based on a form of edge elimination [109]. One example of chordal cellular networks is shown in Fig. 5.3, where the arrows indicate the transmission from sources to destinations. The corresponding network topology is the same graph ignoring link directions.

The message sets follow the general multiple unicast setting mentioned earlier, where each message originates from a unique source and intends for a unique destination. The source S_i with message set $\mathcal{W}(S_i)$ may have

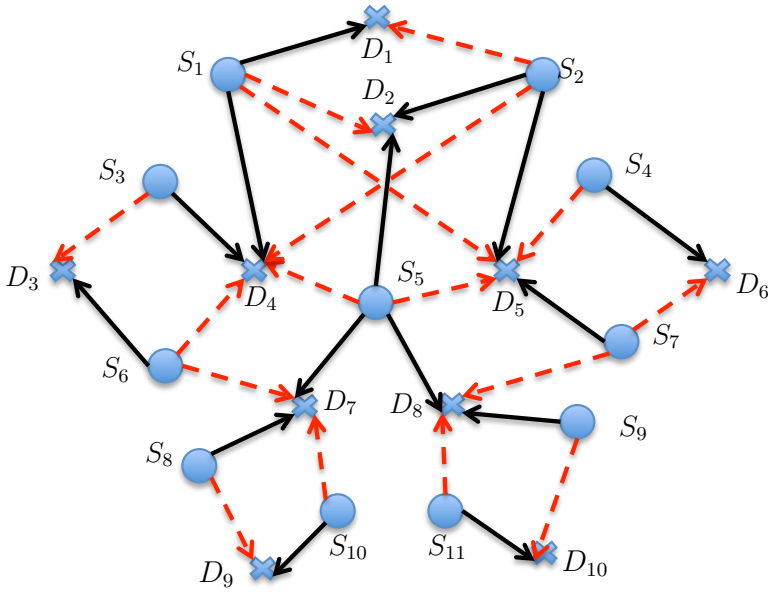


Figure 5.3: A chordal cellular network, which is a two-dimensional convex network but is not necessarily a one-dimensional convex network.

multiple messages to a subset of destinations, and the destination D_j may desire multiple messages, i.e., $\mathcal{W}(D_i)$.

Let $W_{\mathcal{M}}$ be a set of messages desired by some destinations. For instance, if $W_{ij} \in W_{\mathcal{M}}$ is a message desired by destination D_j and originated from source S_i, then $S_i \to D_j$ is a desired link. Otherwise, the connected links are interfering links.

Before presenting the results, we introduce some useful definitions in graph theory.

Definition 5.7 (Basic Graph Theoretic Definitions [106])**.**

- *A* **line graph** *of $\mathcal{G}$ is another graph, denoted by $\mathcal{G}_e$, that represents the adjacencies of the edges in $\mathcal{G}$.*
- *The* **square** *of a graph $\mathcal{G}$, denoted by $\mathcal{G}^2$, is another graph with the same vertex set, but in which two vertices are adjacent when the distance between them is no more than 2. The* **distance** *between two vertices is the number of edges in a shortest path connecting them.*
- *A subgraph of $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ containing a subset of vertices $\mathcal{S}$ ($\mathcal{S} \subseteq \mathcal{V}$) is said to be an* **induced subgraph***, denoted by $\mathcal{G}[\mathcal{S}]$, if for any pair of*

vertices u and v in $\mathcal{S}$, uv is an edge of $\mathcal{G}[\mathcal{S}]$ if and only if uv is an edge of $\mathcal{G}$.

Lemma 5.1. *For the network topology graph $\mathcal{G}$, given the message demand set $\mathcal{M}$, the induced subgraph $\mathcal{G}_e^2[\mathcal{M}]$ is the message conflict graph of $\mathcal{G}$ with respect to message set $\mathcal{M}$, in which $\mathcal{G}_e^2$ is the square of line graph of $\mathcal{G}$.*

Proof. See Section 5.4.3. □

By message conflict graphs $\mathcal{G}_e^2[\mathcal{M}]$, we obtain the following results.

Theorem 5.2. *Orthogonal access is DoF optimal for multiple unicast TIM problems in all chordal cellular networks with arbitrary message demands. The optimal symmetric DoF, sum DoF and DoF region achieved by orthogonal schemes are given by*

$$d_{\text{sym}} = \frac{1}{\chi(\mathcal{G}_e^2[\mathcal{M}])} \tag{5.11}$$

$$d_{\text{sum}} = \alpha(\mathcal{G}_e^2[\mathcal{M}]) \tag{5.12}$$

$$\mathcal{D} = \left\{ (d_1, \ldots, d_{|\mathcal{M}|}) \in \mathbb{R}_+^{|\mathcal{M}|} \;\middle|\; \sum_{i \in \mathscr{C}_j} d_i \leq 1, \forall\, \mathscr{C}_j \in \mathfrak{C}(\mathcal{G}_e^2[\mathcal{M}]) \right\} \tag{5.13}$$

where $\mathcal{M}$ is the desired message set, $\mathcal{G}_e^2[\mathcal{M}]$ is the induced subgraph of $\mathcal{G}_e^2$ associated with the message set $\mathcal{M}$, and $\mathfrak{C}(\mathcal{G}_e^2[\mathcal{M}])$ is the collection of all possible cliques in $\mathcal{G}_e^2[\mathcal{M}]$.

Proof. See Section 5.4.4. □

Remark 5.2. *This theorem is the more general setting on both network topology and message set. The interference channel is a special case, where the message set $\mathcal{M}$ is selected to be $\{W_i | S_i \to D_i, \forall\, i\}$. The X network is also a special case, where the message sets is corresponding to the network topology, i.e., $\{W_{ij} | S_i \rightarrowtail D_j, \forall\, j \in \mathcal{R}_i\}$.*

Remark 5.3. *As long as $\mathcal{G}$ is a chordal bipartite graph, then orthogonal schemes are DoF optimal regardless of message demands. For instance, no matter whether in interference channels or X networks, no matter whether with convex message set as in [104] or non-convex message set, the orthogonal schemes are DoF optimal.*

Example 5.2. *For ease of illustration, we consider a subset of messages* $\mathcal{M} = \{W_{31}, W_{42}, W_{13}, W_{54}, W_{25}, W_{55}, W_{37}\}$ *in the cellular network of Fig. 5.4 involving the source and destination pairs* $S_3 \to D_1, S_4 \to D_2, S_1 \to D_3, S_5 \to D_4, S_2 \to D_5, S_5 \to D_5, S_3 \to D_7$. *According to Theorem 5.2, we have* $d_{\text{sym}} = \frac{1}{3}$, $d_{\text{sum}} = 3$, *and DoF region*

$$\mathcal{D} = \left\{ d_{\mathcal{M}} \in \mathbb{R}_+^7 \;\middle|\; \begin{array}{l} d_{42} + d_{13} \le 1, d_{54} + d_{25} \le 1 \\ d_{31} + d_{13} + d_{37} \le 1, d_{31} + d_{54} + d_{37} \le 1 \\ d_{54} + d_{55} + d_{37} \le 1, d_{25} + d_{54} + d_{55} \le 1 \end{array} \right\} \tag{5.14}$$

where d_{ij} *is the DoF associated with message* W_{ij}.

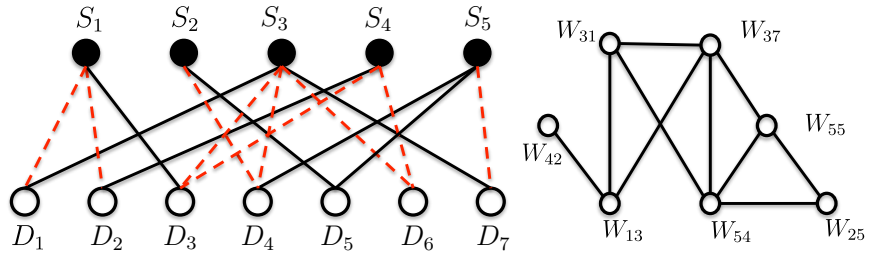


Figure 5.4: A chordal cellular network. The left is the network topology graph $\mathcal{G}$ with solid black lines being desired links and dashed red lines being interfering links, and the right is its conflict graph $\mathcal{G}_e^2[\mathcal{M}]$.

In fact, the one-dimensional convex cellular networks in [104], and its relaxed version with one side convexity, are special cases of the chordal cellular networks. Thus, we have the following corollary, which is the direct generalization of Theorem 5.1 with arbitrary message demands.

Corollary 5.1. *Orthogonal access achieves optimal symmetric DoF, optimal sum DoF, and optimal DoF region of all the one-dimensional convex cellular networks with general message demands and source and/or destination convexity.*

Proof. This corollary can be directly obtained from Theorem 5.2, owing to the fact that convex and biconvex bipartite graphs are subclass of chordal bipartite graphs. The detailed proof will be presented in Section 5.4.5. □

Besides convex/biconvex bipartite graphs, the chordal bipartite graphs include forests (graphs without cycles), bipartite permutation graphs, bipartite distance hereditary graphs, and difference graphs [108].

5.4 Proofs

5.4.1 Preliminaries

Definition 5.8 (More Graph Theoretic Definitions [106–108, 110])**.**

- *A graph $\mathcal{G}$ is said to be $n : m$-**colorable** if each vertex in $\mathcal{G}$ can be assigned a set of m colors in which the colors are drawn from a palette of n colors, such that any adjacent vertices have no colors in common. Denote by $\chi_m(\mathcal{G})$ the minimum required number, such that the* **fractional chromatic number** *$\chi_f(\mathcal{G})$ can be defined as*

$$\chi_f(\mathcal{G}) = \lim_{m\to\infty} \frac{\chi_m(\mathcal{G})}{m} = \inf_m \frac{\chi_m(\mathcal{G})}{m}. \tag{5.15}$$

- *A* **maximum clique** *is the a clique with the maximum possible size in $\mathcal{G}$. The* **clique number** *of $\mathcal{G}$, denoted by $\omega(\mathcal{G})$, is the number of vertices in the maximum clique.*

- *The* **clique covering** *of $\mathcal{G}$ is a set of cliques such that every vertex of G is a member of at least one clique. The* **clique covering number** *of $\mathcal{G}$ is the minimum number of cliques in $\mathcal{G}$ required to cover the vertex set of $\mathcal{G}$.*

- *The* **perfect graph** *is a graph $\mathcal{G}$ in which the chromatic number of every induced subgraph $\mathcal{H}$ of $\mathcal{G}$ equals to the clique number of this subgraph, i.e., $\chi(\mathcal{H}) = \omega(\mathcal{H})$.*

- *A* **chordless cycle** *is a cycle with no edges between any non-consecutive vertices. A* **hole** *is a chordless cycle with five or more vertices, and an* **antihole** *is the complement of a hole.*

- *A graph is* **chordal** *(or triangulated) if there is no induced subgraph with chordless cycles of length greater than three, i.e., every cycle with length greater than three has a chord.*

- *A graph is* **weakly chordal** *(or weakly triangulated) if it is hole-free and antihole-free in its induced graph.*

- *The* **complement** *of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, denoted by $\bar{\mathcal{G}}$, is another graph containing the same vertices set $\mathcal{V}$, but the edge uv ($\forall\ u, v \in \mathcal{V}$) is in $\bar{\mathcal{G}}$ if and only if uv is not in $\mathcal{G}$.*

- *A* **separator** *of an edge e in $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a vertex subset $\mathcal{S} \subset \mathcal{V}$ whose removal partitions the graph $\mathcal{G}[\mathcal{V}\backslash\mathcal{S}]$ into at least two connected components, and the edge e belongs to one of these components. A* **minimal separator** *is a separator with minimum number of vertices such that there do not exist any smaller size of vertices in it as a separator.*

- *A* **neighborhood** *of a vertex s, denoted by $\mathcal{N}(s)$, is a set of vertices adjacent to s. A* **neighborhood** *of an edge e, denoted by $\mathcal{N}(e)$, is a set of vertices, any of which can be seen*[1] *by either endpoint of e. The neighborhood of a set of vertices $\mathcal{S}$ is defined as $\mathcal{N}(\mathcal{S}) \triangleq \{\mathcal{N}(s), s \in \mathcal{S}\}\backslash\mathcal{S}$.*

- *With regard to a set $\mathcal{S}$ of vertices in graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, $\bar{\mathcal{G}}[\mathcal{S}]$ may have a number of disconnected components, denoted by $\mathcal{S}_j$ the j-th component. An edge e of $\mathcal{G}[\mathcal{V}\backslash\mathcal{S}]$ is said to be* **$\mathcal{S}$-saturating** *if for each component $\mathcal{S}_j$, at least one endpoint of e sees all vertices of $\mathcal{S}_j$.*

- *In $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, an edge $e \in \mathcal{E}$ is* **LB-simplicial***, if one of the following conditions holds*

  - $e \cup \mathcal{N}(e) = \mathcal{V}$
  - *For each minimal separator $\mathcal{S}$ included in the neighborhood of e, e is $\mathcal{S}$-saturating.*

By these definitions, we have the following lemma to recognize the weakly chordal graph.

Lemma 5.2 ([111])**.** *A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is weakly chordal if every edge of $\mathcal{E}$ is LB-simplicial.*

Proof. We sketch the proof here, for the details please refer to [111]. To prove this lemma, two facts were introduced.

- Given a graph $\mathcal{G}$, an edge belongs to a hole cannot be LB-simplicial.
- Given a graph $\mathcal{G}$, each antihole contains an edge that is not LB-simplicial.

By contraposition, if every edge is LB-simplicial, then holes and antiholes should not be presented in $\mathcal{G}$. Therefore it yields the conclusion. □

[1] A vertex a "sees" another vertex b refers to the adjacency of these two vertices.

Definition 5.9 (Message Demand Graph)**.** *The* **message demand graph** *is a directed graph defined as a bipartite graph with messages at one side and destinations at the other side, with a directed edge from message to a destination if this message intends for this destination, and with a directed edge from destination to message if this destination has side information of this message, i.e., the source from which this message originates is not connected to the destination.*

Lemma 5.3 ([49, 52])**.** *The symmetric DoF of TIM problems are bounded as*

$$d_{\text{sym}} \leq \frac{1}{\Psi} \tag{5.16}$$

where Ψ is the maximum cardinality of a message set that forms an acyclic message demand graph.

5.4.2 Proof of Theorem 5.1

Proof of Symmetric DoF

To prove the optimality of symmetric DoF, we have the following four key steps, which are outlined below.

1. The symmetric DoF value achieved by the orthogonal scheme is the inverse of the fractional chromatic number of the conflict graph $\mathcal{G}$, i.e., $d_{\text{sym}} \geq \frac{1}{\chi_f(\mathcal{G})}$.

2. Because the network topology is convex, the message conflict graph $\mathcal{G}$ is a weakly chordal graph, which is perfect [110]. As such, the chromatic number is equal to clique number, i.e., $\chi(\mathcal{G}) = \omega(\mathcal{G})$.

3. For each clique in the conflict graph $\mathcal{G}$, the involved messages form an acyclic demand graph, yielding the outer bound $d_{\text{sym}} \leq \frac{1}{\omega(\mathcal{G})}$.

4. By the relation $\chi(\mathcal{G}) \geq \chi_f(\mathcal{G}) = \omega_f(\mathcal{G}) \geq \omega(\mathcal{G})$, it follows that $\chi(\mathcal{G}) = \chi_f(\mathcal{G}) = \omega_f(\mathcal{G}) = \omega(\mathcal{G})$ for perfect graph $\mathcal{G}$, and thus the achievability coincides with the outer bound.

Given the conflict graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of the one-dimensional convex interference channels, we present the detailed proof in the following.

Step I

By the definition of message conflict graph, any two vertices joint with an edge in $\mathcal{G}$ are conflicting, such that the associated messages cannot be delivered simultaneously without causing mutual interference. Thus, the adjacent vertices (messages) in $\mathcal{G}$ should be activated (delivered) in different time slots, and thus they should be assigned with different colors from a graph coloring viewpoint.

A proper color assignment strategy in which no adjacent vertices in $\mathcal{G}$ have any colors in common corresponds to an orthogonal scheduling scheme, i.e., the number of colors assigned to one vertex is the number of delivered symbols and the total number of required colors is the total number of required time slots. As such, the maximum symmetric DoF achieved by orthogonal schemes turn out to be the maximum ratio between the number of colors assigned to each vertex and the total required number of colors. According to the definition of fractional coloring, it is exactly the inverse of fractional chromatic number, i.e., $\frac{1}{\chi_f(\mathcal{G})}$. Thus, $d_{\text{sym}} = \frac{1}{\chi_f(\mathcal{G})}$ is best achievable symmetric DoF by orthogonal schemes.

Step II

To prove that the conflict graph $\mathcal{G}$ is perfect such that $\chi(\mathcal{G}) = \omega(\mathcal{G})$, we first verify that $\mathcal{G}$ is a weakly chordal graph by the lemma bellow.

Lemma 5.4. *The conflict graphs of one-dimensional convex interference channels are weakly chordal.*

Proof. Let us denote the network topology by $\mathcal{H}$ for presentational convenience, which is a convex bipartite graph, and its conflict graph by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. We consider an arbitrary edge $e_{ij} \in \mathcal{E}$ connecting two messages W_i and W_j in conflict graph $\mathcal{G}$, in which these two messages are conflicting, denoted by $W_i \leftrightarrow W_j$. In $\mathcal{H}$, accordingly, there exist desired links $S_i \to D_i$ and $S_j \to D_j$ in $\mathcal{H}$, with either $S_i - D_j$ or $S_j - D_i$.

In what follows, we will complete the proof according to Lemma. 5.2. If $e_{ij} \cup \mathcal{N}(e_{ij}) = \mathcal{V}$, then e_{ij} is LB-simplicial by definition. Thus, we hereafter focus on $e_{ij} \cup \mathcal{N}(e_{ij}) \subset \mathcal{V}$, and check if e_{ij} is LB-simplicial as well.

Let us consider a minimal separator of e_{ij} denoted by $\mathcal{S} \triangleq \{s_1, s_2, \ldots, s_{|\mathcal{S}|}\} \subset \mathcal{V}$ in $\mathcal{G}$ containing a subset of messages $W_{\mathcal{S}}$, which by definition must be in the neighborhood of e_{ij}, i.e., $\mathcal{S} \subseteq \mathcal{N}(e_{ij})$. Obviously, $W_i, W_j \notin W_{\mathcal{S}}$. Otherwise, $\mathcal{S}$ is not the separator of e_{ij}. Denote the involved sources and destinations of these messages in $\mathcal{H}$ by $S_{\mathcal{S}}$ and $D_{\mathcal{S}}$, respectively.

Consider the convexity at the sources, and let the sources in $S_{\mathcal{S}}$ be arbitrarily placed and the destinations in $D_{\mathcal{S}}$ be arranged from left to right, without loss of generality, as

$$D_{s_1} \leq D_{s_2} \leq \cdots \leq D_{s_{|\mathcal{S}|}} \tag{5.17}$$

where $S_{s_i} \to D_{s_i}$ is the desired link delivering the message W_{s_i}, for all $s_i \in \mathcal{S}$.

By the minimal separator $\mathcal{S}$, the induced subgraph $\mathcal{G}[\mathcal{V}\backslash\mathcal{S}]$ is partitioned into at least two connected components. Denote by $\mathcal{G}[\mathcal{A}]$ the connected component to which e_{ij} belongs, and by $\mathcal{G}[\mathcal{B}]$ the rest of the components. As such, $\mathcal{B} \neq \emptyset$ and $\mathcal{V} = \mathcal{A} \cup \mathcal{B} \cup \mathcal{S}$. Since $\mathcal{S}$ is the minimal separator and lies in the neighborhood of e_{ij}, we have

$$\mathcal{S} \subseteq \mathcal{N}(e_{ij}), \quad \mathcal{S} \subseteq \mathcal{N}(\mathcal{B}), \tag{5.18}$$

$$\mathcal{B} \nsubseteq \mathcal{N}(e_{ij}), \quad \{i, j\} \nsubseteq \mathcal{N}(\mathcal{B}). \tag{5.19}$$

It follows that, for any $s \in \mathcal{S}$, it must see at least one endpoint of e_{ij}, i.e., $W_s \leftrightarrow W_i$ (or $W_s \leftrightarrow W_j$) and at least one vertex (e.g., b) in $\mathcal{B}$, i.e., $W_s \leftrightarrow W_b$, otherwise $\mathcal{S}$ is not the minimal separator, because $\mathcal{S}\backslash s$ also separates e_{ij} from b.

In general, the relative placement of the destinations in $D_{\mathcal{S}}$ has three possibilities:

- $D_{\mathcal{S}_1}$: to the left of the destinations D_i and D_j
- $D_{\mathcal{S}_2}$: between the destinations D_i and D_j
- $D_{\mathcal{S}_3}$: to the right of the destinations D_i and D_j

where $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$.

Since $W_i \leftrightarrow W_j$ with $S_i - D_j$ or $S_j - D_i$, it follows that for any D_s between D_i and D_j, we have either $S_i - D_s$ or $S_i - D_s$ by source convexity, and thus $s \in \mathcal{N}(e_{ij})$. As $\mathcal{B} \nsubseteq \mathcal{N}(e_{ij})$, for any $b \in \mathcal{B}$, D_b should not be placed between D_i and D_j.

As such, the destinations of interest are located without loss of generality as follows

$$\underbrace{D_{s_1} \leq \cdots \leq D_{s_k}}_{D_{\mathcal{S}_1}} \leq D_i \leq \underbrace{D_{s_{k+1}} \leq \cdots \leq D_{s_l}}_{D_{\mathcal{S}_2}} \tag{5.20}$$

$$\leq D_j < \underbrace{D_{s_{l+1}} \leq \cdots \leq D_{s_{|\mathcal{S}|}}}_{D_{\mathcal{S}_3}} < D_{\mathcal{B}}. \tag{5.21}$$

It is readily verified that destinations in $D_{\mathcal{B}}$ should be placed to the right of all the destinations in $D_{\mathcal{S}}$, otherwise $\mathcal{S}$ would not be the minimal separator.

In what follows, we consider the messages in $\mathcal{S}$ regarding the above placement of their destinations.

- Case 1 - messages set $W_{\mathcal{S}_1 \cup \mathcal{S}_2}$: Since $\mathcal{S}_1 \cup \mathcal{S}_2 \subseteq \mathcal{S} \subseteq \mathcal{N}(\mathcal{B})$, there exists a destination D_b ($b \in \mathcal{B}$) such that for each $s_n \in \mathcal{S}_1 \cup \mathcal{S}_2$, we have $S_{s_n} - D_b$ to guarantee $W_{s_n} \leftrightarrow W_b$. It is because $S_b - D_{s_n}$ will result in $W_j \leftrightarrow W_b$, which contradicts $\mathcal{B} \nsubseteq \mathcal{N}(e_{ij})$. Specifically, if $S_b - D_{s_n}$, then by source convexity, we have $S_b - D_{s_m}$ for all s_m satisfying $D_{s_n} \leq D_{s_m} \leq D_b$, and then $S_b - D_j$, which lead to $W_j \leftrightarrow W_b$ and the contradiction.

  Given $S_{s_n} - D_b$, by source convexity, it follows that $S_{s_n} - D_{s_m}$ for all s_m satisfying $D_{s_n} \leq D_{s_m} \leq D_b$. Accordingly in conflict graph $\mathcal{G}$, we have $W_{s_n} \leftrightarrow W_{\{j\} \cup \mathcal{S} \backslash \{s_1, \ldots, s_n\}}$ for each s_n. It follows that $\mathcal{G}[\mathcal{S}_1 \cup \mathcal{S}_2]$ is a clique and every vertex in $\mathcal{S}_1 \cup \mathcal{S}_2$ sees $\{j\} \cup \mathcal{S} \backslash \{\mathcal{S}_1 \cup \mathcal{S}_2\}$. Thus, in $\bar{\mathcal{G}}[\mathcal{S}]$, each vertex in $\mathcal{S}_1 \cup \mathcal{S}_2$ is isolated from all others and seen by W_j.

- Case 2 - messages set $W_{\mathcal{S}_3}$: Let us start with the farthest destination in $D_{\mathcal{S}_3}$ from D_i, i.e., $D_{s_{|\mathcal{S}|}}$, followed by the closer ones in $\{D_{s_{|\mathcal{S}|-1}}, \ldots, D_{l+1}\}$ one by one recursively.

  Since $s_{|\mathcal{S}|} \in \mathcal{N}(e_{ij})$ in $\mathcal{G}$, i.e., $W_{s_{|\mathcal{S}|}} \leftrightarrow W_i$ (or $W_{s_{|\mathcal{S}|}} \leftrightarrow W_j$), there are two possible connectivities between desired links $S_i \rightarrow D_i$ (or $S_j \rightarrow D_j$) and $S_{s_{|\mathcal{S}|}} \rightarrow D_{s_{|\mathcal{S}|}}$ in $\mathcal{H}$.

  (a) When $S_{s_{|\mathcal{S}|}} - D_i$ (or $S_{s_{|\mathcal{S}|}} - D_j$): With source convexity, $S_{s_{|\mathcal{S}|}} - D_{s_m}$, and thus, $W_{s_{|\mathcal{S}|}} \leftrightarrow W_{s_m}$, for all $s_m \in \mathcal{S}_3$. As such, $W_{s_{|\mathcal{S}|}}$ is isolated component in $\bar{\mathcal{G}}(\mathcal{S})$, and seen by either W_i or W_j.

  (b) When $S_i - D_{s_{|\mathcal{S}|}}$ (or $S_j - D_{s_{|\mathcal{S}|}}$): Owing to the source convexity property, $S_i - D_{s_m}$ (or $S_j - D_{s_m}$), for all $s_m \in \mathcal{S}_3$. Therefore, $W_i \leftrightarrow W_{s_m}$ (or $W_j \leftrightarrow W_{s_m}$), for all $s_m \in \mathcal{S}_3$. In this case, no matter whether $\bar{\mathcal{G}}[\mathcal{S}_3]$ is connected, disconnected or empty, W_i (or W_j) will see all the elements in $W_{\mathcal{S}_3}$.

  For the rest vertices in $\mathcal{S}_3$, we consider two aforementioned cases whether S_i (or S_j) $- D_{s_n}$ or $S_{s_n} - D_i$ (or D_j) for $s_n = \{s_{|\mathcal{S}|-1}, \ldots, s_{l+1}\}$. As a result, for any $s_n \in \mathcal{S}_3$, there are two categories:

  – W_{s_n} is isolated component in $\bar{\mathcal{G}}[\mathcal{S}]$, and seen by W_i or W_j.

  – $W_{\{s_n, s_{n-1}, \ldots, s_{l+1}\}}$ are seen by W_i or W_j regardless of the connectivity in $\bar{\mathcal{G}}[\mathcal{S}]$.

As such, in $\bar{\mathcal{G}}[\mathcal{S}]$, the vertices in $\mathcal{S}_3$ either are isolated components with one vertex or belong to a connected component with multiple vertices.

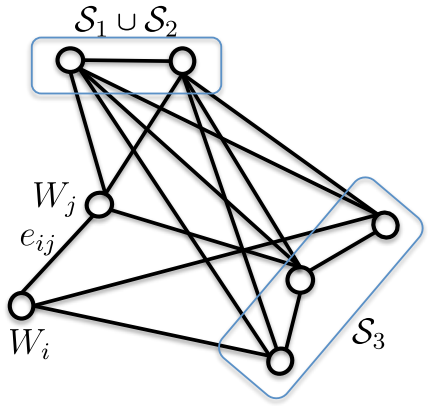


Figure 5.5: Illustration of minimal separator $\mathcal{S}$ in $\mathcal{G}$.

To sum up, $\mathcal{G}[\mathcal{S}_1 \cup \mathcal{S}_2]$ is a clique and each vertex in $\mathcal{S}_1 \cup \mathcal{S}_2$ sees the vertices in $\{j\} \cup \mathcal{S}_3$, such that each vertex in $\mathcal{S}_1 \cup \mathcal{S}_2$ is an isolated component in $\bar{\mathcal{G}}[\mathcal{S}]$ and seen by W_j. Although $\mathcal{G}[\mathcal{S}_3]$ is not generally a clique, the vertices in $\mathcal{S}_3$ are either isolated components with single vertex in $\bar{\mathcal{G}}[\mathcal{S}]$ seen by one endpoint of e_{ij}, or in a connected component with multiple vertices in $\bar{\mathcal{G}}[\mathcal{S}]$ in which all vertices are seen by one endpoint of e_{ij}. Hence, in each component in $\bar{\mathcal{G}}(\mathcal{S})$, all vertices are seen by at least one endpoint of e_{ij}, and thus e_{ij} is $\mathcal{S}$-saturating. An illustration is shown in Fig. 5.5. Due to $\mathcal{S} \subseteq \mathcal{N}(e_{ij})$, e_{ij} is LB-simplicial. This applies to any edge in $\mathcal{E}$. Hence, $\mathcal{G}$ is a weakly chordal graph by Lemma 5.2. □

As weakly chordal graphs are perfect graphs [108,110], it follows that the chromatic number equals to the clique number for every induced subgraph. Thus, we have $\chi(\mathcal{G}) = \omega(\mathcal{G})$.

Step III

For each clique in the conflict graph $\mathcal{G}$, the involved messages form an acyclic demand graph, whose cardinality yields an outer bound for the symmetric DoF by Lemma 5.3. Thus, among all cliques, the tightest outer bound is given by the clique with the maximum size, i.e., clique number. As such, we have the outer bound

$$d_{\text{sym}} \leq \frac{1}{\omega(\mathcal{G})} \tag{5.22}$$

This is corroborated by the following lemma.

Lemma 5.5. *For each clique in conflict graph $\mathcal{G}$, the involved messages form an acyclic demand graph, such that the sum DoF value of involved messages is no more than 1.*

Proof. Consider without loss of generality a clique in $\mathcal{G}$ with vertices $\mathcal{Q}_n = \{q_1, \ldots, q_{|\mathcal{Q}_n|}\}$, which is an induced graph of $\mathcal{G}$, i.e., $\mathcal{G}[\mathcal{Q}_n]$. In $\mathcal{G}[\mathcal{Q}_n]$, any two messages interfere one another. In the original network topology graph $\mathcal{H}$, the message W_{q_i} originates from S_{q_i} and intends for D_{q_i}. Assume without loss of generality the destinations $D_{\mathcal{Q}_n}$ are arranged from left to right as

$$D_{q_1} \leq D_{q_2} \leq \cdots \leq D_{q_{|\mathcal{Q}_n|}}. \tag{5.23}$$

For instance, because of $W_{q_1} \leftrightarrow W_{q_{|\mathcal{Q}_n|}}$, there are two interfering possibilities between desired links $S_{q_1} \to D_{q_1}$ and $S_{q_{|\mathcal{Q}_n|}} \to D_{q_{|\mathcal{Q}_n|}}$, i.e., $S_{q_1} - D_{q_{|\mathcal{Q}_n|}}$ or $S_{q_{|\mathcal{Q}_n|}} - S_{q_1}$. By source convexity, it follows that $S_{q_1} - D_k$ (or $S_{q_{|\mathcal{Q}_n|}} - S_k$), for all $k \in \mathcal{Q}_n \backslash \{q_1\}$ (or $k \in \mathcal{Q} \backslash \{q_{|\mathcal{Q}_n|}\}$). Correspondingly, according to the definition of demand graph, either of the following two directed edges is missing:

- from D_{q_k} to W_{q_1}, $\forall\ q_k = \{q_2, \ldots, q_{|\mathcal{Q}_n|}\}$, or
- from D_{q_k} to $W_{q_{|\mathcal{Q}_n|}}$, $\forall\ q_k = \{q_1, \ldots, q_{|\mathcal{Q}_n|-1}\}$.

By this, we give the proof by contradiction. Assume there is a cycle in the demand graph involving messages $W_{\mathcal{Q}_n}$, thus at least one of W_{q_1} and $W_{q_{|\mathcal{Q}_n|}}$ should not be in the cycle, because there is no incoming edge at message W_{q_1} (or $W_{q_{|\mathcal{Q}_n|}}$). Thus, we remove one of q_1 and $q_{|\mathcal{Q}_n|}$ from $\mathcal{Q}_n$ and form a new message set, denoted by $W_{\mathcal{Q}_{n-1}}$. The induced subgraph $\mathcal{G}[\mathcal{Q}_{n-1}]$ is still a clique. Following the same argument above, at least one message in $W_{\mathcal{Q}_{n-1}}$ should not be included in a cycle, and therefore we reduce the size of message set to $n-2$, yielding $W_{\mathcal{Q}_{n-2}}$ with cardinality $n-2$. By induction, we finally reduce the message set to $\mathcal{Q}_2$, where only two messages are involved. By assumption, there is a cycle in demand graph with involving messages $W_{\mathcal{Q}_2}$, and thus these two messages should be disconnected in $\mathcal{G}[\mathcal{Q}_2]$, which is contradict with the fact that $\mathcal{G}[\mathcal{Q}_i]$ is a clique.

As such, there should not be a cycle in the demand graph involving messages $W_{\mathcal{Q}}$, if the induced subgraph $\mathcal{G}[\mathcal{Q}]$ is a clique. By [104, Lemma 1], it follows that

$$\sum_{q \in \mathcal{Q}} d_q \leq 1. \tag{5.24}$$

This completes the proof.

□

Step IV

By the definition of perfect graph $\mathcal{G}$, we have $\chi(\mathcal{G}) = \omega(\mathcal{G})$. Because fractional chromatic (resp. clique) number $\chi_f(\mathcal{G})$ [resp. $\omega_f(\mathcal{G})$] is the linear programming relaxation of chromatic number $\chi(\mathcal{G})$ [resp. $\omega(\mathcal{G})$], and $\chi_f(\mathcal{G})$ and $\omega_f(\mathcal{G})$ are solutions of the dual problems [106], we have

$$\omega(\mathcal{G}) \leq \omega_f(\mathcal{G}) = \chi_f(\mathcal{G}) \leq \chi(\mathcal{G}) \tag{5.25}$$

For the perfect graph $\mathcal{G}$, all these four quantities are equal. Hence, the achievability and the outer bound coincide.

All in all, we conclude that the symmetric DoF are achievable by orthogonal schemes, yielding $d_{\text{sym}} = \frac{1}{\chi(\mathcal{G})} = \frac{1}{\omega(\mathcal{G})}$. This completes the proof of the symmetric DoF.

Proof of Sum DoF

Achievability

The message vertices that are not adjacent to one another in the conflict graph $\mathcal{G}$ can be scheduled at the same time. This set of vertices forms an independent set. Thus, the largest achievable sum DoF value is the size of the largest independent set, i.e., independent set number $\alpha(\mathcal{G})$, such that we have the sum DoF inner bound

$$d_{\text{sum}} \geq \alpha(\mathcal{G}). \tag{5.26}$$

Outer Bound

As in Lemma 5.5, each clique in conflict graph leads to an acyclic demand graph, such that for the messages involved in the clique, the sum DoF are bounded by 1, i.e.,

$$\sum_{i \in \mathscr{C}(\mathcal{G})} d_i \leq 1, \tag{5.27}$$

where $\mathscr{C}(\mathcal{G})$ is a clique in conflict graph $\mathcal{G}$.

Given a set of cliques $\mathfrak{C}$ in conflict graph $\mathcal{G}$, in which each vertex is included at least once, it follows that the union of these cliques gives the collection of all messages. As such, the sum DoF can be bounded by

$$d_{\text{sum}} = \sum_{\mathscr{C}_j \in \mathfrak{C}} \sum_{i \in \mathscr{C}_j(\mathcal{G})} d_i \leq |\mathfrak{C}| \tag{5.28}$$

where $\mathscr{C}_j(\mathcal{G})$ is one clique in $\mathfrak{C}$.

According to the definition of clique covering, the tightest outer bound will be given by the clique covering number:

$$d_{\text{sum}} = \sum_{\mathscr{C}_j \in \mathfrak{C}} \sum_{i \in \mathscr{C}_j(\mathcal{G})} d_i \leq \theta(\mathcal{G}). \tag{5.29}$$

Optimality

Due to the duality of the maximum clique (resp. maximum independent set) problem and the minimum proper coloring (resp. minimum clique covering) problem, we have

$$\alpha(\mathcal{G}) = \omega(\bar{\mathcal{G}}), \quad \theta(\mathcal{G}) = \chi(\bar{\mathcal{G}}). \tag{5.30}$$

According the weakly perfect graph theorem [108], a graph is perfect if and only if its complement is perfect. Thus, the complement of conflict graph $\bar{\mathcal{G}}$ is perfect and therefore $\omega(\bar{\mathcal{G}}) = \chi(\bar{\mathcal{G}})$. It follows immediately that $\alpha(\mathcal{G}) = \theta(\mathcal{G})$, indicating that the achievability coincides with the outer bound, i.e., the optimal sum DoF are $d_{\text{sum}} = \alpha(\mathcal{G}) = \theta(\mathcal{G})$.

Proof of DoF Region

Consider the conflict graph $\mathcal{G}$, which is a weakly chordal graph and thus a perfect graph.

Achievability

Given an independent (or stable) set S with $S \in \mathcal{S}(\mathcal{G})$, where $\mathcal{S}(\mathcal{G})$ is the collection of all independent sets in $\mathcal{G}$, let us define the incidence vector $\boldsymbol{x}^S \in \mathbb{R}^{|\mathcal{V}|}$ with i-th element given by

$$x_i^S = \begin{cases} 1, & \text{if } i \in S \\ 0, & \text{otherwise.} \end{cases} \tag{5.31}$$

Accordingly, we define the independent set polytope [112][2] as the convex hull of the incidence vectors of its independent sets, i.e.,

$$\mathcal{K}(\mathcal{G}) = \text{conv}\{\boldsymbol{x}^S | \forall\ S \text{ independent set of } \mathcal{G}.\} \tag{5.32}$$

[2]A polytope is the convex hull of its vertices.

which has an equivalent formulation

$$\mathcal{K}(\mathcal{G}) = \left\{ \boldsymbol{x} \in \{0,1\}^{|\mathcal{V}|} \middle| \sum_{i \in Q} x_i \leq 1, Q \text{ clique in } \mathcal{G} \right\}. \tag{5.33}$$

The incidence vector $\boldsymbol{x}^S$ can be associated with a one-shot proper coloring strategy, in which the vertices in independent set S are assigned with same colors while the rest ones are not assigned with any colors. This color assignment yields an achievable DoF tuple with $d_i = x_i^S$ for all $i \in \mathcal{V}$.

Since every incidence vector (i.e., the vertices of independent set polytope) can be associated with a proper vertex coloring, the whole independent set polytope can be achievable by time sharing, and thus offers us an achievable DoF region with orthogonal access.

Outer Bound

From Lemma 5.5, for every clique in $\mathcal{G}$, the involved messages form an acyclic demand graph, and thus offer an outer bound. Thus, the overall outer bound region can be given by

$$\mathcal{D}_o = \left\{ (d_1, \ldots, d_{|\mathcal{V}|}) \in \mathbb{R}_+^{|\mathcal{V}|} \middle| \sum_{i \in \mathscr{C}_j} d_i \leq 1, \forall\, \mathscr{C}_j \in \mathfrak{C}(\mathcal{G}) \right\}. \tag{5.34}$$

where $\mathfrak{C}(\mathcal{G})$ is the collection of all cliques in $\mathcal{G}$.

The outer bound region is exactly a relaxation of the independent set polytope by replacing $0-1$ with real numbers, i.e.,

$$\mathcal{K}_f(\mathcal{G}) = \left\{ \boldsymbol{x} \in \mathbb{R}_+^{|\mathcal{V}|} \middle| \sum_{i \in Q} x_i \leq 1, Q \text{ clique in } \mathcal{G} \right\} \tag{5.35}$$

which is also called fractional independent set polytope. It has been shown by Chvátal [113] that if $\mathcal{G}$ is perfect, then $\mathcal{K}(\mathcal{G}) = \mathcal{K}_f(\mathcal{G})$, which means the vertices of polytope $\mathcal{K}_f(\mathcal{G})$ are integral, i.e., all the corner points have integral coordinates.

Clearly, it follows immediately that $\mathcal{D}_o = \mathcal{K}(\mathcal{G})$, which means achievability coincides with outer bound. As such, the orthogonal schemes achieve the optimal DoF region.

All in all, orthogonal access achieves the optimal symmetric DoF, optimal sum DoF and optimal DoF region, and therefore it is DoF optimal for all one-dimensional convex interference channels.

5.4.3 Proof of Lemma 5.1

We show by construction that $\mathcal{G}_e^2[\mathcal{M}]$ is the message conflict graph of $\mathcal{G}$ with message set $\mathcal{M}$.

First, we consider an arbitrary cellular network topology $\mathcal{G}$, where each source has one message to each connected destination, i.e., $\mathcal{W}_{S_i} = \{W_{\mathcal{R}_i}\}$ and $\mathcal{W}(D_j) = \{W_{\mathcal{T}_j}\}$. Each edge (e.g., $S_i \to D_j$) in $\mathcal{G}$ corresponds to a message. As such, its line graph $\mathcal{G}_e$ represents the adjacency of messages, where the messages (i.e., edges in $\mathcal{G}$) with common sources or destinations are adjacent in $\mathcal{G}_e$. These adjacent messages cannot be orthogonally delivered at the same time and hence are conflicting. In addition, for two messages W_{ij} ($S_i \to D_j$ in $\mathcal{G}$) and W_{kl} ($S_k \to D_l$ in $\mathcal{G}$), either $S_i \to D_l$ or $S_k \to D_j$ will lead to conflict between W_{ij} and W_{kl}, because these two messages cannot be delivered together without mutual interference. Hence, any two edges connected with one common edge in $\mathcal{G}$, i.e., any two messages with distance of 2 in $\mathcal{G}_e$ are conflicting as well. As such, connecting any two messages with distance of 2 in $\mathcal{G}_e$ by an edge gives us the square of $\mathcal{G}_e$, i.e., $\mathcal{G}_e^2$. In $\mathcal{G}_e^2$, any two adjacent messages conflict one another, and in turn $\mathcal{G}_e^2$ is the message conflict graph with overall message set.

Second, the conflict between messages is only determined by the network topology and irrelative to the message set. If some edges in $\mathcal{G}$ are interfering links, then they will have not impact on the conflict of other messages. Removing the messages (i.e., vertices) associated with these interfering links from $\mathcal{G}_e^2$ will not break the conflict condition. As such, given the message set $\mathcal{M}$, the vertices associated with messages out of $W_{\mathcal{M}}$ together with the involved edges will be removed from $\mathcal{G}_e^2$, which yields an induced subgraph $\mathcal{G}_e^2[\mathcal{M}]$. The conflict between messages $W_{\mathcal{M}}$ is inherited. Thus, $\mathcal{G}_e^2[\mathcal{M}]$ is the conflict graph with respect to messages $W_{\mathcal{M}}$.

5.4.4 Proof of Theorem 5.2

Proof of Symmetric DoF

In general, Theorem 5.2 is a wider and stronger generalization of Theorem 5.1. In what follows, we first sketch the main steps in the proof, which is similar to that of Theorem 5.1, and more details will be given afterwards.

1. As $\mathcal{G}_e^2[\mathcal{M}]$ is the induced message conflict subgraph with message set $\mathcal{M}$, the symmetric DoF value achieved by orthogonal schemes is the inverse of the fractional chromatic number of this conflict graph, i.e., $d_{\text{sym}} = \frac{1}{\chi_f(\mathcal{G}_e^2[\mathcal{M}])}$.

2. Since $\mathcal{G}$ is a chordal bipartite graph, which is weakly chordal [108], then the conflict graph $\mathcal{G}_e^2$ is also weakly chordal [114] and therefore perfect. According to the definition of perfect graphs, for any induced subgraph with vertex set $\mathcal{M}$, we have $\chi(\mathcal{G}_e^2[\mathcal{M}]) = \omega(\mathcal{G}_e^2[\mathcal{M}])$.

3. For every clique in induced conflict subgraph $\mathcal{G}_e^2[\mathcal{M}]$, the involved messages form an acyclic demand graph. The clique with largest size dominates the symmetric DoF. As such, the symmetric DoF are outer bounded by the inverse of clique number of $\mathcal{G}_e^2[\mathcal{M}]$, i.e., $d_{\text{sym}} \leq \frac{1}{\omega(\mathcal{G}_e^2[\mathcal{M}])}$.

4. Due to the relation $\omega \leq \omega_f = \chi_f \leq \chi$ and $\chi = \omega$ for perfect graphs $\mathcal{G}_e^2[\mathcal{M}]$, it follows that $\omega = \omega_f = \chi_f = \chi$ and thus the achievability coincides with the outer bound.

The detailed proofs are presented as follows.

Step I

As proved in Lemma 5.1, $\mathcal{G}_e^2[\mathcal{M}]$ is the message conflict graph of $\mathcal{G}$ with regard to message set $\mathcal{M}$. Following the same arguments in the proof of Theorem 5.1, the adjacent vertices (messages) in $\mathcal{G}_e^2[\mathcal{M}]$ should be assigned with different colors. A proper color assignment strategy corresponds to an orthogonal scheduling scheme. As such, the maximum symmetric DoF achieved by orthogonal schemes turn out to be the inverse of fractional chromatic number, i.e., $\frac{1}{\chi_f(\mathcal{G}_e^2[\mathcal{M}])}$. Thus, $d_{\text{sym}} = \frac{1}{\chi_f(\mathcal{G}_e^2[\mathcal{M}])}$ is best achievable symmetric DoF by the orthogonal scheme.

Step II

First, the topology graph $\mathcal{G}$ is a chordal bipartite graph, and thus it is weakly chordal [108] [3].

Then, we show that the message conflict graph $\mathcal{G}_e^2$ of chordal cellular networks is weakly chordal as well with aid of the following lemma.

Lemma 5.6 ([114])**.** *Given a graph $\mathcal{G}$, if $\mathcal{G}$ is weakly chordal, then $\mathcal{G}_e^2$ is also weakly chordal.*

Proof. We sketch the proof in short and the detailed proof can be found in [114]. The proof is based on the following two facts:

[3]Note that the chordal bipartite graph is bipartite and weakly chordal, but not chordal. Although this terminology has caused confusion, it becomes well-accepted by the graph theory community.

- If $\mathcal{G}$ has no induced cycles on at least k vertices, where $k \geq 4$, then $\mathcal{G}_e^2$ has no induced cycles on at least k vertices.
- If $\mathcal{G}$ is weakly chordal, then $\mathcal{G}_e^2$ does not contain the antihole as an induced subgraph.

If $\mathcal{G}$ is weakly chordal, then $\mathcal{G}$ contains neither a hole nor an antihole, and thus $\mathcal{G}_e^2$ is also weakly chordal according to the above two facts. □

Lastly, weakly chordal graphs are perfect, such that for any induced subgraph $\mathcal{G}_e^2[\mathcal{M}]$ by $\mathcal{M}$, we have $\chi(\mathcal{G}_e^2[\mathcal{M}]) = \omega(\mathcal{G}_e^2[\mathcal{M}])$.

Step III

For a given induced subgraph $\mathcal{G}_e^2[\mathcal{M}]$, every clique with involved messages forms an acyclic demand graph, whose cardinality yields an outer bound for the symmetric DoF according to Lemma 5.3. Thus, among all cliques, the tightest outer bound is given by the clique with the maximum size, i.e., clique number. As such, we have the outer bound

$$d_{\text{sym}} \leq \frac{1}{\omega(\mathcal{G}_e^2[\mathcal{M}])} \tag{5.36}$$

This is corroborated by the following lemma.

Lemma 5.7. *Given a chordal bipartite graph $\mathcal{G}$ with message set $\mathcal{M}$ and its induced conflict subgraph $\mathcal{G}_e^2[\mathcal{M}]$, the demand graph of the associated messages of every clique is acyclic, such that the sum DoF value of these involved messages is no more than 1.*

Proof. Let us consider the whole conflict graph $\mathcal{G}_e^2$ first. Given a set of messages $W_{\mathcal{L}} = \{W_{i_1 i_2}, W_{i_3 i_4}, \ldots, W_{i_{n-1} i_n}\}$, the associated vertices form a clique in $\mathcal{G}_e^2$. Suppose there exist a shortest chordless cycle in demand graph, which is corresponding to the following connectivity in network topology

$$S_{i_1} \to D_{i_2} \nrightarrow S_{i_3} \to D_{i_4} \nrightarrow S_{i_5} \to \cdots \to D_{i_n} \nrightarrow S_{i_1}. \tag{5.37}$$

Because of the chordless cycle, all the sources are distinct, and so are the destinations. Otherwise, there should exist a shorter cycle. It is readily verified that n has to be even and $n \geq 6$. It is because the demand graphs are bipartite graph, and for any two messages $W_{i_k i_{k+1}}$ and $W_{i_j i_{j+1}}$ in $W_{\mathcal{L}}$, $D_{i_{j+1}} \nrightarrow S_{i_k}$ and $D_{i_{k+1}} \nrightarrow S_{i_j}$ should not present simultaneously, i.e., $n \neq 4$, otherwise two messages $W_{i_k i_{k+1}}$ and $W_{i_j i_{j+1}}$ are not adjacent in $\mathcal{G}_e^2$, which results in a contradiction.

When $n = 6$, according to the cycle (5.37), it follows that

$$S_{i_1} \rightarrowtail D_{i_4}, \quad S_{i_3} \rightarrowtail D_{i_6}, \quad S_{i_5} \rightarrowtail D_{i_2}. \tag{5.38}$$

Together with the desired links

$$S_{i_1} \to D_{i_2}, \quad S_{i_3} \to D_{i_4}, \quad S_{i_5} \to D_{i_6} \tag{5.39}$$

we conclude that there is a chordless cycle in $\mathcal{G}$, namely,

$$S_{i_1} \to D_{i_2} \leftarrowtail S_{i_5} \to D_{i_6} \leftarrowtail S_{i_3} \to D_{i_4} \leftarrowtail S_{i_1} \tag{5.40}$$

with length six.

When $n \geq 8$, according to the cycle (5.37), it follows that

$$S_{i_1} \rightarrowtail D_{i_4}, D_{i_6}, D_{i_8}, \ldots, \quad S_{i_3} \rightarrowtail D_{i_6}, D_{i_8}, \ldots \tag{5.41}$$

$$S_{i_5} \rightarrowtail D_{i_8}, \ldots, D_{i_2}, \quad S_{i_7} \rightarrowtail \ldots, D_{i_2}, D_{i_4}. \tag{5.42}$$

Considering $D_{i_2}, S_{i_3}, D_{i_4}, S_{i_5}, D_{i_6}, S_{i_7}$, we have a chordless cycle with length six, i.e.,

$$S_{i_3} \to D_{i_4} \leftarrowtail S_{i_7} \rightarrowtail D_{i_2} \leftarrowtail S_{i_5} \to D_{i_6} \leftarrowtail S_{i_3}. \tag{5.43}$$

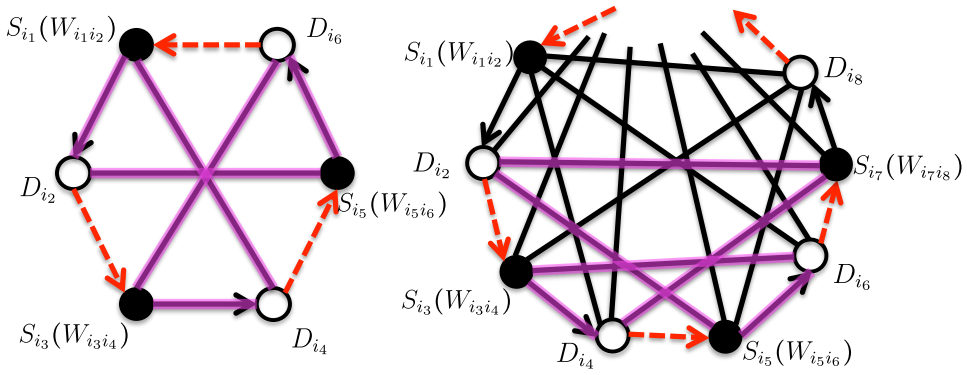


Figure 5.6: Illustration of the proof for Lemma 5.7. The black and red dashed arrows form a cyclic demand graph, the black arrows and lines are edges in topology graph, and the purple lines indicate a cycle with length six in topology graph.

Thus, if there is a cycle in the demand graph corresponding to a clique in $\mathcal{G}_e^2$ as shown in Fig. 5.6, then there is at least a cycle with length six in

topology graph $\mathcal{G}$. By contraposition, for a chordal bipartite graph $\mathcal{G}$ where no cycles with length greater than four exist, every clique with messages $W_{\mathcal{L}}$ in the conflict graph $\mathcal{G}_e^2$ leads to an acyclic demand graph with regard to $W_{\mathcal{L}}$.

In any subgraph induced by $\mathcal{M}$, e.g., $\mathcal{G}_e^2[\mathcal{M}]$, the clique in $\mathcal{G}_e^2[\mathcal{M}]$ is still a clique in $\mathcal{G}_e^2$. Thus, the above argument still holds, i.e., each clique in $\mathcal{G}_e^2[\mathcal{M}]$ corresponds to an acyclic demand graph. By [104, Lemma 1], it follows that the sum DoF value of the corresponding messages involved in the clique is no more than 1. This completes the proof. □

In fact, based on the proof of Lemma 5.7, we can make a stronger claim.

Lemma 5.8. *Any clique $\mathcal{Q}$ in $\mathcal{G}_e^2$ serves as an outer bound, i.e.,*

$$\sum_{q\in\mathcal{Q}} d_q \leq 1, \tag{5.44}$$

if the network topological graph $\mathcal{G}$ is C_6-free, i.e., there do not exist induced subgraphs with length-six cycles in $\mathcal{G}$.

Proof. The proof of this lemma can be directly extended from the above proof. Since if there exists a cycle in demand graphs, then these should exist a length-six cycle in $\mathcal{G}$. By contraposition, if $\mathcal{G}$ is C_6-free, then cliques in $\mathcal{G}_e^2$ correspond to acyclic demand graphs, and thus serve as outer bounds. □

By Lemma 5.3 and Lemma 5.7, the tightest outer bound on symmetric DoF is determined by the clique with the largest size. Thus, the clique number offers an outer bound $d_{\text{sym}} \leq \frac{1}{\omega(\mathcal{G}_e^2[\mathcal{M}])}$.

Step IV

Since $\mathcal{G}_e^2$ is weakly chordal and therefore perfect, every induced subgraph $\mathcal{G}_e^2[\mathcal{M}]$ is also perfect [108]. By the definition of perfect graph, we have $\chi(\mathcal{G}_e^2[\mathcal{M}]) = \omega(\mathcal{G}_e^2[\mathcal{M}])$. Because fractional chromatic (resp. clique) number $\chi_f(\mathcal{G}_e^2[\mathcal{M}])$ [resp. $\omega_f(\mathcal{G}_e^2[\mathcal{M}])$] is the linear programming relaxation of chromatic number $\chi(\mathcal{G}_e^2[\mathcal{M}])$ [resp. $\omega(\mathcal{G}_e^2[\mathcal{M}])$], and $\chi_f(\mathcal{G}_e^2[\mathcal{M}])$ and $\omega_f(\mathcal{G}_e^2[\mathcal{M}])$ are solutions of the dual problems [106], we have

$$\omega(\mathcal{G}_e^2[\mathcal{M}]) \leq \omega_f(\mathcal{G}_e^2[\mathcal{M}]) = \chi_f(\mathcal{G}_e^2[\mathcal{M}]) \leq \chi(\mathcal{G}_e^2[\mathcal{M}]) \tag{5.45}$$

For the perfect graph $\mathcal{G}_e^2[\mathcal{M}]$, all these four quantities are equal. As such, orthogonal schemes achieve the optimal symmetric DoF $d_{\text{sym}} = \frac{1}{\chi(\mathcal{G}_e^2[\mathcal{M}])} = \frac{1}{\omega(\mathcal{G}_e^2[\mathcal{M}])}$.

Proof of Sum DoF

The proof of sum DoF is similar to that of Theorem 5.1. Given an induced conflict subgraph $\mathcal{G}_e^2[\mathcal{M}]$, the messages in any independent set can be scheduled at a single time slot without causing any mutual interference. As such, the independent set number yields the maximum achievable sum DoF

$$d_{\text{sum}} = \alpha(\mathcal{G}_e^2[\mathcal{M}]). \tag{5.46}$$

On the other hand, by Lemma 5.7, each clique in $\mathcal{G}_e^2[\mathcal{M}]$ leads to an acyclic demand graph outer bound, such that for the messages involved in the clique, the sum DoF are bounded by 1, i.e.,:

$$\sum_{i \in \mathscr{C}_j} d_i \leq 1, \qquad \forall\, \mathscr{C}_j \in \mathfrak{C}(\mathcal{G}_e^2[\mathcal{M}]) \tag{5.47}$$

where $\mathfrak{C}(\mathcal{G}_e^2[\mathcal{M}])$ is the collection of all possible cliques in $\mathcal{G}_e^2[\mathcal{M}]$. Suppose there exists a set of cliques $\mathfrak{C}_c(\mathcal{G}_e^2[\mathcal{M}])$ such that each vertex is included at least once. Thus, we have

$$d_{\text{sum}} = \sum_{\mathscr{C}_j \in \mathfrak{C}_c} \sum_{i \in \mathscr{C}_j} d_i \leq |\mathfrak{C}_c|. \tag{5.48}$$

According to the definition of clique covering, the tightest outer bound is given by the clique covering number, i.e.,

$$d_{\text{sum}} = \sum_{\mathscr{C}_j \in \mathfrak{C}_c} \sum_{i \in \mathscr{C}_j} d_i \leq \theta(\mathcal{G}_e^2[\mathcal{M}]) \tag{5.49}$$

Due to the duality between the maximum independent set (resp. maximum clique) problem and minimal clique covering (resp. minimum proper coloring) problem, in the induced subgraph $\mathcal{G}_e^2[\mathcal{M}]$, we have

$$\alpha(\mathcal{G}_e^2[\mathcal{M}]) = \omega(\overline{\mathcal{G}_e^2[\mathcal{M}]}), \quad \theta(\mathcal{G}_e^2[\mathcal{M}]) = \chi(\overline{\mathcal{G}_e^2[\mathcal{M}]}) \tag{5.50}$$

where $\overline{\mathcal{G}_e^2[\mathcal{M}]}$ is the complement graph of $\mathcal{G}_e^2[\mathcal{M}]$. Because $\mathcal{G}_e^2$ is perfect, then its complement $\overline{\mathcal{G}_e^2}$ is also perfect [108], such that

$$\chi(\overline{\mathcal{G}_e^2}[\mathcal{M}]) = \omega(\overline{\mathcal{G}_e^2}[\mathcal{M}]) \tag{5.51}$$

where $\overline{\mathcal{G}_e^2}[\mathcal{M}]$ is the induced subgraph of $\overline{\mathcal{G}_e^2}$. It is obvious that $\overline{\mathcal{G}_e^2[\mathcal{M}]}$ and $\overline{\mathcal{G}_e^2}[\mathcal{M}]$ yield the same subgraph. Thus, $\alpha(\mathcal{G}_e^2[\mathcal{M}]) = \theta(\mathcal{G}_e^2[\mathcal{M}])$, showing that the achievability and outer bound coincide. This completes the proof of sum DoF.

Proof of DoF Region

Consider the conflict graph $\mathcal{G}_e^2[\mathcal{M}]$ with a message set $\mathcal{M}$. Because $\mathcal{G}_e^2$ is a weakly chordal graph and thus perfect, its induced subgraph $\mathcal{G}_e^2[\mathcal{M}]$ is perfect as well [108]. The achievability and outer bound proofs are similar to those of Theorem 5.1.

Achievability

Regarding the graphs $\mathcal{G}_e^2[\mathcal{M}]$, given an independent (or stable) set S with $S \in \mathcal{S}(\mathcal{G}_e^2[\mathcal{M}])$, where $\mathcal{S}(\mathcal{G}_e^2[\mathcal{M}])$ is the collection of all independent sets in $\mathcal{G}_e^2[\mathcal{M}]$, let us define the incidence vector $\boldsymbol{x}^S \in \mathbb{R}^{|\mathcal{M}|}$ with i-th element given by

$$x_i^S = \begin{cases} 1, & \text{if } i \in S \\ 0, & \text{otherwise.} \end{cases} \tag{5.52}$$

Similarly, we have the independent set polytope [112]

$$\mathcal{K}(\mathcal{G}_e^2[\mathcal{M}]) = \left\{ \boldsymbol{x} \in \{0,1\}^{|\mathcal{M}|} \middle| \sum_{i \in Q} x_i \leq 1, Q \text{ clique in } \mathcal{G}_e^2[\mathcal{M}] \right\} \tag{5.53}$$

which gives the achievable region.

Outer Bound

From Lemma 5.7, for every clique in $\mathcal{G}$, the involved messages form an acyclic demand graph, and thus offer an outer bound. The outer bound region is exactly a relaxation of the independent set polytope by replacing $0-1$ with real numbers, i.e., fractional independent set polytope,

$$\mathcal{K}_f(\mathcal{G}_e^2[\mathcal{M}]) = \left\{ \boldsymbol{x} \in \mathbb{R}_+^{|\mathcal{M}|} \middle| \sum_{i \in Q} x_i \leq 1, Q \text{ clique in } \mathcal{G}_e^2[\mathcal{M}] \right\}. \tag{5.54}$$

It has been shown by Chvátal [113] that if $\mathcal{G}_e^2[\mathcal{M}]$ is perfect, $\mathcal{K}(\mathcal{G}_e^2[\mathcal{M}]) = \mathcal{K}_f(\mathcal{G}_e^2[\mathcal{M}])$, which means the vertices of polytope $\mathcal{K}_f(\mathcal{G}_e^2[\mathcal{M}])$ are integral, i.e., all the corner points have integral coordinates.

Clearly, it follows immediately that achievability coincides with outer bound. As such, the orthogonal schemes achieve the optimal DoF region.

All in all, orthogonal access has been shown to achieve the optimal symmetric DoF, optimal sum DoF and optimal DoF region, and therefore it is DoF optimal for all the chordal cellular networks.

5.4.5 Proof of Corollary 5.1

What is left to prove is to verify that convex bipartite graphs are chordal bipartite graphs. It is only need to show that there do not exist chordless cycles with length more than four.

We consider the one side convexity at sources, and the proof with regard to the convexity at destinations follows similarly. Assume by contradiction that there exists a chordless cycle with length n

$$S_1 \rightarrowtail D_2 \leftarrowtail S_3 \rightarrowtail \cdots \rightarrowtail D_n \leftarrowtail S_1 \tag{5.55}$$

where $n \geq 6$ and n is even. We hereafter focus on the induced subgraph by these sources and destinations, denoted by $\mathcal{G}[1:n]$, and consider the relative placement of $\{D_2, D_4, \ldots, D_n\}$.

According to the one-dimensional convexity property at sources, the destinations connected to a source are consecutive in $\mathcal{G}[1:n]$. It follows that every source S_k ($k \in \{1, 3, \ldots, n-1\}$) is adjacent to two consecutive destinations D_{k-1} and D_{k+1} in $\mathcal{G}[1:n]$, where the indices are taken modulo n. As such, any two destinations in $\{D_2, D_4, \ldots, D_n\}$ are consecutive. It is impossible for the one-dimensional convex networks, because any placement of destinations has an order such that the head and tail destinations are not consecutive, for more than two destinations, i.e., $n \geq 6$. Consequently, there should not exist a chordless cycle with length greater than four, and therefore by definition convex bipartite graphs are chordal bipartite graphs.

5.5 Summary and Discussions

5.5.1 Summary

Orthogonal access has been shown to achieve the optimal symmetric DoF, optimal sum DoF, and optimal DoF region of multiple unicast TIM problems with arbitrary message sets in chordal cellular networks in which these do not exist cycles with length greater than four. According to the equivalence between TIM and index coding problems with regard to linear solutions [49], the orthogonal schemes achieve symmetric capacity, sum capacity, and capacity region of the corresponding index coding problems. Remarkably, as demonstrated by the powerful graph theoretic tools, the physical convexity in cellular network, generally speaking the absence of long cycles (i.e., with length no less than six) in network topology graphs, turns out to be the fundamental limitation.

5.5.2 Discussions

In fact, orthogonal access is equivalent to finding the induced matchings [115] in the network topology graph, in which no edges in an induced matching are joint with an edge. Finding the induced matchings of $\mathcal{G}$ can be translated to finding the conventional matching of its square of line graph, i.e., $\mathcal{G}_e^2$. As claimed earlier, the symmetric DoF can be done by finding the maximal clique, the sum DoF by finding the maximum independent set, and the DoF region characterization by finding all cliques. As such, the computational complexity of characterizing DoF in chordal cellular networks is polynomial-time. On the other hand, given a set of messages $\mathcal{M}$, calculating the weighted sum DoF of these messages is equivalent to finding the maximum weighted independent set of $\mathcal{G}_e^2[\mathcal{M}]$, where the weights of vertices in $\mathcal{G}_e^2[\mathcal{M}]$ can be the weights of individual DoF.

The optimality of orthogonal access works for arbitrary message set, because of the special property of chordal cellular networks, whose conflict graphs are weakly chordal. The message conflict graph associated with any message set can be regarded as the induced subgraph of the square of line graph of network topology. By the definition of (weakly) chordal graphs, any induced subgraph of (weakly) chordal graphs are still (weakly) chordal. As such, the optimality holds for any message set.

As demonstrated in [116–118], finding a minimum vertex coloring of a weakly chordal graph runs in $O(mn)$, where n is the number of vertices and m is the number of edges. As in [104], the greedy algorithm is optimal for biconvex network, because the conflict graph of biconvex networks is chordal, which is perfectly orderable, implying the optimality of greedy algorithms according to the definition of perfectly orderable graphs [119]. The conflict graph of biconvex bipartite complies with perfect elimination order. Weakly chordral graphs are in general not perfectly orderable. But P5-free weakly chordal graphs are perfect orderable [120]. As such, if the conflict graphs of chordal cellular networks are P5-free, then greedy coloring (i.e., edge scheduling) schemes will lead to optimal solutions.

In fact, some established general solutions to TIM and index coding problems [49, 98, 103] can be applied here, so that new insights emerge taking into account some properties of chordal cellular networks. For instance, if the transmitters are equipped with multiple antenna to form MISO cellular networks, according to the conclusion in [103], then the additional transmit antennas are useless in terms of DoF and the optimality of orthogonal scheme still holds in MISO chordal cellular networks. Due to the uplink and downlink duality of multiple unicast TIM problems with linear schemes [49],

the optimality of orthogonal access still holds for the uplink channel of chordal cellular networks, in which the conflict graphs of uplink and downlink channels are exactly the same.

5.5.3 Insight to Practical Scenarios

As argued earlier, in practice, the wireless networks are always fully connected, no matter how weak the links are. One interesting question is how much insight obtained in TIM settings can be translated to practical scenarios. This question is still challenging at this stage. As a first step, we perform some simulations to see the behavior of the DoF-optimal schemes in the finite SNR regime.

As an illustrative example, we consider an interference channel with 10 transmitter-receiver pairs. The channel coefficient is modeled taking both path loss and small-scale fading into account, as

$$h_{ji} = c \cdot \sqrt{d_{ji}^{-3}} \cdot h_w \tag{5.56}$$

where c is a constant, d_{ji} is the distance between the transmitter i to the receiver j, and h_w is Gaussian distributed with zero mean and unit variance. The transmitters and receivers are located in two parallel straight lines. Every two transmitters/receivers are spaced with 500 meters and each transmitter/receiver pair is spaced with 500 meters as well. At the receiver side, interference is treated as noise.

For this interference channel with no CSIT, conventional wisdom relying on fully connected topological graphs suggests either activating all transmitters simultaneously and treating interference as noise at receivers (denoted by "TIN") or scheduling one transmitter per time slot so as to completely avoid interference at each receiver (referred to as "TDMA"). Under the TIM setting, we distinguish the weak channels from the strong ones by setting a threshold, such that the channel is modeled to be present if the path loss value is beyond this threshold and to be absent if otherwise. By adjusting the threshold, the connectivity graphs can be distinct, and so are the corresponding message conflict graphs. Intuitively, if the threshold is higher, the topology can be denser, while the topology graph will be sparser if it is lower. Since the cellular networks of interest are one-dimensional convex, orthogonal access built on resultant conflict graphs is the DoF-optimal strategy for TIM setting.

Fig. 5.7 shows the throughput performance per user at finite SNR with different transmission strategies and different threshold values for the TIM

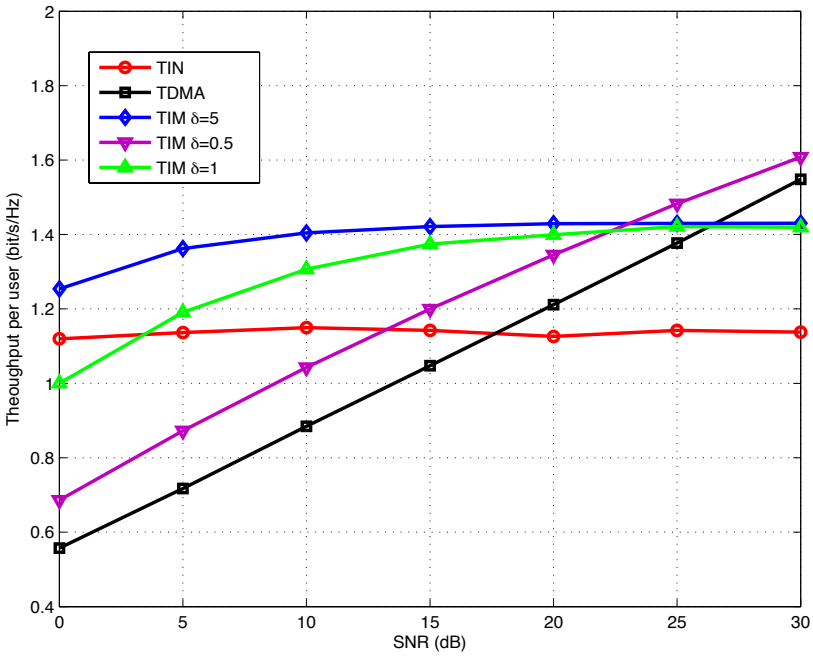


Figure 5.7: Throughput versus SNR at finite SNR with different thresholds.

setting. When the threshold is set to be infinity, the message conflict graph is fully connected, and thus orthogonal access degrades to TDMA. On the other hand, if the threshold is set to be zero, the message conflict graph is empty, such that orthogonal access is incline to activate all transmitters simultaneously, which results in the TIN scheme. Nevertheless, with a fairly reasonable threshold, the network is partially connected. With respect to the partial connectivity, the transmitters are activated according to conflict graphs, such that orthogonal access achieves optimal DoF at infinite SNR and attains some gains at finite SNR. As shown in Fig. 5.7, within a range of SNR (i.e., 5dB - 25dB) in which the real systems usually operate, orthogonal access suggested under TIM settings outperforms both TIN and TDMA schemes. That is what the TIM results suggest to practical system design.

Chapter 6

Topological Interference Management with Transmitter Cooperation

As shown in the previous chapter, the original multiple-unicast TIM problems mainly focus on cellular networks with no CSIT except for the knowledge of the connectivity graph via topological feedback. In this chapter, the emphasis will be placed specifically on interference channels, where one transmitter servers one single receiver. We consider a similar problem with sole topological knowledge but in a distributed broadcast channel setting, i.e. a network where transmitter cooperation is enabled.

We show that the topological information can also be exploited in this case to strictly improve DoF as long as the network is not fully connected, which is a reasonable assumption in practice. Three inner bounds based on graph coloring, interference alignment, and hypergraph covering, are proposed, together with three outer bounds built upon generator sequence, compound settings, and index coding, by all of which we characterize the symmetric DoF for so-called regular networks with constant number of interfering links, and identify the sufficient and/or necessary conditions for the arbitrary network topologies to achieve a certain amount of symmetric DoF.

6.1 Introduction

Recently, the benefit of topological information to network performance was intensively revealed by [49–51, 101, 104, 121–123], in the context of interference and X channels with topology information from different perspectives, focusing on symmetric DoF. These various topological interference management approaches arrived at a common conclusion that the symmetric DoF can be significantly improved with the sole topological feedback beyond time-division, provided that the network is partially connected. In [49], the TIM problem was bridged with the index coding problem [52, 96–98, 124, 125], so that the optimal solution to the latter is the outer bound of the former, and the linear solution to the former is automatically transferrable to the latter. The ensuing extension in [98] that considered the TIM problem as an application of index coding problems and attacked it from an interference alignment perspective, covers a wider class of network topologies, partly settling the problem for the sparse networks with each receiver interfered by at most two interfering links.

Given such promising results in interference and X channels, a logical question is whether the TIM framework can somehow be exploited in the context of an interference channel where a message exchange mechanism between transmitters pre-exists. For instance, in future LTE-A cellular networks, a backhaul routing mechanism ensures that base stations selected to cooperate under the coordinated multiple point (CoMP) framework receive a copy of the messages to be transmitted. Still, the exchange of timely CSI is challenging due to the rapid obsolescence of instantaneous CSI and the latency of backhaul signaling links. In this case, a broadcast channel over distributed transmitters (a.k.a. network MIMO) ensues, with a lack of instantaneous CSIT. The problem raised by this chapter concerns the use of topology information in this setting. We follow the same strategy as [49, 50] in targeting the symmetric DoF as a simple figure of merit. By resorting to interference avoidance and alignment techniques, we characterize the achievable and/or optimal symmetric DoF of this distributed BC with topology information in several scenarios of interest. More specifically, our contributions are as follows:

- We propose an interference avoidance approach built upon fractional selective graph coloring over the square of line graph of the original network topology. In doing so, the optimal symmetric DoF of three-cell networks with all possible topologies are determined, by a new outer bound on the basis of the concept of generator sequence.

- We propose an interference alignment based approach to identify the achievable symmetric DoF of so-called *regular* networks, which correspond to networks with constant number of interfering links. The optimal symmetric DoF of the Wyner-type regular networks (i.e., with only one interfering link) are characterized with the aid of an outer bound derived based on compound settings.

- We show the sufficient conditions for arbitrary network topologies to achieve a certain amount of symmetric DoF. These sufficient conditions are in the form of the relations of transmit sets (i.e., the indices of transmitters that connected to one receiver) associated with some receivers, implying the alignment feasibility of the messages desired by these receivers. Such relations of transmit sets lead us to the construction of a hypergraph, by which a symmetric DoF inner bound is consequently established via hypergraph covering.

- We also bridge our problem to index coding problems, letting outer bounds of the latter serve our problem, by which we identify the sufficient and necessary condition when time division is symmetric DoF optimal.

The rest of this chapter is organized as follows. The system model and some basic definitions are given in the coming section. In Section 6.3, we summarize the main results with three outer bounds and several inner bounds of achievability schemes built upon interference avoidance and alignment techniques, followed by in Section 6.4 two illustrative examples to show the tightness of outer and inner bounds. The general proofs are presented in Section 6.5. We conclude the chapter in Section 6.6.

Notation: Throughout this chapter, we define $\mathcal{K} \triangleq \{1, 2, \ldots, K\}$, and $[n] \triangleq \{1, 2, \ldots, n\}$ for any positive integer n. Let A, $\mathcal{A}$, and $\boldsymbol{A}$ represent a variable, a set, and a matrix/vector, respectively. In addition, $\mathcal{A}^c$ is the complementary set of $\mathcal{A}$, and $|\mathcal{A}|$ is the cardinality of the set $\mathcal{A}$. $\boldsymbol{A}_{ij}$ or $[\boldsymbol{A}]_{ij}$ presents the ij-th entry of the matrix $\boldsymbol{A}$, and $\boldsymbol{A}_i$ or $[\boldsymbol{A}]_i$ is the i-th row of $\boldsymbol{A}$. $A_{\mathcal{S}} \triangleq \{A_i, i \in \mathcal{S}\}$, $\mathcal{A}_{\mathcal{S}} \triangleq \cup_{i\in\mathcal{S}}\mathcal{A}_i$, and $\boldsymbol{A}_{\mathcal{S}}$ denotes the submatrix of $\boldsymbol{A}$ with the rows out of $\mathcal{S}$ removed. Define $\mathcal{A}\backslash a \triangleq \{x|x \in \mathcal{A}, x \neq a\}$ and $\mathcal{A}_1\backslash\mathcal{A}_2 \triangleq \{x|x \in \mathcal{A}_1, x \notin \mathcal{A}_2\}$. We use $\mathbf{I}_M$ to denote an $M \times M$ identity matrix where the dimension is omitted whenever the confusion is not probable. $\mathbf{1}(\cdot)$ is the indicator function with values 1 when the parameter is true and 0 otherwise. $O(\cdot)$ follows the standard Landau notation. Logarithms are in base 2.

6.2 System Model

6.2.1 Channel Model

In this chapter, we focus on a network modeled by a K-user partially connected interference channel, in which each transmitter (e.g. base station) is equipped with one antenna and serves one single-antenna receiver (e.g., user). In this partially connected network, similarly to that in the previous chapter, the received signal for Receiver j at time instant t can be modeled by

$$Y_j(t) = \sum_{i \in \mathcal{T}_j} h_{ji}(t) X_i(t) + Z_j(t) \tag{6.1}$$

where the parameters are recalled here.

- $h_{ji}(t)$ is the channel coefficient between Transmitter i and Receiver j at time instant t and the nonzero channel coefficients drawn from a continuous distribution are i.i.d.;
- the transmitted signal $X_i(t)$ is subject to the average power constraint, i.e., $\mathbb{E}\left(|X_i(t)|^2\right) \leq P$, with P being the average transmit power;
- $Z_j(t)$ is the additive Gaussian noise with zero-mean and unit-variance and is independent of transmitted signals and channel coefficients;
- $\mathcal{T}_k$ is the transmit set containing the indices of transmitters that are *connected* to Receiver k;
- $\mathcal{R}_k$ is the receive set consisting of the indices of receivers that are *connected* to Transmitter k, for $k \in \{1, 2, \ldots, K\}$.

Being consistent with TIM framework, the actual channel realizations are not available at the transmitters, yet the network topology (i.e., $\mathcal{T}_k, \mathcal{R}_k, \forall k$) is known by all transmitters and receivers. A typical transmitter cooperation is enabled, where every transmitter is endowed the messages desired by its connected receivers, i.e., the Transmitter k has access to a subset of messages $W_{\mathcal{R}_k}$, where W_j ($j \in \mathcal{R}_k$) denotes the message desired by Receiver j. We refer to TIM problem with transmitter cooperation in this chapter as "TIM-CoMP" problem. Each message may originate from multiple transmitters but is intended for one unique receiver. Note that direct links are not required to be present thanks to transmitter cooperation. We consider a block fading channel, where the channel coefficients stay constant during a coherence time τ_c but vary to independent realizations in the next coherence time.

The coherence time is set to $\tau_c = 1$ unless otherwise specified. The network topology is fixed throughout the communication.

While message sharing creates the opportunity of transmitter cooperation, it also imposes some challenges. For the partially connected IC or X networks [49–51], each message has a unique source and a unique destination that are determined *a priori* such that the desired and interfering links are known. By contrast, with transmitter cooperation, the message can be sent from *any* source that has access to this message. Consequently, the approaches developed for IC and X networks cannot be directly applied here, as the desired and interference links are not able to be predetermined.

For notational convenience, we define $\mathcal{H} \triangleq \{h_{ji}, \forall\, i, j\}$ as the ensemble of channel coefficients, and denote by $\mathcal{G}$ the network topology known by all transmitters and receivers.

6.2.2 Definitions

In addition to the definition in the previous chapter, a few basic definitions dedicated to this chapter are now recalled, while some more specific definitions will be given in later sections when needed.

Definition 6.1 (Basic Graph Theoretic Definitions)**.**

- *The* **distance** *between two vertices in a graph is the minimum number of edges connecting them.*
- *A* (K, d)**-regular bipartite graph** $\mathcal{G} = (\mathcal{U}, \mathcal{V}, \mathcal{E})$ *is such that* $|\mathcal{U}| = |\mathcal{V}| = K$ *and* $|\mathcal{T}_k| = |\mathcal{R}_k| = d,\ \forall\ k$.
- *For a bipartite graph* $\mathcal{G} = (\mathcal{U}, \mathcal{V}, \mathcal{E})$ *with edges* $\{e_{ij} \in \mathcal{E}\}$ *where* $i \in \mathcal{U}$ *and* $j \in \mathcal{V}$*, the* **biadjacency matrix**[1] $\boldsymbol{B}$ *is defined as*

$$[\boldsymbol{B}]_{ji} = \begin{cases} 1, & e_{ij} \in \mathcal{E} \\ 0, & otherwise \end{cases} . \tag{6.2}$$

  As a **reference graph***, the regular bipartite graph* $\mathcal{G}_r = (\mathcal{U}_r, \mathcal{V}_r, \mathcal{E}_r)$ *with biadjacency matrix* $\boldsymbol{B}_r$ *is characterized by*

$$[\boldsymbol{B}_r]_{ji} = \begin{cases} 1, & 0 \le i - j \le d - 1 \\ 0, & otherwise \end{cases} , \tag{6.3}$$

[1] It is slightly different from the classic definition, where the biadjacency matrix is in fact $\boldsymbol{B}^{\mathsf{T}}$ in classic definitions. We abuse this definition for the sake of presentational convenience. This matrix is also referred to as the "topology matrix" in [49].

which implies $\mathcal{T}_j = \{j, j+1, \ldots, j+d-1\}$. *Two bipartite graphs are said to be* **similar**, *denoted as* $\mathcal{G} \simeq \mathcal{G}_r$, *if their biadjacency matrices* $\boldsymbol{B}$ *and* $\boldsymbol{B}_r$ *satisfy* $\boldsymbol{B} = \boldsymbol{P}^T \boldsymbol{B}_r \boldsymbol{Q}$, *where* $\boldsymbol{P}$ *and* $\boldsymbol{Q}$ *are permutation matrices. Accordingly, it implies that* $\mathcal{U}$ *and* $\mathcal{V}$ *in* $\mathcal{G}$ *can be obtained by reordering the vertices of* $\mathcal{U}_r$ *and* $\mathcal{V}_r$ *in* $\mathcal{G}_r$ *with* $\mathcal{U} = \mathcal{U}_r$ *and* $\mathcal{V} = \mathcal{V}_r$, *respectively.*

- *A* **Hamiltonian cycle** *for a graph is a cycle that visits all vertices exactly once.*

- *A* **matching** *of the graph is a set of edges with no common vertices between any two edges. A* **perfect matching** *is a matching contains all vertices.*

Some special network topologies are therefore defined as follows accordingly.

Definition 6.2 (Some Special Network Topologies)**.** *A* (K, d)**-regular network** *refers to the* K*-cell network that representable by a regular bipartite graph* $\mathcal{G} = (\mathcal{U}, \mathcal{V}, \mathcal{E})$ *(i.e.,* $|\mathcal{U}| = |\mathcal{V}| = K$ *and* $|\mathcal{T}_k| = |\mathcal{R}_k| = d$, $\forall\ k$*), where each transmitter (resp. receiver) has* d *outgoing (resp. incoming) links. The reference network represented by* $\mathcal{G}_r$ *therefore corresponds to the* K*-cell network where each receiver will overhear the signals from the transmitter with the same index as well as the successive* $d-1$ *ones, i.e.,* $\mathcal{T}_j = \{j, j+1, \ldots, j+d-1\}$. *A* **triangular network** *refers to a category of cellular networks with* $\mathcal{T}_j = \{1, \ldots, j\}$ *(i.e.,* $\boldsymbol{B}$ *is a lower triangular matrix) or* $\mathcal{T}_j = \{j, \ldots, K\}$ *(i.e.,* $\boldsymbol{B}$ *is an upper triangular matrix), as well as those whose topology graphs are similar to either one.*

A rate tuple $(R_1, \ldots, R_K)$ is said to be *achievable* to the partially connected BC with no CSIT beyond topology (i.e., TIM-CoMP problem), if these exists a $(2^{nR_1}, \ldots, 2^{nR_K}, n)$ code scheme including the following elements:

- K message sets $\mathcal{W}_k \triangleq [1 : 2^{nR_k}]$, from which the message W_k is uniformly chosen, $\forall\ k \in \mathcal{K}$;

- one encoding function for Transmitter i ($\forall\ i \in \mathcal{K}$):

$$X_i(t) = f_{i,t}\left(W_{\mathcal{R}_i},\ \mathcal{G}\right), \tag{6.4}$$

  where only a subset of messages $W_{\mathcal{R}_i}$ is available at Transmitter i for encoding;

- one decoding function for Receiver j ($\forall\ j \in \mathcal{K}$):

$$\hat{W}_j = g_j\left(Y_j^n,\ \mathcal{H}^n,\ \mathcal{G}\right), \tag{6.5}$$

such that the average decoding error probability is vanishing as the code length n tends to infinity. The capacity region $\mathcal{C}$ is defined as the set of all achievable rate tuples.

In this chapter, we only consider the symmetric DoF as our main figure of merit [49–51, 97, 98, 121].

Definition 6.3 (Symmetric DoF)**.**

$$d_{\text{sym}} = \limsup_{P\to\infty} \sup_{(R_{\text{sym}},\ldots,R_{\text{sym}})\in\mathcal{C}} \frac{R_{\text{sym}}}{\log P} \tag{6.6}$$

where P is the average transmit power.

6.3 Main Results

In what follows, we will present a summary of our results by the following theorems.

- As a baseline, an interference avoidance approach with the aid of fractional graph coloring is first presented in Theorem 6.1 for the general topologies, followed by an outer bound in Theorem 6.2 built upon the concept of generator. With the inner and outer bounds we are able to characterize the optimal symmetric DoF of three-cell networks with arbitrary topologies.

- To gain further improvements, an interference alignment based approach is proposed in Theorem 6.3, by which the achievable symmetric DoF of the regular networks are identified. A new outer bound with the help of compound settings are derived in Theorem 6.4, with which the optimal symmetric DoF of Wyner-type networks with only one interfering link are characterized.

- This interference alignment based technique is further generalized to the arbitrary network topologies in Theorem 6.5 to identify the sufficient conditions achieving a certain amount of symmetric DoF. The conditions are based on the relations of transmit sets, which leads us to the construction of a hypergraph and hence an inner bound via hypergraph covering in Theorem 6.6.

- Later on, we bridge our problem to index coding problems in Theorem 6.7, showing that our problem is outer bounded by a corresponding index coding problem, by which the sufficient and necessary condition of the optimality of time-division is identified.

Before presenting the interference avoidance approach, we first introduce the following definition generalized from the standard graph coloring.

Definition 6.4 (Fractional Selective Graph Coloring)**.** *Consider an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with a partition $\mathbb{V} = \{\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_p\}$ where $\cup_{i=1}^{p} \mathcal{V}_i = \mathcal{V}$ and $\mathcal{V}_i \cap \mathcal{V}_j = \emptyset \ \forall \ i \neq j$. The portion $\mathcal{V}_i$ $(i \in \{1, 2, \ldots, p\})$ is called a cluster. A graph with the partition $\mathbb{V}$ is said to be* **selectively $n : m$-colorable**, *if*

- *each cluster $\mathcal{V}_i$ $(\forall \ i)$ is assigned a set of m colors drawn from a palette of n colors, no matter which vertex in the cluster receives;*
- *any two adjacent vertices have no colors in common.*

Denote by $s\chi_f(\mathcal{G}, \mathbb{V})$ the **fractional selective chromatic number** *of $n : m$-coloring over a graph with the partition $\mathbb{V}$, defined as*

$$s\chi_f(\mathcal{G}, \mathbb{V}) = \lim_{m \to \infty} \frac{s\chi_m(\mathcal{G}, \mathbb{V})}{m} = \inf_{m} \frac{s\chi_m(\mathcal{G}, \mathbb{V})}{m} \tag{6.7}$$

where $s\chi_m(\mathcal{G}, \mathbb{V})$ is the minimum n for the selective $n : m$-coloring regarding the partition $\mathbb{V}$.

Remark 6.1. *If $m = 1$, fractional selective graph coloring will be reduced to standard selective graph coloring (a.k.a. partition coloring) [126, 127]. If $|\mathcal{V}_i| = 1 \ \forall \ i \in \{1, 2, \ldots, p\}$, then fractional selective graph coloring will be reduced to standard fractional graph coloring.*

Theorem 6.1 (Achievable DoF via Graph Coloring)**.** *For the TIM-CoMP problem, the symmetric DoF*

$$d_{\text{sym}} = \frac{1}{s\chi_f(\mathcal{G}_e^2, \mathbb{V}_e)} \tag{6.8}$$

can be achieved by an interference avoidance approach built upon fractional selective graph coloring, where

- *$\mathcal{G}_e$: the line graph of $\mathcal{G} = (\mathcal{U}, \mathcal{V}, \mathcal{E})$, where the vertices in $\mathcal{G}_e$ are edges of $\mathcal{G}$;*

- $\mathbb{V}_e$*: a vertex partition of* $\mathcal{G}_e$*, where vertices in* $\mathcal{G}_e$ *associated with the edges in* $\mathcal{G}$ *that share the same vertex* $v_j \in \mathcal{V}$ $(\forall\ j)$ *form a cluster;*
- $\mathcal{G}_e^2$*: the square of* $\mathcal{G}_e$*, in which any two vertices in* $\mathcal{G}_e$ *with distance no more than 2 are joint with an edge;*
- $s\chi_f$*: fractional selective chromatic number of selective graph coloring, as defined in Definition 6.4.*

Proof. See Section 6.5.1 for a general proof and Section 6.4.1 for an illustrative example. □

By connecting the achievable symmetric DoF of our problem to the fractional selective chromatic number, we are able to calculate the former by computing the latter with rich toolboxes developed in graph theory.

To see how tight this interference avoidance scheme is, we provide an outer bound based on the concept of generator [51]. For simplicity of presentation, we introduce an index function $\mathbf{f}_{\text{idx}}$, which is defined as $\mathbf{f}_{\text{idx}} : \mathcal{B} \mapsto \{0,1\}^K$, to map the position indicated by $\mathcal{B} \subseteq \mathcal{K}$ to a $K \times 1$ binary vector with the corresponding position being 1, and 0 otherwise, e.g., $\mathbf{f}_{\text{idx}}(\{1,3,5\}) = [1\,0\,1\,0\,1\,0]^\intercal$ with $K = 6$. Thus, we have the following definition.

Definition 6.5 (Generator Sequence)**.** *Given* $\mathcal{S} \subseteq \mathcal{K}$*, a sequence* $\{\mathcal{I}_0, \mathcal{I}_1, \ldots, \mathcal{I}_S\}$ *is called a* **generator sequence***, if it is a partition of* $\mathcal{S}$ *(i.e.,* $\cup_{s=0}^{S} \mathcal{I}_s = \mathcal{S}$ *and* $\mathcal{I}_i \cap \mathcal{I}_j = \emptyset$*,* $\forall\ i \neq j$*), such that*

$$\boldsymbol{B}_{\mathcal{I}_s} \subseteq^{\pm} \mathit{rowspan}\left\{\boldsymbol{B}_{\mathcal{I}_0}, \mathbf{I}_{\mathcal{A}_s}\right\}, \quad \forall\ s = 1, \ldots, S \tag{6.9}$$

where $\boldsymbol{B}_{\mathcal{I}}$ *is the submatrix of* $\boldsymbol{B}$ *with rows of indices in* $\mathcal{I}$ *selected,* $\mathcal{A}_s \triangleq \{i | [\boldsymbol{B}^\intercal]_i \cdot \mathbf{f}_{\text{idx}}(\cup_{r=0}^{s-1}\mathcal{I}_r) = |\mathcal{R}_i \backslash \mathcal{S}^c|\}$ *with* $[\boldsymbol{B}^\intercal]_i$ *being the* i*-th row of* $\boldsymbol{B}^\intercal$ *(i.e.,* i*-th column of* $\boldsymbol{B}$*), and* $\mathbf{I}_{\mathcal{A}_s}$ *denotes a submatrix of* $\mathbf{I}_K$ *with the rows in* $\mathcal{A}_s$ *selected.* $\boldsymbol{A}_1 \subseteq^{\pm} \mathit{rowspan}\{\boldsymbol{A}_2\}$ *is such that two matrices* $\boldsymbol{A}_1 \in \mathbb{C}^{m_1 \times n}$ *and* $\boldsymbol{A}_2 \in \mathbb{C}^{m_2 \times n}$ *satisfy* $\boldsymbol{A}_1 = \boldsymbol{C}\boldsymbol{A}_2\mathbf{I}^{\pm}$*, where* $\boldsymbol{C} \in \mathbb{C}^{m_1 \times m_2}$ *can be any full rank matrix,* $\mathbf{I}^{\pm}$ *is as same as the identity matrix up to the sign of elements. This implies that the row of* $\boldsymbol{A}_1$ *can be represented by the rows of* $\boldsymbol{A}_2$ *with possible difference of signs of elements. We refer to* $\mathcal{I}_0$ *as the initial generator with regard to* $\mathcal{S}$*, and denote by* $\mathcal{J}(\mathcal{S})$ *all the possible initial generators.*

Theorem 6.2 (Outer Bound via Generator Sequence)**.** *The symmetric DoF of the* K*-cell TIM-CoMP problem are upper bounded by*

$$d_{\text{sym}} \leq \min_{\mathcal{S} \subseteq \mathcal{K}} \min_{\mathcal{I}_0 \subseteq \mathcal{J}(\mathcal{S})} \frac{|\mathcal{I}_0|}{|\mathcal{S}|} \tag{6.10}$$

where $\mathcal{I}_0$ is the initial generator, from which a sequence can be initiated and generated subsequently as defined in Definition 6.5.

Proof. See Section 6.5.2 for a general proof and Section 6.4.1 for an illustrative example. □

By this outer bound and the interference avoidance approach, we characterize the optimal symmetric DoF of three-cell networks with arbitrary topologies as well as the triangular networks below.

Corollary 6.1 (Optimal DoF for Three-cell Networks)**.** *The optimal symmetric DoF of the three-cell TIM-CoMP problem can be achieved by interference avoidance.*

Proof. See Section 6.5.3. □

Corollary 6.2 (Optimal DoF for Triangular Networks)**.** *For the K-cell triangular networks, the optimal symmetric DoF value of the TIM-CoMP problem is $\frac{1}{K}$.*

Proof. See Section 6.5.4. □

The above interference avoidance approach is in general efficient to the sparse networks, whereas the graph coloring complexity increases dramatically as the number of edges increases, which motivates ourselves to resort to advanced interference management techniques, such as interference alignment. Building upon the concept of interference alignment, we identify the achievable symmetric DoF of regular networks as follows.

Theorem 6.3 (Achievable DoF for Regular Networks)**.** *For a (K, d)-regular network represented by a regular bipartite graph $\mathcal{G}$, as long as it is similar to the reference bipartite graph $\mathcal{G}_r$, i.e., $\mathcal{G} \simeq \mathcal{G}_r$, the symmetric DoF*

$$d_{\mathrm{sym}}(K,d) = \begin{cases} \frac{2}{d+1}, & d \leq K-1 \\ \frac{1}{K}, & d = K \end{cases} \tag{6.11}$$

can be achieved by interference alignment approaches, when the channel coherence time satisfies $\tau_c \geq d+1$.

Proof. See Section 6.5.5 for a general proof and Section 6.4.2 for an illustrative example. □

For the regular networks, the outer bound via generator sequence becomes loose. This urges us to find another bounding techniques. By generalizing and extending the idea in [50], we obtain in what follows a new outer bound with the aid of compound settings.

Theorem 6.4 (Outer Bound via Compound Settings)**.** *The symmetric DoF of K-cell TIM-CoMP problems are upper bounded by the solution of the following optimization problem:*

$$\min_{\mathcal{S}\subseteq\mathcal{K}} \quad \frac{K-|\mathcal{S}'|}{2K-|\mathcal{S}'|-|\mathcal{S}|} \tag{6.12}$$

$$s.t. \quad \mathcal{S}'=\{i|[\boldsymbol{B}^{\mathsf{T}}]_i\cdot\mathbf{f}_{\text{idx}}(\mathcal{S})=|\mathcal{R}_i|\} \tag{6.13}$$

$$\cup_{j\in\mathcal{S}}\mathcal{T}_j=\mathcal{K} \tag{6.14}$$

where $[\boldsymbol{B}^{\mathsf{T}}]_i$ is the i-th row of $\boldsymbol{B}^{\mathsf{T}}$ (i.e., i-th column of $\boldsymbol{B}$).

Proof. See Section 6.5.6 for a general proof and Section 6.4.2 for an illustrative example. □

By the above outer bound, we are able to characterize the optimal symmetric DoF of a subset of regular networks.

Corollary 6.3 (Optimal DoF of Cyclic Wyner-type Networks)**.** *For a $(K,2)$-regular network, e.g., a cyclic Wyner-type network, the optimal symmetric DoF are*

$$d_{\text{sym}}(K,2)=\begin{cases}\frac{1}{2}, & K=2\\ \frac{2}{3}, & K\geq 3\end{cases} \tag{6.15}$$

if the coherence time $\tau_c\geq 3$ when $K\geq 3$.

Proof. See Section 6.5.7. □

As the previous theorems identify the achievable symmetric DoF of regular networks, we may wonder to what conditions the network is subject for some symmetric DoF targets. In what follows, we define alignment-feasible graph, proper partition, and alignment non-conflict matrix, and identify the sufficient conditions to achieve a certain amount of symmetric DoF.

Definition 6.6 (Alignment-Feasible Graph)**.** *The* **alignment-feasible graph (AFG)***, denoted by $\mathcal{G}_{AFG}$, refers to a graph with vertices representing the messages and with edges between any two messages indicating if they are alignment-feasible. The two messages W_i and W_j are said to be alignment-feasible, denoted by $i\leftrightarrow j$, if*

$$\mathcal{T}_i\nsubseteq\mathcal{T}_j, \quad \textit{and} \quad \mathcal{T}_j\nsubseteq\mathcal{T}_i. \tag{6.16}$$

The condition in (6.16) implies the alignment feasibility, that is, if two messages are alignment-feasible, then whenever one of them is desired by one receiver, the other one must be absent to this receiver. The similar insight was also revealed in [97] in the context of index coding. While the alignment feasibility condition in (6.16) is given with regard to the transmit sets, it can be also expressed by the receive sets, i.e., W_i and W_j are alignment-feasible, if $\mathcal{R}_i \nsubseteq \mathcal{R}_j$ and $\mathcal{R}_j \nsubseteq \mathcal{R}_i$. This alignment feasibility can also be generated to multiple messages, as shown in the following definition.

Definition 6.7 (Proper Partition)**.** *A partition $\mathcal{K} = \{\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_\kappa\}$ with size κ, where $\cup_{i=1}^{\kappa} \mathcal{P}_i = \mathcal{K}$ and $\mathcal{P}_i \cap \mathcal{P}_j = \emptyset \ \forall \ i \neq j$, is called a* **proper partition***, if for every portion $\mathcal{P}_i = \{i_1, i_2, \ldots, i_{p_i}\}$ with $p_i \triangleq |\mathcal{P}_i|$ $(i \in [\kappa])$, we have*

$$\mathcal{T}_{i_k} \bigcap \left(\bigcup_{i_j \in \mathcal{P}_i \backslash i_k} \mathcal{T}_{i_j} \right)^c \neq \emptyset, \quad \forall \ i_k \in \mathcal{P}_i. \tag{6.17}$$

Definition 6.8 (Alignment Non-Conflict Matrix)**.** *Regarding an alignment-feasible cycle $i_1 \leftrightarrow i_2 \leftrightarrow \cdots \leftrightarrow i_K \leftrightarrow i_1$, we construct a $K \times K$ binary matrix $\boldsymbol{A}^{AFG}$, referred to as* **alignment non-conflict matrix***, with element $\boldsymbol{A}_{kj}^{AFG} = 1$ $(j, k \in \mathcal{K})$, if*

$$\mathcal{T}_{i_j} \cap \mathcal{T}_{i_{j+1}}^c \nsubseteq \mathcal{T}_{i_k}, \quad \textit{and} \quad \mathcal{T}_{i_{j+1}} \cap \mathcal{T}_{i_j}^c \nsubseteq \mathcal{T}_{i_k}, \tag{6.18}$$

and with $\boldsymbol{A}_{kj}^{AFG} = 0$ otherwise. Further, we reset $\boldsymbol{A}_{kj}^{AFG} = 0$, if

$$\mathcal{T}_{i_j} \bigcap \mathcal{T}_{i_{j+1}}^c \bigcap_{k:\boldsymbol{A}_{kj}^{AFG}=1} \mathcal{T}_{i_k}^c = \emptyset, \quad \textit{or} \quad \mathcal{T}_{i_{j+1}} \bigcap \mathcal{T}_{i_j}^c \bigcap_{k:\boldsymbol{A}_{kj}^{AFG}=1} \mathcal{T}_{i_k}^c = \emptyset. \tag{6.19}$$

Similarly, for a proper partition $\{\mathcal{P}_1, \ldots, \mathcal{P}_\kappa\}$, we construct a $\kappa \times \kappa$ binary matrix $\boldsymbol{A}^{PP}$, with $\boldsymbol{A}_{ij}^{PP} = 1$ $(j, i \in [\kappa])$, if

$$\mathcal{T}_{j_t} \bigcap \left(\bigcup_{j_s \in \mathcal{P}_j \backslash j_t} \mathcal{T}_{j_s} \right)^c \nsubseteq \mathcal{T}_{i_k}, \quad \forall \ j_t \in \mathcal{P}_j, \forall \ i_k \in \mathcal{P}_i \tag{6.20}$$

and with $\boldsymbol{A}_{ij}^{PP} = 0$ otherwise. Further, we reset $\boldsymbol{A}_{ij}^{PP} = 0$, if there exist $j_t \in \mathcal{P}_j$ and $i_k \in \mathcal{P}_i$, such that

$$\mathcal{T}_{j_t} \bigcap \left(\bigcup_{j_s \in \mathcal{P}_j \backslash j_t} \mathcal{T}_{j_s} \right)^c \bigcap_{i:\boldsymbol{A}_{ij}^{PP}=1} \mathcal{T}_{i_k}^c = \emptyset. \tag{6.21}$$

Either the pairs in alignment-feasible graph or the elements in each portion of proper partition imply that the corresponding messages are able to align in the same subspace, while the alignment non-conflict matrix identifies the subspace that can be absent to some receivers. As such, the sufficient conditions to achieve a certain amount of symmetric DoF are presented as follows.

Theorem 6.5 (Sufficient Conditions for Arbitrary Networks)**.** *For a K-cell cellular network with arbitrary topologies, with sufficient coherence time, the following symmetric DoF are achievable:*

- $d_{\text{sym}} = \frac{2}{K}$*, if there exists a Hamiltonian cycle or a perfect matching in* $\mathcal{G}_{AFG}$*;*
- $d_{\text{sym}} = \frac{2}{K-q}$*, if there exists a Hamiltonian cycle in graph* $\mathcal{G}_{AFG}$*, say* $i_1 \leftrightarrow i_2 \leftrightarrow \cdots \leftrightarrow i_K \leftrightarrow i_1$*, associated with an alignment non-conflict matrix* $\boldsymbol{A}^{AFG}$*, such that*

$$q \triangleq \min_k \sum_j \boldsymbol{A}_{kj}^{AFG}; \tag{6.22}$$

- $d_{\text{sym}} = \frac{1}{\kappa}$*, if there exists a proper partition with size* κ*;*
- $d_{\text{sym}} = \frac{1}{\kappa-q}$*, if there exists a proper partition with size* κ*, say* $\{\mathcal{P}_1, \ldots, \mathcal{P}_\kappa\}$*, associated with an alignment non-conflict matrix* $\boldsymbol{A}^{PP}$*, such that*

$$q \triangleq \min_i \sum_j \boldsymbol{A}_{ij}^{PP}. \tag{6.23}$$

Proof. See Section 6.5.8 for general proofs and illustrative examples. □

There is a very interesting observation. The alignment-feasible condition in (6.16) or a proper portion in (6.17) implies the feasibility of a proper distance-2 graph coloring. The two messages satisfy (6.16) means there exist two vertices in two clusters i and j of $\mathcal{G}_e$ that can be assigned the same color, so do the messages in (6.17). It follows that the interference alignment is a general form of interference avoidance, in agreement with the observation in [50]. Specifically, any two (or more) vertices in clusters $j_{\mathcal{S}}$ ($\mathcal{S} \subseteq \mathcal{K}$) in $\mathcal{G}_e$ (corresponding to edges in $\mathcal{G}$ connecting Transmitter i_s to Receiver j_s ($\forall\ s \in \mathcal{S}$)) that receive the same color are scheduled in a single time slot with interference avoided, implying that the transmitted signals in the form of $\{X_{i_s}(W_{j_s}), s \in \mathcal{S}\}$ are alignment-feasible in the same subspace.

Thus, interference alignment provides at least the same performance as interference avoidance. Even better, one advantage of interference alignment over interference avoidance is that, the number of dimensions of the subspace to make interference alignment feasible could be less than the total number of colors (i.e., the total number of time slots to schedule links), as some subspaces may be absent to some receivers so as to decrease the number of required dimensions.

From the previous theorem, we observe that the messages connected by an edge in $\mathcal{G}_{AFG}$ or lied in the same portion of a proper partition are able to be scheduled at the same time slot or be aligned at the same direction. Inspired by this observation, we construct a hypergraph and translate our problem into a covering problem of this hypergraph. We start with the construction of this hypergraph as follows.

Definition 6.9 (Hypergraph Covering)**.** *A hypergraph $\mathcal{H}_\mathcal{G} = (\mathcal{S}, \mathcal{X})$ associated with $\mathcal{G}$ is composed of the vertex set $\mathcal{S} \subseteq \mathcal{K}$ being a finite set, and the hyperedge set $\mathcal{X}$ being a family of subsets of $\mathcal{S}$, where $\mathcal{X}_i \triangleq \{x_{i_1}, x_{i_2}, \ldots, x_{i_{|\mathcal{X}_i|}}\} \subseteq \mathcal{S}$ is called a hyperedge, i.e., $\mathcal{X}_i \in \mathcal{X}$, as long as*

$$\mathcal{T}_{x_{i_k}} \bigcap \left(\bigcup_{x_{i_j} \in \mathcal{X}_i \backslash x_{i_k}} \mathcal{T}_{x_{i_j}} \right)^c \neq \emptyset, \quad \forall\ x_{i_k} \in \mathcal{X}_i. \tag{6.24}$$

A covering of a hypergraph $\mathcal{H}_\mathcal{G}$ is a collection of hyperedges $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_\tau$ such that $\mathcal{S} \subseteq \cup_{j=1}^{\tau} \mathcal{X}_j$, and the least number of τ is called the **hypergraph covering number***, denoted by $\tau(\mathcal{H}_\mathcal{G})$. A t-fold covering is a multiset $\{\mathcal{X}_1, \ldots, \mathcal{X}_\tau\}$ such that each $s \in \mathcal{S}$ is in at least t of the $\mathcal{X}_i$'s, and correspondingly $\tau_t(\mathcal{H}_\mathcal{G})$ is referred to as the t-fold covering number. Accordingly, the* **hypergraph fractional covering number** *is defined to be*

$$\tau_f(\mathcal{H}_\mathcal{G}) \triangleq \lim_{t \to \infty} \frac{\tau_t(\mathcal{H}_\mathcal{G})}{t} = \inf_t \frac{\tau_t(\mathcal{H}_\mathcal{G})}{t}. \tag{6.25}$$

Theorem 6.6 (Achievable DoF via Hypergraph Covering)**.** *For the TIM-CoMP problem, the symmetric DoF*

$$d_{\text{sym}} = \frac{1}{\tau_f(\mathcal{H}_\mathcal{G})} \tag{6.26}$$

are achievable, where $\tau_f(\mathcal{H}_\mathcal{G})$ is the fractional covering number of the hypergraph $\mathcal{H}_\mathcal{G} = (\mathcal{K}, \mathcal{X})$ constructed in Definition 6.9 with the vertex set $\mathcal{K}$ and the hyperedge set $\mathcal{X}$ including all satisfactory subsets $\mathcal{X}_i \subseteq \mathcal{K}$.

Proof. See Section 6.5.9. □

The characterization of the fractional hypergraph covering number $\tau_f(\mathcal{H}_\mathcal{G})$ can also be performed by the following integer linear programming relaxation

$$\tau_f(\mathcal{H}_\mathcal{G}) = \min \quad \sum_{i\in\mathcal{K}} \rho_i \tag{6.27}$$

$$s.t. \quad \sum_{i\in\mathcal{K}:j\in\mathcal{X}_i} \rho_i \geq 1, \quad \forall\ j \in \mathcal{K} \tag{6.28}$$

$$\rho_i \in [0,1], \quad \forall\ i \in \mathcal{K} \tag{6.29}$$

where ρ_i is an indicator variable associated with the hyperedge $\mathcal{X}_i \in \mathcal{X}$ with value between 0 and 1 indicating the weight assigned to $\mathcal{X}_i$ accounts for the total weight, the first constraint ensures that every vertex in $\mathcal{K}$ is covered at least once, and the last constraint specifies a fractional ρ_i, which is the relaxation of integers $\{0,1\}$. Although the optimization of this linear program is NP-hard, the connection of our problem and hypergraph covering bridges the fields of information theory and graph theory, such that the progress on one problem is automatically transferrable to the other one.

Knowing that the TIM problem was nicely bridged to the index coding problem [49], one may wonder if there exist relations between our problem and index coding. Indeed, our problem can be also related to the index coding problem. Before presenting this relation, we first define the index coding problem and its demand graph similarly to those in [49, 128].

Definition 6.10 (Index Coding)**.** *A multiple unicast* **index coding** *problem, denoted as* $\mathbf{IC}(k|\mathcal{S}_k)$*, is comprised of a transmitter who wants to send* K *messages* $W_k, k \in \mathcal{K}$ *to their respective receivers over a noiseless link, and* K *receivers, each of which has prior knowledge of* $W_{\mathcal{S}_k}$ *with* $\mathcal{S}_k \subseteq \mathcal{K}\backslash k$*. Its* **demand graph** *is a directed bipartite graph* $\mathcal{G}_d = (\mathcal{W}, \mathcal{K}, \mathcal{E})$ *with vertices of Message* $W_k \in \mathcal{W}$ *and Receiver* k $(k \in \mathcal{K})$*, and there exists a directed forward edge* $i \to j$ *from Message* W_i *to Receiver* j *if* W_i *is demanded by Receiver* j *and a backward edge* $k \leftarrow j$ *from Receiver* j *to Message* W_k *if Receiver* j *has the knowledge of* W_k *as side information.*

Theorem 6.7 (Outer Bound via Index Coding)**.** *For the TIM-CoMP problem, given the topology information* $\{\mathcal{T}_k, \mathcal{R}_k, \forall\ k \in \mathcal{K}\}$*, the DoF region is outer bounded by the capacity region of a multiple unicast index coding problem* $\mathbf{IC}(k|\mathcal{S}_k)$*, where*

$$\mathcal{S}_k \triangleq \bigcup_{j\in\mathcal{T}_k^c} \mathcal{R}_j. \tag{6.30}$$

Proof. See Section 6.5.10 for the general proof and examples. □

The above theorem implies that the outer bounds of the multiple unicast index coding problem in literature still hold for our problem, but with the modified side information sets. Together with Corollary 7 in [49], stating that the symmetric DoF value of $\frac{1}{K}$ is optimal to a K-unicast index coding problem if and only if its demand graph is acyclic, we obtain in the following corollary to identify the sufficient condition of the optimality of symmetric DoF $\frac{1}{K}$ in our problem.

Corollary 6.4. *For TIM-CoMP problems, the symmetric DoF value* $d_{\text{sym}} = \frac{1}{K}$ *is optimal if the demand graph associated with the index coding problem* $\mathbf{IC}(k|\bigcup_{j\in\mathcal{T}_k^c}\mathcal{R}_j)$ *is acyclic.*

Proof. See Section 6.5.11. □

While the DoF region of TIM problem is outer bounded by the capacity region of the index coding problem $\mathbf{IC}(k|\mathcal{T}_k^c)$, our problem with transmitter cooperation is outer bounded by $\mathbf{IC}(k|\cup_{j\in\mathcal{T}_k^c}\mathcal{R}_j)$. In general, this bound is loose, because the side information might be over-endowed to the receivers. Fortunately, it can be used to identify the sufficient and necessary condition to achieve $d_{\text{sym}} = \frac{1}{K}$.

The sufficient condition provided in Corollary 6.4 on the optimality of time division through a reconstructed index coding problem is nonetheless implicit, yet the necessity is still unclear. Considering both the definition of $\mathcal{G}_{AFG}$ and Corollary 6.4, we arrive at the sufficient and necessary condition of the symmetric DoF optimality of $\frac{1}{K}$.

Corollary 6.5. *For the* K*-cell TIM-CoMP problem, the symmetric DoF value* $d_{\text{sym}} = \frac{1}{K}$ *is optimal, if and only if* $\mathcal{G}_{AFG}$ *is an empty graph.*

Proof. See Section 6.5.12. □

Remark 6.2. *For the triangular network, the alignment feasible graph is empty and thus the symmetric DoF value is* $\frac{1}{K}$*, which coincides with Corollary 6.2. Note that this triangular network is the minimum graph with empty alignment-feasible graph.*

6.4 Illustrative Examples

In what follows, we start with presenting two illustrative examples, with which the basic ideas of Theorems 6.1-6.4 are explained here. The general proofs are relegated to the next section.

6.4.1 Example 1

In this example, we focus on the network topology studied in [51], as shown in Fig. 6.1, but with a key difference that message sharing across transmitters is enabled. As in [49,51,98], the optimal symmetric DoF value is pessimistically $\frac{1}{3}$ without message sharing. In contrast, if transmitter cooperation is allowed, the achievable symmetric DoF can be remarkably improved to $\frac{2}{5}$ even with a simple interference avoidance scheme according to Theorem 6.1. An outer bound with $d_{\text{sym}} \leq \frac{1}{2}$ based on generator sequence is also presented according to Theorem 6.2.

Inner Bound via Interference Avoidance

Without message sharing, the interference avoidance scheme consists in scheduling transmitters to avoid mutual interferences. For instance, by delivering W_1, Transmitter 1 will cause interferences to Receivers 2 and 3, and consequently Transmitters 2 and 3 should be deactivated, because W_2 and W_3 cannot be delivered to Receivers 2 and 3 free of interference. In contrast, with message sharing, the desired message W_1 can be sent either from Transmitter 1 or 4. Hence, scheduling can be done across links rather than across transmitters. For instance, if the link Transmitter 4 $\rightarrow$ Receiver 1 (denoted by e_{41}) is scheduled, the links adjacent to e_{41} (i.e., e_{11}, e_{42}, and e_{44}) as well as the links adjacent to e_{11}, e_{42} and e_{44} (i.e., e_{12}, e_{13}, e_{22}, e_{32}, e_{34} and e_{54}) should not be scheduled, because activating Transmitter 1 will interfere Receivers 1, and Receivers 2 and 4 will overhear interferences from Transmitter 4 such that any delivery from Transmitter 1 or to Receivers 2 and 4 causes mutual interferences. A possible link scheduling associated with Fig. 6.1 is shown in Table 6.1. It can be found that each message is able to be independently delivered twice during five time slots, and hence symmetric DoF of $\frac{2}{5}$ are achievable.

Table 6.1: Link Scheduling

Slot	Scheduled Links (e_{ij}: TX $i \rightarrow$ RX j)	Delivered Messages
A	e_{41}, e_{55}, e_{66}	W_1, W_5, W_6
B	e_{12}, e_{54}, e_{66}	W_2, W_4, W_6
C	e_{13}, e_{54}	W_3, W_4
D	e_{41}, e_{33}	W_1, W_3
E	e_{12}, e_{55}	W_2, W_5

Although the above link scheduling solution provides an achievable scheme

for the example in Fig. 6.1, the generalization is best undertaken by reinterpreting the link scheduling into a graph coloring problem, such that the rich graph theoretic toolboxes can be directly utilized to solve our problem. The nature of our problem calls for a distance-2 fractional clustered-graph coloring scheme, which consists of the following ingredients:

- *Distance-2 fractional coloring*: Both the adjacent links and the adjacency of the adjacent links (resp. edges less than two hops) should be scheduled in difference time slots (resp. assigned with different colors).
- *Clustered-graph coloring*: Only the total number of messages delivered by links with the common receiver (resp. colors assigned to the edges with the same vertex) matters. Thus, the number of assigned colors should be counted by clusters where the edges with common vertices are grouped together.

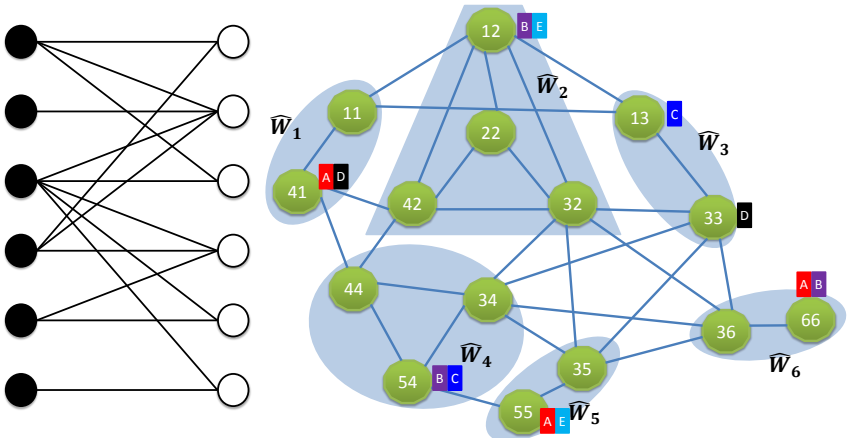


Figure 6.1: An instance of TIM-CoMP problem ($K = 6$). On the left is the network topology graph $\mathcal{G}$, and on the right is its line graph $\mathcal{G}_e$. The distance-2 fractional coloring is performed to offer each cluster two out of in total five colors, where any two vertices that receives the same color are set apart with distance no less than 2.

In what follows, we reinterpret the link scheduling as a distance-2 fractional clustered-graph coloring, and defer the elaboration of the relation with fractional selective graph coloring to the general proof. To ease presentation, we transform graph edge-coloring into graph vertex-coloring of its line graph. As shown in Fig. 6.1, we first transform the topology graph $\mathcal{G}$ (left) into its line graph $\mathcal{G}_e$ (right) and map the links connected to each receiver in $\mathcal{G}$ to the vertices in $\mathcal{G}_e$. For instance, the four links to Receiver 2 in $\mathcal{G}$ are

mapped to Vertices $e_{12}, e_{22}, e_{32}, e_{42}$ in $\mathcal{G}_e$. Then, we group relevant vertices in $\mathcal{G}_e$ as clusters, e.g., Vertices $e_{12}, e_{22}, e_{32}, e_{42}$ in $\mathcal{G}_e$ corresponding to the links to Receiver 2 are grouped as one cluster. By now, a clustered-graph is generated. The graph coloring can be performed as follows. For instance, if Vertex e_{41} in $\mathcal{G}_e$ receives a color indicated by 'A', then Vertices e_{55} and e_{66} can receive the same color, because the distance between any two of them is more than 2. Try any possible color assignment until we obtain a proper one, where each cluster receives m distinct colors out of total n ones, such that any two vertices with distance no more than 2 receive distinct colors. There may exist many proper color assignments.

This distance-2 fractional clustered graph coloring over the line graph of network topology, coincides with the definition of fractional selective graph coloring over the square of line graph given a certain vertex clustering. The fractional selective chromatic number $s\chi_f(\mathcal{G}_e^2)$ refers to the minimum of $\frac{n}{m}$ among all proper color assignments. In this example, we have $m = 2$ and $n = 5$. The vertices (i.e., links in $\mathcal{G}$) with the same color can be scheduled in the same time slot. Accordingly, each cluster receives two out of five colors means every message is scheduled twice during five time slots, yielding the symmetric DoF of $\frac{2}{5}$. According to the connection between link scheduling and graph coloring, the inverse of the fractional selective chromatic number, i.e., $\frac{1}{s\chi_f(\mathcal{G}_e^2)}$, can serve as an inner bound of symmetric DoF of the general cellular networks, although its computation is still NP-hard.

Outer Bound via Generator Sequence

According to the topology of the above cellular network, we have the transmit sets $\mathcal{T}_1 = \{1, 4\}$, $\mathcal{T}_2 = \{1, 2, 3, 4\}$, $\mathcal{T}_3 = \{1, 3\}$, $\mathcal{T}_4 = \{3, 4, 5\}$, $\mathcal{T}_5 = \{3, 5\}$, $\mathcal{T}_6 = \{3, 6\}$ and receive sets $\mathcal{R}_1 = \{1, 2, 3\}$, $\mathcal{R}_2 = \{2\}$, $\mathcal{R}_3 = \{2, 3, 4, 5, 6\}$, $\mathcal{R}_4 = \{1, 2, 4\}$, $\mathcal{R}_5 = \{4, 5\}$, $\mathcal{R}_6 = \{6\}$. With the message sharing strategy mentioned earlier, the messages $W_{\mathcal{R}_i}$ are accessible at Transmitter i.

Before proceeding further, we define the following virtual signals

$$\tilde{Y}_1^n \triangleq h_1^n X_1^n + h_4^n X_4^n + \tilde{Z}_1^n \tag{6.31}$$

$$\tilde{Y}_4^n \triangleq h_3^n X_3^n + h_4^n X_4^n + h_5^n X_5^n + \tilde{Z}_4^n, \tag{6.32}$$

where h_k^n $(k = 1, \ldots, 6)$ is assumed to be independent and identically distributed as h_{ji}^n when there is a link between Transmitter i and Receiver j, and the noise terms $\{\tilde{Z}_1^n, \tilde{Z}_4^n\}$ are identically distributed as Z_j with zero-mean and unit-variance. Given the fact that the distribution of channel gain is

symmetric around zero, it follows that $\{\tilde{Y}_1^n, \tilde{Y}_4^n\}$ are statistically equivalent to $\{Y_1^n, Y_4^n\}$, respectively. From both $\{\tilde{Y}_1^n, \tilde{Y}_4^n\}$ and $\{Y_1^n, Y_4^n\}$, the corresponding messages $\{\hat{W}_1, \hat{W}_4\}$ can be decoded with error probability tends to 0 as $n \to \infty$.

As symmetric DoF metric is considered, the DoF outer bound regarding any subset of messages serves as one candidate in general. Let us consider a subset of messages $W_{\mathcal{S}}$ with $\mathcal{S} = \{1, 3, 4, 5\}$, where W_i ($i \in \mathcal{S}^c = \{2, 6\}$) are set to be deterministic. Note that eliminating some messages or setting them to be deterministic does not hurt the maximum achievable rate of remaining messages. Thus, the sum rate of users in $\mathcal{S}$ can be upper bounded as

$$
\begin{aligned}
n \sum_{i \in \mathcal{S}} R_i &= H(W_{\mathcal{S}} | \mathcal{H}^n, \mathcal{G}) && (6.33)\\
&= I(W_{\mathcal{S}}; \tilde{Y}_{1,4}^n | \mathcal{H}^n, \mathcal{G}) + H(W_{\mathcal{S}} | \tilde{Y}_{1,4}^n, \mathcal{H}^n, \mathcal{G}) && (6.34)\\
&= I(W_{\mathcal{S}}; \tilde{Y}_{1,4}^n | \mathcal{H}^n, \mathcal{G}) + H(W_{1,4} | \tilde{Y}_{1,4}^n, \mathcal{H}^n, \mathcal{G}) \\
&\quad + H(W_{\mathcal{S}\setminus\{1,4\}} | W_{1,4}, \tilde{Y}_{1,4}^n, \mathcal{H}^n, \mathcal{G}) && (6.35)\\
&\le 2n \log P + H(W_{\mathcal{S}\setminus\{1,4\}} | W_{1,4}, \tilde{Y}_{1,4}^n, \mathcal{H}^n, \mathcal{G}) + n \cdot O(1) + n\epsilon_n && (6.36)
\end{aligned}
$$

where the last inequality is obtained by Fano's inequality, and $n\epsilon_n \triangleq 1 + nRP_e^{(n)}$ tends to zero as $n \to \infty$ by the assumption that $\lim_{n\to\infty} P_e^{(n)} = 0$. Since the transmitted signal X_i^n is encoded from the messages $W_{\mathcal{R}_i}$ ($\forall\ i$), it suffices to reproduce X_4^n and X_5^n from W_1, W_4 and W_4, W_5, respectively, with W_2, W_6 switched off (i.e., being set to be deterministic). Thus, we have

$$
\begin{aligned}
&H(W_{\mathcal{S}\setminus\{1,4\}} | W_{1,4}, \tilde{Y}_{1,4}^n, \mathcal{H}^n, \mathcal{G}) && (6.37)\\
&\quad = H(W_{3,5} | W_{1,4}, X_4^n, \tilde{Y}_{1,4}^n, \mathcal{H}^n, \mathcal{G}) && (6.38)\\
&\quad = H(W_5 | W_{1,4}, X_4^n, \tilde{Y}_{1,4,5}^n, \mathcal{H}^n, \mathcal{G}) + H(W_3 | W_{1,4,5}, X_4^n, \tilde{Y}_{1,4}^n, \mathcal{H}^n, \mathcal{G}) && (6.39)\\
&\quad \le H(W_5 | \tilde{Y}_5^n, \mathcal{H}^n) + H(W_3 | W_{1,4,5}, X_4^n, X_5^n, \tilde{Y}_{1,4}^n, \mathcal{H}^n, \mathcal{G}) && (6.40)\\
&\quad \le n\epsilon_n + H(W_3 | W_{1,4,5}, X_4^n, X_5^n, \tilde{Y}_{1,4}^n, \mathcal{H}^n, \mathcal{G}) && (6.41)\\
&\quad = n\epsilon_n + H(W_3 | W_{1,4,5}, X_4^n, X_5^n, \tilde{Y}_{1,3,4}^n, \mathcal{H}^n, \mathcal{G}) && (6.42)\\
&\quad \le n\epsilon_n + H(W_3 | \tilde{Y}_3^n, \mathcal{H}^n, \mathcal{G}) && (6.43)\\
&\quad \le n\epsilon_n && (6.44)
\end{aligned}
$$

where (6.38) is from the fact that it suffices to reproduce X_4^n from $W_{1,4}$, (6.39) is because of the chain rule of entropy and the fact that $\tilde{Y}_5^n = \tilde{Y}_4^n - h_4^n X_4^n =$

$h_3^n X_3^n + h_5^n X_5^n + \tilde{Z}_4^n$ can be generated from $\tilde{Y}_4^n$ and X_4^n, (6.40) is due to 1) removing condition does not reduce entropy, and 2) X_5^n can be obtained given the messages $W_{4,5}$, (6.42) comes from the generator sequence where $\tilde{Y}_3^n = \tilde{Y}_1^n - \tilde{Y}_4^n + h_5^n X_5^n = h_1^n X_1^n - h_3^n X_3^n + \tilde{Z}_1^n - \tilde{Z}_4^n$ can be generated from $\tilde{Y}_{1,4}^n$ and X_5^n, (6.43) is due to removing condition does not decrease entropy, and inequality (6.41) and the last inequalities are due to Fano's inequality, where $\tilde{Y}_5^n$ and $\tilde{Y}_3^n$ are statistically equivalent to Y_5^n and Y_3^n respectively, with bounded difference of noise variance, such that both W_5 and W_3 can be decoded respectively with negligible errors. Hence, we have

$$n \sum_{i \in \mathcal{S}} R_i \leq 2n \log P + n \cdot O(1) + n\epsilon_n \tag{6.45}$$

which leads to one possible outer bound for symmetric DoF

$$d_{\text{sym}} \leq \frac{1}{2}. \tag{6.46}$$

In contrast, by compound settings it gives a looser outer bound $d_{\text{sym}} \leq \frac{4}{7}$. Seemingly the generator sequence bounding is more suitable to asymmetric network topologies.

To summarize, we first generate two statistically equivalent signals $\{\tilde{Y}_1^n, \tilde{Y}_4^n\}$, then with the messages W_1, W_4, we reconstruct X_4^n, and then generate $\tilde{Y}_5^n$ from $\tilde{Y}_4^n$. Finally, $\tilde{Y}_3^n$ can be generated by knowing $\{\tilde{Y}_1^n, \tilde{Y}_4^n\}$ and X_5^n encoded from W_4, W_5. As such, the generation sequence is $\{\{1,4\},\{5\},\{3\}\}$, initiated from $\mathcal{I}_0 = \{1,4\}$. With $\mathcal{S} = \{1,3,4,5\}$, according to Definition 6.5, we have

$$\boldsymbol{B} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^{\mathsf{T}}, \quad \boldsymbol{B}_{\{1,4\}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}^{\mathsf{T}} \quad \boldsymbol{B}_5 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}^{\mathsf{T}} \quad \boldsymbol{B}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}^{\mathsf{T}} \tag{6.47}$$

and $\mathcal{A}_1 = \{4\}$, $\mathcal{A}_2 = \{4,5\}$, and $\mathcal{A}_3 = \{1,3,4,5\}$. It is readily verified that $\boldsymbol{B}_5 \subseteq^{\pm} rowspan\{\boldsymbol{B}_{\{1,4\}}, \mathbf{I}_{\mathcal{A}_1}\}$ and $\boldsymbol{B}_3 \subseteq^{\pm} rowspan\{\boldsymbol{B}_{\{1,4\}}, \mathbf{I}_{\mathcal{A}_2}\}$.

One may notice that the above outer bound derivation has common properties as that in [51], the differences however are two-fold: 1) due to transmitter cooperation, the transmitted signal is encoded from multiple messages, instead of the single message in the TIM setting, and 2) when we

switch off some messages (e.g., by setting them to be deterministic), we only eliminate them from the message set $\mathcal{R}_i$ of X_i^n, instead of switching off X_i^n as did in [51].

6.4.2 Example 2

Let us consider another example of a $(5,3)$-regular network as shown in Fig. 6.2. By enabling transmitter cooperation, the achievable symmetric DoF are improved from $\frac{2}{5}$ (as reported in [50]) to $\frac{1}{2}$ according to Theorem 6.3. In what follows, we will show an interference alignment scheme to achieve this and an outer bound with $d_{\text{sym}} \leq \frac{5}{8}$ according to Theorem 6.4.

Inner Bound via Interference Alignment

According to the network topology, we have transmit and receive sets $\mathcal{T}_1 = \mathcal{R}_1 = \{1,3,4\}, \mathcal{T}_2 = \mathcal{R}_2 = \{2,4,5\}, \mathcal{T}_3 = \mathcal{R}_3 = \{1,3,5\}, \mathcal{T}_4 = \mathcal{R}_4 = \{1,2,4\}, \mathcal{T}_5 = \mathcal{R}_5 = \{2,3,5\}$. For notational convenience, we denote by a, b, c, d, e the messages desired by five receivers, with the subscript distinguishing different symbols for the same receiver. We consider a multiple time-slotted protocol, in which a space is spanned such that the symbols will be sent in certain subspaces. Given five random vectors $\boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{V}_3, \boldsymbol{V}_4, \boldsymbol{V}_5 \in \mathbb{C}^{4\times 1}$, any four of which are linearly independent, the transmitters send signals with precoding

$$\boldsymbol{X}_1 = \boldsymbol{V}_1 c_1 + \boldsymbol{V}_3 d_1, \quad \boldsymbol{X}_2 = \boldsymbol{V}_2 d_2 + \boldsymbol{V}_4 e_1 \tag{6.48}$$

$$\boldsymbol{X}_3 = \boldsymbol{V}_5 a_1 + \boldsymbol{V}_3 e_2, \quad \boldsymbol{X}_4 = \boldsymbol{V}_4 a_2 + \boldsymbol{V}_1 b_2 \tag{6.49}$$

$$\boldsymbol{X}_5 = \boldsymbol{V}_5 b_1 + \boldsymbol{V}_2 c_2 \tag{6.50}$$

within four time slots, where $\boldsymbol{X}_i \in \mathbb{C}^{4\times 1}$ is the vector of the concatenated transmit signals from Transmitter i, with each element being the transmitted signal at each corresponding time slot.

We assume the coherence time $\tau_c \geq 4$, during which the channel coefficients keep constant. The received signal at Receiver 1 for example within four time slots, with $\mathcal{T}_1 = \{1,3,4\}$, can be written as

$$\boldsymbol{Y}_1 = \sum_{i\in\mathcal{T}_1} h_{1i}\boldsymbol{X}_i + \boldsymbol{Z}_1 \tag{6.51}$$

$$= h_{11}\boldsymbol{X}_1 + h_{13}\boldsymbol{X}_3 + h_{14}\boldsymbol{X}_4 + \boldsymbol{Z}_1 \tag{6.52}$$

$$= \underbrace{h_{13}\boldsymbol{V}_5 a_1 + h_{14}\boldsymbol{V}_4 a_2}_{\text{desired signal}} + \underbrace{\boldsymbol{V}_1(h_{11}c_1 + h_{14}b_2) + \boldsymbol{V}_3(h_{11}d_1 + h_{13}e_2)}_{\text{aligned interferences}} + \boldsymbol{Z}_1. \tag{6.53}$$

Recall that $\{\boldsymbol{V}_i, i = 1, \ldots, 5\}$ are 4×1 linearly independent vectors spanning four-dimensional space, by which it follows that the interferences are aligned in the two-dimensional subspace spanned by $\boldsymbol{V}_1$ and $\boldsymbol{V}_3$, leaving two-dimensional interference-free subspace spanned by $\boldsymbol{V}_4$ and $\boldsymbol{V}_5$ to the desired symbols a_1, a_2. Hence, the desired messages of Receiver 1 can be successfully recovered, almost surely. In doing so, all receivers can decode two messages within four time slots, yielding the symmetric DoF of $\frac{1}{2}$.

To illustrate the interference alignment, we describe the transmitted signals geometrically as shown in Fig. 6.2. In this figure, we depict the subspace spanned by $\{\boldsymbol{V}_i, i = 1, \ldots, 5\}$ as a four-dimensional space, where any four of them suffice to represent this space. We also denote by $X_i(W_j)$ the message W_j sent from Transmitter i. Let us still take Receiver 1 for example. Because of $\mathcal{T}_1 = \{1, 3, 4\}$, the transmitted signals from the transmitters that do not belong to $\mathcal{T}_1$ will not reach Receiver 1, and hence the vector $\boldsymbol{V}_2$ is absent to Receiver 1. In addition, we have the interference-free signals in the directions of $\boldsymbol{V}_4$ and $\boldsymbol{V}_5$, and the aligned interferences carrying messages other than a_1, a_2 in the subspace spanned by $\boldsymbol{V}_1$ and $\boldsymbol{V}_3$. Recall that vectors $\{\boldsymbol{V}_1, \boldsymbol{V}_3, \boldsymbol{V}_4, \boldsymbol{V}_5\}$ are linearly independent, almost surely, so that the interference alignment is feasible at Receiver 1, and it can also be checked to be feasible at other receivers.

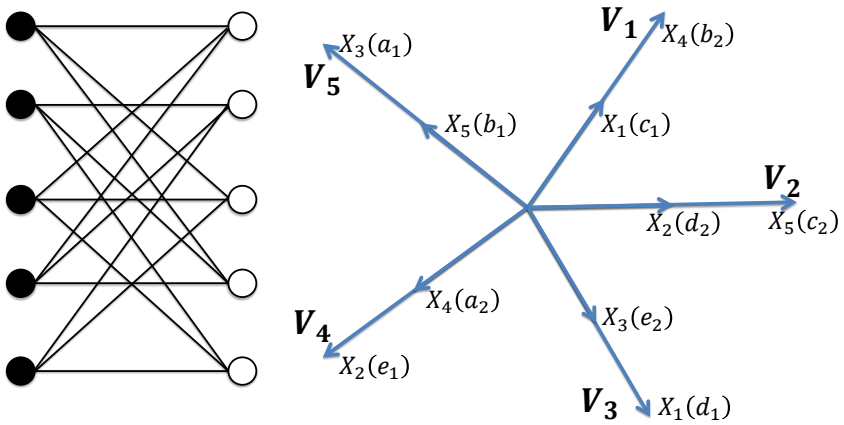


Figure 6.2: A $(5, 3)$-regular cellular network. With the interference alignment scheme illustrated on the right, a four-dimensional space is sufficient to deliver every message twice free of interference.

Outer Bound via Compound Settings

In what follows, we derive an outer bound with compound settings [50]. A more general version will be presented in Section 6.5.6. By Fano's inequality, we have

$$\begin{aligned} n(R_1 - \epsilon_n) &\leq I(W_1, Y_1^n | \mathcal{H}^n, \mathcal{G}) && (6.54)\\ &= h(Y_1^n | \mathcal{H}^n, \mathcal{G}) - h(Y_1^n | W_1, \mathcal{H}^n, \mathcal{G}) && (6.55)\\ &\leq n \log P - h(Y_1^n | W_1, \mathcal{H}^n, \mathcal{G}) + n \cdot O(1). && (6.56) \end{aligned}$$

Assuming there are two compound receivers demanding the same message W_1, we have two compound signals Y_1', Y_1'', which are also the linear combinations of X_1, X_3, X_4 as Y_1, yet with independent channel coefficients. Thus, these three received signals are linearly independent with regard to X_1, X_3, X_4, almost surely, and are statistically equivalent, which results in the same achievable rate R_1. Similarly, we have

$$\begin{aligned} n(R_1 - \epsilon_n) &\leq n \log P - h(Y_1'^n | W_1, \mathcal{H}^n, \mathcal{G}) + n \cdot O(1) && (6.57)\\ n(R_1 - \epsilon_n) &\leq n \log P - h(Y_1''^n | W_1, \mathcal{H}^n, \mathcal{G}) + n \cdot O(1). && (6.58) \end{aligned}$$

For Receiver 2, we consider the statistically equivalent received signals Y_2 by itself and Y_2' by a compound receiver, and have

$$\begin{aligned} n(R_2 - \epsilon_n) &\leq n \log P - h(Y_2^n | W_2, \mathcal{H}^n, \mathcal{G}) + n \cdot O(1) && (6.59)\\ n(R_2 - \epsilon_n) &\leq n \log P - h(Y_2'^n | W_2, \mathcal{H}^n, \mathcal{G}) + n \cdot O(1). && (6.60) \end{aligned}$$

Combining all above inequalities, we have

$$\begin{aligned} & n(3R_1 + 2R_2 - \epsilon_n) && (6.61)\\ & \quad \leq 5n \log P - h(Y_1^n, Y_1'^n, Y_1''^n, Y_2^n, Y_2'^n | W_1, W_2, \mathcal{H}^n, \mathcal{G}) + n \cdot O(1) && (6.62)\\ & \quad = 5n \log P - h(\{X_i^n + \bar{Z}_i^n, i = 1, \ldots, 5\} | W_1, W_2, \mathcal{H}^n, \mathcal{G}) + n \cdot O(1) && (6.63)\\ & \quad \leq 5n \log P - n(R_3 + R_4 + R_5) + n \cdot O(1) && (6.64) \end{aligned}$$

where $Y_1, Y_1', Y_1'', Y_2, Y_2'$ are linearly independent with regard to $\{X_i, i = 1, 2, 3, 4, 5\}$, by which the noisy versions of $\{X_i, i = 1, 2, 3, 4, 5\}$, i.e., $X_i^n + \bar{Z}_i^n$ with $\bar{Z}_i$ being bounded noise term, can be recovered, almost surely; the last inequality due to the fact that the decode of W_3, W_4, W_5 is possible only if

$$n(R_3 + R_4 + R_5) \leq I(W_3, W_4, W_5; \{X_i^n + \bar{Z}_i^n, i = 1, \ldots, 5\} | W_1, W_2, \mathcal{H}^n, \mathcal{G}) \quad (6.65)$$

$$= h(\{X_i^n + \bar{Z}_i^n, i = 1, \ldots, 5\}|W_1, W_2, \mathcal{H}^n, \mathcal{G}) + n \cdot O(1). \tag{6.66}$$

By now, according to the definition of symmetric DoF, it follows that

$$d_{\text{sym}} \leq \frac{5}{8}. \tag{6.67}$$

In contrast, generator sequence offers the best possible outer bound $d_{\text{sym}} \leq \frac{4}{5}$, which is looser. This confirms again that compound setting bounding is more suitable to regular networks.

6.5 General Proofs

6.5.1 Proof of Theorem 6.1

To prove this achievability, we first build a connection between interference avoidance of TIM-CoMP problems and link scheduling problems, and then solve the link scheduling problems through graph coloring.

Enabling transmitter cooperation, it requires to schedule links rather than transmitters to avoid mutual interference. Without transmitter cooperation, the message W_j can only be sent from Transmitter j for all j, whose activation will cause interferences to Receiver k ($k \in \mathcal{R}_j$), and consequently inactivate Transmitter k ($k \in \mathcal{R}_j$), because W_k cannot be delivered from Transmitter k to Receiver k free of interference. The interference avoidance in this case is a matter of activating or inactivating transmitters. In contrast, with transmitter cooperation (i.e., message sharing), the message W_j can be sent from any Transmitter i with $i \in \mathcal{T}_j$, and thus, it is not sufficient to schedule transmitters only. In fact, the link - rather than the transmitter - scheduling is of interest, because both the scheduling of the transmitters and the receivers does matter.[2] For instance, if the link e_{ij} (i.e., from Transmitter i to Receiver j) is scheduled, the links adjacent to e_{ij} (i.e., e_{ik_1} and e_{k_2j} with $k_1 \in \mathcal{T}_i \backslash j$ and $k_2 \in \mathcal{R}_j \backslash i$) as well as the links adjacent to e_{ik_1} and e_{k_2j} should not be scheduled, because activating Transmitter k_2 will interfere Receiver j and Receiver k_1 will overhear interferences from Transmitter i, such that any delivery from Transmitter k_2 or to Receiver k_1 causes mutual interferences.

Such a link scheduling problem is usually solved through graph edge-coloring, while the nature of our problem calls for a more specific graph

[2] In fact, transmitter scheduling can also be regarded as link scheduling, yet only the direct links (i.e., the links from Transmitter j to Receiver j) are candidates of link scheduling.

coloring solution. Let us represent the cellular network as a bipartite graph $\mathcal{G} = (\mathcal{U}, \mathcal{V}, \mathcal{E})$, and assign the links that are scheduled at different time slots with distinct colors. Suppose the edge $e_{ij} \in \mathcal{E}$ receives a color. Analogously, the edges e_{ik_1} and $e_{k_2 j}$ with $k_1 \in \mathcal{T}_i \backslash j$ and $k_2 \in \mathcal{R}_j \backslash i$ should not be assigned the same color. Moreover, the edges adjacent to e_{ik_1} and $e_{k_2 j}$ should not receive the same color either. In a word, the edges within two-hop should be assigned with distinct colors. In addition, as we aim at symmetric DoF, the total number of scheduled times of the links connecting a common receiver is of interest. Thus, the number of colors should be counted by clusters where the edges with the common vertex $v_j \in \mathcal{V}$ are grouped together to form a cluster.

To ease our presentation, we further transfer the edge-coloring of $\mathcal{G}$ to vertex-coloring of its line graph $\mathcal{G}_e$. Accordingly, we group the vertices in $\mathcal{G}_e$ for which the corresponding edges in $\mathcal{G}$ have a vertex $v_j \in \mathcal{V}$ in common as a cluster, such that the number of colors is counted by clusters in $\mathcal{G}_e$. The above two-hop condition is therefore transferred to a distance-2 constraint, where two vertices in $\mathcal{G}_e$ with distance no more than 2 should receive different colors. Thus, the above link scheduling problem is transferable to a distance-2 selective graph vertex coloring problem of its line graph $\mathcal{G}_e$, and thus to a selective graph vertex coloring problem of the square of its line graph, i.e., $\mathcal{G}_e^2$, in which the vertices are clustered into $\mathbb{V}_e = \{\mathcal{V}_1, \ldots, \mathcal{V}_K\}$ with $\mathcal{V}_k = \{e_{jk},\ j \in \mathcal{T}_k\}$. Specifically, a proper selective coloring of $\mathcal{G}_e^2$ over $\mathbb{V}_e$ is a proper color assignment such that each cluster $\mathcal{V}_i$ receives m colors out of in total n colors. As such, $\mathcal{G}_e^2$ is selectively $n:m$ colorable over $\mathbb{V}_e$, indicating that the links in each cluster can be scheduled m times within overall n time slots without causing mutual interference. Consequently, according to Definition 6.4, an achievability of symmetric DoF can be given by

$$d_{\text{sym}} = \sup_m \frac{m}{s\chi_m(\mathcal{G}_e^2, \mathbb{V}_e)} = \frac{1}{s\chi_f(\mathcal{G}_e^2, \mathbb{V}_e)} \tag{6.68}$$

where $s\chi_f$ is the fractional selective chromatic number as in Definition 6.4.

6.5.2 Proof of Theorem 6.2

According to the definition of symmetric DoF, the outer bound of symmetric DoF obtained for any subset of receivers should serve as the outer bound in general. In other words, the general outer bound is the minimum value of all possible outer bounds for any subset of receivers.

Let us take a subset of receivers $\mathcal{S} \subseteq \mathcal{K}$ with received signals $Y_{\mathcal{S}}$ into account. For those receivers who are not considered, we switch off their

desired messages from the transmitted signal, i.e., the constituent messages in transmitted signal X_i^n is now comprised of message W_j where $j \in \mathcal{R}_i \backslash \mathcal{S}^c$. Define $\tilde{\boldsymbol{X}}^\intercal \triangleq \begin{bmatrix} h_1 X_1 & \dots & h_K X_K \end{bmatrix}$, where h_i $(i \in \mathcal{K})$ is independent and identically distributed as the nonzero h_{ji}, and a set of virtual signals in the compact form

$$\tilde{\boldsymbol{Y}}_{\mathcal{I}} \triangleq \boldsymbol{B}_{\mathcal{I}} \tilde{\boldsymbol{X}} + \tilde{\boldsymbol{Z}}_{\mathcal{I}} \tag{6.69}$$

$$\bar{\boldsymbol{Y}}_{\mathcal{I}} \triangleq \boldsymbol{B}_{\mathcal{I}} \mathbf{I}^{\pm} \tilde{\boldsymbol{X}} + \tilde{\boldsymbol{Z}}_{\mathcal{I}} \tag{6.70}$$

for a set of receivers in $\mathcal{I}$, where $\boldsymbol{B}_{\mathcal{I}}$ is the submatrix of $\boldsymbol{B}$ with the rows out of $\mathcal{I}$ removed, $\mathbf{I}^{\pm}$ is the same as the identity matrix up to the sign of elements, and $\tilde{\boldsymbol{Y}}_{\mathcal{I}}$, $\bar{\boldsymbol{Y}}_{\mathcal{I}}$, $\tilde{\boldsymbol{Z}}_{\mathcal{I}}$ are vectors compacted by $\tilde{Y}_{\mathcal{I}}$, $\bar{Y}_{\mathcal{I}}$, and $\tilde{Z}_{\mathcal{I}}$, respectively. Note that $\tilde{Y}_{\mathcal{I}}$ and $\bar{Y}_{\mathcal{I}}$ are statistically equivalent to $Y_{\mathcal{I}}$, because the distribution of channel gain is symmetric around zero. We assume there exists a generator sequence $\{\mathcal{I}_0, \mathcal{I}_1, \dots, \mathcal{I}_S\}$ with $\cup_{s=0}^{S} \mathcal{I}_s = \mathcal{S}$ and $\mathcal{I}_i \cap \mathcal{I}_j = \emptyset \ \forall\ i \neq j$, such that

$$\boldsymbol{B}_{\mathcal{I}_s} \subseteq^{\pm} rowspan\left\{\boldsymbol{B}_{\mathcal{I}_0}, \mathbf{I}_{\mathcal{A}_s}\right\}, \quad \forall\ s = 1, \dots, S. \tag{6.71}$$

This implies that there exist $\boldsymbol{C}_s \in \mathbb{C}^{|\mathcal{I}_s| \times |\mathcal{I}_0|}$ and $\boldsymbol{D}_s \in \mathbb{C}^{|\mathcal{I}_s| \times |\mathcal{A}_s|}$, such that

$$\boldsymbol{B}_{\mathcal{I}_s} = (\boldsymbol{C}_s \boldsymbol{B}_{\mathcal{I}_0} + \boldsymbol{D}_s \mathbf{I}_{\mathcal{A}_s}) \mathbf{I}^{\pm}. \tag{6.72}$$

Multiplying $\mathbf{I}^{\pm} \tilde{\boldsymbol{X}}$ at both sides yields

$$\boldsymbol{B}_{\mathcal{I}_s} \mathbf{I}^{\pm} \tilde{\boldsymbol{X}} = \boldsymbol{C}_s \boldsymbol{B}_{\mathcal{I}_0} \tilde{\boldsymbol{X}} + \boldsymbol{D}_s \mathbf{I}_{\mathcal{A}_s} \tilde{\boldsymbol{X}} \tag{6.73}$$

$$\Rightarrow \bar{\boldsymbol{Y}}_{\mathcal{I}_s} = \boldsymbol{C}_s \tilde{\boldsymbol{Y}}_{\mathcal{I}_0} + \boldsymbol{D}_s \mathbf{I}_{\mathcal{A}_s} \tilde{\boldsymbol{X}} + \tilde{\boldsymbol{Z}}_{\mathcal{I}_s} - \boldsymbol{C}_s \tilde{\boldsymbol{Z}}_{\mathcal{I}_0} \tag{6.74}$$

$$= \boldsymbol{C}_s \tilde{\boldsymbol{Y}}_{\mathcal{I}_0} + \boldsymbol{D}_s \tilde{\boldsymbol{X}}_{\mathcal{A}_s} + \tilde{\boldsymbol{Z}}_{\mathcal{I}_s} - \boldsymbol{C}_s \tilde{\boldsymbol{Z}}_{\mathcal{I}_0} \tag{6.75}$$

$$= \boldsymbol{C}_s \tilde{\boldsymbol{Y}}_{\mathcal{I}_0} + \boldsymbol{D}_s \tilde{\boldsymbol{X}}_{\mathcal{A}_s} - \bar{\boldsymbol{Z}}_s \tag{6.76}$$

with $\bar{\boldsymbol{Z}}_s \triangleq \boldsymbol{C}_s \boldsymbol{Z}_{\mathcal{I}_0} - \boldsymbol{Z}_{\mathcal{I}_s}$ being the entropy-bounded noise term [51]. Thus, according to the mapping $\mathbf{f}_{\text{idx}} : \mathcal{B} \mapsto \{0,1\}^K$ and the definition of $\mathcal{A}_s$, we have

$$H(W_{\mathcal{I}_s} | \tilde{\boldsymbol{Y}}_{\mathcal{I}_0}^n, \cup_{r=0}^{s-1} W_{\mathcal{I}_r}, \mathcal{H}^n, \mathcal{G}) = H(W_{\mathcal{I}_s} | \tilde{\boldsymbol{Y}}_{\mathcal{I}_0}^n, \cup_{r=0}^{s-1} W_{\mathcal{I}_r}, X_{\mathcal{A}_s}, \mathcal{H}^n, \mathcal{G}) \tag{6.77}$$

$$= H(W_{\mathcal{I}_s} | \tilde{\boldsymbol{Y}}_{\mathcal{I}_0}^n, \bar{\boldsymbol{Y}}_{\mathcal{I}_s}^n + \bar{\boldsymbol{Z}}_s^n, \cup_{r=0}^{s-1} W_{\mathcal{I}_r}, X_{\mathcal{A}_s}, \mathcal{H}^n, \mathcal{G}) \tag{6.78}$$

$$\leq H(W_{\mathcal{I}_s} | \bar{\boldsymbol{Y}}_{\mathcal{I}_s}^n + \bar{\boldsymbol{Z}}_s^n, \mathcal{H}^n, \mathcal{G}) \tag{6.79}$$

$$= H(W_{\mathcal{I}_s} | \tilde{\boldsymbol{Y}}_{\mathcal{I}_s}^n + \bar{\boldsymbol{Z}}_s^n, \mathcal{H}^n, \mathcal{G}) \tag{6.80}$$

$$\leq n\epsilon_n + n \cdot O(1) \tag{6.81}$$

where (6.77) is due to the fact that X_i^n is encoded only from $W_{\mathcal{R}_i \setminus \mathcal{S}^c}$, (6.78) comes from (6.76) where $\tilde{\boldsymbol{X}}_{\mathcal{A}_s}$ can be constructed from $X_{\mathcal{A}_s}$, (6.79) is because removing conditioning does not reduce entropy, (6.80) is due to the argument that $\tilde{Y}$ and $\bar{Y}$ are statistically equivalent, and the last inequality is obtained by following the fact that $H(W|Y^n+\bar{Z}^n) \leq n\epsilon_n + n\cdot O(1)$, if $H(W|Y^n+Z^n) \leq n\epsilon_n$ [51], since $\bar{\boldsymbol{Z}}_s$ is bounded noise term. Further, we have

$$
\begin{aligned}
n\sum_{i\in\mathcal{S}} R_i &= H(W_{\mathcal{S}}|\mathcal{H}^n, \mathcal{G}) && (6.82)\\
&= I(W_{\mathcal{S}}; \tilde{\boldsymbol{Y}}_{\mathcal{I}_0}^n|\mathcal{H}^n, \mathcal{G}) + H(W_{\mathcal{S}}|\tilde{\boldsymbol{Y}}_{\mathcal{I}_0}^n, \mathcal{H}^n, \mathcal{G}) && (6.83)\\
&= I(W_{\mathcal{S}}; \tilde{\boldsymbol{Y}}_{\mathcal{I}_0}^n|\mathcal{H}^n, \mathcal{G}) + H(W_{\mathcal{I}_0}|\tilde{\boldsymbol{Y}}_{\mathcal{I}_0}^n, \mathcal{H}^n, \mathcal{G}) + H(W_{\mathcal{S}\setminus\mathcal{I}_0}|\tilde{\boldsymbol{Y}}_{\mathcal{I}_0}^n, W_{\mathcal{I}_0}, \mathcal{H}^n, \mathcal{G}) && (6.84)\\
&\leq n|\mathcal{I}_0|\log P + n\cdot O(1) + n\epsilon_n + \sum_{s=1}^{S} H(W_{\mathcal{I}_s}|\tilde{\boldsymbol{Y}}_{\mathcal{I}_0}^n, \cup_{r=0}^{s-1} W_{\mathcal{I}_r}, \mathcal{H}^n, \mathcal{G}) && (6.85)\\
&\leq n|\mathcal{I}_0|\log P + n\cdot O(1) + n\epsilon_n. && (6.86)
\end{aligned}
$$

By the definition of symmetric DoF, we have

$$
d_{\text{sym}} \leq \lim_{P\to\infty} \frac{R_{\text{sym}}}{\log P} = \frac{|\mathcal{I}_0|}{|\mathcal{S}|}. \tag{6.87}
$$

Among all possible subsets of $\mathcal{S}$ and initial generator $\mathcal{I}_0$, the symmetric DoF should be outer-bounded by the minimum of them. Thus, we have

$$
d_{\text{sym}} \leq \min_{\mathcal{S}\subseteq\mathcal{K}} \min_{\mathcal{I}_0\subseteq\mathcal{J}(\mathcal{S})} \frac{|\mathcal{I}_0|}{|\mathcal{S}|}. \tag{6.88}
$$

6.5.3 Proof of Corollary 6.1

Enumerating all the possible topologies of three-cell networks, we verify the optimality of symmetric DoF by comparing the inner and outer bounds according to Theorem 6.1 and Theorem 6.2. It is readily verified that all but two topologies have enhanced symmetric DoF, compared to the case without transmitter cooperation [49, 51, 129]. As shown in Fig. 6.3, message sharing improves the symmetric DoF from $\frac{1}{2}$ to $\frac{2}{3}$ for the topology (i) and from $\frac{1}{3}$ to $\frac{1}{2}$ for the topology (m).

For the achievability, two graph coloring realizations are illustrated in Fig. 6.4 concerning the topologies of (i) and (m). Specifically, every cluster receives two out of three colors in total in (i), and one out of two colors in

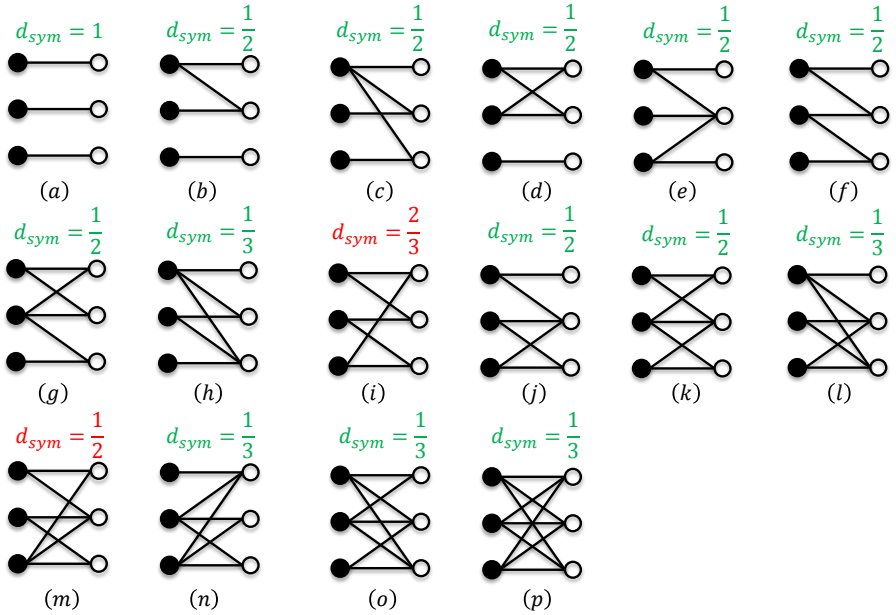


Figure 6.3: The three-cell TIM-CoMP problem, where all non-isomorphic topologies are enumerated. The symmetric DoF improvement over the noncooperation case is due to topologies (i) and (m).

(m), where the conditions of distance-2 fractional selective graph coloring are satisfied, yielding symmetric DoF inner bounds $d_{\text{sym}} = \frac{2}{3}$ and $d_{\text{sym}} = \frac{1}{2}$, respectively. For other topologies, the achievability can be similarly obtained.

Regarding the converse, we apply the outer bound via generator sequence here. Again, we take those two topologies for example. For topology-(i), we have a generator sequence $\{\{1,2\},\{3\}\}$ with $\mathcal{I}_0 = \{1,2\}$ and $\mathcal{I}_1 = \{3\}$. By generating the virtual signals $\tilde{Y}_1^n = h_1^n X_1^n + h_3^n X_3^n + \tilde{Z}_1^n$ and $\tilde{Y}_2^n = h_1^n X_1^n + h_2^n X_2^n + \tilde{Z}_2^n$, which are statistically equivalent to Y_1^n and Y_2^n respectively, we obtain $\tilde{Y}_3^n = \tilde{Y}_1^n - \tilde{Y}_2^n = h_3^n X_3^n - h_2^n X_2^n + \tilde{Z}_1^n - \tilde{Z}_2^n$ that is statistically equivalent to Y_3^n with a bounded noise difference [51]. Thus, according to Theorem 6.2, we have $d_{\text{sym}} \leq \frac{|\mathcal{I}_0|}{|\mathcal{S}|} = \frac{2}{3}$. Similarly for topology-(m), we have a generator sequence $\{\{2\},\{1\}\}$ with $\mathcal{I}_0 = \{2\}$ and $\mathcal{I}_1 = \{1\}$. Note that we ignore the received signal at Receiver 3, and therefore eliminate the message W_3 from the message sets of the respective transmitted signals. Thus, the message sets of Transmitters 1, 2, and 3 become $\{W_1, W_2\}, \{W_2\}$, and $\{W_1, W_2\}$, respectively. Following the generator sequence approach, we initiate the

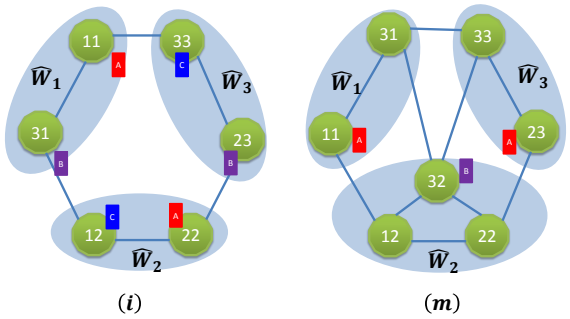


Figure 6.4: Fractional selective graph coloring of the topologies (i) and (m). It requires three colors to ensure every cluster receive two in (i), and two colors are sufficient to offer every cluster one color in (m).

generator sequence by a virtual signal $\tilde{Y}_2^n = h_1^n X_1^n + h_2^n X_2^n + h_3^n X_3^n + \tilde{Z}_2^n$, and successively generate $\tilde{Y}_1^n = \tilde{Y}_2^n - h_2^n X_2^n$, where X_2 can be encoded from the message W_2. Hence, the symmetric DoF outer bound is $d_{\text{sym}} \leq \frac{|\mathcal{I}_0|}{|\mathcal{S}|} = \frac{1}{2}$.

Aware of the coincidence of the inner and outer bounds, we conclude that the interference avoidance achieves the optimal symmetric DoF. The optimality verification of other topologies can be similarly done.

6.5.4 Proof of Corollary 6.2

For the converse proof, since the lower and upper triangular matrices are similar, it suffices to consider the lower triangular matrix $\boldsymbol{B}$ without loss of generality, where $\mathcal{T}_j = \{1, \ldots, j\}$ for all $j \in \mathcal{K}$. Thus, the message sets to X_j with transmitter cooperation are comprised of $W_{\{j,\cdots,K\}}$. It is readily verified that $\{\{K\}, \{K-1\}, \ldots, \{1\}\}$ forms a generator sequence with $\mathcal{I}_0 = \{K\}$ and $\mathcal{S} = \mathcal{K}$. Thus, we have the outer bound $d_{\text{sym}} \leq \frac{|\mathcal{I}_0|}{|\mathcal{S}|} = \frac{1}{K}$, which is achievable by time division. This completes the proof.

6.5.5 Proof of Theorem 6.3

According to the definition of (K, d)-regular networks, we have $|\mathcal{T}_j| = d$, $\forall\ j \in \mathcal{K}$. As we know, when $d = K$, the network is fully connected and therefore the optimal symmetric DoF value is $\frac{1}{K}$ by time division. So, in what follows, we will consider the general achievability proof when $d \leq K-1$.

Since the cellular network graph is assumed to be similar to the reference one by reordering the transmitters and/or receivers, we directly consider the referred network topology, because they are equivalent in terms of symmetric

DoF with transmitter cooperation. For the referred network topology, the transmit set of Receiver j is given by

$$\mathcal{T}_j = \{j, j+1, \ldots, j+d-1\}, \tag{6.89}$$

where all the receiver indices are modulo K, e.g., $j - K = j$ and $0 = K$. Thus, at Transmitter i we send symbols with careful design

$$\boldsymbol{X}_i = \boldsymbol{V}_{i+1} X_i(W_i^1) + \boldsymbol{V}_{i+2} X_i(W_{i-d+1}^2), \forall\ i = 1, \ldots, K$$

where $\{\boldsymbol{V}_i,\ i = 1, \ldots, K\}$ are $(d+1) \times 1$ random vectors, and linearly independent among any $(d+1)$ vectors, almost surely, $X_i(W_j)$ is the signal transmitted from Transmitter i carrying on message W_j, and W_j^1, W_j^2 are two realizations (symbols) of message W_j. The signals at Receiver j during $d+1$ time slots, with coherence time $\tau_c \geq d+1$, can be compacted as

$$\begin{aligned}
\boldsymbol{Y}_j &= \sum_{i \in \mathcal{T}_j} h_{ji} \boldsymbol{X}_i + \boldsymbol{Z}_j \\
&= \sum_{i=j}^{j+d-1} h_{ji} (\boldsymbol{V}_{i+1} X_i(W_i^1) + \boldsymbol{V}_{i+2} X_i(W_{i-d+1}^2)) + \boldsymbol{Z}_j \\
&= h_{j,j} \boldsymbol{V}_{j+1} X_j(W_j^1) + h_{j,j+d-1} \boldsymbol{V}_{j+d+1} X_{j+d-1}(W_j^2) \\
&\qquad + \sum_{i=j+1}^{j+d-1} h_{ji} \boldsymbol{V}_{i+1} X_i(W_i^1) + \sum_{i=j}^{j+d-2} h_{ji} \boldsymbol{V}_{i+2} X_i(W_{i-d+1}^2) + \boldsymbol{Z}_j \\
&= \underbrace{h_{j,j} \boldsymbol{V}_{j+1} X_j(W_j^1) + h_{j,j+d-1} \boldsymbol{V}_{j+d+1} X_{j+d-1}(W_j^2)}_{\text{desired signal}} \\
&\qquad + \underbrace{\sum_{i=j}^{j+d-2} \boldsymbol{V}_{i+2} (h_{j,i+1} X_{i+1}(W_{i+1}^1) + h_{j,i} X_i(W_{i-d+1}^2))}_{\text{aligned interferences}} + \boldsymbol{Z}_j.
\end{aligned}$$

It is readily shown that the interferences occupy $d-1$ dimensional subspace out of the total $d+1$ dimensional space, leaving 2-dimensional interference-free subspace spanned by $\{\boldsymbol{V}_{j+1}, \boldsymbol{V}_{j+d+1}\}$ to the desired signals, such that the desired messages for Receiver j, W_j^1 and W_j^2, can be successfully recovered. This philosophy applies to all other receivers. During $d+1$ time slots, every receiver can decode two messages, yielding symmetric DoF of $\frac{2}{d+1}$.

Geometrically, the interference alignment can be shown in Fig. 6.5, and also interpreted as follows. Transmitted signals $X_{j-1}(W_{j-1}^1)$ and $X_{j-2}(W_{j-d-1}^2)$

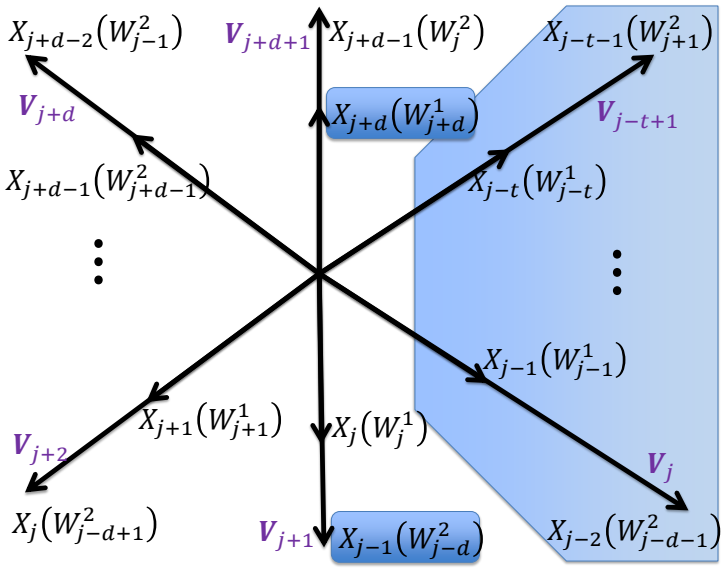


Figure 6.5: Interference alignment for the general (K, d) regular cellular networks.

are aligned in the same subspace spanned by vector $\boldsymbol{V}_j$, which is absent to Receiver k ($k \in \{j, \ldots, j+K-3\}$). Note that $t \triangleq K-d-1$ and $j-t = j+d+1$ modulo K. By deduction, the subspace spanned by $\{\boldsymbol{V}_{j+d+2}, \ldots, \boldsymbol{V}_j\}$ are absent to Receiver j (i.e., the shadow in Fig. 6.5), leaving $d+1$ linearly independent vectors $\{\boldsymbol{V}_{j+1}, \ldots, \boldsymbol{V}_{j+d+1}\}$ to span the space. As such, the signal carrying $X_j(W_j^1)$ is aligned with $X_{j-1}(W_{j-d}^2)$ in the subspace spanned by $\boldsymbol{V}_{j+1}$, and $X_{j+d-1}(W_j^2)$ is aligned with $X_{j+d}(W_{j+d}^1)$ in the subspace spanned by $\boldsymbol{V}_{j+d+1}$. Note that the signals from Transmitter $j-1$ and $j+d$ cannot be heard by Receiver j according to the network topology, such that W_j^1 and W_j^2 are free of interference, and retrievable from overall $d+1$ dimensional subspace.

It is worth noting that, although the message W_j is shared among the transmitters i ($\forall\, i \in \mathcal{T}_j$), its two realizations W_j^1 and W_j^2 are only utilized in this scheme by Transmitter j and Transmitter $(j-d+1)$, respectively.

6.5.6 Proof of Theorem 6.4

In what follows, we present an outer bound with the aid of compound settings. As illustrated in Example 2, it is necessary to determine the least required compound receivers such that the noisy versions of X_i can be recovered. Thus, we first look into this problem, given that a subset of messages is

known *a priori*.

Consider a set of receivers $\mathcal{S} \subseteq \mathcal{K}$ satisfying $\cup_{j\in\mathcal{S}}\mathcal{T}_j = \mathcal{K}$. The received signals Y_j ($j \in \mathcal{S}$) at Receiver j is a linear combination of $\{X_i, i \in \mathcal{T}_j\}$ polluted by noise. To recover the noisy versions of $\{X_i, i \in \mathcal{K}\}$, it requires at least $K - |\mathcal{S}|$ extra linear independent equations, which can be provided by compound receivers that are assume to be possessing the same topology as the original receivers and demanding the same messages. In the rest of the proof, we do not distinguish the original from the compound receivers explicitly.

In fact, in the present of a set of messages $W_{\mathcal{S}}$, the required number of compound receivers can be further reduced. According to transmitter cooperation, the transmitted signal X_i^n is encoded with the messages $W_{\mathcal{R}_i}$. Thus, we have

$$\mathcal{S}' = \{i | [\boldsymbol{B}^{\mathsf{T}}]_i \cdot \mathbf{f}_{\text{idx}}(\mathcal{S}) = |\mathcal{R}_i|\}, \tag{6.90}$$

which implies that the knowledge of $W_{\mathcal{S}}$ is equivalent to the knowledge of $X_{\mathcal{S}'}^n$. Knowing $X_{\mathcal{S}'}^n$, we can remove their contributions from the received signals. Denote by $Y_{j,i}$ and $\tilde{Y}_{j,i}$ the received signals of compound receiver i before and after removing the contribution of $X_{\mathcal{S}'}$, respectively, i.e.,

$$Y_{j,i} = \sum_{k\in\mathcal{T}} h_{j,i,k} X_k + Z_{j,i} \tag{6.91}$$

$$\tilde{Y}_{j,i} = \sum_{k\in\mathcal{T}\setminus\mathcal{S}'} h_{j,i,k} X_k + Z_{j,i}. \tag{6.92}$$

Let $\mathcal{T}_j'$ be the set of the least required compound receivers associated with Receiver j. Thus, we collect all the compound signals and compact them as

$$\tilde{\boldsymbol{Y}}_{\mathcal{T}_{\mathcal{S}}'} = \boldsymbol{H}_{\mathcal{T}_{\mathcal{S}}'} \boldsymbol{X}_{\mathcal{K}\setminus\mathcal{S}'} + \boldsymbol{Z}_{\mathcal{T}_{\mathcal{S}}'} \tag{6.93}$$

where $\boldsymbol{H}_{\mathcal{T}_{\mathcal{S}}'} \in \mathbb{C}^{\sum_{j\in\mathcal{S}}|\mathcal{T}_j'|\times(K-|\mathcal{S}'|)}$ is the reduced channel matrix with the columns indexed by $\mathcal{S}'$ removed. It suffices to recover $X_{\mathcal{K}\setminus\mathcal{S}'}$ from $\tilde{\boldsymbol{Y}}_{\mathcal{T}_{\mathcal{S}}'}$ as long as $\sum_{j\in\mathcal{S}}|\mathcal{T}_j'| \geq K - |\mathcal{S}'|$. We conclude that the required number of compound receivers can be reduced to $K - |\mathcal{S}'| - |\mathcal{S}|$, given the knowledge of $W_{\mathcal{S}}$.

Secondly, we proceed to present the outer bound of achievable rates of compound receivers. For the compound receiver i, by Fano's inequality, we have

$$n(R_{j,i} - \epsilon_n) \leq I(W_j, Y_{j,i}^n | \mathcal{H}^n, \mathcal{G}) \tag{6.94}$$

$$= h(Y_{j,i}^n|\mathcal{H}^n,\mathcal{G}) - h(Y_{j,i}^n|W_j,\mathcal{H}^n,\mathcal{G}) \tag{6.95}$$

$$\leq n\log P - h(Y_{j,i}^n|W_j,\mathcal{H}^n,\mathcal{G}) + n\cdot O(1) \tag{6.96}$$

where $R_{j,i}$ denotes the achievable rate of the compound receiver i, and is the same as R_j. Let $\sum_{j\in\mathcal{S}}|\mathcal{T}_j'| = K - |\mathcal{S}'|$. By adding all achievable rates of all compound receivers, we have

$$n\left(\sum_{j\in\mathcal{S}}\sum_{i\in\mathcal{T}_j'} R_{j,i} - \epsilon_n\right) \tag{6.97}$$

$$\leq n\sum_{j\in\mathcal{S}}|\mathcal{T}_j'|\log P - h(\{Y_{j,i}^n, j\in\mathcal{S}, i\in\mathcal{T}_j'\}|W_{\mathcal{S}},\mathcal{H}^n,\mathcal{G}) + n\cdot O(1) \tag{6.98}$$

$$= n\sum_{j\in\mathcal{S}}|\mathcal{T}_j'|\log P - h(\{Y_{j,i}^n, j\in\mathcal{S}, i\in\mathcal{T}_j'\}|W_{\mathcal{S}},X_{\mathcal{S}'}^n,\mathcal{H}^n,\mathcal{G}) + n\cdot O(1) \tag{6.99}$$

$$= n\sum_{j\in\mathcal{S}}|\mathcal{T}_j'|\log P - h(\tilde{\boldsymbol{Y}}_{\mathcal{T}_{\mathcal{S}}'}^n|W_{\mathcal{S}},X_{\mathcal{S}'}^n,\mathcal{H}^n,\mathcal{G}) + n\cdot O(1) \tag{6.100}$$

$$= n\sum_{j\in\mathcal{S}}|\mathcal{T}_j'|\log P - h(\boldsymbol{X}_{\mathcal{K}\setminus\mathcal{S}'}^n + \boldsymbol{H}_{\mathcal{T}_{\mathcal{S}}'}^{-1}\boldsymbol{Z}_{\mathcal{T}_{\mathcal{S}}'}^n|W_{\mathcal{S}},X_{\mathcal{S}'}^n,\mathcal{H}^n,\mathcal{G}) + n\cdot O(1) \tag{6.101}$$

$$= n\sum_{j\in\mathcal{S}}|\mathcal{T}_j'|\log P - h(X_{\mathcal{K}\setminus\mathcal{S}'}^n + \bar{Z}_{\mathcal{K}\setminus\mathcal{S}'}^n|W_{\mathcal{S}},X_{\mathcal{S}'}^n,\mathcal{H}^n,\mathcal{G}) + n\cdot O(1) \tag{6.102}$$

$$\leq n\sum_{j\in\mathcal{S}}|\mathcal{T}_j'|\log P - n\sum_{j\in\mathcal{S}^c} R_j + n\cdot O(1) \tag{6.103}$$

where (6.99) is due to the fact that the knowledge of $W_{\mathcal{S}}$ is equivalent to the knowledge of $X_{\mathcal{S}'}$ given topological information, (6.100) is because translation does not change the differential entropy, (6.101) is obtained because $\boldsymbol{H}_{\mathcal{T}_{\mathcal{S}}'}$ is a square matrix and has full rank almost surely, in (6.102), $\bar{Z}_{\mathcal{K}\setminus\mathcal{S}'}^n$ is the bounded noise terms, and the last inequality is from the decodable condition similar to that in (6.66). By the definition of the symmetric DoF, it follows that

$$d_{\text{sym}} \leq \frac{\sum_{j\in\mathcal{S}}|\mathcal{T}_j'|}{\sum_{j\in\mathcal{S}}|\mathcal{T}_j'| + |\mathcal{S}^c|} \tag{6.104}$$

$$= \frac{K - |\mathcal{S}'|}{2K - |\mathcal{S}'| - |\mathcal{S}|}. \tag{6.105}$$

Among all the possible $\mathcal{S}$, we have the outer bound of symmetric DoF

$$d_{\text{sym}} \leq \min_{\mathcal{S}\subseteq\mathcal{K}} \frac{K - |\mathcal{S}'|}{2K - |\mathcal{S}'| - |\mathcal{S}|} \tag{6.106}$$

where $\mathcal{S}$ and $\mathcal{S}'$ are subject to two constraints: $\cup_{j\in\mathcal{S}}\mathcal{T}_j = \mathcal{K}$ and $\mathcal{S}' = \{i|[\boldsymbol{B}^{\mathsf{T}}]_i \cdot \mathbf{f}_{\text{idx}}(\mathcal{S}) = |\mathcal{R}_i|\}$.

6.5.7 Proof of Corollary 6.3

When $K = 2$, the network is fully connected and $d_{\text{sym}} = \frac{1}{2}$ is optimal. So, in the rest of the proof, we focus on $K \geq 3$. From the graph theoretic perspective, any two $(K, 2)$-regular networks are similar, because they are in fact the same cycle with rearranged vertices. Hence, it suffices to consider one typical topology of the $(K, 2)$-regular networks, e.g., a K-cell cyclic Wyner network, for the convenience of presentation. The received signal at Receiver j of the K-cell cyclic Wyner model can be given as

$$Y_j = h_{j,j-1}X_{j-1} + h_{j,j}X_j + Z_j \tag{6.107}$$

where the indices are modulo K, and W_i, W_{i+1} are the only accessible messages to Transmitter i. In what follows, we will present first the converse, followed by the achievability proof.

Converse

We consider two cases when K is even or odd.

- K is even: Let $\mathcal{S} = \{1, 3, \ldots, K-1\}$ and $\mathcal{S}' = \emptyset$. Consider the received signals $Y_{\mathcal{S}}$ and the signals of their respective compound receivers $\tilde{Y}_{\mathcal{S}}$. Following the proof of the general case, we have

$$2n\sum_{j\in\mathcal{S}} R_j \leq nK\log P - h(Y_{\mathcal{S}}^n, \tilde{Y}_{\mathcal{S}}^n|W_{\mathcal{S}}, \mathcal{H}^n, \mathcal{G}) \tag{6.108}$$

$$= nK\log P - n(R_2 + R_4 + \cdots + R_K) + n\cdot O(1) \tag{6.109}$$

  where the noisy version $\{X_i^n + \bar{Z}_i^n, i \in \mathcal{K}\}$ can be recovered from K linearly independent equations. Thus, with $|\mathcal{S}| = \frac{K}{2}$ and $|\mathcal{S}'| = 0$, it follows that

$$d_{\text{sym}} \leq \frac{K}{K + K/2} = \frac{2}{3}. \tag{6.110}$$

- K is odd: Let $\mathcal{S} = \{1, 3, \ldots, K-2, K\}$ and $\mathcal{S}' = \{K\}$. Consider here the received signals $Y_{\mathcal{S}}$ and the signals of their respective compound receivers $\tilde{Y}_{\mathcal{S}\backslash\{K-2,K\}}$. Similarly, we have

$$\begin{aligned} 2n \sum_{j\in\mathcal{S}\backslash\{K-2,K\}} R_j + nR_{K-2} + nR_K & \quad (6.111)\\ \leq (K-1)\log P - h(Y_{\mathcal{S}}^n, \tilde{Y}_{\mathcal{S}\backslash\{K-2,K\}}^n | W_{\mathcal{S}}, \mathcal{H}^n, \mathcal{G}) & \quad (6.112)\\ = n(K-1)\log P - h(Y_{\mathcal{S}}^n, \tilde{Y}_{\mathcal{S}\backslash\{K-2,K\}}^n | W_{\mathcal{S}}, X_K^n, \mathcal{H}^n, \mathcal{G}) & \quad (6.113)\\ = n(K-1)\log P - n(R_2 + R_4 + \cdots + R_{K-1}) & \quad (6.114) \end{aligned}$$

where X_K^n is reproducible with W_1 and W_k, and the noisy version $\{X_i^n + \bar{Z}_i^n, i \in \mathcal{K}\backslash K\}$ can be recovered from $K-1$ linearly independent equations. Thus, with $|\mathcal{S}| = \frac{K+1}{2}$ and $|\mathcal{S}'| = 1$, it follows that

$$d_{\text{sym}} \leq \frac{K-1}{K-1+\frac{K-1}{2}} = \frac{2}{3}. \quad (6.115)$$

To sum up, we have $d_{\text{sym}} \leq \frac{2}{3}$ whenever K is even or odd.

Achievability

Although the general achievability proof has been presented with general d, we make it concrete here for $d = 2$. During three time slots, we send at Transmitter i

$$\boldsymbol{X}_i = \boldsymbol{V}_{i-1} X_i(W_{i+1}^1) + \boldsymbol{V}_i X_i(W_i^2) \quad (6.116)$$

where $\{\boldsymbol{V}_i, i = 1, \ldots, n\}$ are 3×1 vectors satisfy that any three of them are linearly independent, almost surely. At Receiver j, we have

$$\begin{aligned} \boldsymbol{Y}_j &= h_{j,j-1}\boldsymbol{X}_{j-1} + h_{j,j}\boldsymbol{X}_j + \boldsymbol{Z}_j & (6.117)\\ &= h_{j,j-1}(\boldsymbol{V}_{j-2}X_{j-1}(W_j^1) + \boldsymbol{V}_{j-1}X_{j-1}(W_{j-1}^2)) \\ &\quad + h_{j,j}(\boldsymbol{V}_{j-1}X_j(W_{j+1}^1) + \boldsymbol{V}_j X_j(W_j^2)) + \boldsymbol{Z}_j & (6.118)\\ &= \underbrace{h_{j,j-1}\boldsymbol{V}_{j-2}X_{j-1}(W_j^1) + h_{j,j}\boldsymbol{V}_j X_j(W_j^2)}_{\text{desired signal}} \\ &\quad + \underbrace{\boldsymbol{V}_{j-1}(h_{j,j-1}X_{j-1}(W_{j-1}^2) + h_{j,j}X_j(W_{j+1}^1))}_{\text{aligned interferences}} + \boldsymbol{Z}_j. & (6.119) \end{aligned}$$

The interferences carrying messages W_{j-1} and W_{j+1} are aligned together in the direction of $\boldsymbol{V}_{j-1}$, leaving two-dimensional interference-free subspace for desired signals carrying on message realizations W_j^1 and W_j^2. Therefore, two messages are delivered during three time slots, yielding $\frac{2}{3}$ DoF per user, which coincides with the outer bound. This completes the proof of optimality.

6.5.8 Proof of Theorem 6.5

$d_{\text{sym}} = \frac{2}{K}$ is achievable

First, we consider the case when there exists a Hamiltonian cycle $i_1 \leftrightarrow i_2 \leftrightarrow \cdots \leftrightarrow i_K \leftrightarrow i_1$ in the alignment-feasible graph ($\mathcal{G}_{AFG}$). According to the definition of $\mathcal{G}_{AFG}$, it follows that, there exist z_j^1 and z_{j+1}^2, such that

$$z_j^1 \in \mathcal{T}_{i_j} \cap \mathcal{T}_{i_{j+1}}^c, \quad \text{and} \quad z_{j+1}^2 \in \mathcal{T}_{i_{j+1}} \cap \mathcal{T}_{i_j}^c \tag{6.120}$$

with $z_j^1, z_j^2 \in \mathcal{T}_{i_j}$ and $z_{j-1}^1, z_{j+1}^2 \notin \mathcal{T}_{i_j}$, for $j \in \mathcal{K}$. Thus, we send along the direction $\boldsymbol{V}_j \in \mathbb{C}^{K\times 1}$ two signals $X_{z_j^1}(W_{i_j}^1)$ and $X_{z_{j+1}^2}(W_{i_{j+1}}^2)$ from Transmitter z_j^1 and Transmitter z_{j+1}^2, respectively, for $j \in \mathcal{K}$. The received signals at Receiver i_j during K time slots can be given as a compact form by

$$\boldsymbol{Y}_{i_j} = \sum_{s=1}^{K} \boldsymbol{V}_s \left(h_{i_j,z_s^1} X_{z_s^1}(W_{i_s}^1)\mathbf{1}(z_s^1 \in \mathcal{T}_{i_j}) + h_{i_j,z_{s+1}^2} X_{z_{s+1}^2}(W_{i_{s+1}}^2)\mathbf{1}(z_{s+1}^2 \in \mathcal{T}_{i_j}) \right) \tag{6.121}$$

$$= \underbrace{\boldsymbol{V}_j h_{i_j,z_j^1} X_{z_j^1}(W_{i_j}^1) + \boldsymbol{V}_{j-1} h_{i_j,z_j^2} X_{z_j^2}(W_{i_j}^2)}_{\text{desired signal}} + \underbrace{\sum_{s=1,s\neq j-1,j}^{K} \boldsymbol{V}_s \left(h_{i_j,z_s^1} X_{z_s^1}(W_{i_s}^1)\mathbf{1}(z_s^1 \in \mathcal{T}_{i_j}) + h_{i_j,z_{s+1}^2} X_{z_{s+1}^2}(W_{i_{s+1}}^2)\mathbf{1}(z_{s+1}^2 \in \mathcal{T}_{i_j}) \right)}_{\text{aligned interferences}} \tag{6.122}$$

where $\mathbf{1}(\cdot)$ is the indicator function with value 1 if the parameter is true and 0 otherwise. It is readily verified that two symbols $W_{i_j}^1$ and $W_{i_j}^2$ can be retrieved almost surely, yielding symmetric DoF of $\frac{2}{K}$. An example is shown in Fig. 6.6, where $1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow 4 \leftrightarrow 5 \leftrightarrow 1$ forms a Hamiltonian cycle.

Second, we consider a perfect matching in $\mathcal{G}_{AFG}$ where K is even, say $i_1 \leftrightarrow i_2, \ldots, i_{K-1} \leftrightarrow i_K$. Similarly, there exist z_j and z_{j+1}, such that

$$z_j \in \mathcal{T}_{i_j} \cap \mathcal{T}_{i_{j+1}}^c, \quad \text{and} \quad z_{j+1} \in \mathcal{T}_{i_{j+1}} \cap \mathcal{T}_{i_j}^c, \quad j = 1, 3, \ldots, K-1 \tag{6.123}$$

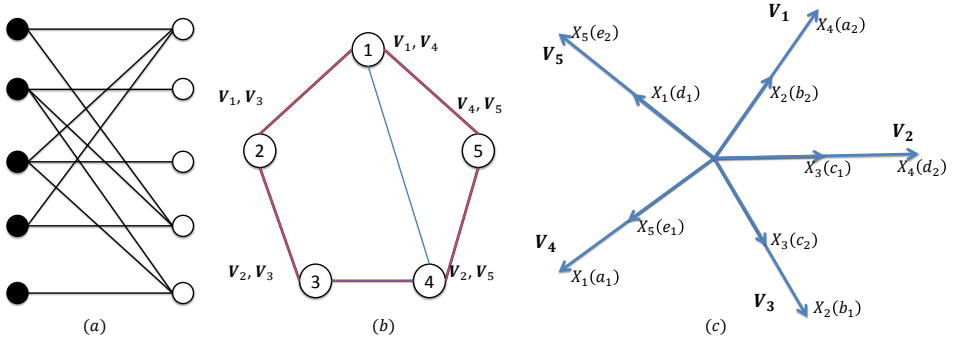


Figure 6.6: (a) An instance of TIM-CoMP problem ($K = 5$), (b) alignment-feasible graph, where the messages connected with an edge are alignment-feasible and can be aligned in the same subspace, and the red edges indicate a Hamiltonian cycle, and (c) an interference alignment scheme, where every message appears twice.

with $z_j \in \mathcal{T}_{i_j}$ and $z_{j+1} \notin \mathcal{T}_{i_j}$. Thus, during in total $\frac{K}{2}$ time slots, we send two signals $X_{z_j}(W_{i_j})$ and $X_{z_{j+1}}(W_{i_{j+1}})$ from Transmitter z_j and Transmitter z_{j+1}, respectively, with the same precoder $\boldsymbol{V}_j \in \mathbb{C}^{\frac{K}{2}\times 1}$. The received signals at Receiver i_j during $\frac{K}{2}$ time slots can be similarly written as

$$\boldsymbol{Y}_{i_j} = \sum_{s=1}^{\frac{K}{2}} \boldsymbol{V}_s \left(h_{i_j,z_s} X_{z_s}(W_{i_s}) \mathbf{1}(z_s \in \mathcal{T}_{i_j}) + h_{i_j,z_{s+1}} X_{z_{s+1}}(W_{i_{s+1}}) \mathbf{1}(z_{s+1} \in \mathcal{T}_{i_j}) \right) \tag{6.124}$$

$$= \underbrace{\boldsymbol{V}_j h_{i_j,z_j} X_{z_j}(W_{i_j})}_{\text{desired signal}} + \underbrace{\sum_{s=1,s\neq j}^{\frac{K}{2}} \boldsymbol{V}_s \left(h_{i_j,z_s} X_{z_s}(W_{i_s}) \mathbf{1}(z_s \in \mathcal{T}_{i_j}) + h_{i_j,z_{s+1}} X_{z_{s+1}}(W_{i_{s+1}}) \mathbf{1}(z_{s+1} \in \mathcal{T}_{i_j}) \right)}_{\text{aligned interferences}} \tag{6.125}$$

with which the message W_{i_j} is recovered, yielding $\frac{2}{K}$ DoF per user. This completes the proof.

$d_{\text{sym}} = \frac{2}{K-q}$ is achievable

The achievability is similar to the previous case, but the duration of transmission is shortened. Without loss of generality, we assume the Hamiltonian

cycle $1 \leftrightarrow 2 \leftrightarrow \cdots \leftrightarrow K \leftrightarrow 1$ for the brevity of presentation. According to the definition of alignment-feasible graph, there exist z_s^1 and z_{s+1}^2, such that

$$z_s^1 \in \mathcal{T}_s \cap \mathcal{T}_{s+1}^c, \quad \text{and} \quad z_{s+1}^2 \in \mathcal{T}_{s+1} \cap \mathcal{T}_s^c \tag{6.126}$$

with $z_s^1 \in \mathcal{T}_s$ and $z_{s+1}^2 \notin \mathcal{T}_s$, for $s \in \mathcal{K}$. Assuming

$$k_0 \in \arg\min_k \sum_j \boldsymbol{A}_{kj}, \tag{6.127}$$

we have $\sum_j \boldsymbol{A}_{k_0 j} = q$ and thus

$$\mathbf{f}_{\text{idx}}^{-1}(\boldsymbol{A}_{k_0}^{\mathsf{T}}) = \{j_1, \ldots, j_q\} \triangleq \mathcal{J}_q \tag{6.128}$$

where $\mathbf{f}_{\text{idx}}^{-1} : \{0,1\}^K \mapsto \mathcal{B}$ is the inverse function of $\mathbf{f}_{\text{idx}}$.

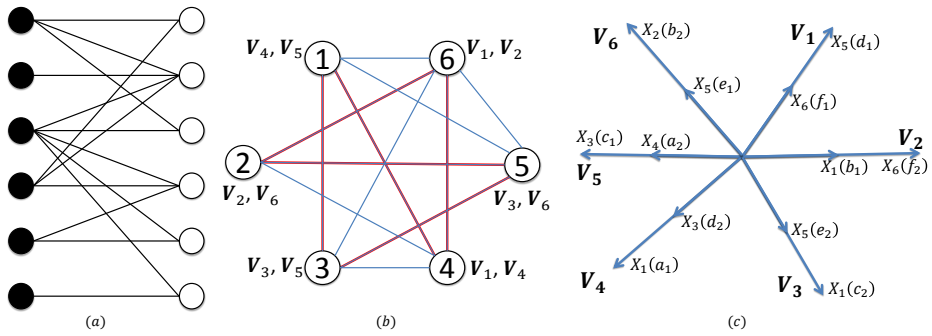


Figure 6.7: (a) An instance of TIM-CoMP problem ($K = 6$), and (b) alignment-feasible graph, in which there may exist many Hamiltonian cycles, and the cycle with red edges is one of them. (c) An interference alignment scheme, where every message appears twice, and for each receiver there exists at least one absent subspace.

According to the definition of alignment non-conflict matrix, if $\boldsymbol{A}_{k_0 j} = 1$, then

$$\mathcal{T}_{i_j} \bigcap \mathcal{T}_{i_{j+1}}^c \bigcap_{k:\boldsymbol{A}_{kj}=1} \mathcal{T}_{k_0}^c \neq \emptyset, \quad \text{and} \quad \mathcal{T}_{i_{j+1}} \bigcap \mathcal{T}_{i_j}^c \bigcap_{k:\boldsymbol{A}_{kj}=1} \mathcal{T}_{k_0}^c \neq \emptyset, \tag{6.129}$$

meaning that there is non-conflict to make W_{i_j} and $W_{i_{j+1}}$ aligned with the occupied subspace absent to Receiver k_0. It follows that, there exist $z_{j_t}^1$ and $z_{j_t+1}^2$ ($j_t \in \mathcal{J}_q$), such that

$$z_{j_t}^1 \in \mathcal{T}_{j_t} \cap \mathcal{T}_{j_t+1}^c \cap \mathcal{T}_{k_0}^c, \quad \text{and} \quad z_{j_t+1}^2 \in \mathcal{T}_{j_t+1} \cap \mathcal{T}_{j_t}^c \cap \mathcal{T}_{k_0}^c \tag{6.130}$$

with $z^1_{j_t}, z^2_{j_t+1} \notin \mathcal{T}_{k_0}$. We send $X_{z^1_s}(W^1_s)$ and $X_{z^2_{s+1}}(W^2_{s+1})$ at Transmitter z^1_s and Transmitter z^2_{s+1}, respectively, along with the subspace spanned by $\boldsymbol{V}_s \in \mathbb{C}^{(K-q)\times 1}$, giving the received signal at Receiver k_0 as

$$\begin{aligned}\boldsymbol{Y}_{k_0} &= \sum_{s=1}^{K} \boldsymbol{V}_s \left(h_{k_0,z^1_s} X_{z^1_s}(W^1_s)\mathbf{1}(z^1_s \in \mathcal{T}_{k_0}) + h_{k_0,z^2_{s+1}} X_{z^2_{s+1}}(W^2_{s+1})\mathbf{1}(z^2_{s+1} \in \mathcal{T}_{k_0}) \right) \\ &= \sum_{s=1, s\notin\mathcal{J}_q}^{K} \boldsymbol{V}_s \left(h_{k_0,z^1_s} X_{z^1_s}(W^1_s)\mathbf{1}(z^1_s \in \mathcal{T}_{k_0}) + h_{k_0,z^2_{s+1}} X_{z^2_{s+1}}(W^2_{s+1})\mathbf{1}(z^2_{s+1} \in \mathcal{T}_{k_0}) \right) \qquad (6.131) \\ &= \boldsymbol{V}_{k_0} h_{k_0,z^1_{k_0}} X_{z^1_{k_0}}(W^1_{k_0}) + \boldsymbol{V}_{k_0-1} h_{k_0,z^2_{k_0}} X_{z^2_{k_0}}(W^2_{k_0}) \\ &+ \sum_{\substack{s=1, s\notin\mathcal{J}_q, \\ s\neq k_0-1, k_0}}^{K} \boldsymbol{V}_s \left(h_{k_0,z^1_s} X_{z^1_s}(W^1_s)\mathbf{1}(z^1_s \in \mathcal{T}_{k_0}) + h_{k_0,z^2_{s+1}} X_{z^2_{s+1}}(W^2_{s+1})\mathbf{1}(z^2_{s+1} \in \mathcal{T}_{k_0}) \right) \qquad (6.132)\end{aligned}$$

It follows that the desired messages by Receiver k_0 can be recovered in a $K-q$ dimensional space with two interference-free subspace and $K-q-2$ dimensional subspace with interferences aligned. According to the definition of q, we conclude that the overall $K-q$ dimensional space is sufficient to support other receivers with $\sum_j \boldsymbol{A}_{kj} \geq q$. As such, the symmetric DoF $\frac{2}{K-q}$ are achievable. An example can be found in Fig. 6.7 with $K=6$ and $q=1$. For a Hamiltonian cycle $1 \leftrightarrow 3 \leftrightarrow 5 \leftrightarrow 2 \leftrightarrow 6 \leftrightarrow 4 \leftrightarrow 1$, the alignment non-conflict matrix is

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \qquad (6.133)$$

by which k_0 can be 2 or 4.

$d_{\text{sym}} = \frac{1}{\kappa}$ is achievable

According to the definition of proper partition, for a portion $\mathcal{P}_i = \{i_1, \ldots, i_{p_i}\}$, we assume with $k = 1, \ldots, p_i$ that

$$z_{i_k} \in \mathcal{T}_{i_k} \bigcap \left(\bigcup_{i_j \in \mathcal{P}_i \setminus i_k} \mathcal{T}_{i_j} \right)^c, \quad \forall\ i_k \in \mathcal{P}_i. \qquad (6.134)$$

with $z_{i_k} \in \mathcal{T}_{i_k}$ and $z_{i_k} \notin \mathcal{T}_{i_j}$, $\forall\ j \neq k$. Thus we send $\{X_{z_{i_k}}(W_{i_k}), k = 1, \ldots, p_i\}$ at Transmitter z_{i_k} via the same precoder $\boldsymbol{V}_i \in \mathbb{C}^{\kappa \times 1}$, yielding the receiver signal at Receiver i_k as

$$\boldsymbol{Y}_{i_k} = \sum_{j=1}^{\kappa} \boldsymbol{V}_j \left(\sum_{s=1}^{p_j} h_{i_k, z_{j_s}} X_{z_{j_s}}(W_{j_s}) \mathbf{1}(z_{j_s} \in \mathcal{T}_{i_k}) \right) \tag{6.135}$$

$$= \boldsymbol{V}_i h_{i_k, z_{i_k}} X_{z_{i_k}}(W_{i_k}) + \boldsymbol{V}_i \left(\sum_{s=1, s\neq k}^{p_i} h_{i_k, z_{i_s}} X_{z_{i_s}}(W_{i_s}) \mathbf{1}(z_{i_s} \in \mathcal{T}_{i_k}) \right) + \sum_{j=1, j\neq i}^{\kappa} \boldsymbol{V}_j \left(\sum_{s=1}^{p_j} h_{i_k, z_{j_s}} X_{z_{j_s}}(W_{j_s}) \mathbf{1}(z_{j_s} \in \mathcal{T}_{i_k}) \right) \tag{6.136}$$

$$= \underbrace{\boldsymbol{V}_i h_{i_k, z_{i_k}} X_{z_{i_k}}(W_{i_k})}_{\text{desired signal}} + \underbrace{\sum_{j=1, j\neq i}^{\kappa} \boldsymbol{V}_j \left(\sum_{s=1}^{p_j} h_{i_k, z_{j_s}} X_{z_{j_s}}(W_{j_s}) \mathbf{1}(z_{j_s} \in \mathcal{T}_{i_k}) \right)}_{\text{aligned interferences}} \tag{6.137}$$

with which the desired signal can be retrieved with high probability during κ time slots. This applies all messages that offers $\frac{1}{\kappa}$ DoF per user. An example is shown in Fig. 6.8(b), where a proper partition can be $\{\{1, 3, 5\}, \{2, 4, 6\}\}$.

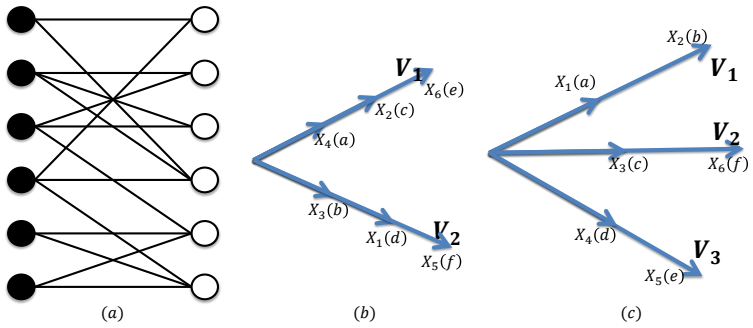


Figure 6.8: (a) An instance of TIM-CoMP problem ($K = 6$). The network topology has two different proper partitions as shown in (b) and (c), where the messages in each portion are alignment-feasible and can be aligned in the same subspace. Both partitions give the same symmetric DoF of $\frac{1}{2}$.

$d_{\text{sym}} = \frac{1}{\kappa - q}$ is achievable

The achievability is similar to the previous case, but the required number of subspace dimension is reduced. Assume similarly

$$m \in \arg\min_i \sum_j \boldsymbol{A}_{ij}, \tag{6.138}$$

we have $\sum_j \boldsymbol{A}_{mj} = q$ and $\mathbf{f}_{\text{idx}}^{-1}(\boldsymbol{A}_m^{\mathsf{T}}) = \mathcal{J}_q$.

According to the definition of proper partition, there exists z_{i_k} with $i \in \{1, \ldots, \kappa\}$ such that

$$z_{i_k} \in \mathcal{T}_{i_k} \bigcap \left(\bigcup_{i_j \in \mathcal{P}_i \backslash i_k} \mathcal{T}_{i_j} \right)^c, \quad \forall\ i_k \in \mathcal{P}_i \tag{6.139}$$

with $z_{i_k} \in \mathcal{T}_{i_k}, \forall\ i$, and according to the alignment non-conflict matrix, if $\boldsymbol{A}_{mj} = 1$, then

$$\mathcal{T}_{j_t} \bigcap \left(\bigcup_{j_s \in \mathcal{P}_j \backslash j_t} \mathcal{T}_{j_s} \right)^c \bigcap_{i:\boldsymbol{A}_{mj}=1} \mathcal{T}_{m_k}^c \neq \emptyset, \quad \forall\ j_t \in \mathcal{P}_j, \forall\ m_k \in \mathcal{P}_m, \tag{6.140}$$

meaning that this is non-conflict to make the messages in portion $\mathcal{P}_j$ aligned with the spanned subspace absent to all the receivers in $\mathcal{P}_m$. It follows that, there exists z_{j_t} with $j \in \mathcal{J}_q$, such that

$$z_{j_t} \in \mathcal{T}_{j_t} \bigcap \left(\bigcup_{j_s \in \mathcal{P}_j \backslash j_t} \mathcal{T}_{j_s} \right)^c \bigcap \mathcal{T}_{m_k}^c, \quad \forall\ m_k \in \mathcal{P}_m, j_t \in \mathcal{P}_j \tag{6.141}$$

with $z_{j_t} \notin \mathcal{T}_{m_k}, \forall\ m_k \in \mathcal{P}_m, j_t \in \mathcal{P}_j$. The received signal at Transmitter m_k with $m_k \in \mathcal{P}_m$ can be given as

$$\begin{aligned}
\boldsymbol{Y}_{m_k} &= \sum_{l=1}^{\kappa} \boldsymbol{V}_l \left(\sum_{s=1}^{p_l} h_{m_k, z_{l_s}} X_{z_{l_s}}(W_{l_s}) \mathbf{1}(z_{l_s} \in \mathcal{T}_{m_k}) \right) && (6.142) \\
&= \sum_{l=1, l \notin \mathcal{J}_q}^{\kappa} \boldsymbol{V}_l \left(\sum_{s=1}^{p_l} h_{m_k, z_{l_s}} X_{z_{l_s}}(W_{l_s}) \mathbf{1}(z_{l_s} \in \mathcal{T}_{m_k}) \right) && (6.143) \\
&= \boldsymbol{V}_m h_{m_k, z_{m_k}} X_{z_{m_k}}(W_{m_k}) \mathbf{1}(z_{m_k} \in \mathcal{T}_{m_k}) \\
&\quad + \boldsymbol{V}_m \left(\sum_{s=1, s \neq k}^{p_m} h_{m_k, z_{m_s}} X_{z_{m_s}}(W_{m_s}) \mathbf{1}(z_{m_s} \in \mathcal{T}_{m_k}) \right)
\end{aligned}$$

$$+\sum_{l=1,l\neq m,l\notin\mathcal{J}_q}^{\kappa}\boldsymbol{V}_l\left(\sum_{s=1}^{p_l}h_{m_k,z_{l_s}}X_{z_{l_s}}(W_{l_s})\mathbf{1}(z_{l_s}\in\mathcal{T}_{m_k})\right) \tag{6.144}$$

$$=\boldsymbol{V}_m h_{m_k,z_{m_k}}X_{z_{m_k}}(W_{m_k}) +\sum_{l=1,l\neq m,l\notin\mathcal{J}_q}^{\kappa}\boldsymbol{V}_l\left(\sum_{s=1}^{p_l}h_{m_k,z_{l_s}}X_{z_{l_s}}(W_{l_s})\mathbf{1}(z_{l_s}\in\mathcal{T}_{m_k})\right) \tag{6.145}$$

where $\boldsymbol{V}_l\in\mathbb{C}^{(\kappa-q)\times 1}$ is sufficient to recover desired message W_{m_k}, yielding $\frac{1}{\kappa-q}$ DoF. According to the definition of q, this $\kappa-q$ dimensional space suffices to support all other receivers. Thus, symmetric DoF of $\frac{1}{\kappa-q}$ are achievable, almost surely. An example is shown in Fig. 6.8(c), where there exists another proper partition $\{\{1,2\},\{3,6\},\{4,5\}\}$ with $\kappa=3$ and $q=1$. With different partition, we have the same achievable symmetric DoF.

6.5.9 Proof of Theorem 6.6

In this theorem, we represent the inner bound of the TIM-CoMP problem by a graph-theoretic parameter, i.e., fractional covering number. To this end, we will bridge our problem to the hypergraph fractional covering problem, which is in general a set covering problem.

First of all, we construct such a hypergraph $\mathcal{H}_{\mathcal{G}}$ according to the network topology. From the definition of proper partition and the proof in 6.5.8, it follows that if a set $\mathcal{X}_i\triangleq\{x_{i_1},x_{i_2},\ldots,x_{i_{|\mathcal{X}_i|}}\}\subseteq\mathcal{S}$ satisfies

$$\mathcal{T}_{x_{i_k}}\bigcap\left(\bigcup_{x_{i_j}\in\mathcal{X}_i\backslash x_{i_k}}\mathcal{T}_{x_{i_j}}\right)^c\neq\emptyset,\quad\forall\ x_{i_k}\in\mathcal{X}_i, \tag{6.146}$$

then any two messages in $W_{\mathcal{X}_i}$ are mutually alignment-feasible. The messages $\{W_{x_{i_k}},x_{i_k}\in\mathcal{X}_i\}$ can be sent from the transmitters $\{z_{i_k}\}$ in the form of $X_{z_{i_k}}(W_{x_{i_k}})$ with the same precoding vector $\boldsymbol{V}_i$,[3], where

$$z_{i_k}\in\mathcal{T}_{x_{i_k}}\bigcap\left(\bigcup_{x_{i_j}\in\mathcal{X}_i\backslash x_{i_k}}\mathcal{T}_{x_{i_j}}\right)^c. \tag{6.147}$$

As such, only one transmitted signal $X_{z_{i_k}}(W_{x_{i_k}})$ is active in subspace spanned by $\boldsymbol{V}_i$ at Receiver x_{i_k}, and hence $W_{x_{i_k}}$ is recoverable from this subspace. The

[3]Alternatively, the links from Transmitter z_{i_k} to Receiver x_{i_k} $(k=1,\ldots,|\mathcal{X}_i|)$ can be scheduled at the same time slot.

presence of the subspace spanned by $\boldsymbol{V}_i$ carrying on messages with indices in $\mathcal{X}_i$ guarantees the successful delivery of the messages in $W_{\mathcal{X}_i}$. Thus, the set $\mathcal{X}_i$ can serve as a hyperedge of $\mathcal{H}_{\mathcal{G}}$. Any set of elements in $\mathcal{K}$ that satisfies the condition (6.146) serves as a hyperedge. As a result, the hypergraph $\mathcal{H}_{\mathcal{G}}$ is constructed with vertex set $\mathcal{K}$ and hyperedge set being enumeration of all possible sets of elements satisfy condition (6.146).

Our problem to find the symmetric DoF is equivalent to the covering problem of this hypergraph to find the minimum number of hyperedges $\{\mathcal{X}_i, i = 1, 2, \ldots, \tau_t\}$ such that each $j \in \mathcal{K}$ appears at least t of the $\mathcal{X}_i$'s. According to the definition of hypergraph covering in Definition 6.9, the minimum number of hyperedges that meets the covering problem can be represented by the t-fold covering number. With hyperedge cover of $\tau_t(\mathcal{H}_{\mathcal{G}})$ times, each vertex in $\mathcal{K}$ is covered at least t times, meaning that within a $\tau_t(\mathcal{H}_{\mathcal{G}})$ dimensional subspace spanned by $\boldsymbol{V}_i \in \mathbb{C}^{\tau_t \times 1}, i = 1, \ldots, \tau_t$, each $W_j, j \in \mathcal{K}$ can be delivered t times free of interference. As a consequence, the achievable symmetric DoF can be represented by

$$d_{\text{sym}} = \sup_t \frac{t}{\tau_t(\mathcal{H}_{\mathcal{G}})} = \frac{1}{\tau_f(\mathcal{H}_{\mathcal{G}})}, \tag{6.148}$$

where $\tau_f(\mathcal{H}_{\mathcal{G}})$ is the hypergraph fractional covering number as defined in Definition 6.9.

6.5.10 Proof of Theorem 6.7

The proof follows the channel enhancement approach in [49] with slight modification by taking transmitter cooperation (i.e., message sharing) into account. We brief the steps of the channel enhancement as follows.

- Denote by $\mathcal{C}_1$ the capacity region of the TIM-CoMP problem, where Transmitter i is endowed with the messages desired by its connected receivers, i.e., $W_{\mathcal{R}_i}$, for all $i \in \mathcal{K}$.

- $\forall\ k, j \in \mathcal{K}$, if $j \in \mathcal{T}_k$, we specify

$$h_{kj} = \sqrt{\frac{SNR}{P_j}} \tag{6.149}$$

  which will not impact on the reliability of the capacity-achieving coding scheme.

- $\forall\, k, j \in \mathcal{K}$, if $j \notin \mathcal{T}_k$, we provide $W_{\mathcal{R}_j}$ to Receiver k as side information, and connect the missing link by setting the channel coefficient as a non-zero value

$$h_{kj} = \sqrt{\frac{SNR}{P_j}}, \tag{6.150}$$

  where the newly enabled interferences from Transmitter j can be eliminated given the side information $W_{\mathcal{R}_j}$.

- Allowing full transmitter cooperation and full CSIT, the channel turns out an MISO channel to each receiver, where all received signals are statistically equivalent. Denote by $\mathcal{C}_2$ the capacity region of current channel. The capacity region is not diminished, i.e., $\mathcal{C}_1 \subseteq \mathcal{C}_2$.

- With the network equivalence theorem [130], the MISO channel can be replaced by a noise-free link with finite capacity, as the bottleneck link of index coding problem with capacity region $\mathcal{C}_3$.

It is noticed that all the above steps do not reduce the capacity region, i.e., $\mathcal{C}_1 \subseteq \mathcal{C}_2 \subseteq \mathcal{C}_3$, such that the capacity region of the index coding problem with side information $\cup_{j\in\mathcal{T}_k^c} W_{\mathcal{R}_j}$ can serve as an outer bound of our problem.

6.5.11 Proof of Corollary 6.4

The achievability is due to time division, whereas the converse comes from the outer bound of the index coding problem with side information $\cup_{j\in\mathcal{T}_k^c} W_{\mathcal{R}_j}$. From Corollary 7 in [49], it is necessary and sufficient to achieve the symmetric capacity of $\frac{1}{K}$ per message that the demand graph of the index coding problem is acyclic. Thus, if the demand graph of index coding problem $\mathbf{IC}(k|\mathcal{S}_k)$ with $\mathcal{S}_k = \cup_{j\in\mathcal{T}_k^c}\mathcal{R}_j$ is acyclic, it suffices that the TIM-CoMP problem is upper bounded by $\frac{1}{K}$, which is also achievable.

As an example, we illustrate in Fig. 6.9(a) a four-cell network with transmit sets $\mathcal{T}_1 = \mathcal{T}_2 = \{1,2\}, \mathcal{T}_3 = \mathcal{T}_4 = \{1,2,3,4\}$ and receive sets $\mathcal{R}_1 = \mathcal{R}_2 = \{1,2,3,4\}, \mathcal{R}_3 = \mathcal{R}_4 = \{3,4\}$. By providing Receivers 1 and 2 with $W_{3,4}$, we connect the missing links as shown in Fig. 6.9(c) without reducing the capacity region. Allowing full CSIT, the problem now is equivalent to the index coding problem (as in Fig. 6.9(c)) where messages $W_{1,2,3,4}$ are sent from one transmitter to Receiver j ($j = 1,2,3,4$) who demands W_j, and both Receivers 1 and 2 have the side information $W_{3,4}$. This index coding problem has no cycles in its demand graph as shown in Fig. 6.9(d), such that the optimal symmetric DoF value is $\frac{1}{K}$.

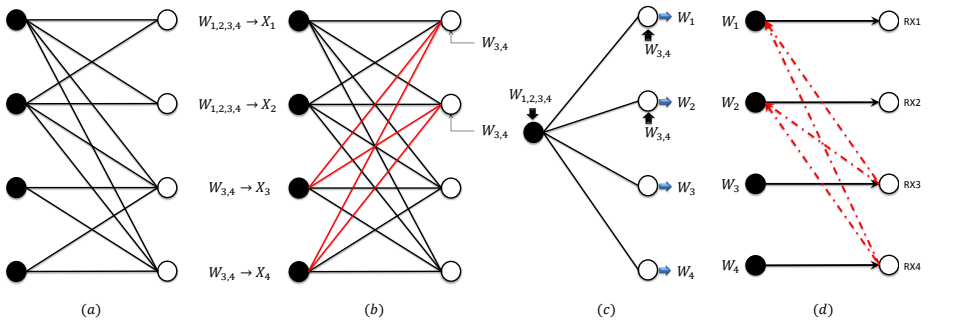


Figure 6.9: (a) An instance of TIM-CoMP problem ($K = 4$). By providing the side information $W_{3,4}$ to Receivers 1 and 2, the network becomes fully connected as shown in (b). Thus, the DoF region is outer bounded by an index coding problem with side information as in (c), whose corresponding directed demand graph is shown in (d). There exist no cycles in this directed graph in (d).

6.5.12 Proof of Corollary 6.5

First, we prove the necessary condition that, if the optimal symmetric DoF value is $\frac{1}{K}$, then $\mathcal{G}_{AFG}$ is an empty graph. We achieve this goal by constructing a proof by contraposition, i.e., if $\mathcal{G}_{AFG}$ is not empty, then the optimal symmetric DoF value is not $\frac{1}{K}$. Assume there exists an edge e_{ij} in $\mathcal{G}_{AFG}$, which imply W_i and W_j are alignment feasible, i.e., $\mathcal{T}_i \nsubseteq \mathcal{T}_j$ and $\mathcal{T}_i \nsubseteq \mathcal{T}_j$. According to the definition of proper partition, we have a proper partition with size $K-1$ where the indices i and j associate with W_i and W_j respectively are in the same portion and the rest $K-2$ messages form $K-2$ portions, such that $d_{\text{sym}} = \frac{1}{K-1}$ is achievable. Had proven the contraposition, the original statement automatically implies.

Second, we prove the sufficient condition that, if the AFG is an empty graph, then the optimal symmetric DoF value is $\frac{1}{K}$. Aware of the fact that the symmetric DoF $\frac{1}{K}$ can be trivially achieved by time division, we only have to prove $\frac{1}{K}$ is also an outer bound given that $\mathcal{G}_{AFG}$ is an empty graph. According to Theorem 6.7 and Corollary 6.4, our goal can be reached if the following statement is proved:

$$\text{if } \mathcal{G}_{AFG} = \emptyset, \text{ then } \mathbf{IC}(k|\mathcal{S}_k) \text{ is acyclic.}$$

We construct the proof of this statement by contraposition, i.e.,

$$\text{if the demand graph of the index coding problem } \mathbf{IC}(k|\mathcal{S}_k) \text{ is not acyclic,}$$
$$\text{then } \mathcal{G}_{AFG} \neq \emptyset, \text{ i.e., } \exists\ i,\ j, \text{ such that } \mathcal{T}_i \nsubseteq \mathcal{T}_j \text{ and } \mathcal{T}_j \nsubseteq \mathcal{T}_i.$$

To prove this contraposition, we first assume there exists a cycle involving only two messages, e.g., W_m and W_n, in the demand graph. Thus, we have $m \in \mathcal{S}_n = \cup_{j\in\mathcal{T}_n^c}\mathcal{R}_j$ and $n \in \mathcal{S}_m = \cup_{j\in\mathcal{T}_m^c}\mathcal{R}_j$, while $m \notin \mathcal{S}_m = \cup_{j\in\mathcal{T}_m^c}\mathcal{R}_j$ and $n \notin \mathcal{S}_n = \cup_{j\in\mathcal{T}_n^c}\mathcal{R}_j$, such that there exist $j_1 \in \mathcal{T}_n^c$ and $j_2 \in \mathcal{T}_m^c$ where $m \in \mathcal{R}_{j_1}$ and $n \in \mathcal{R}_{j_2}$ whereas $m \notin \mathcal{R}_{j_2}$ and $n \notin \mathcal{R}_{j_1}$. This leads to $\mathcal{R}_{j_1} \nsubseteq \mathcal{R}_{j_2}$ and $\mathcal{R}_{j_2} \nsubseteq \mathcal{R}_{j_1}$. Equivalently, there exist $t_1 \in \mathcal{R}_{j_1}$ and $t_2 \in \mathcal{R}_{j_2}$, such that $\mathcal{T}_{t_1} \nsubseteq \mathcal{T}_{t_2}$ and $\mathcal{T}_{t_2} \nsubseteq \mathcal{T}_{t_1}$, because both conditions imply the same alignment feasibility, where $X_{j_1}(W_{t_1})$ and $X_{j_2}(W_{t_2})$ can be aligned in the same subspace. Consequently, two messages W_{t_1} and W_{t_2} are alignment-feasible, and therefore connected in $\mathcal{G}_{AFG}$. Thus, $\mathcal{G}_{AFG} \neq \emptyset$ is proven.

Furthermore, we assume without loss of generality the smallest cycle involving more than two messages, i.e., $i_1 \to i_2 \to \cdots \to i_s \to i_1$ with directed edge from Message i_m to Receiver i_m then via Message i_{m+1} to Receiver i_{m+1} and so on, for $m = 1, 2, \ldots, s$, with modulo applied to the indices. Given the smallest cycle in the directed demand graph, we have

$$i_{m+1} \in \mathcal{S}_{i_m} = \cup_{j\in\mathcal{T}_{i_m}^c}\mathcal{R}_j, \tag{6.151}$$

$$i_{m+1} \notin \mathcal{S}_{i_n} = \cup_{j\in\mathcal{T}_{i_n}^c}\mathcal{R}_j, \quad \forall\, n \in \{1, 2, \ldots, s\}, \quad n \neq m. \tag{6.152}$$

From (6.151), it is readily verified that there must exist $j_m \in \mathcal{T}_{i_m}^c$, such that $i_{m+1} \in \mathcal{R}_{j_m}$. By set $n = m-1$ and $n = m+1$ respectively in (6.152), we have

$$\forall\, j_{m-1} \in \mathcal{T}_{i_{m-1}}^c, \quad i_{m+1} \notin \mathcal{R}_{j_{m-1}}, \tag{6.153}$$

$$\forall\, j_{m+1} \in \mathcal{T}_{i_{m+1}}^c, \quad i_{m+1} \notin \mathcal{R}_{j_{m+1}}. \tag{6.154}$$

It follows that $\mathcal{T}_{i_m}^c \nsubseteq \mathcal{T}_{i_{m-1}}^c$ and $\mathcal{T}_{i_m}^c \nsubseteq \mathcal{T}_{i_{m+1}}^c$ for all m, and in turn

$$\mathcal{T}_{i_{m-1}} \nsubseteq \mathcal{T}_{i_m}, \quad \text{and} \quad \mathcal{T}_{i_{m+1}} \nsubseteq \mathcal{T}_{i_m}, \quad \forall\, m. \tag{6.155}$$

Otherwise, it results in contradictions with $i_{m+1} \in \mathcal{R}_{j_m}$. Recalling that $i_1 \to i_2 \to \cdots \to i_s \to i_1$ forms a cycle, we conclude that $\mathcal{R}_{i_m} \nsubseteq \mathcal{R}_{i_{m+1}}$ and $\mathcal{R}_{i_{m+1}} \nsubseteq \mathcal{R}_{i_m}$ for all $m \in \{1, 2, \ldots, s\}$. Thus, we conclude that, if there exist a cycle in demand graph, then $\mathcal{G}_{AFG} \neq \emptyset$. Consequently, its contraposition is equivalently proved: if $\mathcal{G}_{AFG} = \emptyset$, then the corresponding demand graph is acyclic, and optimal symmetric DoF value is $\frac{1}{K}$.

Given the proof of necessity and sufficiency, we complete the proof.

6.6 Summary

The TIM problem with transmitter cooperation (i.e., TIM-CoMP) where a subset of messages is routed to transmitters has been considered in this

chapter. Three interference management techniques built upon fractional graph coloring, interference alignment, and hypergraph covering have been proposed to exploit the benefits of both topological knowledge and transmitter cooperation, with which the achievable symmetric DoF are identified for a class of network topologies. Three outer bounds built upon the concepts of generator sequence, compound settings, and the equivalence to index coding are also derived to show the optimality of symmetric DoF for all the three-cell networks and the cyclic Wyner-type networks.

Yet, fundamental limits of transmitter cooperation in TIM settings are far less understood. Note also that transmitter cooperation only enables a subset of messages to be routed to the transmitters. A natural question then arises as to whether further transmitter cooperation (e.g., more messages shared among transmitters) offers more gain. In this regard, we give a conjecture in the negative.

Conjecture 6.1. *The transmitter cooperation with Transmitter i ($\forall$ i) only endowed with messages $W_{\mathcal{R}_i}$ is sufficient in the sense of DoF, and further transmitter cooperation does not offer more gains on DoF.*

The intuition is that, if Transmitter-i is not connected to Receiver-j, the presence of W_j at Transmitter-i does not increase DoF of W_j, and the transmission of W_j at Transmitter-i will interfere other receivers connected to Transmitter-i due to lack of channel knowledge. A future interesting direction would be to prove or disprove this conjecture, which will pave the way for other TIM-CoMP related problems.

Chapter 7

Conclusion and Future Work

Focusing on the channel uncertainty in the forms of delayed and sole topological feedback, this thesis has studied interference management of wireless networks with channel uncertainty from both signal processing and information theoretic perspectives, unveiling the fundamental limitation of channel uncertainty in terms of sum rate and degrees of freedom. These results also shed light on how to design robust precoders and transmission protocols to overcome the potential channel degradation due to channel uncertainty.

The first part of this thesis has focused on the channel uncertainty caused by feedback delay in the form of delayed CSIT. In Chapter 3, we considered throughput maximization of time i.i.d. MISO BC at finite SNR with delayed CSIT. We proposed a first construction for the precoders which reaches a useful trade-off between interference alignment and signal enhancement at finite SNR, allowing for significant throughput improvement in practical settings. In Chapter 4, we considered time-correlated MIMO networks (BC and IC) where the transmitter(s) has/have delayed and imperfect current CSI. The DoF regions for two-user broadcast and interference MIMO networks with general antenna configuration under such conditions have been fully characterized. Specifically, a simple unified framework has been proposed, allowing us to attain optimal DoF region for general antenna configurations and current CSIT qualities.

The focus of the second part was placed on the channel uncertainty due to the sole topological feedback under the framework of TIM settings. In Chapter 5, the optimal DoF of the multiple unicast TIM problem have been fully characterized via simple orthogonal schemes for a subclass of cellular network topologies. In particular, it was shown that orthogonal access achieves the optimal symmetric DoF, sum DoF, and DoF region of the cellular networks

without cycles of length larger than four in connectivity graphs. In Chapter 6, we considered the TIM setting where a typical transmitter cooperation is enabled. We showed that the sole topological information can be also exploited in this case to strictly improve DoF when the network is not fully connected. In particular, the symmetric DoF for regular networks were characterized as well as the conditions to achieve a certain amount of DoF for arbitrary network topologies were also identified.

While some interesting results have been established in the past three years in delayed feedback networks, there are still some interesting open problems left. Most existing works considered delayed CSIT with infinite preciseness by assuming that either the CSI feedback could be tolerant to sufficient delays or the feedback links are with infinite capacity. Nevertheless, the practical systems have neither unlimited delay tolerance nor unbounded capacity of feedback links. As such, considering delayed feedback in MIMO networks with finite precision is of practical relevance and interest. Further, the extension to the general K-user MIMO case is also an interesting open problem. Besides the lack of tight outer bound, a better achievability scheme is also needed to reduce the gap between the existing inner and outer bounds [80].

While DoF characterization made a first-order approximation of channel capacity in the infinite SNR regime, capacity characterization at any SNR is still missing. In Chapter 3, we made the first progress toward this target, followed by the recent progress in [131] which gave a constant gap of capacity approximation. Nevertheless, the problem of characterizing the sum capacity or the capacity region with both delayed and imperfect current CSIT is challenging and still open. Our previous results on DoF characterization in Chapter 4 shed light on the achievability scheme design, but the optimization over a number of parameters makes this problem challenging at this moment.

Regarding the topological interference management problems, there are still lots of interesting open problems for future work. TIM problems can be regarded as applications of index coding problems in wireless networks, such that the linear solutions of the former can be transferrable to the latter. As many researchers were trying to attack index coding problems from different perspectives, such as network coding [132], graph theory [99,100], interference alignment [49, 97, 98], linear programming [125], random coding [124], rate distortion theory [133], to name a few, the graph theoretic and interference alignment perspectives appear to be the most promising ones. While graph theoretic tools are specialized in describing general results with the aid of some well-defined graph parameters (such as fractional chromatic number, local chromatic number, stable set number), the interference alignment

perspective arms us the powerful lens to reveal the underlying insights and to allow for much more intuitively transparent statements. Remarkably, their interplay enables us to gain insights from the latter and leads us to general solutions from the former. Differently from the index coding problems where all the messages are encoded together, an additional constraint that the encoding of messages must be done in a distributed manner is imposed on TIM problems. So, how much index coding results can be translated into TIM problems is our future work.

Although some preliminary results have been reported in Chapter 6, the fundamental limits of transmitter cooperation under TIM settings are far less understood. The tighter relation between TIM-CoMP and index coding problems is still unclear. Interesting extensions including reducing the gap between inner and outer bounds, and fully characterizing the symmetric DoF of arbitrary topologies are still challenging. A closer examination of TIM-CoMP problem via combinatorial optimization [112] is a promising direction. Another research avenue is to investigate how much the DoF results can translate into practical scenarios. For example, the partial connectivity in practice is relevant to the thresholds distinguishing strong links from weak ones. Different threshold levels result in different topologies. As we know, a lower threshold level leads to a denser connectivity and hence offers limited degrees of freedom to play with topological information, while a higher threshold level renders SNR suffered although the flexibility of scheme design is increased with sparser connectivity. So, how to choose the optimal threshold level to achieve the capacity? Some preliminary simulation results was given in Chapter 5, but the theoretic analysis is still challenging and is our ongoing work.

Bibliography

[1] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[2] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge University Press, 2003.

[3] D. Gesbert, M. Shafi, D.-s. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas in Commun.*, vol. 21, no. 3, pp. 281–302, 2003.

[4] D. Gesbert, M. Kountouris, R. Heath, C. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Sep. 2007.

[5] G. Caire and S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691 – 1706, Jul. 2003.

[6] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

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

[8] S. A. Jafar, *Interference alignment: A new look at signal dimensions in a communication network*. Now Publishers Inc, 2011.

[9] T. Gou and S. A. Jafar, "Degrees of freedom of the K user MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6040–6057, 2010.

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

[11] C. Wang, H. Sun, and S. A. Jafar, "Genie chains: Exploring outer bounds on the degrees of freedom of MIMO interference networks," *arXiv preprint arXiv:1404.2258*, 2014.

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

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

[14] C. Suh and D. Tse, "Interference alignment for cellular networks," in *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*. IEEE, 2008, pp. 1037–1044.

[15] C. Suh, M. Ho, and D. N. Tse, "Downlink interference alignment," *IEEE Trans. Communications*, vol. 59, no. 9, pp. 2616–2626, 2011.

[16] V. Ntranos, M. A. Maddah-Ali, and G. Caire, "Cellular interference alignment," *arXiv preprint arXiv:1402.3119*, 2014.

[17] T. Gou, S. A. Jafar, C. Wang, S.-W. Jeon, and S.-Y. Chung, "Aligned interference neutralization and the degrees of freedom of the $2 \times 2 \times 2$ interference channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 4381–4395, 2012.

[18] N. Lee and C. Wang, "Aligned interference neutralization and the degrees of freedom of the two-user wireless networks with an instantaneous relay," *IEEE Trans. Communications,*, vol. PP, no. 99, pp. 1–9, 2013.

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

[20] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer

bounds on the capacity region," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3032–3050, 2009.

[21] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum–rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689–699, 2009.

[22] C. Geng, N. Naderializadeh, A. S. Avestimehr, and S. A. Jafar, "On the optimality of treating interference as noise," *arXiv preprint arXiv:1305.4610*, 2013.

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

[24] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas in Commun.*, vol. 26, no. 8, pp. 1341 –1365, Oct. 2008.

[25] A. Lozano, R. Heath, and J. Andrews, "Fundamental limits of cooperation," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5213–5226, 2013.

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

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

[28] S. Jafar and M. Fakhereddin, "Degrees of freedom for the MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2637–2642, July 2007.

[29] S. A. Jafar and A. J. Goldsmith, "Isotropic fading vector broadcast channels: The scalar upper bound and loss in degrees of freedom," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 848–857, 2005.

[30] C. Vaze and M. Varanasi, "The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5354–5374, Aug. 2012.

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

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

[33] S. Jafar, "Blind interference alignment," *IEEE J. Sel. Topics in Signal Processing*, vol. 6, no. 3, pp. 216–227, 2012.

[34] A. Lapidoth, S. Shamai, and M. Wigger, "On the capacity of fading MIMO broadcast channels with imperfect transmitter side-information," *arXiv preprint cs/0605079*, 2006.

[35] T. Gou, S. Jafar, and C. Wang, "On the degrees of freedom of finite state compound wireless networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3286–3308, Jun. 2011.

[36] M. A. Maddah-Ali, "On the degrees of freedom of the compound MISO broadcast channels with finite states," in *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*. IEEE, 2010, pp. 2273–2277.

[37] A. G. Davoodi and S. A. Jafar, "Aligned image sets under channel uncertainty: Settling a conjecture by Lapidoth, Shamai and Wigger on the collapse of degrees of freedom under finite precision CSIT," *arXiv preprint arXiv:1403.1541*, 2014.

[38] G. Caire, N. Jindal, and S. Shamai, "On the required accuracy of transmitter channel state information in multiple antenna broadcast channels," in *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*. IEEE, 2007, pp. 287–291.

[39] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[40] H. Bolcskei and I. Thukral, "Interference alignment with limited feedback," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*. IEEE, 2009, pp. 1759–1763.

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

[42] H. Maleki, S. Jafar, and S. Shamai, "Retrospective interference alignment over interference networks," *IEEE J. Sel. Topics in Signal Processing*, vol. 6, no. 3, pp. 228–240, 2012.

[43] C. Vaze and M. Varanasi, "The degrees of freedom region and interference alignment for the MIMO interference channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4396–4417, Jul. 2012.

[44] M. Abdoli, A. Ghasemi, and A. Khandani, "On the degrees of freedom of K-user SISO interference and X channels with delayed CSIT," *arXiv preprint arXiv:1109.4314*, 2011.

[45] A. Ghasemi, A. Motahari, and A. Khandani, "On the degrees of freedom of X channel with delayed CSIT," in *2011 IEEE International Symposium on Information Theory Proceedings (ISIT)*, 31 2011-Aug. 5 2011, pp. 767–770.

[46] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Capacity results for binary fading interference channels with delayed CSIT," *arXiv preprint arXiv:1301.5309*, 2013.

[47] S. Yang, M. Kobayashi, P. Piantanida, and S. Shamai Shitz, "Secrecy degrees of freedom of MIMO broadcast channels with delayed CSIT," *Information Theory, IEEE Transactions on*, vol. 59, no. 9, pp. 5244–5256, 2013.

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

[49] S. A. Jafar, "Topological interference management through index coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 529–568, Jan. 2014.

[50] ——, "Elements of cellular blind interference alignment—aligned frequency reuse, wireless index coding and interference diversity," *arXiv preprint arXiv:1203.2384*, 2012.

[51] N. Naderializadeh and A. S. Avestimehr, "Interference networks with no CSIT: Impact of topology," *submitted to IEEE Trans. Inf. Theory, arXiv:1302.0296*, Feb. 2013.

[52] Z. Bar-Yossef, Y. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1479–1494, March 2011.

[53] M. Kobayashi, S. Yang, D. Gesbert, and X. Yi, "On the degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, 2012, pp. 2501–2505.

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

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

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

[57] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "Interference alignment for the MIMO interference channel with delayed local CSIT," *arXiv preprint arXiv:1102.5673*, 2011.

[58] J. Xu, J. Andrews, and S. Jafar, "MISO broadcast channels with delayed finite-rate feedback: Predict or observe?" *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1456 –1467, Apr. 2012.

[59] A. Adhikary, H. Papadopoulos, S. Ramprashad, and G. Caire, "Multi-user MIMO with outdated CSI: training, feedback and scheduling," in *Proc. Allerton 2011*, 2011.

[60] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3309–3322, 2011.

[61] S. W. Peters and R. W. Heath, "Interference alignment via alternating minimization," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 2445–2448.

[62] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2009.

[63] D. P. Palomar and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 141–154, 2006.

[64] A. C. Aitken, "Determinants and matrices," 1958.

[65] Z. K. M. Ho and D. Gesbert, "Balancing egoism and altruism on interference channel: The MIMO case," in *Communications (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.

[66] D. A. Schmidt, W. Utschick, and M. L. Honig, "Beamforming techniques for single-beam MIMO interference networks," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*. IEEE, 2010, pp. 1182–1187.

[67] S. W. Peters and R. W. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Vehicular Technology*, vol. 60, no. 1, pp. 206–218, 2011.

[68] P. Cao, E. Jorswieck, and S. Shi, "Pareto boundary of the rate region for single-stream MIMO interference channels: Linear transceiver design," *IEEE Trans. Signal Processing*, vol. 61, no. 20, pp. 4907–4922, Oct 2013.

[69] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Communications Letters*, vol. 5, no. 9, pp. 369–371, 2001.

[70] C. Xiao, J. Wu, S.-Y. Leong, Y. R. Zheng, and K. B. Letaief, "A discrete-time model for triply selective MIMO rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1678–1688, 2004.

[71] X. Yi and D. Gesbert, "Precoding on the broadcast MIMO channel with delayed CSIT: The finite SNR case," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, 2012, pp. 2933–2936.

[72] ——, "Precoding methods for the MISO broadcast channel with delayed CSIT," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 1–11, May 2013.

[73] J. Wang, M. Matthaiou, S. Jin, and X. Gao, "Precoder design for multiuser miso systems exploiting statistical and outdated csit," *IEEE Trans. Communications*, vol. 61, no. 11, pp. 4551–4564, November 2013.

[74] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2740–2746, 2007.

[75] C. S. Vaze and M. K. Varanasi, "The degrees of freedom regions of two-user and certain three-user MIMO broadcast channels with delayed CSIT," *ArXiv:1101.0306*, Dec. 2011.

[76] N. Lee and R. Heath Jr, "Not too delayed CSIT achieves the optimal degrees of freedom," in *Proc. Allerton 2012, arXiv:1207.2211*, 2012.

[77] C. Hao and B. Clerckx, "Degrees-of-freedom region of time correlated MISO broadcast channel with perfect delayed CSIT and asymmetric partial current CSIT," *arXiv preprint arXiv:1211.4381*, 2012.

[78] Y. Lejosne, D. Slock, and Y. Yuan-Wu, "Degrees of freedom in the MISO BC with delayed-CSIT and finite coherence time: A simple optimal scheme," in *2012 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC)*. IEEE, 2012, pp. 180–185.

[79] X. Yi, S. Yang, D. Gesbert, and M. Kobayashi, "The degrees of freedom region of temporally-correlated MIMO networks with delayed CSIT," *arXiv:1211.3322v1*, Oct. 2012.

[80] P. de Kerret, X. Yi, and D. Gesbert, "On the degrees of freedom of the k-user time correlated broadcast channel with delayed csit," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, 2013, pp. 624–628.

[81] J. Chen and P. Elia, "Degrees-of-freedom region of the MISO broadcast channel with general mixed-CSIT," *ITA 2013, arXiv:1205.3474*, 2012.

[82] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, 1979.

[83] L. H. Ozarow and S. K. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 667–671, 1984.

[84] C. Suh and D. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2667 –2685, May 2011.

[85] F. Willems and E. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. 31, no. 3, pp. 313 – 327, May 1985.

[86] R. Tandon, M. Maddah-Ali, A. Tulino, H. Poor, and S. Shamai, "On fading broadcast channels with partial channel state information at the transmitter," in *2012 International Symposium on Wireless Communication Systems (ISWCS)*, 2012, pp. 1004–1008.

[87] K. Mohanty, C. S. Vaze, and M. K. Varanasi, "The degrees of freedom region for the MIMO interference channel with hybrid CSIT," *arXiv preprint arXiv:1209.0047*, 2012.

[88] J. Chen and P. Elia, "Toward the performance versus feedback tradeoff for the two-user MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8336–8356, 2013.

[89] T. M. Cover and J. A. Thomas, *Elements of information theory*. Wiley-interscience, 2006.

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

[91] H. Weingarten, T. Liu, S. Shamai, Y. Steinberg, and P. Viswanath, "The capacity region of the degraded multiple-input multiple-output compound broadcast channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5011–5023, 2009.

[92] J. Chen, S. Yang, and P. Elia, "On the fundamental feedback-vs-performance tradeoff over the MISO-BC with imperfect and delayed CSIT," *ISIT 2013, arXiv preprint arXiv:1302.0806*, 2013.

[93] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 2012.

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

[95] G. Caire, S. A. Ramprashad, and H. C. Papadopoulos, "Rethinking network MIMO: Cost of CSIT, performance analysis, and architecture comparisons," in *ITA2010*, 2010.

[96] Y. Birk and T. Kol, "Informed-source coding-on-demand (ISCOD) over broadcast channels," in *Proc. INFOCOM'98, Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3, 1998, pp. 1257–1264.

[97] H. Maleki, V. Cadambe, and S. Jafar, "Index coding: An interference alignment perspective," in *2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2012, pp. 2236–2240.

[98] H. Sun and S. A. Jafar, "Index coding capacity: How far can one go with only Shannon inequalities?" *arXiv preprint arXiv:1303.7000*, 2013.

[99] K. Shanmugam, A. G. Dimakis, and M. Langberg, "Local graph coloring and index coding," in *2013 IEEE International Symposium on Information Theory Proceedings (ISIT)*. IEEE, 2013, pp. 1152–1156.

[100] ——, "Graph theory versus minimum rank for index coding," in *2014 IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2014.

[101] H. Sun, C. Geng, and S. A. Jafar, "Topological interference management with alternating connectivity," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2013.

[102] S. Gherekhloo, A. Chaaban, and A. Sezgin, "Resolving entanglements in topological interference management with alternating connectivity," *arXiv preprint arXiv:1404.7827*, 2014.

[103] H. Sun and S. A. Jafar, "Topological interference management with multiple antennas," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2014.

[104] H. Maleki and S. Jafar, "Optimality of orthogonal access for one-dimensional convex cellular networks," *IEEE Communications Letters*, vol. 17, no. 9, pp. 1770–1773, Sept. 2013.

[105] J. A. Bondy and U. S. R. Murty, *Graph theory with applications*. Macmillan London, 1976, vol. 290.

[106] D. B. West *et al.*, *Introduction to graph theory*. Prentice hall Englewood Cliffs, 2001, vol. 2.

[107] E. R. Scheinerman and D. H. Ullman, *Fractional graph theory*. Dover-Publications. com, 2011.

[108] M. C. Golumbic, "Algorithmic graph theory and perfect graphs," *Annals of Discrete Mathematics*, vol. 57, 2004.

[109] M. C. Golumbic and C. F. Goss, "Perfect elimination and chordal bipartite graphs," *Journal of Graph Theory*, vol. 2, no. 2, pp. 155–163, 1978.

[110] R. B. Hayward, "Weakly triangulated graphs," *Journal of Combinatorial Theory, Series B*, vol. 39, no. 3, pp. 200–208, 1985.

[111] A. Berry, J.-P. Bordat, and P. Heggernes, "Recognizing weakly triangulated graphs by edge separability," *Nordic Journal of Computing*, vol. 7, no. 3, pp. 164–177, 2000.

[112] A. Schrijver, *Combinatorial optimization: polyhedra and efficiency*. Springer, 2003, vol. 24.

[113] V. Chvátal, "On certain polytopes associated with graphs," *Journal of Combinatorial Theory, Series B*, vol. 18, no. 2, pp. 138–154, 1975.

[114] K. Cameron, R. Sritharan, and Y. Tang, "Finding a maximum induced matching in weakly chordal graphs," *Discrete Mathematics*, vol. 266, no. 1, pp. 133–142, 2003.

[115] K. Cameron, "Induced matchings," *Discrete Applied Mathematics*, vol. 24, no. 1, pp. 97–102, 1989.

[116] R. Hayward, C. Hoàng, and F. Maffray, "Optimizing weakly triangulated graphs," *Graphs and Combinatorics*, vol. 5, no. 1, pp. 339–349, 1989.

[117] R. B. Hayward, J. P. Spinrad, and R. Sritharan, "Improved algorithms for weakly chordal graphs," *ACM Transactions on Algorithms (TALG)*, vol. 3, no. 2, p. 14, 2007.

[118] R. B. Hayward, J. Spinrad, and R. Sritharan, "Weakly chordal graph algorithms via handles," in *Proceedings of the eleventh annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 2000, pp. 42–49.

[119] C. T. Hoàng and B. A. Reed, "Some classes of perfectly orderable graphs," *Journal of Graph Theory*, vol. 13, no. 4, pp. 445–463, 1989.

[120] R. B. Hayward, "Meyniel weakly triangulated graphs—I: co-perfect orderability," *Discrete Applied Mathematics*, vol. 73, no. 3, pp. 199–210, 1997.

[121] T. Gou, C. da Silva, J. Lee, and I. Kang, "Partially connected interference networks with no CSIT: Symmetric degrees of freedom and multicast across alignment blocks," *IEEE Communications Letters*, vol. 17, no. 10, pp. 1893–1896, 2013.

[122] S. Gherekhloo, A. Chaaban, and A. Sezgin, "Topological interference management with alternating connectivity: The Wyner-type three user interference channel," *arXiv preprint arXiv:1310.2385*, 2013.

[123] C. Geng, H. Sun, and S. A. Jafar, "Multilevel topological interference management," in *Information Theory Workshop (ITW), 2013 IEEE*, 2013, pp. 1–5.

[124] F. Arbabjolfaei, B. Bandemer, Y.-H. Kim, E. Sasoglu, and L. Wang, "On the capacity region for index coding," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, July 2013, pp. 962–966.

[125] A. Blasiak, R. Kleinberg, and E. Lubetzky, "Index coding via linear programming," *arXiv preprint arXiv:1004.1379*, 2010.

[126] G. Li and R. Simha, "The partition coloring problem and its application to wavelength routing and assignment," in *Proceedings of the First Workshop on Optical Networks*. Citeseer, 2000, p. 1.

[127] M. Demange, T. Ekim, B. Ries, and C. Tanasescu, "On some applications of the selective graph coloring problem," *European Journal of Operational Research*, 2014.

[128] M. Neely, A. Tehrani, and Z. Zhang, "Dynamic index coding for wireless broadcast networks," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7525–7540, Nov. 2013.

[129] V. Aggarwal, A. Avestimehr, and A. Sabharwal, "On achieving local view capacity via maximal independent graph scheduling," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2711–2729, 2011.

[130] R. Koetter, M. Effros, and M. Medard, "A theory of network equivalence – part I: Point-to-point channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 972–995, Feb 2011.

[131] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Approximate capacity of the two-user MISO broadcast channel with delayed CSIT," *arXiv preprint arXiv:1405.1143*, 2014.

[132] M. Effros, S. E. Rouayheb, and M. Langberg, "An equivalence between network coding and index coding," *arXiv preprint arXiv:1211.6660*, 2012.

[133] S. Unal and A. B. Wagner, "A rate-distortion approach to index coding," in *Information Theory and Applications Workshop (ITA), 2014*. IEEE, 2014, pp. 1–5.