Non-coherent wireless communications : fundamental limits and system design

Khac Hoang Ngo

université
PARIS-SACLAY

Noncoherent Wireless Communications: Fundamental Limits and System Design

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 580 Sciences et Technologies de l'Information et de la Communication (STIC)
Spécialité de doctorat: Réseaux, Information et Communications
Unité de recherche: Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, 91190, Gif-sur-Yvette, France
Référent: CentraleSupélec

Thèse présentée et soutenue à Gif-sur-Yvette, France, le 30 Juin 2020, par

Khac-Hoang NGO

Composition du jury:

Philippe CIBLAT Professeur, Telecom ParisTech, France	Président
Giuseppe DURISI Professeur, Université de Technologie de Chalmers, Suède	Rapporteur & Examinateur
Stefan M. MOSER Professeur, ETH Zürich, Suisse	Rapporteur & Examinateur
Iryna ANDRIYANOVA Professeur, Université de Cergy-Pontoise, France	Examinatrice
Marco DI RENZO Directeur de Recherche, L2S, CNRS, France	Examinateur
Petar POPOVSKI Professeur, Université d'Aalborg, Danemark	Examinateur
Sheng YANG Professeur, CentraleSupélec, France	Directeur de thèse
Maxime GUILLAUD Chercheur Principal, Huawei Technologies, France	Co-directeur de thèse

Thèse de doctorat

NNT: 2020UPASC031

université PARIS-SACLAY

CentraleSupélec

NNT : 2020UPASC031

Noncoherent Wireless Communications: Fundamental Limits and System Design

THESIS

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
from University of Paris-Saclay
prepared at CentraleSupélec

Doctoral School n°580 Information and Communication Sciences and Technologies (STIC)
Specialization: Networks, Information and Communications

Thesis presented and defended in Gif-sur-Yvette, France, on 30 June 2020, by

Khac-Hoang NGO

Composition of the Jury :

Philippe CIBLAT Professor, Telecom ParisTech, France	President
Giuseppe DURISI Professor, Chalmers University of Technology, Sweden	Reviewer & Examiner
Stefan M. MOSER Professor, ETH Zürich, Switzerland	Reviewer & Examiner
Iryna ANDRIYANOVA Professor, University of Cergy-Pontoise, France	Examiner
Marco DI RENZO Research Director, L2S, CNRS, France	Examiner
Petar POPOVSKI Professor, Aalborg University, Denmark	Examiner
Sheng YANG Professor, CentraleSupélec, France	Thesis supervisor
Maxime GUILLAUD Principal Researcher, Huawei Technologies, France	Thesis co-supervisor

To my family and Thao...

"*You see, wire telegraph is a kind of a very, very long cat. You pull his tail in New York and his head is meowing in Los Angeles. Do you understand this? And radio operates exactly the same way: you send signals here, they receive them there. The only difference is that there is no cat.*"

Albert Einstein

"*We live on an island surrounded by a sea of ignorance. As our island of knowledge grows, so does the shore of our ignorance.*"

John Archibald Wheeler

"*But the absence of fighting or hatred or desire also means the opposites do not exist either. No joy, no communion, no love. Only where there is disillusionment and depression and sorrow does happiness arise; without the despair of loss, there is no hope.*"

Haruki Murakami

Acknowledgements

On my adventures in research, I have the privilege to meet, interact with and learn from many wonderful people. I wish to fully appreciate their help, and for the space limit, particularly mention a few of them.

First, I would like to express my deepest gratitude to my supervisors Prof. Sheng Yang and Dr. Maxime Guillaud. Their patience, timely comments, and continuous guidance and support have made this thesis possible. In particular, Prof. Sheng Yang guided me through my first research papers since my master, and taught me to try to pin down the essence of a problem through an as-simple-as-possible, but not simpler, example, without getting lost in a maze of details. Dr. Maxime Guillaud showed me the benefit of scientific interaction between individuals with diverse but complementary expertise and perspectives. He has always put me the position where I could make the most of the working environment. Working with them, I not only learn to be a decent researcher, but also grow up in many other aspects. Thank you so much, Sheng and Maxime!

I am thankful to Prof. Philippe Ciblat, Prof. Giuseppe Durisi, Prof. Stefan Moser, Prof. Iryna Andriyanova, Prof. Marco Di Renzo, and Prof. Petar Popovski for serving in my committee, reviewing the first version of my thesis, giving me insightful comments and suggestions, and raising many interesting questions and remarks during my defense.

Many thanks go to my collaborators whose joint effort produced many results in the thesis: Prof. Philip Schniter and Subrata Sarkar from the Ohio State University, Prof. Aria Nosratinia and Fan Zhang from University of Texas at Dallas, and especially my colleague Alexis Decurninge in Huawei Paris Research Center. Discussions with and helps from other colleagues in Huawei, including Paul Ferrand, Ingmar Land, Valerio Bioglio, and Prof. Jean-Claude Belfiore, have been very useful.

I would also like to extend my warmest thanks to my colleagues and friends in both Huawei and CentraleSupélec for maintaining a very enjoyable atmosphere, in particular: Van-Minh Nguyen, Khoa Le-Trung, Quoc-Thai Nguyen, Luis G. Ordóñez, Marco Maso in Huawei and Richard Combes in CentraleSupélec. I have shared a good time in and out of the office with my fellow Ph.D. students, including Arvid Jakobsson, Juan Carlos Bucheli, Filip Pawlowski, Thibaut Tachon, Pierre Leca, Pierre Escamilla, Anthony Palmieri, Apostolos Avranas, Charles Pillet, and Yigit Ugur in Huawei; Zheng Li, Asma Ghorbel, Qifa Yan (office-mates), Chao He, Hafiz Tiomoko Ali, Salah Eddine Hajri, Matha Deghel, and Clément Feutry in CentraleSupélec. I would like to acknowledge the help of the staffs Huu-Hung Vuong, Catherine Magnet, Jose Fonseca, Anne Batalie in CentraleSupélec and Zhenrong Liu in Huawei.

My life in France has been pleasant thanks to my Vietnamese seniors and friends: Van-Ly Nguyen, Duc-Tuyen Ta, Viet-Dung Nguyen, Thi-Hao Nguyen, Duy-Cuong Pham, Quang-Linh Nguyen, Duc-Nghia Tran, Lynh Hoang, Quan Do, Quang-Trung Luu, Minh-Tri Nguyen, Quang-Binh Ngo, Tan-Quang Duong, Lam-Thanh Tu, Trung-Dung Le, Thi-Huong Au, Hoang-Phuong Thai, Ngoc-Diep Lai, Loan Truong, Toan Pham, Quang-Huy Trinh, Viet-Bac Le, and many others. I am deeply grateful to Mme. Raymonde Huguet for her kindness and generosity towards me and many other Vietnamese.

In retrospect, I am indebted to all the people who got me engaged in research and looked after my first steps in this endeavour. Special thanks to Prof. Aaron Danner, who hosted me in 2012 in the National University of Singapore for my very first experience in an international research environment. I would like to thank Prof. Linh-Trung Nguyen

and Prof. Quoc-Tuan Nguyen for supervising my bachelor thesis. Prof. Linh-Trung Nguyen passed on to me the lesson: never avoid the difficulty in research, keep going straight and raze it. I also very much appreciate the guidance of Prof. Mari Kobayashi during my master internship.

Last but not least, I wish to express my heartfelt gratitude to my family members, who have always been encouraging me to pursue my passion, trusting in me, and showing me the value of knowledge and diligence. Sweetest thanks and a huge hug to my girlfriend Thao for her love and patience.

Thank you! / Merci! / Cảm ơn!

Khac-Hoang Ngo

Résumé

Dans les communications sans fil sur des canaux à évanouissement, en particulier à antennes multiples, la connaissance instantanée des coefficients de canal, appelés informations d'état de canal (CSI), est essentielle car elle permet d'adapter la transmission et la réception aux conditions actuelles du canal. La communication avec CSI *a priori* au niveau du récepteur est dite *cohérente*. En pratique, cependant, le CSI n'est pas disponible avant la communication et doit être estimé à un coût qui ne doit pas être ignoré, en particulier dans un environnement hautement mobile. Ainsi, la communication sans CSI *a priori*, également appelée communication *noncohérente*, est un cadre plus pratique et général. Cette thèse contribue à la compréhension des limites théoriques des communications noncohérentes, ainsi qu'à la conception d'un système de communication pratique noncohérent à évanouissement par bloc. Nous considérons trois scénarios: le canal point à point (P2P), le canal à accès multiple (MAC) et le canal de diffusion (BC).

Dans la première partie, nous étudions les limites fondamentales des communications noncohérentes en termes de débit de données et de degrés de liberté (DoF) réalisables. Nous considérons un évanouissement par bloc générique dans lequel le canal a une entropie différentielle finie et un second moment fini. Nous établissons d'abord le DoF optimal pour le canal P2P noncohérent à entrées multiples et sorties multiples (MIMO) en utilisant l'approche de la dualité pour borner les informations mutuelles. Deuxièmement, en utilisant une approche de dualité similaire, nous dérivons la région de DoF optimale pour le MAC SIMO à deux utilisateurs, qui peut être obtenue par partage de temps entre des schémas à pilotes simples. Troisième, nous dérivons les régions débit et DoF réalisable pour le BC MIMO noncohérent avec un évanouissement spatialement corrélé en exploitant la diversité de corrélation de transmission, qui est la différence entre la corrélation subie par différents utilisateurs. Ce faisant, nous concevons soigneusement des schémas de transmission basés sur des pilotes et sur le partage de débit, la superposition de produits et une combinaison de ceux-ci pour transmettre efficacement des signaux dans les parties communes et mutuellement exclusives des sous-espaces de corrélation.

Dans la deuxième partie, nous concevons la constellation et les schémas de détection efficaces pour les communications noncohérentes sur le canal d'évanouissement de type Rayleigh par bloc. Premièrement, nous proposons une constellation Grassmannienne structurée pour le canal P2P SIMO qui est simple à générer, a une efficacité d'empilement élevée, admet un étiquetage binaire simple mais efficace et permet une détection efficace douce et dure. Deuxièmement, nous étudions la conception de constellation conjointe pour le MAC MIMO. Nous introduisons des critères de conception simples et efficaces afin de minimiser l'erreur de détection conjointe et proposons quelques constructions de constellation simples. Troisièmement, nous proposons un schéma de détection souple multi-utilisateurs noncohérent pour le MAC SIMO à l'évanouissement de Rayleigh corrélé spatialement basé sur l'inférence approximative par propagation d'espérance. Ce schéma présente une complexité polynomiale dans la dimension du canal tout en produisant des marginaux postérieurs approximatifs par utilisateur précis conduisant à des performances d'erreur quasi-optimales.

Mots clés: communications noncohérentes, évanouissement par bloc, information sur l'état du canaux, degrés de liberté, constellations Grassmanniennes, détection multi-utilisateurs

Abstract

In wireless communication over fading channels, especially multiple-antenna communication, the instantaneous knowledge of channel coefficients, so-called channel state information (CSI), is critical because it enables to adapt the transmission and reception to current channel conditions. The communication with *a priori* CSI at the receiver is said to be *coherent*. In practice, however, CSI is not granted for free prior to communication and needs to be estimated at a cost that should not be ignored, especially in a highly mobile environment. Thus, communication without *a priori* CSI, also known as *noncoherent* communication, is a more practical and general framework. This thesis contributes to the understanding of the theoretical limits of noncoherent communications, as well as the design of a practical noncoherent communication system in block fading. We consider three scenarios: the point-to-point (P2P) channel, the multiple-access channel (MAC), and the broadcast channel (BC).

In the first part, we study the fundamental limits of noncoherent communications in terms of achievable data rate and degrees of freedom (DoF). We consider generic block fading in which the channel has finite differential entropy and finite second moment. First, we derive the optimal DoF for the noncoherent multiple-input multiple-output (MIMO) P2P channel by using the duality approach to bound the input-output mutual information. Second, using a similar duality approach, we derive the optimal DoF region for the two-user noncoherent single-input multiple-output (SIMO) MAC, which can be achieved by time sharing between simple pilot-based schemes. Third, we derive achievable rate and DoF regions for the noncoherent MIMO BC with spatially correlated fading by exploiting the transmit correlation diversity, which is the difference between the correlation experienced by different users. In doing so, we carefully design pilot-based transmission schemes based on rate splitting, product superposition, and a combination of them to effectively transmit signals in both the common and mutually exclusive parts of the correlation subspaces.

In the second part, we design the constellation and efficient detection schemes for noncoherent communications over Rayleigh block fading channel. First, we propose a structured Grassmannian constellation for the SIMO P2P channel that is simple to generate, has high packing efficiency, admits a simple yet effective binary labeling, and allows for efficient soft and hard detection. Second, we investigate joint constellation design for the MIMO MAC. We introduce some simple and effective design criteria so as to minimize the joint detection error, and propose some simple constellation constructions. Third, we propose a noncoherent multi-user soft detection scheme for the SIMO MAC in spatially correlated Rayleigh fading based on expectation propagation approximate inference. This scheme has polynomial complexity in the channel dimension while producing accurate approximate per-user posterior marginals leading to near-optimal error performance.

Keywords: noncoherent communications, block fading, channel state information, degrees of freedom, Grassmannian constellations, multi-user detection

Contents

Acronyms

The meaning of an acronym is usually indicated once, when it first occurs in the text. The English acronyms are also used for the French summary.

AWGN	additive white Gaussian noise
BC	broadcast channel
BER	bit error rate
bpcu	bits per channel use
BSTM	Beta-variate space-time modulation
CDF	commulative distribution function
CDI	channel distribution information
CDMA	code-division multiple access
CSI	channel state information
CSIR	channel state information at the receiver
CSIT	channel state information at the transmitter
DoF	degrees of freedom
EP	expectation propagation
EPAK	expectation propagation with approximate Kronecker
FDMA	frequency-division multiple access
GMI	generalized mutual information
IID	independent and identically distributed
JSDM	joint spatial division multiplexing
KL	Kullback-Leibler
LLR	log-likelihood ratio
LS	least-squares
LTE	long-term evolution
MAC	multiple-access channel
MGF	moment generation function
MIMO	multiple-input multiple-output
MISO	multiple-input single-output

ML	maximum likelihood
MLC	multilevel coding
MMSE	minimum mean square error
MSD	multistage decoding
MSE	mean square error
NND	nearest-neighbor decoding
NP	non-deterministic polynomial-time
OFDM	orthogonal frequency division multiplexing
P2P	point-to-point
PAM	pulse amplitude modulation
PDF	probability density function
PEP	pairwise error probability
PMF	probability mass function
POCIS	Projection onto the Orthogonal Complement of the Interference Subspace
PSK	phase shift keying
QAM	quadrature amplitude modulation
RWBS	repeated weighted boosting search
SER	symbol error rate
SESD	Schnorr-Euchner sphere decoder
SIA	successive interference approximation
SIMO	single-input multiple-output
SISO	single-input single-output
SNR	signal-to-noise ratio
SVD	singular value decomposition
TDM	time division multiplexing
TDMA	time-division multiple access
URLLC	ultra-reliable low-latency communication
USTM	unitary space-time modulation
w.l.o.g.	without loss of generality
w.r.t.	with respect to
ZF	zero forcing

Notation

In general, we denote scalars with normal-font letters, e.g., x, vectors with lower-case bold-font letters, e.g., $\boldsymbol{x}$, and matrices with upper-case bold-font letters, e.g., $\boldsymbol{X}$. For random quantities, we use nonitalic letters with sans-serif fonts, e.g., a scalar x, a vector $\mathbf{v}$, and a matrix $\mathbf{M}$. Deterministic quantities are denoted with italic letters, e.g., a scalar x, a vector $\boldsymbol{v}$, and a matrix $\boldsymbol{M}$. We adopt the column convention for vectors.

Linear Algebra

$\|\boldsymbol{v}\|$	Euclidean norm of a vector $\boldsymbol{v}$
$\|\boldsymbol{M}\|_{\mathrm{F}}$	Frobenius norm of a matrix $\boldsymbol{M}$
$\boldsymbol{M}^*$	conjugate of $\boldsymbol{M}$
$\boldsymbol{M}^{\mathsf{T}}$ (resp. $\boldsymbol{M}^{-\mathsf{T}}$)	transpose of $\boldsymbol{M}$ (resp. $\boldsymbol{M}^{-1}$)
$\boldsymbol{M}^{\mathsf{H}}$ (resp. $\boldsymbol{M}^{-\mathsf{H}}$)	conjugate transpose of $\boldsymbol{M}$ (resp. $\boldsymbol{M}^{-1}$)
$\boldsymbol{M}^\dagger$	Moore-Penrose (pseudo) inverse of $\boldsymbol{M}$, i.e., $\boldsymbol{M}^\dagger = (\boldsymbol{M}^{\mathsf{H}}\boldsymbol{M})^{-1}\boldsymbol{M}^{\mathsf{H}}$
$\mathrm{tr}(\boldsymbol{M})$	trace of $\boldsymbol{M}$
$\mathrm{rank}(\boldsymbol{M})$	rank of $\boldsymbol{M}$
$\det(\boldsymbol{M})$	determinant of $\boldsymbol{M}$
$\sigma_i(\boldsymbol{M})$	the i-th eigenvalue of a matrix $\boldsymbol{M}$ in decreasing order, unless otherwise specified
$\boldsymbol{M} \succeq 0$	$\boldsymbol{M}$ is positive semi-definite
$\boldsymbol{M}_1 \succeq \boldsymbol{M}_2$ or $\boldsymbol{M}_2 \preceq \boldsymbol{M}_1$	$\boldsymbol{M}_1 - \boldsymbol{M}_2 \succeq 0$
$\lvert\boldsymbol{M}\rvert$	a matrix having the same singular vectors as $\boldsymbol{M}$ and singular values as the absolute value of the singular values of $\boldsymbol{M}$, i.e., if $\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathsf{H}}$, then $\lvert\boldsymbol{M}\rvert = \boldsymbol{U}\tilde{\boldsymbol{\Sigma}}\boldsymbol{V}^{\mathsf{H}}$ where $\tilde{\boldsymbol{\Sigma}}$ is the component-wise absolute value of $\boldsymbol{\Sigma}$
$[\boldsymbol{M}]_{i,j}$	(i,j)-th element of $\boldsymbol{M}$
$\boldsymbol{M}_{[i]}$	i-th column of $\boldsymbol{M}$
$\boldsymbol{M}_{[i:j]}$	sub-matrix containing the i-th to j-th columns of $\boldsymbol{M}$
$\mathrm{diag}(x_1, \ldots, x_N)$	diagonal matrix with diagonal entries $x_1, \ldots, x_N$
$\mathrm{Span}(\boldsymbol{U})$	the column space of a "tall" matrix $\boldsymbol{U}$
$\boldsymbol{X} \otimes \boldsymbol{Y}$	Kronecker product of $\boldsymbol{X}$ and $\boldsymbol{Y}$
$\boldsymbol{I}_m$	$m \times m$ identity matrix (the dimension m is omitted when confusion is not likely)
$\boldsymbol{e}_i$	i-th column of $\boldsymbol{I}$, i.e., the canonical basis vector with 1 at position i and 0 elsewhere, where the dimension will be specified

$\mathbf{0}_m$ (resp. $\mathbf{1}_m$)	$m \times 1$ all-zero (resp. all-one) vector (the dimension m is omitted when confusion is not likely)
$\mathbf{0}_{m\times n}$ (resp. $\mathbf{1}_{m\times n}$)	$m \times n$ all-zero (resp. all-one) matrix (the dimensions m and n are omitted when confusion is not likely)

Analysis

$\mathbb{R}$ (resp. $\mathbb{C}$)	set of real (resp. complex) numbers
$\jmath = \sqrt{-1}$	imaginary unit
$\mathrm{Re}(z)$ (resp. $\mathrm{Im}(z)$)	real (resp. imaginary) part of $z \in \mathbb{C}$
$[n]$	set of integers $\{1, 2, \ldots, n\}$
$\lvert x \rvert$	absolute value of x
$\propto$	proportional to
$\prod, \times$	the conventional or Cartesian product (depending on the factors)
$A := B$ or $B =: A$	A is defined by B
$\mathbb{1}\{A\} := \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{if } A \text{ is false} \end{cases}$	indicator function
$\delta(t)$	$\mathbb{1}\{t = 0\}$
$\binom{n}{k} := \frac{n!}{k!(n-k)!}\mathbb{1}\{n \geq k\}$	binomial coefficient, i.e., the number of k-element subsets of an n-element set [1, (3.1.2)]
$f^{-1}(\cdot)$	the inverse of a function $f(\cdot)$
$\gamma \approx 0.5772$	Euler's constant [1, (6.1.3)]
$e = \sum_{n=1}^{\infty} \frac{1}{n!} \approx 2.71828$	Euler's number [1, (4.1.17)]
$(x)^+$	$\max\{x, 0\}$
$\exp(\cdot) \equiv e^{(\cdot)} := \sum_{k=0}^{\infty} \frac{1}{k!}(\cdot)^k$	the exponential function where the argument is a scalar or a square matrix [1, (4.2.1)]
$\ln(x) := \int_1^z \frac{\mathrm{d}t}{t}$	natural (base-e) logarithm of x [1, (4.1.1)]
$\log_a(x) := \frac{\ln x}{\ln a}$	base-a logarithm of x; a is omitted if the base is immaterial
$\log^+(x)$	$\log(\max\{x, 1\})$
$\psi(k) := -\gamma + \sum_{n=1}^{k-1} \frac{1}{n}$	Euler's (digamma) psi-function [1, (6.3.2)]
$\Gamma(x) := \int_0^{\infty} z^{x-1}e^{-z}\,\mathrm{d}z,\ x > 0$	the Gamma function [1, (6.1.1)]
$\Gamma(x, y) := \int_y^{\infty} z^{x-1}e^{-z}\,\mathrm{d}z,\ x > 0, y \geq 0$	the upper incomplete Gamma function [1, (6.5.3)]
$\gamma(x, y) := \int_0^{y} z^{x-1}e^{-z}\,\mathrm{d}z,\ x > 0, y \geq 0$	the lower incomplete Gamma function [1, (6.5.2)]
$\Gamma_m(a) := \pi^{m(m-1)/2}\prod_{k=1}^{m}\Gamma(a - k + 1)$	the complex multivariate Gamma function
$E_1(x) := \int_x^{\infty} \frac{e^{-t}}{t}\,\mathrm{d}t$	the exponential integral function [1, (5.1.1)]
$f(x) = O(g(x))$	there exists a constant $c > 0$ and some x_0 such that $\lvert f(x) \rvert \leq c\lvert g(x) \rvert, \forall x \geq x_0$
$f(x) = \Theta(g(x))$	$f(x) = O(g(x))$ and $g(x) = O(f(x))$
$f(x) = o(g(x))$	$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$

Probability

$\sim$	distributed as
$\mathbb{P}(\mathcal{A})$	the probability of an event $\mathcal{A}$
$\mathbb{E}_{\mathsf{x}}[\cdot]$ (resp. $\mathbb{E}_{\mathcal{P}}[\cdot]$)	the expected value of a random quantity calculated w.r.t. random variable x (resp. distribution $\mathcal{P}$)
$\mathrm{Gamma}(k, \theta)$	Gamma distribution with shape k and scale θ
$\mathcal{N}_{\mathbb{C}}(\mu, \sigma^2)$	circularly symmetric complex Gaussian distribution of mean μ and variance σ^2
$\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	circularly symmetric complex Gaussian vector distribution of mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$

Information Theoretic Functions (see Appendix 1.A)

$H(\cdot)$ (resp. $h(\cdot)$)	entropy (resp. differential entropy) function
$D(P\|\|Q)$	the Kullback-Leibler divergence from a distribution P to a distribution Q
$I(\mathsf{x};\mathsf{y})$	mutual information between x and y

Topological Space

The Grassmann manifold $G(\mathbb{K}^T, M)$ is defined as the space of M-dimensional subspaces in $\mathbb{K}^T$ with $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$. In particular, $G(\mathbb{K}^T, 1)$ is the Grassmannian of lines. We use a truncated unitary matrix $\boldsymbol{U} \in \mathbb{C}^{T \times M}$, i.e., $\boldsymbol{U}^{\mathsf{H}}\boldsymbol{U} = \boldsymbol{I}_M$, to represent its column span, which is a point in $G(\mathbb{K}^T, M)$. The chordal distance between two points represented by $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ is $d(\boldsymbol{U}_1, \boldsymbol{U}_2) = \sqrt{M - \|\boldsymbol{U}_1^{\mathsf{H}}\boldsymbol{U}_2\|_{\mathrm{F}}^2}$. (See Appendix 1.B.)

Commonly Used Symbols

The following symbols are used repeatedly in the thesis, possibly with a user index.

M	number of transmit antennas
N	number of receive antennas
K	number of users
T	coherence interval
C	capacity
R	achievable rate
d	distance / degrees of freedom
P	power
B	number of bits
P_e	symbol error probability
$\mathcal{X}$/ $\mathcal{C}$	constellation
X/ x/ x	transmitted signal
Y/ y/ y	received signal
H/ h/ h	channel
Z/ z/ z	noise
$\boldsymbol{R}$	correlation/covariance matrix
r	matrix rank

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Wireless Communications

1.1.1 The History and Evolution of Wireless Communications

Wireless communication[1] is commonly understood as the *electromagnetic* transfer of information between points that are not connected by an electrical conductor [6]. In the history, the early communications between separated points were done by means of signals (i.e., changes in a physical quality) that human senses can directly perceive, e.g., auditorily such as shouts and drums, or visually such as smokes, flashes and flag semaphore. The electromagnetic transfer of information started when James C. Maxwell postulated the transmission of electromagnetic waves in 1864, and then Heinrich Hertz verified and demonstrated it in 1880 and 1887, respectively. This effect allows to use electromagnetic waves as information carriers with the help of electronic devices at the transmitter (source) and the receiver (destination). Marconi implemented a wireless telegraph and patented a complete wireless system in 1897. With the development of integrated circuits, electromagnetic wireless communication grew rapidly as radio and television broadcasting became widespread worldwide. Wireless systems evolved from the transmission of analog signals to the transmission of digital signals composed of bits, which was rooted in the seminal work of Claude Shannon in 1948 [7] and deployed in the 1980s.

Since the 1980s, the evolution of mobile wireless systems have been progressing from one generation to the next every nearly ten years. Each generation features different regulation, services, and innovations [8], [9]. In the first generation (1G), each country developed its own (analog or digital) system which provided only basic voice calling and fax service at a high price. The second generation (2G) was deployed in the 1990s and provided improved voice calls, short messaging service (SMS), and low-rate digital data services, such as e-mail and personal information management. The dominant standard was the GSM (first stood for Groupe Spécial Mobile and then Global System for Mobile communications) started in Europe, which was based on narrowband frequency-division multiple access (FDMA)/time-division multiple access (TDMA). Other standards included the American IS95 based on code-division multiple access (CDMA) and the Japanese Personal Digital Cellular (PDC) based on TDMA. In preparation for the third generation

[1] Standard texts on wireless communications include [2]–[5].

(3G), the Third Generation Partnership Project (3GPP) was established for the specification of the Universal Mobile Telecommunication System (UMTS), whose main air interface was wideband CDMA (WCDMA). It provided higher-rate data services, but encountered fierce competition from wireless local area network (WLAN). The fourth generation (4G) in the 2010s saw the success of long-term evolution (LTE) and LTE-Advanced, which are based on orthogonal frequency division multiplexing (OFDM) using multiple carriers. It achieves a high speed in both downlink and uplink, and facilitates a wide range of services that bring ever-increasing revenue.

The fifth generation (5G) [10], [11] has been under development since early 2010s. 5G networks are envisioned to support a large number and heterogeneity of end devices, i.e., the internet of things (IoT) [12]. Three main use cases targeted in 5G are enhanced mobile broad band (eMBB), ultra-reliable low-latency communication (URLLC), and massive machine-type communications (mMTC) [13]. 5G networks have been tested in many countries and are now in the early stage of commercial deployment [14]. Meanwhile, the research activities towards the sixth generation (6G) have been initiated. Although it remains unclear what 6G will be, many (speculative) visions for 6G have been provided from different perspectives in, e.g., [15]–[20].

1.1.2 Wireless Propagation

In this subsection, we review the propagation effect of electromagnetic waves through a wireless channel. Electromagnetic waves are emitted by an antenna at the transmitter and intercepted by an antenna at the receiver. In principle, one could solve the electromagnetic field Maxwell's equations to find the received waves at the receive antenna. However, since this is too complex, simpler models are used to approximate the signal propagation. A transmitted wave is reflected, scattered, and diffracted when it interacts with objects in the environment as it propagates toward the receiver. Therefore, the receiver observes multiple copies of this wave at different delays, each experiences a specific attenuation. In other words, the wave travels through multiple paths, gets attenuated in each path, and arrives at the receive antenna at different delays. Furthermore, due to the relative movement of the transmitter, receiver, and objects in the environment, the attenuation factors, i.e., channel gains, vary over time. This effect is called *fading*. Let $a_n(t)$ be the complex channel gain of path n at time t. Consider a transmission of a signal $x(t)$ at time t and carrier frequency f_c, the received signal is the superposition of the all the multipath components as [4, Eq.(3.2)]

$$y(t) = \sum_{n=0}^{N_p(t)} a_n(t) x(t - \tau_n(t)) e^{\jmath 2\pi f_c (t - \tau_n(t))}, \tag{1.1}$$

where $n = 0$ corresponds to the line-of-sight path, $N_p(t)$ is the number of resolvable multipath components and $\tau_n(t)$ is the delay of the n-th path at time t. The received signal can also be expressed as [4, Eq.(3.5)]

$$y(t) = \left(\int_{-\infty}^{\infty} h(\tau, t) x(t - \tau) \, \mathrm{d}\tau \right) e^{\jmath 2\pi f_c t}, \tag{1.2}$$

where [4, Eq.(3.6)]

$$h(\tau, t) := \sum_{n=0}^{N_p(t)} e^{-\jmath 2\pi f_c \tau_n(t)} a_n(t) \delta(\tau - \tau_n(t)) \tag{1.3}$$

is the channel impulse response at time t and delay τ, i.e., the response of the channel at time t to an impulse transmitted at time $t-\tau$. In this way, the Maxwell's equations are replaced by an input/output relation of a linear time-varying system.

In general, due to the random nature of the environment, the attenuation gain $a_n(t)$, the delay $\tau_n(t)$, and the number of paths $N_p(t)$ are random. Therefore, the channel impulse response $h(\tau,t)$ is modeled as realization of a random variable $\mathsf{h}(\tau,t)$. It is commonly assumed that $\mathsf{h}(\tau,t)$ is a Gaussian process, when the number of multipath components $N_p(t)$ is large, by evoking the central-limit theorem. Another common assumption is that the phase of each multipath component is uniformly distributed. Under these assumptions, the channel impulse response statistics are fully characterized by the mean $\mathbb{E}[h(\tau,t)]$ and autocorrelation function [4, Eq.(3.50)]

$$A_{\mathsf{h}}(\tau_1,\tau_2;t_1,t_2) := \mathbb{E}[\mathsf{h}^*(\tau_1,t_1)\mathsf{h}(\tau_2,t_2)]. \tag{1.4}$$

If the joint statistics of $\mathsf{h}(\tau_1,t_1)$ and $\mathsf{h}(\tau_2,t_2)$, in particular $A_{\mathsf{h}}(\tau_1,\tau_2;t_1,t_2)$, depend only on the time difference $\Delta t = t_2 - t_1$, the channel is said to be wide-sense stationary (WSS). Furthermore, if the channel response of a given multipath component at different delays τ are uncorrelated, the channel has uncorrelated scattering (US). WSS and US are reasonable assumptions in practice.

In the frequency domain, the channel frequency response is given by the Fourier transform of $\mathsf{h}(\tau,t)$ with respect to (w.r.t.) τ [4, Eq.(3.57)]

$$\mathsf{H}(f,t) := \int_{-\infty}^{\infty} \mathsf{h}(t,\tau)e^{-\jmath 2\pi f t}\,\mathrm{d}\tau. \tag{1.5}$$

Given that $\mathsf{h}(t,\tau)$ is a WSSUS (i.e., WSS and US) Gaussian process, $\mathsf{H}(f,t)$ is also a WSSUS Gaussian process with autocorrelation function

$$A_{\mathsf{H}}(\Delta f;\Delta t) := \mathbb{E}[\mathsf{H}^*(f,t)\mathsf{H}(f+\Delta f,t+\Delta t)], \tag{1.6}$$

which depends only one the time difference Δt and frequency difference Δ_f. Two channel frequency responses at the same time instant, i.e. $\Delta t = 0$, and frequency separation Δf are approximately independent if $A_{\mathsf{H}}(\Delta f;0) \approx 0$. The frequency B_c where $A_{\mathsf{H}}(\Delta f;0) \approx 0$ for all $\Delta f > B_c$ is called the *coherence bandwidth* of the channel. B_c is inversely proportional to the delay spread of $\mathsf{h}(\tau,t)$ in time. In general, if the transmitted signal has narrow bandwidth $B \ll B_c$, then the channel response is roughly constant across the entire signal bandwidth. This is referred to as *flat fading*. On the contrary, if $B \gg B_c$, then the channel response varies widely across the signal bandwidth. This is referred to as *frequency-selective fading*. On the other hand, the autocorrelation $A_{\mathsf{H}}(\Delta f;\Delta t)$ for $\Delta f = 0$ characterizes how the channel responses decorrelates over time. If $A_{\mathsf{H}}(0;\Delta t) \approx 0$, then the channel measured at time instants separated by Δt are approximately uncorrelated and therefore independent. The time duration T_c where $A_{\mathsf{H}}(0;\Delta t) \approx 0$ for all $\Delta t > T_c$ is called the *coherence time* of the channel. It is inversely proportional to the Doppler spread. A block of coherence bandwidth B_c and coherence time T_c is called a *coherence block*, and the total length $T = T_c B_c$ of a coherence block is called *coherence interval*.

By a shift in the frequency domain (i.e, down-converting) of the received signal (1.2), we have an equivalent baseband representation. Furthermore, by sampling and considering an additive noise, we obtain a discrete-time baseband model given by [5, Eq.(2.39)]

$$\mathsf{y}[m] = \sum_l \mathsf{h}_l[m]\mathsf{x}[m-l] + \mathsf{z}[m], \tag{1.7}$$

where at sampling instant m/B (B is the bandwidth), $\mathsf{y}[m]$ and $\mathsf{x}[m]$ are respectively the samples of the received and transmitted signal in baseband, $\mathsf{z}[\mathsf{m}]$ is the low-pass filtered noise, and

$$\mathsf{h}_l[m] := \sum_{n=1}^{N_p(t)} e^{-\jmath\pi f_c \tau_n (m/B)} \mathsf{a}_n(m/B)\ \mathrm{sinc}(l - \tau_n(m/B)B), \tag{1.8}$$

with $\mathrm{sinc}(t) := \frac{\sin(\pi t)}{\pi t}$, is the l-th complex channel filter tap. $\mathsf{h}_l[m]$ and $\mathsf{z}[m]$ are also normally assumed to be (discrete) Gaussian process.

1.1.3 The MIMO Channel

We now assume that the antenna arrays at the transmitter and the receiver have multiple elements. Specifically, the transmitter is equipped with M antennas and the receiver with N antennas. We assume a frequency-flat channel and use a statistical channel model as in the previous subsection with a single-tap discrete-time baseband representation.[2] At time m, the transmitter sends a signal $\mathsf{x}_j[m]$ from the j-th antenna. The received signal at the i-th antenna is

$$\mathsf{y}_i[m] = \sum_{j=1}^{M} \mathsf{h}_{ij}[m]\mathsf{x}_j[m] + \mathsf{z}_i[m], \quad i \in [N], \tag{1.9}$$

where h_{ij} is the channel coefficient from the j-th transmit antenna to the i-th receive antenna with arbitrary distribution, and $\mathsf{z}_i[m]$ is the additive white Gaussian noise (AWGN) following the $\mathcal{N}_{\mathbb{C}}(0,1)$ distribution. Using the vector/matrix representation

$$\mathbf{x}[m] = \begin{bmatrix} \mathsf{x}_1[m] \\ \mathsf{x}_2[m] \\ \cdots \\ \mathsf{x}_M[m] \end{bmatrix}, \quad \mathbf{y}[m] = \begin{bmatrix} \mathsf{y}_1[m] \\ \mathsf{y}_2[m] \\ \cdots \\ \mathsf{y}_N[m] \end{bmatrix}, \quad \mathbf{z}[m] = \begin{bmatrix} \mathsf{z}_1[m] \\ \mathsf{z}_2[m] \\ \cdots \\ \mathsf{z}_N[m] \end{bmatrix},$$
$$\mathbf{H}[m] = \begin{bmatrix} \mathsf{h}_{11}[m] & \mathsf{h}_{12}[m] & \ldots & \mathsf{h}_{1M}[m] \\ \mathsf{h}_{21}[m] & \mathsf{h}_{22}[m] & \ldots & \mathsf{h}_{2M}[m] \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{h}_{N1}[m] & \mathsf{h}_{N2}[m] & \ldots & \mathsf{h}_{NM}[m] \end{bmatrix}, \tag{1.10}$$

we obtain the following channel model

$$\mathbf{y}[m] = \mathbf{H}[m]\mathbf{x}[m] + \mathbf{z}[m]. \tag{1.11}$$

This channel is referred to as a multiple-input multiple-output (MIMO) point-to-point (P2P) channel. A conceptual illustration is given in Fig. 1.1.

1.1.3.a Fundamental Limits of MIMO Channel

At the transmitter, using a *channel coding scheme*, information is encoded into messages $\{\mathcal{M}\}$ where each message is represented by a channel codeword. Then each codeword is mapped into a sequence of N_s transmitted complex vector-valued symbols $\mathbf{x}[1], \mathbf{x}[2], \ldots, \mathbf{x}[N_s]$. Due to the regulation and hardware limits, the transmitted signal is subject to a power constraint as

$$\frac{1}{N_s} \sum_{i=1}^{N_s} \left\|\mathbf{x}[i]\right\|^2 \leq P. \tag{1.12}$$

[2] In the remainder of the thesis, we consider the baseband representation of the channel model.

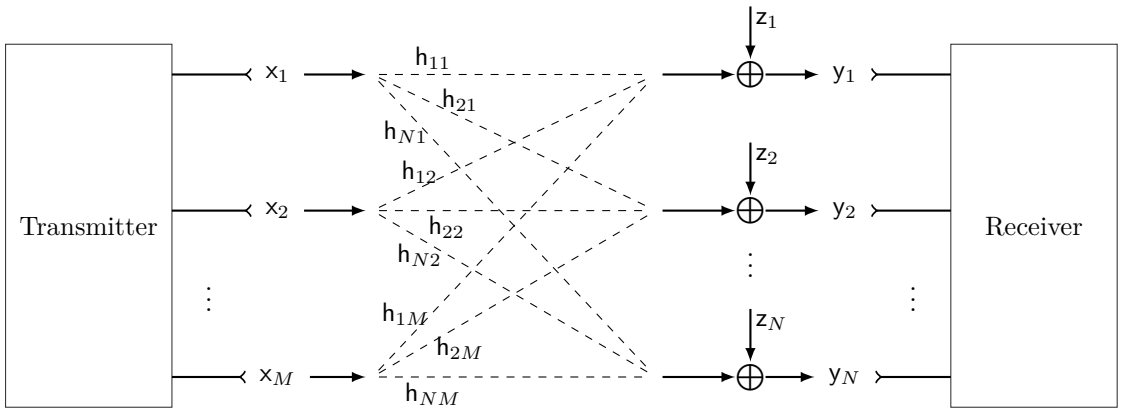


FIGURE 1.1: Conceptual illustration of a MIMO channel with M transmit antennas and N receive antennas.

That is, P is the maximal total transmit power from M transmit antenna in average. Since the noise has unit power, P is also the average power ratio between the transmitted signal and the noise at each receive antenna, and is referred to as the signal-to-noise ratio (SNR) of the channel. The receiver detects the transmitted symbols using an maximum likelihood (ML) detector as

$$\{\hat{\mathbf{x}}[i]\}_{i=1}^{N_s} = \arg \max_{\{\mathbf{x}[i]\}_{i=1}^{N_s}} p(\{\mathbf{y}[i]\}_{i=1}^{N_s} \mid \{\mathbf{x}[i]\}_{i=1}^{N_s}), \tag{1.13}$$

where the maximization is over all possible codewords subject to the power constraint (1.12), $\{\mathbf{y}[i]\}_{i=1}^{N_s}$ are the corresponding received signals, and the conditional probability $p(\{\mathbf{y}[i]\}_{i=1}^{N_s} \mid \{\mathbf{x}[i]\}_{i=1}^{N_s})$ is implicitly conditioned on the knowledge about the channel and noise available at the receiver.[3] Then, the receiver decodes the detected symbol sequence to obtain an estimate $\hat{\mathcal{M}}$ of the message $\mathcal{M}$. The reliability of this information transmission is captured by the error probability $\mathbb{P}\left(\hat{\mathcal{M}} \neq \mathcal{M}\right)$ which depends on the codelength N_s. One would be interested in finding the maximal rate at which information can be transmitted and received reliably at the receiver. This quantity is the called the (Shannon) *capacity* of the channel and defined as follows.

Definition 1.1 (Channel capacity)**.** *The capacity C of the channel is the maximal data rate, i.e., the maximal number of information units needed to represent a message $\mathcal{M}$ normalized by the communication time, such that there exists a channel coding scheme achieving arbitrarily low error probability $\mathbb{P}\left(\hat{\mathcal{M}} \neq \mathcal{M}\right)$ as the codelength N_s goes to infinity.*

The channel capacity is normally measured in bits per channel use (bpcu). From the definition, reliable communication is not possible, even with infinite codelength, if one transmits at a rate above the channel capacity.

Since our considered channel (1.9) is memoryless, i.e., the output $\mathbf{y}$ depends only on the input $\mathbf{x}$ at the time and is independent of all previous inputs, the channel capacity is given by [21], [22]

$$C(P) = \max_{p_{\mathbf{x}}:\, \mathbb{E}[\|\mathbf{x}\|^2] \leq P} I(\mathbf{x};\mathbf{y}), \tag{1.14}$$

where $I(\mathbf{x};\mathbf{y})$ is the mutual information (see Appendix 1.A) between two distributions $p_{\mathbf{x}}$ and $p_{\mathbf{y}}$ implicitly conditioned on the available knowledge about the channel and noise.

[3] We assume that the distribution of the channel and the noise are known. The availability of the instantaneous value of the channel matrix will be discussed in subsequent subsections.

A *capacity-achieving input distribution* is the solution to the maximization in (1.14). If one transmits with a given input distribution $p_{\mathbf{x}}$ satisfying the power constraint but not necessarily capacity achieving, then the mutual information

$$R(P) = I(\mathbf{x};\mathbf{y}) \tag{1.15}$$

is an *achievable rate* of the channel. In many situations, deriving the channel capacity or achievable rate is very challenging, and one can resort to an asymptotic coarse representation of these quantities given by the degrees of freedom (DoF). An achievable DoF and the optimal DoF of the channel are respectively defined by

$$d_{\text{achievable}} := \lim_{P\to\infty} \frac{R(P)}{\log_2(P)} \quad \text{and} \quad d_{\text{optimal}} := \lim_{P\to\infty} \frac{C(P)}{\log_2(P)}. \tag{1.16}$$

With this, the achievable rate and capacity behave in the high-SNR regime as $R(P) = d_{\text{achievable}} \log_2(P) + o(\log_2 P)$ and $C(P) = d_{\text{optimal}} \log_2(P) + o(\log_2 P)$. Therefore, the DoF is also called the *pre-log* factor of the rate/capacity. Roughly speaking, the DoF is the number of additional bits that can be transmitted reliably when the signal power is doubled.

1.1.3.b Practical Design for a MIMO Channel

Assuming that the channel $\mathbf{H}$ is independent over time, one can simplify the ML detection (1.13) to the symbol-by-symbol detection

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{y}\,|\,\mathbf{x}). \tag{1.17}$$

Constellation Design

Although the capacity-achieving input distribution is often continuous, in practice, the transmitted signal $\mathbf{x}$ is normally drawn from a finite discrete set to reduce complexity. This set is referred to as the *constellation* and each element is called a *constellation point/symbol*. Consider a constellation $\mathcal{X} := \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{|\mathcal{X}|}\}$ with $\frac{1}{|\mathcal{X}|}\sum_{i=1}^{|\mathcal{X}|} \|\boldsymbol{x}_i\|^2 \le P$. Let $\mathbf{x}$ be uniformly drawn from $\mathcal{X}$, i.e., the input distribution is $p_{\mathbf{x}}(\boldsymbol{x}) = \frac{1}{|\mathcal{X}|}\mathbb{1}\{\boldsymbol{x} \in \mathcal{X}\}$, then the achievable data rate is given as

$$\begin{aligned} R(P,\mathcal{X}) &= I(\mathbf{x};\mathbf{y}) && (1.18)\\ &= \mathbb{E}\left[\log_2 \frac{p(\mathbf{y}|\mathbf{x})}{\frac{1}{|\mathcal{X}|}\sum_{\boldsymbol{x}\in\mathcal{X}} p(\mathbf{y}|\mathbf{x}=\boldsymbol{x})}\right] && (1.19)\\ &= \log_2|\mathcal{X}| - \mathbb{E}\left[\log_2 \frac{\sum_{\boldsymbol{x}\in\mathcal{X}} p(\mathbf{y}|\mathbf{x}=\boldsymbol{x})}{p(\mathbf{y}|\mathbf{x})}\right] \quad \text{bpcu}, && (1.20)\end{aligned}$$

where the expectation is over the joint distribution of the input and output. Here, $\log_2|\mathcal{X}|$—the number of bits required to represent a constellation symbol—is the rate achievable in the noiseless case, and $\mathbb{E}\left[\log_2 \frac{\sum_{\boldsymbol{x}\in\mathcal{X}} p(\mathbf{y}|\mathbf{x}=\boldsymbol{x})}{p(\mathbf{y}|\mathbf{x})}\right]$ is the rate loss due to noise. The constellation $\mathcal{X}$ should be designed so as to maximize the achievable rate:[4]

$$\mathcal{X}^* = \arg\max_{\mathcal{X}} R(P,\mathcal{X}), \tag{1.21}$$

[4]In principle, the constellation and the probability mass function (PMF) of the transmitted signal in this constellation should be jointly optimized. The optimization of this PMF for a fixed constellation is called signal shaping [23].

where the optimization space can be the Cartesian product of M instances of the unit complex disc

$$\mathcal{D}(0,1) := \{z \in \mathbb{C} : \ |z|^2 \leq 1\}, \tag{1.22}$$

scaled by the transmit power per antenna such that the power constraint (1.12) holds. In the large constellation regime $|\mathcal{X}| \to \infty$, the uniform distribution of the signal $\mathbf{x}$ in $\mathcal{X}$ converges to a continuous distribution. The constellation should be designed such that this limit distribution is close to the capacity-achieving input distribution, and as a consequence, the achievable rate converges to the channel capacity.

Another constellation design criterion is to minimize the detection error $P_e(\mathcal{X}) = \mathbb{P}(\hat{\mathbf{x}} \neq \mathbf{x}) = \frac{1}{|\mathcal{X}|}\sum_{i=1}^{|\mathcal{X}|} \mathbb{P}(\hat{\mathbf{x}} \neq \boldsymbol{x}_i \,|\, \mathbf{x} = \boldsymbol{x}_i)$:

$$\mathcal{X}^* = \arg\min_{\mathcal{X}} P_e(\mathcal{X}). \tag{1.23}$$

Since the error event $\{\hat{\mathbf{x}} \neq \boldsymbol{x}_i \,|\, \mathbf{x} = \boldsymbol{x}_i\}$ is the union of the pairwise error events $\{p(\mathbf{y} \,|\, \boldsymbol{x}_j) > p(\mathbf{y} \,|\, \boldsymbol{x}_i) \,|\, \mathbf{x} = \boldsymbol{x}_i\}$ for all $j \neq i$, we have the following union bound

$$P_e(\mathcal{X}) \leq \frac{1}{|\mathcal{X}|}\sum_{i=1}^{|\mathcal{X}|}\sum_{j=1, j\neq i}^{|\mathcal{X}|} \mathbb{P}(p(\mathbf{y} \,|\, \boldsymbol{x}_j) > p(\mathbf{y} \,|\, \boldsymbol{x}_i) \,|\, \mathbf{x} = \boldsymbol{x}_i). \tag{1.24}$$

Therefore, the criterion (1.23) amounts to

$$\mathcal{X}^* = \arg\min_{\mathcal{X}} \max_{1\leq i<j\leq|\mathcal{X}|} \mathbb{P}(p(\mathbf{y} \,|\, \boldsymbol{x}_j) > p(\mathbf{y} \,|\, \boldsymbol{x}_i) \,|\, \mathbf{x} = \boldsymbol{x}_i). \tag{1.25}$$

Efficient Detection Design

Due to the discrete domain of the symbols, the ML detection (1.17) is often non-deterministic polynomial-time (NP) hard [24]. To solve it, one has to enumerate the whole constellation, which is cumbersome if the constellation size is large. Therefore, from a practical point of view, it is favorable to use a sub-optimal detection with low complexity. If the likelihood function metric $p(\mathbf{y} \,|\, \mathbf{x})$ is replaced with a sub-optimal detection metric $\hat{p}(\mathbf{y} \,|\, \mathbf{x})$ (which might be easier to compute) in (1.17), then according to [25, Section II], the highest data rate reliably achievable—so-called the *mismatched rate*—is lower bounded by the generalized mutual information (GMI) given by

$$R_{\mathrm{GMI}}(P, \mathcal{X}, \hat{p}) = \sup_{s\geq 0} \mathbb{E}\left[\log_2 \frac{\hat{p}(\mathbf{y} \,|\, \mathbf{x})^s}{\sum_{\boldsymbol{x}\in\mathcal{X}} \mathbb{P}(\mathbf{x} = \boldsymbol{x})\hat{p}(\mathbf{y} \,|\, \boldsymbol{x})^s}\right] \tag{1.26}$$

$$= \sup_{s\geq 0} \mathbb{E}\left[\log_2 |\mathcal{X}| - \log_2 \frac{\sum_{\boldsymbol{x}\in\mathcal{X}} \hat{p}(\mathbf{y} \,|\, \boldsymbol{x})^s}{\hat{p}(\mathbf{y} \,|\, \mathbf{x})^s}\right] \tag{1.27}$$

$$= \log_2 |\mathcal{X}| - \inf_{s\geq 0} \mathbb{E}\left[\log_2 \frac{\sum_{\boldsymbol{x}\in\mathcal{X}} \hat{p}(\mathbf{y} \,|\, \boldsymbol{x})^s}{\hat{p}(\mathbf{y} \,|\, \mathbf{x})^s}\right] \quad \text{bpcu}, \tag{1.28}$$

where the expectation is over the joint distribution of $\mathbf{x}$ and $\mathbf{y}$, and the second equality follows from the uniformity of $\mathbf{x}$ in $\mathcal{X}$. If the constellation $\mathcal{X}$ has a structure, one can also exploit this structure to design efficient detection by, e.g., decoupling (1.17) into the detection of each component of $\mathbf{x}$.

1.1.4 Coherent Communications: The Role of Channel State Information

The instantaneous value of $\mathbf{H}$ is referred to as channel state information (CSI). If the transmitter and/or receiver has this information, the communication is said to be *coherent*. We assume that the channel is independent and identically distributed (IID) Rayleigh fading so that the channel components h_{ij} are independent and follow $\mathcal{N}_{\mathbb{C}}(0,1)$. Let the receiver have CSI. In this case, if the transmitter only knows the channel distribution, the capacity was derived by Telatar [26] and Foschini and Gans [27] as

$$C(P) = \mathbb{E}_{\mathbf{H}}\Big[\log_2 \det\Big(\boldsymbol{I}_N + \tfrac{P}{M}\mathbf{H}\mathbf{H}^{\mathsf{H}}\Big)\Big] \quad \text{bpcu}, \tag{1.29}$$

and the capacity-achieving input distribution is the circularly symmetric complex vector Gaussian $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \frac{P}{M}\boldsymbol{I}_M)$. Furthermore, if CSI is also available at the transmitter (CSIT), then the capacity is given by

$$C(P) = \mathbb{E}_{\mathbf{H}}\left[\max_{\boldsymbol{Q}\in\mathbb{C}^{M\times M}:\ \boldsymbol{Q}\succeq 0, \mathrm{tr}(\boldsymbol{Q})=P} \log_2 \det\big(\boldsymbol{I}_N + \mathbf{H}\boldsymbol{Q}\mathbf{H}^{\mathsf{H}}\big)\right] \quad \text{bpcu}, \tag{1.30}$$

and the capacity-achieving input distribution is adapted to each realization of $\mathbf{H}$ as $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{Q}_{\mathrm{opt}}(\boldsymbol{H}))$ where $\boldsymbol{Q}_{\mathrm{opt}}(\boldsymbol{H})$ is the solution to the maximization in (1.30) when $\mathbf{H} = \boldsymbol{H}$. The eigenvectors of $\boldsymbol{Q}_{\mathrm{opt}}(\boldsymbol{H})$ are the same as the right singular vectors of $\boldsymbol{H}$, while the eigenvalues of $\boldsymbol{Q}_{\mathrm{opt}}(\boldsymbol{H})$ are obtained from the singular values of $\boldsymbol{H}$ by the waterfilling power allocation [26], [28]. In both cases, the capacity scales as

$$C(P) = \min\{M, N\} \log_2 P + O(1), \tag{1.31}$$

thus the optimal DoF is given by $\min\{M, N\}$. In fact, for each realization of $\mathbf{H}$, the MIMO channel can be converted to $\min\{M, N\}$ parallel, non-interfering single-input single-output (SISO) channels through a singular value decomposition (SVD) of $\mathbf{H}$. Each parallel channel has gain corresponding to a singular value of $\mathbf{H}$ and has one DoF. Therefore, with CSI, the capacity scales up linearly with the number of antennas.

In coherent communications, one normally uses a scalar constellation $\mathcal{C}$ containing a set of points in the disc $\mathcal{D}(0,1)$ (1.22) for each component of $\mathbf{x}$ (thus $\mathcal{X} = \{\rho[x_1 \ \ldots \ x_M]^{\mathsf{T}} : x_i \in \mathcal{C}, i \in [M]\}$ where ρ is a scaling factor). The most common scalar constellations are pulse amplitude modulation (PAM), phase shift keying (PSK), and quadrature amplitude modulation (QAM), where the constellation points differ respectively in their amplitude, phase, and both amplitude and phase. A diagram of these constellations is shown in Fig. 1.2.

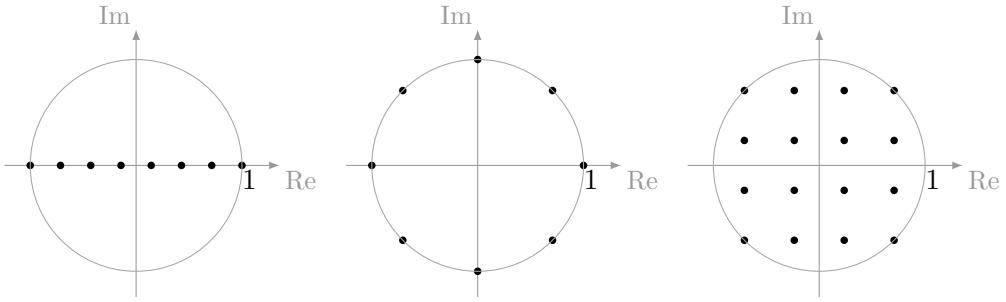
Given the IID Rayleigh fading matrix $\mathbf{H} = \boldsymbol{H}$ and $\mathbf{x} = \boldsymbol{x}$, the channel output $\mathbf{y}$ is a Gaussian vector with mean $\boldsymbol{H}\boldsymbol{x}$ and covariance matrix $\boldsymbol{I}_N$, thus the likelihood function is given by

$$p_{\mathbf{y}\,|\,\mathbf{x},\mathbf{H}}(\boldsymbol{y}\,|\,\boldsymbol{x},\boldsymbol{H}) = \frac{1}{\pi^N} \exp\Big(-\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2\Big). \tag{1.32}$$

Therefore, the ML detection (1.17) given $\mathbf{y} = \boldsymbol{y}$ and $\mathbf{H} = \boldsymbol{H}$ is equivalent to the least-squares (LS) problem

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}\in\mathcal{X}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2. \tag{1.33}$$

This is also referred to as nearest-neighbor decoding (NND) since it looks for the closest (in terms of Euclidean distance) symbol $\boldsymbol{x}$ to $\boldsymbol{y}$ in the subspace of $\boldsymbol{H}$. The coherent detection problem (1.33) has been investigated extensively in the literature. Since it is NP-hard [24], many sub-optimal schemes have been proposed to reduce complexity, including:



(a) 8-PAM (b) 8-PSK (c) 16-QAM

FIGURE 1.2: Diagram of three common scalar constellations: PAM, PSK, and QAM. The circle is the boundary of the disc $\mathcal{D}(0,1)$.

- linear detectors consisting in component-wise demapper of $\hat{\boldsymbol{x}} = \boldsymbol{T}\boldsymbol{y}$, such as matched-filter with $\boldsymbol{T} = \boldsymbol{H}^{\mathsf{H}}$, zero forcing (ZF) with $\boldsymbol{T} = (\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathsf{H}} =: \boldsymbol{H}^{\dagger}$ (assuming $N \geq M$), and minimum mean square error (MMSE) with $\boldsymbol{T} = (\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H} + \boldsymbol{I}_N)^{-1}\boldsymbol{H}^{\mathsf{H}}$;
- interference cancellation aided detectors, such as successive/parallel/multi-stage/decision-feedback interference cancellation;
- tree-search based detectors, such as sphere decoders;
- lattice-reduction aided detectors.

In short, CSI enables to adapt the transmission and reception to the current fading state. Comprehensive surveys of coherent MIMO transmission and detection can be found respectively in [29] and [30].

1.1.5 The Cost of Acquiring CSI

Although coherent communication exploits effectively the extra spatial resources of a MIMO channel to increase the spectral efficiency, it relies on the availability of CSI. In practice, since the channel matrix is random and fades over time and frequency, its value is not given *a priori* and has to be estimated. Typically, channel estimation is carried out by sending reference symbols, so-called *pilots*, known to the receiver in some channel uses of a coherence block. The receiver estimates the channel in these channel uses using the known pilots, then inter/extrapolates them to infer the channel gains of the remaining channel uses within the coherence block. By treating the channel estimate as the known channel, coherent communication can be performed in these remaining channel uses. This is called a *pilot-based scheme/approach* [31]. To properly analyze the system performance, one needs to take into account the cost of channel estimation and the channel estimation error.

Since the pilot symbols are known to the receiver, they do not carry any information. On the other hand, they occupy a fraction of communication time/frequency resource. In a MIMO channel with M transmit antennas, one would need to send at least M pilot symbols for the receiver to determine M channel vectors corresponding to the M antennas [31]. Let $T = T_c B_c$ be the coherence interval, then the fraction of resource spent for channel estimation is $\frac{M}{T}$ and there remains a $1 - \frac{M}{T}$ fraction of the coherence block for coherent data transmission. In a highly mobile environment where the channel state changes rapidly, the coherence interval T is short, and the fraction of pilot transmission can be disproportionate to data transmission, especially if the number of antennas is large.

Channel estimation error has negative impact on both channel throughput and error performance. On one hand, if one treats the channel estimate as the true channel and disregard any inaccuracy, the optimal detector under this assumption is a mismatched detector for the channel with channel estimation error, and the channel throughput is determined by the mismatched rate. On the other hand, even if the statistics of the channel estimation error is taken into account, this residual error imposes a secondary noise which increases the total noise power and reduces the channel capacity for a given signal power. The effect of imperfect channel knowledge on channel capacity has been investigated in, e.g., [32], [33] and on detection error in, e.g., [34], [35].

In some scenarios, pilot-based channel estimation becomes difficult or even impossible. For example, in the uplink of a multi-user system, pilot sequences are assigned per user and orthogonally across users. If the total number of users is larger than the coherence interval (but probably only a random number of users are active at a time), pre-assigning mutually orthogonal pilot sequences to every user present in the system is not possible. One can consider non-orthogonal pilots, but accurate CSI acquisition is still challenging.

Communication without *a priori* CSI at any node is said to be *noncoherent*. This framework provides a more realistic standpoint to system analysis and design. Noncoherent communication is also a more general framework since, as aforementioned, one can first estimate the channel and then perform coherent communication. Nevertheless, sequential/joint channel estimation and coherent data detection might not be optimal. When the cost of channel estimation is significant, the channel estimation error is severe, or pilot-based channel estimation becomes inconvenient/impossible, it might be beneficial to refrain from doing it by using a communication scheme that does not rely on the knowledge of CSI. Noncoherent communication accounts for such a scheme.

In this thesis, we focus on noncoherent communications. We review the state-of-the-art of noncoherent communications in the next section.

1.2 Noncoherent Wireless Communications

In this section, we provide an overview of the state-of-the-art of noncoherent wireless communications, as well as the questions that we want to answer in this thesis.

1.2.1 Assumptions

Throughout this section, we consider the following assumptions on the fading coefficients.

1.2.1.a Channel Information Availability

- *Channel Distribution Information:* The distribution of the channel coefficients, so-called channel distribution information (CDI), is assumed to be known to all the communicating nodes. This is because CDI characterizes the macroscopic propagation effects that are stable, and thus can be tracked with negligible communication resources. In particular, the channel statistics, such as the mean and covariance, are assumed to be known.
- *No Channel State Information:* We assume that the instantaneous value of the channel coefficients, i.e. CSI, is not known at any node.

1.2.1.b Channel Variation

- *Stationary Fast Fading:* In an extremely highly mobile environment, the channel coefficients can be assumed to be mutually uncorrelated. In this case, the random channel process is stationary [36].
- *Block Fading:* The block fading assumption refers to the case where the channel coefficients within a coherence block of length $T = T_c B_c$ are highly correlated such that they can be considered unchanged within the block. Furthermore, the channel coefficients in different blocks are assumed to be independent and identically distributed. The block fading model approximates, in a tractable manner, a continuous fading process, such as Jakes' [2], by a piecewise-constant process. The inter-block independence is also relevant for a system with, e.g., block interleaving, random TDMA, frequency hopping, or sporadic transmissions, in which the gap between successive transmissions (to the same receiver) is large or indefinite. The special case $T = 1$ coincides with stationary fast fading. In the multi-user case, we assume that the coherence blocks have equal length and are synchronous across users.[5] Because the channel coefficients are statistically identical between different coherence blocks, one can focus on a single representative coherence block.

1.2.1.c Channel Distribution

For a channel matrix $\mathbf{H}$, either of following assumptions on its distribution will be considered.

- *IID Rayleigh Fading:* the entries of $\mathbf{H}$ are independent and identically Gaussian distributed.

[5] The case where the coherence blocks of different users are asynchronous and unequal in length, so-called coherence diversity, was investigated in [37].

- *Generic Fading:* The fading is said to be *generic* if the channel matrix has finite differential entropy[6] and finite second moment:

$$h(\mathbf{H}) > -\infty, \quad \mathbb{E}\left[\|\mathbf{H}\|_{\mathrm{F}}^2\right] < \infty. \tag{1.34}$$

  The IID Rayleigh fading is a special case of generic fading.

- *Spatial Correlated Fading:* In practice, the channel coefficients between different antennas are often correlated, i.e., the channel gains and directions are not independent: the received signal gain in some spatial directions are larger than in the others. The spatial correlation can stem from the propagation environment or the spatially dependent patterns of the antennas. We consider the Kronecker (separable) correlation model

$$\mathbf{H} = \boldsymbol{R}_r^{\frac{1}{2}} \breve{\mathbf{H}} \boldsymbol{R}_t^{\frac{1}{2}}, \tag{1.35}$$

  where $\boldsymbol{R}_r \in \mathbb{C}^{N\times N}$ and $\boldsymbol{R}_t \in \mathbb{C}^{M\times M}$ are deterministic and known correlation matrices at the receiver and transmitter sides, respectively; $\breve{\mathbf{H}}$ may have IID Gaussian entries (spatially correlated Rayleigh fading) or follow a generic distribution as in (1.34) (spatially correlated generic fading).

In the following, we focus on the P2P channel, the multiple-access channel (MAC), and the broadcast channel (BC). For other channels, refer to, e.g., [39] for the interference channel, [40], [41] for the relay channel, and [42] for the two-hop diamond network.

1.2.2 The Point-to-Point Channel

1.2.2.a Channel Model

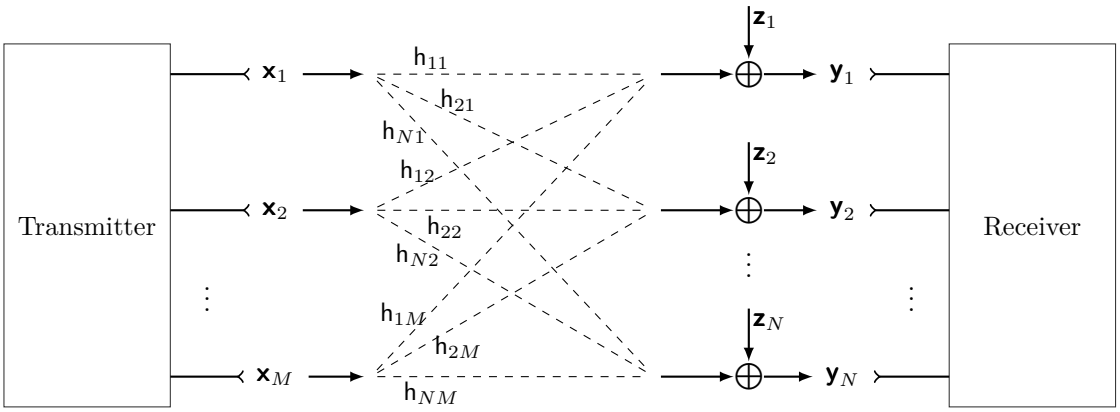


FIGURE 1.3: Conceptual illustration of a MIMO P2P channel with M transmit antennas and N receive antennas in block fading. Each transmitted signal vector $\mathbf{x}_i \in \mathbb{C}^T$ contains the scalar symbols transmitted from antenna i during a coherence block of length T.

We consider an $M \times N$ MIMO P2P channel as in Section 1.1.3 but with joint transmission within each coherence block as illustrated in Fig. 1.3. During a coherence block b, the transmitter sends an $M \times T$ signal matrix $\mathbf{X}[b]$ across M antennas and T channel uses, and the receiver receives an $N \times T$ signal matrix

$$\mathbf{Y}[b] = \mathbf{H}[b]\mathbf{X}[b] + \mathbf{Z}[b], \quad b = 1, 2, \ldots, \tag{1.36}$$

[6]The fading in which the channel matrix has (negative) finite differential entropy is so called *regular* fading [38].

where $\mathbf{Z}[b] \in \mathbb{C}^{N\times T}$ is the AWGN with IID $\mathcal{N}_{\mathbb{C}}(0,1)$ entries independent of $\mathbf{H}$. We consider the power constraint

$$\frac{1}{\nu}\sum_{b=1}^{\nu} \|\mathbf{X}[b]\|_{\mathrm{F}}^2 \leq PT, \tag{1.37}$$

where ν is the number of the blocks spanned by a channel codeword. P is the SNR of the channel.

1.2.2.b Fundamental Limits and Capacity-Achieving Signal

The capacity of the noncoherent P2P channel has been investigated in a large body of work. In general, the explicit capacity is unknown and only approximations in the extreme-SNR regime are known in some settings. A commonly considered setting is the IID Rayleigh fading in which the components of $\mathbf{H}$ are IID Gaussian random variables.

$T = 1$ (stationary fast fading)

Under the IID Rayleigh fading assumption for the SISO channel ($M = N = 1$) in fast fading ($T = 1$), Richters conjectured in 1967 that, although the channel is continuous, the capacity-achieving input distribution is *discrete* [43]. Later, Abou-Faycal *et al.* proved this conjecture and showed that the capacity-achieving input distribution has a finite number of mass points with one of them located at the origin [44]. Under the same setting, Taricco and Elia showed that the capacity scales as $\alpha \log\log P \leq C \leq \log\log P + O(1)$ for some $\alpha \in (0,1)$ when $P \to \infty$ [45].

The double-logarithmic scaling of the capacity w.r.t. the SNR was proved rigorously in a more general setting by Lapidoth and Moser in [36]. They considered a MIMO channel under *generic* fading (see (1.34)). To bound the channel capacity, they used a dual expression and replaced the maximization over channel input distribution (as in (1.14)) by a minimization over channel output distribution (see Appendix 1.A.4). With this *duality* approach, they proved that the capacity scales as

$$C = \log\log P + \chi(\mathbf{H}) + o(1), \tag{1.38}$$

where $\chi(\mathbf{H})$ is the so-called *fading number* of the channel. Therefore, under fast fading, communication at rates significantly higher than the fading number is extremely power inefficient. Upper and lower bounds on the fading number are derived for various channels. In particular, the fading number for IID Rayleigh fading is bounded as

$$-\gamma - 1 \leq \chi(\mathbf{H}) \leq -N - \log\Gamma(N) + N\log \inf_{\boldsymbol{A}\in\mathbb{C}^{N\times N}:\ \det(\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}})=1} \mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}}), \tag{1.39}$$

where γ is Euler's constant and $\Gamma(\cdot)$ is the Gamma function. A sharper result for the capacity with IID Rayleigh fading is [36, Section IV-F.2]

$$\begin{aligned} C = \inf_{\alpha,\beta>0}\ \inf_{\delta>0} \Bigg\{ &-N + (N-\alpha)\psi(N) - \log\Gamma(N) + \alpha\log\beta + \log\Gamma\Big(\alpha, \tfrac{\delta}{\beta}\Big) \\ &+ \frac{N}{\beta}(1+P) + \frac{\delta}{\beta} + \frac{1-\alpha}{N-1}\delta\mathbb{1}\{\alpha \leq 1\}\Bigg\} + o(1) \end{aligned} \tag{1.40}$$

where $\psi(\cdot), \Gamma(\cdot)$, and $\Gamma(\cdot,\cdot)$ denote Euler's psi-function, the Gamma function, and the upper incomplete Gamma function, respectively. Note that in both (1.39) and (1.40), any choice of the arguments in the infima leads to valid upper bounds.

Under spatially correlated Rayleigh fast fading as in (1.35) (where $\breve{\mathbf{H}}$ has IID $\mathcal{N}_{\mathbb{C}}(0,1)$ entries), Jafar and Goldsmith showed that the capacity is a Schur-concave function of the vector of eigenvalues of $\boldsymbol{R}_t$ [46]. They also showed that the maximal possible capacity gain due to transmit correlations w.r.t. independent fading is $10\log_{10} M$ dB.

$T \geq 2$ (block fading)

The characteristics of the capacity scaling and capacity-achieving input distribution become rather different as soon as the channel remains constant for at least two channel uses. One of the first studies to address the capacity in this case was done by Marzetta and Hochward [47]. Assuming IID Rayleigh fading, they proved two important results on the capacity and the structure of the capacity-achieving signal as follows.

- The capacity obtained with $M > T$ is the same as the capacity obtained with $M = T$ transmit antennas for any T and N and arbitrary SNR. This is in contrast to the unbounded linear growth of capacity with $\min\{M, N\}$ when CSI is available (see (1.31)).
- The signal that achieves capacity can be represented as

$$\mathbf{X} = \mathbf{D}\boldsymbol{\Phi}, \tag{1.41}$$

  where $\mathbf{D}$ is an $M \times M$ real, non-negative, diagonal matrix, and $\boldsymbol{\Phi}$ is an independent $M \times T$ isotropically distributed truncated unitary matrix. That is, the probability density of $\boldsymbol{\Phi}$ is unchanged when postmultiplied by a $T \times T$ unitary matrix.

They also derived the capacity for the SISO case ($M = N = 1$), and showed that the noncoherent capacity approaches the coherent counterpart as $T \to \infty$ for a fixed number of antennas. This can be intuitively interpreted as when T is large, one can spend a negligible number of channel uses to estimate accurately the channel and then communicate coherently. The probability density of the received signal $\mathbf{Y}$ with isotropically distributed truncated unitary input signal was then derived in closed form by Hassibi and Marzetta in [48], enabling to evaluate the input-output mutual information with this input signal. It was shown numerically that at high SNR, the mutual information is maximized for $M = \min\{N, \frac{T}{2}\}$.

These results were later generalized and sharpened by Zheng and Tse in [49] for the $T \geq \min\{M, N\}+N$ case and by Yang *et al.* in [50] for the $T \leq M+N, M \leq \min\{N, \lfloor T/2 \rfloor\}$ case. These papers collectively showed that, under IID Rayleigh fading, for $T \geq 2\min\{M, N\}$, the capacity in the high-SNR regime is

$$C = M'\left(1 - \frac{M'}{T}\right)\log P + c(T, M, N) + o(1), \tag{1.42}$$

where $M' := \min\{M, N\}$ and $c(T, M, N)$ is a constant independent of the SNR given in [49, Eq.(24)] and [50, Eq.(9)] as

$$\begin{aligned} c(T, M, N) := \frac{1}{T}\log\left(\frac{\Gamma_{M'}(M')\Gamma_{M'}(\underline{L})}{\Gamma_{M'}(N)\Gamma_{M'}(T)}\right) + M'\left(1 - \frac{M'}{T}\right)\log\frac{T}{M'} \\ + \frac{M'\underline{L}}{T}\log\frac{N}{\underline{L}} + \frac{\overline{L}}{T}\left(\sum_{i=1}^{M'}\psi(N - i + 1) - M'\right) \end{aligned} \tag{1.43}$$

where $\underline{L} := \min\{N, T - M\}$, $\overline{L} := \max\{N, T - M\}$, $\psi(\cdot)$ is Euler's psi-function, and $\Gamma_m(a)$ is the complex multivariate Gamma function. The input distribution that achieves the capacity (1.42) inherits the structure in (1.41) with specified distributions of $\mathbf{D}$ and $\boldsymbol{\Phi}$.[7]

[7] Gaussian signaling, which was optimal had CSI been available, is sub-optimal and its achievable mutual information was evaluated in [51], [52].

In general, $\mathbf{\Phi}$ is a truncated unitary and isotropically distributed matrix. Whereas the distribution of $\mathbf{D}$ needs to be tailored according to the relation of the coherence interval and the number of antennas.

- If $T \geq M' + N$, it is optimal to let the first M' diagonal elements of $\mathbf{D}$ equal $\sqrt{\frac{PT}{M'}}$ and the remaining diagonal elements equal 0 with probability 1 [49]. That is to say, only M' transmit antennas are used to transmit the first M' rows of $\mathbf{\Phi}$ with equal power per antenna. The resulting input distribution is referred to as unitary space-time modulation (USTM) [53].
- If $T < M+N$ and $M \leq \min\{N, \lfloor T/2 \rfloor\}$, the squared diagonal elements of the optimal $\mathbf{D}$ have the same joint distribution as the ordered eigenvalues of a positive-definite $M \times M$ Beta-distributed random matrix [50]. That is, a power control is needed for the signals transmitted from different transmit antennas. The resulting input distribution is referred to as Beta-variate space-time modulation (BSTM). Note that in this case, the achievable rate with USTM input is at a constant gap below the capacity achieved with BSTM input. A closed-form PDF of the channel output with BSTM input was derived in [52], allowing to evaluate the mutual information at finite SNR.

When $1 < T < 2\min\{M, N\}$, only the pre-log factor of the capacity, so-called the DoF, was pointed out (without a rigorous proof) in [49, Section IV-D] to be $\lfloor T/2 \rfloor \left(1 - \frac{\lfloor T/2 \rfloor}{T}\right)$, i.e., the capacity scales as

$$C = \lfloor T/2 \rfloor \left(1 - \frac{\lfloor T/2 \rfloor}{T}\right) \log P + O(1). \tag{1.44}$$

Combining these cases, we have that the optimal DoF of the noncoherent MIMO block fading channel is given by $M^* \left(1 - \frac{M^*}{T}\right)$ with $M^* := \min\{M, N, \lfloor T/2 \rfloor\}$. Therefore, in the high SNR regime, using more than M^* transmit antennas does not increase, and in fact may penalize, the capacity.

Takeuchi *et al.* derived an achievable rate for the noncoherent MIMO block fading channel with successive decoding in the large system limit [54]. The paper [55] addresses the noncoherent MIMO channel with asymmetric link strengths where the channel coefficients are still independently zero-mean Gaussian distributed, but have different variances that can scale differently with the power P. In this case, the capacity-achieving input signal is similar to (1.35) with $\mathbf{D}$ replaced by a lower triangular matrix. The generalized DoF were derived in terms of the coherence interval and the exponents in the SNR-scaling of the channel gains for single-input multiple-output (SIMO), multiple-input single-output (MISO), and 2×2 MIMO channels. In particular, the optimal generalized DoF can be achieved for the SIMO and MISO channel by using only the statistically best receive and transmit antenna, respectively.

When the channel is in spatially correlated Rayleigh fading with the Kronecker model given by (1.35), Jafar and Goldsmith showed that the channel capacity depends on $\boldsymbol{R}_t$ through only the $\min\{T, M\}$ largest eigenvalues of $\boldsymbol{R}_t$ and is independent of the eigenvectors of $\boldsymbol{R}_t$ and $\boldsymbol{R}_r$ [46]. In contrast to the result for the IID Rayleigh fading where using more than T transmit antennas does not increase the capacity, [46] showed that the channel capacity increases almost surely with M as long as the channel coefficients are spatially correlated. As compared to (1.41), the capacity-achieving input is further premultiplied with the matrix containing the eigenvectors of $\boldsymbol{R}_t$ as its columns. Gohary *et al.* derived the high-SNR capacity with spatially correlated Rayleigh fading within an SNR-independent gap in [56].

Since most of the aforementioned works in the $T \geq 2$ case rely on the IID Rayleigh fading assumption, the following question is still open:

Question 1.1. *What is the capacity limit of the noncoherent MIMO P2P channel in* generic *block fading* (1.34)*?*

1.2.2.c Constellation Design

In this subsection, we focus on the constellation design for the $T \geq 2$ case and, motivated by the optimal DoF, assume that $M \leq \min\{N, \lfloor T/2 \rfloor\}$. We consider the USTM input

$$\mathbf{X} = \sqrt{\frac{PT}{M}}\mathbf{\Phi}, \quad \text{where } \mathbf{\Phi}\mathbf{\Phi}^{\mathsf{H}} = \boldsymbol{I}_M. \tag{1.45}$$

From the previous subsection, we know that, for IID Rayleigh fading channel at high SNR, USTM achieves a vanishing gap from the channel capacity if $T \geq M + N$, a constant gap from the channel capacity if $2M \leq T < M + N$, and the optimal DoF if $T \leq 2M$. In addition, constellation design according to USTM has a nice geometric interpretation as will be seen shortly.

With USTM, the input signal is invariant to rotation from the right and information is embedded in the row space of the matrix $\mathbf{\Phi}$. The intuition behind the optimality of USTM is that the channel matrix $\mathbf{H}$ only scales and rotates the bases of the transmitted signal matrix $\mathbf{X}$ without changing its subspace since the row spaces of $\mathbf{X}$ and the noise-free observation $\mathbf{HX}$ are the same. Information is carried in the position of the row-space of $\mathbf{\Phi}$ in the Grassmann manifold $G(\mathbb{C}^T, M)$, which is the space of M-dimensional subspaces of $\mathbb{C}^T$.[8] By definition, we see that $\mathbf{X}$ and $\mathbf{HX}$ represent the same element of $G(\mathbb{C}^T, M)$. At high SNR, the additive noise has low impact on the subspace of the output signal, and the subspace of $\mathbf{X}$ can be accurately recovered from the subspace of the noisy output.

From the above observation, a constellation for noncoherent MIMO communication over block fading can be designed as a set of representatives of M-dimensional subspaces in $\mathbb{C}^T$. Equivalently, these constellations represent a set of points on the Grassmann manifold $G(\mathbb{C}^T, M)$, i.e., $\mathbf{\Phi}^{\mathsf{T}}$ is drawn from

$$\mathcal{X} := \{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_{|\mathcal{X}|} \in \mathbb{C}^{T\times M} : \boldsymbol{X}_i^{\mathsf{H}}\boldsymbol{X}_i = \boldsymbol{I}_M, i \in [|\mathcal{X}|]\} \tag{1.46}$$

where we have that

$$\Big\{\mathrm{Span}(\boldsymbol{X}_1), \mathrm{Span}(\boldsymbol{X}_2), \ldots, \mathrm{Span}(\boldsymbol{X}_{|\mathcal{X}|})\Big\} \subset G(\mathbb{C}^T, M). \tag{1.47}$$

We let the truncated unitary matrix $\boldsymbol{X}_i \in \mathbb{C}^{T\times M}$ represent its column space $\mathrm{Span}(\boldsymbol{X}_i)$, which is a subspace in $\mathbb{C}^T$ and a point in $G(\mathbb{C}^T, M)$. Therefore, we write $\boldsymbol{X}_i \in G(\mathbb{C}^T, M)$ and $\mathcal{X} \subset G(\mathbb{C}^T, M)$. We refer to constellations of this kind as *Grassmannian constellations*.

Given the constellation size, the Grassmannian constellation construction can be interpreted as a packing of points in the Grassmann manifold. The ultimate packing criteria is to minimize the detection error under noisy observation. It was shown that the pair-wise error probability between two symbols $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$ decreases as any of the singular values of the matrix $\boldsymbol{X}_i^{\mathsf{H}}\boldsymbol{X}_j$ decreases [53, Theorem 5]. On the other hand, these singular values are also the principal angles between $\mathrm{Span}(\boldsymbol{X}_i)$ and $\mathrm{Span}(\boldsymbol{X}_j)$. Several distance metrics between two points in the Grassmannian represented by $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$ are defined as decreasing

[8] For a detailed review of the Grassmann manifold, see Appendix 1.B.

functions of these principal angles (see Appendix 1.B.2). Therefore, the error-minimization packing criterion typically amounts to maximizing the minimum pairwise distance between the constellation points, i.e.,

$$\mathcal{X} = \arg \max_{\{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_{|\mathcal{X}|}\}\subset G(\mathbb{C}^T,M)} \min_{1\leq i<j\leq|\mathcal{X}|} d(\boldsymbol{X}_i,\boldsymbol{X}_j) \tag{1.48}$$

where $d(\boldsymbol{X}_i,\boldsymbol{X}_j)$ is the considered distance. A commonly used distance measure is the *chordal distance* defined as

$$d(\boldsymbol{X}_i,\boldsymbol{X}_j) := \sqrt{M - \mathrm{tr}\{\boldsymbol{X}_i^{\mathsf{H}}\boldsymbol{X}_j\boldsymbol{X}_i^{\mathsf{H}}\boldsymbol{X}_j\}}. \tag{1.49}$$

Other distance measures include the geodesic distance, Fubini-Study distance, the chordal 2-norm, and the projection 2-norm (see, e.g., [57], [58]).[9] To assess the optimality of a packing, one can compare it to the packing efficiency limits in terms of the maximal minimum pairwise distance for a given constellation size derived in, e.g., [58], [60]–[62].

Based on this Grassmannian packing interpretation, a number of Grassmannian constellations have been proposed with different criteria, constellation generation, and detection methods. They follow two main approaches.

- The first approach uses numerical optimization tools to solve the sphere-packing problem on the Grassmannian by maximizing the minimum symbol pairwise distance (1.48) [63]–[67] or directly minimizing the error probability upper bound [68], [69]. This results in constellations with a good distance spectrum but no particular structure. Due to the lack of structure, this kind of constellation needs to be stored at both the transmitter and receiver, and decoded with the high-complexity ML decoder, which limits practical use to only small constellations.
- The second approach imposes particular structure on the constellation based on, e.g., algebraic construction [70]–[72], parameterized mappings of unitary matrices [73]–[75], concatenation of PSK and coherent space-time codes [76], or geometric motion on the Grassmannian [77]. The pilot-data structured input of a pilot-based scheme can also be seen as a noncoherent code [78]. The constellation structure facilitates low complexity constellation mapping and, probably, demapping.

Given a Grassmannian constellation, one can further optimize the input probabilities and per-antenna amplitudes for the constellation points at a given SNR as done in [79]. Grassmannian constellations can be used for a bit-interleaved coded modulation scheme with iterative soft demapping and decoding [80]. The performance of Grassmannian signaling and some other differential schemes over temporally-correlated channels has been investigated in [81]. Therein, it was shown numerically that noncoherent communication has clear advantage over coherent schemes in medium to high mobility scenarios, i.e., for short coherence intervals. Grassmannian signaling was validated on a software-defined radio testbed in [82], showing that noncoherent techniques are more robust to system impairments than the coherent Alamouti approach. Furthermore, Grassmannian signaling was shown to achieve tight bounds on the ergodic high-SNR capacity of the noncoherent MIMO full-duplex relay channel [83].

In this thesis, we would like to give our answer to the following question.

[9] The Kullback-Leibler (KL) distance metric is used in [59] for constellation points belonging to hyperspheres of different radii and not necessarily the Grassmann manifold.

Question 1.2. *How to design a Grassmannian constellation that has a high packing efficiency while being simple to generate (thus available for large constellation size and large symbol length, and not requiring to be stored), admitting a simple and effective labeling scheme, and allowing for an efficient detection scheme?*

1.2.3 The Multiple-Access Channel

1.2.3.a Channel Model

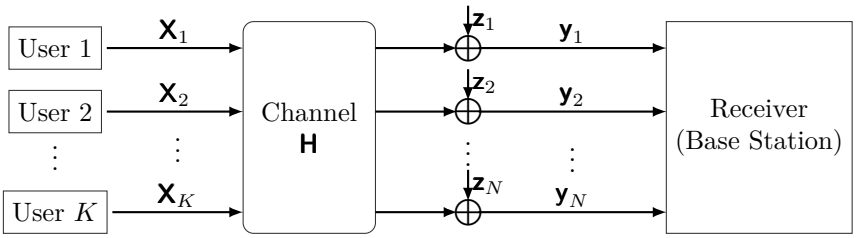


FIGURE 1.4: Conceptual illustration of a K-user MIMO MAC with N receive antennas in block fading.

We consider a MIMO MAC with K transmitters (so-called users) transmitting to a receiver (so-called base station), as illustrated in Fig. 1.4. User k is equipped with M_k antennas, $k \in [K]$, while the receiver has N antennas. The channel of user k is represented by the matrix $\mathbf{H}_k \in \mathbb{C}^{N \times M_k}$ whose (i,j)-th element is the random fading coefficient between the i-th receive antenna and the j-th transmit antenna of user k. The channels are in block fading with equal-length and synchronous coherence blocks (across users) of length T. Only the channel distribution is known to the transceiver. During a coherence block b, user k sends an $M_k \times T$ signal matrix $\mathbf{X}_k[b]$ across M_k antennas and T channel uses, and the receiver receives an $N \times T$ signal matrix

$$\mathbf{Y}[b] = \sum_{k=1}^{K} \mathbf{H}_k[b]\mathbf{X}_k[b] + \mathbf{Z}[b], \quad b = 1, 2, \ldots, \tag{1.50}$$

where $\mathbf{Z}[b] \in \mathbb{C}^{N \times T}$ is AWGN with IID $\mathcal{N}_{\mathbb{C}}(0,1)$ entries independent of $\mathbf{H}$. Unless stated otherwise, we consider the power constraint

$$\frac{1}{\nu} \sum_{b=1}^{\nu} \|\mathbf{X}_k[b]\|_{\mathrm{F}}^2 \leq PT, \quad k \in [K], \tag{1.51}$$

where ν is the number of the blocks spanned by a channel codeword. P is the SNR of the channel. The most important difference of the MAC to the P2P channel is the independence between the signals transmitted from the antennas of different users.

1.2.3.b Fundamental Limits

Shamai and Marzetta studied the capacity of the SIMO MAC ($M_k = 1, k \in [K]$) in IID Rayleigh fading in [84]. For fast fading ($T = 1$), they showed that the sum capacity for $K > 1$ users is equal to the capacity for $K = 1$ user, thus TDMA is optimal. For block fading ($T > 1$), they conjectured that the maximum sum capacity can be achieved by no more than $K = T$ users, which is supported by some asymptotic analysis, namely, high SNR and large T, for a fixed M/T ratio. In the same setting with $T > 1$, Gopalan *et al.* derived a constructive lower bound on the sum capacity from a successive decoding

scheme [85]. Since the sum capacity of the MAC can be upper bounded by the capacity of the P2P channel by allowing user cooperation, the MAC sum capacity under fast generic fading can be shown to scale as a double logarithm of the SNR plus a fading number as in (1.38). The fading number of the MISO MAC ($N = 1$) in Rician fading has been derived by Lin and Moser in [86]. It is identical to the fading number of the single-user Rician fading channel obtained when only the user with the best channel is activated and all other users are silent. This holds also with individual per-user peak power constraint. Devassy *et al.* provided non-asymptotic upper and lower bounds on the sum capacity of the MIMO MAC under Rayleigh fading in [87]. Therein, the lower bound derived with independent Grassmannian signaling from each user has a small gap to the upper bound even at moderate SNR.

The aforementioned works address the sum capacity of the noncoherent MAC. The full capacity region is unknown, and only some achievable DoF regions have been proposed. An *achievable* DoF region for the two-user MIMO MAC under IID Rayleigh block fading was proposed in [88] using a geometric approach. Specifically, assuming that $M_1 + M_2 \leq N$ and $T \geq M_1 + M_2 + N$, this achievable DoF region is the convex hull of the origin and three DoF pairs

$$\left(M_1\left(1 - \frac{M_1}{T}\right), 0\right), \tag{1.52a}$$

$$\left(0, M_2\left(1 - \frac{M_2}{T}\right)\right), \tag{1.52b}$$

$$\text{and } \left(M_1\left(1 - \frac{M_1 + M_2}{T}\right), M_2\left(1 - \frac{M_1 + M_2}{T}\right)\right). \tag{1.52c}$$

A generalization of this achievable DoF region to the K-user case is presented in [37, Theorem 5] as

$$d_j = M'_j\left(1 - \frac{\sum_{j\in\mathcal{J}} M'_j}{T}\right), \quad \text{for } j \in \mathcal{J} = \{k_1, k_2, \ldots, k_J\} \subseteq [K], \tag{1.53}$$

where $M'_j = \min\left\{M_j, \left[N - \sum_{m=1}^{j-1} M'_{k_m}\right]^+\right\}$ and $T \geq 2N$. These achievable regions can be achieved by a simple pilot-based scheme. A cooperative outer bound for the optimal DoF region was also given in [37, Section VI-B] as

$$\sum_{j\in\mathcal{J}} d_j \leq \min\left\{N, \sum_{j\in\mathcal{J}} M_j\right\}\left(1 - \frac{\min\left\{N, \sum_{j\in\mathcal{J}} M_j\right\}}{T}\right), \quad \forall \mathcal{J} \subseteq [K]. \tag{1.54}$$

It can be seen that the achievable DoF regions (1.52) and (1.53) are sum-DoF optimal.

In this thesis, we will investigate the *optimal* DoF region of the noncoherent MAC and address the following open question.

Question 1.3. *What is the optimal DoF region for the noncoherent MIMO MAC in* generic *block fading?*

1.2.3.c Constellation Design

As in the single-user (P2P) case, the transmitted signals $\mathbf{X}_k$ of user k are normally drawn from a finite discrete constellation $\mathcal{X}_k$, so $\mathcal{X} := \{[\boldsymbol{X}_1\ \boldsymbol{X}_2\ \ldots\ \boldsymbol{X}_K] : \boldsymbol{X}_k \in \mathcal{X}_k, k \in [K]\}$ is the joint constellation for the MAC.

A straightforward extension of the pilot-based scheme for the single-user case is to divide the coherent block into two parts: 1) the training part in which mutually orthogonal pilot sequences are sent to estimate the CSI for each user, and 2) the data transmission part in which different users communicate simultaneously [89] using a scalar constellation (e.g., PAM, QAM, PSK). The optimality of this approach in terms of achievable rate and detection error remains unclear. Furthermore, it is not always possible to assign orthogonal pilots to all the users in the system before the transmission, as discussed in Section 1.1.5.

An amplitude-based encoding scheme was proposed in [90], [91], but the accompanying energy detector relies on a large number of receive antennas so that the average received power across all antennas concentrates. Also with a massive receive antenna array, some differential encoding schemes were investigated based on PSK [92], [93] or QAM [94]. Another line of work is based on PSK constellations which are absolutely additively uniquely decomposable, i.e., each individual PSK symbol can be uniquely decoded from any linear combination of two PSK constellation points with positive weights [95], [96]. In this scheme, the unique decodability of the signal matrix relies on the asymptotic orthogonality between the users' channels in the massive MIMO regime. A similar uniquely decomposable property was also exploited for QAM-based multi-user space-time modulation [97]. In short, these schemes assume a large number of receive antennas.

Unlike the single-user case where a good constellation design criterion is known to be the sphere packing in the Grassmann manifold, for the noncoherent MIMO MAC in flat and block fading, a simple and effective constellation design criterion remains unclear. In general, the joint constellation $\mathcal{X}$ for the MAC should be designed so as to minimize the symbol detection error. If the users could cooperate, the system could be seen as a $\left(\sum_{k=1}^{K} M_k\right) \times N$ MIMO P2P noncoherent channel, for which the high-SNR optimal input is uniformly distributed on the Grassmannian $G(\mathbb{C}^T, \sum_{k=1}^{K} M_k)$ [49]. Inspired by this, the joint constellation for the MAC can be treated as a Grassmannian constellation in $G(\mathbb{C}^T, \sum_{k=1}^{K} M_k)$, which leads to a design criterion mimicking sphere packing in this Grassmannian by maximizing the minimum pairwise chordal distance. Brehler and Varanasi derived the error probability of the ML detector for the MIMO MAC in [98] and analyzed the high-SNR asymptotics. However, an explicit constellation design criterion was not given and the analysis led to a design for the MAC with *cooperating* users only. With cooperating users, the design criterion is similar to that for a single-user channel proposed in [68, Eq.(8)] by the same authors, which is different from the max-min pairwise chordal distance criterion. This criterion can be used for the non-cooperating users case by modifying the optimization space. It guarantees the full diversity order in the single-user case [68], but depends on the number of receive antennas. The pairwise error exponent can be shown to converge to the KL divergence between the output distributions conditioned on either of the symbols transmitted [59]. Based on this, a criterion consisting in maximizing the minimum KL divergence was proposed for the MAC in [97]. However, this work focuses on QAM-based space-time modulation and only uses the criterion to optimize the transmit powers and the sub-constellation assignment.

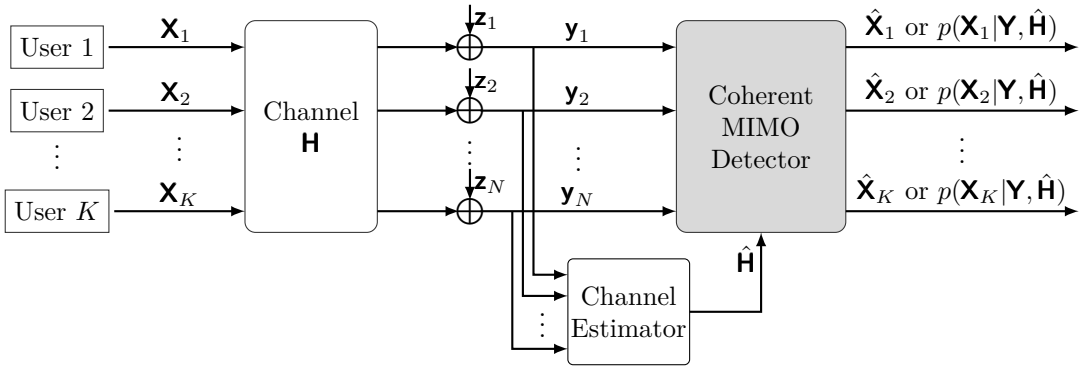
In the thesis, we would like to answer the following question to a further extent:

Question 1.4. *How to design effective joint constellation for the noncoherent MIMO MAC in order to achieve a low symbol error rate (SER)?*

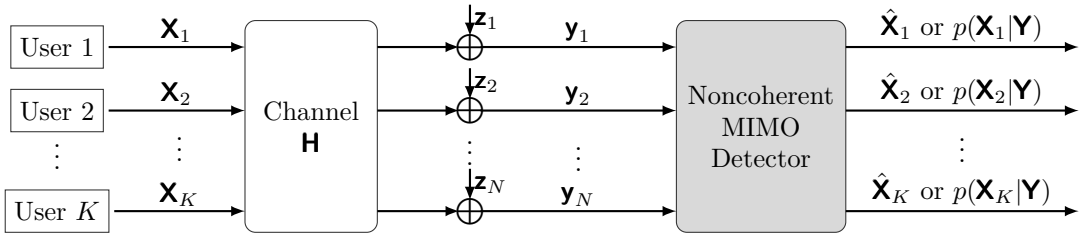
1.2.3.d Multi-User Detection

We now move the focus to the receiver, whose task is to detect the transmitted symbols $\mathbf{X}_k$ (or rather the underlying bits) based on the noisy observation $\mathbf{Y}$ (1.50). Since the optimal ML coherent detection problem is NP-hard, the receiver can use a sub-optimal detector to reduce complexity.

If the transmitted signals contain pilots, the receiver can estimate (normally imperfectly) the channel based on the pilot symbols, then perform coherent detection based on the channel estimate, as illustrated in Fig. 1.5(a). This approach leverages a rich literature of coherent detection, in which many sub-optimal algorithms have been proposed, as listed in Section 1.1.4. Channel estimation and data detection can also be done iteratively [99], [100], or jointly based on tree search [101], [102]. These schemes require pilot transmission for an initial channel estimate or to guarantee the identifiability of the data symbols. On the other hand, with pilot-free transmission, the noncoherent detector does not perform explicit channel estimation and exploits the statistical knowledge of $\mathbf{H}$ only, as shown in Fig. 1.5(b).



(a) Coherent detection with channel estimation



(b) Noncoherent detection

FIGURE 1.5: Conceptual illustration of coherent and noncoherent detection schemes for a K-user MIMO MAC with N receive antennas in block fading. The outputs of the detector are the hard-detected symbols or their marginal posteriors.

The receiver might be interested in not only the hard detection of the symbols, but also their posterior marginal PMF. This "soft" information is needed, for example, when computing the bit-wise log-likelihood ratios (LLRs) required for soft-input soft-output channel decoding. Computing an exact marginal PMF would require enumerating all possible combinations of other-user signals, which is infeasible with many users, many antennas, or large constellations. Thus, as for hard detection, a sub-optimal scheme is needed. In contrast to probabilistic coherent MIMO detection, for which many schemes have been proposed (e.g., [103]–[105]), the probabilistic noncoherent MIMO detection under

general signaling, and Grassmannian signaling in particular, has not been well investigated. Therefore, in this thesis, we would like to answer the following question:

Question 1.5. *How to efficiently detect the symbols and estimate their marginal posteriors in the noncoherent MIMO MAC with general constellations?*

1.2.4 The Broadcast Channel

1.2.4.a Channel Model

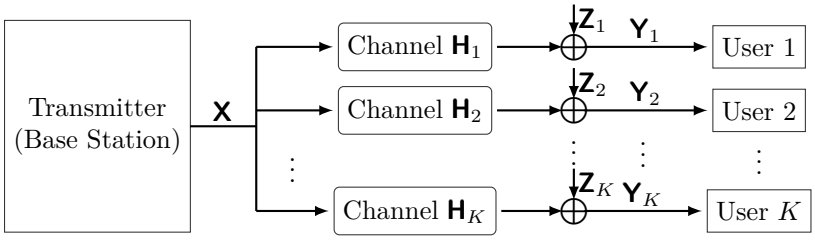


FIGURE 1.6: Conceptual illustration of a K-user MIMO BC with M transmit antennas in block fading.

We consider a MIMO BC in which an M-antenna base station transmit to K users as illustrated in Fig. 1.6. User k is equipped with N_k antennas, $k \in [K]$. The channel $\mathbf{H}_k \in \mathbb{C}^{N_k \times M}$ between the base station and user k is flat and block fading with equal-length and synchronous coherence block (across the users) of length T. Let the matrix $\mathbf{X}[b] \in \mathbb{C}^{M \times T}$ be the transmitted signal from the M antennas during the coherence block b. The received signal matrix at user k during interval b is

$$\mathbf{Y}_k[b] = \mathbf{H}_k[b]\mathbf{X}[b] + \mathbf{Z}_k[b], \quad k \in [K], \tag{1.55}$$

where $\mathbf{Z}_k[b] \in \mathbb{C}^{N_k \times T}$ is the AWGN with $\mathcal{N}_{\mathbb{C}}(0,1)$ entries. The input is subject to the power constraint

$$\frac{1}{\nu}\sum_{b=1}^{\nu} \|\mathbf{X}[b]\|_{\mathrm{F}}^2 \leq PT, \tag{1.56}$$

where ν is the number of blocks spanned by a channel codeword. P is the SNR of the channel.

1.2.4.b Fundamental Limits and Transmission Schemes

Most studies on the capacity limits of the BC assume perfect channel state information at the receivers (CSIRs) and different levels of channel state information at the transmitter (CSIT), namely, perfect CSIT [106], [107] (with dirty paper coding (DPC)), imperfect (partial) CSIT [108]–[110], and no CSIT [111], [112]. The noncoherent (no CSIT, no CSIR) MISO BC with $\mathbf{H}_k$ isotropically distributed was mentioned briefly by Jafar and Goldsmith in [111, Section VII-C]. They showed that their proposed *scalar upper bound* on the capacity region, which was successfully applied to the perfect CSIR case, becomes loose for the noncoherent BC since it fails to account for the DoF loss due to the lack of CSIR. Under IID Rayleigh block fading, Fadel and Nosratinia found the optimal DoF region of the noncoherent MIMO BC given by [37, Theorem 1]

$$\mathcal{D} = \left\{ (d_1, \ldots, d_K) \in \mathbb{R}_+^K : \sum_{k=1}^{K} \frac{d_k}{N_k^*(1 - \frac{N_k^*}{T})} \leq 1 \right\}, \tag{1.57}$$

where $N_k^* := \min\{M, N_k, \lfloor T/2 \rfloor\}$ and $\mathbb{R}_+$ denotes the set of non-negative real numbers. This is achievable with a TDMA scheme and the transmitter uses N_k^* antennas while transmitting to user k.

The aforementioned results assume the statistical independence between each pair of transmit and receive antennas. In practice, however, the channels between different antennas are often correlated because the propagation environment often causes stronger received signal gains in some spatial directions, and also due to the spatially dependent patterns of the antennas. Since the users are not co-located, they may have non-identical correlation matrices. In this case, a useful tool leveraging the difference between the spatial correlations observed by different users is *transmit correlation diversity*. For transmit spatial correlation matrices that have mutually exclusive eigenspaces, transmit correlation diversity can be harvested using a joint spatial division multiplexing (JSDM) transmission scheme [113], [114] that reduces the overhead needed for channel estimation. The main idea of JSDM is to partition the users into groups with approximately the same channel correlation eigenspace. Another transmission scheme in *product superposition* proposed by Li and Nosratinia in [115], [116] for the BC with mixed static (with CSIR) and dynamic (no CSIR) users and then applied to spatially correlated noncoherent BCs with fully overlapping correlation eigenspaces [117].

Note that the JSDM transmission scheme attempts to exploit the *non-overlapping* (mutually exclusive) parts of the correlation eigenspaces, while product superposition exploits the *overlapping* parts. In this thesis, we would like to exploit both parts by answering the following question.

Question 1.6. *What are the fundamental limits of the noncoherent spatially correlated BC in generic block fading with* partially *overlapping correlation eigenspaces and how to effectively exploit transmit correlation diversity under this condition?*

1.3 Thesis Description

This thesis is a contribution to noncoherent wireless communications. We assume throughout that the channel distribution is known, but the CSI is not known by any communicating node, and channel coefficients are in block fading with coherence interval T. The assumption on channel distribution (Rayleigh/generic, correlated/uncorrelated) will be specified in each chapter. We study the P2P channel, the MAC, and the BC in terms of fundamental limits (capacity, achievable rate, DoF) and transceiver design (constellation design, efficient detection design). The main contribution of the thesis is the answer or extension of existing answers to the questions posed in the previous section. The detailed outline and contributions of the thesis are presented next.

1.3.1 Thesis Outline and Contributions

The remainder of the thesis is organized as follows. At the end of this chapter, we provide the common mathematical preliminaries. After that, the thesis comprises two main parts addressing respectively the fundamental limits and transceiver design aspects of noncoherent communications, followed by the conclusion and outlook.

Part I: Fundamental Limits of Noncoherent Communications

In *Chapter 2*, we partly answer *Question 1.1* by studying the optimal DoF of the noncoherent MIMO P2P channel in generic block fading (1.34). We show that the optimal DoF for the MIMO P2P channel under IID Rayleigh fading (found in [49], [50]) also holds under generic fading. We introduce a novel converse proof technique based on a genie-aided bound and the duality approach [36]. The results in this chapter are presented in [118], [119].

In *Chapter 3*, we respond to *Question 1.3*, i.e., the optimal DoF region for the noncoherent MAC in generic block fading, in the SIMO case. We prove that the achievable DoF region in (1.52) when $M_1 = M_2 = 1$ is optimal. The converse proof follows by showing that the average power constraint can be replaced by a peak power constraint without changing the optimal DoF region, together with a similar genie-aided bound and the duality approach as in Chapter 2. The results in this chapter are presented in [118], [119].

Chapter 4 presents our answer to *Question 1.6*. Considering a spatially correlated MIMO BC with partially overlapping correlation eigenspaces, we exploit transmit correlation diversity by carefully designing transmission schemes based on rate splitting, product superposition [115], and a hybrid version of them. In doing so, we find some achievable rate and DoF regions for the BC in the two-user case, and some achievable DoF regions in the K-user case. The achievable DoF regions significantly improve over TDMA, which was shown to be DoF-optimal for uncorrelated fading [37]. As a by-product, we also derive an achievable rate for the MIMO P2P channel under spatially correlated fading. The result in this chapter was published in [120] and [121].

Part II: Transceiver Design for Noncoherent Communications

Chapter 5 gives our answer to *Question 1.2* in the SIMO case. We propose a structured constellation in the Grassmannian of lines so-called *cube-split constellation*. It is generated by partitioning the Grassmannian of lines into a collection of bent hypercubes and defining a mapping onto each of these bent hypercubes such that the resulting symbols are approximately uniformly distributed on the Grassmannian. This constellation fulfills all the desired characteristics in Question 1.2: it has a high packing efficiency represented by the minimum pairwise chordal distance while being simple to generate (thus available

for large constellation size and large symbol length, and not requiring to be stored); it admits a simple and effective binary labeling scheme; and it allows for efficient hard and soft detection. These advantages over the pilot-based scheme and other structured Grassmannian constellations are more pronounced in the regime of short coherence interval and large constellation size, as verified by numerical results. The result in this chapter was published in [122], [123], and [124].

In *Chapter 6*, we present a joint constellation design for the noncoherent MIMO MAC in IID Rayleigh block fading, thus give our answer to *Question 1.4*. We analyze the ML detection error and introduce novel design criteria so as to minimize the error probability. We further simplify the metrics by a high-SNR analysis. Our metrics can be used for joint constellation construction by (numerically) optimizing them over the set of signal matrices. Moreover, based on these metrics, we propose two simple constructions consisting respectively in partitioning a single-user constellation or precoding single-user constellations of lower dimension. We investigate the option of building each individual constellation as a Grassmannian constellation scaled by the respective transmit power. Numerical results show that our proposed metrics are meaningful, and the resulting constellations perform better, for the same transmission rate and power, than a pilot-based scheme and the constellations optimized with existing metrics. The results in this chapter are presented collectively in [125], [124], [126], and [127].

In *Chapter 7*, we focus on the receiver side of a MAC and answer *Question 1.5*. We propose a noncoherent multi-user soft detection scheme for the SIMO MAC under spatially correlated Rayleigh block fading. Our detector is based on expectation propagation (EP) approximate inference and has polynomial complexity in the number of users, number of receive antennas and coherence interval. We also propose two simplifications of this detector with reduced complexity. The proposed detectors can be used for general signaling with vector-valued symbols transmitted over each coherence block. In this chapter, we also propose an efficient detection scheme for the precoding-based constellation in Chapter 6, which has lower complexity but performs inferior to the EP detector. With pilot-assisted signaling, the EP detector outperforms, in terms of symbol error rate, some conventional coherent pilot-based detectors, namely, a linear MMSE decoder, a sphere decoder and a joint channel estimation–data detection scheme. Our EP-based detectors produce accurate approximates of the true posterior leading to high achievable sum-rates. The gains of these detectors are further observed in terms of the bit error rate when using their soft outputs for a turbo channel decoder. The results in this chapter were published in [128], [125] and [129].

Conclusion and Outlook

We conclude the thesis in *Chapter 8* which summarizes the main results and provides an outlook to future work. We will put noncoherent communications, especially what we have developed in this thesis, in the context of other (emerging) topics in wireless communications.

Finally, a French summary is provided in Appendix A.

We summarize the main results of the thesis in Table 1.1.

1.3.2 List of Publications

The publications included in the main result of this thesis are listed below.

TABLE 1.1: The main results of the thesis

Channel	Part I: Fundamental Limits		Part II: Transceiver Design	
	Achievable rate/DoF	Optimal DoF	Constellation design	Detection design
P2P	Chapter 4	Chapter 2	Chapter 5	
MAC	—	Chapter 3	Chapter 6	Chapter 7
BC	Chapter 4	—	—	—

Journal Articles

[119] K.-H. Ngo, S. Yang, and M. Guillaud, "The optimal degrees of freedom for the point-to-point and multiple-access channels in generic block fading," *in preparation to submit to IEEE Trans. Inf. Theory*, 2020.

[127] K.-H. Ngo, S. Yang, M. Guillaud, and A. Decurninge, "Joint constellation design for the noncoherent MIMO multiple-access channel," *in preparation to submit to IEEE Trans. Wireless Commun.*, 2020.

[121] F. Zhang, K.-H. Ngo, S. Yang, and A. Nosratinia, "Transmit correlation diversity: Generalization, new techniques, and improved bounds," *submitted to IEEE Trans. Inf. Theory*, 2020.[10]

[129] K.-H. Ngo, M. Guillaud, A. Decurninge, S. Yang, and P. Schniter, "Multi-user detection based on expectation propagation for the noncoherent SIMO multiple access channel," *IEEE Trans. Wireless Commun.*, 2020, (arXiv preprint arXiv:1905.11152).

[123] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Cube-split: A structured Grassmannian constellation for noncoherent SIMO communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1948–1964, Mar. 2020.

Preprint

[126] K.-H. Ngo, S. Yang, M. Guillaud, and A. Decurninge, "Joint constellation design for the two-user noncoherent multiple-access channel," arXiv preprint arXiv:2001.04970, 2020.

Conference Papers

[128] K.-H. Ngo, M. Guillaud, A. Decurninge, S. Yang, S. Sarkar, and P. Schniter, "Non-coherent multi-user detection based on expectation propagation," in *53rd Asilomar Conference on Signals, Systems, and Computers*, CA, USA, Nov. 2019.

[125] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "A multiple access scheme for noncoherent SIMO communications," in *52nd Asilomar Conference on Signals, Systems, and Computers*, CA, USA, Oct. 2018, pp. 1846–1850.

[118] K.-H. Ngo, S. Yang, and M. Guillaud, "The optimal DoF region for the two-user noncoherent SIMO multiple-access channel," in *IEEE Information Theory Workshop (ITW)*, arXiv preprint arXiv:1806.04102, Guangzhou, China, Nov. 2018.

[120] K.-H. Ngo, S. Yang, and M. Guillaud, "An achievable DoF region for the two-user noncoherent MIMO broadcast channel with statistical CSI," in *IEEE Information Theory Workshop (ITW)*, Nov. 2017, pp. 604–608.

[10] F. Zhang and K.-H. Ngo contributed equally to the technical content of this paper.

[122] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Design and analysis of a practical codebook for noncoherent communications," in *51st Asilomar Conference on Signals, Systems, and Computers*, CA, USA, Oct. 2017, pp. 1237–1241.

Patent

[124] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Transmitter and receiver communication apparatus for noncoherent communication," English, European patent, Application number 6 860 390, Filed on 30 Oct. 2018.

1.3.3 Contributions Outside the Scope of the Thesis

During the course of the thesis, we have also published some other contributions which are not included in the main result of this thesis:

[130] K.-H. Ngo, S. Yang, and M. Guillaud, "Generalized Gaussian model for data-driven learning in communications," in *International Zurich Seminar on Information and Communication (IZS)*, poster, Zurich, Switzerland, Feb. 2020.

[131] K.-H. Ngo, S. Yang, and M. Kobayashi, "Scalable content delivery with coded caching in multi-antenna fading channels," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 548–562, Jan. 2018.

[132] A. Ghorbel, K.-H. Ngo, R. Combes, M. Kobayashi, and S. Yang, "Opportunistic content delivery in fading broadcast channels," in *IEEE Global Communications Conference (GLOBECOM)*, Singapore, Dec. 2017.

[133] T. T. Q. Tran, V.-L. Nguyen, K.-H. Ngo, L.-T. Nguyen, Q.-T. Nguyen, E. Bastug, S. Azarian, M. Debbah, and P. Duhamel, "Network coding and information security in industry 4.0," in *1st ASEAN IVO Workshop on Cybersecurity and Information Security in Industry 4.0*, poster, Hanoi, Vietnam, Mar. 2019.

In [130], we propose a *generalized Gaussian model* for data-driven model identification in MIMO communications. [131] and [132] are extensions of my master thesis and address scalable content delivery with coded caching in fading broadcast channel by using multiple antennas or opportunistically scheduling the users. [133] contains an extension of my bachelor thesis on software-defined radio implementation of OFDM-based network coding and cognitive radio.

1.A Appendix: Information Theoretic Functions and the Duality Approach

In this appendix, based on [22, Chapter 2], we define and enlist some properties of some relevant information theoretic functions. Then, we review the duality approach in [36].

1.A.1 Entropy and Differential Entropy

Entropy

Let x be a discrete random variable with domain $\mathcal{X}$ and PMF p_{x}. The *entropy* of x, which is the uncertainty about the outcome of x, is defined as

$$H(\mathsf{x}) := -\sum_{x\in\mathcal{X}} p_{\mathsf{x}}(x)\log p_{\mathsf{x}}(x) = -\mathbb{E}_{\mathsf{x}}[\log p_{\mathsf{x}}(\mathsf{x})], \tag{1.58}$$

where the base of the log is arbitrary and determines the information unit. In this thesis, we consider base 2, i.e., information is measured in bits. $H(\mathsf{x})$ is a concave function in p_{x} and satisfies

$$0 \le H(\mathsf{x}) \le \log|\mathcal{X}|. \tag{1.59}$$

Let x be an arbitrary random variable and $\mathsf{y}\,|\,\{\mathsf{x}=x\}$ be discrete for every x, the *conditional entropy* of y given x is given by

$$H(\mathsf{y}\,|\,\mathsf{x}) := -\mathbb{E}_{\mathsf{x},\mathsf{y}}[\log p_{\mathsf{y}\,|\,\mathsf{x}}(\mathsf{y}\,|\,\mathsf{x})]. \tag{1.60}$$

It holds that

$$H(\mathsf{y}\,|\,\mathsf{x}) \le H(\mathsf{y}), \tag{1.61}$$

i.e., conditioning reduces entropy. The *joint entropy* of a pair of discrete random variables $(\mathsf{x},\mathsf{y}) \sim p_{\mathsf{x},\mathsf{y}}$ is defined as

$$H(\mathsf{x},\mathsf{y}) := -\mathbb{E}_{\mathsf{x},\mathsf{y}}[\log p_{\mathsf{x},\mathsf{y}}(\mathsf{x},\mathsf{y})]. \tag{1.62}$$

A chain rule for joint entropy follows

$$H(\mathsf{x},\mathsf{y}) = H(\mathsf{x}) + H(\mathsf{y}\,|\,\mathsf{x}) = H(\mathsf{y}) + H(\mathsf{x}\,|\,\mathsf{y}). \tag{1.63}$$

Applying inductively this chain rule, the entropy of a discrete random vectors $\mathbf{x} = \{\mathsf{x}_i\}_{i=1}^n$ is defined as

$$\begin{aligned} H(\mathbf{x}) &= H(\mathsf{x}_1) + H(\mathsf{x}_2\,|\,\mathsf{x}_1) + \cdots + H(\mathsf{x}_n\,|\,\mathsf{x}_1,\ldots,\mathsf{x}_{n-1}) && (1.64)\\ &= \sum_{i=1}^{n} H(\mathsf{x}_i\,|\,\mathsf{x}_1,\ldots,\mathsf{x}_{i-1}). && (1.65) \end{aligned}$$

Differential Entropy

Let x be a continuous random variable with domain $\mathcal{X}$ and PDF p_{x}. The *differential entropy* of x is defined as

$$h(\mathsf{x}) := -\int_{\mathcal{X}} p_{\mathsf{x}}(x)\log p_{\mathsf{x}}(x) = -\mathbb{E}_{\mathsf{x}}[\log p_{\mathsf{x}}(\mathsf{x})]. \tag{1.66}$$

Unlike the entropy, the differential entropy can be negative. Similar as for the entropy, we can define the *conditional different entropy* as

$$h(\mathsf{y}\,|\,\mathsf{x}) := -\mathbb{E}_{\mathsf{x},\mathsf{y}}[\log p_{\mathsf{y}\,|\,\mathsf{x}}(\mathsf{y}\,|\,\mathsf{x})] \leq h(\mathsf{y}), \tag{1.67}$$

the *joint differential entropy* of $\mathbf{x} = \{\mathsf{x}_i\}_{i=1}^n$ with joint PDF $p_{\mathbf{x}}$ as

$$h(\mathbf{x}) := -\mathbb{E}_{p_{\mathbf{x}}}\big[\log p_{\mathbf{x}}(\mathbf{x})\big] = \sum_{i=1}^{n} h(\mathsf{x}_i\,|\,\mathsf{x}_1,\ldots,\mathsf{x}_{i-1}). \tag{1.68}$$

Furthermore, the differential entropy is invariant under translation:

$$h(\mathbf{x}+\boldsymbol{a}) = h(\mathbf{x}), \tag{1.69}$$

for any deterministic vector $\boldsymbol{a}$, but not under scaling:

$$h(\boldsymbol{A}\mathbf{x}) = h(\mathbf{x}) + \log|\det(\boldsymbol{A})|^2, \tag{1.70}$$

for any non-singular deterministic *complex* matrix $\boldsymbol{A}$. In particular, if $\boldsymbol{A}$ is a unitary matrix, then $|\det(\boldsymbol{A})|^2 = \det(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}) = \det(\boldsymbol{I}) = 1$, thus $h(\boldsymbol{A}\mathbf{x}) = h(\mathbf{x})$, i.e., *rotating does not change differential entropy*. The following lemma is obtained by generalizing the real counterpart in [22, Chapter 2.2] to the complex case.

Lemma 1.1 (Maximum differential entropy)**.** *For a complex random vector* $\mathbf{x} = \{\mathsf{x}_i\}_{i=1}^n$*,*

$$h(\mathbf{x}) \leq \log\det(\pi e\mathbb{E}[\mathbf{x}\mathbf{x}^{\mathsf{H}}]) \leq \sum_{i=1}^{n}\log\Big(\pi e\mathbb{E}\Big[|\mathsf{x}_i|^2\Big]\Big), \tag{1.71}$$

where the first inequality holds with equality if and only if $\mathbf{x}$ *is a zero-mean Gaussian vector, and the second inequality holds with equality if and only if the entries of* $\mathbf{x}$ *are mutually independent.*

1.A.2 Kullback-Leibler Divergence

Let P and Q be two (discrete or continuous) probability measures such that P is absolutely continuous w.r.t. Q, then the KL divergence (also known as relative entropy) from P to Q is defined as

$$D(P\|Q) = \int \log\frac{\mathrm{d}P}{\mathrm{d}Q}\,\mathrm{d}P, \tag{1.72}$$

where $\mathrm{d}P/\,\mathrm{d}Q$ is the Radon-Nikodym derivative [134]. Note that the KL divergence is asymmetric (hence the term "from P to Q"). Let $p(\cdot)$ and $q(\cdot)$ be respectively the PDF/PMF of P and Q. For a random vector $\mathbf{x}$ with domain $\mathcal{X}$,

$$D(P\|Q) := \begin{cases} \int_{\mathcal{X}} p(\boldsymbol{x})\log\frac{p(\boldsymbol{x})}{q(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} & \text{if } \mathcal{X} \text{ is continuous,} \\ \sum_{\mathbf{x}\in\mathcal{X}} p(\mathbf{x})\log\frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} & \text{if } \mathcal{X} \text{ is discrete.} \end{cases} \tag{1.73}$$

We have that

$$D(P\|Q) \geq 0 \tag{1.74}$$

with equality if and only if $p(\boldsymbol{x}) = q(\boldsymbol{x})$ almost everywhere.

1.A.3 Mutual Information

For two random variables x and y, the *mutual information* $I(\mathsf{x};\mathsf{y})$ is defined as the KL divergence from the joint measure to the product measure of x and y:

$$I(\mathsf{x};\mathsf{y}) := D(p_{\mathsf{x},\mathsf{y}} \| p_{\mathsf{x}} p_{\mathsf{y}}). \tag{1.75}$$

Thus the mutual information is non-negative and becomes zero if and only if $p_{\mathsf{x},\mathsf{y}}(x,y) = p_{\mathsf{x}}(x)p_{\mathsf{y}}(y), \forall x, y$, i.e., x and y are independent. Let the domain of x and y be respectively $\mathcal{X}$ and $\mathcal{Y}$. The mutual information can be expanded as

$$\begin{aligned} I(\mathsf{x};\mathsf{y}) &= \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} p_{\mathsf{x},\mathsf{y}}(x,y) \log \frac{p_{\mathsf{x},\mathsf{y}}(x,y)}{p_{\mathsf{x}}(x)p_{\mathsf{y}}(y)} && (1.76)\\ &= H(\mathsf{x}) - H(\mathsf{x}\,|\,\mathsf{y}) && (1.77)\\ &= H(\mathsf{y}) - H(\mathsf{y}\,|\,\mathsf{x}) && (1.78)\\ &= H(\mathsf{x}) + H(\mathsf{y}) - H(\mathsf{x},\mathsf{y}) && (1.79) \end{aligned}$$

if x and y are discrete;

$$\begin{aligned} I(\mathsf{x};\mathsf{y}) &= \int_{\mathcal{X}\times\mathcal{Y}} p_{\mathsf{x},\mathsf{y}}(x,y) \log \frac{p_{\mathsf{x},\mathsf{y}}(x,y)}{p_{\mathsf{x}}(x)p_{\mathsf{y}}(y)} \,\mathrm{d}x\,\mathrm{d}y && (1.80)\\ &= h(\mathsf{x}) - h(\mathsf{x}\,|\,\mathsf{y}) && (1.81)\\ &= h(\mathsf{y}) - h(\mathsf{y}\,|\,\mathsf{x}) && (1.82)\\ &= h(\mathsf{x}) + h(\mathsf{y}) - h(\mathsf{x},\mathsf{y}) && (1.83) \end{aligned}$$

if x and y are continuous; and

$$I(\mathsf{x};\mathsf{y}) = h(\mathsf{y}) - h(\mathsf{y}\,|\,\mathsf{x}) = H(\mathsf{x}) - H(\mathsf{x}\,|\,\mathsf{y}) \tag{1.84}$$

if x is discrete and $\mathsf{y}\,|\,\{\mathsf{x}=x\}$ is continuous for every x. The *conditional mutual information* is given by

$$\begin{aligned} I(\mathsf{x};\mathsf{y}\,|\,\mathsf{z}) &= H(\mathsf{x}\,|\,\mathsf{z}) - H(\mathsf{x}\,|\,\mathsf{y},\mathsf{z}) && (1.85)\\ &= H(\mathsf{y}\,|\,\mathsf{z}) - H(\mathsf{y}\,|\,\mathsf{x},\mathsf{z}) && (1.86)\\ &= H(\mathsf{x}\,|\,\mathsf{z}) + H(\mathsf{y}\,|\,\mathsf{z}) - H(\mathsf{x},\mathsf{y}\,|\,\mathsf{z}) && (1.87) \end{aligned}$$

for discrete random variables, and similarly for continuous random variables. $I(\mathsf{x};\mathsf{y}\,|\,\mathsf{z})$ is non-negative and becomes zero if and only if x and y are independent given z, i.e., $\mathsf{x} \leftrightarrow \mathsf{z} \leftrightarrow \mathsf{y}$ forms a Markov chain.

Lemma 1.2 (Data processing inequality)**.** *If* $\mathsf{x} \leftrightarrow \mathsf{y} \leftrightarrow \mathsf{z}$ *forms a Markov chain, then*

$$I(\mathsf{x};\mathsf{z}) \le I(\mathsf{x};\mathsf{y}). \tag{1.88}$$

The mutual information can be extended straightforwardly to random vectors. In particular, we have chain rule

$$I(\{\mathsf{x}_i\}_{i=1}^n;\mathsf{y}) = \sum_{i=1}^n I(\mathsf{x}_i;\mathsf{y}\,|\,\mathsf{x}_1,\ldots,\mathsf{x}_{i-1}). \tag{1.89}$$

1.A.4 The Duality Approach

Consider a discrete memoryless channel with input $\mathsf{x} \in \mathcal{X}$, output $\mathsf{y} \in \mathcal{Y}$, and the channel law $p_{\mathsf{y}|\mathsf{x}}$. The Shannon capacity of this channel is given by [21, Chapter 7]

$$C = \max_{p_{\mathsf{x}} \in \mathcal{P}(\mathcal{X})} I(\mathsf{x};\mathsf{y}), \tag{1.90}$$

where $\mathcal{P}(\mathcal{X})$ is the set of all possible distributions in $\mathcal{X}$ (satisfying all the underlying, e.g. power, constraints), and the distribution of the output y is $\mathbb{E}_{\mathsf{x}}[p_{\mathsf{y}|\mathsf{x}}]$. A dual expression for the channel capacity is [135]

$$C = \min_{p_{\mathsf{y}} \in \mathcal{P}(\mathcal{Y})} \max_{x \in \mathcal{X}} D(p_{\mathsf{y}|\mathsf{x}=x} \| p_{\mathsf{y}}) \tag{1.91}$$

where $\mathcal{P}(\mathcal{Y})$ is the set of all possible distributions in $\mathcal{Y}$. Any choice of auxiliary output distribution p_{y} leads to an upper bound on the channel capacity:

$$C \leq \max_{x \in \mathcal{X}} D(p_{\mathsf{y}|\mathsf{x}=x} \| p_{\mathsf{y}}). \tag{1.92}$$

This bound can be tightened as [36, Eq.(11)]

$$I(\mathsf{x};\mathsf{y}) \leq \sum_{x \in \mathcal{X}} p_{\mathsf{x}}(x) D(p_{\mathsf{y}|\mathsf{x}=x} \| p_{\mathsf{y}}) \tag{1.93}$$

for any $\mathsf{x} \sim p_{\mathsf{x}} \in \mathcal{P}(\mathcal{X})$ and $\mathsf{y} \sim p_{\mathsf{y}} \in \mathcal{P}(\mathcal{Y})$. Therefore, the maximization over input distribution in (1.90) is replaced by a minimization over output distribution in (1.91). One can bound the capacity and the achievable rate by choosing an auxiliary output distribution p_{y} in (1.92) and (1.93), respectively. This is referred to as the *duality approach* and was extended to continuous alphabets in [36], [38].

We can also interpret the duality approach by looking at the expression $I(\mathsf{x};\mathsf{y}) = h(\mathsf{y}) - h(\mathsf{y}|\mathsf{x})$. While $h(\mathsf{y}|\mathsf{x})$ can be computed easily for many channels, the differential entropy $h(\mathsf{y})$ is normally hard to compute. Assume that p_{y} is the true output distribution, we consider another auxiliary output distribution q_{y} and have that

$$\begin{aligned} h(\mathsf{y}) &= -\mathbb{E}[\log p_{\mathsf{y}}(\mathsf{y})] && (1.94)\\ &= -\mathbb{E}[\log q_{\mathsf{y}}(\mathsf{y})] - D(p_{\mathsf{y}} \| q_{\mathsf{y}}) && (1.95)\\ &\leq -\mathbb{E}[\log q_{\mathsf{y}}(\mathsf{y})], && (1.96) \end{aligned}$$

due to the non-negativity of the KL divergence $D(p_{\mathsf{y}} \| q_{\mathsf{y}})$. Note that the expectation is w.r.t. p_{y}. This imposes an upper bound on the channel capacity.

Using the duality approach to upper bound the capacity, one should first "guess" the capacity-achieving output distribution, then choose an auxiliary output distribution to be close to that while guaranteeing that (1.92), (1.93), and (1.96) are tractable. The next lemma introduces a family of vector-valued distribution that will be used to define auxiliary output distributions in Chapter 2 and Chapter 3.

Lemma 1.3. *Let $\mathbf{y} \in \mathbb{C}^N$ be a random vector with distribution $\mathcal{P}$. Consider another family of distributions $\mathcal{R}$ whose densities are given by*

$$r_{\mathbf{y}}(\boldsymbol{y}) = \frac{\Gamma(N)|\det \boldsymbol{A}|^2}{\pi^N \beta^\alpha \Gamma(\alpha)} \|\boldsymbol{A}\boldsymbol{y}\|^{2(\alpha-N)} \exp\left(-\frac{\|\boldsymbol{A}\boldsymbol{y}\|^2}{\beta}\right), \quad \boldsymbol{y} \in \mathbb{C}^N, \tag{1.97}$$

where $\alpha, \beta > 0$, $\boldsymbol{A}$ is any non-singular deterministic $N \times N$ complex matrix. When $\beta = \mathbb{E}_{\mathcal{P}}[\|\boldsymbol{A}\mathbf{y}\|^2]$ and $\alpha = 1/\log(\beta) = 1/\log(\mathbb{E}_{\mathcal{P}}[\|\boldsymbol{A}\mathbf{y}\|^2])$, denote this distribution as $\mathcal{R}(N, \boldsymbol{A})$. In this case,

$$\mathbb{E}_{\mathcal{P}}[-\log(r_{\mathbf{y}}(\mathbf{y}))] = -\log|\det \boldsymbol{A}|^2 + N\mathbb{E}_{\mathcal{P}}[\log\|\boldsymbol{A}\mathbf{y}\|^2] + O\Big(\log\log\left(\mathbb{E}_{\mathcal{P}}[\|\boldsymbol{A}\mathbf{y}\|^2]\right)\Big). \tag{1.98}$$

Note that if $\boldsymbol{A} = \boldsymbol{I}_N$, (1.98) becomes

$$\mathbb{E}_{\mathcal{P}}[-\log(r_{\mathbf{y}}(\mathbf{y}))] = N\mathbb{E}_{\mathcal{P}}\big[\log\|\mathbf{y}\|^2\big] + O\Big(\log\log\big(\mathbb{E}_{\mathcal{P}}\big[\|\mathbf{y}\|^2\big]\big)\Big).$$

Proof. In this proof, all expectations are implicitly w.r.t. $\mathcal{P}$. A direct calculation from (1.97) yields

$$\begin{aligned}\mathbb{E}[-\log(r_{\mathbf{y}}(\mathbf{y}))] = &-\log|\det\boldsymbol{A}|^2 + (N-\alpha)\mathbb{E}\Big[\log\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\Big] + \frac{\mathbb{E}\big[\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\big]}{\beta} \\ &+ \log\Gamma(\alpha) + \log\beta^{\alpha} + \log\frac{\pi^N}{\Gamma(N)}. \end{aligned} \tag{1.99}$$

When $\beta = \mathbb{E}\big[\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\big]$ and $\alpha = \frac{1}{\log(\beta)} = \frac{1}{\log(\mathbb{E}\left[\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\right])}$, this becomes

$$\begin{aligned}\mathbb{E}[-\log(r_{\mathbf{y}}(\mathbf{y}))] = &-\log|\det\boldsymbol{A}|^2 + N\mathbb{E}\Big[\log\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\Big] - \frac{\mathbb{E}\big[\log\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\big]}{\log(\mathbb{E}\left[\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\right])} \\ &+ \log\Gamma\left(\frac{1}{\log(\mathbb{E}\left[\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\right])}\right) + \log\frac{e\pi^N}{\Gamma(N)} \qquad (1.100) \\ = &-\log|\det\boldsymbol{A}|^2 + N\mathbb{E}\Big[\log\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\Big] + O\Big(\log\log\Big(\mathbb{E}\Big[\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\Big]\Big)\Big), \qquad (1.101)\end{aligned}$$

where the last equality is because $0 < \frac{\mathbb{E}\left[\log\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\right]}{\log(\mathbb{E}\left[\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\right])} < 1$ (due to Jensen's inequality) and

$$\log\Gamma\left(\frac{1}{\log(\mathbb{E}\left[\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\right])}\right) - \log\log\Big(\mathbb{E}\Big[\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\Big]\Big) \to 0 \quad \text{as } \mathbb{E}\Big[\|\boldsymbol{A}\mathbf{y}\|_{\mathrm{F}}^2\Big] \to \infty \tag{1.102}$$

due to

$$\begin{aligned}\lim_{x\to\infty}\log\Gamma\left(\frac{1}{x}\right) - \log x &= \lim_{x\to\infty}\log\left(\frac{1}{x}\Gamma\left(\frac{1}{x}\right)\right) \\ &= \lim_{x\to\infty}\log\left(\Gamma\left(1+\frac{1}{x}\right)\right) \\ &= \log(\Gamma(1)) \\ &= 0. \end{aligned} \tag{1.103}$$

This concludes the proof. □

This family of distribution $\mathcal{R}$ is a special case of the one primarily defined in [36, Section IV-A]. If we take $\mathbf{y}$ as the channel output (of a MIMO channel), as long as $\mathbb{E}\big[\|\boldsymbol{A}\mathbf{y}\|^2\big] \leq \mathrm{SNR}^{c_0}$ for any constant c_0 whose value only depends on the channel statistics, the term $O(\log\log(\mathbb{E}_{\mathcal{P}}[\|\boldsymbol{A}\mathbf{y}\|^2]))$ scales double-logarithmically with SNR. Therefore, in the DoF sense, it is enough to consider only the first two terms in (1.98).

1.B Appendix: Grassmann Manifold

1.B.1 Definition and Invariant Measure

Consider a T-dimensional Euclidean space $\mathbb{K}^T$, where $\mathbb{K}$ can be $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.2 (Grassmann manifold)**.** *The* Grassmann manifold *(so-called Grassmannian) $G(\mathbb{K}^T, M)$ is the set of all M-dimensional vector subspaces of $\mathbb{K}^T$ with $M < T$.*

The Grassmannian is a compact Riemann manifold. For example, when $M = 1$, $G(\mathbb{K}^T, 1)$ is the set of one-dimensional subspaces, i.e. lines, in $\mathbb{K}^T$, and is thus referred to as the *Grassmannian of lines*. An illustration of $G(\mathbb{R}^2, 1)$ is given in Fig. 1.7.

The Grassmann manifold is isomorphic to a quotient of unitary/orthogonal groups:

$$G(\mathbb{K}^T, M) \cong U(T)/(U(M) \times U(T-M)), \tag{1.104}$$

where the division slash denotes a quotient space and the unitary/orthogonal group $U(n)$ is the set of all $n \times n$ unitary/orthogonal matrices, i.e.,

$$U(n) := \{\boldsymbol{U} \in \mathbb{K}^{n\times n} \ : \ \boldsymbol{U}^{\mathsf{H}}\boldsymbol{U} = \boldsymbol{I}_n\}. \tag{1.105}$$

This equivalence is interpreted as follows. Each element of $U(T)$ is an orthonormal basis of $\mathbb{K}^T$ and can be split into M vectors spanning an M-dimensional subspace and the remaining $T - M$ vectors spanning the orthogonal complement of that subspace. The M-dimensional subspace and its orthogonal complement are invariant to unitary rotation of the set of M vectors and the remaining $T-M$ vectors, respectively. Therefore, to uniquely represent an M-dimensional subspace, which is an element in $G(\mathbb{K}^T, M)$, we need to divide $U(T)$ by the space of these rotations, which are $U(M)$ and $U(T-M)$, respectively [136]. The Grassmann manifold can also be represented as a space of equivalence classes in the Stiefel manifold, which is the space of $T \times M$ truncated unitary matrices in $\mathbb{K}^T$, i.e.,

$$S(\mathbb{K}^T, M) := \{\boldsymbol{Q} \in \mathbb{K}^{T\times M} \ : \ \boldsymbol{Q}^{\mathsf{H}}\boldsymbol{Q} = \boldsymbol{I}_M\}. \tag{1.106}$$

Defining the equivalence between two elements $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ of $S(\mathbb{K}^T, M)$ as $\boldsymbol{Q}_1 = \boldsymbol{Q}_2\boldsymbol{U}$ for some $\boldsymbol{U} \in U(M)$, we have that each element in the Grassmann manifold $G(\mathbb{K}^T, M)$ is an equivalence class in the Stiefel manifold $S(\mathbb{K}^T, M)$.

From the quotient space representation (1.104), the real dimension of the Grassmannian $G(\mathbb{K}^T, M)$ can be calculated as

$$\begin{aligned} \dim(G(\mathbb{K}^T, M)) &= \dim(U(T)) - \dim(U(M)) - \dim(U(T-M)) && (1.107)\\ &= \beta M(T-M), && (1.108) \end{aligned}$$

where $\beta = 1$ if $\mathbb{K} = \mathbb{R}$ and $\beta = 2$ if $\mathbb{K} = \mathbb{C}$. A uniform measure in $G(\mathbb{K}^T, M)$ can be induced from the Haar measure on the unitary/orthogonal group as follows. Let $\mu_{\mathrm{H}}(\cdot)$ be the unit Haar measure on the unitary/orthogonal group $U(T)$ and fix a point $\mathcal{V}$ in $G(\mathbb{K}^T, M)$. For any set $\mathcal{M} \subseteq G(\mathbb{K}^T, M)$, the measure $\mu(\mathcal{M})$ on $G(\mathbb{K}^T, M)$ is defined as

$$\mu(\mathcal{M}) = \mu_{\mathrm{H}}\{\boldsymbol{U} \in U(T) \ : \ \boldsymbol{U}\mathcal{V} \in \mathcal{M}\}. \tag{1.109}$$

This measure is invariant under actions from the group $U(T)$, i.e., $\mu(\boldsymbol{U}\mathcal{M}) = \mu(\mathcal{M})$ for any $\boldsymbol{U} \in U(T)$. With this measure, the volume of a Grassmann manifold is given by

$$|G(\mathbb{K}^T, M)| = \frac{\prod_{i=T-M+1}^{T} \frac{2\pi^i}{(i-1)!}}{\prod_{i=1}^{M} \frac{2\pi^i}{(i-1)!}}. \tag{1.110}$$

An isotropically distributed matrix in $\mathbb{K}^{T\times M}$ has subspace uniformly distributed on the Grassmannian $G(\mathbb{K}^T, M)$ w.r.t. this invariant measure.

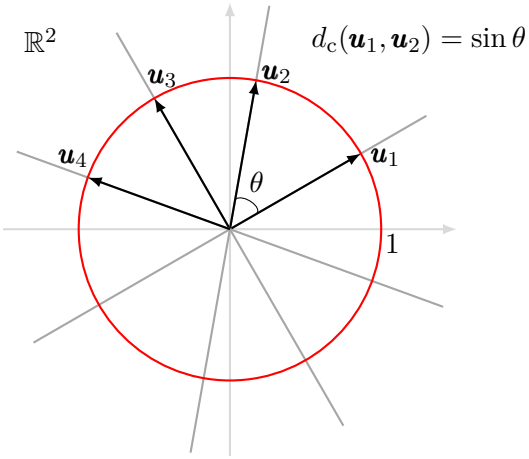


FIGURE 1.7: An illustration of four elements of the Grassmannian of lines $G(\mathbb{R}^2, 1)$ in the plane $\mathbb{R}^2$. Each element i is a set $\{\lambda \boldsymbol{u}_i \ : \lambda \in \mathbb{R}, \boldsymbol{u}_i \in \mathbb{R}^2, \|\boldsymbol{u}_i\| = 1\}$, which is a line in $\mathbb{R}^2$. It is represented by the unit vector $\boldsymbol{u}_i$, i.e., the intersection of the line and the unit circle. Note that $\boldsymbol{u}_i$ and $-\boldsymbol{u}_i$ represent the same Grassmannian point. The principal angle between two Grassmannian points represented by $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ is the acute angle θ between two corresponding lines. The chordal distance between these two points is simply the sine of that angle, i.e., $d_c(\boldsymbol{u}_1, \boldsymbol{u}_2) = \sin\theta$.

1.B.2 Principal Angles and Metrics

The Grassmannian is a metric space with the metric defined through the *principal angles* between the points on the manifold. The principal angles between two M-dimensional subspaces $\mathcal{U}$ and $\mathcal{V}$ of $\mathbb{K}^T$, which are two points on $G(\mathbb{K}^T, M)$, are defined in a sequential manner as follows. The first angle is the smallest angle between two unit vectors in $\mathcal{U}$ and $\mathcal{V}$, respectively, i.e., $\theta_1 = \arccos |\boldsymbol{u}_1^{\mathsf{H}} \boldsymbol{v}_1|$ with[11]

$$(\boldsymbol{u}_1, \boldsymbol{v}_1) = \arg \max_{\substack{\boldsymbol{u}\in\mathcal{U}, \|\boldsymbol{u}\|=1 \\ \boldsymbol{v}\in\mathcal{V}, \|\boldsymbol{v}\|=1}} |\boldsymbol{u}^{\mathsf{H}} \boldsymbol{v}|. \tag{1.111}$$

The next principal angle is the smallest angle between two unit vectors in $\mathcal{U}$ and $\mathcal{V}$ and orthogonal to $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$, respectively, i.e., $\theta_2 = \arccos |\boldsymbol{u}_2^{\mathsf{H}} \boldsymbol{v}_2|$ with

$$(\boldsymbol{u}_2, \boldsymbol{v}_2) = \arg \max_{\substack{\boldsymbol{u}\in\mathcal{U}, \|\boldsymbol{u}\|=1, \boldsymbol{u}^{\mathsf{H}}\boldsymbol{u}_1=0 \\ \boldsymbol{v}\in\mathcal{V}, \|\boldsymbol{v}\|=1, \boldsymbol{v}^{\mathsf{H}}\boldsymbol{v}_1=0}} |\boldsymbol{u}^{\mathsf{H}} \boldsymbol{v}|. \tag{1.112}$$

Inductively, the i-th principal angle is the smallest angle between two unit vectors in $\mathcal{U}$ and $\mathcal{V}$ orthogonal to $\boldsymbol{u}_j$ and $\boldsymbol{v}_j$ for all $1 \leq j < i$, respectively, i.e., $\theta_i = \arccos |\boldsymbol{u}_i^{\mathsf{H}} \boldsymbol{v}_i|$ with

$$(\boldsymbol{u}_i, \boldsymbol{v}_i) = \arg \max_{\substack{\boldsymbol{u}\in\mathcal{U}, \|\boldsymbol{u}\|=1, \boldsymbol{u}^{\mathsf{H}}\boldsymbol{u}_j=0, \forall 1\leq j<i \\ \boldsymbol{v}\in\mathcal{V}, \|\boldsymbol{v}\|=1, \boldsymbol{v}^{\mathsf{H}}\boldsymbol{v}_j=0, \forall 1\leq j<i}} |\boldsymbol{u}^{\mathsf{H}} \boldsymbol{v}|, \tag{1.113}$$

for $i \in [M]$. The principal angles are between 0 and $\frac{\pi}{2}$. For example, the principal angle between two points in $G(\mathbb{R}^2, 1)$ is simply the acute angle between two corresponding lines in $\mathbb{R}^2$, as illustrated in Fig. 1.7.

Let

$$\boldsymbol{\theta} := [\theta_1 \ \theta_2 \ \ldots \ \theta_M], \quad \text{and} \quad \sin\boldsymbol{\theta} := [\sin\theta_1 \ \sin\theta_2 \ \ldots \ \sin\theta_M]. \tag{1.114}$$

There are several distance metrics defined between $\mathcal{U}$ and $\mathcal{V}$, such as the geodesic distance $d_g(\mathcal{U}, \mathcal{V}) := \|\boldsymbol{\theta}\|$, the chordal distance $d_c(\mathcal{U}, \mathcal{V}) := \|\sin\boldsymbol{\theta}\|$, the Fubini-Study distance

[11] The conjugate transpose is implicitly replaced by the transpose for the real Grassmannian.

$d_{\mathrm{FS}}(\mathcal{U},\mathcal{V}) := \arccos\left(\prod_{i=1}^{M}\cos\theta_i\right)$, and the chordal Frobenius norm $d_{\mathrm{cF}}(\mathcal{U},\mathcal{V}) := \left\|2\sin\frac{\boldsymbol{\theta}}{2}\right\|$ (see, e.g., [57], [58]). Among them, the chordal distance is widely used and has some advantages, such as it allows an isometric embedding of the Grassmannian in a sphere [60]. Its values range from 0 to $\sqrt{M}$. For example, the chordal distance between two points in $G(\mathbb{R}^2,1)$ is the sine of the acute angle between two corresponding lines in $\mathbb{R}^2$, as illustrated in Fig. 1.7.

The principal angles can be computed from the matrix representatives of the subspaces. Let $\boldsymbol{U}$ and $\boldsymbol{V}$ be $T\times M$ matrices with orthonormal columns spanning the subspaces $\mathcal{U}$ and $\mathcal{V}$, respectively. Then the singular values of the product $\boldsymbol{U}^{\mathsf{H}}\boldsymbol{V}$ are the cosines of the principal angles between $\mathcal{U}$ and $\mathcal{V}$. Therefore, these principal angles can be found by a SVD of $\boldsymbol{U}^{\mathsf{H}}\boldsymbol{V}$. The distance between $\mathcal{U}$ and $\mathcal{V}$ can also be expressed in terms of $\boldsymbol{U}$ and $\boldsymbol{V}$. For example, the chordal distance is written as

$$d_{\mathrm{c}}(\mathcal{U},\mathcal{V}) = \|\boldsymbol{U}^{\mathsf{H}}\boldsymbol{U} - \boldsymbol{V}^{\mathsf{H}}\boldsymbol{V}\|_{\mathrm{F}} = \sqrt{M - \mathrm{tr}(\boldsymbol{U}^{\mathsf{H}}\boldsymbol{V}\boldsymbol{V}^{\mathsf{H}}\boldsymbol{U})}. \tag{1.115}$$

For this reason, we also denote the chordal distance as $d_{\mathrm{c}}(\boldsymbol{U},\boldsymbol{V})$ for convenience, although its value does not depend on the choice of matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ that represent the two subspaces. From now on, we consider the chordal distance and omit the subscript "c" whenever confusion is not likely. Letting a matrix $\boldsymbol{U}\in\mathbb{K}^{T\times M}$ represent its column span, we write with a slight abuse of notation that $\boldsymbol{U}\in G(\mathbb{K}^T,M)$.

1.B.3 Sphere Packing in the Grassmannian

Following [62, Corollary 1], the next lemma gives the volume of a metric ball in the Grassmannian.

Lemma 1.4 (The volume of a metric ball in $G(\mathbb{K}^T,M)$)**.** *With the invariant measure $\mu(\cdot)$, the volume of a metric ball $\mathcal{B}(\delta)$ of radius δ (defined with chordal distance) in the Grassmann manifold $G(\mathbb{K}^T,M)$ is given by*

$$\mu(\mathcal{B}(\delta)) = c_{T,M,\beta}\,\delta^{\beta M(T-M)}, \tag{1.116}$$

where

$$c_{T,M,\beta} := \frac{1}{\Gamma\left(\frac{\beta}{2}M(T-M)+1\right)}\prod_{i=1}^{\min\{M,T-M\}}\frac{\Gamma\left(\frac{\beta}{2}(T-i+1)\right)}{\Gamma\left(\frac{\beta}{2}(\min\{M,T-M\}-i+1)\right)}, \tag{1.117}$$

where $\beta=1$ if $\mathbb{K}=\mathbb{R}$ and $\beta=2$ if $\mathbb{K}=\mathbb{C}$.

A packing $\mathcal{C}$ in $G(\mathbb{K}^T,M)$ is a finite-size discrete subset of $G(\mathbb{K}^T,M)$. For example, Fig. 1.7 depicts a packing of four points in $G(\mathbb{R}^2,1)$. The sphere packing problem in the Grassmannian has a variety of applications in wireless communications. It provides a geometric interpretation of the capacity of a noncoherent MIMO block-fading channel [49, Section III-C]. A Grassmannian packing can be used as a constellation for such channel [53], [64] (see Section 1.2.2.c of this thesis). It can also be used as a CSI quantization or precoding codebook in limited-feedback communication systems [137]. The set of the intersections of the Grassmannian constellation symbols and the unit sphere is also called an antipodal spherical code, which has applications in designing measurement matrix for compressive sensing [138].

Let us focus on the complex Grassmannian $G(\mathbb{C}^T,M)$. A common sphere packing criterion is the maximization of minimum pairwise distance between the elements/symbols

for a given packing size, i.e.,

$$\mathcal{C} = \arg \max_{\{\boldsymbol{C}_1,\ldots,\boldsymbol{C}_{|\mathcal{C}|}\}\subset G(\mathbb{C}^T,M)} \min_{1\le i<j\le|\mathcal{C}|} d(\boldsymbol{C}_i,\boldsymbol{C}_j). \tag{1.118}$$

The sphere packing bounds characterize the fundamental limits regarding the size of the packing and the minimum distance. Using upper bounds on spherical codes and the embedding of the Grassmannian in a sphere, Conway *et al.* derived the well-known Rankin bounds for real Grassmannian packing [60]. These bounds naturally generalize to the complex case as

$$\delta(\mathcal{C}) \le \begin{cases} \sqrt{\frac{M(T-M)}{T}\frac{|\mathcal{C}|}{|\mathcal{C}|-1}}, & \text{if } |\mathcal{C}| \le \frac{T(T+1)}{2}, \\ \sqrt{\frac{M(T-M)}{T}}, & \text{if } |\mathcal{C}| > \frac{T(T+1)}{2}. \end{cases} \tag{1.119}$$

There exist sequences of small-size packings that meet the Rankin bounds [139]. However, the Rankin bounds quickly become loose as the packing size grows. Another approach relates the packing size to the volume of the metric ball of radius equal to the minimum distance δ. Specifically, for any positive integer K such that $K\mu(\mathcal{B}(\delta)) \le 1$, there exists a packing $\mathcal{C}$ of size $K+1$ with minimum distance δ. This is captured by the Gilbert-Varshamov lower bound [62, Eq.(2)]: there exists a packing with minimum distance δ and size

$$|\mathcal{C}| > \frac{1}{\mu(\mathcal{B}(\delta))}. \tag{1.120}$$

On the other hand, the total volume of the balls of radius $\delta/2$ centered at each symbol is upper bounded by the volume of the Grassmannian. This implies the Hamming upper bound [62, Eq.(3)] as

$$|\mathcal{C}| \le \frac{1}{\mu(\mathcal{B}(\delta/2))}. \tag{1.121}$$

Substituting the volume formula in Lemma 1.4 into the Gilbert-Varshamov lower bound and the Hamming upper bound, the next lemma follows readily.

Lemma 1.5 (Grassmannian sphere packing bounds)**.** *The minimum distance $\delta(\mathcal{C}_{\text{opt}})$ of an optimal packing $\mathcal{C}_{\text{opt}}$ of cardinality $|\mathcal{C}|$ on the complex Grassmannian $G(\mathbb{C}^T, M)$ is bounded by*

$$\min\left\{\sqrt{M}, 2(|\mathcal{C}|c_{T,M})^{-\frac{1}{2M(T-M)}}\right\} \ge \delta(\mathcal{C}_{\text{opt}}) \ge (|\mathcal{C}|c_{T,M})^{-\frac{1}{2M(T-M)}}, \tag{1.122}$$

where $c_{T,M} := \frac{1}{(M(T-M))!}\prod_{i=1}^{\min\{M,T-M\}}\frac{(T-i)!}{(\min\{M,T-M\}-i)!}$.

The sphere packing problem can be solved numerically by, e.g., Newton's or (conjugate) gradient descent as presented in the next subsection. An alternative projection method was proposed in [136], but works well for sparse packings (small $|\mathcal{C}|$, large T) only.

1.B.4 Manifold-Constrained Optimization

Consider a minimization of a function $f(\boldsymbol{C})$ defined on the Grassmann manifold

$$\min_{\boldsymbol{C}\in G(\mathbb{C}^T,M)} f(\boldsymbol{C}). \tag{1.123}$$

To apply standard optimization methods, such as Newton's and the (conjugate) gradient descent, one needs to derive the Riemannian (conjugate) gradient, the Hessian of the function $f(\boldsymbol{C})$ and the motion of a point on the Grassmannian.

The motion on the Grassmann manifold can be obtained using the differential-geometric methods [57]. Consider a moving point on the Grassmannian, which is located at $\boldsymbol{C}(0) = \boldsymbol{C} \in G(\mathbb{C}^T, M)$ at time $t = 0$. Let this point move along a geodesic in the direction $\boldsymbol{H} = \boldsymbol{C}_\perp \boldsymbol{B}$ where $\boldsymbol{C}_\perp \in \mathbb{C}^{T\times(T-M)}$ has orthonormal columns spanning the orthogonal complement of $\mathrm{Span}(\boldsymbol{C})$ in $\mathbb{C}^T$ and $\boldsymbol{B}$ is an arbitrary $(T-M)\times M$ matrix. Then, its location at an arbitrary time instant t is given by

$$\boldsymbol{C}(t) = [\boldsymbol{C}\, \boldsymbol{C}_\perp] \exp\left(t \begin{bmatrix} \boldsymbol{0} & -\boldsymbol{B}^{\mathsf{H}} \\ \boldsymbol{B} & \boldsymbol{0} \end{bmatrix} \right) \begin{bmatrix} \boldsymbol{I}_M \\ \boldsymbol{0} \end{bmatrix}. \tag{1.124}$$

According to [57, Section 2.5.3] and [140, Section 3.6], the Riemannian gradient of f at $\boldsymbol{C}$ is defined as the tangent vector $\nabla_R f(\boldsymbol{C})$ such that $\mathrm{tr}(\nabla_E^{\mathsf{H}} f(\boldsymbol{C})\Delta) = \mathrm{tr}(\nabla_R^{\mathsf{T}} f(\boldsymbol{C})\Delta)$ for all tangent vectors Δ at $\boldsymbol{C}$, where $\nabla_E f(\boldsymbol{C})$ is the Euclidean gradient of $f(\boldsymbol{C})$, i.e.,

$$[\nabla_E f(\boldsymbol{C})]_{i,j} = \frac{\partial f}{\partial [\boldsymbol{C}]_{i,j}}. \tag{1.125}$$

It follows that

$$\nabla_R f(\boldsymbol{C}) = (\boldsymbol{I}_T - \boldsymbol{C}\boldsymbol{C}^{\mathsf{H}}) \nabla_E f(\boldsymbol{C}). \tag{1.126}$$

The Riemannian Hessian of $f(\boldsymbol{C})$ is defined as the quadratic form $\mathrm{Hess}_R f(\Delta, \Delta) = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\big|_{t=0} f(\boldsymbol{C}(t))$ where $\boldsymbol{C}(t)$ is a geodesic with tangent Δ at $\boldsymbol{C}(0)$. It is computed as

$$\mathrm{Hess}_R f(\Delta_1, \Delta_2) = \sum_{i,j,k,l} \left[\mathrm{Hess}_E f(\Delta_1, \Delta_2)\right]_{ij,kl} [\Delta_1]_{i,j} [\Delta_2]_{k,l} - \mathrm{tr}(\Delta_1^{\mathsf{H}} \Delta_2 \boldsymbol{C}^{\mathsf{H}} \nabla_E^{\mathsf{H}} f(\boldsymbol{C})), \tag{1.127}$$

where $\mathrm{Hess}_E f(\Delta_1, \Delta_2)$ is the Euclidean Hessian of $f(\boldsymbol{C})$, i.e.,

$$\left[\mathrm{Hess}_E f(\Delta_1, \Delta_2)\right]_{ij,kl} = \frac{\partial^2 f}{\partial [\boldsymbol{Y}]_{i,j} \partial [\boldsymbol{Y}]_{k,l}}. \tag{1.128}$$

Finally, Newton's and the (conjugate) gradient descent methods to solve the minimization (1.123) on the Grassmannian are summarized in [57], [140].

When the optimization is over a set of points on the Grassmannian, such as the sphere packing problem (1.118), one can consider the set of optimization points as a single element of a manifold in larger dimension. For example, one can construct the block diagonal matrix $\bar{\boldsymbol{C}}$ with $\boldsymbol{C}_1, \ldots, \boldsymbol{C}_{|\mathcal{C}|}$ as diagonal blocks, which is a single point in a sub-manifold of $G(\mathbb{C}^{T|\mathcal{C}|}, M|\mathcal{C}|)$. By examining the tangent space of this sub-manifold, the techniques for single-point optimization on the Grassmann manifold can be applied. This was done in [64] to solve (1.118). Another difficulty in solving (1.118) is that the objective function $\min_{1\le i<j\le|\mathcal{C}|} d(\boldsymbol{C}_i, \boldsymbol{C}_j) = -\max_{1\le i<j\le|\mathcal{C}|} -d(\boldsymbol{C}_i, \boldsymbol{C}_j)$ is not smooth. We can approximate it with a smooth function using the well-known approximation

$$\max_i x_i \approx \epsilon \ln \sum_i \exp(x_i/\epsilon) \tag{1.129}$$

for a small "diffusion constant" ϵ. The smaller ϵ is, the more accurate this approximation is. However, if ϵ is too small, the optimization becomes slow due to the small step size and numerical problems, such as floating point overflow errors, may occur.

In the following, we investigate a particular example of sphere packing in the Grassmannian of lines $G(\mathbb{C}^T, 1)$.

Example 1.1 (Sphere packing in the Grassmannian of lines)**.** *Consider the instance of the sphere packing problem* (1.118) *with* $M = 1$*:*

$$\max_{\mathcal{C}=\{\boldsymbol{c}_i\}_{i=1}^{|\mathcal{C}|}\subset G(\mathbb{C}^T,1)} \min_{1\le i<j\le|\mathcal{C}|} \sqrt{1-|\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_j|^2}. \tag{1.130}$$

This optimization is equivalently expressed as

$$\min_{\mathcal{C}=\{\boldsymbol{c}_i\}_{i=1}^{|\mathcal{C}|}\subset G(\mathbb{C}^T,1)} \max_{1\le i<j\le|\mathcal{C}|} |\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_j|, \tag{1.131}$$

For smoothness, we approximate the objective function using (1.129) *to obtain*

$$\min_{\mathcal{C}=\{\boldsymbol{c}_i\}_{i=1}^{|\mathcal{C}|}\subset G(\mathbb{C}^T,1)} \epsilon \ln \sum_{1\le i<j\le|\mathcal{C}|} \exp\left(\frac{|\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_j|}{\epsilon}\right). \tag{1.132}$$

This smooth optimization is, however, jointly over multiple points on the Grassmannian of lines. To tackle this, we construct the matrix $\boldsymbol{C} := [\boldsymbol{c}_1 \dots \boldsymbol{c}_{|\mathcal{C}|}]$*, then* $\boldsymbol{C}$ *belongs to the oblique manifold* $\mathcal{OB}(T, |\mathcal{C}|)$ *defined as*

$$\mathcal{OB}(n,m) := \Big\{\boldsymbol{X} = [\boldsymbol{x}_1 \dots \boldsymbol{x}_m] \in \mathbb{C}^{n\times m} \colon \|\boldsymbol{x}_1\| = \dots = \|\boldsymbol{x}_m\| = 1\Big\}. \tag{1.133}$$

The oblique manifold $\mathcal{OB}(n,m)$ *can be seen as an embedded Riemannian manifold of* $\mathbb{C}^{n\times m}$ *endowed with the usual inner product, or as the product manifold of* m *unit spheres in* $\mathbb{C}^n$ *[140, Section 3.4.1]. Then, the optimization problem* (1.132) *can be reformulated as a single-variable optimization on this oblique manifold as*

$$\min_{\boldsymbol{C}\in\mathcal{O}(T,|\mathcal{C}|)} \underbrace{\epsilon \ln \sum_{1\le i<j\le|\mathcal{C}|} \exp\left(\frac{|\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_j|}{\epsilon}\right)}_{=:g(\boldsymbol{C})}. \tag{1.134}$$

The derivative of $g(\boldsymbol{C})$ *w.r.t. a symbol* $\boldsymbol{c}_m$ *is*

$$\frac{\partial g(\boldsymbol{C})}{\partial \boldsymbol{c}_m} = \epsilon\left(\sum_{1\le i<j\le|\mathcal{C}|} \exp\left(\frac{|\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_j|}{\epsilon}\right)\right)^{-1}\left[\sum_{i<m} \frac{\partial \exp(|\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_m|/\epsilon)}{\partial \boldsymbol{c}_m} + \sum_{j>m} \frac{\partial \exp(|\boldsymbol{c}_m^{\mathsf{H}}\boldsymbol{c}_j|/\epsilon)}{\partial \boldsymbol{c}_m}\right], \tag{1.135}$$

where

$$\frac{\partial \exp(|\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_m|/\epsilon)}{\partial \boldsymbol{c}_m} = \frac{\partial \exp(|\boldsymbol{c}_m^{\mathsf{H}}\boldsymbol{c}_i|/\epsilon)}{\partial \boldsymbol{c}_m} \tag{1.136}$$

$$= \frac{1}{\epsilon}\exp\left(\frac{|\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_m|}{\epsilon}\right)\frac{\partial|\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_m|}{\partial \boldsymbol{c}_m} \tag{1.137}$$

$$= \frac{1}{\epsilon}\exp\left(\frac{|\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_m|}{\epsilon}\right)\frac{\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_m}{|\boldsymbol{c}_i^{\mathsf{H}}\boldsymbol{c}_m|}\boldsymbol{c}_i. \tag{1.138}$$

Consequently, the Euclidean gradient of $g(\boldsymbol{C})$ *w.r.t.* $\boldsymbol{C}$ *is given by*

$$\nabla_E^{\mathsf{T}} g(\boldsymbol{C}) = \left[\frac{\partial g(\boldsymbol{C})}{\partial \boldsymbol{c}_1} \; \frac{\partial g(\boldsymbol{C})}{\partial \boldsymbol{c}_2} \; \cdots \; \frac{\partial g(\boldsymbol{C})}{\partial \boldsymbol{c}_{|\mathcal{C}|}} \right]^{\mathsf{T}} \tag{1.139}$$

$$= \left(\sum_{1 \le i < j \le |\mathcal{C}|} \exp\left(\frac{|\boldsymbol{c}_i^{\mathsf{H}} \boldsymbol{c}_j|}{\epsilon} \right) \right)^{-1} \times \begin{bmatrix} 0 & \exp(|\boldsymbol{c}_2^{\mathsf{H}} \boldsymbol{c}_1|/\epsilon) \frac{\boldsymbol{c}_2^{\mathsf{H}} \boldsymbol{c}_1}{|\boldsymbol{c}_2^{\mathsf{H}} \boldsymbol{c}_1|} & \cdots & \exp(|\boldsymbol{c}_{|\mathcal{C}|}^{\mathsf{H}} \boldsymbol{c}_1|/\epsilon) \frac{\boldsymbol{c}_{|\mathcal{C}|}^{\mathsf{H}} \boldsymbol{c}_1}{|\boldsymbol{c}_{|\mathcal{C}|}^{\mathsf{H}} \boldsymbol{c}_1|} \\ \exp(|\boldsymbol{c}_1^{\mathsf{H}} \boldsymbol{c}_2|/\epsilon) \frac{\boldsymbol{c}_1^{\mathsf{H}} \boldsymbol{c}_2}{|\boldsymbol{c}_1^{\mathsf{H}} \boldsymbol{c}_2|} & 0 & \cdots & \exp(|\boldsymbol{c}_{|\mathcal{C}|}^{\mathsf{H}} \boldsymbol{c}_2|/\epsilon) \frac{\boldsymbol{c}_{|\mathcal{C}|}^{\mathsf{H}} \boldsymbol{c}_2}{|\boldsymbol{c}_{|\mathcal{C}|}^{\mathsf{H}} \boldsymbol{c}_2|} \\ \vdots & \vdots & \ddots & \vdots \\ \exp(|\boldsymbol{c}_1^{\mathsf{H}} \boldsymbol{c}_{|\mathcal{C}|}|/\epsilon) \frac{\boldsymbol{c}_1^{\mathsf{H}} \boldsymbol{c}_{|\mathcal{C}|}}{|\boldsymbol{c}_1^{\mathsf{H}} \boldsymbol{c}_{|\mathcal{C}|}|} & \exp(|\boldsymbol{c}_2^{\mathsf{H}} \boldsymbol{c}_{|\mathcal{C}|}|/\epsilon) \frac{\boldsymbol{c}_2^{\mathsf{H}} \boldsymbol{c}_{|\mathcal{C}|}}{|\boldsymbol{c}_2^{\mathsf{H}} \boldsymbol{c}_{|\mathcal{C}|}|} & \cdots & 0 \end{bmatrix} \boldsymbol{C}^{\mathsf{H}}. \tag{1.140}$$

Then, the Riemannian gradient of $g(\boldsymbol{C})$ *is computed based on* (1.126).

A useful tool for optimization on manifolds is the Manopt toolbox [141] in Matlab. It provides a large library of manifolds and ready-to-use Riemannian optimization algorithms. In this thesis, we resort to this toolbox to solve numerically the manifold-constrained optimizations, which we encounter in Chapter 5, Chapter 6, and Chapter 7. Note that in manifold-constrained optimization, the optimization space is nonlinear, and the objective functions that we consider in this thesis are in general nonconvex. Thus, most descent algorithms only guarantee to return an (approximate) critical point. In order to ensure that this point is a local minimum and not a saddle point, the search direction needs to be carefully constructed. Several rules to construct the new search direction based on a linear combination of the previous search direction and the new (preconditioned) gradient are provided for the Euclidean in [142]. The Manopt toolbox adapts these rules to the Riemannian space. If no descent direction is found, one can restart, i.e., switch to the negative gradient. This is equivalent to resetting the direction to a steepest descent step, which discards the past information. The Manopt toolbox implements the Powell's restart strategy [143].

For further details concerning the Grassmann manifold, sphere packing, and manifold-constrained optimization, refer to, e.g., [57], [60], [140], [144].

Part I

Fundamental Limits of Noncoherent Communications

Chapter 2

The MIMO Point-to-Point Channel

The high-SNR capacity of noncoherent MIMO P2P channel has been derived for the case of IID Rayleigh block fading by exploiting the Gaussianity of the channel matrix. This implies the optimal DoF. Nevertheless, as far as the DoF is concerned, it is apparent that the result holds for a wider class of fading channels. In this chapter, we show that the optimal DoF for the IID Rayleigh block fading channel is also the optimal DoF for a more general class of generic *block fading channels, in which the random channel matrix has finite power and finite differential entropy. In doing so, we introduce a novel converse proof technique based on a genie-aided bound and the duality approach.*

2.1 Overview

As presented in Section 1.2.2.b, for an $M \times N$ MIMO channel in IID Rayleigh block fading with coherence interval $T > 1$, the optimal DoF is given by [49], [50]

$$\text{DoF} = M^*\left(1 - \frac{M^*}{T}\right), \quad \text{with} \quad M^* := \min\{M, N, \lfloor T/2 \rfloor\}. \tag{2.1}$$

The converse of this result was based on the IID Rayleigh fading assumption, using either a direct approximation at high SNR [49] or a duality upper bound with a carefully chosen auxiliary output distribution [50]. On the other hand, the optimal DoF for the case $T = 1$ (fast fading) is known to be zero for a much wider class of fading model, namely, the generic fading in which the channel matrix has finite differential entropy and finite second moment [36].

In this chapter, we generalize the DoF result of [49], [50] to the generic fading model as in [36]. Specifically, we prove that the DoF given in (2.1) is also the optimal DoF under generic block fading. The main technical contribution of this paper lies in the converse proof. Leveraging the duality upper bound [36], we carefully choose an auxiliary output distribution (inspired by a pilot-based scheme) with which we derive a tight DoF upper bound.[1]

[1] Another useful tool for the converse bounds in the literature is the *escape-to-infinity* property [36], which allows one to assume without loss of generality (w.l.o.g.) that the high-SNR capacity-achieving input distribution has no mass in a disk around the origin, whose radius can be made arbitrarily large. This property is necessary to derive the constant term after the logarithmic term in the capacity high-SNR expansion. In our work, since we focus on the pre-log characterization, we do not rely on this property.

The remainder of this chapter is organized as follows. The system model and preliminaries are presented in Section 2.2. In Section 2.3, we provide the main result and the achievability proof. Then, we present the converse proof for the SIMO and MIMO cases in Section 2.4 and Section 2.5, respectively. Some concluding remarks are given in Section 2.6. Some detailed proofs are deferred to the appendices.

2.2 System Model and Preliminaries

We consider an $M \times N$ MIMO P2P channel as in Section 1.2.2. The channel between the transmitter and the receiver is flat and block fading with coherence interval T channel uses. The distribution of the channel matrix $\mathsf{H} \in \mathbb{C}^{N\times M}$ is assumed to be known, but its realizations are *unknown* to both the transmitter and the receiver. During a coherence block b, the received signal is

$$\mathsf{Y}[b] = \mathsf{H}[b]\mathsf{X}[b] + \mathsf{Z}[b], \quad b = 1, 2, \ldots, \tag{2.2}$$

where $\mathsf{Z}[b] \in \mathbb{C}^{N\times T}$ is the AWGN with IID $\mathcal{N}_{\mathbb{C}}(0,1)$ entries and $\mathsf{X}[b]$ is the transmitted signal satisfying the power constraint

$$\frac{1}{\nu}\sum_{b=1}^{\nu} \|\mathsf{X}[b]\|_{\mathrm{F}}^2 \le PT, \tag{2.3}$$

where ν is the number of blocks spanned by a channel codeword. P is identified with the SNR of the channel. Hereafter, we omit the block index b whenever confusion is not likely.

If the rate $R(P)$ (in bpcu) is achievable at SNR P, $\forall P > 0$, i.e., inferior to the channel capacity, we say that d is an achievable DoF with

$$d := \lim_{P\to\infty} \frac{R(P)}{\log_2(P)}. \tag{2.4}$$

The optimal DoF is the supremum of the achievable DoF over all possible input distribution satisfying the power constraint. We assume that the channel matrix H is drawn from a generic distribution satisfying the following conditions:

$$h(\mathsf{H}) > -\infty, \quad \mathbb{E}\Big[\|\mathsf{H}\|_{\mathrm{F}}^2\Big] < \infty. \tag{2.5}$$

That is, the channel matrix has finite different entropy and finite second moment.

2.2.1 Mathematical Preliminaries

The following results are useful for our main analysis.

Lemma 2.1. *Let $\boldsymbol{A} \in \mathbb{C}^{m\times t}$ have full column rank, $\mathsf{W} \in \mathbb{C}^{n\times m}$ be a random matrix such that $h(\mathsf{W}) > -\infty$ and $\mathbb{E}[\|\mathsf{W}\|_{\mathrm{F}}^2] < \infty$, then we have*

$$h(\mathsf{W}\boldsymbol{A}) = n \log\det(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}) + c_0 \tag{2.6}$$

where c_0 is bounded by some constant that only depends on the statistics of W.

Proof. The proof is provided in Appendix 2.A.1. □

Lemma 2.2. *Let $\mathsf{u} \ge 0$ be some random variable such that $\mathbb{E}[\mathsf{u}] < \infty$ and $h(\mathsf{u}/\mathbb{E}[\mathsf{u}]) > -\infty$. Then, for any $0 < \alpha < 1$ and $\beta > 0$,*

$$\mathbb{E}[\log(\beta + \mathsf{u})] \ge \alpha \log(\beta + \mathbb{E}[\mathsf{u}]) + c_0 \tag{2.7}$$

where $c_0 > -\infty$ is some constant that only depends on α, β, and $h(\mathsf{u}/\mathbb{E}[\mathsf{u}])$.

Proof. The proof is provided in Appendix 2.A.2. □

From the above result and the upper bound $\frac{\mathbb{E}[\log(\beta+\mathsf{u})]}{\log(\beta+\mathbb{E}[\mathsf{u}])} \leq 1$ (from Jensen's inequality), we observe that when $\mathbb{E}[\mathsf{u}] \to \infty$, $\frac{\mathbb{E}[\log(\beta+\mathsf{u})]}{\log(\beta+\mathbb{E}[\mathsf{u}])} \approx 1$ since we can let α be arbitrarily close to 1.

Lemma 2.3. *Consider an $m \times n$ complex matrix $\boldsymbol{X}$ with $n \geq m$. Let $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ be the eigenvalues of $\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}$ sorted in decreasing order. For any $m' \leq m$, there exist index sets $\boldsymbol{s} = \{s_1, s_2, \ldots, s_{m'}\} \subset [n]$, $s_1 < \cdots < s_{m'}$, such that*

$$\log\det\left(\boldsymbol{I}_{m'} + \boldsymbol{X}_{[\boldsymbol{s}]}^{\mathsf{H}}\boldsymbol{X}_{[\boldsymbol{s}]}\right) - \sum_{i=1}^{m'} \log(1+\lambda_i) = O(1), \tag{2.8}$$

where $\boldsymbol{X}_{[\boldsymbol{s}]} := \left[\boldsymbol{X}_{[s_1]}\ \boldsymbol{X}_{[s_2]} \ldots \boldsymbol{X}_{[s_{m'}]}\right]$ and $O(1)$ denotes a constant independent of $\{\lambda_1, \ldots, \lambda_n\}$.

Proof. The proof is provided in Appendix 2.A.3. □

Hereafter, unless otherwise specified, $\{\sigma_i(\boldsymbol{M})\}_{i=1}^n$ denote n eigenvalues in decreasing order of a matrix $\boldsymbol{M} \in \mathbb{C}^{n\times n}$.

Lemma 2.4 (Inclusion Principle [145, Theorem 4.3.28])**.** *Let $\boldsymbol{A}$ be an $n \times n$ Hermitian matrix and $\boldsymbol{B}$ be an $m \times m$ block along the diagonal of $\boldsymbol{A}$.*[2] *Then*

$$\sigma_i(\boldsymbol{A}) \geq \sigma_i(\boldsymbol{B}) \geq \sigma_{i+n-m}(\boldsymbol{A}), \quad i \in [m]. \tag{2.9}$$

Corollary 2.1. *Let $\boldsymbol{A}$ be an $n \times n$ Hermitian matrix and $\boldsymbol{B}$ be an $m \times m$ block along the diagonal of $\boldsymbol{A}$. Then*

$$\log\det(\boldsymbol{I}_n + \boldsymbol{A}) \geq \log\det(\boldsymbol{I}_m + \boldsymbol{B}). \tag{2.10}$$

2.3 Main Result

The optimal DoF of the noncoherent MIMO generic block fading channel described above is stated in Theorem 2.1.

Theorem 2.1. *For the $M \times N$ noncoherent channel in generic, flat, and block fading with coherence interval T, if $T = 1$, the optimal DoF is zero; otherwise, the optimal DoF is given by*

$$d_{\text{opt}} = M^*\left(1 - \frac{M^*}{T}\right) \quad \textit{with} \quad M^* := \min\{M, N, \lfloor T/2 \rfloor\}. \tag{2.11}$$

The zero optimal DoF result for $T = 1$ has been shown in [36] and is included in Theorem 2.1 for completeness.

Corollary 2.2. *In the SIMO/MISO/SISO case ($\min\{M, N\} = 1$) or the $T = 2$ case, the optimal DoF is $d_{\text{opt}} = 1 - \frac{1}{T}$.*

Remark 2.1. *The DoF in* (2.11) *was shown to be optimal for the noncoherent IID Rayleigh block fading channel in [49], [50].*[3] *Theorem 2.1 generalizes this result to generic block fading.*

[2] $\boldsymbol{B}$ is called a principal sub-matrix of $\boldsymbol{A}$—see Definition 2.1 in Appendix 2.A.3.

[3] The converse proof for the case $1 < T < 2\min\{M, N\}$ given in [49, Section IV-D] was, however, based on some heuristic arguments and not rigorous.

For $T \geq 2$, the optimal DoF is achieved by using only M^* antennas and a simple pilot-based scheme: let the transmitter send pilot symbols in the first M^* channel uses, then data symbols in the remaining $T - M^*$ channel uses; the receiver estimates the channel based on the received pilot symbols and detects coherently the data symbols. In the following, we present the converse proof, first for the SIMO case, and then for the general MIMO case.[4]

2.4 The SIMO ($M = 1$) case

We first consider the SIMO case with $M = 1$ and $T \geq 2$. The received signal is

$$\mathbf{Y} = \mathbf{h}\mathbf{x}^{\mathsf{T}} + \mathbf{Z}, \tag{2.12}$$

where $\mathbf{h} \in \mathbb{C}^{N\times 1}$ and $\mathbf{x} \in \mathbb{C}^{T\times 1}$. If the channel is IID Rayleigh fading, i.e., $\mathbf{h} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{I}_N)$, it was shown that the optimal DoF is $1 - \frac{1}{T}$ and can be achieved with either a pilot-based scheme [31] or well-designed space time modulations [49], [50], [53]. For the converse of the high-SNR capacity (which implies the converse of the DoF), while $h(\mathbf{Y}|\mathbf{x})$ can be calculated easily, the upper bound for $h(\mathbf{Y})$ is much more involved [49], [50]. In this section, we provide a simpler proof for the converse of the DoF, which holds for generic fading, using the duality approach (see Appendix 1.A.4) as in [50] but with a simple choice of auxiliary output distribution.

First, let us define a random variable v as the index of the strongest input component, i.e.,[5]

$$\mathsf{v} := \arg\max_{i\in[T]} |\mathsf{x}_i|^2. \tag{2.13}$$

Thus, x_{v} denotes the entry in $\mathbf{x}$ with the largest magnitude. Let the genie give v to the receiver,[6] we have

$$\begin{aligned} I(\mathbf{x};\mathbf{Y}) &\leq I(\mathbf{x};\mathbf{Y},\mathsf{v}) && (2.14)\\ &= I(\mathbf{x};\mathbf{Y}|\mathsf{v}) + I(\mathbf{x};\mathsf{v}) && (2.15)\\ &\leq h(\mathbf{Y}|\mathsf{v}) - h(\mathbf{Y}|\mathbf{x},\mathsf{v}) + H(\mathsf{v}) && (2.16)\\ &\leq h(\mathbf{Y}|\mathsf{v}) - h(\mathbf{Y}|\mathbf{x}) + \log_2(T), && (2.17)\end{aligned}$$

where the last inequality holds because we have the Markov chain $\mathsf{v} \leftrightarrow \mathbf{x} \leftrightarrow \mathbf{Y}$ and $H(\mathsf{v}) \leq \log_2(T)$. For each given $\mathbf{x} = \boldsymbol{x}$, we can apply Lemma 2.1 with $\mathbf{W} = [\mathbf{h}\ \mathbf{Z}]$ and $\boldsymbol{A} = [\boldsymbol{x}\ \boldsymbol{I}_T]^{\mathsf{T}}$ to obtain

$$\begin{aligned} h(\mathbf{Y}|\mathbf{x}) &= N\mathbb{E}\left[\log_2 \det\left(\boldsymbol{I}_T + \mathbf{x}^*\mathbf{x}^{\mathsf{T}}\right)\right] + O(1) && (2.18)\\ &= N\mathbb{E}\left[\log_2\left(1 + \|\mathbf{x}\|^2\right)\right] + O(1). && (2.19)\end{aligned}$$

To bound $h(\mathbf{Y}|\mathsf{v})$, we use the duality approach [36] (see Appendix 1.A.4) as follows

$$\begin{aligned} h(\mathbf{Y}|\mathsf{v}) &= \mathbb{E}[-\log_2 p(\mathbf{Y}|\mathsf{v})]\\ &= \mathbb{E}[-\log_2 q(\mathbf{Y}|\mathsf{v})] - \mathbb{E}_{\mathsf{v}}[D(p_{\mathbf{Y}|\mathsf{v}}\|q_{\mathbf{Y}|\mathsf{v}})]\\ &\leq \mathbb{E}[-\log_2 q(\mathbf{Y}|\mathsf{v})], && (2.20)\end{aligned}$$

[4] Under the assumption (2.5), we assume w.l.o.g. that $\mathbb{E}\left[\|\mathbf{h}_k\|^2\right] = N$ for any column $\mathbf{h}_k$ of $\mathbf{H}$ for convenience.

[5] When there are more than one such components, we pick an arbitrary one.

[6] This technique of giving the index of the strongest input component to the receiver was initially proposed in [146] for phase noise channel, which is also a type of noncoherent channel.

due to the nonnegativity of the KL divergence $D(p_{\mathbf{Y}|\mathsf{v}}\|q_{\mathbf{Y}|\mathsf{v}})$. Here, conditioned on v, the distribution $p_{\mathbf{Y}|\mathsf{v}}$ with PDF $p(\cdot)$ is imposed by the input, channel, and noise distributions, while $q_{\mathbf{Y}|\mathsf{v}}$ is any distribution in $\mathbb{C}^{N\times T}$ with the PDF $q(.)$. Note that a proper choice of $q_{\mathbf{Y}|\mathsf{v}}$ is the key to a tight upper bound. Our choice is inspired by a pilot-based scheme. Specifically, if we send a pilot symbol at channel use $v \in [T]$, then the output vector $\mathbf{Y}_{[v]}$ being the sum of $\mathbf{h}$ and $\mathbf{Z}_{[v]}$ should have comparable power in each direction since $\mathbf{h}$ is generic by assumption. Therefore, it is reasonable (in the DoF sense) to let

$$\mathbf{Y}_{[v]} \sim \mathcal{R}(N, \boldsymbol{I}_N), \tag{2.21}$$

where the family of distributions $\mathcal{R}(N, \boldsymbol{A})$ is defined in Lemma 1.3. Now, $\mathbf{Y}_{[v]}$ should provide a rough estimate of the direction of the channel vector $\mathbf{h}$. Based on such an observation, it is also reasonable to assume that, given $\mathbf{Y}_{[v]}$, all other $\mathbf{Y}_{[i]}$, $i \neq v$, are mutually independent and follow

$$\mathbf{Y}_{[i]} \sim \mathcal{R}\Big(N, \big(\boldsymbol{I}_N + \mathbf{Y}_{[v]}\mathbf{Y}_{[v]}^{\mathsf{H}}\big)^{-\frac{1}{2}}\Big), \quad \forall i \neq v. \tag{2.22}$$

We thus obtain a "guess" of the auxiliary joint distribution $q_{\mathbf{Y}|\mathsf{v}=v}$.

Proposition 2.1. *With the above choice of auxiliary output distribution, we have the following bound for* $\mathbb{E}[-\log_2 q(\mathbf{Y}|\mathsf{v})]$ *and* $h(\mathbf{Y}|\mathsf{v})$*:*

$$\begin{aligned} h(\mathbf{Y}|\mathsf{v}) \le \mathbb{E}[-\log_2 q(\mathbf{Y}|\mathsf{v})] &\le (N+T-1)\mathbb{E}\Big[\log_2(1+|\mathsf{x}_\mathsf{v}|^2)\Big] \\ &\quad + N\mathbb{E}\left[\sum_{i=1, i\neq \mathsf{v}}^{T} \log_2\left(1+\frac{|\mathsf{x}_i|^2}{1+|\mathsf{x}_\mathsf{v}|^2}\right)\right] + O(\log\log P). \end{aligned} \tag{2.23}$$

Proof. The proof is provided in Appendix 2.A.4. □

Plugging (2.19) and (2.23) into (2.17), we obtain

$$\begin{aligned} I(\mathbf{x};\mathbf{Y}) &\le (T-1)\mathbb{E}\Big[\log_2(1+|\mathsf{x}_\mathsf{v}|^2)\Big] + N\mathbb{E}\left[\log_2\frac{1+|\mathsf{x}_\mathsf{v}|^2}{1+\|\mathbf{x}\|^2}\right] \\ &\quad + N\mathbb{E}\left[\sum_{i=1, i\neq \mathsf{v}}^{T} \log_2\left(1+\frac{|\mathsf{x}_i|^2}{1+|\mathsf{x}_\mathsf{v}|^2}\right)\right] + O(\log\log P) && (2.24) \\ &\le (T-1)\log_2\Big(1+\mathbb{E}\Big[|\mathsf{x}_\mathsf{v}|^2\Big]\Big) + O(\log\log P) && (2.25) \\ &\le (T-1)\log_2^+ P + O(\log\log P), && (2.26) \end{aligned}$$

where we have applied Jensen's inequality and the fact that $|\mathsf{x}_i|^2 \le |\mathsf{x}_\mathsf{v}|^2 \le \|\mathbf{x}\|^2$, $\forall i \neq \mathsf{v}$. Thus, the DoF is upper bounded by $\frac{T-1}{T}$, which is tight.

2.5 The MIMO Case

We now consider the general MIMO case and focus on the setting $T \ge 2\min\{M, N\} \ge 2$, i.e., $M^* = \min\{M, N\}$. The converse proof for $2 < T < 2\min\{M, N\}$ can be found in [119].

Let us fix a value $N_0 \in [\min\{M, N\}]$. Let $\lambda_1, \lambda_2, \ldots, \lambda_M$ be the random eigenvalues of $\mathbf{X}^{\mathsf{H}}\mathbf{X}$ sorted in decreasing order. According to Lemma 2.3, for any realization $\boldsymbol{X}$ of $\mathbf{X}$, there exist index vectors $\boldsymbol{s}$ of length N_0 with elements in $[\min\{M, N\}]$ such that

$$\log_2\det\Big(\boldsymbol{I}_{N_0} + \boldsymbol{X}_{[\boldsymbol{s}]}^{\mathsf{H}}\boldsymbol{X}_{[\boldsymbol{s}]}\Big) - \sum_{i=1}^{N_0}\log_2(1+\lambda_i) = O(1). \tag{2.27}$$

We define a random vector v whose values are drawn uniformly from all the possible $\boldsymbol{s}$ for each realization of X such that (2.27) holds. We have the Markov chain $\mathsf{v} \leftrightarrow \mathsf{X} \leftrightarrow \mathsf{Y}$.

Letting the genie give v to the receiver, and using the duality bound as in (2.17) and (2.20) in Section 2.4, we have

$$I(\mathsf{X};\mathsf{Y}) \leq \mathbb{E}[-\log_2 q(\mathsf{Y}|\mathsf{v})] - h(\mathsf{Y}|\mathsf{X}) + \log_2 \binom{T}{N_0}, \tag{2.28}$$

where $q(.)$ is the PDF of an auxiliary output distribution $q_{\mathsf{Y}|\mathsf{v}}$.

Given v, we can always permute the columns of X such that $\mathsf{X}_{[\mathsf{v}_1]}, \mathsf{X}_{[\mathsf{v}_2]}, \ldots, \mathsf{X}_{[\mathsf{v}_{N_0}]}$ become the first to N_0-th columns. For notational simplicity, we assume implicitly that X (and hence Y) is the permuted version. By using Lemma 2.1 with $\mathsf{W} = [\mathsf{H}\ \mathsf{Z}]$ and $\boldsymbol{A} = [\boldsymbol{X}\ \boldsymbol{I}_T]^{\mathsf{T}}$ for each realization $\boldsymbol{X}$ of X, the entropy $h(\mathsf{Y}|\mathsf{X})$ is given by

$$h(\mathsf{Y}|\mathsf{X}) = N\mathbb{E}\big[\log_2 \det(\boldsymbol{I}_T + \mathsf{X}^{\mathsf{H}}\mathsf{X})\big] + O(1). \tag{2.29}$$

Next, for $\mathbb{E}[-\log_2 q(\mathsf{Y}|\mathsf{v})]$, we choose the auxiliary PDF $q_{\mathsf{Y}|\mathsf{v}}$ such that

$$\mathsf{Y}_{[i]} \sim \mathcal{R}(N, \boldsymbol{I}_N), \quad \forall i \in [N_0], \tag{2.30}$$

and given $\mathsf{Y}_{[1:N_0]}$, the other columns are independent and follow

$$\mathsf{Y}_{[j]} \sim \mathcal{R}\Big(N, (\boldsymbol{I}_N + \mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}})^{-\frac{1}{2}}\Big), \quad j = N_0+1, \ldots, T. \tag{2.31}$$

As in the SIMO case, this choice is inspired by a pilot-based scheme that transmits pilot symbols in the first N_0 channel uses and data symbols in the remaining $T - N_0$ channel uses of a coherence block.

From Lemma 1.3 and Jensen's inequality, we have that

$$\begin{aligned}
&\mathbb{E}[-\log_2 q(\mathsf{Y}|\mathsf{v})] \\
&= N\sum_{i=1}^{N_0} \mathbb{E}\Big[\log_2 \|\mathsf{Y}_{[i]}\|^2\Big] + N\sum_{i=N_0+1}^{T} \mathbb{E}\left[\log_2 \left\|(\boldsymbol{I}_N + \mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}})^{-\frac{1}{2}}\mathsf{Y}_{[i]}\right\|^2\right] \\
&\quad + (T-N_0)\mathbb{E}\Big[\log_2 \det\Big(\boldsymbol{I}_N + \mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}}\Big)\Big] + O(\log\log P) \qquad (2.32) \\
&\leq N\sum_{i=1}^{N_0} \mathbb{E}\Big[\log_2 \left(\|\mathsf{X}_{[i]}\|^2 + 1\right)\Big] + N\sum_{i=N_0+1}^{T} \mathbb{E}\Big[\log_2\Big(\mathsf{Y}_{[i]}^{\mathsf{H}}(\boldsymbol{I}_N + \mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}})^{-1}\mathsf{Y}_{[i]}\Big)\Big] \\
&\quad + (T-N_0)\mathbb{E}\Big[\log_2 \det\Big(\boldsymbol{I}_{N_0} + \mathsf{X}_{[1:N_0]}^{\mathsf{H}}\mathsf{X}_{[1:N_0]}\Big)\Big] + O(\log\log P). \qquad (2.33)
\end{aligned}$$

Plugging the bounds into (2.28), we get that

$$\begin{aligned}
I(\mathsf{X};\mathsf{Y}) \leq\ & N\sum_{i=1}^{N_0} \mathbb{E}\Big[\log_2 \left(1 + \|\mathsf{X}_{[i]}\|^2\right)\Big] + (T-N_0)\mathbb{E}\Big[\log_2 \det\Big(\boldsymbol{I}_{N_0} + \mathsf{X}_{[1:N_0]}^{\mathsf{H}}\mathsf{X}_{[1:N_0]}\Big)\Big] \\
&+ N\sum_{i=N_0+1}^{T} \mathbb{E}\Big[\log_2\Big(\mathsf{Y}_{[i]}^{\mathsf{H}}(\boldsymbol{I}_N + \mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}})^{-1}\mathsf{Y}_{[i]}\Big)\Big] \\
&- N\mathbb{E}\big[\log_2 \det(\boldsymbol{I}_T + \mathsf{X}^{\mathsf{H}}\mathsf{X})\big] + O(\log\log P). \qquad (2.34)
\end{aligned}$$

Proposition 2.2. *With the chosen auxiliary output distribution, the mutual information* $I(\mathbf{X};\mathbf{Y})$ *can be bounded as*

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) \leq\ & NN_0\mathbb{E}[\log_2(1+\lambda_1)] + (T-N_0)\sum_{i=1}^{N_0}\mathbb{E}[\log_2(1+\lambda_i)] \\ & + N(T-N_0)\mathbb{E}\left[\sum_{i=1}^{\min\{M,N,N_0+1\}}\log_2(1+\lambda_i) - \sum_{i=1}^{N_0}\log_2(1+\lambda_i)\right] \\ & - N\sum_{i=1}^{M}\mathbb{E}[\log_2(1+\lambda_i)] + O(\log\log P), \end{aligned} \tag{2.35}$$

for any $N_0 \leq \min\{M,N\}$.

Proof. The proof is provided in Appendix 2.A.5. □

2.5.1 The Case $T = 2, M \geq 2, N \geq 2$

In this case, $\lambda_i = 0$ for all $i > 2$. We let $N_0 = T/2 = 1$ in (2.35) and obtain the bound

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &\leq N\mathbb{E}[\log_2(1+\lambda_1)] + \mathbb{E}[\log_2(1+\lambda_1)] + N\mathbb{E}[\log_2(1+\lambda_2)] \\ &\quad - N\sum_{i=1}^{2}\mathbb{E}[\log_2(1+\lambda_i)] + O(\log\log P) && (2.36) \\ &= \mathbb{E}[\log_2(1+\lambda_1)] + O(\log\log P) && (2.37) \\ &\leq \log_2(1+\mathbb{E}[\lambda_1]) + O(\log\log P) && (2.38) \\ &\leq \log_2^+ P + O(\log\log P) && (2.39) \end{aligned}$$

where (2.38) follows from Jensen's inequality and the last inequality holds because $\mathbb{E}[\lambda_1] \leq \mathbb{E}\big[\|\mathbf{X}\|_{\mathrm{F}}^2\big] \leq PT$. Therefore, the optimal DoF is upper bounded by $1/2$.

2.5.2 The Case $M \geq N, T \geq 2N$

In this case, we let $N_0 = N$ in (2.35) and obtain

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &\leq N^2\mathbb{E}[\log_2(1+\lambda_1)] + (T-2N)\sum_{i=1}^{N}\mathbb{E}[\log_2(1+\lambda_i)] + O(\log\log P) && (2.40) \\ &\leq (N^2 + N(T-2N))\mathbb{E}[\log_2(1+\lambda_1)] + O(\log\log P) && (2.41) \\ &\leq N(T-N)\log_2(1+\mathbb{E}[\lambda_1]) + O(\log\log P) && (2.42) \\ &\leq N(T-N)\log_2^+ P + O(\log\log P), && (2.43) \end{aligned}$$

where (2.41) holds because $T - 2N \geq 0$ and $\lambda_i \leq \lambda_1$, $\forall i$; (2.42) follows from Jensen's inequality; and the last inequality holds because $\mathbb{E}[\lambda_1] \leq \mathbb{E}\big[\|\mathbf{X}\|_{\mathrm{F}}^2\big] \leq PT$. Therefore, the optimal DoF is upper bounded by $N\left(1 - \frac{N}{T}\right)$.

2.5.3 The Case $M \leq N, T \geq 2M$

In this case, $\mathbf{H}$ is a tall matrix and can be written as

$$\mathbf{H} = \begin{bmatrix}\mathbf{H}_1 \\ \mathbf{H}_2\end{bmatrix}, \tag{2.44}$$

where $\mathbf{H}_1 \in \mathbb{C}^{M\times M}$ and $\mathbf{H}_2 \in \mathbb{C}^{(N-M)\times M}$. Because $\mathbb{E}\big[\|\mathbf{H}\|_{\mathrm{F}}^2\big] < \infty$, we have that $\mathbb{E}\big[\|\mathbf{H}_1\|_{\mathrm{F}}^2\big] < \infty$ and $\mathbb{E}\big[\|\mathbf{H}_2\|_{\mathrm{F}}^2\big] < \infty$. Furthermore, since $\mathbf{H}$ is generic, it is full rank almost surely (since otherwise, the differential entropy of $\mathbf{H}$ approaches $-\infty$), and therefore, we may w.l.o.g. assume that $\mathbf{H}_1$ is full rank almost surely. Therefore, there exists almost surely a spanning matrix $\tilde{\mathbf{H}} \in \mathbb{C}^{(N-M)\times M}$ such that

$$\mathbf{H}_2 = \tilde{\mathbf{H}}\mathbf{H}_1. \tag{2.45}$$

The matrix $\tilde{\mathbf{H}}$ must satisfy $\mathbb{E}\Big[\|\tilde{\mathbf{H}}\|_{\mathrm{F}}^2\Big] < \infty$ (because otherwise, $\mathbb{E}\big[\|\mathbf{H}_2\|_{\mathrm{F}}^2\big] = \infty$), and as a consequence, $h(\tilde{\mathbf{H}}) < \infty$. Then we have that

- $h(\mathbf{H}_1 \,|\, \tilde{\mathbf{H}}) = h(\mathbf{H}_1, \tilde{\mathbf{H}}) - h(\tilde{\mathbf{H}}) = h(\mathbf{H}) - h(\tilde{\mathbf{H}}) > -\infty$ because $h(\mathbf{H}) > -\infty$ and $h(\tilde{\mathbf{H}}) < \infty$;
- $\mathbb{E}\Big[\|\mathbf{H}_1\|_{\mathrm{F}}^2 \Big|\, \tilde{\mathbf{H}}\Big] < \infty$ almost surely because if $\mathbb{P}\Big(\tilde{\mathbf{H}} \in \mathcal{H}\Big) > 0$ where

  $$\mathcal{H} := \Big\{\tilde{\mathbf{H}} \in \mathbb{C}^{(N-M)\times M} : \mathbb{E}\Big[\|\mathbf{H}_1\|_{\mathrm{F}}^2 \Big|\, \tilde{\mathbf{H}}\Big] = \infty\Big\}$$

  then $\mathbb{E}\big[\|\mathbf{H}_1\|_{\mathrm{F}}^2\big] = \infty$, which is not true.

This implies that conditioned on $\tilde{\mathbf{H}}$, $\mathbf{H}_1$ is generic.

The channel output can be expressed as

$$\mathbf{Y} = \begin{bmatrix}\mathbf{Y}_1 \\ \mathbf{Y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{H}_1\mathbf{X} + \mathbf{Z}_1 \\ \mathbf{H}_2\mathbf{X} + \mathbf{Z}_2\end{bmatrix}, \tag{2.46}$$

where $\mathbf{Z}_1$ and $\mathbf{Z}_2$ contain the first M rows and the remaining $N-M$ rows of $\mathbf{Z}$, respectively. We have that

$$\begin{aligned}
I(\mathbf{X};\mathbf{Y}) &= I(\mathbf{X};\mathbf{Y}_1,\mathbf{Y}_2) && (2.47)\\
&\le I(\mathbf{X};\mathbf{Y}_1,\mathbf{Y}_2,\tilde{\mathbf{H}}) && (2.48)\\
&= I(\mathbf{X};\mathbf{Y}_1,\tilde{\mathbf{H}}) + I(\mathbf{X};\mathbf{Y}_2|\mathbf{Y}_1,\tilde{\mathbf{H}}) && (2.49)\\
&= I(\mathbf{X};\mathbf{Y}_1,\tilde{\mathbf{H}}) + h(\mathbf{Y}_2|\mathbf{Y}_1,\tilde{\mathbf{H}}) - h(\mathbf{Y}_2|\mathbf{X},\mathbf{Y}_1,\tilde{\mathbf{H}}) && (2.50)\\
&\le I(\mathbf{X};\mathbf{Y}_1|\tilde{\mathbf{H}}) + O(1), && (2.51)
\end{aligned}$$

where the last inequality follows because

$$\begin{aligned}
h(\mathbf{Y}_2|\mathbf{X},\mathbf{Y}_1,\tilde{\mathbf{H}}) &= h(\tilde{\mathbf{H}}(\mathbf{Y}_1 - \mathbf{Z}_1) + \mathbf{Z}_2|\mathbf{X},\mathbf{Y}_1,\tilde{\mathbf{H}}) && (2.52)\\
&\ge h(\mathbf{Z}_2 - \tilde{\mathbf{H}}\mathbf{Z}_1|\mathbf{X},\mathbf{Y}_1,\tilde{\mathbf{H}},\mathbf{Z}_1) && (2.53)\\
&= h(\mathbf{Z}_2) && (2.54)\\
&> 0, && (2.55)
\end{aligned}$$

and

$$\begin{aligned}
h(\mathbf{Y}_2|\mathbf{Y}_1,\tilde{\mathbf{H}}) &= h(\mathbf{Y}_2 - \tilde{\mathbf{H}}\mathbf{Y}_1 \,|\, \mathbf{Y}_1,\tilde{\mathbf{H}}) && (2.56)\\
&= \mathbb{E}_{\tilde{\mathbf{H}}}\big[h(\mathbf{Z}_2 - \tilde{\mathbf{H}}\mathbf{Z}_1|\, \mathbf{Y}_1,\tilde{\mathbf{H}} = \tilde{\boldsymbol{H}})\big] && (2.57)\\
&\le \mathbb{E}_{\tilde{\mathbf{H}}}\big[h(\mathbf{Z}_2 - \tilde{\mathbf{H}}\mathbf{Z}_1 \,|\, \tilde{\mathbf{H}} = \tilde{\boldsymbol{H}})\big] && (2.58)\\
&= M\mathbb{E}_{\tilde{\mathbf{H}}}\Big[\log_2\det\Big(\boldsymbol{I}_M + \tilde{\mathbf{H}}^{\mathsf{H}}\tilde{\mathbf{H}}\Big)\Big] + O(1) && (2.59)\\
&= M\sum_{i=1}^{M}\mathbb{E}\Big[\log_2\left(1 + \sigma_i(\tilde{\mathbf{H}}^{\mathsf{H}}\tilde{\mathbf{H}})\right)\Big] + O(1) && (2.60)\\
&\le M^2\log_2\Big(1 + \mathbb{E}\Big[\|\tilde{\mathbf{H}}\|_{\mathrm{F}}^2\Big]\Big) + O(1) && (2.61)\\
&= O(1). && (2.62)
\end{aligned}$$

Here, (2.58) holds because removing condition does not reduce entropy ; (2.59) holds because given $\tilde{\mathbf{H}}$, $\mathbf{Z}_2 - \tilde{\mathbf{H}}\mathbf{Z}_1$ is a Gaussian matrix with M independent columns having the same covariance matrix $\boldsymbol{I}_{N-M} + \tilde{\mathbf{H}}\tilde{\mathbf{H}}^{\mathsf{H}}$; (2.61) holds due to $\sigma_i(\tilde{\mathbf{H}}^{\mathsf{H}}\tilde{\mathbf{H}}) \leq \|\tilde{\mathbf{H}}\|_{\mathrm{F}}^2$ and Jensen's inequality; and the last equality is because $\mathbb{E}\left[\|\tilde{\mathbf{H}}\|_{\mathrm{F}}^2\right] < \infty$ by assumption.

Following (2.51), we aim to find an upper bound on $I(\mathbf{X};\mathbf{Y}_1|\tilde{\mathbf{H}})$. Observe that given $\tilde{\mathbf{H}}$, $\mathbf{Y}_1$ is the output of a $M \times M$ generic fading channel (since given $\tilde{\mathbf{H}}$, $\mathbf{H}_1$ is generic), which falls into the case of Section 2.5.2. Thus by the same argument as in Section 2.5.2, we can show that $I(\mathbf{X};\mathbf{Y}_1|\tilde{\mathbf{H}}) \leq M(T-M)\log_2^+ P + O(\log\log P)$. Plugging this into (2.51) yields

$$I(\mathbf{X};\mathbf{Y}) \leq M(T-M)\log_2^+ P + O(\log\log P), \tag{2.63}$$

thus the optimal DoF is upper bounded by $M\left(1-\frac{M}{T}\right)$.

Summarizing the results in Section 2.5.2 and Section 2.5.3, we conclude that the optimal DoF of the channel is upper bounded by $\min\{M,N\}\left(1-\frac{\min\{M,N\}}{T}\right)$ if $T \geq 2\min\{M,N\}$.

2.6 Closing Remarks

TABLE 2.1: The optimal DoF of the noncoherent $M \times N$ MIMO P2P channel in block fading with coherent interval T (the gray-colored cells contain the results of this chapter)

Assumption	$T=1$	$T \geq 2$
Generic fading	0 [36]	$\min\{M,N,\lfloor T/2\rfloor\}\left(1-\frac{\min\{M,N,\lfloor T/2\rfloor\}}{T}\right)$
IID Rayleigh fading		$\min\{M,N,\lfloor T/2\rfloor\}\left(1-\frac{\min\{M,N,\lfloor T/2\rfloor\}}{T}\right)$ [49], [50]

In this chapter, we study the optimal DoF of the noncoherent MIMO P2P channel with generic block fading. The findings are summarized, together with the known results, in Table 2.1. Based on the duality approach, we proposed an upper bound which coincides with the inner bound achieved by a simple pilot-based scheme. Our results generalize the optimal DoF of the Rayleigh fading case.

2.A Appendices

2.A.1 Proof of Lemma 2.1

Consider the eigendecomposition $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}$, where $\boldsymbol{U} \in \mathbb{C}^{m\times m}$ and $\boldsymbol{V} \in \mathbb{C}^{t\times t}$ are unitary matrices, and $\boldsymbol{\Sigma} = \begin{bmatrix}\boldsymbol{\Sigma}'\\ \mathbf{0}\end{bmatrix}$ with $\boldsymbol{\Sigma}' \in \mathbb{C}^{t\times t}$ a diagonal matrix containing the singular values of $\boldsymbol{A}$. Let $\mathbf{W}' = \mathbf{W}\boldsymbol{U}$, we have

$$\begin{aligned}
h(\mathbf{W}\boldsymbol{A}) &= h(\mathbf{W}\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}) && (2.64)\\
&= h(\mathbf{W}'\boldsymbol{\Sigma}) && (2.65)\\
&= h(\mathbf{W}'_{[1:t]}\boldsymbol{\Sigma}') && (2.66)\\
&= h(\mathbf{W}'_{[1:t]}) + n\log|\det(\boldsymbol{\Sigma}')|^2 && (2.67)\\
&= h(\mathbf{W}'_{[1:t]}) + n\log\det(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}), && (2.68)
\end{aligned}$$

where the second equality is because rotationing does not change differential entropy and (2.67) follows from (1.70). It remains to bound $h(\mathbf{W}'_{[1:t]})$. Since the average powers of $\mathbf{W}'_{[1:t]}$ and $\mathbf{W}'_{[t+1:m]}$ are bounded by $\mathbb{E}\big[\|\mathbf{W}'\|_{\mathrm{F}}^2\big] = \mathbb{E}\big[\|\mathbf{W}\|_{\mathrm{F}}^2\big] < \infty$, we have that

$$h(\mathbf{W}'_{[1:t]}) < \infty \tag{2.69}$$

and $h(\mathbf{W}'_{[t+1:m]}) < \infty$ by evoking the maximum differential entropy lemma in Appendix 1.A.1. Furthermore, it follows from $h(\mathbf{W}'_{[1:t]}) + h(\mathbf{W}'_{[t+1:m]}) \geq h(\mathbf{W}') = h(\mathbf{W}) > -\infty$ that

$$h(\mathbf{W}'_{[1:t]}) > -\infty - h(\mathbf{W}'_{[t+1:m]}) > -\infty. \tag{2.70}$$

Therefore, $h(\mathbf{W}'_{[1:t]})$ is bounded by some constant that only depends on the statistics of $\mathbf{W}$. This concludes the proof.

2.A.2 Proof of Lemma 2.2

If $\mathbb{E}[\mathsf{u}] \leq \beta$, we have that $\mathbb{E}[\log(\beta + \mathsf{u})] \geq \log(\beta)$ and $\alpha \log(\beta + \mathbb{E}[\mathsf{u}]) \leq \alpha \log(2\beta)$, so (2.7) holds with $c_0 = \log(\beta) - \alpha \log(2\beta) = (1-\alpha)\log\beta - \alpha \log 2$.

Let $p(\cdot)$ be the PDF of u. We introduce an auxiliary distribution with density

$$q(u) = \left(\frac{1}{\alpha} - 1\right)\beta^{\frac{1}{\alpha}-1}(\beta + u)^{-1/\alpha}, \quad u \geq 0, \tag{2.71}$$

with parameters $\alpha < 1, \beta > 0$. Then it follows that $h(\mathsf{u}) + \mathbb{E}[\log(q(\mathsf{u}))] = -D(p\|q) \leq 0$, which yields

$$\mathbb{E}[\log(\beta + \mathsf{u})] \geq \alpha h(\mathsf{u}) + \alpha \log\left(\frac{1}{\alpha} - 1\right) + (1-\alpha)\log\beta. \tag{2.72}$$

If $\mathbb{E}[\mathsf{u}] > \beta$, we have

$$\begin{aligned} h(\mathsf{u}) &= h\left(\mathbb{E}[\mathsf{u}]\frac{\mathsf{u}}{\mathbb{E}[\mathsf{u}]}\right) && (2.73)\\ &= \log(\mathbb{E}[\mathsf{u}]) + h\left(\frac{\mathsf{u}}{\mathbb{E}[\mathsf{u}]}\right) && (2.74)\\ &= \log(2\mathbb{E}[\mathsf{u}]) - \log 2 + h\left(\frac{\mathsf{u}}{\mathbb{E}[\mathsf{u}]}\right) && (2.75)\\ &> \log(\beta + \mathbb{E}[\mathsf{u}]) - \log 2 + h\left(\frac{\mathsf{u}}{\mathbb{E}[\mathsf{u}]}\right), && (2.76) \end{aligned}$$

then applying (2.72), (2.7) holds with

$$c_0 = \alpha \log\left(\frac{1}{\alpha} - 1\right) + (1-\alpha)\log\beta - \alpha\log 2 + \alpha h\left(\frac{\mathsf{u}}{\mathbb{E}[\mathsf{u}]}\right) > -\infty.$$

2.A.3 Proof of Lemma 2.3

The following results are necessary for our proof.

Definition 2.1 (Principal sub-matrix and principal minor [145, Section 0.7])**.** *Consider a matrix $\boldsymbol{M} \in \mathbb{C}^{n\times n}$. For each index set $\boldsymbol{s} = \{s_1, s_2, \ldots, s_t\} \subset [n]$, $t \leq n$, let $\boldsymbol{M}_{[\boldsymbol{s},\boldsymbol{s}]}$ denote the sub-matrix of $\boldsymbol{M}$ obtained by deleting all the rows $i \notin \boldsymbol{s}$ and columns $j \notin \boldsymbol{s}$ of $\boldsymbol{M}$. Then $\boldsymbol{M}_{[\boldsymbol{s},\boldsymbol{s}]}$ is called a* principal sub-matrix *of $\boldsymbol{M}$, and* $\det\left(\boldsymbol{M}_{[\boldsymbol{s},\boldsymbol{s}]}\right)$ *a* principal minor *of order t of $\boldsymbol{M}$.*

Lemma 2.5. *Consider a matrix $\boldsymbol{M} \in \mathbb{C}^{n\times n}$ with n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ sorted in decreasing order. It holds that*

$$\sum_{1\le i_1<\cdots<i_t\le n} \prod_{j=1}^{t} \lambda_{i_j} = \sum_{\boldsymbol{s}\subset[n]:\ |\boldsymbol{s}|=t} \det(\boldsymbol{M}_{[\boldsymbol{s},\boldsymbol{s}]}), \quad \forall t \in [n], \tag{2.77}$$

and

$$\det(\boldsymbol{I}_n + \boldsymbol{M}) = 1 + \sum_{t=1}^{n} \sum_{\boldsymbol{s}\subset[n]:\ |\boldsymbol{s}|=t} \det(\boldsymbol{M}_{[\boldsymbol{s},\boldsymbol{s}]}). \tag{2.78}$$

Proof. (2.77) follows from [145, Thm. 1.2.16]. Using (2.77) and the factorization

$$\det(\boldsymbol{I}_n + \boldsymbol{M}) = (1+\lambda_1)(1+\lambda_2)\ldots(1+\lambda_n) \tag{2.79}$$

$$= 1 + \sum_{t=1}^{n} \sum_{1\le i_1<\cdots<i_t\le n} \prod_{j=1}^{t} \lambda_{i_j}, \tag{2.80}$$

we obtain (2.78). □

We prove Lemma 2.3 by construction. Let $\boldsymbol{M} := \boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}$, then $\boldsymbol{M}_{[\boldsymbol{s},\boldsymbol{s}]} := \boldsymbol{X}^{\mathsf{H}}_{[\boldsymbol{s}]}\boldsymbol{X}_{[\boldsymbol{s}]}$ is a principal sub-matrix of $\boldsymbol{M}$. Let $\boldsymbol{s}_* \subset [n]$ be the index set corresponding to the highest principal minor of order at most m', i.e.,

$$\det\Big(\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}\Big) = \max_{\boldsymbol{s}\subset[n]:|\boldsymbol{s}|\le m'} \det\Big(\boldsymbol{M}_{[\boldsymbol{s},\boldsymbol{s}]}\Big). \tag{2.81}$$

Consider an index set $\boldsymbol{s}_\dagger \subset [n]$, $|\boldsymbol{s}_\dagger| = m'$, such that $\boldsymbol{s}_* \subset \boldsymbol{s}_\dagger$. Applying (2.78) in Lemma 2.5 to $\boldsymbol{X}^{\mathsf{H}}_{[\boldsymbol{s}_\dagger]}\boldsymbol{X}_{[\boldsymbol{s}_\dagger]}$, we obtain

$$\det\Big(\boldsymbol{I}_{m'} + \boldsymbol{X}^{\mathsf{H}}_{[\boldsymbol{s}_\dagger]}\boldsymbol{X}_{[\boldsymbol{s}_\dagger]}\Big) = 1 + \sum_{t=1}^{m'} \sum_{\boldsymbol{s}\subset\boldsymbol{s}_\dagger:|\boldsymbol{s}|=t} \det\Big(\boldsymbol{X}^{\mathsf{H}}_{[\boldsymbol{s}]}\boldsymbol{X}_{[\boldsymbol{s}]}\Big) \tag{2.82}$$

and thus

$$1 + \det\Big(\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}\Big) \le \det(\boldsymbol{I}_{m'} + \boldsymbol{X}^{\mathsf{H}}_{[\boldsymbol{s}_\dagger]}\boldsymbol{X}_{[\boldsymbol{s}_\dagger]}) \tag{2.83}$$

$$\le 1 + \sum_{t=1}^{m'} \binom{m'}{t} \det\Big(\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}\Big). \tag{2.84}$$

On the other hand, using (2.77) in Lemma 2.5, we have that

$$\prod_{i=1}^{m'}(1+\lambda_i) = 1 + \sum_{t=1}^{m'} \sum_{1\le i_1<\cdots<i_t\le m'} \prod_{j=1}^{t} \lambda_{i_j} \tag{2.85}$$

$$\le 1 + \sum_{t=1}^{m'} \sum_{1\le i_1<\cdots<i_t\le n} \prod_{j=1}^{t} \lambda_{i_j} \tag{2.86}$$

$$= 1 + \sum_{t=1}^{m'} \sum_{\boldsymbol{s}\subset[n]:|\boldsymbol{s}|=t} \det(\boldsymbol{M}_{[\boldsymbol{s},\boldsymbol{s}]}) \tag{2.87}$$

$$\le 1 + \sum_{t=1}^{m'} \binom{n}{t} \det\Big(\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}\Big), \tag{2.88}$$

where (2.86) holds because $\lambda_i \geq 0, \forall i \in [n]$ (since $\boldsymbol{M} = \boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}$ is positive semi-definite). Let r be the rank of $\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}$. Since $\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}$ is a principle sub-matrix of $\boldsymbol{M}$, from the inclusion principle (Lemma 2.4), we have that the r eigenvalues of $\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}$ sorted in decreasing order are respectively smaller than the corresponding r largest eigenvalues of $\boldsymbol{M}$. Thus $\det\Big(\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}\Big) \leq \prod_{i=1}^{r} \lambda_i$, which implies that

$$1 + \det\Big(\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}\Big) \leq \prod_{j=1}^{m'}(1+\lambda_j). \tag{2.89}$$

Therefore,

$$\begin{aligned}
\frac{1}{1+\sum_{t=1}^{m'}\binom{n}{t}} &\leq \frac{1+\det\Big(\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}\Big)}{1+\sum_{t=1}^{m'}\binom{n}{t}\det\Big(\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}\Big)} && (2.90)\\
&\leq \frac{\det\Big(\boldsymbol{I}_{m'} + \boldsymbol{X}^{\mathsf{H}}_{[\boldsymbol{s}_\dagger]}\boldsymbol{X}_{[\boldsymbol{s}_\dagger]}\Big)}{\prod_{i=1}^{m'}(1+\lambda_i)} && (2.91)\\
&\leq \frac{1+\sum_{t=1}^{m'}\binom{m'}{t}\det\Big(\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}\Big)}{1+\det\Big(\boldsymbol{M}_{[\boldsymbol{s}_*,\boldsymbol{s}_*]}\Big)} && (2.92)\\
&\leq 1+\sum_{t=1}^{m'}\binom{m'}{t}, && (2.93)
\end{aligned}$$

where (2.91) follows from (2.83) and (2.88); (2.92) from (2.84) and (2.89); (2.90) and (2.93) are due to $\frac{1+ax}{1+x} \leq 1+a$ for any $x>0, a>0$. Taking the logarithm, this implies (2.8) with $\boldsymbol{s} = \boldsymbol{s}_\dagger$. Thus, we complete the proof.

2.A.4 Proof of Proposition 2.1

Using Lemma 1.3, it follows that

$$\begin{aligned}
&\mathbb{E}[-\log_2 q(\mathbf{Y}|\mathsf{v}=v)]\\
&= N\mathbb{E}\Big[\log_2\|\mathbf{Y}_{[v]}\|^2\Big]\\
&\quad + \sum_{i=1,i\neq v}^{T}\mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_N + \mathbf{Y}_{[v]}\mathbf{Y}^{\mathsf{H}}_{[v]}\Big) + N\log_2\Big\|\Big(\boldsymbol{I}_N + \mathbf{Y}_{[v]}\mathbf{Y}^{\mathsf{H}}_{[v]}\Big)^{-\frac{1}{2}}\mathbf{Y}_{[i]}\Big\|^2\Big]\\
&\quad + O(\log\log P) && (2.94)\\
&= N\mathbb{E}\Big[\log_2\|\mathbf{Y}_{[v]}\|^2\Big] + \sum_{i=1,i\neq v}^{T}\mathbb{E}\left[\log_2\Big(1+\|\mathbf{Y}_{[v]}\|^2\Big) + N\log_2\left(\|\mathbf{Y}_{[i]}\|^2 - \frac{|\mathbf{Y}^{\mathsf{H}}_{[i]}\mathbf{Y}_{[v]}|^2}{1+\|\mathbf{Y}_{[v]}\|^2}\right)\right]\\
&\quad + O(\log\log P) && (2.95)\\
&= (N+T-1)\mathbb{E}\Big[\log_2(1+\|\mathbf{Y}_{[v]}\|^2)\Big]\\
&\quad + N\sum_{i=1,i\neq v}^{T}\mathbb{E}\Big[\log_2(\|\mathbf{Y}_{[i]}\|^2 + \|\mathbf{Y}_{[i]}\|^2\|\mathbf{Y}_{[v]}\|^2 - |\mathbf{Y}^{\mathsf{H}}_{[i]}\mathbf{Y}_{[v]}|^2) - \log_2(1+\|\mathbf{Y}_{[v]}\|^2)\Big]\\
&\quad + O(\log\log P), && (2.96)
\end{aligned}$$

where in the second equality, we used the identities $\det(\boldsymbol{I}+\boldsymbol{u}\boldsymbol{v}^{\mathsf{H}}) = 1+\boldsymbol{v}^{\mathsf{H}}\boldsymbol{u}$ and $\|(\boldsymbol{I}+\boldsymbol{u}\boldsymbol{u}^{\mathsf{H}})^{-1/2}\boldsymbol{x}\|^2 = \boldsymbol{x}^{\mathsf{H}}(\boldsymbol{I}+\boldsymbol{u}\boldsymbol{u}^{\mathsf{H}})^{-1}\boldsymbol{x} = \boldsymbol{x}^{\mathsf{H}}\Big(\boldsymbol{I} - \frac{\boldsymbol{u}\boldsymbol{u}^{\mathsf{H}}}{1+\boldsymbol{u}^{\mathsf{H}}\boldsymbol{u}}\Big)\boldsymbol{x} = \|\boldsymbol{x}\|^2 - \frac{|\boldsymbol{x}^{\mathsf{H}}\boldsymbol{u}|^2}{1+\|\boldsymbol{u}\|^2}$.

By expanding $\mathbf{Y}_{[1]}, \ldots, \mathbf{Y}_{[T]}$, we get that, given $\mathbf{x}$ and for a fixed $v \in [T]$,

$$\mathbb{E}_{\mathbf{h},\mathbf{Z}}\left[\|\mathbf{Y}_{[i]}\|^2\right] = N(1 + |\mathsf{x}_i|^2), \quad \forall i, \tag{2.97}$$

$$\mathbb{E}_{\mathbf{h},\mathbf{Z}}\left[\|\mathbf{Y}_{[i]}\|^2\|\mathbf{Y}_{[v]}\|^2 - |\mathbf{Y}_{[i]}^{\mathsf{H}}\mathbf{Y}_{[v]}|^2\right] = (N^2 - N)(1 + |\mathsf{x}_v|^2 + |\mathsf{x}_i|^2), \quad i \neq v. \tag{2.98}$$

Then, it follows that

$$\begin{aligned}
&\mathbb{E}[-\log_2 q(\mathbf{Y}|\mathsf{v} = v)] \\
&\leq (N + T - 1)\mathbb{E}\Big[\log_2(1 + N + N|\mathsf{x}_v|^2)\Big] \\
&\quad + N \sum_{i=1, i\neq \mathsf{v}}^{T} \mathbb{E}\left[\log_2 \frac{N + N|\mathsf{x}_i|^2 + (N^2 - N)(1 + |\mathsf{x}_v|^2 + |\mathsf{x}_i|^2)}{1 + N + N|\mathsf{x}_v|^2}\right] \\
&\quad + O(\log\log P) \qquad (2.99) \\
&= (N + T - 1)\mathbb{E}\Big[\log_2(1 + |\mathsf{x}_v|^2)\Big] + N \sum_{i=1, i\neq v}^{T} \mathbb{E}\left[\log_2\left(1 + \frac{|\mathsf{x}_i|^2}{1 + |\mathsf{x}_v|^2}\right)\right] \\
&\quad + O(\log\log P), \qquad (2.100)
\end{aligned}$$

where (2.99) is obtained by using Jensen's inequality and applying Lemma 2.2 (with $\beta = 1$ and α arbitrarily close to 1) to $\|\mathbf{Y}_{[v]}\|^2$. Then, taking expectation of both sides of (2.100) w.r.t. v, we obtain (2.23).

It remains to show that $\mathsf{u} := \|\mathbf{Y}_{[v]}\|^2$ fulfills the conditions $\mathbb{E}[\mathsf{u}] < \infty$ and $h(\mathsf{u}/\mathbb{E}[\mathsf{u}]) > -\infty$ for Lemma 2.2. To see this, first, $\mathbb{E}[\mathsf{u}] = N(1 + |\mathsf{x}_v|^2)$, so it is finite for any fixed input entry $\mathsf{x}_v = x_v$. Second, we have that[7]

$$\begin{aligned}
h(\mathsf{u}/\mathbb{E}[\mathsf{u}]) &= h(\mathsf{u}) - \log_2(\mathbb{E}[\mathsf{u}]) \qquad (2.101) \\
&= h(\|\mathbf{h}\mathsf{x}_v + \mathbf{Z}_{[v]}\|^2) - \log_2(N(1 + |\mathsf{x}_v|^2)). \qquad (2.102)
\end{aligned}$$

It follows from [36, Lemma 6.17] that

$$h(\|\mathbf{w}\|^2) = h(\mathbf{w}) - h_\lambda\left(\frac{\mathbf{w}}{\|\mathbf{w}\|}\Big|\|\mathbf{w}\|\right) - (N - 1)\mathbb{E}\Big[\log_2(\|\mathbf{w}\|^2)\Big] + 1, \tag{2.103}$$

for a complex random vector $\mathbf{w} \in \mathbb{C}^N$, where $h_\lambda(.)$ is the modified differential entropy function defined for random unit-norm vector, w.r.t. the area measure on the unit sphere. Since the unit sphere in $\mathbb{C}^N$ has finite area $\frac{2\pi^N}{\Gamma(N)}$, we have that

$$h_\lambda\left(\frac{\mathbf{w}}{\|\mathbf{w}\|}\Big|\|\mathbf{w}\|\right) \leq \log_2 \frac{2\pi^N}{\Gamma(N)}. \tag{2.104}$$

[7] In the case of IID Rayleigh fading $\mathbf{h} \sim \mathcal{CN}(0, \boldsymbol{I}_N)$, we can compute explicitly $h(\mathsf{u}/\mathbb{E}[\mathsf{u}]) = N + \log_2 \frac{\Gamma(N)}{N} - (N - 1)\psi(N)$, where $\psi(N)$ is Euler's digamma function. It is obvious that $h(\mathsf{u}/\mathbb{E}[\mathsf{u}]) < \infty$.

Taking $\mathbf{w} = \mathbf{h}\mathsf{x}_v + \mathbf{Z}_{[v]}$ and plugging $h(\|\mathbf{w}\|^2)$ from (2.103) to (2.102) yields

$$
\begin{aligned}
&h(\mathsf{u}/\mathbb{E}[\mathsf{u}]) \\
&\geq h(\mathbf{h}\mathsf{x}_v + \mathbf{Z}_{[v]}) - (N-1)\mathbb{E}\Big[\log_2(\|\mathbf{h}\mathsf{x}_v + \mathbf{Z}_{[v]}\|^2)\Big] - \log_2(N(1+|\mathsf{x}_v|^2)) - \log_2 \frac{\pi^N}{\Gamma(N)} && (2.105)\\
&\geq \max\{h(\mathbf{h}\mathsf{x}_v), h(\mathbf{Z}_{[v]})\} - (N-1)\log_2\Big(\mathbb{E}\Big[\|\mathbf{h}\mathsf{x}_v + \mathbf{Z}_{[v]}\|^2\Big]\Big) - \log_2(N(1+|\mathsf{x}_v|^2)) \\
&\quad - \log_2 \frac{\pi^N}{\Gamma(N)} && (2.106)\\
&= \max\{h(\mathbf{h}) + N\log_2(|\mathsf{x}_v|^2), N\log_2(\pi e)\} - N\log_2(1+|\mathsf{x}_v|^2) - \log_2 \frac{(N\pi)^N}{\Gamma(N)} && (2.107)\\
&= N\log_2\Big(\max\{|\mathsf{x}_v|^2 e^{h(\mathbf{h})/N}, \pi e\}\Big) - N\log_2(1+|\mathsf{x}_v|^2) - \log_2 \frac{(N\pi)^N}{\Gamma(N)} && (2.108)\\
&\geq N\log_2(\pi e) - N\log_2\Big(1 + \pi e^{1-h(\mathbf{h})/N}\Big) - \log_2 \frac{N^N}{\Gamma(N)} && (2.109)\\
&> -\infty, && (2.110)
\end{aligned}
$$

where the second inequality follows from Jensen's inequality and

$$h\Big(\mathbf{h}\mathsf{x}_v + \mathbf{Z}_{[v]}\Big) \geq \max\Big\{h\Big(\mathbf{h}\mathsf{x}_v\Big), h(\mathbf{Z}_{[v]}\Big)\Big\};$$

the second-to-last inequality follows by inspecting two cases $|\mathsf{x}_v|^2 e^{h(\mathbf{h})/N} \geq \pi e$ and $|\mathsf{x}_v|^2 e^{h(\mathbf{h})/N} < \pi e$; and the last inequality holds because $h(\mathbf{h})$ is finite by assumption.

2.A.5 Proof of Proposition 2.2

We bound each term in the right-hand side of (2.34) as follows. The first term is bounded using

$$
\begin{aligned}
\mathbb{E}\Big[\log_2(1+\|\mathbf{X}_{[i]}\|^2)\Big] &\leq \mathbb{E}\Big[\log_2(1+\|\mathbf{X}\|_{\mathrm{F}}^2)\Big] && (2.111)\\
&= \mathbb{E}\left[\log_2\left(1 + \sum_{j=1}^{N_0}\lambda_j\right)\right] && (2.112)\\
&\leq \mathbb{E}[\log_2(1+\lambda_1)] + O(1). && (2.113)
\end{aligned}
$$

For the second term, it follows from (2.27) that

$$\mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_{N_0} + \mathbf{X}_{[1:N_0]}^{\mathsf{H}}\mathbf{X}_{[1:N_0]}\Big)\Big] = \sum_{i=1}^{N_0}\mathbb{E}[\log_2(1+\lambda_i)] + O(1). \tag{2.114}$$

The third term is bounded as

$$\begin{aligned}
&\mathbb{E}\Big[\log_2\Big(\mathsf{Y}_{[i]}^{\mathsf{H}}(\boldsymbol{I}_N+\mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}})^{-1}\mathsf{Y}_{[i]}\Big)\Big]\\
&\le \mathbb{E}\Big[\log_2\Big(1+\mathsf{Y}_{[i]}^{\mathsf{H}}(\boldsymbol{I}_N+\mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}})^{-1}\mathsf{Y}_{[i]}\Big)\Big] && (2.115)\\
&= \mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_N+(\boldsymbol{I}_N+\mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}})^{-1}\mathsf{Y}_{[i]}\mathsf{Y}_{[i]}^{\mathsf{H}}\Big)\Big] && (2.116)\\
&= \mathbb{E}\Big[\log_2\det\Big((\boldsymbol{I}_N+\mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}})^{-1}(\boldsymbol{I}_N+\mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}}+\mathsf{Y}_{[i]}\mathsf{Y}_{[i]}^{\mathsf{H}})\Big)\Big] && (2.117)\\
&= \mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_N+[\mathsf{Y}_{[1:N_0]}\ \mathsf{Y}_{[i]}][\mathsf{Y}_{[1:N_0]}\ \mathsf{Y}_{[i]}]^{\mathsf{H}}\Big)\Big]-\mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_N+\mathsf{Y}_{[1:N_0]}\mathsf{Y}_{[1:N_0]}^{\mathsf{H}}\Big)\Big] && (2.118)\\
&= \mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_N+\mathsf{H}[\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}][\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}]^{\mathsf{H}}\mathsf{H}^{\mathsf{H}}\Big)\Big]\\
&\quad-\mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_N+\mathsf{H}\mathsf{X}_{[1:N_0]}\mathsf{X}_{[1:N_0]}^{\mathsf{H}}\mathsf{H}^{\mathsf{H}}\Big)\Big]+O(1) && (2.119)\\
&= \mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_M+[\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}][\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}]^{\mathsf{H}}\mathsf{H}^{\mathsf{H}}\mathsf{H}\Big)\Big]\\
&\quad-\mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_M+\mathsf{X}_{[1:N_0]}\mathsf{X}_{[1:N_0]}^{\mathsf{H}}\mathsf{H}^{\mathsf{H}}\mathsf{H}\Big)\Big]+O(1), && (2.120)
\end{aligned}$$

where (2.119) is obtained by applying Lemma 2.6 at the end of this appendix, which says that an additive Gaussian noise with bounded variance does not affect the pre-log. We proceed to bound the two expected values in (2.119). For an $n\times n$ matrix $\boldsymbol{M}$, we denote $\sigma_1(\boldsymbol{M}),\sigma_2(\boldsymbol{M}),\ldots,\sigma_n(\boldsymbol{M})$ its n eigenvalues sorted in decreasing order. Noting that the rank of $[\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}][\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}]^{\mathsf{H}}\mathsf{H}^{\mathsf{H}}\mathsf{H}$ is upper bounded by $\min\{M,N,N_0+1\}$, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_M+[\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}][\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}]^{\mathsf{H}}\mathsf{H}^{\mathsf{H}}\mathsf{H}\Big)\Big]\\
&=\sum_{i=1}^{\min\{M,N,N_0+1\}}\mathbb{E}\Big[\log_2\Big(1+\sigma_i([\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}][\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}]^{\mathsf{H}}\mathsf{H}^{\mathsf{H}}\mathsf{H})\Big)\Big] && (2.121)\\
&\le\sum_{i=1}^{\min\{M,N,N_0+1\}}\mathbb{E}\Big[\log_2\Big(1+\sigma_i([\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}][\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}]^{\mathsf{H}})\sigma_1(\mathsf{H}^{\mathsf{H}}\mathsf{H})\Big)\Big] && (2.122)\\
&\le\sum_{i=1}^{\min\{M,N,N_0+1\}}\mathbb{E}_{\mathsf{X}}\Big[\log_2\Big(1+\sigma_i([\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}][\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}]^{\mathsf{H}})\mathbb{E}_{\mathsf{H}}\big[\|\mathsf{H}\|_{\mathrm{F}}^2\big]\Big)\Big] && (2.123)\\
&=\sum_{i=1}^{\min\{M,N,N_0+1\}}\mathbb{E}\Big[\log_2\Big(1+\sigma_i([\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}][\mathsf{X}_{[1:N_0]}\ \mathsf{X}_{[i]}]^{\mathsf{H}})\Big)\Big]+O(1) && (2.124)\\
&\le\sum_{i=1}^{\min\{M,N,N_0+1\}}\mathbb{E}[\log_2(1+\lambda_i)]+O(1), && (2.125)
\end{aligned}$$

where (2.122) follows by applying Lemma 2.7 at the end of this appendix; (2.123) is due to $\sigma_1(\mathsf{H}^{\mathsf{H}}\mathsf{H})\le\|\mathsf{H}\|_{\mathrm{F}}^2$ and Jensen's inequality; (2.124) holds because $\mathbb{E}\big[\|\mathsf{H}\|_{\mathrm{F}}^2\big]<\infty$ by assumption; and (2.125) is due to the inclusion principle (Lemma 2.4). Consider the eigendecomposition $\mathsf{X}_{[1:N_0]}\mathsf{X}_{[1:N_0]}^{\mathsf{H}}=\mathsf{U}\begin{bmatrix}\boldsymbol{\Sigma}&\mathbf{0}\\\mathbf{0}&\mathbf{0}\end{bmatrix}\mathsf{U}^{\mathsf{H}}$ where $\boldsymbol{\Sigma}$ is an $N_0\times N_0$ diagonal matrix,

we have that

$$\log_2 \det\Big(\boldsymbol{I}_M + \mathbf{X}_{[1:N_0]}\mathbf{X}_{[1:N_0]}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{H}\Big) = \log_2 \det\left(\boldsymbol{I}_M + \mathbf{U}\begin{bmatrix}\boldsymbol{\Sigma} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{U}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{H}\right) \tag{2.126}$$

$$= \log_2 \det\left(\boldsymbol{I}_M + \begin{bmatrix}\boldsymbol{\Sigma} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\tilde{\mathbf{H}}^{\mathsf{H}}\tilde{\mathbf{H}}\right) \tag{2.127}$$

$$\geq \log_2 \det\Big(\boldsymbol{I}_{N_0} + \boldsymbol{\Sigma}\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}}\Big) \tag{2.128}$$

$$= \sum_{i=1}^{N_0} \log_2\Big(1 + \sigma_i(\boldsymbol{\Sigma}\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}})\Big) \tag{2.129}$$

$$\geq \sum_{i=1}^{N_0} \log_2\Big(1 + \sigma_i(\mathbf{X}_{[1:N_0]}\mathbf{X}_{[1:N_0]}^{\mathsf{H}})\sigma_{N_0}(\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}})\Big), \tag{2.130}$$

where in (2.127), $\tilde{\mathbf{H}} := \mathbf{H}\mathbf{U}$ is also a generic fading matrix; (2.128) follows from Corollary 2.1 with $\bar{\mathbf{H}}$ containing the first N_0 columns of $\tilde{\mathbf{H}}$; (2.130) is obtained by applying Lemma 2.7 and the fact that $\boldsymbol{\Sigma}$ contains N_0 largest eigenvalues of $\mathbf{X}_{[1:N_0]}\mathbf{X}_{[1:N_0]}^{\mathsf{H}}$. Since $\tilde{\mathbf{H}}$ is generic, $\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}}$ is full rank almost surely, thus $\sigma_{N_0}(\bar{\mathbf{H}}^{\mathsf{H}}\bar{\mathbf{H}}) \geq c_0 > 0$ almost surely. Therefore, taking the expectation, we have that

$$\mathbb{E}\Big[\log_2 \det\Big(\boldsymbol{I}_M + \mathbf{X}_{[1:N_0]}\mathbf{X}_{[1:N_0]}^{\mathsf{H}}\mathbf{H}^{\mathsf{H}}\mathbf{H}\Big)\Big]$$

$$\geq \sum_{i=1}^{N_0} \mathbb{E}\Big[\log_2\Big(1 + \sigma_i(\mathbf{X}_{[1:N_0]}\mathbf{X}_{[1:N_0]}^{\mathsf{H}})\Big)\Big] + O(1) \tag{2.131}$$

$$= \mathbb{E}\Big[\log_2 \det\Big(\boldsymbol{I}_{N_0} + \mathbf{X}_{[1:N_0]}^{\mathsf{H}}\mathbf{X}_{[1:N_0]}\Big)\Big] + O(1) \tag{2.132}$$

$$= \sum_{i=1}^{N_0} \mathbb{E}[\log_2(1+\lambda_i)] + O(1), \tag{2.133}$$

where the last equality follows from (2.27). Plugging (2.125) and (2.133) into (2.118), we have an upper bound on the third term in the right-hand side of (2.34).

The fourth term in the right-hand side of (2.34) is expanded using $\det(\boldsymbol{I}_T + \mathbf{X}^{\mathsf{H}}\mathbf{X}) = \prod_{i=1}^{M}(1+\lambda_i)$.

Finally, substituting the bound/expansion of each term in the right-hand side of (2.34), we obtain (2.35).

Lemma 2.6. *Consider an $m \times n$ random matrix $\mathbf{G} = \hat{\mathbf{G}} + \tilde{\mathbf{G}}$ where $\tilde{\mathbf{G}}$ has IID $\mathcal{N}_{\mathbb{C}}(0,1)$ entries independent of $\hat{\mathbf{G}}$. It holds that*

$$\mathbb{E}_{\tilde{\mathbf{G}}}\big[\log\det(\boldsymbol{I}_m + \mathbf{G}\mathbf{G}^{\mathsf{H}})\big] = \log\det\Big(\boldsymbol{I}_m + \hat{\mathbf{G}}\hat{\mathbf{G}}^{\mathsf{H}}\Big) + o(\log P). \tag{2.134}$$

Proof. Following the footsteps of [147, Lemma 1], we can show that

$$\mathbb{E}_{\tilde{\mathbf{G}}}\big[\log\det(\boldsymbol{I}_m + \mathbf{G}\mathbf{G}^{\mathsf{H}})\big] = \sum_{i=1}^{\tau} \log\left(1 + \sigma_i(\hat{\mathbf{G}}\hat{\mathbf{G}}^{\mathsf{H}})\right) + o(\log P), \tag{2.135}$$

where $\tau \leq \operatorname{rank}(\hat{\mathbf{G}})$ is the number of eigenvalues of $\hat{\mathbf{G}}\hat{\mathbf{G}}^{\mathsf{H}}$ that do not vanish with P, i.e., $\sigma_i(\hat{\mathbf{G}}\hat{\mathbf{G}}^{\mathsf{H}}) = o(1)$ when $P \to \infty$, $\forall i > \tau$. It follows that

$$\mathbb{E}_{\tilde{\mathbf{G}}}\big[\log\det(\boldsymbol{I}_m + \mathbf{G}\mathbf{G}^{\mathsf{H}})\big] = \sum_{i=1}^{\operatorname{rank}(\hat{\mathbf{G}})} \log(1 + \sigma_i(\hat{\mathbf{G}}\hat{\mathbf{G}}^{\mathsf{H}})) + o(\log P) \tag{2.136}$$

since the remaining $\mathrm{rank}(\hat{\mathsf{G}}) - \tau$ eigenvalues do not contribute more than $o(\log P)$ to the expectation. This implies (2.134). □

Lemma 2.7. *If $\boldsymbol{A}$ and $\boldsymbol{B}$ are $n \times n$ Hermitian positive semidefinite matrices, then*

$$\sigma_i(\boldsymbol{A})\sigma_n(\boldsymbol{B}) \leq \sigma_i(\boldsymbol{AB}) \leq \sigma_i(\boldsymbol{A})\sigma_1(\boldsymbol{B}), \quad i \in [n]. \tag{2.137}$$

Proof. The result follows immediately by applying [148, Theorem 3] and [148, Theorem 4] with $k = 1$ therein. □

Chapter 3

The Two-User SIMO Multiple-Access Channel

The optimal DoF region of the non-coherent block-fading MAC is still unknown in general. In this chapter, we make some progress by deriving the entire optimal DoF region in the case of the two-user SIMO MAC in generic block fading. The achievability is based on a simple pilot-based scheme. The novelty of our result lies in the converse using a genie-aided bound and the duality upper bound.

3.1 Overview

For the non-coherent MAC, the sum capacity has been studied in [84]–[87]. The capacity/DoF region is not known, and only some achievable DoF regions achieving the optimal sum DoF in IID Rayleigh block fading have been proposed [37], [88].

In this chapter, we make some progress for the non-coherent SIMO MAC. Specifically, we derive the optimal DoF region in the case of two single-antenna transmitters (users) and a N-antenna receiver in generic block fading channel with coherence interval T. When $N = 1$, the region is achieved with a simple time division multiplexing (TDM) between two users. In this case, letting two users cooperate does not help exploit more DoF and it is optimal to activate only one user at a time to achieve $1-\frac{1}{T}$ DoF for that user. When $N > 1$, a pilot-based scheme can achieve another DoF pair. We let two users send orthogonal pilots for channel estimation in 2 of the channel uses, and send data simultaneously in the remaining $T-2$ channel uses. In this way, each user achieves $1-\frac{2}{T}$ DoF.

The main technical contribution of this chapter lies in the converse proof. Leveraging the duality upper bound [36], we carefully choose an output distribution with which we derive a tight outer bound on the DoF region. This is a generalization of our converse proof technique for the single-user SIMO case in Chapter 2. Unlike previous results such as [37], [84], [85], [87], [88], we do not assume Gaussianity of the channel coefficients, which makes our proof more general.

The remainder of this chapter is organized as follows. The system model and preliminaries are presented in Section 3.2. In Section 3.3, we provide the main result on the optimal DoF region of the two-user MAC, as well as the proof for the case $N = 1$ and the

achievability for the case $N > 1$. We show the tight outer bound for the case $N > 1$ in Section 3.4. Finally, we conclude the chapter in Section 3.5.

3.2 System Model and Preliminaries

We consider a SIMO MAC in which two single-antenna users send their signals to a receiver with N antennas. The channel between the users and the receiver is flat and block fading with equal-length and synchronous coherence interval (across users) of T channel uses. That is, the channel vector $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$, $k = 1, 2$, remains unchanged during each block of length T and changes independently between blocks. The realizations of $\mathbf{h}_1$ and $\mathbf{h}_2$ are *unknown* to both the users and the receiver. The received signal during the coherence block b is

$$\mathbf{Y}[b] = \mathbf{h}_1[b]\,\mathbf{x}_1^{\mathsf{T}}[b] + \mathbf{h}_2[b]\,\mathbf{x}_2^{\mathsf{T}}[b] + \mathbf{Z}[b], \quad b = 1, 2, \ldots, \tag{3.1}$$

where $\mathbf{x}_1[b] \in \mathbb{C}^T$ and $\mathbf{x}_2[b] \in \mathbb{C}^T$ are the transmitted signals from user 1 and user 2, respectively, with the power constraint

$$\frac{1}{\nu_k}\sum_{b=1}^{\nu_k} \|\mathbf{x}_k[b]\|^2 \le PT, \quad k = 1, 2, \tag{3.2}$$

where ν_k is the number of blocks spanned by a channel codeword of user k, $k \in [K]$, and $\mathbf{Z}[b] \in \mathbb{C}^{N\times T}$ is the AWGN with IID $\mathcal{N}_{\mathbb{C}}(0, 1)$ entries. The parameter P is the SNR of the channel. In the remainder of the chapter, we omit the block index b whenever confusion is unlikely.

Since the channel is block memoryless, it is well known that a rate pair $(R_1(P), R_2(P))$ in bpcu is achievable at SNR P, i.e., lies within the capacity region $\mathcal{C}_{\text{Avg}}(P)$, for the MAC if and only if

$$\begin{cases} R_1 & \le \frac{1}{T} I(\mathbf{x}_1; \mathbf{Y} \,|\, \mathbf{x}_2), & \text{(3.3a)} \\ R_2 & \le \frac{1}{T} I(\mathbf{x}_2; \mathbf{Y} \,|\, \mathbf{x}_1), & \text{(3.3b)} \\ R_1 + R_2 & \le \frac{1}{T} I(\mathbf{x}_1, \mathbf{x}_2; \mathbf{Y}), & \text{(3.3c)} \end{cases}$$

for some input distribution subject to the average power constraint P (as the channel codeword length goes to infinity) [22]. Then, we say that (d_1, d_2) is an achievable DoF pair with

$$d_k := \lim_{P\to\infty} \frac{R_k(P)}{\log_2(P)}, \quad k = 1, 2. \tag{3.4}$$

The optimal DoF region $\mathcal{D}_{\text{Avg}}$ is defined as the set of all achievable DoF pairs.

We assume that the channel vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ are independent and drawn from a generic distribution satisfying the following conditions:[1]

$$h(\mathbf{h}_k) > -\infty, \quad \mathbb{E}\left[\|\mathbf{h}_k\|^2\right] < \infty, \quad k = 1, 2. \tag{3.5}$$

If the support of the input distribution is further bounded such that

$$\|\mathbf{x}_k\|^2 \le P, \quad k = 1, 2, \tag{3.6}$$

[1] Under the assumption (3.5), we assume w.l.o.g. that $\mathbb{E}\left[\|\mathbf{h}_k\|^2\right] = N, k = 1, 2$, for convenience.

then we say that the input satisfies the *peak power constraint* P. In this case, the capacity region and DoF region are denoted $\mathcal{C}_{\text{Peak}}(P)$ and $\mathcal{D}_{\text{Peak}}$, respectively. Since the peak power constraint implies the average power constraint, we have that

$$\mathcal{C}_{\text{Peak}}(P) \subseteq \mathcal{C}_{\text{Avg}}(P), \quad \mathcal{D}_{\text{Peak}}(P) \subseteq \mathcal{D}_{\text{Avg}}(P). \tag{3.7}$$

Lemma 3.1. *For any rate pair (R_1, R_2) achievable under the average power constraint P, for any $\beta > 1$, there exists (R_1', R_2') achievable under the peak power constraint P^β, such that*

$$R_k - R_k' = O(P^{1-\beta} \log P^\beta), \quad k = 1, 2. \tag{3.8}$$

In short, with a slight abuse of notation,

$$\mathcal{C}_{\text{Avg}}(P) \subseteq \mathcal{C}_{\text{Peak}}(P^\beta) + O(P^{1-\beta} \log P^\beta), \quad \forall\, \beta > 1. \tag{3.9}$$

Proof. The proof is provided in Appendix 3.A.1 □

Since the pre-log of the gap $P^{1-\beta} \log P^\beta$ is vanishing at high SNR for any $\beta > 1$, we have the DoF region

$$\mathcal{D}_{\text{Avg}}(P) \subseteq \mathcal{D}_{\text{Peak}}(P^\beta) \subseteq \mathcal{D}_{\text{Avg}}(P^\beta), \quad \forall\, \beta > 1. \tag{3.10}$$

Letting β arbitrarily close to 1, we conclude from (3.7) and (3.10) that using the peak power constraint instead of the average power constraint does not change the optimal DoF region. We therefore consider throughout this chapter the peak power constraint, which can simplify considerably the analysis.

3.3 Main Result

The main finding of this chapter is the optimal DoF region of the MAC described above, as stated in Theorem 3.1.

Theorem 3.1. *For the non-coherent MAC with two single-antenna transmitters and an N-antenna receiver in generic flat and block fading with coherence interval T, the optimal DoF region is characterized by*

$$d_1 + d_2 \leq 1 - \frac{1}{T}, \tag{3.11}$$

if $T \leq 2$ or $N = 1$, and

$$\begin{cases} \dfrac{d_1}{T-2} + d_2 & \leq 1 - \dfrac{1}{T}, \quad (3.12a) \\ d_1 + \dfrac{d_2}{T-2} & \leq 1 - \dfrac{1}{T}, \quad (3.12b) \end{cases}$$

otherwise.

Remark 3.1. *When $T \to \infty$, the optimal DoF region approaches the region in the coherent case given by*

$$\begin{cases} d_1 + d_2 \leq 1, & \text{if } N = 1, \\ \max\{d_1, d_2\} \leq 1, & \text{if } N > 1, \end{cases} \tag{3.13}$$

as shown in Fig. 3.1.

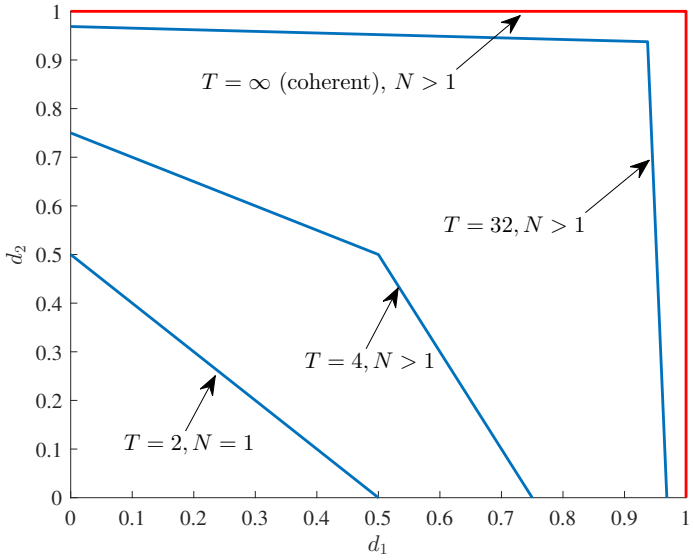


FIGURE 3.1: The optimal DoF region of the two-user SIMO MAC with N receive antennas in generic block fading with coherence interval T.

The case $T = 1$ (stationary fading) is trivial: zero DoF is achievable, even if two users cooperate [36].

If $T = 2$ or $N = 1$, the optimal DoF region is achieved with TDM between the users, noting that the active user can achieve $1 - \frac{1}{T}$ DoF by either a pilot-based scheme [31] or USTM [49], [53]. The tight outer bound follows by letting two users cooperate, then according to Theorem 2.1 in Chapter 2, the optimal total DoF is $1 - \frac{1}{T}$.

When $T \geq 3, N > 1$, the region is the convex hull of the origin and three points: $\left(1 - \frac{1}{T}, 0\right)$, $\left(0, 1 - \frac{1}{T}\right)$, and $\left(1 - \frac{2}{T}, 1 - \frac{2}{T}\right)$. The first two points are achieved by activating only one user. The third point is achieved with a pilot-based scheme: let two users send orthogonal pilots in the two arbitrary channel uses for the receiver to learn their channel, and send data in the remaining $T - 2$ channel uses. The region is then achieved with time sharing between these points. Another achievable scheme using a geometric approach can be found in [88]. There, motivated by the geometric structure of the problem, the transmitted signals $\mathbf{x}_1$ and $\mathbf{x}_2$ are drawn uniformly from quotient spaces of specified unitary groups.

It remains to show the tight outer bound for the case $T \geq 3, N > 1$. To this end, we use the duality approach as done for the single-user case in Section 2.4. The details are presented in the next section.

3.4 The Converse for the Case $T \geq 3, N > 1$

We are going to show that, when $T \geq 3, N > 1$, any achievable DoF pair (d_1, d_2) must satisfy (3.12a) and (3.12b).

3.4.1 The Case $T \geq N + 1 > 2$

Let us first consider the more straightforward case with $T \geq N + 1 > 2$. We first bound R_1 and R_2 using similar techniques as for the single-user case in Section 2.4, and then give the tight outer bound for the DoF region in the following steps.

Step 1: Output Rotation and Genie-Aided Bound

Given $\mathbf{x}_2$, the channel w.r.t. input $\mathbf{x}_1$ has equivalent noise $\mathbf{h}_2\mathbf{x}_2^{\mathsf{T}} + \mathbf{Z}$. Consider the following eigendecomposition

$$\mathbf{x}_2^*\mathbf{x}_2^{\mathsf{T}} = \mathbf{U}\ \mathrm{diag}(0, \ldots, 0, \|\mathbf{x}_2\|^2)\ \mathbf{U}^{\mathsf{H}}, \tag{3.14}$$

for some $T \times T$ unitary matrix $\mathbf{U}$. We consider the rotated output $\tilde{\mathbf{Y}} = \mathbf{Y}\mathbf{U} = \mathbf{h}_1\tilde{\mathbf{x}}_1^{\mathsf{T}} + \tilde{\mathbf{Z}}$, where $\tilde{\mathbf{x}}_1^{\mathsf{T}} = \mathbf{x}_1^{\mathsf{T}}\mathbf{U} = [\tilde{\mathrm{x}}_{11}\ \tilde{\mathrm{x}}_{12}\ \ldots\ \tilde{\mathrm{x}}_{1T}]$ and $\tilde{\mathbf{Z}} = (\mathbf{h}_2\mathbf{x}_2^{\mathsf{T}} + \mathbf{Z})\mathbf{U}$. Note that given $\mathbf{x}_2$, the first $T-1$ columns of the noise $\tilde{\mathbf{Z}}$ are IID Gaussian whereas the last column is stronger as the sum of $\mathbf{h}_2\|\mathbf{x}_2\|$ and a Gaussian noise vector. Thus, from (3.3a) and the fact that unitary rotation does not change the mutual information, we have that

$$TR_1 \leq I(\mathbf{x}_1; \mathbf{Y}|\mathbf{x}_2) = I(\tilde{\mathbf{x}}_1; \tilde{\mathbf{Y}}|\mathbf{x}_2), \tag{3.15}$$

for any achievable rate R_1 of user 1. Let us define the random variable v as the index of the strongest among the first $T-1$ elements of $\tilde{\mathbf{x}}_1$, namely,

$$\mathsf{v} = \arg\max_{i \in [T-1]} |\tilde{\mathrm{x}}_{1i}|^2. \tag{3.16}$$

Let a genie give v to the receiver. With the genie-aided bound, we have that

$$\begin{aligned} I(\tilde{\mathbf{x}}_1; \tilde{\mathbf{Y}}|\mathbf{x}_2) &\leq I(\tilde{\mathbf{x}}_1; \tilde{\mathbf{Y}}, \mathsf{v}|\mathbf{x}_2) && (3.17)\\ &= I(\tilde{\mathbf{x}}_1; \tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v}) + I(\tilde{\mathbf{x}}_1; \mathsf{v}|\mathbf{x}_2) && (3.18)\\ &\leq h(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v}) - h(\tilde{\mathbf{Y}}|\tilde{\mathbf{x}}_1, \mathbf{x}_2, \mathsf{v}) + H(\mathsf{v}) && (3.19)\\ &\leq h(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v}) - h(\tilde{\mathbf{Y}}|\tilde{\mathbf{x}}_1, \mathbf{x}_2) + \log_2(T-1). && (3.20) \end{aligned}$$

where the last inequality is because we have the Markov chain $\mathsf{v} \leftrightarrow (\tilde{\mathbf{x}}_1, \mathbf{x}_2) \leftrightarrow \tilde{\mathbf{Y}}$ and $H(\mathsf{v}) \leq \log_2(T-1)$.

Step 2: Bounding $h(\tilde{\mathbf{Y}}|\tilde{\mathbf{x}}_1, \mathbf{x}_2)$ and $h(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v})$

Given any realizations $\tilde{\boldsymbol{x}}_1$ of $\tilde{\mathbf{x}}_1$ and $\boldsymbol{x}_2$ of $\mathbf{x}_2$, we can apply Lemma 2.1 with $\mathbf{W} = [\mathbf{h}_1\ \mathbf{h}_2\ \mathbf{Z}]$ and $\boldsymbol{A} = [\tilde{\boldsymbol{x}}_1\ \boldsymbol{x}_2\ \boldsymbol{I}_T]^{\mathsf{T}}\mathbf{U}$ to obtain

$$\begin{aligned} h(\tilde{\mathbf{Y}}|\tilde{\mathbf{x}}_1, \mathbf{x}_2) &= N\mathbb{E}[\log_2 \det(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A})] + O(1) && (3.21)\\ &= N\mathbb{E}\left[\log_2\left((1+\|\mathbf{x}_2\|^2)\left(1 + \sum_{i=1}^{T-1}|\tilde{\mathrm{x}}_{1i}|^2\right) + |\tilde{\mathrm{x}}_{1T}|^2\right)\right] + O(1), && (3.22) \end{aligned}$$

where the last equality is obtained by applying $\tilde{\mathbf{x}}_1^{\mathsf{T}} = \mathbf{x}_1^{\mathsf{T}}\mathbf{U}$.

For $h(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v})$, we use the duality upper bound

$$\begin{aligned} h(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v}) &= \mathbb{E}\Big[-\log_2 p(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v})\Big] && (3.23)\\ &= \mathbb{E}\Big[-\log_2 q(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v})\Big] - \mathbb{E}_{\mathsf{v}}[D(p_{\tilde{\mathbf{Y}}|\mathbf{x}_2,\mathsf{v}} \| q_{\tilde{\mathbf{Y}}|\mathbf{x}_2,\mathsf{v}})] && (3.24)\\ &\leq \mathbb{E}\Big[-\log_2 q(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v})\Big], && (3.25) \end{aligned}$$

where the only difference from the single-user case in Section 2.4 is the presence of $\mathbf{x}_2$. Inspired by a pilot-based scheme with pilot transmission in channel use v and data

transmission in all other channel uses, we choose the auxiliary distribution $q_{\tilde{\mathbf{Y}}|\mathbf{x}_2,\mathsf{v}=v}$ as follows. Given $\mathsf{v} = v$, $v \leq T-1$, we let

$$\tilde{\mathbf{Y}}_{[v]} \sim \mathcal{R}(N, \boldsymbol{I}_N) \tag{3.26}$$

where the family of distributions $\mathcal{R}(N, \boldsymbol{A})$ was defined in Lemma 1.3 in Chapter 2, and given $\tilde{\mathbf{Y}}_{[v]}$, the other $\tilde{\mathbf{Y}}_{[i]}$'s are independent and follow

$$\tilde{\mathbf{Y}}_{[i]} \sim \mathcal{R}\Big(N, \Big(\boldsymbol{I}_N + \tilde{\mathbf{Y}}_{[v]}\tilde{\mathbf{Y}}_{[v]}^{\mathsf{H}}\Big)^{-\frac{1}{2}}\Big), \quad i \notin \{v, T\}, \tag{3.27}$$

$$\tilde{\mathbf{Y}}_{[T]} \sim \mathcal{R}\Big(N, \Big((1+\|\mathbf{x}_2\|^2)\boldsymbol{I}_N + \tilde{\mathbf{Y}}_{[v]}\tilde{\mathbf{Y}}_{[v]}^{\mathsf{H}}\Big)^{-\frac{1}{2}}\Big). \tag{3.28}$$

Proposition 3.1. *With the above choice of auxiliary output distributions, we obtain the following upper bound for* $\mathbb{E}\Big[-\log_2 q(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v})\Big]$*, and hence for* $h(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v})$*:*

$$\begin{aligned}\mathbb{E}\Big[-\log_2 q(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v})\Big] \leq\ & (N+T-2)\mathbb{E}\Big[\log_2(1+|\tilde{\mathsf{x}}_{1\mathsf{v}}|^2)\Big] \\ & + N\mathbb{E}\left[\sum_{i=1, i\neq V}^{T-1} \log_2\left(1+\frac{|\tilde{\mathsf{x}}_{1i}|^2}{1+|\tilde{\mathsf{x}}_{1\mathsf{v}}|^2}\right)\right] \\ & + N\mathbb{E}\Big[\log_2(1+\|\mathbf{x}_2\|^2)\Big] + \mathbb{E}\left[\log_2\left(1+\frac{|\tilde{\mathsf{x}}_{1\mathsf{v}}|^2}{1+\|\mathbf{x}_2\|^2}\right)\right] \\ & + N\mathbb{E}\left[\log_2\left(1+\frac{|\tilde{\mathsf{x}}_{1T}|^2}{1+\|\mathbf{x}_2\|^2+|\tilde{\mathsf{x}}_{1\mathsf{v}}|^2}\right)\right] + O(\log\log P). \end{aligned} \tag{3.29}$$

Proof. The proof is provided in Appendix 3.A.2. □

Step 3: Upper Bounds on R_1 and R_2

From (3.15), (3.20), (3.22) and (3.29), we have the bound for R_1

$$TR_1 \leq \mathbb{E}[f(\tilde{\mathbf{x}}_1, \mathbf{x}_2)] + O(\log\log P), \tag{3.30}$$

where

$$\begin{aligned} f(\tilde{\mathbf{x}}_1, \mathbf{x}_2) :=\ & (N+T-2)\log_2\left(1+\max_{i\in[T-1]}|\tilde{\mathsf{x}}_{1i}|^2\right) + \log_2\left(1+\frac{\max\limits_{i\in[T-1]}|\tilde{\mathsf{x}}_{1i}|^2}{1+\|\mathbf{x}_2\|^2}\right) \\ & + N\log_2\left(1+\frac{|\tilde{\mathsf{x}}_{1T}|^2}{1+\|\mathbf{x}_2\|^2+\max\limits_{i\in[T-1]}|\tilde{\mathsf{x}}_{1i}|^2}\right) - N\log_2\left(1+\sum_{i=1}^{T-1}|\tilde{\mathsf{x}}_{1i}|^2+\frac{|\tilde{\mathsf{x}}_{1T}|^2}{1+\|\mathbf{x}_2\|^2}\right). \end{aligned} \tag{3.31}$$

Following the exact same steps by swapping the users' role, we get that

$$TR_2 \leq \mathbb{E}[f(\tilde{\mathbf{x}}_2, \mathbf{x}_1)] + O(\log\log P), \tag{3.32}$$

where $\tilde{\mathbf{x}}_2 := \mathbf{x}_2\mathbf{U}_1$ with $\mathbf{U}_1$ obtained from the decomposition

$$\mathbf{x}_1^*\mathbf{x}_1^{\mathsf{T}} = \mathbf{U}_1\ \mathrm{diag}(0, \ldots, 0, \|\mathbf{x}_1\|^2)\ \mathbf{U}_1^{\mathsf{H}}. \tag{3.33}$$

It follows that, for any $\lambda_1, \lambda_2 \geq 0$, we have the following upper bound on the weighted sum rate

$$\lambda_1 R_1 + \lambda_2 R_2 \leq \frac{1}{T}\mathbb{E}[\lambda_1 f(\tilde{\mathbf{x}}_1, \mathbf{x}_2) + \lambda_2 f(\tilde{\mathbf{x}}_2, \mathbf{x}_1)] + O(\log\log P) \tag{3.34}$$

$$\leq \frac{1}{T}\sup_{\boldsymbol{x}_1, \boldsymbol{x}_2}[\lambda_1 f(\tilde{\boldsymbol{x}}_1, \boldsymbol{x}_2) + \lambda_2 f(\tilde{\boldsymbol{x}}_2, \boldsymbol{x}_1)] + O(\log\log P), \tag{3.35}$$

where the supremum is over all $\boldsymbol{x}_1, \boldsymbol{x}_2$ subject to the peak power constraints $\|\boldsymbol{x}_1\|^2 \leq P$ and $\|\boldsymbol{x}_2\|^2 \leq P$.

Step 4: DoF Upper Bounds

Since we are only interested in the pre-log at high SNR, it is without loss of optimality to let $\|\boldsymbol{x}_1\|^2 = \Theta(P^{\eta_1}), \|\boldsymbol{x}_2\|^2 = \Theta(P^{\eta_2})$ for some $\eta_1, \eta_2 \leq 1$. In addition, we assume that

$$\max_{i\in[T-1]} |\tilde{x}_{1i}|^2 = \Theta(P^{\bar{\eta}_1}), \quad |\tilde{x}_{1T}|^2 = \Theta(P^{\eta_{1T}}), \tag{3.36}$$

$$\max_{i\in[T-1]} |\tilde{x}_{2i}|^2 = \Theta(P^{\bar{\eta}_2}), \quad |\tilde{x}_{2T}|^2 = \Theta(P^{\eta_{2T}}). \tag{3.37}$$

Hence, at high SNR, $\eta_1 = \max\{\bar{\eta}_1, \eta_{1T}\}$, $\eta_2 = \max\{\bar{\eta}_2, \eta_{2T}\}$. From (3.31) and (3.35), we have the weighted sum DoF bound

$$\begin{aligned}\lambda_1 d_1 + \lambda_2 d_2 \leq\ & \lambda_1 \frac{N+T-2}{T}\bar{\eta}_1 + \lambda_1 \frac{1}{T}(\bar{\eta}_1 - \eta_2)^+ \\ & + \lambda_1 \frac{N}{T}(\eta_{1T} - \max\{\bar{\eta}_1, \eta_2\})^+ - \lambda_1 \frac{N}{T}\max\{\bar{\eta}_1, \eta_{1T} - \eta_2\} \\ & + \lambda_2 \frac{N+T-2}{T}\bar{\eta}_2 + \lambda_2 \frac{1}{T}(\bar{\eta}_2 - \eta_1)^+ \\ & + \lambda_2 \frac{N}{T}(\eta_{2T} - \max\{\bar{\eta}_2, \eta_1\})^+ - \lambda_2 \frac{N}{T}\max\{\bar{\eta}_2, \eta_{2T} - \eta_1\},\end{aligned} \tag{3.38}$$

subject to the constraints $\bar{\eta}_1, \eta_{1T} \leq 1$ and $\bar{\eta}_2, \eta_{2T} \leq 1$. Taking (λ_1, λ_2) as $\left(1, \frac{1}{T-2}\right)$ or $\left(\frac{1}{T-2}, 1\right)$, we can verify that, when $3 \leq N+1 \leq T$, (3.12a) and (3.12b) hold for all (d_1, d_2) satisfying (3.38). Thus the optimal DoF region is characterized for this case.

3.4.2 The Case $3 \leq T \leq N$

When $T \leq N$, the above choice of auxiliary output distribution is not sufficient for a tight DoF outer bound. To see this, let us take $(\lambda_1, \lambda_2) = \left(1, \frac{1}{T-2}\right)$, then if $\bar{\eta}_1 + \eta_2 \geq \eta_{1T} = 1$ and $\eta_2 = \bar{\eta}_1$, (3.38) becomes

$$d_1 + \frac{d_2}{T-2} \leq \frac{T-1}{T}\bar{\eta}_1 + \frac{N}{T}(\eta_{1T} - \bar{\eta}_1), \tag{3.39}$$

which is loose since the right-hand side is larger than $1 - \frac{1}{T}$ whenever $N \geq T$. Generally, the bound (3.38) can be loose when $\eta_{1T} > \max\{\bar{\eta}_1, \eta_2\}$ or $\eta_{2T} > \max\{\bar{\eta}_2, \eta_1\}$. To account for such scenarios, we ought to refine our choice of auxiliary output distribution for the duality upper bound. First, given x_2, we define a pair of random variables (v, u) as

$$\mathsf{v} := \arg\max_{i\in[T]} \frac{|\tilde{\mathsf{x}}_{1i}|^2}{\sigma_i^2}, \tag{3.40}$$

where $\sigma_i^2 = 1, \forall i < T$ and $\sigma_T^2 = 1 + \|\mathbf{x}_2\|^2$, and

$$\mathsf{u} := \begin{cases} 1, & \text{if } |\tilde{\mathsf{x}}_{1T}|^2 \geq \max\left\{ \max\limits_{i\in[T-1]} |\tilde{\mathsf{x}}_{1i}|^2, 1 + \|\mathbf{x}_2\|^2 \right\}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.41}$$

Thus, $\tilde{\mathsf{x}}_{1\mathsf{v}}$ is the input entry with the largest instantaneous SNR, and u determines a specific configuration of input entry powers in which the choice of auxiliary output distribution in the previous case possibly fails. Then similarly as for the case $T \geq N+1$, with output rotation, genie-aided bound, and duality upper bound, we have that

$$\begin{aligned} TR_1 &\leq I(\tilde{\mathbf{x}}_1; \tilde{\mathbf{Y}} | \mathbf{x}_2) && (3.42) \\ &\leq \mathbb{E}\Big[-\log_2 q(\tilde{\mathbf{Y}} | \mathbf{x}_2, \mathsf{v}, \mathsf{u}) \Big] - h(\tilde{\mathbf{Y}} | \tilde{\mathbf{x}}_1, \mathbf{x}_2) + \log_2(T) + 1, && (3.43) \end{aligned}$$

where $h(\tilde{\mathbf{Y}}|\tilde{\mathbf{x}}_1, \mathbf{x}_2)$ was calculated in (3.22). For $\mathbb{E}\Big[-\log_2 q(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v}, \mathsf{u}) \Big]$, we choose the auxiliary PDF $q(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v}, \mathsf{u})$ as follows. Given $\mathsf{v} = v$ and $\mathsf{u} = u$, we consider two scenarios:

- If $v = T$ or $\{v < T, u = 0\}$, we let $\tilde{\mathbf{Y}}_{[v]} \sim \mathcal{R}(N, \boldsymbol{I}_N)$ and conditioned on $\tilde{\mathbf{Y}}_{[v]}$, the other $\tilde{\mathbf{Y}}_{[i]}$'s are independent and follow

$$\tilde{\mathbf{Y}}_{[i]} \sim \mathcal{R}\left(N, \left(\sigma_i^2 \boldsymbol{I}_N + \frac{\tilde{\mathbf{Y}}_{[v]} \tilde{\mathbf{Y}}_{[v]}^{\mathsf{H}}}{\sigma_v^2} \right)^{-\frac{1}{2}} \right), \quad i \neq v. \tag{3.44}$$

  This choice is inspired by a pilot-based scheme in which the input symbol with strongest SNR is used as pilot. After some manipulations similar as for Proposition 3.1, we get the bounds

$$\begin{aligned} &\mathbb{E}\Big[-\log_2 q(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v} = v < T, \mathsf{u} = 0) \Big] \\ &\leq (N + T - 2)\mathbb{E}\Big[\log_2(1 + |\tilde{\mathsf{x}}_{1v}|^2) \Big] + N\mathbb{E}\Big[\log_2(1 + \|\mathbf{x}_2\|^2) \Big] \\ &\quad + \mathbb{E}\left[\log_2\left(1 + \frac{|\tilde{\mathsf{x}}_{1v}|^2}{1 + \|\mathbf{x}_2\|^2} \right) \right] + O(\log\log P), \end{aligned} \tag{3.45}$$

  and

$$\begin{aligned} \mathbb{E}\Big[-\log_2 q(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v} = T) \Big] &\leq N\mathbb{E}\Big[\log_2(1 + \|\mathbf{x}_2\|^2 + |\tilde{\mathsf{x}}_{1T}|^2) \Big] \\ &\quad + (T-1)\mathbb{E}\left[\log_2\left(1 + \frac{|\tilde{\mathsf{x}}_{1T}|^2}{1 + \|\mathbf{x}_2\|^2} \right) \right] + O(\log\log P). \end{aligned} \tag{3.46}$$

- If $\{v < T, u = 1\}$, we let $\tilde{\mathbf{Y}}_{[v]} \sim \mathcal{R}(N, \boldsymbol{I}_N)$ and conditioned on $\tilde{\mathbf{Y}}_{[v]}$, the other $\tilde{\mathbf{Y}}_{[i]}$'s are independent with

$$\tilde{\mathbf{Y}}_{[i]} \sim \mathcal{R}\Big(N, \Big(\boldsymbol{I}_N + \tilde{\mathbf{Y}}_{[v]} \tilde{\mathbf{Y}}_{[v]}^{\mathsf{H}} \Big)^{-\frac{1}{2}} \Big), \quad i \notin \{v, T\}, \tag{3.47}$$

$$\tilde{\mathbf{Y}}_{[T]} \sim \mathcal{R}\left(N, \left((1 + \|\mathbf{x}_2\|^2) \boldsymbol{I}_N + \frac{P}{\|\tilde{\mathbf{Y}}_{[v]}\|^2} \tilde{\mathbf{Y}}_{[v]} \tilde{\mathbf{Y}}_{[v]}^{\mathsf{H}} \right)^{-\frac{1}{2}} \right), \tag{3.48}$$

  where the only difference from (3.44) is the presence of the factor $\frac{P}{\|\tilde{\mathbf{Y}}_{[v]}\|^2}$. This factor is added to account for the fact that if $u = 1$, then $|\tilde{\mathsf{x}}_{1v}|^2 < |\tilde{\mathsf{x}}_{1T}|^2$, which can make

the power of $\tilde{\mathbf{Y}}_{[v]}$ inferior to that of $\tilde{\mathbf{Y}}_{[T]}$. In this case, after some manipulations similar as for Proposition 3.1, we have the bound

$$\begin{aligned}
&\mathbb{E}\Big[-\log q(\tilde{\mathbf{Y}}|\mathbf{x}_2, \mathsf{v}=v<T, \mathsf{u}=1)\Big] \\
&\leq (N+T-2)\mathbb{E}\Big[\log(1+|\tilde{\mathrm{x}}_{1v}|^2)\Big] + N\mathbb{E}\left[\log\frac{1+\|\mathbf{x}_2\|^2+|\tilde{\mathrm{x}}_{1T}|^2}{1+\|\mathbf{x}_2\|^2+P}\right] \\
&\quad + N\mathbb{E}\Big[\log(1+\|\mathbf{x}_2\|^2)\Big] + \mathbb{E}\left[\log\left(1+\frac{P}{1+\|\mathbf{x}_2\|^2}\right)\right] + O(\log\log P). \qquad (3.49)
\end{aligned}$$

The bounds (3.45), (3.46), and (3.49), together with (3.22), give us the upper bound on R_1 as

$$TR_1 \leq \mathbb{E}[g(\tilde{\mathbf{x}}_1, \mathbf{x}_2)] + O(\log\log P), \qquad (3.50)$$

where

$$g(\tilde{\mathbf{x}}_1, \mathbf{x}_2) :=$$

$$\begin{cases}
(T-2)\log_2\left(1+\max\limits_{i\in[T-1]}|\tilde{\mathrm{x}}_{1i}|^2\right) + \log_2\left(1+\frac{\max\limits_{i\in[T-1]}|\tilde{\mathrm{x}}_{1i}|^2}{1+\|\mathbf{x}_2\|^2}\right), \\
\qquad \text{if } \frac{|\tilde{\mathrm{x}}_{1T}|^2}{1+\|\mathbf{x}_2\|^2} < \max\limits_{i\in[T-1]}|\tilde{\mathrm{x}}_{1i}|^2 \text{ and } |\tilde{\mathrm{x}}_{1T}|^2 \leq \max\left\{\max\limits_{i\in[T-1]}|\tilde{\mathrm{x}}_{1i}|^2, 1+\|\mathbf{x}_2\|^2\right\}, \\
(T-2)\log_2\left(1+\max\limits_{i\in[T-1]}|\tilde{\mathrm{x}}_{1i}|^2\right) + N\log_2\left(\frac{1+\|\mathbf{x}_2\|^2+|\tilde{\mathrm{x}}_{1T}|^2}{1+\|\mathbf{x}_2\|^2+P}\right) + \log_2\left(1+\frac{P}{1+\|\mathbf{x}_2\|^2}\right), \\
\qquad \text{if } \frac{|\tilde{\mathrm{x}}_{1T}|^2}{1+\|\mathbf{x}_2\|^2} < \max\limits_{i\in[T-1]}|\tilde{\mathrm{x}}_{1i}|^2 \text{ and } |\tilde{\mathrm{x}}_{1T}|^2 > \max\left\{\max\limits_{i\in[T-1]}|\tilde{\mathrm{x}}_{1i}|^2, 1+\|\mathbf{x}_2\|^2\right\}, \\
(T-1)\log_2\left(1+\frac{|\tilde{\mathrm{x}}_{1T}|^2}{1+\|\mathbf{x}_2\|^2}\right), \\
\qquad \text{if } \frac{|\tilde{\mathrm{x}}_{1T}|^2}{1+\|\mathbf{x}_2\|^2} > \max\limits_{i\in[T-1]}|\tilde{\mathrm{x}}_{1i}|^2,
\end{cases} \qquad (3.51)$$

and the similar bound for R_2

$$TR_2 \leq \mathbb{E}[g(\tilde{\mathbf{x}}_2, \mathbf{x}_1)] + O(\log\log P). \qquad (3.52)$$

The rest of the proof follows from a similar weighted sum bound for the rates and the DoFs as done in the previous case.

3.5 Closing Remarks

In this chapter, we have proposed a new tight outer bound on the DoF region of the two-user non-coherent SIMO MAC with generic block fading. The outer bound region coincides with the inner bound region achieved by a simple pilot-based scheme. The resulting optimal DoF region is given in Table 3.1, together with existing results on the DoF/capacity limits of the non-coherent MAC.

3.A Appendices

3.A.1 Proof of Lemma 3.1

We prove the lemma by construction. Consider a rate pair (R_1, R_2) achievable with some input PDF $p_{\mathbf{x}_1}(.)$ and $p_{\mathbf{x}_2}(.)$ satisfying the *average* power constraints P. Let us define a

TABLE 3.1: The DoF/capacity region of the non-coherent MAC in block fading with coherence interval T and N receive antennas (the gray-colored cell contains the result of this chapter)

<table>
<tr><th colspan="2">Assumption</th><th>Two users</th><th>K users</th></tr>
<tr><td rowspan="2">Generic fading: optimal DoF region</td><td>SIMO MAC</td><td>$\begin{cases} d_1+d_2 \le 1-\frac{1}{T}, \\ \quad \text{if } T\le 2 \text{ or } N=1 \\ \frac{d_1}{T-2}+d_2 \le 1-\frac{1}{T}, d_1+\frac{d_2}{T-2} \le 1-\frac{1}{T}, \\ \quad \text{otherwise} \end{cases}$</td><td rowspan="2">Unknown</td></tr>
<tr><td>MIMO MAC</td><td>Unknown</td></tr>
<tr><td rowspan="2">IID Rayleigh fading</td><td>DoF region</td><td>Achievability: (1.52) [88]</td><td>Achievability: (1.53) [37]
Outer bound: (1.54) [37]</td></tr>
<tr><td>Sum capacity</td><td colspan="2">• Fast fading: sum capacity of the SIMO MAC with $K>1$ users equals capacity with 1 user [84]
• Block fading: bounds on the sum capacity for the SIMO MAC [85] and MIMO MAC [87]</td></tr>
<tr><td colspan="2">IID Rician fast fading</td><td colspan="2">Fading number for the sum capacity of the MISO MAC: [86]</td></tr>
</table>

new input $\underline{\mathsf{x}}_k$ with domain $\{\boldsymbol{x}\in\mathbb{C}^T, \|\boldsymbol{x}\|^2 \le P^\beta\}$ and PDF

$$p_{\underline{\mathsf{x}}_k}(\boldsymbol{x}) = \begin{cases} \frac{p_{\mathsf{x}_k}(\boldsymbol{x})}{\mathbb{P}(\|\mathsf{x}_k\|^2\le P^\beta)}, & \text{if } \|\boldsymbol{x}\|^2 \le P^\beta, \\ 0, & \text{if } \|\boldsymbol{x}\|^2 > P^\beta, \end{cases} \tag{3.53}$$

for $k=1,2$, with $\beta>1$. Then the inputs $\underline{\mathsf{x}}_1$ and $\underline{\mathsf{x}}_2$ satisfy the *peak* power constraint P^β. We define a random variable v as

$$\mathsf{v} = \begin{cases} 0, & \text{if } \|\mathbf{x}_1\|^2 \le P^\beta \text{ and } \|\mathbf{x}_2\|^2 \le P^\beta \ , \\ 1, & \text{if } \|\mathbf{x}_1\|^2 \le P^\beta \text{ and } \|\mathbf{x}_2\|^2 > P^\beta, \\ 2, & \text{if } \|\mathbf{x}_1\|^2 > P^\beta. \end{cases} \tag{3.54}$$

Then, $\{\mathbf{x}_k \,|\, \mathsf{v}=0\} \sim p_{\underline{\mathsf{x}}_k}$. By Markov's inequality,

$$\mathbb{P}(\|\mathbf{x}_k\|^2 > P^\beta) \le \frac{\mathbb{E}[\|\mathbf{x}_k\|^2]}{P^\beta} \le TP^{1-\beta}, \quad k=1,2, \tag{3.55}$$

then

$$\mathbb{P}(\mathsf{v}=1) = \mathbb{P}(\|\mathbf{x}_1\|^2 \le P^\beta)\mathbb{P}(\|\mathbf{x}_2\|^2 > P^\beta) \le \mathbb{P}(\|\mathbf{x}_2\|^2 > P^\beta) \le TP^{1-\beta}, \tag{3.56}$$

$$\mathbb{P}(\mathsf{v}=2) = \mathbb{P}(\|\mathbf{x}_1\|^2 > P^\beta) \le TP^{1-\beta}. \tag{3.57}$$

Assume that a genie gives v to the receiver:

$$\begin{aligned} TR_1 &\le I(\mathbf{x}_1;\mathbf{Y}|\mathbf{x}_2) && (3.58)\\ &\le I(\mathbf{x}_1;\mathbf{Y},\mathsf{v}|\mathbf{x}_2) && (3.59)\\ &= I(\mathbf{x}_1;\mathbf{Y}|\mathbf{x}_2,\mathsf{v}) + I(\mathbf{x}_1;\mathsf{v}|\mathbf{x}_2) && (3.60)\\ &\le \mathbb{P}(\mathsf{v}=0)I(\mathbf{x}_1;\mathbf{Y}|\mathbf{x}_2,\mathsf{v}=0) + \mathbb{P}(\mathsf{v}=1)I(\mathbf{x}_1;\mathbf{Y}|\mathbf{x}_2,\mathsf{v}=1) \\ &\quad + \mathbb{P}(\mathsf{v}=2)I(\mathbf{x}_1;\mathbf{Y}|\mathbf{x}_2,\mathsf{v}=2) + \log_2 3, && (3.61)\\ &\le I(\underline{\mathbf{x}}_1;\mathbf{Y}|\underline{\mathbf{x}}_2) + \mathbb{P}(\mathsf{v}=1)I(\mathbf{x}_1;\mathbf{Y}|\mathbf{x}_2,\mathsf{v}=1) + \mathbb{P}(\mathsf{v}=2)I(\mathbf{x}_1;\mathbf{Y}|\mathbf{x}_2,\mathsf{v}=2) + \log_2 3, \\ & && (3.62) \end{aligned}$$

where (3.61) is due to $I(\mathbf{x}_1;\mathsf{v}|\mathbf{x}_2) \le H(\mathsf{v}) \le \log_2(3)$ bits, and the last inequality is because given $\mathsf{v}=0$, $\mathbf{x}_k$ is identically distributed to $\underline{\mathbf{x}}_k, k=1,2$. Next, since removing noise and giving CSI increase the rate, for $v \in \{1,2\}$,

$$\begin{aligned}
&\mathbb{P}(\mathsf{v}=v) I(\mathbf{x}_1; \mathbf{Y} \,|\, \mathbf{x}_2, \mathsf{v}=v) \\
&= \mathbb{P}(\mathsf{v}=v) I(\mathbf{x}_1; \mathbf{H}_1\mathbf{x}_1^\mathsf{T} + \mathbf{H}_2\mathbf{x}_2^\mathsf{T} + \mathbf{Z} \,|\, \mathbf{x}_2, \mathsf{v}=v) && (3.63)\\
&\le \mathbb{P}(\mathsf{v}=v) I(\mathbf{x}_1; \mathbf{H}_1\mathbf{x}_1^\mathsf{T} + \mathbf{Z} \,|\, \mathbf{H}_1, \mathsf{v}=v) && (3.64)\\
&\le N\mathbb{P}(\mathsf{v}=v) \log_2\Big(1 + \mathbb{E}\Big[\|\mathbf{x}_1\|^2 | \mathsf{v}=v\Big]\Big). && (3.65)
\end{aligned}$$

With $v=1$, $\|\mathbf{x}_1\|^2 \le P^\beta$, thus

$$\begin{aligned}
\mathbb{P}(\mathsf{v}=1) \log_2\Big(1 + \mathbb{E}\Big[\|\mathbf{x}_1\|^2 \Big| \mathsf{v}=1\Big]\Big) &\le TP^{1-\beta}\log_2(1+P^\beta) && (3.66)\\
&= O(P^{1-\beta}\log P^\beta). && (3.67)
\end{aligned}$$

With $v=2$, $\|\mathbf{x}_1\|^2 > P^\beta$, we have that

$$\begin{aligned}
\mathbb{E}\Big[\|\mathbf{x}_1\|^2 \Big| \mathsf{v}=2\Big] &= \int_{P^\beta}^{\infty} x \frac{p_{\|\mathbf{x}_1\|^2}(x)}{\mathbb{P}(\|\mathbf{x}_1\|^2 > P^\beta)}\,\mathrm{d}x && (3.68)\\
&\le \frac{\int_0^\infty x p_{\|\mathbf{x}_1\|^2}(x) dx}{\mathbb{P}(\|\mathbf{x}_1\|^2 > P^\beta)} && (3.69)\\
&\le \frac{P}{\mathbb{P}(\|\mathbf{x}_1\|^2 > P^\beta)}. && (3.70)
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{P}(\mathsf{v}=2) \log_2\Big(1 + \mathbb{E}\Big[\|\mathbf{x}_1\|^2 \Big| \mathsf{v}=2\Big]\Big) &\le \mathbb{P}(\|\mathbf{x}_1\|^2 > P^\beta) \log_2\left(1 + \frac{P}{\mathbb{P}(\|\mathbf{x}_1\|^2 > P^\beta)}\right) && (3.71)\\
&\le TP^{1-\beta}\log_2\left(1 + \frac{P}{TP^{1-\beta}}\right) && (3.72)\\
&= O(P^{1-\beta}\log P^\beta), && (3.73)
\end{aligned}$$

where the second inequality is because the function $x\log_2(1+1/x)$ is monotonically increasing in x for $x \ge 0$ and $0 \le \mathbb{P}(\|\mathbf{x}_1\|^2 > P^\beta) \le TP^{1-\beta}$. Plugging (3.67) and (3.73) into (3.65), then (3.65) into (3.62) yields

$$TR_1 \le I(\underline{\mathbf{x}}_1; \mathbf{Y} | \underline{\mathbf{x}}_2) + O(P^{1-\beta}\log P^\beta). \tag{3.74}$$

Following the same steps by swapping the users' roles, we get the bound for R_2

$$TR_2 \le I(\underline{\mathbf{x}}_2; \mathbf{Y} | \underline{\mathbf{x}}_1) + O(P^{1-\beta}\log P^\beta). \tag{3.75}$$

Using similar techniques, we can also show that

$$T(R_1+R_2) \le I(\underline{\mathbf{x}}_1, \underline{\mathbf{x}}_2; \mathbf{Y}) + O(P^{1-\beta}\log P^\beta). \tag{3.76}$$

Therefore, there exists (R_1', R_2') satisfying

$$\begin{cases}
R_1' & \le \dfrac{1}{T} I(\underline{\mathbf{x}}_1; \mathbf{Y} \,|\, \underline{\mathbf{x}}_2), && (3.77a)\\
R_2' & \le \dfrac{1}{T} I(\underline{\mathbf{x}}_2; \mathbf{Y} \,|\, \underline{\mathbf{x}}_1), && (3.77b)\\
R_1' + R_2' & \le \dfrac{1}{T} I(\underline{\mathbf{x}}_1, \underline{\mathbf{x}}_2; \mathbf{Y}), && (3.77c)
\end{cases}$$

i.e., achievable with the constructed inputs $\underline{\mathbf{x}}_1$ and $\underline{\mathbf{x}}_2$ satisfying the peak power constraint P^β, such that (3.8) holds. This concludes the proof.

3.A.2 Proof of Proposition 3.1

We obtain from Lemma 1.3 that

$$\begin{aligned}
&\mathbb{E}\Big[-\log(q(\tilde{\mathbf{Y}}|\mathbf{x}_2,\mathsf{v}=v)\Big] \\
&= N\mathbb{E}\Big[\log\|\tilde{\mathbf{Y}}_{[v]}\|^2\Big] + \sum_{i=1,i\neq v}^{T-1}\mathbb{E}\bigg[\log\det\Big(\boldsymbol{I}_N+\tilde{\mathbf{Y}}_{[v]}\tilde{\mathbf{Y}}_{[v]}^{\mathsf{H}}\Big)+N\log\Big\|\Big(\boldsymbol{I}_N+\tilde{\mathbf{Y}}_{[v]}\tilde{\mathbf{Y}}_{[v]}^{\mathsf{H}}\Big)^{-\frac{1}{2}}\tilde{\mathbf{Y}}_{[i]}\Big\|^2\bigg] \\
&\quad+\mathbb{E}\bigg[\log\det\Big((1+\|\mathbf{x}_2\|^2)\boldsymbol{I}_N+\tilde{\mathbf{Y}}_{[v]}\tilde{\mathbf{Y}}_{[v]}^{\mathsf{H}}\Big)+N\log\Big\|\Big((1+\|\mathbf{x}_2\|^2)\boldsymbol{I}_N+\tilde{\mathbf{Y}}_{[v]}\tilde{\mathbf{Y}}_{[v]}^{\mathsf{H}}\Big)^{-\frac{1}{2}}\tilde{\mathbf{Y}}_{[T]}\Big\|^2\bigg] \\
&\quad+O(\log\log P) \qquad (3.78)\\
&\leq N\mathbb{E}\Big[\log(1+|\tilde{\mathsf{x}}_{1v}|^2)\Big]+\sum_{i=1,i\neq v}^{T}B_i+O(\log\log P), \qquad (3.79)
\end{aligned}$$

where

$$B_i := \mathbb{E}\left[\log\Big(1+\|\tilde{\mathbf{Y}}_{[v]}\|^2\Big)+N\log\left(\|\tilde{\mathbf{Y}}_{[i]}\|^2-\frac{|\tilde{\mathbf{Y}}_{[i]}^{\mathsf{H}}\tilde{\mathbf{Y}}_{[v]}|^2}{1+\|\tilde{\mathbf{Y}}_{[v]}\|^2}\right)\right],\quad i\notin\{v,T\}, \qquad (3.80)$$

$$\begin{aligned}
B_T := \mathbb{E}\Bigg[&\log\left((1+\|\mathbf{x}_2\|^2)^N\left(1+\frac{\|\tilde{\mathbf{Y}}_{[v]}\|^2}{1+\|\mathbf{x}_2\|^2}\right)\right)\\
&+N\log\left(\frac{1}{1+\|\mathbf{x}_2\|^2}\left(\|\tilde{\mathbf{Y}}_{[T]}\|^2-\frac{|\tilde{\mathbf{Y}}_{[T]}^{\mathsf{H}}\tilde{\mathbf{Y}}_{[v]}|^2}{1+\|\mathbf{x}_2\|^2+\|\tilde{\mathbf{Y}}_{[v]}\|^2}\right)\right)\Bigg]. \qquad (3.81)
\end{aligned}$$

By expanding $\tilde{\mathbf{Y}}_{[1]},\ldots,\tilde{\mathbf{Y}}_{[T]}$, we get that, given $\mathbf{x}_1$ and $\mathbf{x}_2$,

$$\mathbb{E}_{\mathbf{H}_1,\mathbf{Z}}\Big[\|\tilde{\mathbf{Y}}_{[v]}\|^2\|\tilde{\mathbf{Y}}_{[i]}\|^2-|\tilde{\mathbf{Y}}_{[i]}^{\mathsf{H}}\tilde{\mathbf{Y}}_{[v]}|^2\Big]=(N^2-N)\Big(1+|\tilde{\mathsf{x}}_{1v}|^2+|\tilde{\mathsf{x}}_{1i}|^2\Big),\quad i\notin\{v,T\}, \qquad (3.82)$$

$$\begin{aligned}
\mathbb{E}_{\mathbf{H}_1,\mathbf{Z}}\Big[\|\tilde{\mathbf{Y}}_{[v]}\|^2\|\tilde{\mathbf{Y}}_{[T]}\|^2-|\tilde{\mathbf{Y}}_{[T]}^{\mathsf{H}}\tilde{\mathbf{Y}}_{[v]}|^2\Big]&=(N^2-N)\Big((1+\|\mathbf{x}_2\|^2)(1+|\tilde{\mathsf{x}}_{1v}|^2)+|\tilde{\mathsf{x}}_{1T}|^2\Big)\\
&\leq(N^2-N)(1+\|\mathbf{x}_2\|^2)(1+|\tilde{\mathsf{x}}_{1v}|^2+|\tilde{\mathsf{x}}_{1T}|^2). \qquad (3.83)
\end{aligned}$$

Then, applying repeatedly Lemma 2.2 (with $\beta=1$ and α arbitrarily close to 1) and Jensen's inequality, we get that

$$\begin{aligned}
B_i&=\mathbb{E}\Big[-(N-1)\log\Big(1+\|\tilde{\mathbf{Y}}_{[v]}\|^2\Big)\Big]\\
&\quad+N\mathbb{E}\Big[\log\Big(\|\tilde{\mathbf{Y}}_{[i]}\|^2+\|\tilde{\mathbf{Y}}_{[i]}\|^2\|\tilde{\mathbf{Y}}_{[v]}\|^2-|\tilde{\mathbf{Y}}_{[i]}^{\mathsf{H}}\tilde{\mathbf{Y}}_{[v]}|^2\Big)\Big]\\
&\leq\mathbb{E}\Big[-(N-1)\log\Big(1+N+N|\tilde{\mathsf{x}}_{1v}|^2\Big)\Big] \qquad (3.84)\\
&\quad+N\mathbb{E}\Big[\log\Big(N^2(1+|\tilde{\mathsf{x}}_{1i}|^2)+(N^2-N)|\tilde{\mathsf{x}}_{1v}|^2\Big)\Big]+O(1) \qquad (3.85)\\
&=\mathbb{E}\Big[\log\Big(1+|\tilde{\mathsf{x}}_{1v}|^2\Big)\Big]+N\mathbb{E}\left[\log\left(1+\frac{|\tilde{\mathsf{x}}_{1i}|^2}{1+|\tilde{\mathsf{x}}_{1v}|^2}\right)\right]+O(1), \qquad (3.86)
\end{aligned}$$

for $i \notin \{v, T\}$. Similarly, by applying repeatedly Lemma 2.2 (with $\beta = 1 + \|\mathbf{x}_2\|^2$ and α arbitrarily close to 1) and Jensen's inequality,

$$
\begin{aligned}
B_T &= N\mathbb{E}\Big[\log\Big(1 + \|\mathbf{x}_2\|^2\Big)\Big] + \mathbb{E}\left[\log\left(1 + \frac{\|\tilde{\mathbf{Y}}_v\|^2}{1 + \|\mathbf{x}_2\|^2}\right)\right] \\
&\quad + N\mathbb{E}\left[\log\left(\|\tilde{\mathbf{Y}}_{[T]}\|^2 + \frac{\|\tilde{\mathbf{Y}}_{[v]}\|^2\|\tilde{\mathbf{Y}}_{[T]}\|^2 - |\tilde{\mathbf{Y}}_{[T]}^{\mathsf{H}}\tilde{\mathbf{Y}}_{[v]}|^2}{1 + \|\mathbf{x}_2\|^2}\right)\right] \\
&\quad - N\mathbb{E}\Big[\log\Big(1 + \|\mathbf{x}_2\|^2 + \|\tilde{\mathbf{Y}}_v\|^2\Big)\Big] \qquad (3.87) \\
&\leq N\mathbb{E}\Big[\log\Big(1 + \|\mathbf{x}_2\|^2\Big)\Big] + \mathbb{E}\left[\log\left(1 + \frac{N + N|\tilde{\mathsf{x}}_{1v}|^2}{1 + \|\mathbf{x}_2\|^2}\right)\right] \\
&\quad + N\mathbb{E}\Big[\log\Big(N(1 + \|\mathbf{x}_2\|^2) + (N^2 - N)(1 + |\tilde{\mathsf{x}}_{1v}|^2) + N^2|\tilde{\mathsf{x}}_{1T}|^2\Big)\Big] \\
&\quad - N\mathbb{E}\Big[\log\Big(1 + \|\mathbf{x}_2\|^2 + N + N|\tilde{\mathsf{x}}_{1v}|^2\Big)\Big] + O(1) \qquad (3.88) \\
&= N\mathbb{E}\Big[\log\Big(1 + \|\mathbf{x}_2\|^2\Big)\Big] + \mathbb{E}\left[\log\left(1 + \frac{|\tilde{\mathsf{x}}_{1v}|^2}{1 + \|\mathbf{x}_2\|^2}\right)\right] \\
&\quad + N\mathbb{E}\left[\log\left(1 + \frac{|\tilde{\mathsf{x}}_{1T}|^2}{1 + \|\mathbf{x}_2\|^2 + |\tilde{\mathsf{x}}_{1v}|^2}\right)\right] + O(1) \qquad (3.89)
\end{aligned}
$$

Note that we have applied Lemma 2.2 with $\mathsf{u} = \|\mathbf{Y}_{[v]}\|^2$ in (3.85) and $\mathsf{u} = \|\tilde{\mathbf{Y}}_{[v]}\|^2$ in (3.88). The conditions $\mathbb{E}[\mathsf{u}] < \infty$ and $h(\mathsf{u}/\mathbb{E}[\mathsf{u}]) > -\infty$ can be verified as done in Appendix 2.A.4.

Plugging (3.85) and (3.89) into (3.79) then taking expectation over v, we obtain (3.29), which concludes the proof.

Chapter 4

The Spatially Correlated MIMO Broadcast Channel

In a MIMO BC, the difference in spatial transmit correlation matrices of different users is called transmit correlation diversity*. This diversity was conceived for channels in which transmit correlation matrices have mutually exclusive eigenspaces, allowing noninterfering pilot and data transmission. This chapter broadens the scope of transmit correlation diversity in a noncoherent MIMO BC to the case of* partially *overlapping eigenspaces and introduces techniques to harvest these generalized gains. For the two-user case, we derive achievable DoF regions and achievable rate regions. We then extend the DoF results to the K-user case by analyzing the interference graph that characterizes the overlapping structure of the correlation eigenspaces. Our achievability results employ a combination of product superposition in the common part of the eigenspaces, and pre-beamforming (rate splitting) to create multiple data streams in nonoverlapping parts of the correlation eigenspaces.*

4.1 Overview

In [37], Fadel and Nosratinia found the optimal DoF region for the noncoherent block-fading BC under the assumption that the channel components are IID, which can be achieved by TDMA. In practice, however, the channels between different antennas are often correlated. The correlation arises from the propagation environment causing the received signal gains to be larger in some spatial directions, and also from the spatially dependent patterns of the antennas. The interest in spatial correlation was sharpened by its experimental validation [149], [150], and more recently by the increasing attention to higher frequencies [151] and larger number of antennas [152].

The effect of spatial correlation on the capacity of MIMO links has been a subject of long-standing interest. Shiu *et al.* proposed an abstract "one-ring" model to determine the spatial fading correlation and studied its effect on the MIMO capacity [153]. In single-user channels with CSIR but no CSIT, channel correlation can boost power but may reduce the DoF [154], [155], thus it can be detrimental at high signal-to-noise ratio (SNR) but a boon at low SNR. Tulino *et al.* characterized analytically the capacity of correlated MIMO channels under a general correlation model in the large antenna array regime in [156]. In [157], Chang *et al.* showed that channel rank deficiency due to spatial correlation lowers

the diversity-multiplexing tradeoff curves from that of uncorrelated channels. Capacity bounds subject to channel estimation errors in correlated fading were characterized in [158], [159].

Under the assumption that all users experience *identical* correlation, Al-Naffouri *et al.* showed that correlation is detrimental to the sum rate scaling of the MIMO BC achieved with various transmission schemes, including the dirty paper coding [160]. However, in practice, the users may have *different* correlation matrices because they are not co-located [161], making it difficult to draw conclusions based on [160]. The sum-rate capacity under user-specific transmit correlations with CSIR was studied in [162], [163]. Furthermore, at higher frequencies or with a large number of antennas, when spatial correlation is unavoidable, comparing capacity against a hypothetically uncorrelated channel may have limited operational impact. Instead, a more relevant question in that case could be: how to maximize performance in the presence of spatial correlation? A useful tool for that purpose is *transmit correlation diversity*, i.e., leveraging the difference between the spatial correlation observed by different users in the system.

Transmit correlation diversity was originally conceived for transmit spatial correlation matrices that have *mutually exclusive* eigenspaces. Under this condition, a JSDM transmission scheme was proposed [113], [114] that reduces the overhead needed for channel estimation. For multi-user networks with orthogonal eigenspace correlation matrices, Adhikary and Caire showed that transmit correlation helps in multi-cell network by partitioning the user spaces into clusters [164]. It was also concluded that transmit correlation benefits the sum rate in the downlink performance of a heterogeneous cellular network (HetNet) where both macro and small cells share the same spectrum [165]. Non-overlapping transmit correlation eigenspaces have also been exploited in a two-tier system where a large number of small cells are deployed under a macro cell [166].

One might ask: *how often do nonoverlapping correlation eigenspaces naturally occur?* Certainly in some scenarios, e.g., severely rank-deficient MIMO links, the occurrence of such link characteristics may be commonplace. However, in many other scenarios, transmit correlation matrices may be different but have *overlapping* eigenspaces, and a natural motivation exists to explore and understand transmit correlation diversity in this more general setting. This chapter significantly broadens the conditions under which correlation diversity can be exploited, and proposes new techniques to harvest these gains. We focus on the noncoherent case, i.e., neither CSIT nor CSIR.

The main contributions of this chapter are summarized as follows.

1. We derive an achievable DoF region for the two-user noncoherent BC in spatially correlated block fading (Theorem 4.1). This region is significantly larger than the TDMA region, especially when the rank r_0 of the overlap between two correlation eigenspaces is large (see Fig. 4.1).

2. Also for the two-user BC, we propose an achievable rate region with rate splitting (pre-beamforming) for arbitrary input distribution satisfying the average power constraint (Theorem 4.3). We characterize this rate region with an explicit input distribution based on orthogonal pilots and Gaussian data symbols (Theorem 4.4). We also derive the rate achieved with product superposition (Theorem 4.5) and a hybrid version of pre-beamforming and product superposition (Theorem 4.6). As a by-product, we find the rate achieved with pilot-based schemes for the P2P channel (Theorem 4.2), which generalizes of the results in [31] to correlated fading.

3. We derive achievable DoF regions for the K-user BC in spatially correlated fading with symmetrically partially overlapping eigenspaces (Theorems 4.7, 4.8) and general correlation structure (Theorem 4.9).

For the achievability results above, we employ rate splitting, product superposition, and a combination of them, thus showing that these transmission techniques are suitable tools to harvest the gain of transmit correlation diversity with partially overlapping eigenspaces. We note that most of our results do not require the fading to be Rayleigh, but rather hold for generic fading.

The remainder of the chapter is organized as follows. First, the system model is introduced in Section 4.2. Then, Section 4.3 and Section 4.4 present respectively the achievable DoF and rate region analysis for the two-user case. Section 4.5 presents the achievable DoF region analysis for the K-user case. Finally, Section 4.6 concludes the chapter.

4.2 System Model

We consider a MIMO BC in which a transmitter equipped with M antennas transmitting to K receivers (users), where user k is equipped with N_k antennas, $k \in [K]$. The channel between the transmitter and user k is flat and block fading with equal-length and synchronous coherence interval (across the users) of T channel uses. That is, the channel propagation matrix $\mathsf{H}_k \in \mathbb{C}^{N_k \times M}$, $k \in [K]$, remains constant during each block of length T symbols and changes independently between blocks. Let the matrix $\mathsf{X}[b] \in \mathbb{C}^{M \times T}$ be the transmitted signal from the M antennas during a coherence block b. The received signal matrix at user k is

$$\mathsf{Y}_k[b] = \mathsf{H}_k[b]\mathsf{X}[b] + \mathsf{Z}_k[b], \quad k \in [K], b = 1, 2, \ldots, \tag{4.1}$$

where $\mathsf{Z}_k[b] \in \mathbb{C}^{N_k \times T}$ is the AWGN with IID $\mathcal{N}_{\mathbb{C}}(0,1)$ entries. The input is subject to the power constraint

$$\frac{1}{\nu} \sum_{b=1}^{\nu} \|\mathsf{X}[b]\|_{\mathrm{F}}^2 \leq PT, \tag{4.2}$$

where ν is the number of blocks spanned by a channel codeword. Therefore, P is the SNR of the channel. Hereafter, we omit the block index b whenever confusion is not likely.

Channel Spatial Correlation

We assume that the channel is spatially correlated according to the Kronecker model (a.k.a. separable model), and focus on the transmit-side correlation. Thus the channel matrices are expressed as

$$\mathsf{H}_k = \breve{\mathsf{H}}_k \boldsymbol{R}_k^{\frac{1}{2}}, \quad k \in [K], \tag{4.3}$$

where $\boldsymbol{R}_k = \frac{1}{N_k}\mathbb{E}[\mathsf{H}_k^{\mathsf{H}}\mathsf{H}_k] \in \mathbb{C}^{M \times M}$, $\mathrm{tr}(\boldsymbol{R}_k) = M$, is the transmit correlation matrix of user k with rank r_k, and $\breve{\mathsf{H}}_k \in \mathbb{C}^{N_k \times M}$ is drawn from a generic distribution satisfying the conditions

$$h(\breve{\mathsf{H}}_k) > -\infty, \quad \mathbb{E}\left[\breve{\mathsf{H}}_k^{\mathsf{H}}\breve{\mathsf{H}}_k\right] = N_k \boldsymbol{I}_M, \quad k \in [K]. \tag{4.4}$$

Since the correlation matrices might be rank-deficient, $\breve{\mathsf{H}}_k$ is not necessarily a minimal representation of the randomness in H_k. The correlation eigenspace of user k is revealed via an eigendecomposition of the correlation matrix:

$$\boldsymbol{R}_k = \boldsymbol{U}_k \boldsymbol{\Sigma}_k \boldsymbol{U}_k^{\mathsf{H}} \tag{4.5}$$

where $\boldsymbol{\Sigma}_k$ is a $r_k \times r_k$ diagonal matrix containing r_k nonzero eigenvalues of $\boldsymbol{R}_k$, and $\boldsymbol{U}_k$ is an $M \times r_k$ matrix whose orthonormal unit column vectors are the eigenvectors of $\boldsymbol{R}_k$ corresponding to the nonzero eigenvalues. The rows of H_k belong to the r_k-dimensional subspace $\mathrm{Span}(\boldsymbol{U}_k)$ of $\boldsymbol{R}_k$, also called as the eigenspace of user k. The channel expression (4.3) can be expanded as

$$\mathsf{H}_k = \breve{\mathsf{H}}_k \boldsymbol{U}_k \boldsymbol{\Sigma}_k^{\frac{1}{2}} \boldsymbol{U}_k^{\mathsf{H}} = \mathsf{G}_k \boldsymbol{\Sigma}_k^{\frac{1}{2}} \boldsymbol{U}_k^{\mathsf{H}}, \tag{4.6}$$

where $\mathsf{G}_k := \breve{\mathsf{H}}_k \boldsymbol{U}_k$ is equivalently drawn from a generic distribution satisfying $h(\mathsf{G}_k) > -\infty$, $\mathbb{E}[\mathsf{G}_k^{\mathsf{H}} \mathsf{G}_k] = N_k \boldsymbol{I}_{r_k}$, $k \in [K]$.

The eigenspaces $\mathrm{Span}(\boldsymbol{U}_k)$ have a prominent role in transmit correlation diversity. For example, methods such as JSDM [113], [114] are critically dependent on finding groups of users whose eigenspaces have no intersection. In contrast, in this chapter, we propose transmission schemes that take advantage of both common and private parts of the eigenspaces. To this end, in several instances, we build an equivalent channel $\bar{\mathsf{H}}_k$ that resides in a *subspace* of the eigenspace $\mathrm{Span}(\boldsymbol{U}_k)$ via the linear transformation

$$\bar{\mathsf{H}}_k = \mathsf{H}_k \boldsymbol{V}_k, \tag{4.7}$$

for some truncated unitary matrix $\boldsymbol{V}_k \in \mathbb{C}^{M \times s_k}$, $s_k \leq r_k$, such that $\mathrm{Span}(\boldsymbol{V}_k) \subset \mathrm{Span}(\boldsymbol{U}_k)$. Unlike $\boldsymbol{U}_k$, $k \in [K]$, that characterize the correlation eigenspaces of the links, the subspaces $\mathrm{Span}(\boldsymbol{V}_k)$ also depend on the proposed transmission schemes and may be customized throughout the chapter.

Channel Information Availability

We assume throughout the chapter that the distribution of H_k, in particular the second-order statistic $\boldsymbol{R}_k$ (and thus $\boldsymbol{\Sigma}_k$ and $\boldsymbol{U}_k$), is known to both the transmitter and user k. This is reasonable because $\boldsymbol{R}_k$ represents long-term behavior of the channel that is stable and can be easily tracked. On the other hand, the realization of H_k is not known *a priori* at any node. User k might attempt to estimate H_k with the help of known pilot symbols inserted in X.

Achievable Rate and DoF

Assuming K independent messages are communicated (no common message), and the corresponding rate tuple $(R_1(P), \ldots, R_K(P))$ is achievable at SNR P, $\forall P \geq 0$, i.e., lie within the capacity region of the channel, then we define an achievable DoF tuple $(d_1, \ldots, d_K)$ as

$$d_k := \lim_{P \to \infty} \frac{R_k(P)}{\log_2 P}, \quad k \in [K]. \tag{4.8}$$

The set of achievable rate (resp., DoF) tuples defines an achievable rate (resp., DoF) region of the channel.

Throughout the chapter, whenever we mention a BC (e.g., in a theorem/proposition), we refer to the noncoherent spatially correlated block-fading BC described in this section, unless otherwise specified. Hereafter, we assume that $T \geq 2\max(r_k, N_k)$ and denote for convenience that $N_k^* := \min(N_k, r_k)$, $k \in [K]$.

4.3 Two-User Broadcast Channel: DoF Analysis

In this section, the DoF performance of the two-user noncoherent BC is presented. If the two eigenspaces are fully overlapping, i.e., the eigenspace of one user is a subspace of the other user's, a DoF pair achievable with product superposition for the BC was proposed in [76] as follows.

Proposition 4.1 ([76, Theorem 4])**.** *For the two-user BC, when the eigenspace of user 2 is a subspace of user 1's (implying $r_2 \leq r_1 \leq M$), the DoF pair $\left(N_1^*(1-\frac{r_1}{T}), N_2^*\frac{r_1-r_2}{T}\right)$ is achievable.*

The DoF pair in Proposition 4.1 is achieved with product superposition [115]. The main idea is to embed information to one user in the pilot for the other user. The details are given in the following to be self-contained.

Proof of Proposition 4.1. There exist transmit eigendirections $\boldsymbol{V}_1 \in \mathbb{C}^{M\times(r_1-r_2)}$ and $\boldsymbol{V}_0 \in \mathbb{C}^{M\times r_2}$ that are aligned with the noncommon and common parts, respectively, of the two channel eigenspaces such that

$$\mathrm{Span}(\boldsymbol{V}_0) = \mathrm{Span}(\boldsymbol{U}_2), \tag{4.9}$$

$$\mathrm{Span}(\boldsymbol{V}_1) = \mathrm{Span}(\boldsymbol{U}_1) \cap \mathrm{Span}(\boldsymbol{U}_2)^{\perp}. \tag{4.10}$$

Define $\boldsymbol{V} := [\boldsymbol{V}_0\ \boldsymbol{V}_1]$. Let the transmitter send the signal $\mathsf{X} = \boldsymbol{V}\mathsf{X}_2\mathsf{X}_1$ during a coherence block, with $\mathsf{X}_1 = [\boldsymbol{I}_{r_1}\ \mathsf{S}_1] \in \mathbb{C}^{r_1\times T}$ and $\mathsf{X}_2 = \begin{bmatrix} \boldsymbol{I}_{r_2} & \mathsf{S}_2 \\ \boldsymbol{0} & \boldsymbol{I}_{r_1-r_2} \end{bmatrix} \in \mathbb{C}^{r_1\times r_1}$, where $\mathsf{S}_1 \in \mathbb{C}^{r_1\times(T-r_1)}$ contains symbols for user 1 and $\mathsf{S}_2 \in \mathbb{C}^{r_2\times(r_1-s_0)}$ contains symbols for user 2. The received signal at user 1 is

$$\mathsf{Y}_1 = \mathsf{H}_1\boldsymbol{V}\mathsf{X}_2\mathsf{X}_1 + \mathsf{Z}_1 = \mathsf{H}_1\boldsymbol{V}\mathsf{X}_2[\boldsymbol{I}_{r_1}\ \mathsf{S}_1] + \mathsf{Z}_1. \tag{4.11}$$

User 1 estimates the equivalent channel $\mathsf{H}_1\boldsymbol{V}\mathsf{X}_2$ and then decodes S_1, achieving $N_1^*(T-r_1)$ DoF. The received signal at user 2 during the first r_1 channel uses is

$$\mathsf{Y}_{2[1:r_1]} = \mathsf{H}_2\boldsymbol{V}\begin{bmatrix} \boldsymbol{I}_{r_2} & \mathsf{S}_2 \\ \boldsymbol{0} & \boldsymbol{I}_{r_1-r_2} \end{bmatrix}\boldsymbol{I}_{r_1} + \mathsf{Z}_{2[1:r_1]} = \mathsf{H}_2\boldsymbol{V}_0[\boldsymbol{I}_{r_2}\quad \mathsf{S}_2] + \mathsf{Z}_{2[1:r_1]}, \tag{4.12}$$

where we used the fact that $\mathsf{H}_2\boldsymbol{V}_1 = \boldsymbol{0}$ due to (4.10). User 2 estimates the equivalent channel $\mathsf{H}_2\boldsymbol{V}_0$, and then decodes S_2, achieving $N_2^*(r_1-r_2)$ DoF. Therefore, the normalized DoF pair $\left(N_1^*(1-\frac{r_1}{T}), N_2^*\frac{r_1-r_2}{T}\right)$ is achievable. □

In the following, we exploit both rate splitting and product superpositions to generalize Proposition 4.1 to a more general setting of partially overlapping eigenspaces.

Theorem 4.1. *For the two-user noncoherent BC and* $\mathrm{rank}(\mathrm{Span}(\boldsymbol{U}_1) \cap \mathrm{Span}(\boldsymbol{U}_2)) =: r_0 \geq 0$*, the DoF pairs $\left(N_1^*(1-\frac{N_1^*}{T}), 0\right)$ and $\left(0, N_2^*(1-\frac{N_2^*}{T})\right)$ are achievable. Furthermore, for any integers (s_1, s_2, s_0) such that $0 \leq s_1 \leq r_1 - r_0$, $0 \leq s_2 \leq r_2 - r_0$, and $0 \leq s_0 \leq r_0$, the DoF pairs*

$$\mathcal{D}_1 = \left(\min(s_0, N_1)\frac{s_2}{T}, \min(s_2+s_0, N_2)\Big(1-\frac{s_2+s_0}{T}\Big)\right), \tag{4.13}$$

$$\mathcal{D}_2 = \left(\min(s_1+s_0, N_1)\Big(1-\frac{s_1+s_0}{T}\Big), \min(s_0, N_2)\frac{s_1}{T}\right) \tag{4.14}$$

are achievable. On top of that, if $s_1 \geq s_2$, the DoF pairs

$$\mathcal{D}_3 = \Bigg(\min(s_1+s_0, N_1)\Big(1-\frac{s_1+s_0}{T}\Big), \\ \min(s_2, N_2)\frac{s_1-s_2}{T} + \min(s_2, (N_2-s_0)^+)\Big(1-\frac{s_1+s_0}{T}\Big)\Bigg), \tag{4.15}$$

$$\mathcal{D}_4 = \Bigg(\min(s_1, (N_1-s_0)^+)\Big(1-\frac{s_1+s_0}{T}\Big), \\ \min(s_2, N_2)\frac{s_1-s_2}{T} + \min(s_2+s_0, N_2)\Big(1-\frac{s_1+s_0}{T}\Big)\Bigg), \tag{4.16}$$

$$\mathcal{D}_5 = \Bigg(\min(s_1+s_0, N_1)\Big(1-\frac{s_1+s_0}{T}\Big), \\ \min(s_2+s_0, N_2)\frac{s_1-s_2}{T} + \min(s_2, N_2)\Big(1-\frac{s_1+s_0}{T}\Big)\Bigg) \tag{4.17}$$

are achievable; if $s_1 \leq s_2$, the DoF pairs

$$\mathcal{D}_3 = \Bigg(\min(s_1, N_1)\frac{s_2-s_1}{T} + \min(s_1, (N_1-s_0)^+)\Big(1-\frac{s_2+s_0}{T}\Big), \\ \min(s_2+s_0, N_2)\Big(1-\frac{s_2+s_0}{T}\Big)\Bigg), \tag{4.18}$$

$$\mathcal{D}_4 = \Bigg(\min(s_1, N_1)\frac{s_2-s_1}{T} + \min(s_1+s_0, N_1)\Big(1-\frac{s_2+s_0}{T}\Big), \\ \min(s_2, (N_2-s_0)^+)\Big(1-\frac{s_2+s_0}{T}\Big)\Bigg), \tag{4.19}$$

$$\mathcal{D}_5 = \Bigg(\min(s_1+s_0, N_1)\frac{s_2-s_1}{T} + \min(s_1, N_1)\Big(1-\frac{s_2+s_0}{T}\Big), \\ \min(s_2+s_0, N_2)\Big(1-\frac{s_2+s_0}{T}\Big)\Bigg) \tag{4.20}$$

are achievable. The convex hull of these DoF pairs (over all feasible values of s_1, s_2, and s_0) and the origin $(0,0)$ is an achievable DoF region for the BC.

Theorem 4.1 coincides with Proposition 4.1 when $r_0 = r_2$.

Remark 4.1. *The parameters s_1, s_2, s_0 represent the allocation of available dimensions to encoding of messages for the two users. By tuning these parameters, we explore the trade-off between the number of data dimensions (indicating the amount of channel uses needed for pilot transmission) and the amount of channel uses for data transmission within each section of the eigenspaces.*

Proof of Theorem 4.1. The DoF pairs $\Big(N_1^*\Big(1-\frac{N_1^*}{T}\Big), 0\Big)$ and $\Big(0, N_2^*\Big(1-\frac{N_2^*}{T}\Big)\Big)$ are achieved by activating only one user according to [117, Theorem 1].

For any nonnegative integers s_1, s_2, s_0 satisfying $s_0 \leq r_0$, $s_1 \leq r_1 - r_0$ and $s_2 \leq r_2 - r_0$, there exist eigendirections $\boldsymbol{V}_0 \in \mathbb{C}^{M\times s_0}, \boldsymbol{V}_1 \in \mathbb{C}^{M\times s_1}, \boldsymbol{V}_2 \in \mathbb{C}^{M\times s_2}$ that are aligned with a

part of the common and noncommon sections of the two channel eigenspaces such that[1]

$$\mathrm{Span}(\boldsymbol{V}_0) \subset \big(\mathrm{Span}(\boldsymbol{U}_1) \cap \mathrm{Span}(\boldsymbol{U}_2)\big), \tag{4.21}$$

$$\mathrm{Span}(\boldsymbol{V}_1) \subset \big(\mathrm{Span}(\boldsymbol{U}_1) \cap \mathrm{Span}(\boldsymbol{U}_2)^{\perp}\big), \tag{4.22}$$

$$\mathrm{Span}(\boldsymbol{V}_2) \subset \big(\mathrm{Span}(\boldsymbol{U}_2) \cap \mathrm{Span}(\boldsymbol{U}_1)^{\perp}\big). \tag{4.23}$$

To achieve $\mathcal{D}_1$, the transmitter employs product superposition and transmits

$$\mathbf{X} = [\boldsymbol{V}_0\ \boldsymbol{V}_1]\mathbf{X}_2\mathbf{X}_1, \tag{4.24}$$

with $\mathbf{X}_1 = [\boldsymbol{I}_{s_1+s_0}\ \mathbf{S}_1]$ and $\mathbf{X}_2 = \begin{bmatrix} \boldsymbol{I}_{s_0} & \mathbf{S}_2 \\ \mathbf{0} & \boldsymbol{I}_{s_1} \end{bmatrix}$, where $\mathbf{S}_1 \in \mathbb{C}^{(s_1+s_0)\times(T-s_1-s_0)}$ and $\mathbf{S}_2 \in \mathbb{C}^{s_0\times s_1}$ contain symbols for user 1 and user 2, respectively. Following steps similar to the proof of Proposition 4.1, it can be shown that this achieves the DoF pair $\mathcal{D}_1$. The DoF pair $\mathcal{D}_2$ can be achieved similarly by switching the users' roles.

When $s_1 \geq s_2$, the pairs $\mathcal{D}_3$ and $\mathcal{D}_4$ are achieved with rate splitting as follows. Let the transmitter send

$$\mathbf{X} = [\boldsymbol{V}_0\ \boldsymbol{V}_1\ \boldsymbol{V}_2] \begin{bmatrix} \boldsymbol{I}_{s_0} & [\mathbf{0}_{s_0\times s_1} & \mathbf{S}_0] \\ \mathbf{0}_{s_1\times s_0} & [\boldsymbol{I}_{s_1} & \mathbf{S}_1] \\ \mathbf{0}_{s_2\times s_0} & [\boldsymbol{I}_{s_2} & \mathbf{S}_2] \end{bmatrix}, \tag{4.25}$$

where $\mathbf{S}_0 \in \mathbb{C}^{s_0\times(T-s_1-s_0)}$ is a common signal to both users while $\mathbf{S}_1 \in \mathbb{C}^{s_1\times(T-s_1-s_0)}$ and $\mathbf{S}_2 \in \mathbb{C}^{s_2\times(T-s_2-s_0)}$ are private signals to user 1 and user 2, respectively. The received signal at user 1 is

$$\begin{aligned} \mathbf{Y}_1 &= \mathbf{H}_1[\boldsymbol{V}_0\ \boldsymbol{V}_1\ \boldsymbol{V}_2]\mathbf{X} + \mathbf{Z}_1 \\ &= \mathbf{H}_1[\boldsymbol{V}_0\ \boldsymbol{V}_1] \begin{bmatrix} \boldsymbol{I}_{s_0} & \mathbf{0} & \mathbf{S}_0 \\ \mathbf{0} & \boldsymbol{I}_{s_1} & \mathbf{S}_1 \end{bmatrix} + \mathbf{Z}_1. \end{aligned} \tag{4.26}$$

User 1 estimates the equivalent channel $\mathbf{H}_1[\boldsymbol{V}_0\ \boldsymbol{V}_1]$ during the first $s_1 + s_0$ channel uses and decodes both $\mathbf{S}_1$ and $\mathbf{S}_0$ during the remaining $T - s_1 - s_0$ channel uses, achieving $\min(s_1 + s_0, N_1)\frac{T-s_1-s_0}{T}$ DoF. The received signal at user 2 is

$$\begin{aligned} \mathbf{Y}_2 &= \mathbf{H}_2[\boldsymbol{V}_0\ \boldsymbol{V}_1\ \boldsymbol{V}_2]\mathbf{X} + \mathbf{Z}_2 \\ &= \mathbf{H}_2[\boldsymbol{V}_0\ \boldsymbol{V}_2] \begin{bmatrix} \boldsymbol{I}_{s_0} & \mathbf{0} & [\mathbf{0}_{s_0\times(s_1-s_2)}\ \mathbf{S}_0] \\ \mathbf{0} & \boldsymbol{I}_{s_2} & \mathbf{S}_2 \end{bmatrix} + \mathbf{Z}_2. \end{aligned} \tag{4.27}$$

User 2 estimates the equivalent channel $\mathbf{H}_2[\boldsymbol{V}_0\ \boldsymbol{V}_2]$ and then decodes $\mathbf{S}_0$ and $\mathbf{S}_2$, achieving $\min(s_2, N_2)\frac{s_1-s_2}{T} + \min(s_2+s_0, N_2)\frac{T-s_1-s_0}{T}$ DoF. By dedicating $\mathbf{S}_0$ to only user 1 or user 2, we can achieve the DoF pairs $\mathcal{D}_3$ and $\mathcal{D}_4$, respectively.

To achieve $\mathcal{D}_5$ (still assuming $s_1 \geq s_2$), we employ a hybrid version of rate splitting and product superposition, referred to as *hybrid superposition*, as follows. The transmitted signal is constructed as

$$\mathbf{X} = [\boldsymbol{V}_0\ \boldsymbol{V}_1]\mathbf{X}_2'\mathbf{X}_1 + \boldsymbol{V}_2\mathbf{X}_2, \tag{4.28}$$

[1] $\boldsymbol{V}_0$ can be calculated from $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ using, e.g., the Zassenhaus algorithm [167]. Specifically, this algorithm uses elementary row operations to transform the $(r_1 + r_2) \times 2M$ matrix $\begin{bmatrix} \boldsymbol{U}_1^{\mathsf{T}} & \boldsymbol{U}_1^{\mathsf{T}} \\ \boldsymbol{U}_2^{\mathsf{T}} & \mathbf{0}_{r_2\times M} \end{bmatrix}$ (or $\begin{bmatrix} \boldsymbol{U}_2^{\mathsf{T}} & \boldsymbol{U}_2^{\mathsf{T}} \\ \boldsymbol{U}_1^{\mathsf{T}} & \mathbf{0}_{r_1\times M} \end{bmatrix}$) to the row echelon form $\begin{bmatrix} \boldsymbol{V}_0^{\mathsf{T}} & * \\ \mathbf{0} & \boldsymbol{V}_0^{\mathsf{T}} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$, where $*$ stands for a matrix which is not of interest. $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ can be found similarly by applying the Zassenhaus algorithm to $\boldsymbol{U}_1$ and $\mathrm{null}(\boldsymbol{U}_2)$, and $\mathrm{null}(\boldsymbol{U}_1)$ and $\boldsymbol{U}_2$, respectively, where $\mathrm{null}(\boldsymbol{U}_k)$ is the matrix such that $[\boldsymbol{U}_k\ \mathrm{null}(\boldsymbol{U}_k)]$ is unitary.

with $\mathbf{X}_2 = [\mathbf{0}_{s_0\times s_0}\ \boldsymbol{I}_{s_2}\ \mathbf{S}_2]$, $\mathbf{X}_1 = [\boldsymbol{I}_{s_1+s_0}\ \mathbf{S}_1]$, and $\mathbf{X}_2' = \begin{bmatrix} \boldsymbol{I}_{s_0} & [\mathbf{0}_{s_0\times s_2}\ \mathbf{S}_2'] \\ \mathbf{0}_{s_1\times s_0} & \boldsymbol{I}_{s_1} \end{bmatrix}$, where $\mathbf{S}_1 \in \mathbb{C}^{(s_1+s_0)\times(T-s_1-s_0)}$ contains symbols intended for user 1 while $\mathbf{S} \in \mathbb{C}^{s_2\times(T-s_2-s_0)}$ and $\mathbf{S}_2' \in \mathbb{C}^{s_0\times(s_1-s_2)}$ contain symbols intended for user 2. The received signal at user 1 is

$$\begin{aligned} \mathbf{Y}_1 &= \mathbf{H}_1[\boldsymbol{V}_0\ \boldsymbol{V}_1]\mathbf{X}_2'\mathbf{X}_1 + \mathbf{Z}_1 \\ &= \mathbf{H}_1[\boldsymbol{V}_0\ \boldsymbol{V}_1]\mathbf{X}_2'[\boldsymbol{I}_{s_1+s_0}\ \mathbf{S}_1] + \mathbf{Z}_1. \end{aligned} \tag{4.29}$$

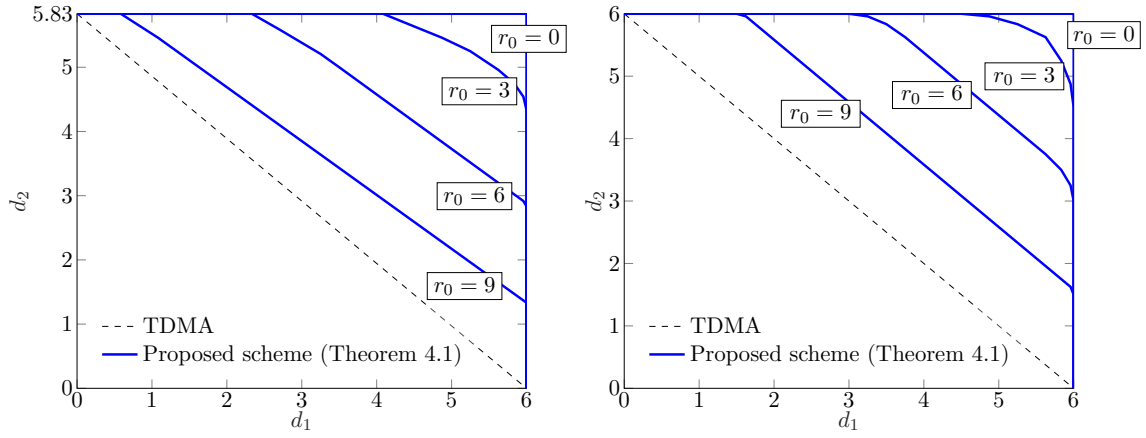
User 1 estimates the equivalent channel $\mathbf{H}_1[\boldsymbol{V}_0\ \boldsymbol{V}_1]\mathbf{X}_2'$, and then decodes $\mathbf{S}_1$ to achieve $\min(s_0+s_1, N_1)\frac{T-s_0-s_1}{T}$ DoF. The received signal at user 2 is

$$\begin{aligned} \mathbf{Y}_2 &= \mathbf{H}_2[\boldsymbol{V}_0\ \boldsymbol{V}_1\ \boldsymbol{V}_2]\mathbf{X} + \mathbf{Z}_2 \\ &= \mathbf{H}_2[\boldsymbol{V}_0\ \boldsymbol{V}_2]\begin{bmatrix} \boldsymbol{I}_{s_0} & \mathbf{0}_{s_2\times s_2} & [\mathbf{S}_2'\ \mathbf{A}] \\ \mathbf{0}_{s_0\times s_0} & \boldsymbol{I}_{s_2} & \mathbf{S}_2 \end{bmatrix} + \mathbf{Z}_2, \end{aligned} \tag{4.30}$$

where $\mathbf{A} := [\boldsymbol{I}_{s_0}\ \mathbf{0}_{s_0\times s_2}\ \mathbf{S}_2']\mathbf{S}_1$. user 2 estimates its equivalent channel $\mathbf{H}_2[\boldsymbol{V}_0\ \boldsymbol{V}_2]$ in the first s_2+s_0 channel uses, and then decodes $\mathbf{S}_2'$ and $\mathbf{S}_2$, achieving $\min(s_2+s_0, N_2)\frac{s_1-s_2}{T} + \min(s_2, N_2)\frac{T-s_1-s_0}{T}$ DoF in total. Therefore, $\mathcal{D}_5$ is achieved.

The proof for the case where $s_1 \geq s_2$ is completed. A similar analysis applies to the case $s_2 \geq s_1$ and completes the proof of Theorem 4.1. □

In Fig. 4.1, the achievable DoF region in Theorem 4.1 is shown for the scenario where $T = 24$, $N_1 = 12$, $N_2 = 12$, $(r_1, r_2) \in \{(12, 10), (12, 12)\}$, and $r_0 \in \{0, 3, 6, 9\}$. We see that exploiting the channel correlation improves significantly the DoF region upon TDMA (which is optimal for uncorrelated Rayleigh fading), especially for small r_0.



(a) $r_1 = 12$, $r_2 = 10$, $r_0 \in \{0, 3, 6, 9\}$ (b) $r_1 = r_2 = 12$, $r_0 \in \{0, 3, 6, 9\}$

FIGURE 4.1: The achievable DoF region for the two-user BC achieved with TDMA or the proposed scheme (Theorem 4.1) for $T = 24$, $N_1 = 12$, $N_2 = 12$, $(r_1, r_2) \in \{(12, 10), (12, 12)\}$, and $r_0 \in \{0, 3, 6, 9\}$.

4.4 Two-User Broadcast Channel: Rate Analysis

In this section, we analyze the achievable rate region of the two-user BC in finite SNR regime. Following are some preliminaries and known results that are useful for our analysis.

Lemma 4.1 (Worst case uncorrelated additive noise [31])**.** *Consider the P2P channel*

$$\mathbf{y} = \sqrt{\frac{P}{M}}\mathbf{H}\mathbf{x} + \mathbf{z}, \tag{4.31}$$

where the channel $\mathbf{H} \in \mathbb{C}^{N\times M}$ *is known to the receiver, and the signal* $\mathbf{x} \in \mathbb{C}^{M\times 1}$ *and the noise* $\mathbf{z} \in \mathbb{C}^{N\times 1}$ *satisfy the power constraints* $\frac{1}{M}\mathbb{E}[\|\mathbf{x}\|^2] = 1$ *and* $\frac{1}{N}\mathbb{E}[\|\mathbf{z}\|^2] = 1$*, are both complex Gaussian distributed, and are uncorrelated, i.e,* $\mathbb{E}[\mathbf{x}\mathbf{z}^{\mathsf{H}}] = \mathbf{0}$*. Let* $\boldsymbol{R}_{\mathbf{x}} := \mathbb{E}[\mathbf{x}\mathbf{x}^{\mathsf{H}}]$ *and* $\boldsymbol{R}_{\mathbf{z}} := \mathbb{E}[\mathbf{z}\mathbf{z}^{\mathsf{H}}]$*. Then the mutual information* $I(\mathbf{y};\mathbf{x}\,|\,\mathbf{H})$ *is lower bounded as*

$$\begin{aligned} I(\mathbf{y};\mathbf{x}\,|\,\mathbf{H}) &\geq \mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_N + \frac{P}{M}\boldsymbol{R}_{\mathbf{z}}^{-1}\mathbf{H}\boldsymbol{R}_{\mathbf{x}}\mathbf{H}^{\mathsf{H}}\Big)\Big], \\ &\qquad \forall\, \boldsymbol{R}_{\mathbf{x}} : \operatorname{tr}(\boldsymbol{R}_{\mathbf{x}}) = M,\ \forall\, \boldsymbol{R}_{\mathbf{z}} :\ \operatorname{tr}(\boldsymbol{R}_{\mathbf{z}}) = N \qquad (4.32) \\ &\geq \min_{\boldsymbol{R}_{\mathbf{z}},\operatorname{tr}(\boldsymbol{R}_{\mathbf{z}})=N} \mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_N + \frac{P}{M}\boldsymbol{R}_{\mathbf{z}}^{-1}\mathbf{H}\boldsymbol{R}_{\mathbf{x}}\mathbf{H}^{\mathsf{H}}\Big)\Big], \quad \forall\, \boldsymbol{R}_{\mathbf{x}} :\ \operatorname{tr}(\boldsymbol{R}_{\mathbf{x}}) = M. \qquad (4.33) \end{aligned}$$

If the distribution of $\mathbf{H}$ *is left rotationally invariant, i.e.,* $p(\boldsymbol{\Theta}\mathbf{H}) = p(\mathbf{H})$ *for any deterministic* $N \times N$ *unitary matrix* $\boldsymbol{\Theta}$*, then the minimizing noise covariance matrix in* (4.33) *is* $\boldsymbol{R}_{\mathbf{z},\text{opt}} = \boldsymbol{I}_N$*.*

Proof. The proof follows from the proof of [31, Theorem 1].[2] Specifically, the mutual information lower bound (4.32) was stated in [31, Eq.(27)]. To show that $\boldsymbol{R}_{\mathbf{z},\text{opt}} = \boldsymbol{I}_N$, we diagonalize $\boldsymbol{R}_{\mathbf{z}}$ using the left rotational invariance of $\mathbf{H}$, and then use the convexity of $\mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_N + \frac{P}{M}\boldsymbol{R}_{\mathbf{z}}^{-1}\mathbf{H}\boldsymbol{R}_{\mathbf{x}}\mathbf{H}^{\mathsf{H}}\Big)\Big]$ in the diagonalized $\boldsymbol{R}_{\mathbf{z}}$. □

The next lemma gives the MMSE estimator used for pilot-based channel estimation.

Lemma 4.2 (MMSE estimator)**.** *Consider the following linear model*

$$\mathbf{Y} = \mathbf{H}\boldsymbol{X} + \mathbf{Z}, \tag{4.34}$$

where $\mathbf{H} \in \mathbb{C}^{N\times M}$ *has correlation matrix* $\boldsymbol{R} = \frac{1}{N}\mathbb{E}[\mathbf{H}^{\mathsf{H}}\mathbf{H}]$*,* $\boldsymbol{X} \in \mathbb{C}^{M\times M}$ *is known, and* $\mathbf{Z} \in \mathbb{C}^{N\times M}$ *has IID* $\mathcal{N}_{\mathbb{C}}(0,1)$ *entries. The linear MMSE estimator for* $\mathbf{H}$ *is given by*

$$\hat{\mathbf{H}} = \mathbf{Y}(\boldsymbol{X}^{\mathsf{H}}\boldsymbol{R}\boldsymbol{X} + \boldsymbol{I}_M)^{-1}\boldsymbol{X}^{\mathsf{H}}\boldsymbol{R}. \tag{4.35}$$

The MMSE estimate $\hat{\mathbf{H}}$ *is also the conditional mean:* $\hat{\mathbf{H}} = \mathbb{E}[\mathbf{H}\,|\,\boldsymbol{X},\mathbf{Y}]$*. The estimate* $\hat{\mathbf{H}}$ *and the estimation error* $\tilde{\mathbf{H}} = \mathbf{H} - \hat{\mathbf{H}}$ *are uncorrelated, have zero mean and row covariance*

$$\frac{1}{N}\mathbb{E}\Big[\hat{\mathbf{H}}^{\mathsf{H}}\hat{\mathbf{H}}\Big] = \boldsymbol{R}\boldsymbol{X}(\boldsymbol{X}^{\mathsf{H}}\boldsymbol{R}\boldsymbol{X} + \boldsymbol{I}_M)^{-1}\boldsymbol{X}^{\mathsf{H}}\boldsymbol{R}, \tag{4.36}$$

$$\frac{1}{N}\mathbb{E}\Big[\tilde{\mathbf{H}}^{\mathsf{H}}\tilde{\mathbf{H}}\Big] = \boldsymbol{R} - \boldsymbol{R}\boldsymbol{X}(\boldsymbol{X}^{\mathsf{H}}\boldsymbol{R}\boldsymbol{X} + \boldsymbol{I}_M)^{-1}\boldsymbol{X}^{\mathsf{H}}\boldsymbol{R}. \tag{4.37}$$

Proof. The linear MMSE channel estimator is given by $\hat{\mathbf{H}} = \mathbf{Y}\boldsymbol{A}$ where $\boldsymbol{A}$ is the minimizer of the mean square error (MSE)

$$\frac{1}{N}\mathbb{E}\Big[\|\mathbf{H} - \hat{\mathbf{H}}\|_{\mathrm{F}}^2\Big] = \operatorname{tr}(\boldsymbol{R}) - \operatorname{tr}(\boldsymbol{R}\boldsymbol{X}\boldsymbol{A}) - \operatorname{tr}(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{X}^{\mathsf{H}}\boldsymbol{R}) + \operatorname{tr}(\boldsymbol{A}^{\mathsf{H}}(\boldsymbol{X}^{\mathsf{H}}\boldsymbol{R}\boldsymbol{X} + \boldsymbol{I}_M)\boldsymbol{A}). \tag{4.38}$$

Solving $\frac{\partial}{\partial \boldsymbol{A}}\frac{1}{N}\mathbb{E}\Big[\|\mathbf{H} - \hat{\mathbf{H}}\|_{\mathrm{F}}^2\Big] = 0$ yields the optimal $\boldsymbol{A}_{\text{opt}} = (\boldsymbol{X}^{\mathsf{H}}\boldsymbol{R}\boldsymbol{X} + \boldsymbol{I}_M)^{-1}\boldsymbol{X}^{\mathsf{H}}\boldsymbol{R}$. Some further simple manipulations give (4.36) and (4.37). □

In the following, we consider partially overlapping eigenspaces and $r_k \leq N_k$, $k \in \{1,2\}$. In addition, we assume w.l.o.g. that $r_1 \geq r_2$.

[2] The argument that the worst case uncorrelated additive noise is Gaussian distributed can also be found in [32].

4.4.1 The Single-User Case

Let us first consider the single-user case and, for simplicity, drop the user's index. The received signal is

$$\mathsf{Y} = \mathsf{H}\mathsf{X} + \mathsf{Z}, \tag{4.39}$$

where the assumptions for the transmitted signal X, the Gaussian noise Z, and the channel H are as before. In particular, H is block fading with coherence interval T and has correlation matrix $\boldsymbol{R} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{U}^{\mathsf{H}}$, thus can be written as $\mathsf{H} = \mathsf{G}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{U}^{\mathsf{H}}$ with $\mathsf{G} \in \mathbb{C}^{N\times r}$ drawn from a generic distribution. The following theorem states the achievable rate in bpcu for this channel.

Theorem 4.2. *For the single-user spatially correlated channel,*

1. *if the transmitter does not exploit $\boldsymbol{R}$, the following rate is achievable with a pilot-based scheme*

$$R = \left(1 - \frac{M}{T}\right)\mathbb{E}\left[\log_2 \det\left(\boldsymbol{I}_N + \frac{P_\delta P_\tau}{P_\delta \mathrm{tr}\left((\boldsymbol{\Sigma}^{-1} + P_\tau \boldsymbol{I}_r)^{-1}\right) + M}\hat{\mathsf{H}}\hat{\mathsf{H}}^{\mathsf{H}}\right)\right], \tag{4.40}$$

   for some power factors P_τ and P_δ satisfying $P_\tau M + P_\delta(T-M) \leq PT$, where $\hat{\mathsf{H}} \in \mathbb{C}^{N\times M}$ is a Gaussian matrix with independent rows following

$$\mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{0}, \boldsymbol{R}(\boldsymbol{I}_M + P_\tau \boldsymbol{R})^{-1}\boldsymbol{R}\Big);$$

2. *if the transmitter exploits $\boldsymbol{R}$, the following rate is achievable with a pilot-based scheme by transmitting in the eigenspace of $\boldsymbol{R}$:*

   - *if the transmitter uses orthogonal pilots:*

$$R = \left(1 - \frac{r}{T}\right)\mathbb{E}\left[\log_2 \det\left(\boldsymbol{I}_N + \frac{P_\delta P_\tau}{P_\delta \mathrm{tr}\left((\bar{\boldsymbol{R}}^{-1} + P_\tau \boldsymbol{I}_r)^{-1}\right) + r}\hat{\boldsymbol{\Omega}}\hat{\boldsymbol{\Omega}}^{\mathsf{H}}\right)\right], \tag{4.41}$$

     where $\hat{\boldsymbol{\Omega}} \in \mathbb{C}^{N\times r}$ is a Gaussian matrix with independent rows following

$$\mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \bar{\boldsymbol{R}}(\boldsymbol{I}_r + P_\tau \bar{\boldsymbol{R}})^{-1}\bar{\boldsymbol{R}}),$$

   - *if the transmitter optimizes the pilots:*

$$R = \left(1 - \frac{r}{T}\right)\mathbb{E}\left[\log_2 \det\left(\boldsymbol{I}_N + \frac{P_\delta}{rP_\delta\left(P_\tau + \frac{1}{r}\mathrm{tr}\left(\bar{\boldsymbol{R}}^{-1}\right)\right)^{-1} + r}\hat{\boldsymbol{\Omega}}\hat{\boldsymbol{\Omega}}^{\mathsf{H}}\right)\right], \tag{4.42}$$

     where $\hat{\boldsymbol{\Omega}} \in \mathbb{C}^{N\times r}$ is a Gaussian matrix with independent rows following

$$\mathcal{N}_{\mathbb{C}}\left(\boldsymbol{0}, \bar{\boldsymbol{R}} - \left(P_\tau + \frac{1}{r}\mathrm{tr}\left(\bar{\boldsymbol{R}}^{-1}\right)\right)^{-1}\boldsymbol{I}_r\right),$$

   for some power factors P_τ and P_δ satisfying $P_\tau r + P_\delta(T-r) \leq PT$, where $\bar{\boldsymbol{R}} := \boldsymbol{V}^{\mathsf{H}}\boldsymbol{R}\boldsymbol{V}$ for a truncated unitary matrix $\boldsymbol{V} \in \mathbb{C}^{M\times r}$ such that $\mathrm{Span}(\boldsymbol{V}) = \mathrm{Span}(\boldsymbol{U})$. The optimal

power allocation maximizing the rate in (4.42) *is characterized by* $P_\tau = \frac{(1-\alpha)PT}{r}$ *and* $P_\delta = \frac{\alpha PT}{T-r}$ *with*

$$\alpha = \begin{cases} \frac{1}{2}, & \text{if } T = 2r, \\ b - \sqrt{b(b-a)}, & \text{if } T > 2r, \end{cases} \tag{4.43}$$

where $a := 1 + \frac{\mathrm{tr}\left(\bar{\boldsymbol{R}}^{-1}\right)}{PT} - \frac{r^2}{PT\mathrm{tr}\left(\bar{\boldsymbol{R}}\right)}$ *and* $b := \frac{T-r}{T-2r}\left(1 + \frac{\mathrm{tr}\left(\bar{\boldsymbol{R}}^{-1}\right)}{PT}\right)$.

Corollary 4.1. *If the channel is uncorrelated, i.e.,* $\boldsymbol{R} = \boldsymbol{I}_M$*, the achievable rate is*

$$R = \left(1 - \frac{M}{T}\right)\mathbb{E}\left[\log_2 \det\left(\boldsymbol{I}_N + \frac{P_\delta P_\tau}{M(1 + P_\delta + P_\tau)}\mathsf{H}\mathsf{H}^{\mathsf{H}}\right)\right] \quad bpcu, \tag{4.44}$$

where $\mathsf{H} \in \mathbb{C}^{N\times M}$ *is the uncorrelated channel matrix. This coincides with [31, Eq.(21)].*

Proof of Theorem 4.2. We prove by constructing pilot-based schemes that can achieve (4.40), (4.41), and (4.42). The proof follows by extending the analysis in [31] to the case of correlated generic fading.

Case 1: The Transmitter Does Not Exploit $\boldsymbol{R}$

In this case, the transmitter ignores $\boldsymbol{R}$ and forms the transmitted signal as if the channel is uncorrelated. Within each coherence block, the transmitter first sends an orthogonal pilot matrix $\boldsymbol{X}_\tau \in \mathbb{C}^{M\times M}$ such that $\boldsymbol{X}_\tau \boldsymbol{X}_\tau^{\mathsf{H}} = M\boldsymbol{I}_M$ during the first M channel uses (this is optimal for uncorrelated fading [31, Sec. III-A]), and then sends IID $\mathcal{N}_{\mathbb{C}}(0,1)$ data matrix $\mathsf{X}_\delta \in \mathbb{C}^{M\times(T-M)}$ during the remaining $T - M$ channel uses. That is,

$$\mathsf{X} = \left[\sqrt{\frac{P_\tau}{M}}\boldsymbol{X}_\tau \ \sqrt{\frac{P_\delta}{M}}\mathsf{X}_\delta\right], \tag{4.45}$$

where P_τ and P_δ are the average power used for training and data phases, respectively, and satisfy the power constraint $P_\tau M + P_\delta(T - M) \le PT$.

In the training phase, the receiver observes $\mathsf{Y}_\tau := \mathsf{Y}_{[1:M]} = \sqrt{\frac{P_\tau}{M}}\mathsf{H}\boldsymbol{X}_\tau + \mathsf{Z}_{[1:M]}$. Following Lemma 4.2, it performs a linear MMSE channel estimator as

$$\hat{\mathsf{H}} = \sqrt{\frac{P_\tau}{M}}\mathsf{Y}_\tau\left(\frac{P_\tau}{M}\boldsymbol{X}_\tau^{\mathsf{H}}\boldsymbol{R}\boldsymbol{X}_\tau + \boldsymbol{I}_M\right)^{-1}\boldsymbol{X}_\tau^{\mathsf{H}}\boldsymbol{R}. \tag{4.46}$$

The estimate $\hat{\mathsf{H}}$ and the estimation error $\tilde{\mathsf{H}} = \mathsf{H} - \hat{\mathsf{H}}$ have zero mean and row covariance

$$\frac{1}{N}\mathbb{E}\left[\hat{\mathsf{H}}^{\mathsf{H}}\hat{\mathsf{H}}\right] = \frac{P_\tau}{M}\boldsymbol{R}\boldsymbol{X}_\tau\left(\frac{P_\tau}{M}\boldsymbol{X}_\tau^{\mathsf{H}}\boldsymbol{R}\boldsymbol{X}_\tau + \boldsymbol{I}_M\right)^{-1}\boldsymbol{X}_\tau^{\mathsf{H}}\boldsymbol{R} = P_\tau\boldsymbol{R}(\boldsymbol{I}_M + P_\tau\boldsymbol{R})^{-1}\boldsymbol{R}, \tag{4.47}$$

$$\frac{1}{N}\mathbb{E}\left[\tilde{\mathsf{H}}^{\mathsf{H}}\tilde{\mathsf{H}}\right] = \boldsymbol{R} - P_\tau\boldsymbol{R}(\boldsymbol{I}_M + P_\tau\boldsymbol{R})^{-1}\boldsymbol{R}. \tag{4.48}$$

In the data transmission phase, the received signal is

$$\mathsf{Y}_\delta := \mathsf{Y}_{[M+1:T]} = \sqrt{\frac{P_\delta}{M}}\mathsf{H}\mathsf{X}_\delta + \mathsf{Z}_{[M+1:T]} = \sqrt{\frac{P_\delta}{M}}\hat{\mathsf{H}}\mathsf{X}_\delta + \mathsf{Z}_\delta, \tag{4.49}$$

where $\mathsf{Z}_\delta := \sqrt{\frac{P_\delta}{M}}\tilde{\mathsf{H}}\mathsf{X}_\delta + \mathsf{Z}_{[M+1:T]}$ is the combined noise consisting of additive noise and channel estimation error. With MMSE estimator, Z_δ and X_δ are uncorrelated because

$$\mathbb{E}[\mathsf{X}_\delta \mathsf{Z}_\delta^{\mathsf{H}} | \boldsymbol{X}_\tau, \mathsf{Y}_\tau] = \mathbb{E}\left[\mathsf{X}_\delta \left(\sqrt{\frac{P_\delta}{M}}\mathsf{X}_\delta^{\mathsf{H}}\tilde{\mathsf{H}}^{\mathsf{H}} + \mathsf{Z}_\delta^{\mathsf{H}}\right) \,\middle|\, \boldsymbol{X}_\tau, \mathsf{Y}_\tau\right] \tag{4.50}$$

$$= \sqrt{\frac{P_\delta}{M}}\mathbb{E}\left[\mathsf{X}_\delta \mathsf{X}_\delta^{\mathsf{H}}(\mathsf{H} - \hat{\mathsf{H}}) \middle| \boldsymbol{X}_\tau, \mathsf{Y}_\tau\right] \tag{4.51}$$

$$= \mathbf{0}, \tag{4.52}$$

since $\mathbb{E}\left[\mathsf{H} - \hat{\mathsf{H}} \middle| \boldsymbol{X}_\tau, \mathsf{Y}_\tau\right] = 0$. From Lemma 4.1, a lower bound on the achievable rate is obtained by replacing Z_δ by IID Gaussian noise with the same variance

$$\sigma^2_{\mathsf{Z}_\delta} = \frac{1}{N(T-M)}\mathrm{tr}\mathbb{E}[\mathsf{Z}_\delta^{\mathsf{H}}\mathsf{Z}_\delta] = \frac{P_\delta}{M}\mathrm{tr}\Big(\boldsymbol{R} - P_\tau \boldsymbol{R}(\boldsymbol{I}_M + P_\tau \boldsymbol{R})^{-1}\boldsymbol{R}\Big) + 1 \tag{4.53}$$

$$= \frac{P_\delta}{M}\mathrm{tr}\Big((\boldsymbol{\Sigma}^{-1} + P_\tau \boldsymbol{I}_r)^{-1}\Big) + 1. \tag{4.54}$$

Thus, the achievable rate is lower bounded by

$$R = \frac{T-M}{T}\mathbb{E}\left[\log_2 \det\left(\boldsymbol{I}_N + \frac{P_\delta}{M\sigma^2_{\mathsf{Z}_\delta}}\hat{\mathsf{H}}\hat{\mathsf{H}}^{\mathsf{H}}\right)\right]. \tag{4.55}$$

From (4.47), $\hat{\mathsf{H}}$ has correlation matrix $P_\tau \boldsymbol{R}(\boldsymbol{I}_M + P_\tau \boldsymbol{R})^{-1}\boldsymbol{R}$. This shows (4.40).

Case 2: The Transmitter Exploits $\boldsymbol{R}$

By exploiting $\boldsymbol{R}$, the transmitter can project the signal onto the eigenspace of $\boldsymbol{R}$ and can also adapt the pilot symbols. The transmitter builds a precoder $\boldsymbol{V} \in \mathbb{C}^{M\times r}$ with r orthonormal columns such that $\mathrm{Span}(\boldsymbol{V}) = \mathrm{Span}(\boldsymbol{U})$. Let $\boldsymbol{\Phi} = \boldsymbol{U}^{\mathsf{H}}\boldsymbol{V}$. The transmitted signal is

$$\mathsf{X} = \boldsymbol{V}\left[\sqrt{\frac{P_\tau}{r}}\boldsymbol{X}_\tau \;\; \sqrt{\frac{P_\delta}{r}}\mathsf{X}_\delta\right] \tag{4.56}$$

where $\boldsymbol{X}_\tau \in \mathbb{C}^{r\times r}$ such that $\mathrm{rank}(\boldsymbol{X}_\tau) = r$ and $\mathrm{tr}(\boldsymbol{X}_\tau^{\mathsf{H}}\boldsymbol{X}_\tau) = r^2$ is the pilot matrix, and $\mathsf{X}_\delta \in \mathbb{C}^{r\times(T-r)}$ is the data matrix containing $\mathcal{N}_{\mathbb{C}}(0,1)$ entries. The average pilot and data powers satisfy $P_\tau r + P_\delta(T-r) \le PT$.

The received signal during the training phase is then $\mathsf{Y}_\tau := \mathsf{Y}_{[1:r]} = \sqrt{\frac{P_\tau}{r}}\mathsf{G}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Phi}\boldsymbol{X}_\tau + \mathsf{Z}_{[1:r]}$. The equivalent channel $\boldsymbol{\Omega} := \mathsf{G}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Phi}$ has correlation matrix $\bar{\boldsymbol{R}} = \boldsymbol{\Phi}^{\mathsf{H}}\boldsymbol{\Sigma}\boldsymbol{\Phi} = \boldsymbol{V}^{\mathsf{H}}\boldsymbol{R}\boldsymbol{V}$. According to Lemma 4.2, the MMSE channel estimate for the equivalent channel $\boldsymbol{\Omega}$ is given by

$$\hat{\boldsymbol{\Omega}} = \sqrt{\frac{P_\tau}{r}}\mathsf{Y}_\tau\left(\frac{P_\tau}{r}\boldsymbol{X}_\tau^{\mathsf{H}}\bar{\boldsymbol{R}}\boldsymbol{X}_\tau + \boldsymbol{I}_r\right)^{-1}\boldsymbol{X}_\tau^{\mathsf{H}}\bar{\boldsymbol{R}}. \tag{4.57}$$

The estimate $\hat{\boldsymbol{\Omega}}$ and the estimation error $\tilde{\boldsymbol{\Omega}} = \mathsf{G}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Phi} - \hat{\boldsymbol{\Omega}}$ have zero mean and row covariance

$$\frac{1}{N}\mathbb{E}\left[\hat{\boldsymbol{\Omega}}^{\mathsf{H}}\hat{\boldsymbol{\Omega}}\right] = \frac{P_\tau}{r}\bar{\boldsymbol{R}}\boldsymbol{X}_\tau\left(\frac{P_\tau}{r}\boldsymbol{X}_\tau^{\mathsf{H}}\bar{\boldsymbol{R}}\boldsymbol{X}_\tau + \boldsymbol{I}_r\right)^{-1}\boldsymbol{X}_\tau^{\mathsf{H}}\bar{\boldsymbol{R}}, \tag{4.58}$$

$$\frac{1}{N}\mathbb{E}\left[\tilde{\boldsymbol{\Omega}}^{\mathsf{H}}\tilde{\boldsymbol{\Omega}}\right] = \bar{\boldsymbol{R}} - \frac{P_\tau}{r}\bar{\boldsymbol{R}}\boldsymbol{X}_\tau\left(\frac{P_\tau}{r}\boldsymbol{X}_\tau^{\mathsf{H}}\bar{\boldsymbol{R}}\boldsymbol{X}_\tau + \boldsymbol{I}_r\right)^{-1}\boldsymbol{X}_\tau^{\mathsf{H}}\bar{\boldsymbol{R}} = \left(\bar{\boldsymbol{R}}^{-1} + \frac{P_\tau}{r}\boldsymbol{X}_\tau\boldsymbol{X}_\tau^{\mathsf{H}}\right)^{-1}. \tag{4.59}$$

In the data transmission phase, the received signal is

$$\mathsf{Y}_\delta := \mathsf{Y}_{[r+1:T]} = \sqrt{\frac{P_\delta}{r}}\mathsf{G}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Phi}\mathsf{X}_\delta + \mathsf{Z}_{[r+1:T]} = \sqrt{\frac{P_\delta}{r}}\hat{\boldsymbol{\Omega}}\mathsf{X}_\delta + \mathsf{Z}_\delta, \tag{4.60}$$

where $\mathsf{Z}_\delta := \sqrt{\frac{P_\delta}{r}}\tilde{\boldsymbol{\Omega}}\mathsf{X}_\delta + \mathsf{Z}_{[r+1:T]}$. From Lemma 4.1, a lower bound on the achievable rate is obtained by replacing Z_δ with IID Gaussian noise with the same variance

$$\sigma^2_{\mathsf{Z}_\delta} = \frac{1}{N(T-r)}\mathrm{tr}\mathbb{E}[\mathsf{Z}_\delta^{\mathsf{H}}\mathsf{Z}_\delta] = \frac{P_\delta}{r}\mathrm{tr}\left(\left(\bar{\boldsymbol{R}}^{-1} + \frac{P_\tau}{r}\boldsymbol{X}_\tau\boldsymbol{X}_\tau^{\mathsf{H}}\right)^{-1}\right) + 1. \tag{4.61}$$

The corresponding achievable rate lower bound is

$$R = \frac{T-r}{T}\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_N + \frac{P_\delta}{r\sigma^2_{\mathsf{Z}_\delta}}\hat{\boldsymbol{\Omega}}\hat{\boldsymbol{\Omega}}^{\mathsf{H}}\right)\right] \tag{4.62}$$

$$= \left(1-\frac{r}{T}\right)\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_N + \frac{P_\delta}{P_\delta\mathrm{tr}(\boldsymbol{B})+r}\hat{\boldsymbol{\Omega}}\hat{\boldsymbol{\Omega}}^{\mathsf{H}}\right)\right] \tag{4.63}$$

where we used (4.61) and $\hat{\boldsymbol{\Omega}}$ has correlation matrix $\bar{\boldsymbol{R}} - \boldsymbol{B}$ with $\boldsymbol{B} := \left(\bar{\boldsymbol{R}}^{-1} + \frac{P_\tau}{r}\boldsymbol{X}_\tau\boldsymbol{X}_\tau^{\mathsf{H}}\right)^{-1}$ according to (4.58).

Taking $\boldsymbol{X}_\tau$ such that $\boldsymbol{X}_\tau\boldsymbol{X}_\tau^{\mathsf{H}} = r\boldsymbol{I}_r$ (i.e., orthogonal pilots), we have $\boldsymbol{B} = (\bar{\boldsymbol{R}}^{-1} + P_\tau\boldsymbol{I}_r)^{-1}$, and the achievable rate R is given in (4.41).

We can also optimize the pilot $\boldsymbol{X}_\tau$ so as to maximize R. The pilot matrix $\boldsymbol{X}_\tau$ affects the achievable rate bound primarily through the effective SNR

$$P_{\text{eff}} = \frac{P_\delta}{r\sigma^2_{\mathsf{Z}_\delta}}\frac{1}{N}\mathbb{E}\mathrm{tr}\left(\hat{\boldsymbol{\Omega}}^{\mathsf{H}}\hat{\boldsymbol{\Omega}}\right) = \frac{P_\delta\mathrm{tr}\left(\bar{\boldsymbol{R}}-\boldsymbol{B}\right)}{P_\delta\mathrm{tr}(\boldsymbol{B})+r} \tag{4.64}$$

which decreases with $\mathrm{tr}(\boldsymbol{B})$. Therefore, to maximize R, we would like to minimize $\mathrm{tr}(\boldsymbol{B})$. That is

$$\min_{\mathrm{tr}\left(\boldsymbol{X}_\tau^{\mathsf{H}}\boldsymbol{X}_\tau\right)=r^2}\mathrm{tr}\left(\left(\bar{\boldsymbol{R}}^{-1} + \frac{P_\tau}{r}\boldsymbol{X}_\tau\boldsymbol{X}_\tau^{\mathsf{H}}\right)^{-1}\right). \tag{4.65}$$

Using Lagrange multiplier λ, we minimize

$$L(\boldsymbol{X}_\tau,\lambda) = \mathrm{tr}\left(\left(\bar{\boldsymbol{R}}^{-1} + \frac{P_\tau}{r}\boldsymbol{X}_\tau\boldsymbol{X}_\tau^{\mathsf{H}}\right)^{-1}\right) + \lambda(\mathrm{tr}(\boldsymbol{X}_\tau\boldsymbol{X}_\tau^{\mathsf{H}}) - r^2). \tag{4.66}$$

Solving $\frac{\partial L(\boldsymbol{X}_\tau,\lambda)}{\partial \boldsymbol{X}_\tau\boldsymbol{X}_\tau^{\mathsf{H}}} = 0$, we obtain the minimizer $\boldsymbol{X}_\tau\boldsymbol{X}_\tau^{\mathsf{H}} = \sqrt{\frac{r}{P_\tau\lambda}}\boldsymbol{I}_r - \frac{r}{P_\tau}\bar{\boldsymbol{R}}^{-1}$. Using the constraint $\mathrm{tr}(\boldsymbol{X}_\tau^{\mathsf{H}}\boldsymbol{X}_\tau) = r^2$, we find that $\frac{P_\tau}{r}\boldsymbol{X}_\tau\boldsymbol{X}_\tau^{\mathsf{H}} = \left(P_\tau + \frac{1}{r}\mathrm{tr}\left(\bar{\boldsymbol{R}}^{-1}\right)\right)\boldsymbol{I}_r - \bar{\boldsymbol{R}}^{-1}$. With this, $\boldsymbol{B} = \left(P_\tau + \frac{1}{r}\mathrm{tr}\left(\bar{\boldsymbol{R}}^{-1}\right)\right)^{-1}\boldsymbol{I}_r$, and R is given in (4.42). The effective SNR is now written as

$$P_{\text{eff}} = \frac{P_\delta}{P_\delta r\left(P_\tau + \frac{1}{r}\mathrm{tr}\left(\bar{\boldsymbol{R}}^{-1}\right)\right)^{-1} + r}\left[\mathrm{tr}\left(\bar{\boldsymbol{R}}\right) - r\left(P_\tau + \frac{1}{r}\mathrm{tr}\left(\bar{\boldsymbol{R}}^{-1}\right)\right)^{-1}\right]. \tag{4.67}$$

Let $P_\tau r = (1-\alpha)PT$ and $P_\delta(T-r) = \alpha PT$ for $\alpha \in (0,1)$, we can derive that

$$P_{\text{eff}} = \frac{PT\mathrm{tr}\left(\bar{\boldsymbol{R}}\right)}{r(T-2r)}\frac{-\alpha^2 + a\alpha}{-\alpha + b} \tag{4.68}$$

where $a := 1 + \frac{\mathrm{tr}\left(\bar{\boldsymbol{R}}^{-1}\right)}{PT} - \frac{r^2}{PT\mathrm{tr}\left(\bar{\boldsymbol{R}}\right)}$ and $b := \frac{T-r}{T-2r}\left(1 + \frac{\mathrm{tr}\left(\bar{\boldsymbol{R}}^{-1}\right)}{PT}\right)$. Noting that $T - 2r \geq 0$, we obtain the optimal value of α that maximizes P_{eff} as given in (4.43). This concludes the proof. □

4.4.2 A Baseline TDMA Scheme

We consider a baseline scheme based on orthogonal transmission, i.e., TDMA that activates only one user at a time. According to Theorem 4.2, the following corollary demonstrates the achievable rate with TDMA.

Corollary 4.2. *For the two-user noncoherent BC, if the transmitter does not exploit $\boldsymbol{R}_k$, the following rate is achievable by activating only user k:*

$$R_k = \left(1 - \frac{M}{T}\right)\mathbb{E}\left[\log_2 \det\left(\boldsymbol{I}_{N_k} + \frac{P_\delta P_\tau}{P_\delta \mathrm{tr}\left(\left(\boldsymbol{\Sigma}_k^{-1} + P_\tau \boldsymbol{I}_{r_k}\right)^{-1}\right) + M}\hat{\mathsf{H}}\hat{\mathsf{H}}^{\mathsf{H}}\right)\right] \tag{4.69}$$

for some power factors P_τ and P_δ satisfying $P_\tau M + P_\delta(T-M) \le PT$, where $\hat{\mathsf{H}} \in \mathbb{C}^{N_k \times M}$ is a Gaussian matrix with independent rows following $\mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{R}_k(\boldsymbol{I}_M + P_\tau \boldsymbol{R}_k)^{-1}\boldsymbol{R}_k)$. If the transmitter transmits in the eigenspace of $\boldsymbol{R}_k$ and optimize the pilot, the following rate is achievable by activating only user k:

$$R_k = \left(1 - \frac{r_k}{T}\right)\mathbb{E}\left[\log_2 \det\left(\boldsymbol{I}_{N_k} + \frac{P_\delta}{r_k P_\delta\left(P_\tau + \frac{1}{r_k}\mathrm{tr}\left(\boldsymbol{\Sigma}_k^{-1}\right)\right)^{-1} + r_k}\hat{\boldsymbol{\Omega}}\hat{\boldsymbol{\Omega}}^{\mathsf{H}}\right)\right] \tag{4.70}$$

for some power factors P_τ and P_δ satisfying $P_\tau r + P_\delta(T-r) \le PT$, where $\hat{\boldsymbol{\Omega}} \in \mathbb{C}^{N_k \times r}$ is a Gaussian matrix with independent rows following $\mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{0}, \boldsymbol{\Sigma}_k - \Big(P_\tau + \frac{1}{r_k}\mathrm{tr}\Big(\boldsymbol{\Sigma}_k^{-1}\Big)\Big)^{-1}\boldsymbol{I}_{r_k}\Big)$. Furthermore, the optimal power allocation for the rate in (4.70) is given by $P_\tau r_k = (1-\alpha)PT$ and $P_\delta(T-r_k) = \alpha PT$ with

$$\alpha = \begin{cases} \frac{1}{2}, & \text{if } T = 2r_k, \\ b - \sqrt{b(b-a)}, & \text{if } T > 2r_k, \end{cases} \tag{4.71}$$

where $a := 1 + \frac{\mathrm{tr}\left(\boldsymbol{\Sigma}_k^{-1}\right)}{PT} - \frac{r_k^2}{PT\mathrm{tr}(\boldsymbol{\Sigma}_k)}$ and $b := \frac{T-r_k}{T-2r_k}\left(1 + \frac{\mathrm{tr}\left(\boldsymbol{\Sigma}_k^{-1}\right)}{PT}\right)$. The convex hull of $(0,0)$, $(R_1, 0)$, and $(0, R_2)$ is an achievable rate region with TDMA for the BC.

Note that to achieve (4.70), the transmitter uses the precoder $\boldsymbol{V}_k = \boldsymbol{U}_k$ and optimizes the pilots.

4.4.3 Rate Splitting

In the following, we analyze the rate achievable with the schemes achieving the DoF region in Theorem 4.1. Recall that for a set of nonnegative integers $s_0 \le r_0$, $s_1 \le r_1 - r_0$, and $s_2 \le r_2 - r_0$, we build the precoding matrices $\boldsymbol{V}_k$, $k \in \{0,1,2\}$, as in (4.21)-(4.23). For $k \in \{1,2\}$, we denote

- $\boldsymbol{\Phi}_k := \boldsymbol{U}_k^{\mathsf{H}}[\boldsymbol{V}_0\ \boldsymbol{V}_k]$, $\boldsymbol{\Phi}_{k0} := \boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{V}_0$, $\boldsymbol{\Phi}_{kk} := \boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{V}_k$ (so $\boldsymbol{\Phi}_k = [\boldsymbol{\Phi}_{k0}\ \boldsymbol{\Phi}_{kk}]$);
- $\bar{\boldsymbol{R}}_k := \boldsymbol{\Phi}_k^{\mathsf{H}}\boldsymbol{\Sigma}_k\boldsymbol{\Phi}_k$, $\bar{\boldsymbol{R}}_{k0} := \boldsymbol{\Phi}_k^{\mathsf{H}}\boldsymbol{\Sigma}_k\boldsymbol{\Phi}_{k0}$, $\bar{\boldsymbol{R}}_{kk} := \boldsymbol{\Phi}_k^{\mathsf{H}}\boldsymbol{\Sigma}_k\boldsymbol{\Phi}_{kk}$ (so $\bar{\boldsymbol{R}} = [\bar{\boldsymbol{R}}_{k0}\ \bar{\boldsymbol{R}}_{kk}]$);
- $\breve{\boldsymbol{R}}_{k0} := \boldsymbol{\Phi}_{k0}^{\mathsf{H}}\boldsymbol{\Sigma}_k\boldsymbol{\Phi}_{k0}$, $\breve{\boldsymbol{R}}_{kk} := \boldsymbol{\Phi}_{kk}^{\mathsf{H}}\boldsymbol{\Sigma}_k\boldsymbol{\Phi}_{kk}$.

With rate splitting, we let the transmitter transmit

$$\mathsf{X} = \boldsymbol{V}_0\mathsf{X}_0 + \boldsymbol{V}_1\mathsf{X}_1 + \boldsymbol{V}_2\mathsf{X}_2, \tag{4.72}$$

where $\mathbf{X}_0$, $\mathbf{X}_1$, and $\mathbf{X}_2$ are independent and satisfy the power constraint

$$\mathbb{E}\Big[\|\mathbf{X}_0\|_{\mathrm{F}}^2 + \|\mathbf{X}_1\|_{\mathrm{F}}^2 + \|\mathbf{X}_2\|_{\mathrm{F}}^2\Big] \leq PT. \tag{4.73}$$

Thanks to the precoders, the private signal $\mathbf{X}_k$ is seen by user k only, while the common signal $\mathbf{X}_0$ is seen by both users. Specifically, the received signals is

$$\mathbf{Y}_1 = \mathbf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{10}\mathbf{X}_0 + \mathbf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{11}\mathbf{X}_1 + \mathbf{Z}_1, \tag{4.74}$$

$$\mathbf{Y}_2 = \mathbf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_{20}\mathbf{X}_0 + \mathbf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_{22}\mathbf{X}_2 + \mathbf{Z}_2, \tag{4.75}$$

where the equivalent channels $\mathbf{G}_k\boldsymbol{\Sigma}_k^{\frac{1}{2}}\boldsymbol{\Phi}_{k0} \in \mathbb{C}^{N_k \times s_0}$ and $\mathbf{G}_k\boldsymbol{\Sigma}_k^{\frac{1}{2}}\boldsymbol{\Phi}_{kk} \in \mathbb{C}^{N_k \times s_k}$, $k \in \{1,2\}$, are correlated and unknown. It can be observed that the received signal at each user is similar to a noncoherent two-user MAC: (4.74) as the MAC 1 with (s_0, s_1) equivalent transmit antennas and N_1 receive antennas, (4.75) as the MAC 2 with (s_0, s_2) equivalent transmit antennas and N_2 receive antennas. The two MACs share a common signal $\mathbf{X}_0$. From the capacity region of the MACs [22], we know that the rate pairs (R_0, R_1^p) and (R_0, R_2^p) are simultaneously achievable for the MAC 1 and MAC 2, respectively, if the rates $R_0 \geq 0, R_1^p \geq 0, R_2^p \geq 0$ satisfy

$$\begin{cases} R_0 & \leq \dfrac{1}{T} I(\mathbf{Y}_1; \mathbf{X}_0 | \mathbf{X}_1), & \text{(4.76a)} \\ R_1^p & \leq \dfrac{1}{T} I(\mathbf{Y}_1; \mathbf{X}_1 | \mathbf{X}_0), & \text{(4.76b)} \\ R_0 + R_1^p & \leq \dfrac{1}{T} I(\mathbf{Y}_1; \mathbf{X}_0, \mathbf{X}_1), & \text{(4.76c)} \\ R_0 & \leq \dfrac{1}{T} I(\mathbf{Y}_2; \mathbf{X}_0 | \mathbf{X}_2), & \text{(4.76d)} \\ R_2^p & \leq \dfrac{1}{T} I(\mathbf{Y}_2; \mathbf{X}_2 | \mathbf{X}_0), & \text{(4.76e)} \\ R_0 + R_1^p & \leq \dfrac{1}{T} I(\mathbf{Y}_2; \mathbf{X}_0, \mathbf{X}_2). & \text{(4.76f)} \end{cases}$$

Then for the BC, user 1 achieves rate R_1^p with private signal $\mathbf{X}_1$, user 2 achieves rate R_2^p with private signal $\mathbf{X}_2$, and both users can achieve rate R_0 with common signal $\mathbf{X}_0$. Let R_{0k} be the user k's share in R_0, then the BC can achieve the rate pair $(R_1, R_2) = (R_{01} + R_1^p, R_{02} + R_2^p)$. Replacing $R_0 = R_{01} + R_{02}$, $R_1^p = R_1 - R_{01}$, and $R_2^p = R_2 - R_{02}$ in (4.76) and applying Fourier-Motzkin elimination, we obtain the following achievable rate region of the BC.

Theorem 4.3. *With rate splitting, the two-user noncoherent BC can achieve any rate pair* (R_1, R_2) *satisfying*

$$\begin{cases} R_1 & \leq \dfrac{1}{T} \min\{I(\mathbf{Y}_1; \mathbf{X}_1, \mathbf{X}_0), I(\mathbf{Y}_1; \mathbf{X}_1 | \mathbf{X}_0) + I(\mathbf{Y}_2; \mathbf{X}_0 | \mathbf{X}_2)\}, & \text{(4.77a)} \\ R_2 & \leq \dfrac{1}{T} \min\{I(\mathbf{Y}_2; \mathbf{X}_2, \mathbf{X}_0), I(\mathbf{Y}_2; \mathbf{X}_2 | \mathbf{X}_0) + I(\mathbf{Y}_1; \mathbf{X}_0 | \mathbf{X}_1)\}, & \text{(4.77b)} \\ R_1 + R_2 & \leq \dfrac{1}{T} \min\{I(\mathbf{Y}_1; \mathbf{X}_1 | \mathbf{X}_0) + I(\mathbf{Y}_2; \mathbf{X}_2, \mathbf{X}_0), I(\mathbf{Y}_1; \mathbf{X}_1, \mathbf{X}_0) + I(\mathbf{Y}_2; \mathbf{X}_2 | \mathbf{X}_0)\}, & \text{(4.77c)} \end{cases}$$

for input distributions $p_{\mathbf{X}_0}$, $p_{\mathbf{X}_1}$, *and* $p_{\mathbf{X}_2}$ *satisfying the power constraint*

$$\mathbb{E}\Big[\|\mathbf{X}_0\|_{\mathrm{F}}^2 + \|\mathbf{X}_1\|_{\mathrm{F}}^2 + \|\mathbf{X}_2\|_{\mathrm{F}}^2\Big] \leq PT.$$

This achievable rate region is fully characterized by the mutual information terms $I(\mathbf{Y}_k; \mathbf{X}_k, \mathbf{X}_0)$, $I(\mathbf{Y}_k; \mathbf{X}_k \,|\, \mathbf{X}_0)$, and $I(\mathbf{Y}_k; \mathbf{X}_0 \,|\, \mathbf{X}_k)$, $k \in \{1,2\}$. By considering an explicit input distribution and deriving these mutual information, we obtain an achievable rate region as follows.

Theorem 4.4. *With rate splitting, the two-user noncoherent BC can achieve any rate pair* (R_1, R_2) *satisfying*

$$\begin{cases} R_1 & \le \min\{R_1', R_1^p + R_{02}\}, \quad (4.78a) \\ R_2 & \le \min\{R_2', R_2^p + R_{01}\}, \quad (4.78b) \\ R_1 + R_2 & \le \min\{R_1^p + R_2', R_1' + R_2^p\}, \quad (4.78c) \end{cases}$$

where

$$R_1' = \Big(1 - \frac{s_1+s_0}{T}\Big)\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_1} + \frac{1}{\mathrm{tr}\Big((\bar{\boldsymbol{R}}_1^{-1}+\boldsymbol{P}_{1\tau})^{-1}\boldsymbol{P}_{1\delta}\Big)+1}\bar{\boldsymbol{\Omega}}_1\bar{\boldsymbol{R}}_1\boldsymbol{P}_{1\delta}\bar{\boldsymbol{R}}_1\bar{\boldsymbol{\Omega}}_1^{\mathsf{H}}\right)\right], \quad (4.79)$$

$$R_1^p = \Big(1 - \frac{s_1+s_0}{T}\Big)\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_1} + \frac{P_{1\delta}}{s_1\Big[\mathrm{tr}\Big((\bar{\boldsymbol{R}}_1^{-1}+\boldsymbol{P}_{1\tau})^{-1}\boldsymbol{P}_{1\delta}\Big)+1\Big]}\bar{\boldsymbol{\Omega}}_1\bar{\boldsymbol{R}}_{11}\bar{\boldsymbol{R}}_{11}^{\mathsf{H}}\bar{\boldsymbol{\Omega}}_1^{\mathsf{H}}\right)\right], \quad (4.80)$$

$$R_{01} = \Big(1 - \frac{s_1+s_0}{T}\Big)\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_1} + \frac{P_{0\delta}}{s_0\Big[\mathrm{tr}\Big((\bar{\boldsymbol{R}}_1^{-1}+\boldsymbol{P}_{1\tau})^{-1}\boldsymbol{P}_{1\delta}\Big)+1\Big]}\bar{\boldsymbol{\Omega}}_1\bar{\boldsymbol{R}}_{10}\bar{\boldsymbol{R}}_{10}^{\mathsf{H}}\bar{\boldsymbol{\Omega}}_1^{\mathsf{H}}\right)\right], \quad (4.81)$$

where

$$\boldsymbol{P}_{1\tau} := \begin{bmatrix} P_{0\tau}\boldsymbol{I}_{s_0} & \boldsymbol{0} \\ \boldsymbol{0} & P_{1\tau}\boldsymbol{I}_{s_1}\end{bmatrix}, \quad \boldsymbol{P}_{1\delta} := \begin{bmatrix} \frac{P_{0\delta}}{s_0}\boldsymbol{I}_{s_0} & \boldsymbol{0} \\ \boldsymbol{0} & \frac{P_{1\delta}}{s_1}\boldsymbol{I}_{s_1}\end{bmatrix},$$

and $\bar{\boldsymbol{\Omega}}_1 \in \mathbb{C}^{N_1\times(s_0+s_1)}$ *is a Gaussian matrix with independent rows following*

$$\mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{0}, \boldsymbol{P}_{1\tau}^{\frac{1}{2}}\Big(\boldsymbol{P}_{1\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_1\boldsymbol{P}_{1\tau}^{\frac{1}{2}} + \boldsymbol{I}_{s_1+s_0}\Big)^{-1}\boldsymbol{P}_{1\tau}^{\frac{1}{2}}\Big);$$

$$\begin{aligned} R_2' = {} & \frac{s_1-s_2}{T}\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_2} + \frac{P_{2\delta}}{P_{2\delta}\mathrm{tr}\Big(\bar{\boldsymbol{R}}_{22}^{\mathsf{H}}(\bar{\boldsymbol{R}}_2+\bar{\boldsymbol{R}}_2\boldsymbol{P}_{2\tau}\bar{\boldsymbol{R}}_2)^{-1}\bar{\boldsymbol{R}}_{22}\Big)+s_2}\bar{\boldsymbol{\Omega}}_2\bar{\boldsymbol{R}}_{22}\bar{\boldsymbol{R}}_{22}^{\mathsf{H}}\bar{\boldsymbol{\Omega}}_2^{\mathsf{H}}\right)\right] \\ & + \Big(1 - \frac{s_1+s_0}{T}\Big)\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_2} + \frac{1}{\mathrm{tr}\Big((\bar{\boldsymbol{R}}_2^{-1}+\boldsymbol{P}_{2\tau})^{-1}\boldsymbol{P}_{2\delta}\Big)+1}\bar{\boldsymbol{\Omega}}_2\bar{\boldsymbol{R}}_2\boldsymbol{P}_{2\delta}\bar{\boldsymbol{R}}_2\bar{\boldsymbol{\Omega}}_2^{\mathsf{H}}\right)\right], \end{aligned} \quad (4.82)$$

$$\begin{aligned} R_2^p = {} & \frac{s_1-s_2}{T}\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_2} + \frac{P_{2\delta}}{P_{2\delta}\mathrm{tr}\Big(\bar{\boldsymbol{R}}_{22}^{\mathsf{H}}(\bar{\boldsymbol{R}}_2+\bar{\boldsymbol{R}}_2\boldsymbol{P}_{2\tau}\bar{\boldsymbol{R}}_2)^{-1}\bar{\boldsymbol{R}}_{22}\Big)+s_2}\bar{\boldsymbol{\Omega}}_2\bar{\boldsymbol{R}}_2\bar{\boldsymbol{R}}_{22}\bar{\boldsymbol{R}}_{22}^{\mathsf{H}}\bar{\boldsymbol{R}}_2\bar{\boldsymbol{\Omega}}_2^{\mathsf{H}}\right)\right] \\ & + \Big(1 - \frac{s_1+s_0}{T}\Big)\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_2} + \frac{P_{2\delta}}{s_2\Big[\mathrm{tr}\Big((\bar{\boldsymbol{R}}_2^{-1}+\boldsymbol{P}_{2\tau})^{-1}\boldsymbol{P}_{2\delta}\Big)+1\Big]}\bar{\boldsymbol{\Omega}}_2\bar{\boldsymbol{R}}_2\bar{\boldsymbol{R}}_{22}\bar{\boldsymbol{R}}_{22}^{\mathsf{H}}\bar{\boldsymbol{R}}_2\bar{\boldsymbol{\Omega}}_2^{\mathsf{H}}\right)\right], \end{aligned} \quad (4.83)$$

$$R_{02} = \Big(1 - \frac{s_1+s_0}{T}\Big)\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_2} + \frac{P_{0\delta}}{s_0\Big[\mathrm{tr}\Big((\bar{\boldsymbol{R}}_2^{-1}+\boldsymbol{P}_{2\tau})^{-1}\boldsymbol{P}_{2\delta}\Big)+1\Big]}\bar{\boldsymbol{\Omega}}_2\bar{\boldsymbol{R}}_2\bar{\boldsymbol{R}}_{20}\bar{\boldsymbol{R}}_{20}^{\mathsf{H}}\bar{\boldsymbol{R}}_2\bar{\boldsymbol{\Omega}}_2^{\mathsf{H}}\right)\right], \quad (4.84)$$

where

$$\boldsymbol{P}_{2\tau} := \begin{bmatrix} P_{0\tau}\boldsymbol{I}_{s_0} & \mathbf{0} \\ \mathbf{0} & P_{2\tau}\boldsymbol{I}_{s_2} \end{bmatrix}, \quad \boldsymbol{P}_{2\delta} := \begin{bmatrix} \frac{P_{0\delta}}{s_0}\boldsymbol{I}_{s_0} & \mathbf{0} \\ \mathbf{0} & \frac{P_{2\delta}}{s_2}\boldsymbol{I}_{s_2} \end{bmatrix},$$

and $\bar{\boldsymbol{\Omega}}_2 \in \mathbb{C}^{N_2\times(s_0+s_2)}$ *is a Gaussian matrix with independent rows following*

$$\mathcal{N}_{\mathbb{C}}\Big(\mathbf{0}, \boldsymbol{P}_{2\tau}^{\frac{1}{2}}\Big(\boldsymbol{P}_{2\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_2\boldsymbol{P}_{2\tau}^{\frac{1}{2}} + \boldsymbol{I}_{s_2+s_0}\Big)^{-1}\boldsymbol{P}_{2\tau}^{\frac{1}{2}}\Big);$$

s_0, s_1, s_2 are integers such that $s_0 \le r_0$, $s_1 \le r_1 - r_0$, and $s_2 \le r_2 - r_0$; and the power components $P_{i\tau}, P_\delta$, $i \in \{0,1,2\}$, satisfy the power constraint

$$P_{0\tau}s_0 + P_{0\delta}(T - s_1 - s_0) + \sum_{i=1}^{2}\left[P_{i\tau}s_i + P_{i\delta}(T - s_i - s_0)\right] \le PT. \tag{4.85}$$

The convex hull of (4.78a), (4.78b), (4.78c) *over all feasible values of s_0, s_1, s_2 and all possible power allocations* (4.85) *is an achievable rate region for the two-user BC.*

Specifically, the input distribution that achieves the rate region in Theorem 4.4 is characterized by

$$\mathbf{X}_0 = \left[\sqrt{P_{0\tau}}\boldsymbol{I}_{s_0}\ \mathbf{0}_{s_0\times s_1}\ \sqrt{\frac{P_{0\delta}}{s_0}}\mathbf{S}_0\right], \tag{4.86}$$

$$\mathbf{X}_1 = \left[\mathbf{0}_{s_1\times s_0}\ \sqrt{P_{1\tau}}\boldsymbol{I}_{s_1}\ \sqrt{\frac{P_{1\delta}}{s_1}}\mathbf{S}_1\right], \tag{4.87}$$

$$\mathbf{X}_2 = \left[\mathbf{0}_{s_2\times s_0}\ \sqrt{P_{2\tau}}\boldsymbol{I}_{s_2}\ \sqrt{\frac{P_{2\delta}}{s_2}}\mathbf{S}_2\right], \tag{4.88}$$

where $\mathbf{S}_0 \in \mathbb{C}^{s_0\times(T-s_1-s_0)}$, $\mathbf{S}_1 \in \mathbb{C}^{s_1\times(T-s_1-s_0)}$, and $\mathbf{S}_2 \in \mathbb{C}^{s_2\times(T-s_2-s_0)}$ are data matrices containing independent $\mathcal{N}_{\mathbb{C}}(0,1)$ symbols and the power components $P_{i\tau}, P_\delta$, $i \in \{0,1,2\}$ satisfy the power constraint (4.85). In the following, we derive the mutual information terms with this input distribution and show that $I(\mathbf{Y}_k;\mathbf{X}_k,\mathbf{X}_0) \ge TR'_k$, $I(\mathbf{Y}_k;\mathbf{X}_k\,|\,\mathbf{X}_0) \ge TR^p_k$, and $I(\mathbf{Y}_k;\mathbf{X}_0\,|\,\mathbf{X}_k) \ge TR_{0k}$, $k \in \{1,2\}$, thus prove Theorem 4.4.

4.4.3.a Rates of User 1

The received signal at user 1 is

$$\mathbf{Y}_1 = \mathbf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_1\begin{bmatrix}\sqrt{P_{0\tau}}\boldsymbol{I}_{s_0} & \mathbf{0} & \sqrt{\frac{P_{0\delta}}{s_0}}\mathbf{S}_0 \\ \mathbf{0} & \sqrt{P_{1\tau}}\boldsymbol{I}_{s_1} & \sqrt{\frac{P_{1\delta}}{s_1}}\mathbf{S}_1\end{bmatrix} + \mathbf{Z}_1 \tag{4.89}$$

$$= \Bigg[\underbrace{\mathbf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_1\boldsymbol{P}_{1\tau}^{\frac{1}{2}} + \mathbf{Z}_{1[1:s_1+s_0]}}_{\mathbf{Y}_{1\tau}} \quad \underbrace{\mathbf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_1\boldsymbol{P}_{1\delta}^{\frac{1}{2}}\begin{bmatrix}\mathbf{S}_0\\ \mathbf{S}_1\end{bmatrix} + \mathbf{Z}_{1[s_1+s_0+1;T]}}_{\mathbf{Y}_{1\delta}}\Bigg] \tag{4.90}$$

where $\boldsymbol{P}_{1\tau} := \begin{bmatrix} P_{0\tau}\boldsymbol{I}_{s_0} & \mathbf{0} \\ \mathbf{0} & P_{1\tau}\boldsymbol{I}_{s_1}\end{bmatrix}$ and $\boldsymbol{P}_{1\delta} := \begin{bmatrix} \frac{P_{0\delta}}{s_0}\boldsymbol{I}_{s_0} & \mathbf{0} \\ \mathbf{0} & \frac{P_{1\delta}}{s_1}\boldsymbol{I}_{s_1}\end{bmatrix}$ are the power matrices for the pilot and data, respectively. The equivalent channel $\boldsymbol{\Omega}_1 := \mathbf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_1$ has correlation matrix $\bar{\boldsymbol{R}}_1$. Following Lemma 4.2, user 1 performs a MMSE channel estimation based on $\mathbf{Y}_{1\tau}$ as

$$\hat{\boldsymbol{\Omega}}_1 = \mathbf{Y}_{1\tau}\big(\boldsymbol{P}_{1\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_1\boldsymbol{P}_{1\tau}^{\frac{1}{2}} + \boldsymbol{I}_{s_1+s_0}\big)^{-1}\boldsymbol{P}_{1\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_1. \tag{4.91}$$

The estimate $\hat{\boldsymbol{\Omega}}_1$ and the estimation error $\tilde{\boldsymbol{\Omega}}_1 = \mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_1 - \hat{\boldsymbol{\Omega}}_1$ have zero mean and row covariance

$$\frac{1}{N}\mathbb{E}\Big[\hat{\boldsymbol{\Omega}}_1^{\mathsf{H}}\hat{\boldsymbol{\Omega}}_1\Big] = \bar{\boldsymbol{R}}_1\boldsymbol{P}_{1\tau}^{\frac{1}{2}}\Big(\boldsymbol{P}_{1\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_1\boldsymbol{P}_{1\tau}^{\frac{1}{2}} + \boldsymbol{I}_{s_1+s_0}\Big)^{-1}\boldsymbol{P}_{1\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_1, \tag{4.92}$$

$$\frac{1}{N}\mathbb{E}\Big[\tilde{\boldsymbol{\Omega}}_1^{\mathsf{H}}\tilde{\boldsymbol{\Omega}}_1\Big] = \bar{\boldsymbol{R}}_1 - \bar{\boldsymbol{R}}_1\boldsymbol{P}_{1\tau}^{\frac{1}{2}}\Big(\boldsymbol{P}_{1\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_1\boldsymbol{P}_{1\tau}^{\frac{1}{2}} + \boldsymbol{I}_{s_1+s_0}\Big)^{-1}\boldsymbol{P}_{1\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_1 = \Big(\bar{\boldsymbol{R}}_1^{-1} + \boldsymbol{P}_{1\tau}\Big)^{-1}. \tag{4.93}$$

A Lower Bound on $I(\mathsf{Y}_1;\mathsf{X}_1,\mathsf{X}_0)$:

The received signal during the data transmission phase can be written as

$$\mathsf{Y}_{1\delta} = \hat{\mathsf{G}}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_1\boldsymbol{P}_{1\delta}^{\frac{1}{2}}\begin{bmatrix}\mathsf{S}_0\\ \mathsf{S}_1\end{bmatrix} + \mathsf{Z}_{1\delta}, \tag{4.94}$$

where $\mathsf{Z}_{1\delta} := \tilde{\boldsymbol{\Omega}}_1\boldsymbol{P}_{1\delta}^{\frac{1}{2}}\begin{bmatrix}\mathsf{S}_0\\ \mathsf{S}_1\end{bmatrix} + \mathsf{Z}_{1[s_1+s_0+1:T]}$ is the combined noise and residual interference due to channel estimation error. By a similar analysis using Lemma 4.1 as for (4.41) in Theorem 4.2, we have

$$I(\mathsf{Y}_1;\mathsf{X}_1,\mathsf{X}_0) = I(\mathsf{Y}_{1\delta};\mathsf{S}_1,\mathsf{S}_0\,|\,\mathsf{Y}_{1\tau}) + \underbrace{I(\mathsf{Y}_{1\tau};\mathsf{S}_1,\mathsf{S}_0)}_{=0} \tag{4.95}$$

$$= I(\mathsf{Y}_{1\delta};\mathsf{S}_1,\mathsf{S}_0\,|\,\hat{\boldsymbol{\Omega}}_1) \tag{4.96}$$

$$\geq TR_1', \tag{4.97}$$

where R_1' is given in (4.79).

A Lower Bound on $I(\mathsf{Y}_1;\mathsf{X}_1\,|\,\mathsf{X}_0)$:

We rewrite $\mathsf{Y}_{1\delta}$ as

$$\mathsf{Y}_{1\delta} = \sqrt{\frac{P_{1\delta}}{s_1}}\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{11}\mathsf{S}_1 + \sqrt{\frac{P_{0\delta}}{s_0}}\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{10}\mathsf{S}_0 + \mathsf{Z}_{1\delta}. \tag{4.98}$$

While decoding S_1, the term $\sqrt{\frac{P_{0\delta}}{s_0}}\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{10}\mathsf{S}_0$ is an interference. Given the knowledge of S_0 and the channel estimate $\hat{\boldsymbol{\Omega}}_1 = [\hat{\boldsymbol{\Omega}}_{10}\ \hat{\boldsymbol{\Omega}}_{11}]$, where $\hat{\boldsymbol{\Omega}}_{10}$ and $\hat{\boldsymbol{\Omega}}_{11}$ are respectively the estimates of $\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{10}$ and $\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{11}$, the receiver can partly remove the interference to obtain

$$\mathsf{Y}_{1\delta} - \sqrt{\frac{P_{0\delta}}{s_0}}\hat{\boldsymbol{\Omega}}_{10}\mathsf{S}_0 = \sqrt{\frac{P_{1\delta}}{s_1}}\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{11}\mathsf{S}_1 + \sqrt{\frac{P_{0\delta}}{s_0}}\Big[\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{10} - \hat{\boldsymbol{\Omega}}_{10}\Big]\mathsf{S}_0 + \mathsf{Z}_{1[s_1+s_0+1:T]} \tag{4.99}$$

$$= \sqrt{\frac{P_{1\delta}}{s_1}}\hat{\boldsymbol{\Omega}}_{11}\mathsf{S}_1 + \mathsf{Z}_{1\delta}. \tag{4.100}$$

With a similar analysis using Lemma 4.1 as for (4.41) in Theorem 4.2, we have the bound

$$\begin{aligned} I(\mathbf{Y}_1;\mathbf{X}_1\,|\,\mathbf{X}_0) &= I(\mathbf{Y}_{1\delta};\mathbf{S}_1\,|\,\mathbf{S}_0,\mathbf{Y}_{1\tau}) && (4.101)\\ &= I(\mathbf{Y}_{1\delta};\mathbf{S}_1\,|\,\mathbf{S}_0,\hat{\mathbf{\Omega}}_1) && (4.102)\\ &= I\left(\mathbf{Y}_{1\delta}-\sqrt{\frac{P_{0\delta}}{s_0}}\hat{\mathbf{\Omega}}_{10}\mathbf{S}_0;\mathbf{S}_1 \;\middle|\; \mathbf{S}_0,\hat{\mathbf{\Omega}}_1\right) && (4.103)\\ &= I\left(\sqrt{\frac{P_{1\delta}}{s_1}}\hat{\mathbf{\Omega}}_{11}\mathbf{S}_1+\mathbf{Z}_{1\delta};\mathbf{S}_1 \;\middle|\; \hat{\mathbf{\Omega}}_{11}\right) && (4.104)\\ &\geq TR_1^p, && (4.105) \end{aligned}$$

where R_1^p is given in (4.80).

A Lower Bound on $I(\mathbf{Y}_1;\mathbf{X}_0\,|\,\mathbf{X}_1)$:

Given $\mathbf{S}_1$ and the channel estimate $\hat{\mathbf{\Omega}}_1=[\hat{\mathbf{\Omega}}_{10}\quad\hat{\mathbf{\Omega}}_{11}]$, the receiver can remove partly the interference in (4.98) to obtain

$$\mathbf{Y}_{1\delta}-\sqrt{\frac{P_{1\delta}}{s_1}}\hat{\mathbf{\Omega}}_{11}\mathbf{S}_1=\sqrt{\frac{P_{0\delta}}{s_0}}\mathbf{G}_1\mathbf{\Sigma}_1^{\frac{1}{2}}\mathbf{\Phi}_{10}\mathbf{S}_0+\sqrt{\frac{P_{1\delta}}{s_1}}\left[\mathbf{G}_1\mathbf{\Sigma}_1^{\frac{1}{2}}\mathbf{\Phi}_{11}-\hat{\mathbf{\Omega}}_{11}\mathbf{S}_1\right]+\mathbf{Z}_{1[s_1+s_0+1:T]} \tag{4.106}$$

$$=\sqrt{\frac{P_{0\delta}}{s_0}}\hat{\mathbf{\Omega}}_{10}\mathbf{S}_0+\mathbf{Z}_{1\delta}. \tag{4.107}$$

With a similar analysis as for (4.41) in Theorem 4.2, we have the bound

$$\begin{aligned} I(\mathbf{Y}_1;\mathbf{X}_0\,|\,\mathbf{X}_1) &= I(\mathbf{Y}_{1\delta};\mathbf{S}_0\,|\,\mathbf{S}_1,\mathbf{Y}_{1\tau}) && (4.108)\\ &= I\left(\mathbf{Y}_{1\delta};\mathbf{S}_0\,|\,\mathbf{S}_1,\hat{\mathbf{\Omega}}_1\right) && (4.109)\\ &= I\left(\mathbf{Y}_{1\delta}-\sqrt{\frac{P_{1\delta}}{s_1}}\hat{\mathbf{\Omega}}_{11}\mathbf{S}_1;\mathbf{S}_0 \;\middle|\; \mathbf{S}_1,\hat{\mathbf{\Omega}}_1\right) && (4.110)\\ &= I\left(\sqrt{\frac{P_{0\delta}}{s_0}}\hat{\mathbf{\Omega}}_{10}\mathbf{S}_0+\mathbf{Z}_{1\delta};\mathbf{S}_0 \;\middle|\; \hat{\mathbf{\Omega}}_{10}\right) && (4.111)\\ &\geq TR_{01}, && (4.112) \end{aligned}$$

where R_{01} is given in (4.81).

4.4.3.b Rates of User 2

The received signal at user 2 is

$$\mathbf{Y}_2=\mathbf{G}_2\mathbf{\Sigma}_2^{\frac{1}{2}}\mathbf{\Phi}_2\begin{bmatrix}\sqrt{P_{0\tau}}\boldsymbol{I}_{s_0} & \mathbf{0} & \mathbf{0}_{s_0\times(s_1-s_2)} & \sqrt{\frac{P_{0\delta}}{s_0}}\mathbf{S}_0\\ \mathbf{0} & \sqrt{P_{2\tau}}\boldsymbol{I}_{s_2} & \sqrt{\frac{P_{2\delta}}{s_2}}\mathbf{S}_{2a} & \sqrt{\frac{P_{2\delta}}{s_2}}\mathbf{S}_{2b}\end{bmatrix}+\mathbf{Z}_2 \tag{4.113}$$

$$=\Bigg[\underbrace{\mathbf{G}_2\mathbf{\Sigma}_2^{\frac{1}{2}}\mathbf{\Phi}_2\boldsymbol{P}_{2\tau}^{\frac{1}{2}}+\mathbf{Z}_{2[1:s_2+s_0]}}_{\mathbf{Y}_{2\tau}}\quad\underbrace{\sqrt{\frac{P_{2\delta}}{s_2}}\mathbf{G}_2\mathbf{\Sigma}_2^{\frac{1}{2}}\mathbf{\Phi}_{22}\mathbf{S}_{2a}+\mathbf{Z}_{2[s_2+s_0+1:s_1+s_0]}}_{\mathbf{Y}_{2\delta a}}$$

$$\underbrace{\mathbf{G}_2\mathbf{\Sigma}_2^{\frac{1}{2}}\mathbf{\Phi}_2\boldsymbol{P}_{2\delta}^{\frac{1}{2}}\begin{bmatrix}\mathbf{S}_0\\ \mathbf{S}_{2b}\end{bmatrix}+\mathbf{Z}_{2[s_1+s_0+1:T]}}_{\mathbf{Y}_{2\delta b}}\Bigg], \tag{4.114}$$

where S_{2a} and S_{2b} contain respectively the first s_1-s_2 columns and the remaining $T-s_1-s_0$ columns of S_2; $\boldsymbol{P}_{2\tau} := \begin{bmatrix} P_{0\tau}\boldsymbol{I}_{s_0} & \mathbf{0} \\ \mathbf{0} & P_{2\tau}\boldsymbol{I}_{s_2} \end{bmatrix}$ and $\boldsymbol{P}_{2\delta} := \begin{bmatrix} \frac{P_{0\delta}}{s_0}\boldsymbol{I}_{s_0} & \mathbf{0} \\ \mathbf{0} & \frac{P_{2\delta}}{s_2}\boldsymbol{I}_{s_2} \end{bmatrix}$ are the power matrices for the pilot and data, respectively. Following Lemma 4.2, user 2 performs a MMSE channel estimation of $\boldsymbol{\Omega}_2 := \mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_2 = [\boldsymbol{\Omega}_{20}\ \boldsymbol{\Omega}_{22}] = \left[\mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_{20}\ \mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_{22}\right]$ based on $\mathsf{Y}_{2\tau}$ as

$$\hat{\boldsymbol{\Omega}}_2 = \mathsf{Y}_{2\tau}\left(\boldsymbol{P}_{2\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_2\boldsymbol{P}_{2\tau}^{\frac{1}{2}} + \boldsymbol{I}_{s_2+s_0}\right)^{-1}\boldsymbol{P}_{2\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_2. \tag{4.115}$$

The estimate $\hat{\boldsymbol{\Omega}}_2 = [\hat{\boldsymbol{\Omega}}_{20}\quad \hat{\boldsymbol{\Omega}}_{22}]$ and the estimation error $\tilde{\boldsymbol{\Omega}}_2 = \mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_2 - \hat{\boldsymbol{\Omega}}_2$ have zero mean and row covariance

$$\frac{1}{N}\mathbb{E}\left[\hat{\boldsymbol{\Omega}}_2^{\mathsf{H}}\hat{\boldsymbol{\Omega}}_2\right] = \bar{\boldsymbol{R}}_2\boldsymbol{P}_{2\tau}^{\frac{1}{2}}\left(\boldsymbol{P}_{2\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_2\boldsymbol{P}_{2\tau}^{\frac{1}{2}} + \boldsymbol{I}_{s_2+s_0}\right)^{-1}\boldsymbol{P}_{2\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_2, \tag{4.116}$$

$$\frac{1}{N}\mathbb{E}\left[\tilde{\boldsymbol{\Omega}}_2^{\mathsf{H}}\tilde{\boldsymbol{\Omega}}_2\right] = \bar{\boldsymbol{R}}_2 - \bar{\boldsymbol{R}}_2\boldsymbol{P}_{2\tau}^{\frac{1}{2}}\left(\boldsymbol{P}_{2\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_2\boldsymbol{P}_{2\tau}^{\frac{1}{2}} + \boldsymbol{I}_{s_2+s_0}\right)^{-1}\boldsymbol{P}_{2\tau}^{\frac{1}{2}}\bar{\boldsymbol{R}}_2 = \left(\bar{\boldsymbol{R}}_2^{-1} + \boldsymbol{P}_{2\tau}\right)^{-1}. \tag{4.117}$$

A Lower Bound on $I(\mathsf{Y}_2;\mathsf{X}_2,\mathsf{X}_0)$:

Using the chain rule, we have that

$$\begin{aligned}
I(\mathsf{Y}_2;\mathsf{X}_2,\mathsf{X}_0) &= I(\mathsf{Y}_{2\tau},\mathsf{Y}_{2\delta a},\mathsf{Y}_{2\delta b};\mathsf{S}_0,\mathsf{S}_{2a},\mathsf{S}_{2b}) && (4.118)\\
&= I(\mathsf{Y}_{2\delta a},\mathsf{Y}_{2\delta b};\mathsf{S}_0,\mathsf{S}_{2a},\mathsf{S}_{2b}\,|\,\mathsf{Y}_{2\tau}) + \underbrace{I(\mathsf{Y}_{2\tau};\mathsf{S}_0,\mathsf{S}_{2a},\mathsf{S}_{2b})}_{=0} && (4.119)\\
&= I(\mathsf{Y}_{2\delta a},\mathsf{Y}_{2\delta b};\mathsf{S}_0,\mathsf{S}_{2a},\mathsf{S}_{2b}\,|\,\hat{\boldsymbol{\Omega}}_2) && (4.120)\\
&= I(\mathsf{Y}_{2\delta a};\mathsf{S}_{2a}\,|\,\hat{\boldsymbol{\Omega}}_2) + \underbrace{I(\mathsf{Y}_{2\delta a};\mathsf{S}_0,\mathsf{S}_{2b}\,|\,\mathsf{S}_{2a},\hat{\boldsymbol{\Omega}}_2)}_{=0}\\
&\quad + \underbrace{I(\mathsf{Y}_{2\delta b};\mathsf{S}_0,\mathsf{S}_{2b}\,|\,\mathsf{Y}_{2\delta a},\hat{\boldsymbol{\Omega}}_2)}_{\geq I(\mathsf{Y}_{2\delta b};\mathsf{S}_0,\mathsf{S}_{2b}\,|\,\hat{\boldsymbol{\Omega}}_2)} + \underbrace{I(\mathsf{Y}_{2\delta b};\mathsf{S}_{2a}\,|\,\mathsf{S}_0,\mathsf{S}_{2b},\mathsf{Y}_{2\delta a},\hat{\boldsymbol{\Omega}}_2)}_{=0} && (4.121)\\
&\geq I(\mathsf{Y}_{2\delta a};\mathsf{S}_{2a}\,|\,\hat{\boldsymbol{\Omega}}_{22}) + I(\mathsf{Y}_{2\delta b};\mathsf{S}_0,\mathsf{S}_{2b}\,|\,\hat{\boldsymbol{\Omega}}_2). && (4.122)
\end{aligned}$$

Following a similar analysis as for (4.41) in Theorem 4.2, we lower bound $\frac{1}{T}I(\mathsf{Y}_{2\delta a};\mathsf{S}_{2a}\,|\,\hat{\boldsymbol{\Omega}}_{22})$ and $\frac{1}{T}I(\mathsf{Y}_{2\delta b};\mathsf{S}_0,\mathsf{S}_{2b}\,|\,\hat{\boldsymbol{\Omega}}_2)$ respectively by the first and the second terms in the right-hand side of (4.82). Therefore, $I(\mathsf{Y}_2;\mathsf{X}_2,\mathsf{X}_0) \geq TR_2'$ with R_2' given in (4.82).

A Lower Bound on $I(\mathsf{Y}_2;\mathsf{X}_2\,|\,\mathsf{X}_0)$:

We write $\mathsf{Y}_{2\delta} := [\mathsf{Y}_{2\delta a}\ \mathsf{Y}_{2\delta b}]$ as

$$\mathsf{Y}_{2\delta} = \sqrt{\frac{P_{2\delta}}{s_2}}\mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_{22}\mathsf{S}_2 + \sqrt{\frac{P_{0\delta}}{s_0}}\mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_{20}[\mathbf{0}\ \mathsf{S}_0] + \mathsf{Z}_{2[s_2+s_0+1:T]}. \tag{4.123}$$

Similar as for $I(\mathsf{Y}_1;\mathsf{X}_1\,|\,\mathsf{X}_0)$, using interference cancellation and worst-case additive noise as for (4.41) in Theorem 4.2, we have the following bound

$$\begin{aligned}
I(\mathsf{Y}_2;\mathsf{X}_2|\mathsf{X}_0) &= I\left(\mathsf{Y}_{2\delta};\mathsf{S}_2\middle|\mathsf{S}_0,\hat{\boldsymbol{\Omega}}_2\right) && (4.124)\\
&= I\left(\mathsf{Y}_{2\delta} - \sqrt{\frac{P_{0\delta}}{s_0}}\hat{\boldsymbol{\Omega}}_{20}[\mathbf{0}\ \mathsf{S}_0];\mathsf{S}_2\ \middle|\ \mathsf{S}_0,\hat{\boldsymbol{\Omega}}_2\right) && (4.125)\\
&\geq TR_2^p, && (4.126)
\end{aligned}$$

where R_2^p is given in (4.83).

A Lower Bound on $I(\mathbf{Y}_2;\mathbf{X}_0\,|\,\mathbf{X}_2)$*:*

Again, using interference cancellation and a similar analysis as for (4.41) in Theorem 4.2, we have the following bound

$$I(\mathbf{Y}_2;\mathbf{X}_0|\mathbf{X}_2) \geq I\Big(\mathbf{Y}_{2\delta b};\mathbf{S}_0\Big|\mathbf{S}_{2b},\hat{\mathbf{\Omega}}_2\Big) \tag{4.127}$$

$$= I\left(\mathbf{Y}_{2\delta b}-\sqrt{\frac{P_{2\delta}}{s_2}}\hat{\mathbf{\Omega}}_{22}\mathbf{S}_{2b};\mathbf{S}_0\;\middle|\;\mathbf{S}_{2b},\hat{\mathbf{\Omega}}_2\right) \tag{4.128}$$

$$\geq TR_{02} \tag{4.129}$$

where R_{02} is given in (4.84).

4.4.4 Product Superposition

Theorem 4.5. *With product superposition, the two-user noncoherent BC can achieve any rate pair* (R_1,R_2) *of the form*

$$R_1=\frac{s_2}{T}\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_1}+\frac{\rho_{1\delta}P_{2\tau}}{s_0+\rho_{1\delta}P_{2\tau}\mathrm{tr}\Big(\Big(\breve{\boldsymbol{R}}_{k0}^{-1}+\rho_{1\tau}P_{2\tau}\boldsymbol{I}_{s_0}\Big)^{-1}\Big)}\hat{\mathbf{\Omega}}_{10}\hat{\mathbf{\Omega}}_{10}^{\mathsf{H}}\right)\right], \tag{4.130}$$

where $\hat{\mathbf{\Omega}}_{10}\in\mathbb{C}^{N_1\times s_0}$ *is a Gaussian matrix with* N_1 *independent rows following*

$$\mathcal{N}_{\mathbb{C}}\Big(\mathbf{0},\rho_{1\tau}P_{2\tau}\breve{\boldsymbol{R}}_{10}\big(\rho_{1\tau}P_{2\tau}\breve{\boldsymbol{R}}_{k0}+\boldsymbol{I}_{s_0}\big)^{-1}\breve{\boldsymbol{R}}_{k0}\Big),$$

and

$$R_2=\left(1-\frac{s_2+s_0}{T}\right)\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_2}+\frac{P_{2\delta}}{s_2+s_0+P_{2\delta}\mathrm{tr}\Big(\big(\boldsymbol{R}_{2e}^{-1}+P_{2\tau}\boldsymbol{I}_{s_2+s_0}\big)^{-1}\Big)}\hat{\mathbf{G}}_{2e}\hat{\mathbf{G}}_{2e}^{\mathsf{H}}\right)\right] \tag{4.131}$$

where $\hat{\mathbf{G}}_{2e}\in\mathbb{C}^{N_2\times(s_2+s_0)}$ *has distribution imposed by*

$$\hat{\mathbf{G}}_{2e}=\sqrt{P_{2\tau}}\Big(\sqrt{P_{2\tau}}\mathbf{G}_2\mathbf{\Sigma}_2^{\frac{1}{2}}\mathbf{\Phi}_2\mathbf{X}_1+\mathbf{Z}_{2[1:s_2+s_0]}\Big)\big(P_{2\tau}\boldsymbol{R}_{2e}+\boldsymbol{I}_{s_2+s_0}\big)^{-1}\boldsymbol{R}_{2e} \tag{4.132}$$

for integers $s_0\leq r_0$*,* $s_2\leq r_2-r_0$ *and the power constraint*

$$(s_0\rho_{1\tau}+s_2(\rho_{1\delta}+\rho_{1a}))\Big(P_{2\tau}+\frac{T-s_2-s_0}{s_2+s_0}P_{2\delta}\Big)\leq PT. \tag{4.133}$$

In (4.132),

$$\mathbf{X}_1:=\begin{bmatrix}\sqrt{\rho_{1\tau}}\boldsymbol{I}_{s_0} & \sqrt{\frac{\rho_{1\delta}}{s_0}}\mathbf{S}_1\\ \mathbf{0} & \sqrt{\rho_{1a}}\boldsymbol{I}_{s_2}\end{bmatrix}$$

where $\mathbf{S}_1\in\mathbb{C}^{s_0\times s_2}$ *and* $\mathbf{Z}_{2[1:s_2+s_0]}\in\mathbb{C}^{N_2\times s_2+s_0}$ *both have IID* $\mathcal{N}_{\mathbb{C}}(0,1)$ *entries; and*

$$\boldsymbol{R}_{2e}:=\begin{bmatrix}\rho_{1\tau}\breve{\boldsymbol{R}}_{20} & \sqrt{\rho_{1\tau}\rho_{1a}}\mathbf{\Phi}_{20}^{\mathsf{H}}\mathbf{\Sigma}_2\mathbf{\Phi}_{22}\\ \sqrt{\rho_{1\tau}\rho_{1a}}\mathbf{\Phi}_{22}^{\mathsf{H}}\mathbf{\Sigma}_2\mathbf{\Phi}_{20} & \frac{\rho_{1\delta}}{s_0}\mathrm{tr}\Big(\breve{\boldsymbol{R}}_{20}\Big)\boldsymbol{I}_{s_2}+\rho_{1a}\breve{\boldsymbol{R}}_{22}\end{bmatrix}. \tag{4.134}$$

By swapping the users' roles, a similar rate pair is achievable. The convex hull of the origin and all these rate pairs over all feasible values of s_0,s_1,s_2 *and all feasible power allocation* (4.133) *is an achievable rate region for the two-user noncoherent BC.*

We consider the input distribution characterized by

$$\mathsf{X} = [\boldsymbol{V}_0\ \boldsymbol{V}_2]\mathsf{X}_1\mathsf{X}_2, \tag{4.135}$$

with

$$\mathsf{X}_1 = \begin{bmatrix} \sqrt{\rho_{1\tau}}\boldsymbol{I}_{s_0} & \sqrt{\frac{\rho_{1\delta}}{s_0}}\mathsf{S}_1 \\ \mathbf{0} & \sqrt{\rho_{1a}}\boldsymbol{I}_{s_2} \end{bmatrix}, \tag{4.136}$$

$$\mathsf{X}_2 = \begin{bmatrix} \sqrt{P_{2\tau}}\boldsymbol{I}_{s_2+s_0} & \sqrt{\frac{P_{2\delta}}{s_2+s_0}}\mathsf{S}_2 \end{bmatrix}, \tag{4.137}$$

where $\mathsf{S}_1 \in \mathbb{C}^{s_0\times s_2}$ and $\mathsf{S}_2 \in \mathbb{C}^{(s_2+s_0)\times(T-s_2-s_0)}$ are the data matrices of user 1 and user 2 respectively, both contain IID $\mathcal{N}_{\mathbb{C}}(0,1)$ symbols. The power constraint $\mathbb{E}\big[\mathrm{tr}(\mathsf{X}^{\mathsf{H}}\mathsf{X})\big] \le PT$ translates to (4.133). In the following, we show that the rate pair (R_1, R_2) defined by (4.130) and (4.131) can be achieved with this input distribution, thus prove Theorem 4.5.

4.4.4.a Rate of User 1

In the first $s_2 + s_0$ channel uses, user 1 receives

$$\mathsf{Y}_{1[1:s_2+s_0]} = \sqrt{P_{2\tau}}\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{10}\begin{bmatrix}\sqrt{\rho_{1\tau}}\boldsymbol{I}_{s_0} & \sqrt{\frac{\rho_{1\delta}}{s_0}}\mathsf{S}_1\end{bmatrix} + \mathsf{Z}_{1[1:s_2+s_0]} \tag{4.138}$$

$$= \Bigg[\underbrace{\sqrt{\rho_{1\tau}P_{2\tau}}\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{10} + \mathsf{Z}_{1[1:s_0]}}_{\mathsf{Y}_{1\tau}} \quad \underbrace{\sqrt{\frac{\rho_{1\delta}P_{2\tau}}{s_0}}\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{10}\mathsf{S}_1 + \mathsf{Z}_{1[s_0+1:s_2+s_0]}}_{\mathsf{Y}_{1\delta}}\Bigg]. \tag{4.139}$$

Following Lemma 4.2, user 1 estimates the equivalent channel $\boldsymbol{\Omega}_{10} = \mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{10}$ using a MMSE estimator based on $\mathsf{Y}_{1\tau}$ as

$$\hat{\boldsymbol{\Omega}}_{10} = \sqrt{\rho_{1\tau}P_{2\tau}}\mathsf{Y}_{1\tau}\Big(\rho_{1\tau}P_{2\tau}\breve{\boldsymbol{R}}_{k0} + \boldsymbol{I}_{s_0}\Big)^{-1}\breve{\boldsymbol{R}}_{k0}. \tag{4.140}$$

The estimate $\hat{\boldsymbol{\Omega}}_{10}$ and the estimation error $\tilde{\boldsymbol{\Omega}}_{10} = \mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_{10} - \hat{\boldsymbol{\Omega}}_{10}$ have zero mean and row covariance

$$\frac{1}{N}\mathbb{E}\Big[\hat{\boldsymbol{\Omega}}_{10}^{\mathsf{H}}\hat{\boldsymbol{\Omega}}_{10}\Big] = \rho_{1\tau}P_{2\tau}\breve{\boldsymbol{R}}_{k0}\Big(\rho_{1\tau}P_{2\tau}\breve{\boldsymbol{R}}_{k0} + \boldsymbol{I}_{s_0}\Big)^{-1}\breve{\boldsymbol{R}}_{k0}, \tag{4.141}$$

$$\frac{1}{N}\mathbb{E}\Big[\tilde{\boldsymbol{\Omega}}_{10}^{\mathsf{H}}\tilde{\boldsymbol{\Omega}}_{10}\Big] = \breve{\boldsymbol{R}}_{k0} - \rho_{1\tau}P_{2\tau}\breve{\boldsymbol{R}}_{k0}\Big(\rho_{1\tau}P_{2\tau}\breve{\boldsymbol{R}}_{k0} + \boldsymbol{I}_{s_0}\Big)^{-1}\breve{\boldsymbol{R}}_{k0}\Big(\breve{\boldsymbol{R}}_{k0}^{-1} + \rho_{1\tau}P_{2\tau}\boldsymbol{I}_{s_0}\Big)^{-1}. \tag{4.142}$$

Using the data processing inequality, we have that

$$I(\mathsf{Y}_1;\mathsf{X}_1) \ge I(\mathsf{Y}_{1[1:s_2+s_0]};\mathsf{X}_1) = I(\mathsf{Y}_{1\delta};\mathsf{S}_1\,|\,\mathsf{Y}_{1\tau}) = I(\mathsf{Y}_{1\delta};\mathsf{S}_1\,|\,\hat{\boldsymbol{\Omega}}_{10}). \tag{4.143}$$

Then, using the worst-case noise argument and Lemma 4.1, we have a lower bound on $I(\mathsf{Y}_{1\delta};\mathsf{S}_1\,|\,\hat{\boldsymbol{\Omega}}_{10})$, which yields the rate R_1 achievable for user 1 defined in (4.130).

4.4.4.b Rate of User 2

The received signal at user 2 is

$$\mathsf{Y}_2 = \mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_2\mathsf{X}_1\left[\sqrt{P_{2\tau}}\boldsymbol{I}_{s_2+s_0} \quad \sqrt{\frac{P_{2\delta}}{s_2+s_0}}\mathsf{S}_2\right] + \mathsf{Z}_2 \tag{4.144}$$

$$= \Bigg[\underbrace{\sqrt{P_{2\tau}}\mathsf{G}_{2e} + \mathsf{Z}_{2[1:s_2+s_0]}}_{\mathsf{Y}_{2\tau}} \quad \underbrace{\sqrt{\frac{P_{2\delta}}{s_2+s_0}}\mathsf{G}_{2e}\mathsf{S}_2 + \mathsf{Z}_{2[s_2+s_0+1:T]}}_{\mathsf{Y}_{2\delta}}\Bigg], \tag{4.145}$$

where $\mathsf{G}_{2e} := \mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_2\mathsf{X}_1$ is the equivalent channel with the correlation matrix

$$\boldsymbol{R}_{2e} := \frac{1}{N_2}\mathbb{E}\big[\mathsf{G}_{2e}^{\mathsf{H}}\mathsf{G}_{2e}\big] = \begin{bmatrix} \rho_{1\tau}\breve{\boldsymbol{R}}_{20} & \sqrt{\rho_{1\tau}\rho_{1a}}\boldsymbol{\Phi}_{20}^{\mathsf{H}}\boldsymbol{\Sigma}_2\boldsymbol{\Phi}_{22} \\ \sqrt{\rho_{1\tau}\rho_{1a}}\boldsymbol{\Phi}_{22}^{\mathsf{H}}\boldsymbol{\Sigma}_2\boldsymbol{\Phi}_{20} & \frac{\rho_{1\delta}}{s_0}\mathrm{tr}\Big(\breve{\boldsymbol{R}}_{20}\Big)\boldsymbol{I}_{s_2} + \rho_{1a}\breve{\boldsymbol{R}}_{22} \end{bmatrix}. \tag{4.146}$$

Following Lemma 4.2, user 2 estimates the equivalent channel G_{2e} using a MMSE estimator based on $\mathsf{Y}_{2\tau}$ as

$$\hat{\mathsf{G}}_{2e} = \sqrt{P_{2\tau}}\mathsf{Y}_{2\tau}(P_{2\tau}\boldsymbol{R}_{2e} + \boldsymbol{I}_{s_2+s_0})^{-1}\boldsymbol{R}_{2e}. \tag{4.147}$$

The estimate $\hat{\mathsf{G}}_{2e}$ and the estimation error $\tilde{\mathsf{G}}_{2e} = \mathsf{G}_{2e} - \hat{\mathsf{G}}_{2e}$ have zero mean and row covariance

$$\frac{1}{N_2}\mathbb{E}\Big[\hat{\mathsf{G}}_{2e}^{\mathsf{H}}\hat{\mathsf{G}}_{2e}\Big] = P_{2\tau}\boldsymbol{R}_{2e}(P_{2\tau}\boldsymbol{R}_{2e} + \boldsymbol{I}_{s_2+s_0})^{-1}\boldsymbol{R}_{2e}, \tag{4.148}$$

$$\frac{1}{N_2}\mathbb{E}\Big[\tilde{\mathsf{G}}_{2e}^{\mathsf{H}}\tilde{\mathsf{G}}_{2e}\Big] = \boldsymbol{R}_{2e} - P_{2\tau}\boldsymbol{R}_{2e}(P_{2\tau}\boldsymbol{R}_{2e} + \boldsymbol{I}_{s_2+s_0})^{-1}\boldsymbol{R}_{2e} = \Big(\boldsymbol{R}_{2e}^{-1} + P_{2\tau}\boldsymbol{I}_{s_2+s_0}\Big)^{-1}. \tag{4.149}$$

We have that

$$I(\mathsf{Y}_2;\mathsf{X}_2) = I(\mathsf{Y}_{2\delta};\mathsf{S}_2\,|\,\mathsf{Y}_{2\tau}) = I(\mathsf{Y}_{2\delta};\mathsf{S}_2\,|\,\hat{\mathsf{G}}_{2e}). \tag{4.150}$$

Then, using the worst-case noise argument and Lemma 4.1, we have the achievable rate R_2 for user 2 as defined in (4.131).

4.4.5 Hybrid Superposition

We recall that *hybrid superposition* refers to a composite scheme that involves both rate splitting and product superposition. An achievable rate region with product superposition for the two-user noncoherent BC is given as follows.

Theorem 4.6. *With hybrid superposition, the two-user noncoherent BC with $r_1 \geq r_2$ can achieve any rate pair (R_1, R_2) of the form*

$$R_1 = \Big(1 - \frac{s_1+s_0}{T}\Big)\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_2} + \frac{P_{1\delta}}{s_1+s_0+P_{1\delta}\mathrm{tr}\Big(\big(\boldsymbol{R}_{1e}^{-1}+P_{1\tau}\boldsymbol{I}_{s_1+s_0}\big)^{-1}\Big)}\hat{\mathsf{G}}_{1e}\hat{\mathsf{G}}_{1e}^{\mathsf{H}}\right)\right], \tag{4.151}$$

where $\hat{\mathsf{G}}_{1e} \in \mathbb{C}^{N_2\times(s_1+s_0)}$ has distribution imposed by

$$\hat{\mathsf{G}}_{1e} = \sqrt{P_{1\tau}}\Big(\sqrt{P_{1\tau}}\mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_1\mathsf{X}_2' + \mathsf{Z}_{1[1:s_1+s_0]}\Big)(P_{1\tau}\boldsymbol{R}_{1e} + \boldsymbol{I}_{s_1+s_0})^{-1}\boldsymbol{R}_{1e}, \tag{4.152}$$

and

$$R_2 = \frac{s_1 - s_2}{T}\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_2} + \frac{1}{\mathrm{tr}\Big(\big(\bar{\boldsymbol{R}}_2^{-1} + \boldsymbol{P}_{2\tau}\big)^{-1}\boldsymbol{P}_{2\delta a}\Big) + 1}\bar{\boldsymbol{\Omega}}_2\boldsymbol{P}_{2\delta a}\bar{\boldsymbol{\Omega}}_2^{\mathsf{H}}\right)\right] + \left(1 - \frac{s_1 + s_0}{T}\right)\mathbb{E}\left[\log_2\det\left(\boldsymbol{I}_{N_2} + \frac{1}{\mathrm{tr}\Big(\big(\bar{\boldsymbol{R}}_2^{-1} + \boldsymbol{P}_{2\tau}\big)^{-1}\boldsymbol{P}_{2\delta b}\Big) + 1}\bar{\boldsymbol{\Omega}}_2\boldsymbol{P}_{2\delta b}\bar{\boldsymbol{\Omega}}_2^{\mathsf{H}}\right)\right] - \left(1 - \frac{s_1 + s_0}{T}\right)\mathbb{E}\Big[\log_2\det\Big(\boldsymbol{I}_{N_2} + P_{1\delta}\Big(\rho_{2\tau} + \rho_{2\delta}\frac{s_1 - s_2}{s_0}\Big)\bar{\boldsymbol{\Omega}}_{20}\bar{\boldsymbol{\Omega}}_{20}^{\mathsf{H}}\Big)\Big], \tag{4.153}$$

for integers $s_0 \le r_0, s_1 \le r_1 - r_0, s_2 \le r_2 - r_0$ *and the power constraint*

$$(s_0\rho_{2\tau} + s_1\rho_{2a} + (s_1 - s_2)\rho_{2\delta})\left(P_{1\tau} + \frac{T - s_1 - s_0}{s_1 + s_0}P_{1\delta}\right) + s_2P_{2\tau} + (T - s_2 - s_0)P_{2\delta} \le PT. \tag{4.154}$$

In (4.152),

$$\mathbf{X}_2' := \begin{bmatrix}\sqrt{\rho_{2\tau}}\boldsymbol{I}_{s_0} & \left[\boldsymbol{0}_{s_0\times s_2}\ \sqrt{\frac{\rho_{2\delta}}{s_0}}\mathbf{S}_2'\right] \\ \boldsymbol{0} & \sqrt{\rho_{2a}}\boldsymbol{I}_{s_1}\end{bmatrix} \in \mathbb{C}^{(s_1+s_0)\times(s_1+s_0)}$$

where $\mathbf{S}_2' \in \mathbb{C}^{s_0\times(s_1-s_2)}$ *and* $\mathbf{Z}_{1[1:s_1+s_0]} \in \mathbb{C}^{N_1\times(s_1+s_0)}$ *both have IID* $\mathcal{N}_{\mathbb{C}}(0,1)$ *entries; and*

$$\boldsymbol{R}_{1e} := \begin{bmatrix}\rho_{2\tau}\breve{\boldsymbol{R}}_{10} & \sqrt{\rho_{2\tau}\rho_{2a}}\boldsymbol{\Phi}_{10}^{\mathsf{H}}\boldsymbol{\Sigma}_1\boldsymbol{\Phi}_{11} \\ \sqrt{\rho_{2\tau}\rho_{2a}}\boldsymbol{\Phi}_{11}^{\mathsf{H}}\boldsymbol{\Sigma}_1\boldsymbol{\Phi}_{10} & \begin{bmatrix}\boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \frac{\rho_{2\delta}}{s_0}\mathrm{tr}\big(\breve{\boldsymbol{R}}_{10}\big)\boldsymbol{I}_{s_1-s_2}\end{bmatrix} + \rho_{2a}\breve{\boldsymbol{R}}_{22}\end{bmatrix}. \tag{4.155}$$

In (4.153),

$$\boldsymbol{P}_{2\delta a} := \begin{bmatrix}\frac{\rho_{2\delta}P_{1\tau}}{s_0}\boldsymbol{I}_{s_0} & \boldsymbol{0} \\ \boldsymbol{0} & \frac{P_{2\delta}}{s_2}\boldsymbol{I}_{s_2}\end{bmatrix}, \quad \boldsymbol{P}_{2\delta b} := \begin{bmatrix}\frac{P_{1\delta}}{T-s_1-s_0}(\rho_{1\tau} + \rho_{2\delta}\frac{s_1-s_2}{s_0})\boldsymbol{I}_{s_0} & \boldsymbol{0} \\ \boldsymbol{0} & \frac{P_{2\delta}}{s_2}\boldsymbol{I}_{s_2}\end{bmatrix};$$

$\bar{\boldsymbol{\Omega}}_2 \in \mathbb{C}^{N_2\times(s_2+s_0)}$ *is a Gaussian matrix with independent rows following*

$$\mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{0}, \bar{\boldsymbol{R}}_2^{\mathsf{H}}(\bar{\boldsymbol{R}}_2 + \boldsymbol{P}_{2\tau}^{-1})^{-1}\bar{\boldsymbol{R}}_2\Big);$$

and $\bar{\boldsymbol{\Omega}}_{20} \in \mathbb{C}^{N_2\times s_0}$ *is a Gaussian matrix with independent rows following* $\mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \breve{\boldsymbol{R}}_{20})$*. The convex hull of the origin and these rate pairs with all feasible values of* s_0, s_1, s_2 *and all feasible power allocation* (4.154) *is an achievable rate region for the two-user noncoherent BC.*

We consider the the input distribution defined by

$$\mathbf{X} = [\boldsymbol{V}_0\ \boldsymbol{V}_1]\,\mathbf{X}_2'\mathbf{X}_1 + \boldsymbol{V}_2\mathbf{X}_2 \tag{4.156}$$

with

$$\mathbf{X}_1 = \left[\sqrt{P_{1\tau}}\boldsymbol{I}_{s_1+s_0}\ \sqrt{\frac{P_{1\delta}}{s_1+s_0}}\mathbf{S}_1\right] \in \mathbb{C}^{(s_1+s_0)\times T}, \tag{4.157}$$

$$\mathbf{X}_2 = \left[\boldsymbol{0}_{s_2\times s_0}\ \sqrt{P_{2\tau}}\boldsymbol{I}_{s_2}\ \sqrt{\frac{P_{2\delta}}{s_2}}\mathbf{S}_2\right] \in \mathbb{C}^{s_2\times T}, \tag{4.158}$$

$$\mathbf{X}_2' = \begin{bmatrix}\sqrt{\rho_{2\tau}}\boldsymbol{I}_{s_0} & \left[\boldsymbol{0}_{s_0\times s_2}\ \sqrt{\frac{\rho_{2\delta}}{s_0}}\mathbf{S}_2'\right] \\ \boldsymbol{0} & \sqrt{\rho_{2a}}\boldsymbol{I}_{s_1}\end{bmatrix} \in \mathbb{C}^{(s_1+s_0)\times(s_1+s_0)}, \tag{4.159}$$

where $\mathbf{S}_1 \in \mathbb{C}^{(s_1+s_0)\times(T-s_1-s_0)}$, $\mathbf{S}_2 \in \mathbb{C}^{s_2\times(T-s_2-s_0)}$, and $\mathbf{S}_2' \in \mathbb{C}^{s_0\times(s_1-s_2)}$ are data matrices containing $\mathcal{N}_{\mathbb{C}}(0,1)$ entries. The power constraint $\mathbb{E}[\mathrm{tr}(\mathbf{X}^{\mathsf{H}}\mathbf{X})] \le PT$ translates to (4.154). In the following, we show that the rate pair (R_1, R_2) given in (4.151) and (4.153) is achievable with this input distribution, thus prove Theorem 4.6.

4.4.5.a Rate of User 1

The received signal at user 1 is

$$\mathsf{Y}_1 = \mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_1\mathsf{X}_2'\left[\sqrt{P_{1\tau}}\boldsymbol{I}_{s_1+s_0}\ \sqrt{\frac{P_{1\delta}}{s_1+s_0}}\mathsf{S}_1\right] + \mathsf{Z}_1 \tag{4.160}$$

$$= \left[\underbrace{\sqrt{P_{1\tau}}\mathsf{G}_{1e} + \mathsf{Z}_{1[1:s_1+s_0]}}_{\mathsf{Y}_{1\tau}}\quad \underbrace{\sqrt{\frac{P_{1\delta}}{s_1+s_0}}\mathsf{G}_{1e}\mathsf{S}_1 + \mathsf{Z}_{1[s_1+s_0+1:T]}}_{\mathsf{Y}_{1\delta}}\right], \tag{4.161}$$

where $\mathsf{G}_{1e} := \mathsf{G}_1\boldsymbol{\Sigma}_1^{\frac{1}{2}}\boldsymbol{\Phi}_1\mathsf{X}_2'$ is the equivalent channel with correlation matrix

$$\boldsymbol{R}_{1e} := \frac{1}{N_1}\mathbb{E}[\mathsf{G}_{1e}^{\mathsf{H}}\mathsf{G}_{1e}] = \begin{bmatrix} \rho_{2\tau}\breve{\boldsymbol{R}}_{10} & \sqrt{\rho_{2\tau}\rho_{2a}}\boldsymbol{\Phi}_{10}^{\mathsf{H}}\boldsymbol{\Sigma}_1\boldsymbol{\Phi}_{11} \\ \sqrt{\rho_{2\tau}\rho_{2a}}\boldsymbol{\Phi}_{11}^{\mathsf{H}}\boldsymbol{\Sigma}_1\boldsymbol{\Phi}_{10} & \begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \frac{\rho_{2\delta}}{s_0}\mathrm{tr}\big(\breve{\boldsymbol{R}}_{10}\big)\boldsymbol{I}_{s_1-s_2}\end{bmatrix} + \rho_{2a}\breve{\boldsymbol{R}}_{22} \end{bmatrix}. \tag{4.162}$$

Following Lemma 4.2, user 1 estimates the equivalent channel G_{1e} using a MMSE estimator based on $\mathsf{Y}_{1\tau}$ as

$$\hat{\mathsf{G}}_{1e} = \sqrt{P_{1\tau}}\mathsf{Y}_{1\tau}(P_{1\tau}\boldsymbol{R}_{1e} + \boldsymbol{I}_{s_1+s_0})^{-1}\boldsymbol{R}_{1e}. \tag{4.163}$$

The estimate $\hat{\mathsf{G}}_{1e}$ and the estimation error $\tilde{\mathsf{G}}_{1e} = \mathsf{G}_{1e} - \hat{\mathsf{G}}_{1e}$ have zero mean and row covariance

$$\frac{1}{N_1}\mathbb{E}\Big[\hat{\mathsf{G}}_{1e}^{\mathsf{H}}\hat{\mathsf{G}}_{1e}\Big] = P_{1\tau}\boldsymbol{R}_{1e}(P_{1\tau}\boldsymbol{R}_{1e} + \boldsymbol{I}_{s_1+s_0})^{-1}\boldsymbol{R}_{1e}, \tag{4.164}$$

$$\frac{1}{N_1}\mathbb{E}\Big[\tilde{\mathsf{G}}_{1e}^{\mathsf{H}}\tilde{\mathsf{G}}_{1e}\Big] = \boldsymbol{R}_{1e} - P_{1\tau}\boldsymbol{R}_{1e}(P_{1\tau}\boldsymbol{R}_{1e} + \boldsymbol{I}_{s_1+s_0})^{-1}\boldsymbol{R}_{1e} = \Big(\boldsymbol{R}_{1e}^{-1} + P_{1\tau}\boldsymbol{I}_{s_1+s_0}\Big)^{-1}. \tag{4.165}$$

Using the worst-case noise argument and Lemma 4.1 as before, we have the achievable rate R_1 for user 1 as defined in (4.151).

4.4.5.b Rate of User 2

After some manipulations, the received signal at user 2 can be written as

$$\mathsf{Y}_2 = \mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_2\begin{bmatrix}\sqrt{\rho_{2\tau}P_{1\tau}}\boldsymbol{I}_{s_0} & \mathbf{0} & \Big[\sqrt{\frac{\rho_{2\delta}P_{1\tau}}{s_0}}\mathsf{S}_2'\ \mathsf{A}\Big] \\ \mathbf{0} & \sqrt{P_{2\tau}}\boldsymbol{I}_{s_2} & \sqrt{\frac{P_{2\delta}}{s_2}}\mathsf{S}_2\end{bmatrix} + \mathsf{Z}_2 \tag{4.166}$$

$$= \Big[\mathsf{Y}_{2\tau}\ \underbrace{\mathsf{Y}_{2\delta a}\ \mathsf{Y}_{2\delta b}}_{\mathsf{Y}_{2\delta}}\Big], \tag{4.167}$$

where $\mathsf{A} := \Big[\sqrt{\rho_{2\tau}}\boldsymbol{I}_{s_0}\ \mathbf{0}\ \sqrt{\frac{\rho_{2\delta}}{s_0}}\mathsf{S}_2'\Big]\sqrt{\frac{P_{1\delta}}{s_1+s_0}}\mathsf{S}_1$ and

$$\mathsf{Y}_{2\tau} := \mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_2\boldsymbol{P}_{2\tau}^{\frac{1}{2}} + \mathsf{Z}_{2[1:s_2+s_0]}, \tag{4.168}$$

$$\mathsf{Y}_{2\delta a} := \mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_2\begin{bmatrix}\sqrt{\frac{\rho_{2\delta}P_{1\tau}}{s_0}}\mathsf{S}_2' \\ \sqrt{\frac{P_{2\delta}}{s_2}}\mathsf{S}_{2[1:s_1-s_2]}\end{bmatrix} + \mathsf{Z}_{2[s_2+s_0+1:s_1+s_0]}, \tag{4.169}$$

$$\mathsf{Y}_{2\delta b} := \mathsf{G}_2\boldsymbol{\Sigma}_2^{\frac{1}{2}}\boldsymbol{\Phi}_2\begin{bmatrix}\mathsf{A} \\ \sqrt{\frac{P_{2\delta}}{s_2}}\mathsf{S}_{2[s_1-s_2+1:T-s_2]}\end{bmatrix} + \mathsf{Z}_{2[s_1+s_0+1:T]}, \tag{4.170}$$

where $\boldsymbol{P}_{2\tau} := \begin{bmatrix} \rho_{2\tau} P_{1\tau} \boldsymbol{I}_{s_0} & \mathbf{0} \\ \mathbf{0} & P_{2\tau} \boldsymbol{I}_{s_2} \end{bmatrix}$. The rate that user 2 can achieve is $\frac{1}{T} I(\mathsf{Y}_2; \mathsf{S}_2', \mathsf{S}_2)$ bits/channel use with

$$
\begin{aligned}
I(\mathsf{Y}_2; \mathsf{S}_2', \mathsf{S}_2) &= I(\mathsf{Y}_{2\tau}, \mathsf{Y}_{2\delta}; \mathsf{S}_2', \mathsf{S}_2) && (4.171)\\
&= \underbrace{I(\mathsf{Y}_{2\tau}; \mathsf{S}_2', \mathsf{S}_2)}_{=0} + I(\mathsf{Y}_{2\delta}; \mathsf{S}_2', \mathsf{S}_2 \,|\, \mathsf{Y}_{2\tau}) && (4.172)\\
&= I(\mathsf{Y}_{2\delta}; \mathsf{S}_2', \mathsf{S}_2, \mathsf{A} \,|\, \mathsf{Y}_{2\tau}) - I(\mathsf{Y}_{2\delta}; \mathsf{A} \,|\, \mathsf{Y}_{2\tau}, \mathsf{S}_2', \mathsf{S}_2) && (4.173)
\end{aligned}
$$

where the second and third equalities follow from the chain rule.

Using the worst-case noise argument and Lemma 4.1 as before, we can bound the term $I(\mathsf{Y}_{2\delta}; \mathsf{S}_2', \mathsf{S}_2, \mathsf{A} \,|\, \mathsf{Y}_{2\tau})$ by the sum of the first two terms in the right-hand side of (4.153). Whereas the term $I(\mathsf{Y}_{2\delta}; \mathsf{A} \,|\, \mathsf{Y}_{2\tau}, \mathsf{S}_2', \mathsf{S}_2)$ can be upper bounded as

$$
\begin{aligned}
&I(\mathsf{Y}_{2\delta}; \mathsf{A} \,|\, \mathsf{Y}_{2\tau}, \mathsf{S}_2', \mathsf{S}_2)\\
&= I(\mathsf{Y}_{2\delta b}; \mathsf{A} \,|\, \mathsf{S}_2', \mathsf{S}_2, \mathsf{Y}_{2\tau}) && (4.174)\\
&= I(\mathsf{Y}_{2\delta b}; \mathsf{A} \,|\, \mathsf{S}_2, \mathsf{Y}_{2\tau}) - I(\mathsf{Y}_{2\delta b}; \mathsf{S}_2' \,|\, \mathsf{S}_2, \mathsf{Y}_{2\tau}) && (4.175)\\
&\le I(\mathsf{Y}_{2\delta b}; \mathsf{A} \,|\, \mathsf{S}_{2[s_1 - s_2 + 1:T - s_2 - s_0]}, \mathsf{Y}_{2\tau}) && (4.176)\\
&= h(\mathsf{A}|\mathsf{S}_{2[s_1 - s_2 + 1:T - s_2 - s_0]}, \mathsf{Y}_{2\tau}) - h(\mathsf{A}|\mathsf{S}_{2[s_1 - s_2 + 1:T - s_2 - s_0]}, \mathsf{Y}_{2\tau}, \mathsf{Y}_{2\delta b}) && (4.177)\\
&\le h(\mathsf{A}|\mathsf{S}_{2[s_1 - s_2 + 1:T - s_2 - s_0]}, \mathsf{Y}_{2\tau}) - h(\mathsf{A}|\mathsf{S}_{2[s_1 - s_2 + 1:T - s_2 - s_0]}, \mathsf{Y}_{2\tau}, \mathsf{Y}_{2\delta b}, \mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_2) && (4.178)\\
&= h(\mathsf{A}|\mathsf{S}_{2[s_1 - s_2 + 1:T - s_2 - s_0]}, \mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_2) - h(\mathsf{A}|\mathsf{S}_{2[s_1 - s_2 + 1:T - s_2 - s_0]}, \mathsf{Y}_{2\delta b}, \mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_2) && (4.179)\\
&= I(\mathsf{Y}_{2\delta b}; \mathsf{A} \,|\, \mathsf{S}_{2[s_1 - s_2 + 1:T - s_2 - s_0]}, \mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_2) && (4.180)\\
&= I\Bigg(\mathsf{Y}_{2\delta b} - \sqrt{\frac{P_{2\delta}}{s_2}} \mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_{22} \mathsf{S}_{2[s_1 - s_2 + 1:T - s_2 - s_0]}; \mathsf{A} \;\Big|\; \mathsf{S}_{2[s_1 - s_2 + 1:T - s_2 - s_0]},\\
&\qquad \mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_{20}, \mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_{22}\Bigg) && (4.181)\\
&= I\Big(\mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_{20} \mathsf{A} + \mathsf{Z}_{2[s_1 + s_0 + 1:T]}; \mathsf{A} \;\Big|\; \mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_{20}\Big) && (4.182)\\
&= (T - s_1 - s_0) \mathbb{E}\Big[\log_2 \det\Big(\boldsymbol{I}_{N_2} + P_{1\delta}\Big(\rho_{2\tau} + \rho_{2\delta} \frac{s_1 - s_2}{s_0}\Big) \bar{\boldsymbol{\Omega}}_{20} \bar{\boldsymbol{\Omega}}_{20}^{\mathsf{H}}\Big)\Big], && (4.183)
\end{aligned}
$$

where $\bar{\boldsymbol{\Omega}}_{20} \in \mathbb{C}^(N_2 \times s_0)$ has independent rows following $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \breve{\boldsymbol{R}}_{20})$. Here, (4.174) and (4.175) follows from the Markov chains $\mathsf{Y}_{2\delta a} \leftrightarrow \mathsf{S}_2' \leftrightarrow \mathsf{A}$ and $\mathsf{Y}_{2\delta b} \leftrightarrow \mathsf{A} \leftrightarrow \mathsf{S}_2'$, respectively; (4.176) holds because mutual information is nonnegative and both $\mathsf{Y}_{2\delta b}$ and A are independent of $\mathsf{S}_{2[1:s_1 - s_2]}$; (4.178) holds because conditioning reduces entropy; (4.179) holds because A is independent of both $\mathsf{Y}_{2\tau}$ and $\mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_2$, while given $\mathsf{Y}_{2\delta b}$, A depends on $\mathsf{Y}_{2\tau}$ only through $\mathsf{G}_2 \boldsymbol{\Sigma}_2^{\frac{1}{2}} \boldsymbol{\Phi}_2$; and in the last equality, we used that $\mathbb{E}[\mathsf{A}\mathsf{A}^{\mathsf{H}}] = P_{1\delta}\Big(\rho_{2\tau} + \rho_{2\delta} \frac{s_1 - s_2}{s_0}\Big) \boldsymbol{I}_{s_0}$.

4.4.6 Numerical Results

In this subsection, we compute numerically the achievable rate regions for the BC with the aforementioned schemes, namely, TDMA, rate splitting, product superposition, and hybrid superposition. We assume Rayleigh fading, i.e., G_k has independent $\mathcal{N}_{\mathbb{C}}(0, 1)$ entries. We generate the correlation matrix $\boldsymbol{R}_k = \boldsymbol{U}_k \boldsymbol{\Sigma}_k \boldsymbol{U}_k^{\mathsf{H}}$, $k \in \{1, 2\}$, as follows:

- The eigenvalues in $\boldsymbol{\Sigma}_k$ are drawn from the joint distribution of the nonzero eigenvalues of a Wishart matrix $\mathsf{B}\mathsf{B}^{\mathsf{H}}$ where B is a $M \times r_k$ random matrix with independent

$\mathcal{N}_{\mathbb{C}}(0,1)$ entries, and normalized such that $\mathrm{tr}(\boldsymbol{\Sigma}_k) = M$. This is suggested by the maximum-entropy channel modeling approach [168], [169].

- The eigenvectors $\boldsymbol{U}_k$ are generated as $\boldsymbol{U}_k = \hat{\boldsymbol{U}}_k \boldsymbol{\Xi}_k$, where $\hat{\boldsymbol{U}}_1$ and $\hat{\boldsymbol{U}}_2$ are drawn respectively by selecting randomly r_1 and r_2 columns of a random unitary matrix $\mathbf{U}$ which is uniformly distributed in the space of $M \times M$ unitary matrix, and the rotation matrix $\boldsymbol{\Xi}_k$ is drawn uniformly from the space of $r_k \times r_k$ unitary matrix.

In Fig. 4.2, we plot the rate regions for the BC achieved with these schemes in a setting of $T = 24$, $M = 16$, $N_1 = N_2 = 12$, $r_1 = 10$, $r_0 = 6$, $r_2 = 4$ and $T = 32$, $M = N_1 = N_2 = 16$, $r_1 = 8$, $r_0 = 7$, $r_2 = 1$, at SNR $P = 30$ dB. We observe that the rate region of TDMA while transmitting in the channel eigenspace $\mathrm{Span}(\boldsymbol{U}_k)$ (see (4.70)) is much larger than the region of TDMA while transmitting in full space $\mathbb{C}^M$ (see (4.69)). This is because the former scheme spends less time (r_k channel uses) for channel estimation than the latter scheme (spending M channel uses), while both schemes essentially estimate the same effective channel. The rate region can be largely improved with the proposed schemes. Rate splitting achieves a large region w.r.t. other schemes, especially when the ranks of the two eigenspaces are close to each other (as $r_1 = 9, r_2 = 7$ in Fig. 4.2(a)). The improvement by product superposition is more pronounced when the rank difference between two eigenspaces is more significant (as $r_1 = 10, r_2 = 6$ in Fig. 4.2(b)) since the gains achieved by product superposition come from the nonoverlapping part of the eigenspaces.

4.5 K-User Broadcast Channel

In this section, we extend our DoF studies to the K-user scenario. In this case, a product superposition scheme was proposed in [76], resulting in an achievable DoF tuple as follows.

Proposition 4.2 ([76, Theorem 5])**.** *For the K-user noncoherent BC where the correlation eigenvectors are nested such that $\boldsymbol{U}_{k-1} = [\bar{\boldsymbol{U}}_k \; \boldsymbol{U}_k]$ with $\bar{\boldsymbol{U}}_k$ being a basis of the complement of $\mathrm{Span}(\boldsymbol{U}_k)$ in $\mathrm{Span}(\boldsymbol{U}_{k-1})$, $k \in \{2, 3, \ldots, K\}$, the DoF tuple $(d_1, \ldots, d_K)$ given by*

$$d_1 = N_1^*\Big(1 - \frac{r_1}{T}\Big) \quad \text{and} \quad d_k = N_k^* \frac{r_{k-1} - r_k}{T}, \quad k \in \{2, 3, \ldots, K\} \tag{4.184}$$

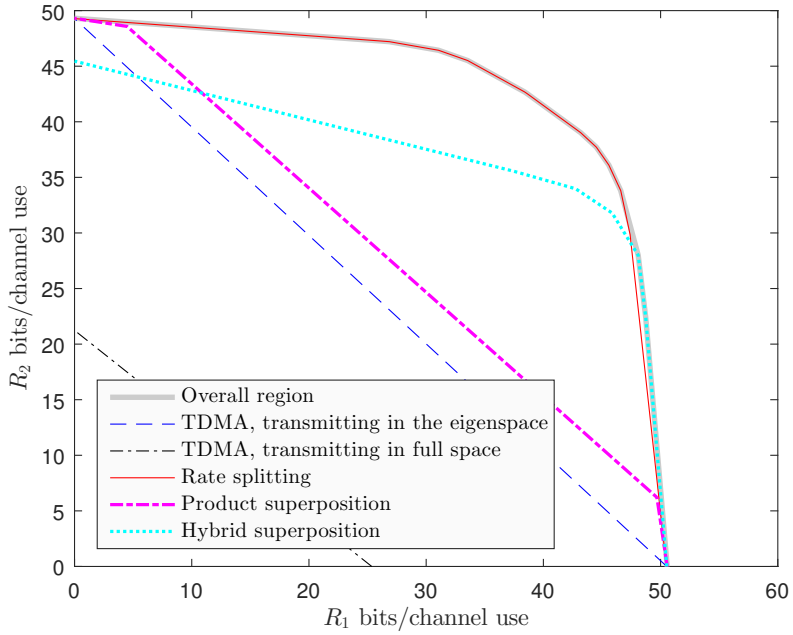
is achievable.

The nested correlation eigen-structure in Proposition 4.2 is a strong assumption. In this section, we consider more general settings. We first describe further the correlation model as follows. Recall that the rows of $\mathbf{H}_k$ belong to the eigenspace $\mathrm{Span}(\boldsymbol{U}_k)$ of $\boldsymbol{R}_k$. Denote the union of all channel eigenspaces as

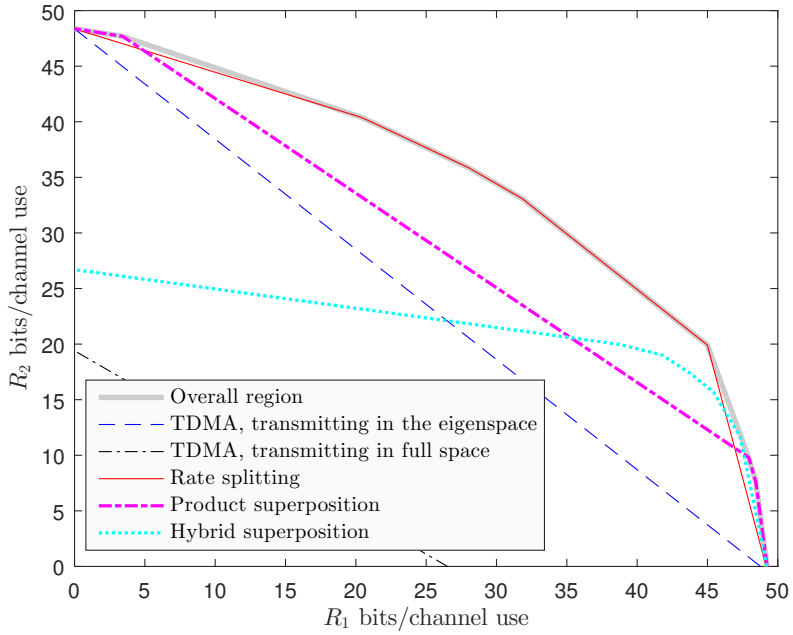
$$\mathcal{V} = \bigcup_{k \in [K]} \mathrm{Span}(\boldsymbol{U}_k). \tag{4.185}$$

$\mathcal{V}$ can be partitioned into $2^K - 1$ subspaces $\mathcal{V}_{\mathcal{J}}$ of dimension $r_{\mathcal{J}}$ whose $r_{\mathcal{J}}$ basis vectors span the channel of every user in a nonempty group $\mathcal{J} \subset [K]$ and are orthogonal to all vectors in $\mathrm{Span}(\boldsymbol{U}_k)$ for $k \in \{[K] \setminus \mathcal{J}\}$. In other words, $\mathcal{V}_{\mathcal{J}} = \bigcap_{k \in \mathcal{J}} \mathrm{Span}(\boldsymbol{U}_k)$. Obviously, $\sum_{\mathcal{J} \subset [K]} r_{\mathcal{J}} = \mathrm{rank}(\mathcal{V}) \leq M$ and $\sum_{\mathcal{J} \subset [K]:\ k \in \mathcal{J}} r_{\mathcal{J}} = \mathrm{rank}(\mathrm{Span}(\boldsymbol{U}_k)) = r_k$. An example of the correlation structure for the case of three-user BC is shown in Fig. 4.3.

In this way, the signal transmitted in the subspace $\mathcal{V}_{\mathcal{J}}$ can be seen by every user in $\mathcal{J}$ and is not seen by all other users. On the other hand, the signals transmitted in $\mathcal{V}_{\mathcal{J}}$ and $\mathcal{V}_{\mathcal{K}}$ interfere each other at every user in $\mathcal{J} \cap \mathcal{K}$. To characterize the interfering relation between signals transmitted in different subspaces, we introduce the concept of *interference graph* as follows:



(a) $r_1 = 9$, $r_2 = 7$, $r_0 = 2$



(b) $r_1 = 10$, $r_2 = 6$, $r_0 = 4$

FIGURE 4.2: The rate regions of various schemes for the spatially correlated BC with $T = 24$, $N_1 = N_2 = 12$, $M = 16$, and at SNR $P = 30$ dB.

Definition 4.1. *For $k \in [K]$, the* interference graph *of order k, denoted by $G(K, k)$, is an undirected graph for which:*

- *the set of vertices is the set of unordered subsets of cardinality k of $[K]$, i.e., $\mathcal{J} \subset [K] : |\mathcal{J}| = k$, hence a vertex is also denoted by a subset $\mathcal{J}$;*
- *there exists an edge between two vertices $\mathcal{J}$ and $\mathcal{K}$ if and only if $\mathcal{J} \cap \mathcal{K} \neq \emptyset$.*

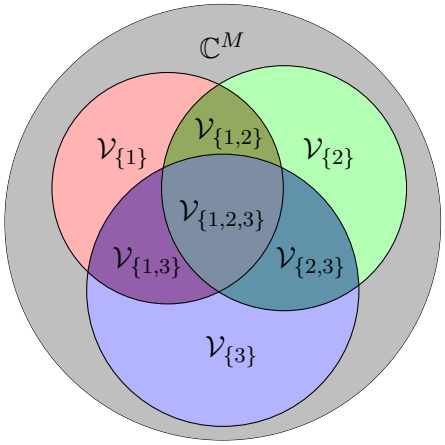


FIGURE 4.3: The channel eigenspace overlapping structure of the three-user BC.

The interference graph $G(K,k)$ has $\binom{K}{k}$ vertices. It is a regular graph [170, Sec. 1.2] of degree $\binom{K}{k} - \binom{K-k}{k} - 1$. Let $\chi(G(K,k))$ denote the chromatic number of $G(K,k)$, i.e., the minimum number of colors to color all the vertices such that adjacent vertices have different colors. We have the following property.

Property 4.1 (The chromatic number of the interference graph)**.** *It holds that*

$$\chi(G(K,k)) \begin{cases} = 1 & \text{if } k = 1, \\ \leq \binom{K}{k} - \binom{K-k}{k} - 1 & \text{if } 1 < k \leq \lfloor K/2 \rfloor, \\ = \binom{K}{k} & \text{if } k > \lfloor K/2 \rfloor. \end{cases} \tag{4.186}$$

Proof. $\chi(G(K,1)) = 1$ since $G(K,1)$ is edgeless. $\chi(G(K,k)) = \binom{K}{k}$ when $k > \lfloor K/2 \rfloor$ because in this case, $G(K,k)$ is complete. The results for the case $1 < k \leq \lfloor K/2 \rfloor$ follows from Brook's theorem [170, Theorem 5.2.4]. □

4.5.1 The Symmetric K-user BC: An Achievable DoF Region with Rate Splitting

We first consider the symmetric case:

$$r_{\mathcal{J}_i} = r_{\mathcal{J}_j}, \quad \forall \mathcal{J}_i, \mathcal{J}_j \subset [K] : |\mathcal{J}_i| = |\mathcal{J}_j|. \tag{4.187}$$

That is, the rank of the common channel eigenspace $\mathcal{V}_{\mathcal{J}}$ is the same for all groups $\mathcal{J}$ containing the same number of users. (In the two-user case, this corresponds to $r_1 = r_2$.) Define

$$p_k = r_{\mathcal{J}}, \quad \forall \mathcal{J} \subset [K] : |\mathcal{J}| = k, \tag{4.188}$$

for $k \in [K]$. Then the set of parameters $(p_1, \ldots, p_K)$ characterizes the correlation structure of the K-user symmetric BC. Furthermore, we assume that $r_k \leq N_k$, $\forall k$.

Theorem 4.7. *The K-user symmetric noncoherent BC characterized by $(p_1, \ldots, p_K)$ can achieve any permutation of the DoF tuple $D_{K,L}(p_1, \ldots, p_K) = (d_1, d_2, \ldots, d_K)$, for any $L \in \{0, 1, \ldots, K-1\}$, defined by*

$$d_k = \frac{1}{T} \sum_{i=1}^{K-\max(k-1,L)} \min\left(\binom{K-i}{i-1} p_i, N_k\right) \left(T - T_\tau(K,L) + \sum_{j=\lfloor K/2 \rfloor + 1}^{K-i} \binom{K-i}{k} p_j\right), \tag{4.189}$$

for $k \in [K]$, *where* $T_\tau(K,L) := \sum_{k=1}^{K-L} \chi(G(K,k))p_k$.

Let us first describe the achievable scheme in the 3-user case for clarity, then go for the K-user case.

Example 4.1 (Achievable scheme for Theorem 4.7 for $K = 3$)**.** When $K = 3$, the correlation structure is illustrated in Fig. 4.3. Under the symmetry assumption, we have $r_{\{1\}} = r_{\{2\}} = r_{\{3\}} := p_1$, $r_{\{1,2\}} = r_{\{1,3\}} = r_{\{2,3\}} := p_2$, $r_{\{1,2,3\}} := p_3$. The achievable scheme for

$$D_{3,0}(p_1,p_2,p_3) = \left((p_1+2p_2+p_3)\left(1-\frac{T_\tau}{T}\right)+\frac{p_1p_2}{T}, (p_1+p_2)\left(1-\frac{T_\tau}{T}\right)+\frac{p_1p_2}{T}, p_1\left(1-\frac{T_\tau}{T}\right)+\frac{p_1p_2}{T}\right), \tag{4.190}$$

is based on rate splitting and channel training as illustrated in Table 4.1.

TABLE 4.1: Illustration of pilot and data alignment for the scheme achieving $D_{3,0}(p_1,p_2,p_3)$

$\mathcal{V}_{\{1\}}$	Pilot		Data			Data
$\mathcal{V}_{\{2\}}$	Pilot			Data		Data
$\mathcal{V}_{\{3\}}$	Pilot	Data				Data
$\mathcal{V}_{\{1,2\}}$		Pilot				Data
$\mathcal{V}_{\{2,3\}}$			Pilot			Data
$\mathcal{V}_{\{1,3\}}$				Pilot		Data
$\mathcal{V}_{\{1,2,3\}}$					Pilot	Data
	$\leftarrow p_1 \rightarrow$	$\leftarrow p_2 \rightarrow$	$\leftarrow p_2 \rightarrow$	$\leftarrow p_2 \rightarrow$	$\leftarrow p_3 \rightarrow$	$\longleftarrow T-(p_1+3p_2+p_3) \longrightarrow$

Owing to linear precoding using a basis $\boldsymbol{V}_\mathcal{J}$ of $\mathcal{V}_\mathcal{J}$, pilots can be sent in each subspace $\mathcal{V}_\mathcal{J}, \mathcal{J} \subset [3]$, so that all users in $\mathcal{J}$ can learn the channel directions in $\mathcal{V}_\mathcal{K}$. User 1 needs to learn channel directions in $\mathcal{V}_{\{1\}}$, $\mathcal{V}_{\{1,2\}}$, $\mathcal{V}_{\{1,3\}}$, and $\mathcal{V}_{\{1,2,3\}}$, and similar for user 2 and user 3. To avoid pilot interference, the pilots in $\mathcal{V}_{\{1,2\}}$, $\mathcal{V}_{\{2,3\}}$, and $\mathcal{V}_{\{1,3\}}$ must be sent in different channel uses. Whereas pilots in $\mathcal{V}_{\{1\}}$, $\mathcal{V}_{\{2\}}$, and $\mathcal{V}_{\{3\}}$ can be sent simultaneously because they do not interfere with each other. This alignment of pilots can be interpreted as interference graph coloring: pilots can be transmitted simultaneously in both $\mathcal{V}_\mathcal{J}$ and $\mathcal{V}_\mathcal{K}$ if $\mathcal{J} \cap \mathcal{K} = \emptyset$, i.e., the vertices corresponding to $\mathcal{J}$ and $\mathcal{K}$ are not connected and thus assigned different colors. The interference graph $G(3,1)$ has chromatic number $\chi(G(3,1)) = 1$, which is also the amount of time, normalized by p_1, needed for pilot transmission without interference in $\mathcal{V}_{\{1\}}$, $\mathcal{V}_{\{2\}}$, and $\mathcal{V}_{\{3\}}$. Similarly, it takes $\chi(G(3,2))p_2 = 3p_2$ channel uses to transmit pilot interference-free in $\mathcal{V}_{\{1,2\}}$, $\mathcal{V}_{\{2,3\}}$, and $\mathcal{V}_{\{1,3\}}$, and takes $\chi(G(3,3))p_3 = p_3$ channel uses for pilot transmission in $\mathcal{V}_{\{1,2,3\}}$.

In this way, the total time for channel training is $T_\tau = \sum_{k=1}^{3} \chi(G(3,k))p_k = p_1+3p_2+p_3$ channel uses and there remains $T - T_\tau$ channel uses for simultaneous data transmission in all subspaces. By dedicating the data transmitted in $\mathcal{V}_{\{1,2\}}$, $\mathcal{V}_{\{1,3\}}$, and $\mathcal{V}_{\{1,2,3\}}$ to user 1, user 1 achieves $(p_1+2p_2+p_3)(1-\frac{T_\tau}{T})$ DoF. By dedicating the data transmitted in $\mathcal{V}_{\{2,3\}}$ to user 2, user 2 achieves $(p_1+p_2)(1-\frac{T_\tau}{T})$ DoF. User 3 achieves $p_1(1-\frac{T-\tau}{T})$ DoF from the data transmitted in $\mathcal{V}_{\{3\}}$. On top of that, the transmitter can transmit additional data to user 3 in $\mathcal{V}_{\{3\}}$ by superimposing it with the pilot for user 1 and user 2 in $\mathcal{V}_{\{1,2\}}$ without interference. Similarly, user 1 and user 2 can also receive additional data. With these additional data, each user achieves $\frac{p_1p_2}{T}$ DoF. Therefore, $D_{3,0}(p_1,p_2,p_3)$ is achieved.

To achieve $D_{3,1}(p_1,p_2,p_3)$, which is

$$\left((p_1+2p_2)\left(1-\frac{p_1+3p_2}{T}\right)+\frac{p_1p_2}{T},(p_1+p_2)\left(1-\frac{p_1+3p_2}{T}\right)+\frac{p_1p_2}{T},p_1\left(1-\frac{p_1+3p_2}{T}\right)+\frac{p_1p_2}{T}\right), \tag{4.191}$$

we simply ignore the subspace $\mathcal{V}_{\{1,2,3\}}$. Then, we do not send pilot in this subspace and have more time to send data in all other subspaces. As a price for that, we lose the data we could send in $\mathcal{V}_{\{1,2,3\}}$ during the last $T-T\tau$ channel uses. When $\text{rank}\left(\mathcal{V}_{\{1,2,3\}}\right)=p_3$ is small enough, this loss is not significant and we can gain DoF. The achievable scheme is illustrated in Table 4.2.

TABLE 4.2: Illustration of pilot and data alignment for the scheme achieving $D_{3,1}(p_1,p_2,p_3)$

$\mathcal{V}_{\{1\}}$	Pilot		Data		Data
$\mathcal{V}_{\{2\}}$	Pilot			Data	Data
$\mathcal{V}_{\{3\}}$	Pilot	Data			Data
$\mathcal{V}_{\{1,2\}}$		Pilot			Data
$\mathcal{V}_{\{2,3\}}$			Pilot		Data
$\mathcal{V}_{\{1,3\}}$				Pilot	Data
$\mathcal{V}_{\{1,2,3\}}$					
	$\leftarrow\ p_1\ \rightarrow$	$\leftarrow\ p_2\ \rightarrow$	$\leftarrow\ p_2\ \rightarrow$	$\leftarrow\ p_2\ \rightarrow$	$\longleftarrow\ T-(p_1+3p_2)\ \longrightarrow$

Similarly, $D_{3,2}(p_1,p_2,p_3)=\left(p_1\left(1-\frac{p_1}{T}\right),p_2\left(1-\frac{p_2}{T}\right),p_3(1-\frac{p_3}{T})\right)$ can be achieved by ignoring $\mathcal{V}_{\{1,2\}}$, $\mathcal{V}_{\{2,3\}}$, $\mathcal{V}_{\{1,3\}}$, and $\mathcal{V}_{\{1,2,3\}}$, as illustrated in Table 4.3.

TABLE 4.3: Illustration of pilot and data alignment for the scheme achieving $D_{3,2}(p_1,p_2,p_3)$

$\mathcal{V}_{\{1\}}$	Pilot	Data
$\mathcal{V}_{\{2\}}$	Pilot	Data
$\mathcal{V}_{\{3\}}$	Pilot	Data
$\mathcal{V}_{\{1,2\}}$		
$\mathcal{V}_{\{2,3\}}$		
$\mathcal{V}_{\{1,3\}}$		
$\mathcal{V}_{\{1,2,3\}}$		
	$\leftarrow\ p_1\ \rightarrow$	$\longleftarrow\ T-p_1\ \longrightarrow$

Due to symmetry, any permutation of $D_{3,L}$, $L\in\{0,1,2\}$ is achieved by permuting the users' indices.

Proof of Theorem 4.7. We first show the achievable scheme for $D_{K,0}(p_1,\ldots,p_K)$ given by

$$d_k=\frac{1}{T}\sum_{i=1}^{K-k+1}\binom{K-i}{i-1}p_i\left(T-T_\tau(K,0)+\sum_{j=\lfloor K/2\rfloor+1}^{K}\binom{K-i}{j}p_j\right). \tag{4.192}$$

The scheme is based on rate splitting and channel training with two key elements: alignment of pilots in different subspaces, and superposition of additional data on top of pilots without causing interference.

User k needs to learn the channel directions in all subspaces $\mathcal{V}_\mathcal{J}$ such that $k\in\mathcal{J}$ and is oblivious to signals (pilot or data) transmitted in other subspaces. To avoid pilot interference, the pilots in $\mathcal{V}_\mathcal{J}$ and $\mathcal{V}_\mathcal{K}$ need to be orthogonal in time if $\mathcal{J}\cap\mathcal{K}=\emptyset$, i.e, $\mathcal{J}$

and $\mathcal{K}$ are connected in the interference graph. Pilots in $\mathcal{V}_\mathcal{J}$ and $\mathcal{V}_\mathcal{K}$ can be transmitted simultaneously if $\mathcal{J} \cap \mathcal{K} = \emptyset$, i.e., $\mathcal{J}$ and $\mathcal{K}$ are not connected. Therefore, the problem of pilot alignment can be interpreted as interference graph coloring: pilots can be transmitted at the same time without interference in the subspaces corresponding to vertices with the same color. The minimum total amount of time for pilot transmission, normalized by the subspace dimension, is thus the minimum number of colors, which is the chromatic number of the graph. That is, the total training time is $T_\tau(K,0) = \sum_{k=1}^{K} \chi(G(K,k))p_k$ channel uses. In the remaining $T - T_\tau(K,0)$ channel uses, data is transmitted in all subspaces. The DoF that user k, $k \in \mathcal{J}$, can achieve with the message transmitted in $\mathcal{V}_\mathcal{J}$ is $\frac{1}{T}p_{|\mathcal{V}_\mathcal{J}|}(T - T_\tau(K,0))$.

Notice that for any $l > \lfloor K/2 \rfloor$, the interference graph $G(K,l)$ is fully connected, the pilots in subspaces $\mathcal{V}_\mathcal{K}$ for $|\mathcal{K}| = l$ cannot be transmitted at the same time. However, additional data can be transmitted in any subspace $\mathcal{V}_\mathcal{J}$ such that $\mathcal{V}_\mathcal{K} \cap \mathcal{V}_\mathcal{J} = \emptyset$. In this way, during the training of all subspaces $\mathcal{V}_\mathcal{K}$ with $|\mathcal{K}| = l$, for each subset $\mathcal{J}$ which does not intersect with $|\mathcal{K}|$, additional data can be transmitted in $\binom{K-|\mathcal{J}|}{l}p_l$ channel uses, enabling each user in $\mathcal{J}$ to achieve $\frac{1}{T}\binom{K-|\mathcal{J}|}{l}p_\mathcal{J}p_l$ more DoF.

Summing up the DoF, the number of DoF that each user in $\mathcal{J}$ can obtain from the message transmitted in $\mathcal{V}_\mathcal{J}$ is

$$\frac{1}{T}p_{|\mathcal{J}|}\Big(T - T_\tau(K,0)\Big) + \frac{1}{T}\sum_{l=\lfloor K/2 \rfloor+1}^{K}\binom{K-|\mathcal{J}|}{l}p_{|\mathcal{J}|}p_l$$
$$= \frac{1}{T}p_{|\mathcal{J}|}\left(T - T_\tau(K,0) + \sum_{l=\lfloor K/2 \rfloor+1}^{K}\binom{K-|\mathcal{J}|}{l}p_l\right). \quad (4.193)$$

By dedicating all the messages transmitted in $\mathcal{V}_\mathcal{K}$ such that $k \in \mathcal{J}$ and $\mathcal{J} \cap [k-1] = \emptyset$ to user k, user k achieves d_k DoF where d_k is given in (4.192). Then $D_{K,0}(p_1,\ldots,p_K)$ is achievable.

Similar to the 3-user case, $D_{K,L}(p_1,\ldots,p_K)$ with $L \in [K-1]$ is achieved by ignoring all the subspaces $\mathcal{V}_\mathcal{J}$ with $|\mathcal{J}| > K - L$. Finally, due to symmetry, any permutation of $D_{K,L}(p_1,\ldots,p_K)$ with $L = 1,\ldots,K-1$ can be achieved by permuting the users' indices. □

Remark 4.2. *We can improve the achievable scheme by sending additional data during the training of $\mathcal{V}_\mathcal{K}$ with $|\mathcal{K}| \leq \lfloor K/2 \rfloor$ also. However, the possibility for this additional data depends on the actual coloring of the interference graph and would not admit nice expressions of achievable DoF tuples. We therefore do not follow this direction in the interest of developing closed-form expressions.*

Computing the chromatic number $\chi(G(K,k))$ is NP-complete in general [171]. Therefore, one might confine to the achievable DoF tuples in the following corollary.

Corollary 4.3. *The K-user symmetric noncoherent BC can achieve the DoF tuple $D_{K,l}(p_1,\ldots,p_K)$ given in Theorem 4.7, with $T_\tau(K,L)$ replaced by*

$$\sum_{k=1}^{K-L}\left(\binom{K}{k} - \binom{K-k}{k} - \mathbb{1}\{1 < k \leq \lfloor K/2 \rfloor\}\right)p_k.$$

This corollary follows from Theorem 4.7 and Property 4.1.

Based on Theorem 4.7, we have the following achievable DoF region for the symmetric K-user BC.

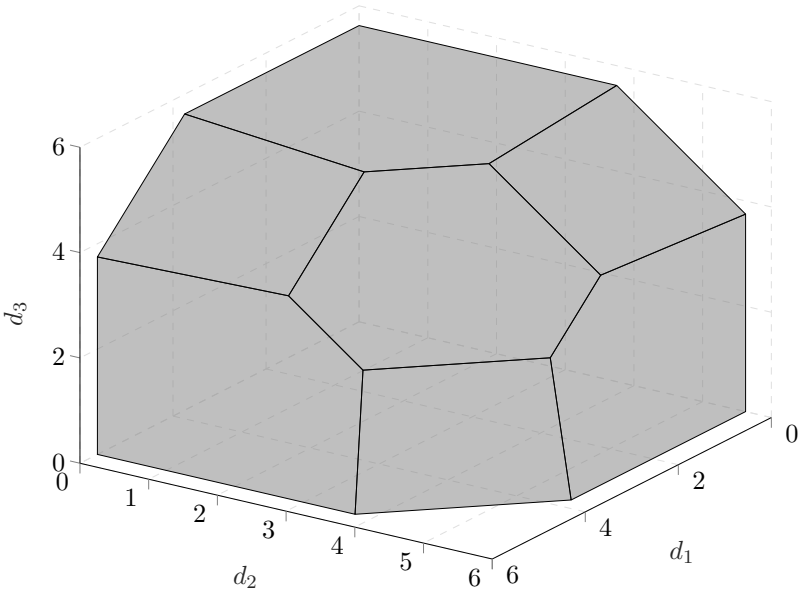


FIGURE 4.4: An achievable DoF region of the symmetric three-user noncoherent BC with spatial correlation with $T = 24, r_{\{1\}} = r_{\{2\}} = r_{\{3\}} =: p_1 = 4, r_{\{1,2\}} = r_{\{1,3\}} = r_{\{2,3\}} =: p_2 = 2$, and $r_{\{1,2,3\}} =: p_3 = 1$.

Theorem 4.8. *The K-user symmetric noncoherent BC characterized by $(p_1, \ldots, p_K)$ can achieve the convex hull of all permutations of any DoF tuple of the form*

$$\Big(D_{K,L}(p_1^*, \ldots, p_k^*), 0, \ldots, 0\Big), \quad \text{for } k \in [K], L \in \{0, \ldots, k-1\}, \tag{4.194}$$

with $D_{K,L}(\cdot)$ defined in (4.189) *and $p_l^* = \sum_{i=0}^{K-k} \binom{K-k}{i} p_{l+i}$ for $l \in [k]$.*

Proof. When $k = K$, (4.194) becomes $D_{K,L}(p_1, \ldots, p_k)$, which can be achieved as stated in Theorem 4.7.

When $k < K$, by ignoring the last $K - k$ users, we construct a new symmetric BC with k users. For example, by ignoring user 3 in the symmetric 3-user BC, we obtain a two-user BC in which the private subspace of user 1 and user 2 are $\mathcal{V}_{\{1\}} \cap \mathcal{V}_{\{1,3\}}$ and $\mathcal{V}_{\{2\}} \cap \mathcal{V}_{\{2,3\}}$, respectively, both of rank $p_1^* = p_1 + p_2$; whereas the common subspace of two users is $\mathcal{V}_{\{1,2\}} \cap \mathcal{V}_{\{1,2,3\}}$ of rank $p_2^* = p_2 + p_3$. In general, the new K-user BC is characterized by the new set of parameters $(p_1^*, \ldots, p_k^*)$, where $p_l^* = \sum_{i=0}^{K-k} \binom{K-k}{i} p_{l+i}$, $l \in [k]$. Then, applying Theorem 4.7 to this k-user symmetric BC, we have that $D_{k,L}(p_1^*, \ldots, p_k^*)$ is achievable. Therefore, $(D_{k,L}(p_1', \ldots, p_k'), 0, \ldots, 0)$ is achievable for the original K-user symmetric BC. Any permutation of (4.194) can be achieved by permuting the users' indices. □

In Fig. 4.4, we plot the achievable DoF region for the symmetric 3-user BC given in Theorem 4.8 with $T = 24, r_{\{1\}} = r_{\{2\}} = r_{\{3\}} = 4, r_{\{1,2\}} = r_{\{1,3\}} = r_{\{2,3\}} = 2$, and $r_{\{1,2,3\}} = 1$.

4.5.2 The General K-User BC: An Extension of Hybrid Superposition

For the general K-user BC, an achievable DoF region with hybrid superposition is stated in the next theorem.

Theorem 4.9. *The K-user noncoherent BC can achieve the DoF tuple* $(d_1, \ldots, d_K)$ *given by*

$$d_k = \sum_{\mathcal{J}\subset[k]:\ k\in\mathcal{J}} r_{\mathcal{J}}\Big(1-\frac{r_k}{T}\Big) + \sum_{l=k+1}^{K} \sum_{\mathcal{J}\subset[K]:\ k\in\mathcal{J},|\{k+1,\ldots,K\}\cap\mathcal{J}|<2} r_{\mathcal{J}}\frac{r_l-r_k}{T}, \tag{4.195}$$

where it is assumed w.l.o.g. that $r_K \geq r_{K-1} \geq \cdots \geq r_1$.

Proof. For simplicity, let us focus on the 3-user case. We assume w.l.o.g. that $r_3 \geq r_2 \geq r_1$. For each partition $\mathcal{V}_{\mathcal{J}}$, $\mathcal{J} \subset [3]$, we build a precoder $\boldsymbol{V}_{\mathcal{J}} \in \mathbb{C}^{M\times r_{\mathcal{J}}}$ as an orthonormal basis of $\mathcal{V}_{\mathcal{J}}$, thus $\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{V}_{\mathcal{J}} = \boldsymbol{0}$, $\forall k \notin \mathcal{J}$, and $\mathrm{rank}(\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{V}_{\mathcal{J}}) = r_{\mathcal{J}}$, $\forall k \in \mathcal{J}$. To combine rate splitting and product superposition, the transmitted signal is

$$\mathbf{X} = [\boldsymbol{V}_{\{1,2,3\}}\ \boldsymbol{V}_{\{2,3\}}\ \boldsymbol{V}_{\{1,3\}}\ \boldsymbol{V}_{\{3\}}]\tilde{\mathbf{S}}_2\mathbf{S}_3 + [\boldsymbol{V}_{\{1,2,3\}}\ \boldsymbol{V}_{\{2\}}]\tilde{\mathbf{S}}_1\mathbf{S}_2 + \boldsymbol{V}_{\{1\}}\mathbf{S}_1, \tag{4.196}$$

with

$$\mathbf{S}_3 = [\boldsymbol{I}_{r_3}\ \mathbf{S}_{d,3}] \in \mathbb{C}^{r_3\times T}, \tag{4.197}$$

$$\mathbf{S}_2 = [\boldsymbol{0}_{(r_{\{1,2\}}+r_{\{2\}})\times(r_{\{1,2,3\}}+r_{\{2,3\}})}\ \boldsymbol{I}_{r_{\{1,2\}}+r_{\{2\}}}\ \mathbf{S}_{d2}] \in \mathbb{C}^{(r_{\{1,2\}}+r_{\{2\}})\times T}, \tag{4.198}$$

$$\mathbf{S}_1 = [\boldsymbol{0}_{r_{\{1\}}\times(r_{\{1,2,3\}}+r_{\{1,3\}}+r_{\{1,2\}})}\ \boldsymbol{I}_r\ \mathbf{S}_{d1}] \in \mathbb{C}^{r_{\{1\}}\times T}, \tag{4.199}$$

$$\tilde{\mathbf{S}}_2 = \begin{bmatrix} \boldsymbol{I}_{r_{\{1,2,3\}}+r_{\{2,3\}}} & \boldsymbol{0}_{(r_{\{1,2,3\}}+r_{\{2,3\}})\times(r_{\{1,2\}}+r_{\{2\}})} & \tilde{\mathbf{S}}_{d2} \\ \boldsymbol{0}_{r_{\{1,3\}}\times r_{\{1,2,3\}}} & \boldsymbol{I}_{r_{\{1,3\}}} & \boldsymbol{0}_{r_{\{1,3\}}\times(r_{\{1,2\}}+r_{\{1\}})}\ \tilde{\mathbf{S}}_{d21} \\ & \bar{\mathbf{S}}_2 & \end{bmatrix} \in \mathbb{C}^{r_3\times r_3}, \tag{4.200}$$

$$\tilde{\mathbf{S}}_1 = \begin{bmatrix} \boldsymbol{0}_{r_{\{1,2\}}\times(r_{\{1,3\}}-r_{\{2,3\}})}\ \boldsymbol{I}_{r_{\{1,2\}}} & \boldsymbol{0}_{r_{\{1,2\}}\times r_{\{1\}}} & \tilde{\mathbf{S}}_{d1} \\ \bar{\mathbf{S}}_1 & & \end{bmatrix} \in \mathbb{C}^{(r_{\{1,2\}}+r_{\{2\}})\times(r_{\{1,2\}}+r_{\{2\}})}, \tag{4.201}$$

where $\bar{\mathbf{S}}_2$ and $\bar{\mathbf{S}}_1$ are designed to guarantee that $\tilde{\mathbf{S}}_2$ and $\tilde{\mathbf{S}}_1$ are respectively nonsingular.

The received signal at user 3 is

$$\mathbf{Y}_3 = \mathbf{H}_3[\boldsymbol{V}_{\{1,2,3\}}\ \boldsymbol{V}_{\{2,3\}}\ \boldsymbol{V}_{\{1,3\}}\ \boldsymbol{V}_{\{3\}}]\tilde{\mathbf{S}}_2[\boldsymbol{I}_{r_3}\ \mathbf{S}_{d3}] + \mathbf{Z}_3. \tag{4.202}$$

User 3 estimates the equivalent channel $\mathbf{H}_3[\boldsymbol{V}_{\{1,2,3\}}\ \boldsymbol{V}_{\{2,3\}}\ \boldsymbol{V}_{\{1,3\}}\ \boldsymbol{V}_{\{3\}}]\tilde{\mathbf{S}}_2$ in the first r_3 channel uses and then decode $\mathbf{S}_{d3}$ to achieve full individual DoF $r_3(1-\frac{r_3}{T})$.

The received signal at user 2 is

$$\mathbf{Y}_2 = \mathbf{H}_2[\boldsymbol{V}_{\{1,2,3\}}\ \boldsymbol{V}_{\{2,3\}}][\boldsymbol{I}_{r_{\{1,2,3\}}+r_{\{2,3\}}}\ \boldsymbol{0}\ \tilde{\mathbf{S}}_{d2}]\mathbf{S}_3 + \mathbf{H}_2[\boldsymbol{V}_{\{1,2\}}\ \boldsymbol{V}_{\{2\}}]\tilde{\mathbf{S}}_1\mathbf{S}_2 + \mathbf{Z}_2 \tag{4.203}$$

$$= \mathbf{H}_2[\boldsymbol{V}_{\{1,2,3\}}\ \boldsymbol{V}_{\{2,3\}}\ [\boldsymbol{V}_{\{1,2\}}\ \boldsymbol{V}_{\{2\}}]\tilde{\mathbf{S}}_1]\begin{bmatrix} \boldsymbol{I}_{r_{\{1,2,3\}}\times r_{\{2,3\}}} & \boldsymbol{0} & [\tilde{\mathbf{S}}_{d2}\ \mathbf{B}] \\ \boldsymbol{0} & \boldsymbol{I}_{r_{\{1,2\}}+r_{\{2\}}} & \mathbf{S}_{d2} \end{bmatrix} + \mathbf{Z}_2, \tag{4.204}$$

where $\mathbf{B} := [\boldsymbol{I}_{r_{\{1,2,3\}}+r_{\{2,3\}}}\ \boldsymbol{0}\ \tilde{\mathbf{S}}_{d2}]\mathbf{S}_{d3}$. User 2 can learn the equivalent channel

$$\mathbf{H}_2[\boldsymbol{V}_{\{1,2,3\}}\ \boldsymbol{V}_{\{2,3\}}\ [\boldsymbol{V}_{\{1,2\}}\ \boldsymbol{V}_{\{2\}}]\tilde{\mathbf{S}}_1]$$

in the first r_2 channel uses and then decode both $\tilde{\mathbf{S}}_{d2}$ and $\mathbf{S}_{d2}$ to achieve $(r_{\{1,2,3\}} + r_{\{2,3\}})\frac{r_3-r_2}{T} + (r_{\{1,2\}} + r_{\{2\}})(1-\frac{r_2}{T})$ DoF in total.

The received signal at user 1 is

$$\begin{aligned}\mathbf{Y}_1 &= \mathbf{H}_1[\boldsymbol{V}_{\{1,2,3\}}\ \boldsymbol{V}_{\{1,3\}}]\begin{bmatrix}\boldsymbol{I}_{r_{\{1,2,3\}}} & \mathbf{0}_{r_{\{1,2,3\}}\times(r_{\{1,3\}}+r_{\{1,2\}}+r_{\{2\}})}\tilde{\mathbf{S}}_{d2[1:r_{\{1,2,3\}}]}\\ \mathbf{0}_{r_{\{1,3\}}\times r_{\{1,2,3\}}} & \boldsymbol{I}_{r_{\{1,3\}}}\ \mathbf{0}_{r_{\{1,3\}}\times(r_{\{1,2\}}+r_{\{1\}})}\ \tilde{\mathbf{S}}_{d21}\end{bmatrix}\mathbf{S}_3\\ &\quad+\mathbf{H}_1\boldsymbol{V}_{\{1,2\}}[\mathbf{0}_{r_{\{1,2\}}\times(r_{\{1,3\}}-r_{\{2,3\}})}\ \boldsymbol{I}_{r_{\{1,2\}}}\mathbf{0}_{r_{\{1,2\}}\times r_{\{1\}}}\ \tilde{\mathbf{S}}_{d1}]\mathbf{S}_2+\mathbf{H}_1\boldsymbol{V}_{\{1\}}\mathbf{S}_1+\mathbf{Z}_1 \qquad (4.205)\\ &=\mathbf{H}_1[\boldsymbol{V}_{\{1,2,3\}}\boldsymbol{V}_{\{1,3\}}\boldsymbol{V}_{\{1,2\}}\boldsymbol{V}_{\{1\}}]\begin{bmatrix}\boldsymbol{I}_{\{1,2,3\}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & [\mathbf{0}_{r_{\{1,2,3\}}\times(r_2-r_1)}\ \tilde{\mathbf{S}}_{d2[1:r_{\{1,2,3\}}]}\ \mathbf{C}]\\ \mathbf{0} & \boldsymbol{I}_{r_{\{1,3\}}} & \mathbf{0} & \mathbf{0} & [\tilde{\mathbf{S}}_{d21}\ \mathbf{D}]\\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}_{r_{\{1,2\}}} & \mathbf{0} & [\tilde{\mathbf{S}}_{d1}\ \mathbf{E}]\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_{r_{\{1\}}} & \mathbf{S}_{d1}\end{bmatrix}\\ &\quad+\mathbf{Z}_1. \qquad (4.206)\end{aligned}$$

where $\mathbf{C} := [\boldsymbol{I}_{r_{\{1,2,3\}}}\ \mathbf{0}\,\tilde{\mathbf{S}}_{d2[1:r_{\{1,2,3\}}]}]\mathbf{S}_{d3}$, $\mathbf{D} := [\mathbf{0}_{r_{\{1,3\}}\times r_{\{1,2,3\}}}\ \boldsymbol{I}_{r_{\{1,3\}}}\ \mathbf{0}_{r_{\{1,3\}}\times(r_{\{1,2\}}+r_{\{1\}})}\ \tilde{\mathbf{S}}_{d21}]\mathbf{S}_{d3}$, and $\mathbf{E} := [\mathbf{0}_{r_{\{1,2\}}\times(r_{\{1,3\}}-r_{\{2,3\}})}\ \boldsymbol{I}_{r_{\{1,2\}}}\ \mathbf{0}_{r_{\{1,2\}}\times r_{\{1\}}}\ \tilde{\mathbf{S}}_{d2}]$. User 1 learns the equivalent channel $\mathbf{H}_1[\boldsymbol{V}_{\{1,2,3\}}\ \boldsymbol{V}_{\{1,3\}}\ \boldsymbol{V}_{\{1,2\}}\ \boldsymbol{V}_{\{1\}}]$ in the first r_1 channel uses then decode $\tilde{\mathbf{S}}_{d21}$, $\tilde{\mathbf{S}}_{d1}$ and $\mathbf{S}_{d1}$ to achieve $r_{\{1,3\}}\frac{r_3-r_1}{T}+r_{\{1,2\}}\frac{r_2-r_1}{T}+r_{\{1\}}(1-\frac{r_1}{T})$ DoF in total. Therefore, the 3-user BC can achieve the DoF triple

$$\begin{aligned}\Bigg(r_3\Big(1-\frac{r_3}{T}\Big),&\Big(r_{\{1,2,3\}}+r_{\{2,3\}}\Big)\frac{r_3-r_2}{T}+\Big(r_{\{1,2\}}+r_{\{2\}}\Big)\Big(1-\frac{r_2}{T}\Big),\\ &r_{\{1,3\}}\frac{r_3-r_1}{T}+r_{\{1,3\}}\frac{r_3-r_1}{T}+r_{\{1\}}\Big(1-\frac{r_1}{T}\Big)\Bigg). \qquad (4.207)\end{aligned}$$

Following the same line, for the general K-user case such that $r_K \geq r_{K-1} \geq \cdots \geq r_1$, we can achieve the DoF tuple given in (4.195). □

4.6 Closing Remarks

TABLE 4.4: The achievable DoF and rate regions of the noncoherent MIMO BC in spatially correlated generic block fading

Scenario	Result	Stated in
Singe user	Achievable rate	Theorem 4.2
Two users	Achievable DoF region	Theorem 4.1
	Achievable rate region with rate splitting	Theorem 4.3, Theorem 4.4
	Achievable rate region with product superposition	Theorem 4.5
	Achievable rate region with hybrid superposition	Theorem 4.6
K users	Achievable DoF region with rate splitting	Theorems 4.7, 4.8
	Achievable DoF region with hybrid superposition	Theorem 4.9

In this chapter, we study the noncoherent MIMO BC with spatial correlation. It is shown that transmit correlation diversity, i.e., the difference in spatial correlations experienced by different users, is beneficial. We study various scenarios with mutually nonexclusive eigenspaces of the channel correlation matrices and propose pilot-based schemes that achieve gains over conventional transmissions. The proposed schemes use pre-beamforming (with rate splitting), product superposition, and a composition thereof. The main idea

is to allow users to communicate in the nonoverlapping eigen-directions simultaneously without causing interference and get some additional gain through the overlapping part. Our schemes provide methods to exploit transmit correlation diversity to improve the achievable DoF and rate of a MIMO BC using only statistical CSI. The proposed achievable DoF and rate regions are summarized in Table 4.4.

Part II

Transceiver Design for Noncoherent Communications

Chapter 5

Cube-Split: A Structured Grassmannian Constellation

In this chapter, we propose a practical structured Grassmannian constellation for noncoherent communication with a single transmit antenna over a Rayleigh flat and block fading channel. The constellation is generated by partitioning the Grassmannian of lines into a collection of bent hypercubes and defining a mapping onto each of these bent hypercubes such that the resulting symbols are approximately uniformly distributed on the Grassmannian. With a reasonable choice of parameters, this so-called cube-split constellation has higher packing efficiency, represented by the minimum distance, than the existing structured constellations. Furthermore, exploiting the constellation structure, we propose low-complexity greedy symbol decoder and LLR computation, as well as an efficient way to associate it to a multilevel code with multistage decoding. Numerical results show that the performance of the cube-split constellation is close to that of a numerically optimized constellation and better than other structured constellations. It also outperforms a coherent pilot-based scheme in terms of error probability and achievable data rate in the regime of short coherence interval and large constellation size.

5.1 Overview

We consider noncoherent single-user SIMO communication in which a single-antenna transmitter transmits to an N-antenna receiver. We assume flat and block fading channel, i.e., the channel vector remains constant within each coherence block of T channel uses with $T \geq 2$ and changes independently between blocks. According to [49], the optimal strategy achieving the high-SNR capacity in IID Rayleigh fading is to transmit isotropically distributed vectors on $\mathbb{C}^T$ belonging to the Grassmannian of lines, which is the space of one-dimensional subspaces in $\mathbb{C}^T$ [144], and use the span of these vectors to carry information.[1] Thus, the constellation design for this channel can be formulated as sphere packing on the Grassmann manifold.[2] The ultimate packing criteria is to minimize the

[1] When $T < N + 1$, a further condition for achieving the capacity is that the input norm square is beta distributed; the rate achieved with *constant-norm* isotropically distributed input approaches the capacity within a constant factor of $O(\frac{\ln N}{T})$ [50].

[2] See Section 1.2.2 in the introduction for a review of the capacity-achieving signal and Grassmannian constellation design for the general MIMO case.

detection error under noisy observation. This typically amounts to maximizing the distance between the constellation points.

Although a numerical sphere-packing optimization on the Grassmannian results in constellations with a good distance spectrum [63]–[66], [68], [69], this kind of constellation has limited practical application due to the storage and ML-decoder complexity requirements. In this chapter, we go for another approach that imposes particular structure on the constellation, thus facilitates low complexity constellation mapping and demapping. We introduce a novel fully structured Grassmannian constellation. This constellation is structurally generated by partitioning the Grassmannian of lines with a collection of bent cubes and mapping the symbol's coordinates in the Euclidean space onto one of these bent cubes.[3] The main advantages of our so-called *cube-split constellation* are as follows:

1. It has a good packing efficiency: its minimum distance is larger than that of existing structured constellation and compares well with the fundamental limits.
2. It allows for a systematic decoder which has low complexity, hence can be easily implemented in practice, yet achieves near-ML performance.
3. It admits a very simple yet effective binary labeling which leads to a low bit error rate (BER).
4. It allows for an accurate LLR approximation which can be efficiently computed, and can be efficiently associated to a multilevel coding-multistage decoding (multilevel coding (MLC)-multistage decoding (MSD)) [173] scheme.

We verify by simulation that under IID Rayleigh block fading channel, in terms of error probability (with or without channel codes) and achievable data rate, our cube-split constellation achieves performance close to the numerically optimized constellation, and outperforms existing structured Grassmannian constellations and a (coherent) pilot-based scheme in the regime of short coherence interval and large constellation size.

The remainder of this chapter is organized as follows. The system model is presented and Grassmannian constellations are overviewed in Section 5.2. We describe the construction and labeling of our cube-split constellation in Section 5.3. We next propose low-complexity decoder and LLR computation, and a MLC-MSD scheme in Section 5.4. Numerical results on the error rates and achievable data rate are provided in Section 5.5. Section 5.6 concludes the chapter. We discuss the extension to the MIMO case and present the preliminaries and proofs in the appendices.

5.2 System Model and Grassmannian Constellations

5.2.1 System Model

We consider a SIMO noncoherent channel in which a single-antenna transmitter transmits to a receiver equipped with N antennas. The channel between the transmitter and the receiver is assumed to be flat and block fading with coherence interval T symbol periods ($T \geq 2$). That is, the channel vector $\mathsf{h} \in \mathbb{C}^N$ remains constant during each coherence block of T symbols, and changes to an independent realization in the next block, and so on. The distribution of h is known, but its realizations are *unknown* to both ends of the link. We assume IID Rayleigh fading, i.e., $\mathsf{h} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{I}_N)$. Within a coherence block, the transmitter

[3] Our constellation was used in [172] as a quantization codebook on the Grassmannian of lines. Although the constellation structure is similar, the labeling and LLR computation presented here do not appear in the quantization problem.

sends a signal $\mathbf{x} \in \mathbb{C}^T$, and the receiver receives the $T \times N$ signal matrix

$$\mathbf{Y} = \sqrt{PT}\mathbf{x}\mathbf{h}^{\mathsf{T}} + \mathbf{Z}, \tag{5.1}$$

where $\mathbf{Z} \in \mathbb{C}^{T\times N}$ is the additive noise with IID $\mathcal{N}_{\mathbb{C}}(0,1)$ entries independent of $\mathbf{h}$, and the block index is omitted for simplicity. We consider the power constraint $\mathbb{E}[\|\mathbf{x}\|^2] = 1$, so that the transmit power P is identified with the SNR at each receive antenna. From [50, Eq.(9)] (see (1.42) in Chapter 1), the high-SNR capacity of this channel is given by

$$C(P,N,T) = \left(1 - \frac{1}{T}\right)\log_2 P + c(N,T) + o(1) \quad \text{bpcu} \tag{5.2}$$

as $P \to \infty$, with

$$\begin{aligned} c(N,T) = \frac{1}{T}\log_2 \frac{(\underline{L}-1)!}{(N-1)!(T-1)!} + \left(1-\frac{1}{T}\right)\log_2 T \\ + \frac{\underline{L}}{T}\log_2 \frac{N}{\underline{L}} + \frac{\overline{L}}{T}(\psi(N)-1)\log_2 e, \end{aligned} \tag{5.3}$$

where $\underline{L} := \min\{N, T-1\}$, $\overline{L} := \max\{N, T-1\}$, and $\psi(\cdot)$ is Euler's digamma function. Note that this generalizes and coincides with [49, Eq.(24)] when $T \geq N+1$.

We assume that the input $\mathbf{x}$ is taken from a finite constellation $\mathcal{X}$ of size $|\mathcal{X}|$. Given an observation $\mathbf{Y} = \boldsymbol{Y}$, the ML decoder is

$$\hat{\boldsymbol{x}}^{\mathrm{ML}} = \arg\max_{\boldsymbol{x}\in\mathcal{X}} p_{\mathbf{Y}|\mathbf{x}}(\boldsymbol{Y}|\boldsymbol{x}). \tag{5.4}$$

Conditioned on $\mathbf{x} = \boldsymbol{x}$, $\mathbf{Y}$ is a Gaussian matrix with independent columns having the same covariance matrix $\boldsymbol{I}_T + PT\boldsymbol{x}\boldsymbol{x}^{\mathsf{H}}$, hence the likelihood function $p_{\mathbf{Y}|\mathbf{x}}(\boldsymbol{Y}|\boldsymbol{x})$ can be derived as

$$p_{\mathbf{Y}|\mathbf{x}}(\boldsymbol{Y}|\boldsymbol{x}) = \frac{\exp(-\mathrm{tr}\{\boldsymbol{Y}^{\mathsf{H}}(\boldsymbol{I}_T + PT\boldsymbol{x}\boldsymbol{x}^{\mathsf{H}})^{-1}\boldsymbol{Y}\})}{\pi^T \det(\boldsymbol{I}_T + PT\boldsymbol{x}\boldsymbol{x}^{\mathsf{H}})} \tag{5.5}$$

$$= \frac{\exp\left(-\|\boldsymbol{Y}\|_{\mathrm{F}}^2 + \frac{PT}{1+PT\|\boldsymbol{x}\|^2}\|\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{x}\|^2\right)}{\pi^T(1+PT\|\boldsymbol{x}\|^2)}. \tag{5.6}$$

Thus, for unit-norm input $\|\mathbf{x}\| = 1$, the ML decoder (5.4) is simply

$$\hat{\boldsymbol{x}}^{\mathrm{ML}} = \arg\max_{\boldsymbol{x}\in\mathcal{X}} \|\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{x}\|^2. \tag{5.7}$$

Assuming that all constellation symbols are equally likely to be transmitted, i.e., the input law is $p_{\mathbf{x}}(\boldsymbol{x}) = \frac{1}{|\mathcal{X}|}\mathbb{1}\{\boldsymbol{x} \in \mathcal{X}\}$, the achievable data rate is given as

$$R = \frac{1}{T} I(\mathbf{x};\mathbf{Y}) \tag{5.8}$$

$$= \frac{1}{T}\mathbb{E}\left[\log_2 \frac{p_{\mathbf{Y}|\mathbf{x}}(\mathbf{Y}|\mathbf{x})}{\frac{1}{|\mathcal{X}|}\sum_{\boldsymbol{x}\in\mathcal{X}} p_{\mathbf{Y}|\mathbf{x}}(\mathbf{Y}|\boldsymbol{x})}\right] \tag{5.9}$$

$$= \frac{\log_2|\mathcal{X}|}{T} - \frac{1}{T}\mathbb{E}\left[\log_2 \frac{\sum_{\boldsymbol{x}\in\mathcal{X}} p_{\mathbf{Y}|\mathbf{x}}(\mathbf{Y}|\boldsymbol{x})}{p_{\mathbf{Y}|\mathbf{x}}(\mathbf{Y}|\mathbf{x})}\right] \quad \text{bpcu}. \tag{5.10}$$

Here, $\frac{\log_2|\mathcal{X}|}{T}$ is the rate achievable in the noiseless case, and $\frac{1}{T}\mathbb{E}\left[\log_2 \frac{\sum_{\boldsymbol{x}\in\mathcal{X}} p_{\mathbf{Y}|\mathbf{x}}(\mathbf{Y}|\boldsymbol{x})}{p_{\mathbf{Y}|\mathbf{x}}(\mathbf{Y}|\mathbf{x})}\right]$ is the rate loss due to noise. In the large constellation regime, the achievable rate converges to the channel capacity $C = \max_{p_{\mathbf{x}}} \frac{1}{T} I(\mathbf{x};\mathbf{Y})$ as the considered constellation gets close to the optimal constellation. Thus it can be used as a performance metric as done in [70, Section V]. The expectation in (5.10) does not have a closed form in general, and we resort to the Monte Carlo method to compute R.

5.2.2 Constellations on the Grassmannian of Lines

It was shown that the high-SNR capacity (5.2) is achieved with isotropically distributed input $\mathbf{x}$ such that its distribution is invariant under rotation, i.e., $p_{\mathbf{x}}(\boldsymbol{x}) = p_{\mathbf{x}}(\boldsymbol{Q}\boldsymbol{x})$ for any $T \times T$ deterministic unitary matrix $\boldsymbol{Q}$ [49], [50]. The span of such $\mathbf{x}$ is uniformly distributed on the Grassmannian of lines $G(\mathbb{C}^T, 1)$ [144]. Motivated by this, the constellation $\mathcal{X}$ can be designed by choosing $|\mathcal{X}|$ elements of $G(\mathbb{C}^T, 1)$, represented by $|\mathcal{X}|$ unit-norm vectors $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{|\mathcal{X}|}\}$. Both $\mathbf{x}$ and the noise-free observation $\mathsf{h}_j\mathbf{x}$ at receive antenna $j \in [N]$ belong to the set $\{\lambda\mathbf{x}, \lambda \in \mathbb{C}\}$. Thus, by definition, $\mathbf{x}$ and $\mathsf{h}_j\mathbf{x}$ represent the same symbol in $G(\mathbb{C}^T, 1)$. Therefore, Grassmannian signaling guarantees error-free detection *without* CSI in the noiseless case if the symbols are not colinear. When the noise $\mathbf{Z}$ is present, since its columns are almost surely not aligned with the signal $\mathbf{x}$, the one-dimensional span of the received signal at each receive antenna deviates from that of $\mathbf{x}$ w.r.t. a distance measure,[4] leading to a detection error if $\mathbf{Y}$ is outside the decision region of the transmitted symbol. With the chordal distance $d(\boldsymbol{x}, \boldsymbol{y}) := \sqrt{1 - |\boldsymbol{x}^{\mathsf{H}}\boldsymbol{y}|^2}$, the decision regions of the optimal ML decoder (5.7) in the case $N = 1$ are the Voronoi regions defined for symbol $\boldsymbol{x}_j$, $j \in [|\mathcal{X}|]$, as

$$V_j = \Big\{\boldsymbol{c} \in G(\mathbb{C}^T, 1) \colon d(\boldsymbol{c}, \boldsymbol{x}_j) \leq d(\boldsymbol{c}, \boldsymbol{x}_l), \forall l \neq j\Big\}. \tag{5.11}$$

The constellation $\mathcal{X}$ should be designed so as to minimize the probability of decoding error.

Following the footsteps of [53], we can derive the pairwise error probability (PEP) of mistaking a symbol $\boldsymbol{x}_j$ for another symbol $\boldsymbol{x}_l$ of the ML decoder as

$$P_{j,l}^{\mathrm{ML}} = \mathbb{P}\Big(\|\mathbf{Y}^{\mathsf{H}}\boldsymbol{x}_l\|^2 > \|\mathbf{Y}^{\mathsf{H}}\boldsymbol{x}_j\|^2 \;\Big|\; \mathbf{x} = \boldsymbol{x}_j\Big) \tag{5.12}$$

$$= \frac{1}{2}\left[1 - \left(1 + \frac{4(1+PT)}{(d(\boldsymbol{x}_j, \boldsymbol{x}_l)PT)^2}\right)^{-\frac{1}{2}}\right]. \tag{5.13}$$

We can verify that the PEP is decreasing with the chordal distance. The error probability P_e^{ML} of the ML decoder can be upper bounded in terms of the PEP using the union bound as

$$P_e^{\mathrm{ML}} = \frac{1}{|\mathcal{X}|}\sum_{j=1}^{|\mathcal{X}|}\mathbb{P}(\hat{\mathbf{x}} \neq \mathbf{x} | \mathbf{x} = \boldsymbol{x}_j) \tag{5.14}$$

$$\leq \frac{1}{|\mathcal{X}|}\sum_{j=1}^{|\mathcal{X}|}\sum_{l \neq j} P_{j,l}^{\mathrm{ML}} \tag{5.15}$$

$$\leq \frac{|\mathcal{X}| - 1}{2}\left[1 - \left(1 + \frac{4(1+PT)}{(d_{\min}PT)^2}\right)^{-\frac{1}{2}}\right], \tag{5.16}$$

where $d_{\min} := \min\limits_{1 \leq j < l \leq |\mathcal{X}|} d(\boldsymbol{x}_j, \boldsymbol{x}_l)$ is the minimum pairwise chordal distance of the constellation. Therefore, maximizing the minimum pairwise distance minimizes the union bound. This leads to a commonly used constellation design criteria

$$\max_{\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{|\mathcal{X}|}\}} \; \min_{1 \leq j < l \leq |\mathcal{X}|} d(\boldsymbol{x}_j, \boldsymbol{x}_l). \tag{5.17}$$

[4] As mentioned in Appendix 1.B, there are several choices for the distance measure between subspaces, such as chordal distance, spectral distance, Fubini-Study distance, geodesic distance. For the Grassmannian of lines, these distances are equivalent up to a monotonically increasing transformation. We adopt the commonly used chordal distance.

This optimization problem can be solved numerically. The resulting constellation, however, is hard to exploit in practice due to its lack of structure. In particular, the unstructured constellations are normally used with the high-complexity ML decoder, do not admit a straightforward binary labeling, and need to be stored at both ends of the channel. On the other hand, another approach imposes particular structure on the constellation, thus facilitates low complexity constellation mapping and, probably, demapping.

Let us briefly review some existing structured Grassmannian constellations for the SIMO channel. The Fourier constellation [70] contains the rows of the discrete Fourier transform matrix. It coincides with the algebraically constructed constellation in [71, Section III-A] when $T = 2$. Unfortunately, this design still requires numerical optimization of the Fourier frequencies and needs ML decoding. The exp-map constellation [74] is obtained by mapping each vector $\boldsymbol{q}$ containing $T-1$ QAM symbols into a noncoherent symbol $\boldsymbol{x}$ via the exponential map $\boldsymbol{x} = \left[\cos(\gamma\|\boldsymbol{q}\|) \;\; -\frac{\sin(\gamma\|\boldsymbol{q}\|)}{\|\boldsymbol{q}\|}\boldsymbol{q}^{\mathsf{T}}\right]^{\mathsf{T}}$ with the homothetic factor γ given in [74, Eq.(19)]. The coprime-PSK constellation [76] has symbols of the form $\boldsymbol{x} = [x \; y \; \boldsymbol{z}^{\mathsf{T}}]^{\mathsf{T}}$ where x and y are respectively Q_x-PSK and Q_y-PSK symbols such that Q_x and Q_y are coprime, and $\boldsymbol{z} \in \mathbb{C}^{(T-2)\times 1}$ belongs to a sub-constellation $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{Q_z}\}$ such that $\boldsymbol{z}_j$ and $\boldsymbol{z}_l$ are linearly independent for any $j \neq l$. The drawback of this design is that a good choice for $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{Q_z}\}$ is not specified and one might need to numerically generate it. The constellation in [77] has a multi-layer construction: starting from an initial constellation, the layer-j symbols are generated by moving each previous-layer symbols along a set of K geodesics. Specifically, given $\boldsymbol{x}$, the new symbol $\bar{\boldsymbol{x}}$ is generated as $\bar{\boldsymbol{x}} = \boldsymbol{x}v_k \cos(\phi_j) + \boldsymbol{x}^{\perp}\sin(\phi_j)$ where v_k, $|v_k| = 1$, determines the k-th geodesic, $\sin(\phi_j)$ is the moving distance, and $\boldsymbol{x}^{\perp}$ is orthogonal to $\boldsymbol{x}$. According to [77, Theorem 2], $\{v_k\}_{k=1}^{K}$ should be evenly spread on the unit circle. However, a good choice for $\sin(\phi_j)$ is not known. Both the coprime-PSK constellation [76] and the multi-layer constellation [77] require ML decoding. The constellations in [71, Section III-B] and [72] are designed for the multi-transmit-antenna case only, while the constellation based on the Cayley transform [75] relies partly on numerical optimization of $T \times T$ matrices. The construction in [174] is possible only for constellations of size at most T^2.

In the next section, we design a novel fully structured constellation in the Grassmannian of lines.

5.3 Cube-Split Constellation

5.3.1 Design Approach

Our design approach follows by first partitioning the Grassmannian of lines into a collection of cells, then defining a mapping from an Euclidean space onto one of these cells. The details are as follows.

5.3.1.a Partitioning of the Grassmannian

We consider a set of V Grassmannian points $\{\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_V\}$ and partition the Grassmannian into V cells whereby cell i is defined as

$$S_i := \Big\{\boldsymbol{x} \in G(\mathbb{C}^T, 1) \colon d(\boldsymbol{x}, \boldsymbol{\zeta}_i) < d(\boldsymbol{x}, \boldsymbol{\zeta}_j), \forall j \in [V] \setminus \{i\}\Big\}. \tag{5.18}$$

We ignore the cell boundaries for which $d(\boldsymbol{x}, \boldsymbol{\zeta}_i) = d(\boldsymbol{x}, \boldsymbol{\zeta}_j) \leq d(\boldsymbol{x}, \boldsymbol{\zeta}_k)$ for some $i \neq j$ and any $k \notin \{i, j\}$ since this is a set of measure zero. In this way, a symbol $\boldsymbol{x}$ belongs to cell S_i

if

$$i = \arg\min_{j\in[T]} d(\boldsymbol{x}, \boldsymbol{\zeta}_j), \tag{5.19}$$

that is, $\boldsymbol{\zeta}_i$ is the closest point to $\boldsymbol{x}$ among $\{\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_V\}$. Thus, these cells correspond to the Voronoi regions associated to the initial set of points $\{\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_V\}$.

5.3.1.b Mapping from the Euclidean Space onto a Cell

The Grassmannian of lines $G(\mathbb{C}^T, 1)$ has $T-1$ complex dimensions, i.e. $2(T-1)$ real dimensions, and so do the cells. Therefore, any point on a cell can be parameterized by $2(T-1)$ real coefficients. We choose to define these coefficients in the Euclidean space of $2(T-1)$ real dimensions, and let them determine the symbol through a bijective mapping $\boldsymbol{g}_i(\cdot)$ from this Euclidean space onto the cell S_i. On the other hand, to define a grid in the Euclidean space, let A_j denote finite sets of 2^{B_j} points regularly spread on the interval $(0,1)$ in order to maximize the minimum distance within the set (B_j then denotes the number of bits necessary to characterize a point in A_j) so that

$$A_j = \left\{ \frac{1}{2^{B_j+1}}, \frac{3}{2^{B_j+1}}, \dots, \frac{2^{B_j+1}-1}{2^{B_j+1}} \right\}, \quad j \in [2(T-1)]. \tag{5.20}$$

The Cartesian product of these sets $\bigotimes_{j=1}^{2(T-1)} A_j$ is then a grid in the Cartesian product of $2(T-1)$ intervals $(0,1)$ that we denote by $(0,1)^{2(T-1)}$.

The constellation can be formally described as the collection of the mapping of this grid onto each cell $S_1, \dots, S_V$, i.e.

$$\mathcal{X} = \left\{ \boldsymbol{x} = \boldsymbol{g}_i(\boldsymbol{a}) \colon i \in [V],\ \boldsymbol{a} \in \bigotimes_{j=1}^{2(T-1)} A_j \right\}. \tag{5.21}$$

We refer to i as the *cell index* of symbol $\boldsymbol{x} = \boldsymbol{g}_i(\boldsymbol{a})$ since it identifies the cell which $\boldsymbol{x}$ belongs to, and to the components of $\boldsymbol{a}$ as the *(local) coordinates* of $\boldsymbol{x}$ since they identify the location of $\boldsymbol{x}$ within the cell (i.e. the relative position of $\boldsymbol{x}$ w.r.t. the cell center $\boldsymbol{\zeta}_i$). The constellation $\mathcal{X}$ can be seen as the collection of V deformed lattice constellations containing in total $V \prod_{i=j}^{2(T-1)} 2^{B_j}$ symbols. Note that for asymptotically large B_j, choosing the points in A_j with equal probability defines a distribution that converges weakly to the uniform distribution on $(0,1)$. Therefore, in this regime, the distribution of the points on the Euclidean grid $\bigotimes_{j=1}^{2(T-1)} A_j$ converges to a continuous uniform distribution in $(0,1)^{2(T-1)}$. On the other hand, as mentioned in Section 5.2.2, an optimal constellation has symbols uniformly distributed on the Grassmannian. Therefore, focusing on this asymptotic regime (to support large constellations), we design $\{\boldsymbol{\zeta}_i\}_{i=1}^V$ and $\{\boldsymbol{g}_i\}_{i=1}^V$ such that for any i, the image of the uniform distribution in $(0,1)^{2(T-1)}$ through $\boldsymbol{g}_i$ is uniformly distributed in S_i, i.e., we seek mappings satisfying the following property.

Property 5.1. *Let* a *be a random vector uniformly distributed on* $(0,1)^{2(T-1)}$*, then for any* $i \in [V]$*,* $\boldsymbol{g}_i(\mathsf{a})$ *is uniformly distributed on the cell* S_i *of* $G(\mathbb{C}^T, 1)$*.*

5.3.2 Constellation Specifications

We construct our constellation using the above approach as follows. We first choose the cell centers to be the canonical basis vectors $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_T\}$, so $V = T$. With this choice, the

cells $S_1, \dots, S_T$ are easily characterized since S_i can be defined in terms of the symbol's coordinates as

$$S_i := \Big\{ \boldsymbol{x} \in G(\mathbb{C}^T, 1) \colon \ |x_i| > |x_j|, \forall j \in [T] \setminus \{i\} \Big\}. \tag{5.22}$$

However, the direct manipulation of the Grassmannian cell S_i is not straightforward. Hence, we equivalently define the mapping $\boldsymbol{g}_i$ as

$$\boldsymbol{g}_i(\boldsymbol{a}) = \frac{1}{\sqrt{1 + \sum_{j=1}^{T-1} |t_j|^2}} [t_1 \ \dots \ t_{i-1} \ 1 \ t_i \ \dots \ t_{T-1}]^\mathsf{T}. \tag{5.23}$$

where $\boldsymbol{t} = [t_1 \dots t_{T-1}]^\mathsf{T}$ is obtained from $\boldsymbol{a}$ via a mapping $\boldsymbol{\xi}_{T-1} : \boldsymbol{a} \mapsto \boldsymbol{t} = [t_1 \dots t_{T-1}]^\mathsf{T}$. The rationale of this definition is that the target space of $\boldsymbol{\xi}_{T-1}$ is then an Euclidean space, which can be manipulated in a simpler fashion than the Grassmannian target space of $\boldsymbol{g}_i$. The design of the mapping $\boldsymbol{g}_i$ thus boils down to the design of $\boldsymbol{\xi}_{T-1}$. Note that, from the definition of S_i, the fact that $\boldsymbol{g}_i$ takes values in S_i is equivalent to the fact that $\boldsymbol{\xi}_{T-1}$ takes values in $\mathcal{D}(0,1)^{T-1}$ with $\mathcal{D}(0,1) := \{z \in \mathbb{C} : |z| \leq 1\}$.

Let us first define $\boldsymbol{\xi}_1$ in the case $T = 2$. In this case, $\mathbf{t}$ is a complex scalar. According to Lemma 5.4, Property 5.1 is equivalent to $\boldsymbol{\xi}_1(\boldsymbol{a})$ being distributed as a Cauchy$(0,1)$ truncated in the unit disc $\mathcal{D}(0,1)$. In that case, we show in Lemma 5.6 that Property 5.1 is satisfied by considering

$$\boldsymbol{\xi}_1(\boldsymbol{a}) := \sqrt{\frac{1 - \exp(-\frac{|w|^2}{2})}{1 + \exp(-\frac{|w|^2}{2})}} \frac{w}{|w|} \tag{5.24}$$

with $w = \mathcal{N}^{-1}(a_1) + \jmath \mathcal{N}^{-1}(a_2)$.

When $T > 2$, for $\mathbf{x}$ to be uniformly distributed on S_i, the distribution underlying $\mathbf{t}$ has dependencies which are hard to characterize. We choose to ignore these dependencies and design $\boldsymbol{\xi}_{T-1}$ in a way similar to the case $T = 2$. To this end, we apply $\boldsymbol{\xi}_1$ to $\boldsymbol{a}$ in a component-wise manner

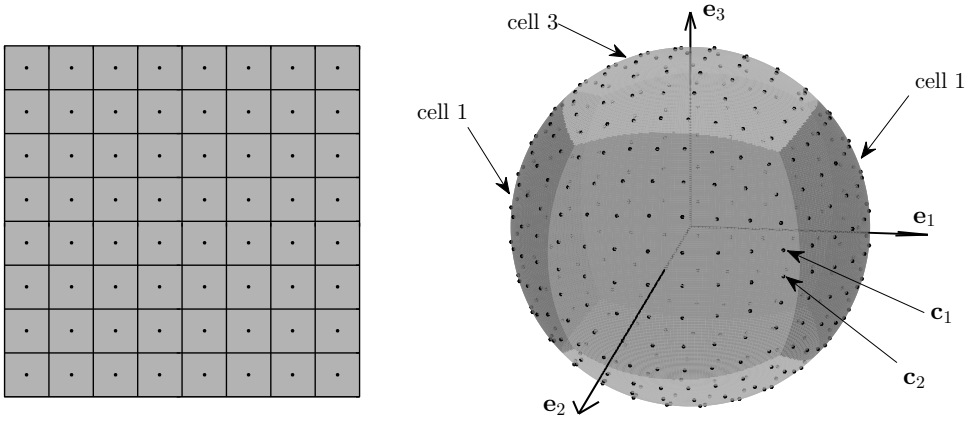
$$\boldsymbol{\xi}_{T-1}(\boldsymbol{a}) = \left[\boldsymbol{\xi}_1([a_1 \ a_2]^\mathsf{T}) \ \dots \ \boldsymbol{\xi}_1([a_{2T-3} \ a_{2T-2}]^\mathsf{T}) \right]^\mathsf{T}. \tag{5.25}$$

However, note that Property 5.1 is not satisfied in that case. Combining (5.23), (5.24) and (5.25) completely defines $\boldsymbol{g}_i$. This mapping is bijective and its inverse $\boldsymbol{a} = \boldsymbol{g}_i^{-1}(\boldsymbol{x}) : S_i \to (0,1)^{2(T-1)}$ is given by $a_{2j-1} = \mathcal{N}(\mathrm{Re}(w_j))$ and $a_{2j} = \mathcal{N}(\mathrm{Im}(w_j)), j \in [T-1]$, where

$$w_j = \sqrt{2 \ln \frac{1 + |t_j|^2}{1 - |t_j|^2}} \frac{t_j}{|t_j|}, \tag{5.26}$$

$$\boldsymbol{t} = \left[\frac{x_1}{x_i}, \dots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \dots, \frac{x_T}{x_i} \right]^\mathsf{T}. \tag{5.27}$$

The constellation is then constructed as in (5.21) where the coordinate sets $\{A_j\}$ is given in (5.20). The grid of symbols in each cell is analogous to a bent hypercube, hence the name *cube-split constellation*. The constellation contains $T \prod_{j=1}^{2(T-1)} 2^{B_j}$ symbols. An example of the grid of points in $(0,1)^{2(T-1)}$ is shown in Fig. 5.1(a) and the resulting cube-split constellation is illustrated in Fig. 5.1(b). For the sake of representation, we plot the constellation built on the *real* Grassmannian $G(\mathbb{R}^T, 1)$ following the same principle.



(a) The grid of points on $(0,1)^2$ (b) The cube-split constellation on $G(\mathbb{R}^3,1)$

FIGURE 5.1: Illustration of the cube-split constellation on $G(\mathbb{R}^3,1)$ for $B_1 = B_2 = 3$ bits. The symbols, represented by the dots, are mapped from a grid on $(0,1)^2$ (left) to one of the cells depicted by different gray levels (right). Note that each symbol in $G(\mathbb{R}^3,1)$ is depicted twice due to the sign indeterminacy associated to the Grassmannian. The constellation defines $T \times 2^{B_1+B_2} = 192$ lines in $\mathbb{R}^3$. Two symbols $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ with minimum distance are in the middle of an edge of a cell.

5.3.3 Minimum Distance

The following corollary of Lemma 1.5 provides theoretical benchmarks for the minimum distance of an optimal constellation in $G(\mathbb{C}^T,1)$ with given size.[5]

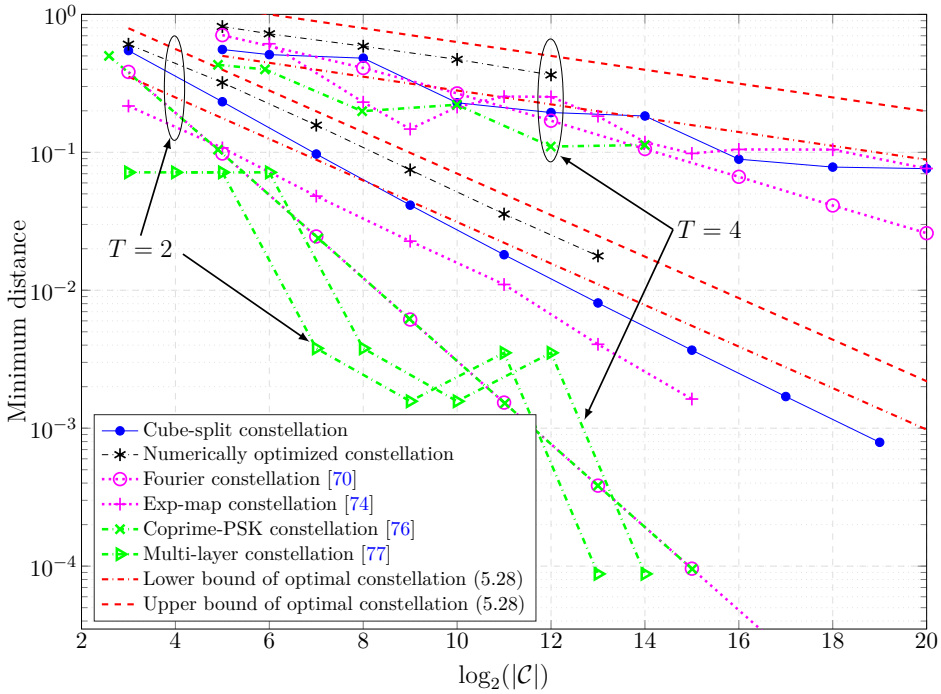
Corollary 5.1. *The minimum distance δ of an optimal constellation $\mathcal{X}_{\text{opt}}$ of cardinality $|\mathcal{X}|$ on the complex Grassmannian of lines $G(\mathbb{C}^T,1)$ is bounded by*

$$\min\left\{1, 2|\mathcal{X}|^{-\frac{1}{2(T-1)}}\right\} \geq \delta \geq |\mathcal{X}|^{-\frac{1}{2(T-1)}}. \tag{5.28}$$

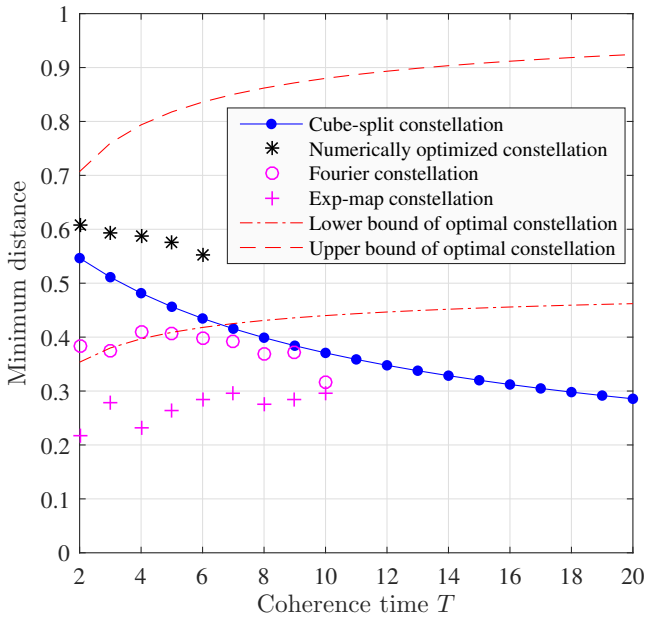
In Fig. 5.2(a), we compare the minimum distance of the cube-split constellation for $T = 2$ and $T = 4$ with these fundamental limits. We also plot the minimum distance of the numerically optimized constellation generated by approximating the optimization (5.17) by $\min_{\mathcal{X}} \ln \sum_{1\leq j<l\leq|\mathcal{X}|} \exp\left(\frac{|\boldsymbol{x}_j^{\mathsf{H}}\boldsymbol{x}_l|}{\epsilon}\right)$ with a small "diffusion constant" ϵ for smoothness, then solving it by conjugate gradient descent on the Grassmann manifold using the Manopt toolbox [141] (see Example 1.1 for details). We also compare with other structured constellations: the Fourier constellation [70], which coincides with the constellation in [71, Section III-A] when $T = 2$; the exp-map constellation [74]; the coprime-PSK constellation [76] where $\boldsymbol{z}$ is taken from a numerically optimized constellation in $G(\mathbb{C}^{T-2},1)$; and the multi-layer constellation [77] in which we use the canonical basis as the initial constellation and adopt the parameters in [77, Section IV]: $L = 3$ layers, the moving distances $\sin(\phi_2) = 0.6$ and $\sin(\phi_3) = 0.35$.

We observe that the cube-split constellation has the largest advantage over other structured constellations when $B_0 := B_1 = \cdots = B_{2(T-1)}$. (For $T = 4$, this corresponds to $\log_2(|\mathcal{X}|) = 8, 14, 20$ bits/symbol.) In this case, all the real dimensions of a cell accommodate the same number of symbols, thus the symbols are more evenly spread. Hereafter, we denote this symmetric cube-split constellation by $CS(T, B_0)$. Let us consider a pair of

[5] In our setting with $B_j \geq 1$, $j \in [2(T-1)]$, the upper bound in (5.28) is tighter than the Rankin bounds $\delta \leq \sqrt{\left(1-\frac{1}{T}\right)\frac{|\mathcal{X}|}{|\mathcal{X}|-1}}$ if $|\mathcal{X}| \leq \frac{T(T+1)}{2}$ and $\delta \leq \sqrt{1-\frac{1}{T}}$ if $|\mathcal{X}| > \frac{T(T+1)}{2}$ [60].



(a) Minimum distance vs. constellation size $|\mathcal{X}|$ for $T \in \{2, 4\}$



(b) Minimum distance vs. symbol length T for $|\mathcal{X}| = T2^{2(T-1)}$ ($B_0 = 1$)

FIGURE 5.2: The minimum distance of the cube-split constellation in comparison with other constellations and the fundamental limits of an optimal constellation given in (5.28).

symbols $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in $CS(T, B_0)$ that are in the same cell of $G(\mathbb{C}^T, 1)$ with respective local

coordinates $\boldsymbol{a}^{(1)}$ and $\boldsymbol{a}^{(2)}$ differing in only one component such that

$$a_{j_0}^{(1)} \neq a_{j_0}^{(2)}, \left\{a_{j_0}^{(1)}, a_{j_0}^{(2)}\right\} = \left\{\frac{1}{2} - \frac{1}{2^{B_0+1}}, \frac{1}{2} + \frac{1}{2^{B_0+1}}\right\}, \tag{5.29}$$

$$a_j^{(1)} = a_j^{(2)} \in \left\{\frac{1}{2^{B_0+1}}, 1 - \frac{1}{2^{B_0+1}}\right\}, \forall j \neq j_0, \tag{5.30}$$

for some $j_0 \in [2(T-1)]$. One such pair of symbols is illustrated in Fig. 5.1. The two symbols are in the middle of an edge of a cell. We conjecture, and have verified with all the cases depicted in Fig. 5.2, that this pair of symbols achieves the minimum distance of $CS(T, B_0)$. While a proof of this conjecture for general B_0 remains elusive, it can be proved for the particular case $B_0 = 1$.

Lemma 5.1. *For the $CS(T, 1)$ constellation, two symbols with minimum distance are in the same cell and have the respective local coordinates given by* (5.29) *and* (5.30) *with $B_0 = 1$. The minimum distance is given by*

$$d_{\min}(T, 1) = \sqrt{1 - \left|1 - \frac{1 + \jmath}{c^{-1} + T - 1}\right|^2}, \tag{5.31}$$

where $c := \frac{1-e^{-m^2}}{1+e^{-m^2}}$ with $m := \mathcal{N}^{-1}(3/4)$.

Proof. The proof is given in Appendix 5.A.3. □

We plot the minimum distance of the $CS(T, 1)$ constellation as a function of T in Fig. 5.2(b). For the cases where we can compute the minimum distance of the Fourier constellation and the exp-map constellation of the same size, we notice that these constellations have smaller minimum distance than that of our constellation, especially for small T.

In general, the conjectured minimum distance over the constellation, i.e., the distance between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ defined by (5.29) and (5.30), is given by

$$\tilde{d}(T, B_0) := \sqrt{1 - \left|1 + \frac{\alpha}{\alpha + \beta}(e^{\jmath 2\varphi} - 1)\right|^2}, \tag{5.32}$$

where $\alpha := \frac{1-\exp\left(-\frac{m_0^2+m_1^2}{2}\right)}{1+\exp\left(-\frac{m_0^2+m_1^2}{2}\right)}$, $\beta := 1 + (T-2)\frac{1-e^{-m_0^2}}{1+e^{-m_0^2}}$, and $\varphi := \arctan\left(\frac{m_1}{m_0}\right)$, with $m_0 := \mathcal{N}^{-1}\left(2^{-B_0-1}\right)$ and $m_1 := \mathcal{N}^{-1}\left(\frac{1}{2} + 2^{-B_0-1}\right)$.

Lemma 5.2. *When $|\mathcal{X}|$ is large, it holds that*

$$\log_2\left(\tilde{d}(T, B_0)\right) = -\frac{1}{2(T-1)}\log_2(|\mathcal{X}|) - \frac{1}{2}\log_2\log_2(|\mathcal{X}|) + O(1). \tag{5.33}$$

Proof. First, it is straightforward to see that, when $|\mathcal{X}|$ is large, i.e. B_0 is large, m_0 goes to $-\infty$ and m_1 goes to 0 and it follows that α goes to 1, β goes to $T-1$, and $\varphi = \frac{m_1}{m_0} + o\left(\frac{m_1}{m_0}\right)$. Then

$$\tilde{d}(T, B_0) = \sqrt{\frac{2\alpha}{\alpha+\beta}(1 - \cos(2\varphi) - \sin(2\varphi)) - \left(\frac{\alpha}{\alpha+\beta}\right)^2 |e^{\jmath 2\varphi} - 1|^2} \tag{5.34}$$

$$= \sqrt{\frac{4}{T}\varphi^2 - \frac{4}{T^2}\varphi^2 + o(\varphi^2)} \tag{5.35}$$

$$= 2\frac{\sqrt{T-1}}{T}|\varphi| + o(|\varphi|). \tag{5.36}$$

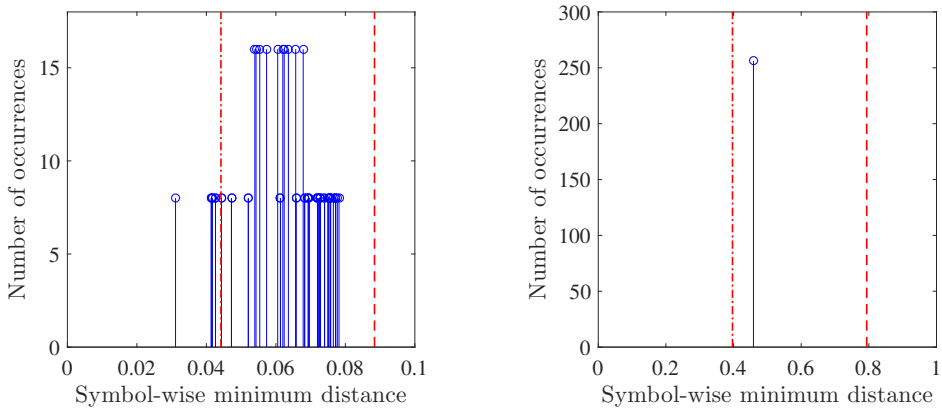
On the other hand, using [1, Section 26.2], it follows that $m_1 = \sqrt{2\pi}2^{-B_0-1} + o\left(2^{-B_0}\right)$ and $|m_0| = \sqrt{2\ln(2^{B_0+1})} + o\left(\sqrt{2\ln(2^{B_0+1})}\right)$. Inserting this and $\varphi = \frac{m_1}{m_0} + o(\frac{m_1}{m_0})$ into (5.36) gives

$$\log_2\left(\tilde{d}(T, B_0)\right) = \log_2(m_1) - \log_2(|m_0|) + O(1) \tag{5.37}$$

$$= -B_0 - \frac{1}{2}\log_2(B_0) + O(1), \tag{5.38}$$

which yields the result. □

Therefore, under the conjecture that $\tilde{d}(T, B_0)$ is the minimum distance of the $CS(T, B_0)$ constellation, the constellation is asymptotically optimal w.r.t. the bounds in Corollary 5.1 up to a log factor and a constant. In Fig. 5.3, we plot the spectrum of the symbol-wise minimum distance, i.e., the distance from each symbol to its nearest neighbor, of $CS(2, 4)$ and $CS(4, 1)$. The symbol-wise minimum distances are concentrated and compare well to the bounds. Every symbol in $CS(4, 1)$ has the same distance to its nearest neighbor, which is $d_{\min}(4, 1)$. This property holds for any $CS(T, 1)$ constellation because for any $\boldsymbol{x}_1 \in CS(T, 1)$, there exists another symbol $\boldsymbol{x}_2$ in the same cell with coordinates satisfying (5.29) and (5.30) for $B_0 = 1$, thus $d(\boldsymbol{x}_1, \boldsymbol{x}_2) = d_{\min}(T, 1)$.



(a) the $CS(2, 4)$ constellation (512 symbols) (b) the $CS(4, 1)$ constellation (256 symbols)

FIGURE 5.3: The symbol-wise minimum distance spectrum of the $CS(2, 4)$ and $CS(4, 1)$ constellations. The dashed and dash-dotted lines are respectively the upper and lower bounds (5.28) of the minimum distance of an optimal constellation of the same size.

5.3.4 Binary Labeling

Another important aspect of designing a constellation is to label each symbol with a binary vectors. If T is a power of 2, the number of bits B required to represent a symbol is

$$B = \log_2(T) + \sum_{j=1}^{2(T-1)} B_j. \tag{5.39}$$

This can be understood as a hierarchical labeling, where T bits indicate the cell index, and the remaining bits indicate the local coordinates of a symbol. If $B - \log_2 T$ is not divisible by $2(T-1)$, the sets $A_j, j \in [2(T-1)]$ cannot be chosen to have equal size. In that case, we can initially let $B_1 = \cdots = B_{2(T-1)} = \left\lfloor \frac{B-\log_2(T)}{2(T-1)} \right\rfloor$, then allocate one more bit to each

of $B - \log_2(T) - 2(T-1)\left\lfloor \frac{B-\log_2(T)}{2(T-1)} \right\rfloor$ randomly selected dimensions. If T is not a power of 2, the constellation size is also not a power of 2, which does not support a convenient bit mapping. Although we can manipulate (e.g., augment or truncate) the constellation such that the size becomes a power of 2, this alters the constellation structure. In the remainder of the chapter, we focus on the case T being a power of 2 whenever the bit mapping is concerned.

The binary labels should be assigned such that a symbol error does not cause many bit errors. This requires that symbols which are likely to be mistaken for each other should differ by a minimal number of bits in their labels. In other words, symbols with small (chordal or Euclidean) distance are given labels with small Hamming distance. This is the principle of Gray labeling which was shown to be optimal in terms of average bit error probability for structured scalar constellations such as PSK, PAM, and QAM [175], and has been widely used, such as in bit-interleaved coded modulation (BICM) [176]. Ideally, a Gray labeling scheme gives the neighboring symbols labels that differ by exactly *one* bit. As shown in [177, Theorem 1], this is possible for a special case of the Grassmannian constellation in [72]. Nevertheless, this is rarely the case in general due to the irregular neighboring properties. When a true Gray labeling is not possible, finding a quasi-Gray one requires an exhaustive search over $|\mathcal{X}|!$ candidate labelings. Therefore, one often resorts to sub-optimal labeling schemes.

An iterative labeling scheme consisting in propagating the labels along the edges of the neighboring graph was proposed in [178]. Unfortunately, building and storing such a graph is possible only for constellations of low dimension and small size. In [179], two matching methods to label a Grassmannian constellation, say $\mathcal{X}$, are proposed. The first, so-called match-and-label algorithm, matches $\mathcal{X}$ to an auxiliary constellation which can be Gray labeled. The second, so-called successive matching algorithm, matches the chordal distance spectrum of $\mathcal{X}$ with the Hamming distance spectrum of an auxiliary Gray label. However, these three schemes still have complexity at least cubic in the constellation size and do not offer any optimality guarantee.

For our cube-split constellation, we introduce a simple yet effective and efficient Gray-like labeling scheme by exploiting the constellation structure. Recall that the number of bits per symbol is $B = \log_2(T) + \sum_{j=1}^{2(T-1)} B_j$, and a symbol is entirely determined by the cell index i and the set of local coordinates $\{a_1, \ldots, a_{2(T-1)}\}$. Our labeling scheme works as follows.

- We let the first $\log_2(T)$ bits represent the cell index i and denote them by *cell bits*. These bits are defined simply as the binary representation of $i-1$. Note that no optimization of the labels of the cell index is possible since each cell have common boundaries with all other cells, as can be seen in Fig. 5.1.
- We let each of the next groups of B_j bits represent the local coordinate $a_j \in A_j$ and denote them by *coordinate bits*. Since A_j is a set of points in one dimension, there always exist Gray labels associated to its elements, in the same manner as a Gray labeling for a PAM constellation. For example, if $B_j = 3$ (as in Fig. 5.1), $A_j = \{\frac{1}{16}, \frac{3}{16}, \frac{5}{16}, \ldots, \frac{15}{16}\}$ and one option to Gray label it is $\{000, 001, 011, 010, 110, 111, 101, 100\}$, as illustrated in Fig. 5.4.

Thanks to the structured constellation design and labeling scheme outlined above, the complexity of generating a constellation symbol $\boldsymbol{x}$ from its binary representation is essentially linear in T. It can be done on-the-fly and requires no storage.

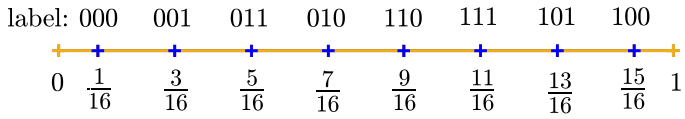


FIGURE 5.4: A visualization of a Gray labeling for the coordinate set A_j with $B_j = 3$.

In Fig. 5.5, we compare the performance in terms of BER of this Gray-like labeling scheme with random labeling, graph propagation labeling [178], match-and-label labeling and successive matching labeling [179]. For the match-and-label scheme, we use the exp-map constellation [74] as the auxiliary constellation. This constellation is mapped from coherent symbols $\boldsymbol{q} \in \mathbb{C}^{T-1}$ containing $T-1$ QAM symbols, and thus can be quasi-Gray labeled by taking the Gray label of $\boldsymbol{q}$. It can be seen that for the considered $CS(2,4)$ and $CS(4,1)$ constellations, our Gray-like labeling scheme, albeit being simpler, outperforms the other considered schemes.

5.4 Low-Complexity Receiver Design

In this section, leveraging the constellation structure, we design efficient symbol decoder and LLR computation from the observation $\mathbf{Y}$.

5.4.1 Low-Complexity Greedy Decoder

In order to avoid the high-complexity ML decoder, we propose to decode the symbol in a greedy manner by estimating sequentially the cell index i and the local coordinates $\mathbf{a}$.

5.4.1.a Step 1 - Denoising

We first use the fact that, by construction, the signal of interest is supported by a rank-1 component of $\mathbf{Y}$ (see (5.1)). We compute the left singular vector $\mathbf{u} = [\mathsf{u}_1\ \mathsf{u}_2 \dots \mathsf{u}_T]^\mathsf{T}$ corresponding to the largest singular value of $\mathbf{Y}$, which is also the solution of

$$\arg \max_{\boldsymbol{u} \in \mathbb{C}^T:\ \|\boldsymbol{u}\|^2=1} \|\mathbf{Y}^\mathsf{H}\boldsymbol{u}\|^2. \tag{5.40}$$

Observe that this is a relaxed version of the ML decoder (5.7) if we disregard the discrete nature of the constellation. Thus, $\mathbf{u}$ serves as a rough estimate of the transmitted symbol $\mathbf{x}$ on the unit sphere.

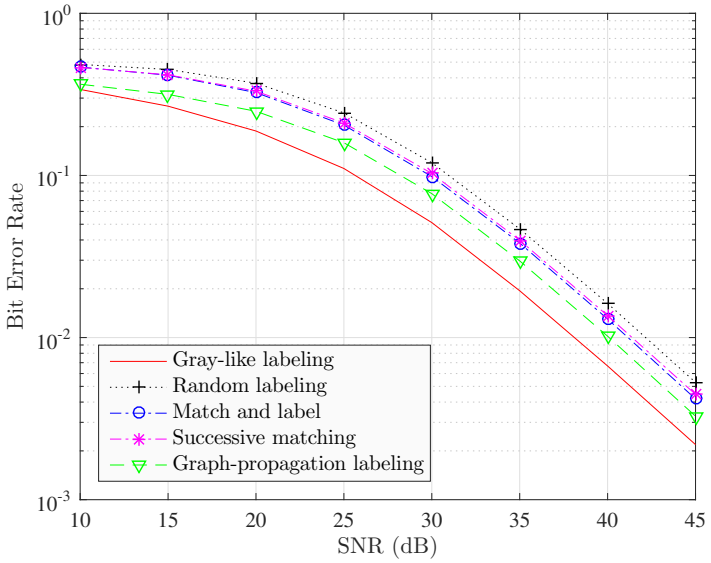
5.4.1.b Step 2 - Estimating the Cell Index and the Local Coordinates

We then find the closest symbol to $\mathbf{u}$ by localizing $\mathbf{u}$ on the system of bent grids defined for the constellation.[6] To do so, we estimate the cell index and the local coordinates. The cell index estimate is obtained as

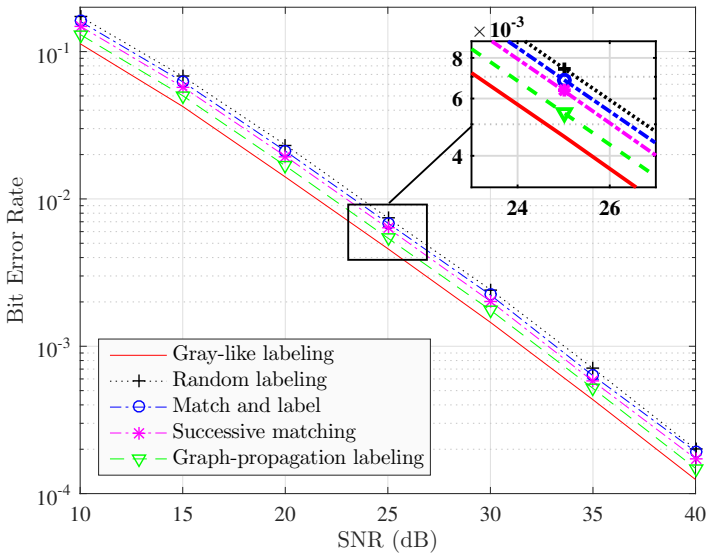
$$\hat{i} = \arg \min_{j \in [T]} d(\mathbf{u}, \boldsymbol{e}_j) = \arg \max_{j \in [T]} |\mathsf{u}_j|. \tag{5.41}$$

In the noise-free case, $\mathbf{Y} = \sqrt{PT}\mathbf{x}\mathbf{h}^\mathsf{T}$ has rank one and $\mathbf{u} = e^{\jmath\theta}\mathbf{x}$ for some $\theta \in [0, 2\pi]$, thus $\hat{i} = i$ since x_i is the strongest component in $\mathbf{x}$ by construction (see (5.23)). The local coordinates are estimated by first applying the inverse mapping $\boldsymbol{g}_{\hat{i}}^{-1}$ (see (5.26) and (5.27))

[6] This is equivalent to the problem of quantizing $\mathbf{u}$ using $\mathcal{X}$ as a quantization codebook on $G(\mathbb{C}^T, 1)$, see [172].



(a) $T = 2, B_0 = 4$ (512 symbols)



(b) $T = 4, B_0 = 1$ (256 symbols)

FIGURE 5.5: The BER of the cube-split constellation with ML decoder and different labeling schemes in a single-receive-antenna system. The proposed Gray-like labeling outperforms the other schemes.

to obtain $\tilde{\mathbf{a}} = [\tilde{\mathsf{a}}_1 \ \ldots \ \tilde{\mathsf{a}}_{2T-2}]^\mathsf{T} = \boldsymbol{g}_{\hat{i}}^{-1}(\mathbf{u})$ and then find the closest point to $\tilde{\mathbf{a}}$ in $\bigotimes_{j=1}^{2(T-1)} A_j$. This can be done component-wise as

$$\hat{\mathsf{a}}_j = \arg\min_{a \in A_j} |\tilde{\mathsf{a}}_j - a|, \quad j \in [2(T-1)]. \tag{5.42}$$

Again, in the absence of noise, $\mathbf{u} = e^{\jmath\theta}\mathbf{x}$, $\hat{i} = i$, and thus $\hat{\mathbf{a}} = \mathbf{a} = \boldsymbol{g}_i^{-1}(\mathbf{x})$.

The decoded symbol $\hat{\mathbf{x}}$ is then identified from the estimated parameters $\{\hat{i}, \hat{\mathbf{a}}\}$ as $\hat{\mathbf{x}} = \boldsymbol{g}_{\hat{i}}(\hat{\mathbf{a}})$. The cell bits are decoded by taking the binary representation of $\hat{i} - 1$. The coordinate bits are demapped from $\hat{\mathsf{a}}_i$ using the Gray code defined for A_j, independently for each real

component $j \in [2(T-1)]$. We observe that the decision regions of the proposed greedy decoder are close to the Voronoi regions (5.11), which are the optimal decision regions, as depicted in Fig. 5.6 for the constellation shown in Fig. 5.1(b) and $N = 1$.

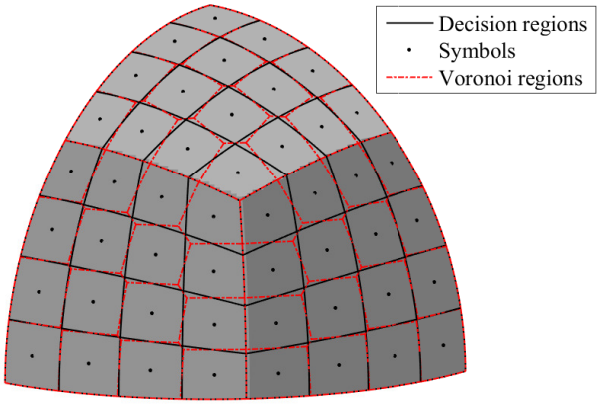


FIGURE 5.6: Illustration of the decision regions of the greedy decoder for a section (around the cell boundaries) of the cube-split constellation on $G(\mathbb{R}^3, 1)$ with $B_1 = B_2 = 3$ bits. These decision regions well match the Voronoi regions (5.11), which are the optimal decision regions.

The complexity of this greedy decoder is $O(NT \min\{N, T\})$ (dominated by the singular value decomposition of $\mathbf{Y}$ to find $\mathbf{u}$), independent of the constellation size. This is significantly reduced w.r.t the ML decoder whose complexity order is $O(NT|\mathcal{X}|)$. Among the other aforementioned structured constellations, only the exp-map constellation admits a simple sub-optimal decoder [74, Section IV-B]. That decoder solves (5.40) to find $\mathbf{u}$, then applies the inverse exponential mapping on $\mathbf{u}$ to find the QAM vector estimate $\hat{\mathbf{q}}$, and a scalar QAM demapping on $\hat{\mathbf{q}}$. It also has complexity order $O(NT \min\{N, T\})$.

5.4.2 Demapping Error Analysis

With the above greedy decoder, two types of error can occur. First, a *cell error* can occur due to false detection of the cell index i. The probability of cell error is

$$\mathbb{P}(\hat{i} \neq i) = \mathbb{P}\Big(\arg\max_j |\mathsf{u}_j| \neq \arg\max_j |\mathsf{x}_j| \Big). \tag{5.43}$$

Second, a *coordinate error* occurs when at least one of the local coordinates in $\mathbf{a}$ is wrongly detected. The probability of a coordinate error given correct cell detection is $\mathbb{P}(\hat{\mathbf{a}} \neq \mathbf{a} | \hat{i} = i)$. Then, the symbol error probability of the greedy decoder is

$$P_e = \mathbb{P}(\hat{i} \neq i) + (1 - \mathbb{P}(\hat{i} \neq i))\mathbb{P}(\hat{\mathbf{a}} \neq \mathbf{a} | \hat{i} = i). \tag{5.44}$$

The error rate can be computed analytically for the $CS(T, 1)$ constellation and $N = 1$ as follows.

Proposition 5.1. *When $N = 1$ and $B_1 = \cdots = B_{2(T-1)} = 1$, the cell error probability is*

$$\mathbb{P}(\hat{i} \neq i) = 1 - \int_0^\infty \int_0^\infty \Big(1 - Q_1(\sqrt{2cP_0x}, \sqrt{2y})\Big)^{T-1} I_0(2\sqrt{P_0xy}) e^{-y-(P_0+1)x} \, \mathrm{d}y \, \mathrm{d}x, \tag{5.45}$$

where $m := \mathcal{N}^{-1}(3/4)$, $c := \frac{1-e^{-m^2}}{1+e^{-m^2}}$, $P_0 := \frac{PT}{1+(T-1)c}$, $I_0(x) := \frac{1}{\pi}\int_0^\pi \exp(x\cos(\theta)) \, \mathrm{d}\theta$ is the modified Bessel function of the first kind at order 0 [1, Eq.(9.6.16)], and $Q_1(a, b) :=$

$\int_b^\infty x \exp\left(-\frac{x^2+a^2}{2}\right) I_0(ax)\,\mathrm{d}x$ *is the Marcum Q-function [180, Eq.(16)] with parameter* 1. *Given correct cell detection, the error probability of one pair of local coordinates is given as*

$$\begin{aligned}
&\mathbb{P}\big(\{\hat{\mathsf{a}}_{2j-1}, \hat{\mathsf{a}}_{2j}\} \neq \{\mathsf{a}_{2j-1}, \mathsf{a}_{2j}\} \big|\ \hat{i} = i\big) \\
&= \mathbb{P}(\hat{\mathsf{q}}_j \neq \mathsf{q}_j | \hat{i} = i) \\
&= 1 - \left(1 + \frac{(1-c)P_0}{\sqrt{(2+(1+c)P_0)^2 - 4cP_0^2}}\right)^{-1} \\
&\quad \times \left(\frac{1}{4} + \frac{\sqrt{2c}P_0 \operatorname{arccot} \frac{1+(c-\sqrt{\frac{c}{2}})P_0}{\sqrt{1+(c+1)P_0+\frac{c}{2}P_0^2}}}{\pi\sqrt{1+(1+c)P_0+\frac{c}{2}P_0^2}} + \frac{(1-c)P_0 \operatorname{arccot} \frac{2+(1-2\sqrt{2c}+c)P_0}{\sqrt{(2+(1+c)P_0)^2-4cP_0^2}}}{\pi\sqrt{(2+(1+c)P_0)^2-4cP_0^2}}\right),
\end{aligned} \tag{5.46}$$

for all $j \in [T-1]$. *The symbol error probability is then bounded by the union bound as*

$$P_e \leq \mathbb{P}(\hat{i} \neq i) + (T-1)(1 - \mathbb{P}(\hat{i} \neq i))\mathbb{P}(\hat{\mathsf{q}}_j \neq \mathsf{q}_j | \hat{i} = i). \tag{5.47}$$

Proof. The proof is given in Appendix 5.A.4. □

In particular, if $T = 2$, the symbol error probability can be computed in closed form as follows.

Corollary 5.2. *When* $N = 1$, $T = 2$ *and* $B_1 = \cdots = B_{2(T-1)} = 1$, *the cell error probability can be written explicitly as*

$$\mathbb{P}(\hat{i} \neq i) = \frac{1}{2}\left(1 - \frac{(1-c)P_0}{\sqrt{(2+(1+c)P_0)^2 - 4cP_0^2}}\right), \tag{5.48}$$

whereas the conditional coordinate error probability $\mathbb{P}(\hat{\mathbf{a}} \neq \mathbf{a} | \hat{i} = i)$ *is exactly the right-hand side of* (5.46). *Accordingly, the symbol error probability* P_e *is*

$$P_e = \frac{7}{8} - \frac{\sqrt{c}P_0 \operatorname{arccot} \frac{1+\left(c-\sqrt{\frac{c}{2}}\right)P_0}{\sqrt{1+(c+1)P_0+\frac{c}{2}P_0^2}}}{\pi\sqrt{2+2(1+c)P_0+cP_0^2}} - \frac{(1-c)P_0 \operatorname{arccot} \frac{2+\left(1-2\sqrt{2c}+c\right)P_0}{\sqrt{(2+(1+c)P_0)^2-4cP_0^2}}}{2\pi\sqrt{(2+(1+c)P_0)^2-4cP_0^2}}. \tag{5.49}$$

Proof. The proof is given in Appendix 5.A.5. □

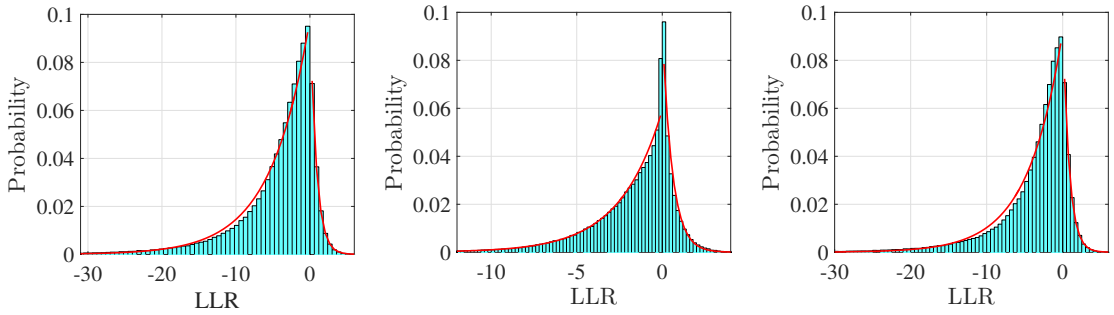
5.4.3 Log-Likelihood Ratio Computation and Code Design

When a channel code is employed, most channel decoders require the LLR of the coded bits as an input. LLR computation is performed independently from the code structure, assuming uniform input probabilities, i.e., all the symbols are equally likely to be transmitted, and so are the bits. Denote the binary label of symbol $\mathbf{x}$ as $\mathbf{b}(\mathbf{x}) = [\mathsf{b}_1(\mathbf{x})\ \mathsf{b}_2(\mathbf{x}) \ldots \mathsf{b}_B(\mathbf{x})]$. The LLR of bit b_i given the observation $\mathbf{Y} = \boldsymbol{Y}$ can be computed using (5.6) as

$$\begin{aligned}
\mathrm{LLR}_j(\boldsymbol{Y}) &= \ln \frac{p_{\mathbf{Y}|\mathsf{b}_j}(\boldsymbol{Y}|1)}{p_{\mathbf{Y}|\mathsf{b}_j}(\boldsymbol{Y}|0)} && (5.50)\\
&= \ln \frac{\sum_{\boldsymbol{\alpha} \in \mathcal{X}_j^{i^{(1)}}} p_{\mathbf{Y}|\mathbf{x}}(\boldsymbol{Y}|\boldsymbol{\alpha})}{\sum_{\boldsymbol{\beta} \in \mathcal{X}_j^{(0)}} p_{\mathbf{Y}|\mathbf{x}}(\boldsymbol{Y}|\boldsymbol{\beta})} && (5.51)\\
&= \ln \frac{\sum_{\boldsymbol{\alpha} \in \mathcal{X}_j^{(1)}} \exp\left(\frac{PT}{1+PT}\|\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{\alpha}\|^2\right)}{\sum_{\boldsymbol{\beta} \in \mathcal{X}_j^{(0)}} \exp\left(\frac{PT}{1+PT}\|\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{\beta}\|^2\right)}, && (5.52)
\end{aligned}$$

where $\mathcal{X}_j^{(b)}$ denotes the set of all possible symbols in $\mathcal{X}$ with bit j being equal to b, i.e., $\mathcal{X}_i^{(b)} := \{\boldsymbol{x} \in \mathcal{X} : \ \mathsf{b}_i(\boldsymbol{x}) = b\}$, for $i \in [B]$ and $b \in \{0, 1\}$.

In general, the LLR distribution differs between the bit positions, thus have different error protection properties. This can be seen in Fig. 5.7, where we depict the LLR histogram of the cell bit and the first two coordinate bits (the remaining two coordinate bits have the same LLR distribution as the first two due to symmetry), given that 0 was sent in that bit, for the $CS(2, 2)$ constellation and single receive antenna. It can also be observed that the LLR distribution truncated on $[0, \infty)$ (or $(-\infty, 0]$) closely fits the exponential distribution (or the flipped exponential distribution, respectively). The LLR distribution fitting can be useful, e.g., to calculate the mutual information $I(\mathsf{b}_i(\mathsf{x}); \mathrm{LLR}_i(\mathsf{Y}))$, which reveals how much information is carried in different bit positions, as well as in the framework of an extrinsic information transfer (EXIT) chart [181] analysis.



(a) 1st bit (cell bit) (b) 2nd bit (coordinate bit) (c) 3rd bit (coordinate bit)

FIGURE 5.7: Histograms of the LLR of the first 3 bits, given that 0 was sent, of the $CS(2, 2)$ constellation (5 bits/symbol), single receive antenna, and SNR = 10 dB. The red solid lines show the fitted exponential distribution of the LLR truncated on $[0, \infty)$ and the fitted flipped exponential distribution of the LLR truncated on $(-\infty, 0]$ obtained by matching the first moment (mean).

5.4.3.a Low-Complexity LLR Computation

Computing the LLR according to (5.52) requires enumerating the whole constellation. To avoid this, we propose a low-complexity LLR computation as follows. First, for any real-valued array $x_1, \ldots, x_n$, denote by $\mathcal{M}_m$ the set of m largest values, i.e., $x_j \leq x_l$ for all $x_l \in \mathcal{M}_m$ and $x_j \notin \mathcal{M}_m$. We have that

$$\ln \sum_{j=1}^{n} e^{x_j} = \ln \sum_{x \in \mathcal{M}_m} e^x + \ln\left(1 + \frac{\sum_{x_j \notin \mathcal{M}_m} e^{x_j}}{\sum_{x \in \mathcal{M}_m} e^x}\right) \tag{5.53}$$

$$\leq \ln \sum_{x \in \mathcal{M}_m} e^x + \ln\left(1 + (n - m) \exp\left(\max_{x \notin \mathcal{M}_m} x - \max_{j \in [n]} x_j\right)\right). \tag{5.54}$$

For a sufficiently large value of $\max_{j\in[n]} x_j - \max_{x\notin\mathcal{M}_m} x$, the second term in the right-hand side vanishes and $\ln \sum_{j=1}^{n} e^{x_j}$ can be well approximated[7] by $\ln \sum_{x\in\mathcal{M}_m} e^x$. Applying this to (5.52) yields

$$\mathrm{LLR}_j(\boldsymbol{Y}) \approx \ln \sum_{\boldsymbol{\alpha} \in \mathcal{M}_{\eta,j}^{(1)}} \exp\left(\frac{PT}{1 + PT} \|\boldsymbol{Y}^{\mathsf{H}} \boldsymbol{\alpha}\|^2\right) - \ln \sum_{\boldsymbol{\beta} \in \mathcal{M}_{\eta,j}^{(0)}} \exp\left(\frac{PT}{1 + PT} \|\boldsymbol{Y}^{\mathsf{H}} \boldsymbol{\beta}\|^2\right), \tag{5.55}$$

[7] When $m = 1$, this approximation coincides with the well-known one $\ln \sum_j e^{x_j} \approx \max_j x_j$.

where $\mathcal{M}_{\eta,j}^{(b)}$ stores the η symbols corresponding to the η largest terms in $\{\|\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{x}\|\}_{\boldsymbol{x}\in\mathcal{X}_j^{(b)}}$ for $b \in \{0,1\}$. Observe that one symbol in either $\mathcal{M}_{\eta,j}^{(1)}$ or $\mathcal{M}_{\eta,j}^{(0)}$ is exactly the output $\hat{\boldsymbol{x}}^{\mathrm{ML}}$ of the ML decoder (5.7). The remaining symbols in $\mathcal{M}_{\eta,j}^{(1)}$ and $\mathcal{M}_{\eta,j}^{(0)}$ are expected to be close to $\hat{\boldsymbol{x}}^{\mathrm{ML}}$ since they also lead to high likelihood of $\boldsymbol{Y}$. Furthermore, $\hat{\boldsymbol{x}}^{\mathrm{ML}}$ can be approximated by the output $\hat{\boldsymbol{x}}$ of the greedy decoder.[8] Thus, the symbols in $\mathcal{M}_{\eta,j}^{(1)}$ and $\mathcal{M}_{\eta,j}^{(0)}$ are expected to be in the neighborhood of the greedy decoder's output $\hat{\boldsymbol{x}}$. Therefore, the LLR can be further approximated by replacing $\mathcal{M}_{\eta,j}^{(1)}$ and $\mathcal{M}_{\eta,j}^{(0)}$ in (5.55) by the sets of the greedy decoder's output $\hat{\boldsymbol{x}}$ and its neighbors as

$$\mathrm{LLR}_j(\boldsymbol{Y}) \approx \ln \sum_{\boldsymbol{\alpha}\in\mathcal{B}_j(\hat{\boldsymbol{x}},1)} \exp\left(\frac{PT}{1+PT}\|\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{\alpha}\|^2\right) - \ln \sum_{\boldsymbol{\beta}\in\mathcal{B}_j(\hat{\boldsymbol{x}},0)} \exp\left(\frac{PT}{1+PT}\|\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{\beta}\|^2\right), \tag{5.56}$$

where the set $\mathcal{B}_j(\boldsymbol{x},b)$ contains η nearest symbols to $\boldsymbol{x}$ (one of them being $\boldsymbol{x}$) with the j-th bit in their labels being equal to b, i.e,

$$\mathcal{B}_j(\boldsymbol{x},b) := \left\{\hat{\boldsymbol{x}}_1,\ldots,\hat{\boldsymbol{x}}_\eta \in \mathcal{X} : \begin{cases} d(\hat{\boldsymbol{x}}_l,\boldsymbol{x}) \le d(\boldsymbol{x},\boldsymbol{x}),\ \forall \boldsymbol{x}\in\mathcal{X}\setminus\{\hat{\boldsymbol{x}}_1,\ldots,\hat{\boldsymbol{x}}_\eta\}, l\in[\eta] \\ \mathsf{b}_j(\hat{\boldsymbol{x}}_1) = \cdots = \mathsf{b}_j(\hat{\boldsymbol{x}}_\eta) = b \end{cases}\right\} \tag{5.57}$$

for $\boldsymbol{x}\in\mathcal{X}$, $b\in\{0,1\}$, and $j\in[B]$.

The sets $\mathcal{B}_j(\boldsymbol{x},b)$ can be precomputed for each $\boldsymbol{x}\in\mathcal{X}$ prior to communication (with negligible complexity) and stored at the receiver. In this way, the complexity of computing the right-hand side of (5.56) is only $O(NT\min\{N,T\}+NT\eta)$ ($O(NT\min\{N,T\})$ for the hard detection to find $\hat{\boldsymbol{x}}$ and $O(NT\eta)$ for the computation of the right-hand side of (5.56)). Alternatively, one can look for an approximation of $\mathcal{B}_j(\hat{\boldsymbol{x}},b)$ (possibly constructed on-the-fly upon detecting $\hat{\boldsymbol{x}}$) when the constellation size is too large. The latter option does not require storage but increases the complexity.

5.4.3.b Multilevel Coding and Multistage Decoding

We propose a MLC-MSD scheme [173] customized for the cube-split constellation as follows. First, each input bit stream is divided into two substreams and each substream is fed into an individual channel encoder. Note that the two encoders can have different code rates. Then, the coded bits are mapped into cube-split symbols by taking the cell bits from the output of the first encoder and the coordinate bits from the output of the second encoder. At the receiver, the cell bits are decoded first (using the exact LLR (5.52) or approximate LLR (5.56)), thus we obtain an estimate of the cell index of each transmitted symbol. Then the LLRs of the coordinate bits are computed based on the received signal and the estimated cell index in a manner similar to (5.56) except that $\mathcal{B}_j(\hat{\boldsymbol{x}},b)$ is replaced by $\mathcal{S}_j(\hat{i},b)$ where $\hat{i}$ is the estimated cell index of the corresponding transmitted symbol and $\mathcal{S}_j(\hat{i},b)$ is the set of constellation symbols in cell $S_{\hat{i}}$ with bit j being equal to b, i.e., $\mathcal{S}_j(\hat{i},b) := \{\boldsymbol{x}\in\mathcal{X}\cap S_{\hat{i}} : b_j(\boldsymbol{x}) = b\}$. This decoding structure is also similar to a turbo equalization receiver [182].

5.5 Performance Evaluation

We evaluate numerically the performance of our cube-split constellation in comparison with other constellations and a baseline (coherent) pilot-based scheme.

[8] We will see in Section 5.5 that the greedy decoder achieves near-ML performance.

5.5.1 A Baseline Pilot-Based Scheme

We consider a baseline scheme based on channel training [31]. The transmitted signal is

$$\mathbf{x} = (PT)^{-\frac{1}{2}}\left[\sqrt{P_\tau} \quad \sqrt{P_d}\mathbf{x}_d^{\mathsf{T}}\right]^{\mathsf{T}}, \tag{5.58}$$

i.e., the first symbol is constant and known to the receiver, the data symbol vector $\mathbf{x}_d = [\mathsf{x}_2 \dots \mathsf{x}_T]^{\mathsf{T}}$ is normalized such that $\mathbb{E}[\mathbf{x}_d\mathbf{x}_d^{\mathsf{H}}] = \boldsymbol{I}_{T-1}$. The power factors P_τ and P_d satisfy $P_\tau + (T-1)P_d = PT$ and can be optimized. The received signal can be written as $\mathbf{Y} = [\mathbf{y}_\tau \ \mathbf{Y}_d^{\mathsf{T}}]^{\mathsf{T}}$ where $\mathbf{y}_\tau = \sqrt{P_\tau}\mathbf{h} + \mathbf{z}_\tau$ and $\mathbf{Y}_d = \sqrt{P_d}\mathbf{x}_d\mathbf{h}^{\mathsf{T}} + \mathbf{Z}_d$ are the received signals in the training phase and data transmission phase, respectively. The receiver uses MMSE channel estimation $\hat{\mathbf{h}} = \frac{\sqrt{P_\tau}}{1+P_\tau}\mathbf{y}_\tau \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \frac{P_\tau}{1+P_\tau}\boldsymbol{I}_N)$. Let $\bar{\mathbf{h}} := \hat{\mathbf{h}}(\frac{1}{N}\mathbb{E}[\|\hat{\mathbf{h}}\|^2])^{-\frac{1}{2}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{I}_N)$ be the normalized estimate. From [31, Theorem 3], a lower bound on the achievable rate of this pilot-based scheme with IID Gaussian input $\mathbf{x}_d \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \frac{1}{T-1}\boldsymbol{I}_{T-1})$ is given as

$$\begin{aligned} &R_{\text{pilot}}(P, N, T) \\ &:= \left(1 - \frac{1}{T}\right)\mathbb{E}\left[\log_2\left(1 + P_{\text{eff}}\|\bar{\mathbf{h}}\|^2\right)\right] \qquad (5.59) \\ &= \frac{T-1}{T}\log_2(e)\sum_{n=1}^{N}\frac{(N-1)!}{(N-n)!}\left(-\frac{1}{P_{\text{eff}}}\right)^{N-n}\left[e^{1/P_{\text{eff}}}E_1\left(\frac{1}{P_{\text{eff}}}\right) + \sum_{m=1}^{N-n}(m-1)!(-P_{\text{eff}})^m\right], \qquad (5.60) \end{aligned}$$

where $P_{\text{eff}} = \frac{P_\tau P_d}{1+P_\tau+P_d}$, $E_1(x) := \int_x^\infty \frac{e^{-t}}{t}\,\mathrm{d}t$ is the exponential integral function, and (5.60) is derived using [183, p. 4.337.5]. The optimal power allocation is given by

$$P_\tau = \begin{cases} P, & \text{if } T = 2, \\ \frac{\sqrt{T-1+PT}\left(\sqrt{(T-1)(1+PT)}-\sqrt{T-1+PT}\right)}{T-2}, & \text{if } T > 2. \end{cases} \tag{5.61}$$

Let $\tilde{\mathbf{h}} = \mathbf{h} - \hat{\mathbf{h}}$ be the channel estimation error, then $\tilde{\mathbf{h}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \frac{1}{1+P_\tau}\boldsymbol{I}_N)$ and $\tilde{\mathbf{h}}$ and $\hat{\mathbf{h}}$ are uncorrelated. The output can be written as $\mathbf{Y}_d = \sqrt{P_d}\mathbf{x}_d\hat{\mathbf{h}}^{\mathsf{T}} + \hat{\mathbf{Z}}_d$, where $\hat{\mathbf{Z}}_d = \sqrt{P_d}\mathbf{x}_d\tilde{\mathbf{h}}^{\mathsf{T}} + \mathbf{Z}_d$. Given the input, the rows of $\hat{\mathbf{Z}}_d$ are independent and the j-th row follows $\mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, (1 + \frac{P_d|\mathsf{x}_j|^2}{1+P_\tau})\boldsymbol{I}_N\right)$, $j \in \{2, \dots, T\}$. Thus, the likelihood function of the output at slot $j \in \{2, \dots, T\}$ is

$$p_{\mathbf{Y}_{[j]}|\mathsf{x}_j,\hat{\mathbf{h}}}(\boldsymbol{Y}_{[j]}|x_j, \hat{\boldsymbol{h}}) = \pi^{-N}\left(1 + \frac{P_d|x_j|^2}{1+P_\tau}\right)^{-N}\exp\left(-\frac{\|\boldsymbol{Y}_{[j]} - \sqrt{P_d}x_j\hat{\boldsymbol{h}}^{\mathsf{T}}\|^2}{1 + \frac{P_d|x_j|^2}{1+P_\tau}}\right), \tag{5.62}$$

where $\mathbf{Y}_{[j]}$ denotes the j-th row of $\mathbf{Y}$.

In practice, the data symbols in $\mathbf{x}_d$ are normally taken from finite scalar constellations such as QAM or PSK in order to reduce the complexity of the ML decoder based on (5.62). A sub-optimal method consists in linear equalization followed by component-wise scalar demapping. With ZF or MMSE equalizer, the equalized symbols are respectively

$$\hat{\mathbf{x}}_d^{\text{zf}} = \frac{\mathbf{Y}_d}{\sqrt{P_d}}\frac{\hat{\mathbf{h}}^*}{\|\hat{\mathbf{h}}\|^2}, \quad \text{or} \quad \hat{\mathbf{x}}_d^{\text{mmse}} = \frac{\mathbf{Y}_d}{\sqrt{P_d}}\frac{\hat{\mathbf{h}}^*}{\|\hat{\mathbf{h}}\|^2 + 1/P_d}. \tag{5.63}$$

The LLR of bit b_i given $\mathbf{Y} = \boldsymbol{Y}$ and the channel estimate $\hat{\mathbf{h}} = \hat{\boldsymbol{h}}$ is calculated as

$$\text{LLR}_j^{\text{pilot}}(\boldsymbol{Y}, \hat{\boldsymbol{h}}) = \ln\frac{\sum_{\alpha\in\mathcal{Q}_j^{(1)}} p_{\mathbf{Y}|\mathsf{x}_{\{j\}},\hat{\mathbf{h}}}(\boldsymbol{Y}|\alpha, \hat{\boldsymbol{h}})}{\sum_{\beta\in\mathcal{Q}_j^{(0)}} p_{\mathbf{Y}|\mathsf{x}_{\{j\}},\hat{\mathbf{h}}}(\boldsymbol{Y}|\beta, \hat{\boldsymbol{h}})}, \tag{5.64}$$

where $\mathcal{Q}_j^{(b)}$, $b \in \{0,1\}$, denotes a subset of the chosen scalar constellation (e.g., QAM) with bit j being b, $\mathsf{x}_{\{j\}}$ denotes the symbol accounting for bit b_j, and $p_{\mathbf{Y}|\mathsf{x}_{\{j\}},\hat{\mathbf{h}}}$ is given in (5.62).[9]

In the remainder of this section, we compare different schemes with the same transmission rate of B bits/symbol. Having observed in Fig. 5.2 that the Fourier constellation [70] and the exp-map constellation [74] have similar or higher packing efficiency than the coprime-PSK constellation [76] and the multi-layer constellation [77], and to keep the comparison clear, we hereafter consider only the two former constellations.

5.5.2 Achievable Data Rate

In Fig. 5.8(a), we compare the achievable rate (computed as in (5.10)) of cube-split constellation with the rate of the numerically optimized constellation and the high-SNR capacity $C(P,N,T)$ given in (5.2) for $T=2$ and single receive antenna. We also include the rate lower bound $R_{\text{pilot}}(P,N,T)$ of a pilot-based scheme with Gaussian input given in (5.60), and the achievable rate of the pilot-based scheme with QAM input. The cube-split constellation can achieve almost the same rate as the numerically optimized constellation and a higher rate than the pilot-based scheme with QAM input at a given SNR. For example, at 25 dB, the cube-split constellation can achieve about 0.3 bits/channel use higher than the rate achieved with the pilot-based scheme. Furthermore, the achievable rate of a large cube-split constellation approaches the high-SNR capacity $C(P,N,T)$.

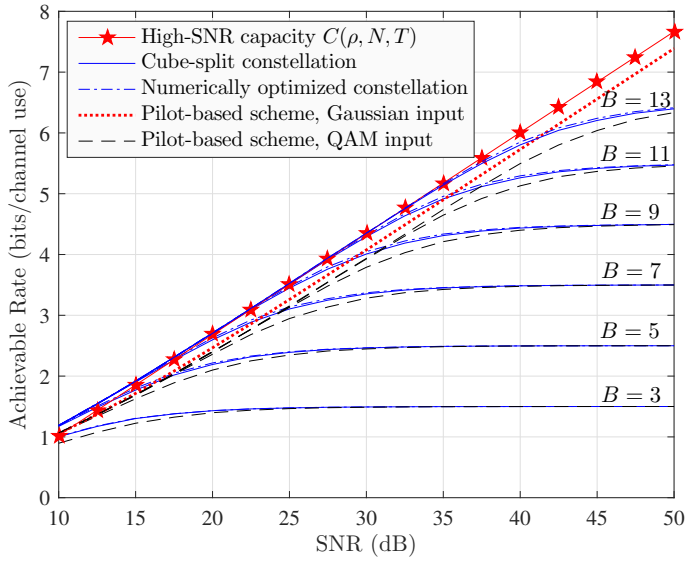
Next, in Fig. 5.8(b), we plot the achievable rate of the cube-split constellation, the numerically optimized constellation, the Fourier constellation [70], and the exp-map constellation [74], and the pilot-based scheme with QAM input for $T=4$ and $N=2$. Again, the rate achieved with cube-split constellation is close to the rate achieved with the numerically optimized constellation and higher than that of other structured constellations and the pilot-based scheme.

5.5.3 Error Rates of Uncoded Constellations

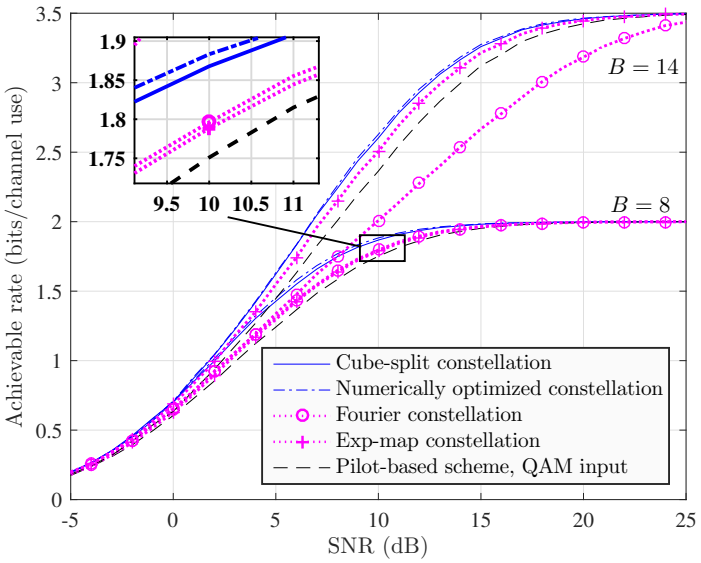
In Fig. 5.9(a) and Fig. 5.9(b), we plot the SER of the cube-split constellation (with ML or greedy decoder), the pilot-based scheme with QAM input, the numerically optimized constellation (with ML decoder), the Fourier constellation [70] (with ML decoder), and the exp-map constellation [74] (with ML or greedy decoder). In Fig. 5.9(c), we show the corresponding BER but omit the Fourier and the numerically optimized constellations for their lack of an effective binary labeling scheme. The cube-split constellation uses the Gray-like labeling in Section 5.3.4, the exp-map constellation takes the Gray label of the QAM vector $\boldsymbol{q}$ for the mapped symbol $\boldsymbol{x}$, and the pilot-based scheme uses Gray labels of the QAM symbols. We observe that the greedy decoder for the cube-split constellation achieves near-ML performance. The cube-split constellation achieves performance close to the numerically optimized constellation, and outperforms other structured constellations and the pilot-based scheme.

In Fig. 5.10, we show the SER and BER of the $CS(T,1)$ constellation, the exp-map constellation, and the pilot-based scheme with QAM input for $T \in \{8,16\}$, $N=T/2$, and $B=\log_2(T)+2(T-1)$. Note that for this large B, the numerically optimized constellation and the Fourier frequencies optimization for the Fourier constellation become infeasible.

[9] One can also compute the LLR based on the equalized symbols $\hat{\mathsf{x}}_j$. When $N=1$, the likelihood function $p_{\hat{\mathsf{x}}_j|\mathsf{x}_j}$ for ZF equalized symbols (5.63) can be derived explicitly using Lemma 5.3 as $p_{\hat{\mathsf{x}}_j^{\text{zf}}|\mathsf{x}_j}(\hat{x}_j^{\text{zf}}|x_j) = \frac{\frac{1+P_\tau}{P_\tau P_d(T-1)}+\frac{|\hat{x}_j^{\text{zf}}|^2}{P_\tau}}{\pi\left(\frac{1+P_\tau}{P_\tau P_d(T-1)}+\frac{|x_j|^2}{P_\tau}+|\hat{x}_j^{\text{zf}}-x_j|^2\right)^2}$, $j \in \{2,\dots,T\}$.



(a) $T = 2$, $N = 1$, $B \in \{3, 5, 7, 9, 11, 13\}$



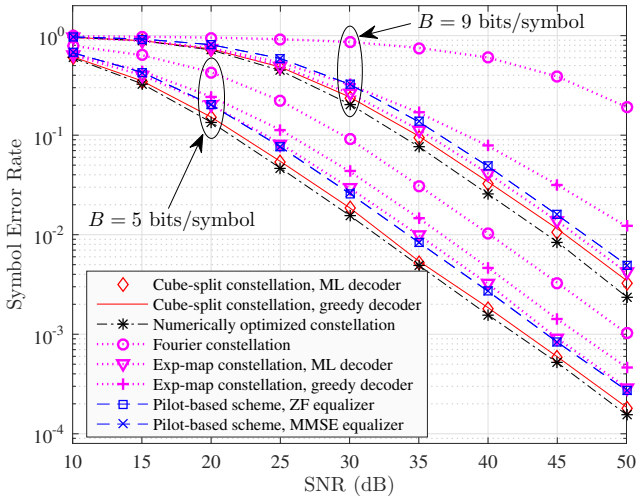
(b) $T = 4$, $N = 2$, $B \in \{8, 14\}$

FIGURE 5.8: The achievable rate (5.10) of the cube-split constellation in comparison with the channel capacity given in (5.2), and the rate achieved with the numerically optimized constellation, other structured constellations, and the pilot-based scheme with Gaussian input (5.60) or QAM input for $T \in \{2, 4\}$, $N = T/2$, and different transmission rate B bits/symbol.

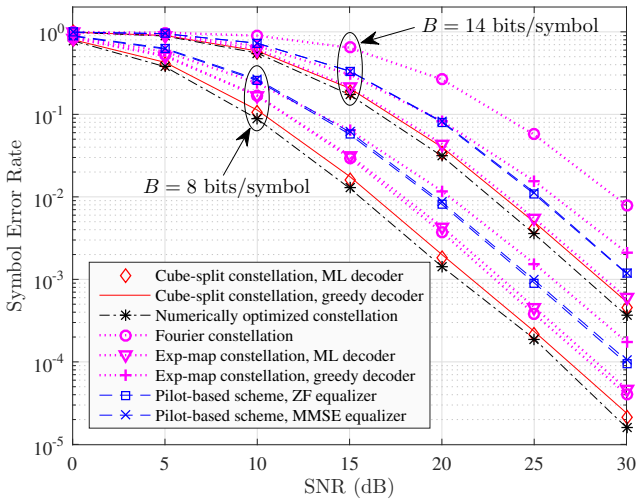
We see that in this regime, the gain of the cube-split constellation w.r.t. the exp-map constellation and the pilot-based scheme is more significant.

5.5.4 Performance with Channel Coding

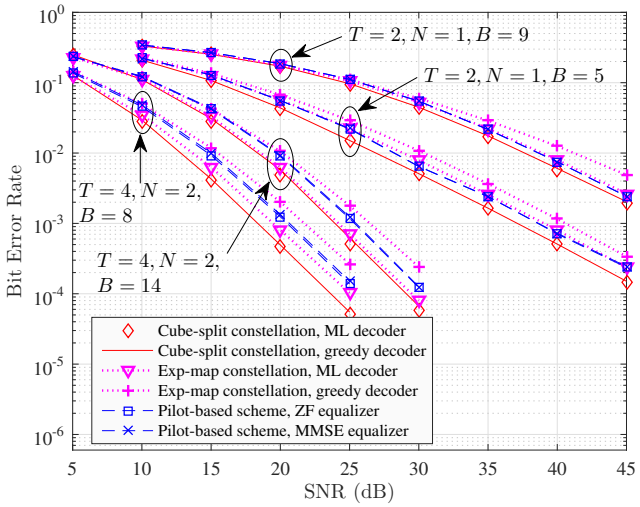
Next, we integrate a systematic parallel concatenated rate-1/3 standard turbo code [184]. The coded bits are mapped into symbols using the Gray-like labeling scheme described in Section 5.3.4. The turbo decoder calculates the LLR of the coded bits as in (5.52) or



(a) Symbol error rate, $T = 2, N = 1$

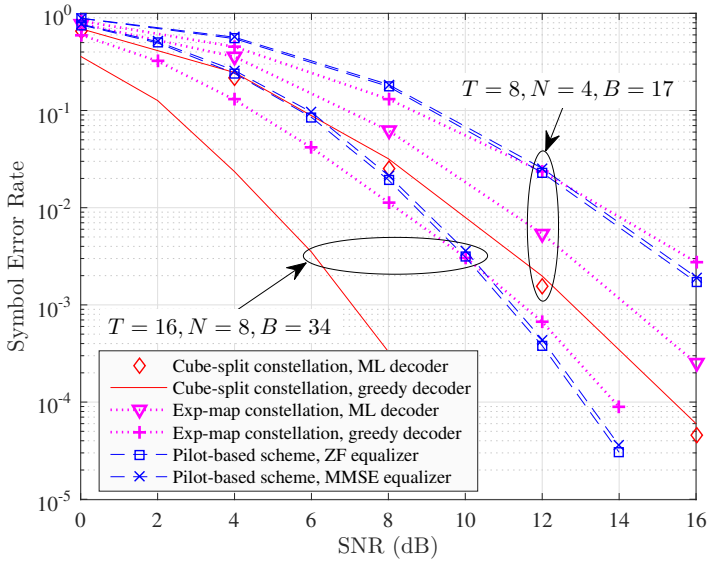


(b) Symbol error rate, $T = 4, N = 2$

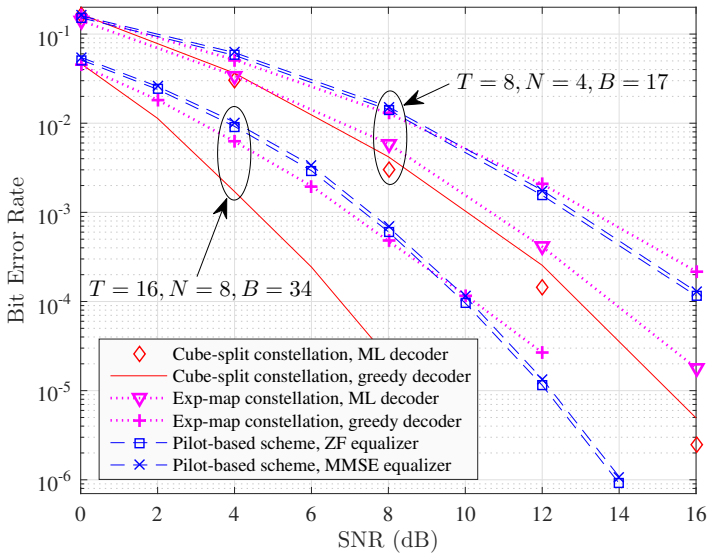


(c) Bit error rate

FIGURE 5.9: The error rates of the cube-split constellation in comparison with the numerically optimized constellation, other structured constellations, and the pilot-based scheme with QAM input for $T \in \{2, 4\}, N = T/2$, and different transmission rate B.



(a) Symbol error rate



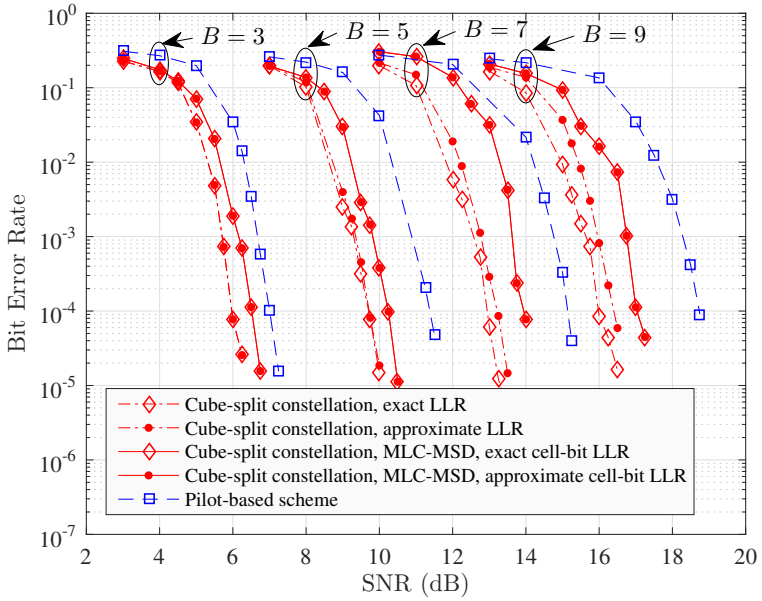
(b) Bit error rate

FIGURE 5.10: The error rates of the $CS(T, 1)$ constellation in comparison with the exp-map constellation and the pilot-based scheme with QAM input for $T \in \{8, 16\}, N = T/2$, and $B = \log_2(T) + 2(T - 1)$.

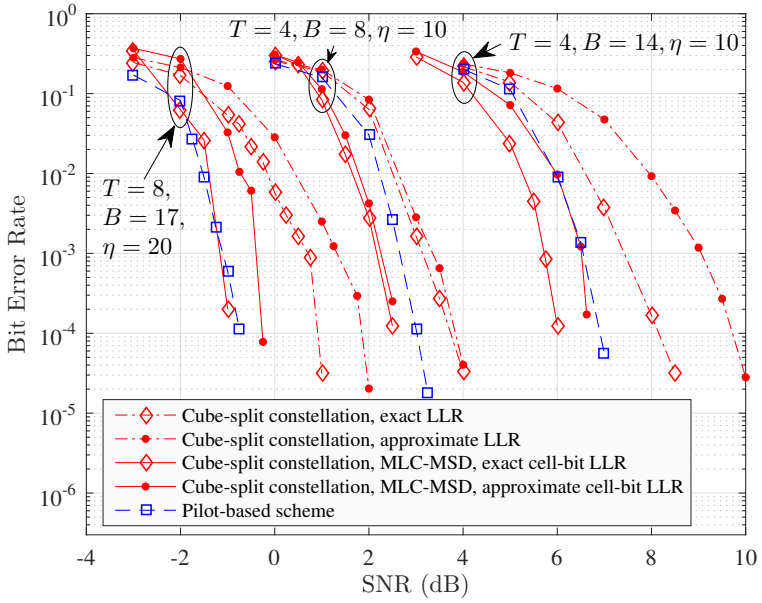
(5.56), and performs 10 decoding iterations for each packet. For the pilot-based scheme, the LLR is computed as in (5.64).

Fig. 5.11 presents the BER of the coded cube-split constellation (with exact LLR (5.52) or approximate LLR (5.56)) as compared to the coded pilot-based scheme when the turbo encoder is applied in each packet of 640 bits. We also consider the MLC-MSD scheme in which the same turbo encoder is used in both coding levels and the exact or approximate cell-bit LLR is used in the first decoding stage. For the $T = 2$ case (Fig. 5.11(a)), the BER of the cube-split constellation with approximate LLR is close to the BER with exact LLR. With the considered turbo code, the cube-split constellation outperforms the pilot-based

scheme: the power gain is about 2.5 dB for the same transmission rate of 9 bits/symbol. On the other hand, for the $T \in \{4, 8\}$ case (Fig. 5.11(b)), with the same (single-level) turbo code, the pilot-based scheme performs better than our cube-split constellation. However, with the MLC-MSD scheme, the performance of the cube-split constellation is greatly improved and can be better than that of the pilot-based scheme. This is because as the number of cells increases, the reliability of the cell bits becomes more crucial to the overall performance, and by protecting the cell bits with an individual code then using the estimated cell bits as a basis for decoding the coordinate bits, the error rate is reduced.



(a) $T = 2, N = 1, \eta = 5$



(b) $T \in \{4, 8\}, N = T/2$

FIGURE 5.11: The BER of the cube-split constellation in comparison with the pilot-based scheme with turbo codes for $T \in \{2, 4, 8\}, N = T/2$, and different transmission rate B.

In the considered setting, a channel codeword spans from 430 to 1288 channel uses. In low-latency communications, which is a relevant use case in 5G, it is favorable to transmit short packets. To this end, we consider the setting $T = 4$, $N = 1$, $B = 8$, and a channel codeword spanning 62 coherence blocks, i.e., 248 channel uses. In Fig. 5.12, we plot the *packet error rate* of our cube-split constellation (with multilevel coding) and compare with that of the exp-map constellation and the pilot-based scheme. For benchmark, we plot the saddle-point approximations of the random coding upper bound and meta converse lower bound on the lowest possible packet error rate for given packet length and transmission rate in [185]. It is observed that our cube-split constellation surpasses the exp-map constellation and the pilot-based scheme, but is still at about 2.2 dB gap from the upper bound for the same packet error rate of 10^{-3}. Note that in the literature, such as [186], [187], the gap between the considered practical coding schemes and the fundamental bounds is typically higher than 2 dB.

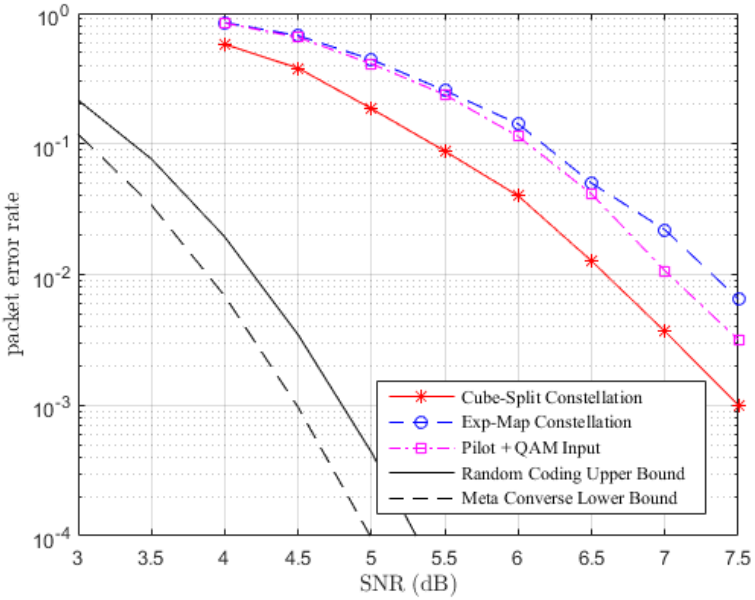


FIGURE 5.12: The packet error rate of the cube-split constellation in comparison with the pilot-based scheme and the exp-map constellation with turbo codes for $T = 4, N = 1, B = 8$ and packet length 248 channel uses.

5.6 Closing Remarks

We proposed a novel Grassmannian constellation for noncoherent SIMO communications. The structure of this constellation allows for on-the-fly symbol generation, a simple yet effective binary labeling, low-complexity symbol decoder and bit-wise LLR computation, and an efficient association with a multilevel coding-multistage decoding scheme. Analytical and numerical results show that this constellation is close to optimality in terms of packing properties, has larger minimum distance than other structured constellations in the literature. For small coherence time/large constellation size, it outperforms the coherent pilot-based approach in terms of error rates with/without channel codes and achievable data rate under Rayleigh block fading channel.

5.A Appendices

5.A.1 Extension to the MIMO Case

In general, if the transmitter has $M \leq \frac{T}{2}$ antennas, we may consider constellation symbols belonging to the Grassmannian $G(\mathbb{C}^T, M)$ represented by $T \times M$ truncated unitary matrices. To extend the cube-split design to the case $M > 1$, we would follow two essential steps: partitioning the Grassmannian into cells and defining a mapping from an Euclidean space onto a cell. To partition $G(\mathbb{C}^T, M)$, we consider a set of reference subspaces $\{\boldsymbol{E}_1, \ldots, \boldsymbol{E}_V\}$ defining an initial constellation in $G(\mathbb{C}^T, M)$ and its associated Voronoi cells

$$S_i := \{\boldsymbol{X} \in G(\mathbb{C}^T, M) : d(\boldsymbol{E}_i, \boldsymbol{X}) \leq d(\boldsymbol{E}_j, \boldsymbol{X}), \forall j \in [V] \setminus \{i\}\} \tag{5.65}$$

$$= \{\boldsymbol{X} \in G(\mathbb{C}^T, M) : \|\boldsymbol{E}_i^{\mathsf{T}}\boldsymbol{X}\|_{\mathrm{F}} \geq \|\boldsymbol{E}_j^{\mathsf{T}}\boldsymbol{X}\|_{\mathrm{F}}, \forall j \in [V] \setminus \{i\}\}, \tag{5.66}$$

where $d(\boldsymbol{Q}_1, \boldsymbol{Q}_2) := \sqrt{M - \mathrm{tr}\{\boldsymbol{Q}_1^{\mathsf{H}}\boldsymbol{Q}_2\boldsymbol{Q}_2^{\mathsf{H}}\boldsymbol{Q}_1\}}$ is the chordal distance between the subspaces spanned by the columns of $T \times M$ truncated unitary matrices $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$. The problem of choosing the initial constellation $\boldsymbol{E}_1, \ldots, \boldsymbol{E}_V$ and designing a mapping from an Euclidean space onto each cell S_i preserving a property similar to Property 5.1 is not evident and left as perspective for future work. In particular, it seems difficult to describe each Voronoi region S_i as in the case for the regions of Grassmannian of lines where the condition $\|\boldsymbol{E}_i^{\mathsf{T}}\boldsymbol{X}\|_{\mathrm{F}} \geq \|\boldsymbol{E}_j^{\mathsf{T}}\boldsymbol{X}\|_{\mathrm{F}}$ expressed on the canonical basis simply translates into a coordinate-wise condition $\left|\frac{x_j}{x_i}\right| < 1$.

5.A.2 Mathematical Preliminaries

Definition 5.1 (Univariate complex Cauchy distribution)**.** *Let $\mu \in \mathbb{C}$ and $\gamma > 0$. The probability distribution with density*

$$f(x) = \frac{1}{\pi\gamma\left[1 + \left(\frac{|x-\mu|}{\gamma}\right)^2\right]^2}, \; x \in \mathbb{C}, \tag{5.67}$$

is called the univariate complex Cauchy distribution with location μ and scale γ, denoted by $\mathrm{Cauchy}(\mu, \gamma)$.

A multivariate version of the complex Cauchy distribution is given in [188, Eq.(2)].

Lemma 5.3. $\mathrm{Cauchy}(\mu, \gamma)$ *is the distribution of the ratio $\frac{\mathsf{u}_1}{\mathsf{u}_2}$ where $[\mathsf{u}_1 \; \mathsf{u}_2]^{\mathsf{T}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = c_0\begin{bmatrix}\gamma + \|\mu\|^2 & \mu \\ \mu^* & 1\end{bmatrix}$ for some constant c_0. Specifically, if $\mathsf{u}_1 \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_1^2)$, $\mathsf{u}_2 \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_2^2)$, and $\frac{\mathbb{E}[\mathsf{u}_1\mathsf{u}_2^*]}{\sigma_1\sigma_2} =: \beta$ then $\mu = \beta\frac{\sigma_1}{\sigma_2}$ and $\gamma = (1 - |\beta|^2)\frac{\sigma_1^2}{\sigma_2^2}$.*

This relation between the Cauchy and Gaussian distributions was stated in, e.g., [188], [189].

Lemma 5.4. *The span of $\mathbf{x} = [\mathsf{x}_1 \; \mathsf{x}_2]^{\mathsf{T}}$ is uniformly distributed in $S_1 := \{\boldsymbol{x} \in G(\mathbb{C}^2, 1) : |x_1| > |x_2|\}$ if and only if the quotient $\frac{\mathsf{x}_2}{\mathsf{x}_1}$ follows a* $\mathrm{Cauchy}(0, 1)$ *distribution truncated on $\{t \in \mathbb{C} : |t| < 1\}$.*

Proof. From Lemma 5.3, if $\frac{\mathsf{x}_2}{\mathsf{x}_1} \sim \mathrm{Cauchy}(0, 1)$ then $\frac{\mathsf{x}_2}{\mathsf{x}_1}$ is identically distributed to $\frac{\mathsf{y}_2}{\mathsf{y}_1}$ with $\mathbf{y} = [\mathsf{y}_1 \; \mathsf{y}_2]^{\mathsf{T}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{I}_2)$. Since Grassmannian symbols are defined up to a scaling, $\mathrm{Span}\,\mathbf{x}$ has the same distribution as $\mathrm{Span}\,\mathbf{y}$. On the other hand, according to [188], $\mathrm{Span}\,\mathbf{y}$ is uniformly distributed in $G(\mathbb{C}^2, 1)$. Therefore, $\mathrm{Span}\,\mathbf{x}$ is uniformly distributed in $G(\mathbb{C}^2, 1)$. Furthermore, $\frac{\mathsf{x}_2}{\mathsf{x}_1} \in \{t \in \mathbb{C} : |t| < 1\}$ means $\mathrm{Span}\,\mathbf{x} \in S_1$. The converse follows from the bijectivity of the mapping $\frac{\mathsf{x}_2}{\mathsf{x}_1} \mapsto \mathrm{Span}\,\mathbf{x}$. □

Definition 5.2 (F(2, 2) distribution [190, Chap. 27])**.** *The probability distribution with PDF* $f(x) = \frac{1}{(1+x)^2}$ *and commulative distribution function (CDF)* $F(x) = \frac{x}{1+x}$, $x \geq 0$, *is called the* F*-distribution with* $(2, 2)$ *DoF denoted by* $\mathrm{F}(2, 2)$. $\mathrm{F}(2, 2)$ *is the distribution of the ratio of two independent chi-square random variables with* 2 *DoF.*

Lemma 5.5. *If* $\mathsf{x} = \mathsf{r}e^{\jmath\theta}$ *where* r *and* θ *are independent,* $\mathsf{r}^2 \sim \mathrm{F}(2, 2)$*, and* θ *is uniformly distributed in* $[0, 2\pi]$*, then* $\mathsf{x} \sim \mathrm{Cauchy}(0, 1)$.

Proof. Since r^2 and θ are independent, $f_{\mathsf{r}^2,\theta}(r^2, \theta) = f_{\mathsf{r}^2}(r^2) f_\theta(\theta) = \frac{1}{(1+r^2)^2} \frac{1}{2\pi}$. A change of variable from x to (r^2, θ) yields $\frac{1}{2} f_{\mathsf{x}}(re^{\jmath\theta}) = f_{\mathsf{r}^2,\theta}(r^2, \theta)$, so $f_{\mathsf{x}}(x) = \frac{1}{\pi(1+|x|^2)^2}$, which is the Cauchy$(0, 1)$ PDF. □

Lemma 5.6. *Let* $[\mathsf{a}_1 \ \mathsf{a}_2]^T$ *be uniformly distributed on* $(0, 1)^2$*. Then* $\mathsf{w} := \mathcal{N}^{-1}(\mathsf{a}_1) + \jmath\mathcal{N}^{-1}(\mathsf{a}_2)$ *follows a* $\mathcal{N}_{\mathbb{C}}(0, 2)$ *distribution and* $\mathsf{t} := \sqrt{\frac{1-\exp(-\frac{|\mathsf{w}|^2}{2})}{1+\exp(-\frac{|\mathsf{w}|^2}{2})}} \frac{\mathsf{w}}{|\mathsf{w}|}$ *follows a* Cauchy$(0, 1)$ *distribution truncated on* $\{t \in \mathbb{C} : \ |t| < 1\}$.

Proof. The Gaussianity of w is a standard result. Then, $|\mathsf{w}|^2$ is independent from $\frac{\mathsf{w}}{|\mathsf{w}|}$ (thus $|\mathsf{t}|^2$ is independent from $\frac{\mathsf{t}}{|\mathsf{t}|}$) and is chi-square distributed with 2 DoF, and $1 - \exp(-\frac{|\mathsf{w}|^2}{2})$ follows a uniform distribution on $(0, 1)$. Then, $\frac{1-\exp(-\frac{|\mathsf{w}|^2}{2})}{1+\exp(-\frac{|\mathsf{w}|^2}{2})}$ has CDF $F(x) = \frac{2x}{1+x}, x \in (0, 1)$. Using Definition 5.2, we see that $\frac{1-\exp(-\frac{|\mathsf{w}|^2}{2})}{1+\exp(-\frac{|\mathsf{w}|^2}{2})}$ follows a $\mathrm{F}(2, 2)$ distribution truncated on $(0, 1)$. Finally, using Lemma 5.5, we conclude that t follows a Cauchy$(0, 1)$ distribution truncated on $\{t \in \mathbb{C} : \ |t| < 1\}$. □

5.A.3 Proof of Lemma 5.1

If $B_1 = \cdots = B_{2(T-1)} = 1$, we have $A_j = \{\frac{1}{4}, \frac{3}{4}\}, \forall j \in [2(T-1)]$. Let $m := \mathcal{N}^{-1}(\frac{3}{4}) = -\mathcal{N}^{-1}(\frac{1}{4})$, then $w_j \in \{\pm m \pm \jmath m\}$ and $|w_j| = m\sqrt{2}$. Then $t_j = \sqrt{c} q_j, j \in [T-1]$, where $c := \frac{1-e^{-m^2}}{1+e^{-m^2}}$ and $q_j \in \{ \pm \frac{1}{\sqrt{2}} \pm \jmath \frac{1}{\sqrt{2}} \}$ is a 4-QAM symbol with unit power. Substituting t_j in (5.23), a constellation symbol can be written simply as

$$\boldsymbol{x}(\boldsymbol{q}) = \left(c^{-1} + T - 1\right)^{-\frac{1}{2}} \left[q_1 \ \ldots \ q_{i-1} \ c^{-\frac{1}{2}} \ q_i \ \ldots \ q_{T-1}\right]^{\mathsf{T}}. \tag{5.68}$$

Consider another symbol $\bar{\boldsymbol{x}}(\bar{\boldsymbol{q}}) \neq \boldsymbol{x}(\boldsymbol{q})$.

5.A.3.a If $\boldsymbol{x}$ and $\bar{\boldsymbol{x}}$ Are in the Same Cell

The correlation between $\boldsymbol{x}$ and $\bar{\boldsymbol{x}}$ is $\boldsymbol{x}^{\mathsf{H}}\bar{\boldsymbol{x}} = (c^{-1} + T - 1)^{-1}(c^{-1} + \sum_{i=j}^{T-1} q_j^* \bar{q}_j)$. Notice that $q_j^* \bar{q}_j \in \{\pm 1, \pm\jmath\}$, we denote by n_a the number of terms $q_j^* \bar{q}_j$ having value $a \in \{\pm 1, \pm\jmath\}$. We have that $n_1 + n_{-1} + n_\jmath + n_{-\jmath} = T - 1$ and $n_1 < T - 1$. We would like to find $\{n_a\}$ that maximize $|\boldsymbol{x}^{\mathsf{H}}\bar{\boldsymbol{x}}|^2 = (c^{-1} + T - 1)^{-2}[(c^{-1} + n_1 - n_{-1})^2 + (n_\jmath - n_{-\jmath})^2]$. The optimal $\{n_a\}$ must satisfy $\{n_1 = 0 \text{ or } n_{-1} = 0\}$ and $\{n_\jmath = 0 \text{ or } n_{-\jmath} = 0\}$ since otherwise, there always exists other $\{n_a\}$ that increases $|\boldsymbol{x}^{\mathsf{H}}\bar{\boldsymbol{x}}|^2$. Specifically, the optimal $\{n_a\}$ falls into one of two cases: $\{n_\jmath = 0 \text{ or } n_{-\jmath} = 0; n_{-1} = 0\}$ or $\{n_\jmath = 0 \text{ or } n_{-\jmath} = 0; n_1 = 0; n_{-1} \geq c^{-1}\}$. By inspecting these cases, we find that the maximal value of $|\boldsymbol{x}^{\mathsf{H}}\bar{\boldsymbol{x}}|^2$ is $(c^{-1} + T - 1)^{-2}[(c^{-1} + T - 2)^2 + 1]$ achieved with $n_1 = T - 2, (n_\jmath; n_{-\jmath}) \in \{(0; 1), (1; 0)\}$.

5.A.3.b If $\boldsymbol{x}$ Is in Cell S_i and $\bar{\boldsymbol{x}}$ Is in Cell $S_{\bar{i}}$ with $\bar{i} \neq i$

We denote $\boldsymbol{r} = [q_1 \dots q_{i-1}\ c^{-\frac{1}{2}}\ q_i \dots q_{T-1}]^{\mathsf{T}}$ and $\bar{\boldsymbol{r}} = [\bar{q}_1 \dots \bar{q}_{\bar{i}-1}\ c^{-\frac{1}{2}}\ \bar{q}_{\bar{i}} \dots \bar{q}_{T-1}]^{\mathsf{T}}$. Then $\boldsymbol{x}^{\mathsf{H}}\bar{\boldsymbol{x}} = (c^{-1}+T-1)^{-1}[c^{-\frac{1}{2}}(r_{\bar{i}}+\bar{r}_i) + \sum_{j\in[T]\setminus\{i,\bar{i}\}} r_j^*\bar{r}_j]$. Observe that $r_{\bar{i}}+\bar{r}_i \in \{0, \pm\sqrt{2}, \pm\jmath\sqrt{2}, \pm\sqrt{2}\pm\jmath\sqrt{2}\}$. By looking at each value of $r_{\bar{i}}+\bar{r}_i$ and inspecting $\{n_a\}$ as done in the previous case, we find that the maximal value of $|\boldsymbol{x}^{\mathsf{H}}\bar{\boldsymbol{x}}|^2$ is: $(c^{-1}+T-1)^{-2}(T-2)^2$ if $r_{\bar{i}}+\bar{r}_i = 0$; $(c^{-1}+T-1)^{-2}(\sqrt{2c^{-1}}+T-2)^2$ if $r_{\bar{i}}+\bar{r}_i \in \{\pm\sqrt{2}, \pm\jmath\sqrt{2}\}$; and $(c^{-1}+T-1)^{-2}(\sqrt{2c^{-1}}+T-2)^2 + 2c^{-1}$ if $r_{\bar{i}}+\bar{r}_i \in \{\pm\sqrt{2}\pm\jmath\sqrt{2}\}$. The maximal value of $|\boldsymbol{x}^{\mathsf{H}}\bar{\boldsymbol{x}}|^2$ among these is $(c^{-1}+T-1)^{-2}(\sqrt{2c^{-1}}+T-2)^2 + 2c^{-1}$.

Comparing the above two cases, we conclude that the overall maximal value of $|\boldsymbol{x}^{\mathsf{H}}\bar{\boldsymbol{x}}|^2$ is $(c^{-1}+T-1)^{-2}[(c^{-1}+T-2)^2+1] = \left|1-\frac{1+\jmath}{c^{-1}+T-1}\right|^2$ attained with $n_1 = T-2$ and $(n_\jmath; n_{-\jmath}) \in \{(0;1),(1;0)\}$, which translates to (5.29) and (5.30) with $B_0 = 1$.

5.A.4 Proof of Proposition 5.1

With $CS(T,1)$, the transmitted signal is

$$\mathsf{x} = \Big(c^{-1}+T-1\Big)^{-\frac{1}{2}}\Big[\mathsf{q}_1 \dots \mathsf{q}_{i-1}\ c^{-\frac{1}{2}}\ \mathsf{q}_i \dots \mathsf{q}_{T-1}\Big]^{\mathsf{T}}, \tag{5.69}$$

where $c = \frac{1-e^{-m^2}}{1+e^{-m^2}}$, $m = \mathcal{N}^{-1}(\frac{3}{4})$, and $\mathsf{q}_j \in \{\pm\frac{1}{\sqrt{2}} \pm \jmath\frac{1}{\sqrt{2}}\}$, $j \in [T-1]$ (see Appendix 5.A.3). The received symbols are $\mathsf{y}_i = \sqrt{P_0}\mathsf{h} + \mathsf{z}_i$, and $\mathsf{y}_j = \sqrt{cP_0}\mathsf{q}_l\mathsf{h} + \mathsf{z}_j$ for $l = j - \mathbb{1}\{j \leq i\}$ if $j \neq i$, where $\mathsf{h} \sim \mathcal{N}_{\mathbb{C}}(0,1)$, $\mathsf{z}_j \sim \mathcal{N}_{\mathbb{C}}(0,1), j \in [T]$, and $P_0 = \frac{PT}{1+(T-1)c}$.

5.A.4.a Cell Error Probability

A cell error occurs if $\exists j \neq i$ such that $|\mathsf{y}_j| > |\mathsf{y}_i|$.[10] Given h and y_i,

$$\begin{aligned}
\mathbb{P}(\hat{i} \neq i \,|\, \mathsf{h}, \mathsf{y}_i) &= \mathbb{P}(\exists j \neq i : |\mathsf{y}_j|^2 > |\mathsf{y}_i|^2 \,|\, \mathsf{h}, \mathsf{y}_i) && (5.70)\\
&= 1 - \textstyle\prod_{j\neq i} \mathbb{P}(|\mathsf{y}_j|^2 \leq |\mathsf{y}_i|^2 \,|\, \mathsf{h}, \mathsf{y}_i) && (5.71)\\
&= 1 - \Big(1 - Q_1(\sqrt{2cP_0}|\mathsf{h}|, \sqrt{2}|\mathsf{y}_i|)\Big)^{T-1}, && (5.72)
\end{aligned}$$

where $Q_1(.,.)$ is the Marcum Q-function with parameter 1. Here, (5.71) holds because conditioned on h and y_i, the events $|\mathsf{y}_j| \leq |\mathsf{y}_i|$ are mutually independent for all $j \neq i$; (5.72) is because given h, the variables $2|\mathsf{y}_j|^2$ for $j \neq i$ are independently noncentral chi-squared distributed with two DoF and noncentrality parameters $2cP_0|\mathsf{h}|^2$, denoted by $\chi_2^2(2cP_0|\mathsf{h}|^2)$.

Next, by averaging $\mathbb{P}(\hat{i} \neq i|\ \mathsf{h}, \mathsf{y}_i)$ over $|\mathsf{y}_i|^2$ and $|\mathsf{h}|^2$, taking into account that $|\mathsf{h}|^2$ is exponentially distributed with mean 1, and given h, $2|\mathsf{y}_i|^2 \sim \chi_2^2(2P_0|\mathsf{h}|^2)$, we get

$$\begin{aligned}
\mathbb{P}(\hat{i} \neq i) &= 1 - \mathbb{E}_{|\mathsf{h}|^2}\mathbb{E}_{|\mathsf{y}_i|^2|\mathsf{h}}\Big[1 - (1 - Q_1(\sqrt{2cP_0}|\mathsf{h}|, \sqrt{2}|\mathsf{y}_i|))^{T-1}\Big]\\
&= 1 - \int_0^\infty \int_0^\infty \Big[1 - (1 - Q_1(\sqrt{2cP_0}|\mathsf{h}|, \sqrt{2}|\mathsf{y}_i|))^{T-1}\Big]\\
&\qquad \cdot \exp\Big(-|\mathsf{y}_i|^2 - (1+P_0)|\mathsf{h}|^2\Big) I_0\Big(2\sqrt{P_0}|\mathsf{y}_i||\mathsf{h}|\Big)\, \mathrm{d}|\mathsf{y}_i|^2\, \mathrm{d}|\mathsf{h}|^2,
\end{aligned}$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind of order 0. From this, a simple change of variables gives (5.45).

[10] Without the additive noise z, $|\mathsf{y}_i|^2 = \frac{1+e^{-m^2}}{1-e^{-m^2}}|\mathsf{y}_j|^2 > |\mathsf{y}_j|^2$ for all $j \neq i$, and therefore, there is no cell error.

5.A.4.b Coordinate Error Probability given Correct Cell Detection

We assume that the cell index i has been correctly decoded, i.e., $\hat{i} = i$, and, without loss of generality, $i = T$. The decoding strategy for the coordinate bits is similar to a 4-QAM demapper on $\mathbf{t} = [\mathsf{t}_1 \ \dots \ \mathsf{t}_{T-1}] = \left[\frac{\mathsf{y}_1}{\mathsf{y}_T} \ \dots \ \frac{\mathsf{y}_{T-1}}{\mathsf{y}_T}\right]$. Given $\mathsf{q}_j = q_j$, we have $\mathsf{y}_j = \sqrt{cP_0}\ q_j\mathsf{h} + \mathsf{z}_j \sim \mathcal{N}_{\mathbb{C}}(0, 1 + cP_0)$ for $j < T$, $\mathsf{y}_T = \sqrt{P_0}\mathsf{h} + \mathsf{z}_T \sim \mathcal{N}_{\mathbb{C}}(0, 1 + P_0)$, and $\mathbb{E}[\mathsf{y}_j\mathsf{y}_T^*] = \sqrt{c}P_0 q_j$. Thus, according to Lemma 5.3, conditioned on $\mathsf{q}_j = q_j$, $\mathsf{t}_j = \frac{\mathsf{y}_j}{\mathsf{y}_T}$ follows the Cauchy$\Big(\frac{\sqrt{c}P_0 q_i}{(1+P_0)^2}, \big(\frac{1+cP_0}{1+P_0}\big)^2 - \frac{cP_0^2|q_i|^2}{(1+P_0)^4}\Big)$ distribution with the PDF

$$f_{\mathsf{t}_j|\mathsf{q}_j}(t|q_j) := \frac{(1+P_0)^2(1+(c+1)P_0)}{\pi\big(1+(c+1)P_0 + |(1+P_0)t - \sqrt{c}P_0 q_j|^2\big)^2}. \tag{5.73}$$

Furthermore, given that $\hat{i} = i = T$, we have $|\mathsf{y}_T| > |\mathsf{y}_j|$, $\forall j < T$. Therefore, the distribution of t_j is further truncated on the unit circle. The conditional PDF of t_j is given by

$$f_{\mathsf{t}_j|\mathsf{q}_j,\hat{i}=i=T}(t|q_j) = \frac{f_{t_j|\mathsf{q}_j}(t|q_j)}{\int_{|x|\le 1} f_{\mathsf{t}_j|\mathsf{q}_j}(x|q_j)\,\mathrm{d}x}. \tag{5.74}$$

An error happens at t_j if $\mathrm{Re}(\mathsf{t}_j)\,\mathrm{Re}(\mathsf{q}_j) < 0$ or $\mathrm{Im}(\mathsf{t}_j)\,\mathrm{Im}(\mathsf{q}_j) < 0$.[11] Therefore,

$$\mathbb{P}(\hat{\mathsf{q}}_j \neq \mathsf{q}_j \,|\, \mathsf{q}_j = q_j, \hat{i} = i = T) = 1 - \frac{\int_{\mathcal{R}_j} f_{t_j|\mathsf{q}_j}(t|q_j)\,\mathrm{d}t}{\int_{|t|\le 1} f_{\mathsf{t}_j|\mathsf{q}_j}(t|q_j)\,\mathrm{d}t}, \tag{5.75}$$

where $\mathcal{R}_j := \{t \in \mathbb{C} : |t| \le 1, \mathrm{Re}(t)\,\mathrm{Re}(\mathsf{q}_j) > 0, \mathrm{Im}(t)\,\mathrm{Im}(\mathsf{q}_j) > 0\}$. Using the polar coordinate, we have that $f_{\mathsf{t}_j|\mathsf{q}_j}(t|q_j)\,\mathrm{d}t = \tilde{f}(r,\theta,q_j) r\,\mathrm{d}r\,\mathrm{d}\theta$, where $\tilde{f}(r,\theta,q_j) := f_{\mathsf{t}_j|\mathsf{q}_j}(re^{\jmath\theta}|q_j)$. Then

$$\mathbb{P}(\hat{\mathsf{q}}_j \neq \mathsf{q}_j \,|\, \mathsf{q}_j = q_j, \hat{i} = i = T) = 1 - \frac{\int_0^1 \int_{\Theta_j} \tilde{f}(r,\theta,q_j) r\,\mathrm{d}\theta\,\mathrm{d}r}{\int_0^1 \int_0^{2\pi} \tilde{f}(r,\theta,q_j) r\,\mathrm{d}\theta\,\mathrm{d}r}, \tag{5.76}$$

where Θ_j is $[0, \pi/2]$ if $q_j = \frac{1}{\sqrt{2}} + \jmath\frac{1}{\sqrt{2}}$, $[\pi/2, \pi]$ if $q_j = -\frac{1}{\sqrt{2}} + \jmath\frac{1}{\sqrt{2}}$, $[\pi, 3\pi/2]$ if $q_j = -\frac{1}{\sqrt{2}} - \jmath\frac{1}{\sqrt{2}}$, and $[3\pi/2, 2\pi]$ if $q_j = \frac{1}{\sqrt{2}} - \jmath\frac{1}{\sqrt{2}}$. After some manipulations, we obtain that

$$\int_0^1 \int_0^{2\pi} \tilde{f}(r,\theta,q_j) r\,\mathrm{d}\theta\,\mathrm{d}r = \frac{1}{2} + \frac{(1-c)P_0}{2\sqrt{(2+(1+c)P_0)^2 - 4cP_0^2}}, \tag{5.77}$$

and for all $q_j \in \{\pm\frac{1}{\sqrt{2}} \pm \jmath\frac{1}{\sqrt{2}}\}$,

$$\begin{aligned}
&\int_0^1 \int_{\Theta_j} \tilde{f}(r,\theta,q_j) r\,\mathrm{d}\theta\,\mathrm{d}r \\
&\quad = \frac{1}{8} + \frac{\sqrt{2c}P_0\,\mathrm{arccot}\,\frac{1+\left(c-\sqrt{\frac{c}{2}}\right)P_0}{\sqrt{1+(c+1)P_0+\frac{c}{2}P_0^2}}}{2\pi\sqrt{1+(1+c)P_0+\frac{c}{2}P_0^2}} + \frac{(1-c)P_0\,\mathrm{arccot}\,\frac{2+\left(1-2\sqrt{2c}+c\right)P_0}{\sqrt{(2+(1+c)P_0)^2-4cP_0^2}}}{2\pi\sqrt{(2+(1+c)P_0)^2-4cP_0^2}}.
\end{aligned} \tag{5.78}$$

Substituting this in (5.76), and using the fact that $\mathbb{P}(\hat{\mathsf{q}}_j \neq \mathsf{q}_j \,|\, \mathsf{q}_j, \hat{i} = i = i_1) = \mathbb{P}(\hat{\mathsf{q}}_j \neq \mathsf{q}_j \,|\, \mathsf{q}_j, \hat{i} = i = i_2)$, for all $i_1, i_2 \in [T]$, we obtain (5.46). The union bound for P_e follows readily.

[11] As for the cell error, there is no coordinate error in the absence of additive noise since in this case, $\mathsf{t}_j = \sqrt{c}\ \mathsf{q}_j$.

5.A.5 Proof of Corollary 5.2

The conditional PDF $f_{\mathsf{t}|\mathsf{q}_1}(t|q_1)$ of $\mathsf{t} = \frac{\mathsf{y}_{2-i}}{\mathsf{y}_i}$ is given in (5.73). The conditional cell error is simply

$$\mathbb{P}(\hat{i} \neq i|\mathsf{q}_1 = q_1) = 1 - \mathbb{P}(|\mathsf{t}| < 1|\mathsf{q}_1 = q_1) = 1 - \int_{|t|\leq 1} f_{\mathsf{t}|\mathsf{q}_1}(t|q_1)\,\mathrm{d}t \tag{5.79}$$

By calculating the integral using polar coordinates (which results in the same value for any $q_1 \in \{\pm\frac{1}{\sqrt{2}} \pm \jmath\frac{1}{\sqrt{2}}\}$) and averaging over q_1, we obtain (5.48). Furthermore, when $T = 2$, the union bound (5.47) of P_e is tight.

Chapter 6

A Joint Constellation Design for the MIMO Multiple-Access Chanel

We consider the joint constellation design problem for the non-coherent MIMO MAC. Based on an analysis of the non-coherent ML detection error, we propose novel design criteria so as to minimize the error probability. Our first criterion is the minimum expected pairwise log-likelihood ratio over the joint constellation. From an analysis of this metric at high signal-to-noise ratio, we obtain further simplified metrics. For any given set of constellation sizes, we can optimize the proposed metrics over the set of signal matrices. Based on these criteria, we propose two simple constructions: partitioning a single-user constellation, which is effective for relatively small constellations, and precoding individual constellations of lower dimension. For a fixed joint constellation, the design metrics can be further optimized over the per-user transmit power, especially when the users transmit at different rates. We investigate the option of building each individual constellation as a Grassmannian constellation scaled by the respective transmit power. Numerical results show that our proposed metrics are meaningful, and can be used as objectives to generate constellations through numerical optimization that perform better, for the same transmission rate and power constraint, than a common pilot-based scheme and the constellations optimized with existing metrics.

6.1 Overview

In this chapter, we consider a K-user MIMO MAC in Rayleigh flat and block fading with coherence interval $T \geq 2$ where user k is equipped with M_k antennas and the receiver with N antennas. We aim to derive simple and effective joint constellation construction criteria so as to minimize the joint maximum likelihood (ML) symbol detection error.

If the users could cooperate, the system could be seen as a $M_{\text{tot}} \times N$ MIMO point-to-point non-coherent channel with $M_{\text{tot}} := \sum_{k=1}^{K} M_k$, for which the high-SNR optimal input is uniformly distributed on the Grassmannian $G(\mathbb{C}^T, M_{\text{tot}})$ [49]. Inspired by this, the joint constellation can be treated as a Grassmannian constellation in $G(\mathbb{C}^T, M_{\text{tot}})$, which leads to a design criterion mimicking sphere packing in this Grassmannian by maximizing the minimum pairwise chordal distance. Brehler and Varanasi derived the error probability of the ML detector for the MIMO MAC in [98] and analyzed the high-SNR asymptotic.

With cooperating users, this analysis led to a design criterion similar to that for a single-user channel proposed in [68, Eq.(8)] by the same authors, which is different from the max-min pairwise chordal distance criterion. However, for non-cooperating users (as we consider in this chapter), no explicit constellation design criterion was given. The joint ML pairwise error exponent can be shown to converge to the KL divergence between the output distributions conditioned on either of the symbols transmitted [59]. Based on this, a criterion consisting in maximizing the minimum KL divergence was proposed for the MAC in [97]. However, that work focuses on QAM-based space-time modulation and only uses the criterion to optimize the transmit powers and the sub-constellation assignment.

Following the approach of [98], we analyze the worst-case pairwise error probability and derive constellation design metrics. The contributions of this chapter are summarized as follows.

1. We propose a constellation design metric for the MIMO MAC, which is the minimum expected pairwise log-likelihood ratio (PLLR) over the joint constellation. This coincides with the minimum KL divergence metric in the massive MIMO regime. Our metric accounts for the fact that the individual constellations can have different sizes and different average symbol power.
2. From the dominant term of the expected PLLR at high SNR, we obtain further simplified metrics. We also propose an alternating optimization consisting in iteratively optimize one user at a time to simplify the optimization.
3. Based on our metrics, we propose two simple constructions. The first one consists in partitioning a single-user constellation and is effective for not-so-large constellations. The second one is based on precoding individual constellations of lower dimension in the SIMO case.
4. Furthermore, for a fixed joint constellation, we investigate the power optimization and find analytically the optimal per-user power optimizing the metrics in the two-user SIMO case.

For any given set of constellation sizes, we can optimize the proposed metrics over the set of signal matrices. Assuming Grassmannian signaling, we numerically solve the metric optimization problem to generate joint constellations, and compare with a pilot-based constellation and constellations optimized with existing metrics induced by the single-user metrics. Numerical results show that our proposed metrics are meaningful and effective, and the resulting constellations outperform the aforementioned baselines.

The remainder of the paper is organized as follows. In Section 6.2, we present the system model and formulate the problem. In Section 6.3, we analyze the detection error probability and derive the design metrics, based on which we propose two simple constellation constructions in Section 6.4. In Section 6.5, we address the transmit power optimization. We present the numerical results in Section 6.6 and conclude the paper in Section 6.7. The preliminaries and proofs can be found in the appendices.

6.2 System Model and Problem Formulation

We consider a MIMO MAC consisting of a receiver equipped with N antennas and K users, user k with M_k antennas, $k \in [K]$. The channel is assumed to be flat and block fading with equal-length and synchronous (across the users) coherence block of length $T \geq 2$. Furthermore, the *distribution* of the channel $\mathbf{H}_k \in \mathbb{C}^{N \times M_k}$ of user k, $k \in [K]$, is assumed to be known, but its *realizations* are unknown to both the users and the receiver. We

consider IID Rayleigh fading, namely, the rows of $\mathsf{H}:=[\mathsf{H}_1\ \mathsf{H}_2\ \dots\ \mathsf{H}_K]$ are independent and follow $\mathcal{N}_{\mathbb{C}}\Big(\mathbf{0}, \boldsymbol{I}_{M_{\text{tot}}}\Big)$. Within a coherence block, each user k sends a signal matrix symbol $\mathsf{X}_k \in \mathbb{C}^{T\times M_k}$, and the receiver observes

$$\mathsf{Y} = \sum_{k=1}^{K} \mathsf{X}_k \mathsf{H}_k^{\mathsf{T}} + \mathsf{Z}, \tag{6.1}$$

where the additive noise $\mathsf{Z} \in \mathbb{C}^{T\times N}$ has IID $\mathcal{N}_{\mathbb{C}}(0,1)$ entries independent of H_k, $k \in [K]$, and the block index is omitted for simplicity. We consider the power constraint $\mathbb{E}\big[\|\mathsf{X}_k\|_{\text{F}}^2\big] \leq PT, k \in [K]$. Thus, P is the SNR of the transmitted signal of each user at each receive antenna.

We assume that the transmitted symbol X_k takes value from a *finite constellation* $\mathcal{X}_k$ with equally likely symbols of fixed size $|\mathcal{X}_k| = 2^{R_k T}$, where R_k bpcu is the transmission rate. To satisfy the power constraint, we assume that $\frac{1}{|\mathcal{X}_k|}\sum_{\boldsymbol{X}_k\in\mathcal{X}_k} \|\boldsymbol{X}_k\|_{\text{F}}^2 = P_k T \leq PT, k \in [K]$. We assume without loss of generality or optimality that $P = \max_k P_k$. Let us rewrite (6.1) as

$$\mathsf{Y} = [\mathsf{X}_1\ \mathsf{X}_2\ \dots\ \mathsf{X}_K][\mathsf{H}_1\ \mathsf{H}_2\ \dots\ \mathsf{H}_K]^{\mathsf{T}} + \mathsf{Z} = \mathsf{X}\mathsf{H}^{\mathsf{T}} + \mathsf{Z}s, \tag{6.2}$$

where the concatenated signal matrix $\mathsf{X} := [\mathsf{X}_1\ \mathsf{X}_2\ \dots\ \mathsf{X}_K]$ takes value from $\mathcal{X} := \{[\boldsymbol{X}_1\ \boldsymbol{X}_2\ \dots\ \boldsymbol{X}_K] :\ \boldsymbol{X}_k \in \mathcal{X}_k\}$, denoted by $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_K = \prod_{k=1}^K \mathcal{X}_k$. Our goal is to derive the desirable properties of the set tuple $(\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_K)$ for a given rate tuple $(R_1, R_2, \dots, R_K)$ to achieve low *symbol detection error probability.*

Remark 6.1. *In the trivial case where only one of the users has non-zero rate, the joint constellation design problem boils down to the single-user constellation design.*

Given $\mathsf{X} = \boldsymbol{X}$, the received signal Y is a Gaussian matrix with independent columns having the same covariance matrix $\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}$. Thus, the likelihood function $p_{\mathsf{Y}|\mathsf{X}}$ is given by

$$p_{\mathsf{Y}|\mathsf{X}}(\boldsymbol{Y}|\boldsymbol{X}) = \frac{\exp(-\text{tr}(\boldsymbol{Y}^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}\boldsymbol{Y}))}{\pi^{NT}\det^N(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})}. \tag{6.3}$$

Therefore, given the received symbol $\mathsf{Y} = \boldsymbol{Y}$, the joint-user ML symbol detector is

$$\Xi(\boldsymbol{Y}) = \arg\max_{\boldsymbol{X}\in\mathcal{X}} \big(-\text{tr}\big((\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}\boldsymbol{Y}\boldsymbol{Y}^{\mathsf{H}}\big) - N\ln\det(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})\big). \tag{6.4}$$

We aim to design the joint constellation $\mathcal{X}$ so as to minimize the ML detection error $P_e(\mathcal{X}) = \mathbb{P}(\Xi(\mathsf{Y}) \neq \mathsf{X})$, i.e.,

$$\mathcal{X}^* = \arg\max_{\mathcal{X}} P_e(\mathcal{X}). \tag{6.5}$$

Since the likelihood function depends on the symbol $\boldsymbol{X}$ only through $\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}$, the following proposition is straightforward.

Proposition 6.1 (Identifiability condition)**.** *With the ML detector, the joint constellation $\mathcal{X}$ must satisfy $\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}} \neq \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$ for any pair of distinct symbols $\boldsymbol{X}$ and $\boldsymbol{X}'$ in $\mathcal{X}$.*

Remark 6.2. *Although we consider IID fading, we remark that if there is correlation between the antennas of the same user,*[1] *namely, the rows of H are independent and follow the same distribution $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{R})$ with $\boldsymbol{R} := \begin{bmatrix} \boldsymbol{R}_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boldsymbol{R}_K \end{bmatrix}$ where $\boldsymbol{R}_k$ is a $M_k \times M_k$ positive definite matrix, the solution to (6.5) can be expressed as $\bar{\mathcal{X}}_k = \{\boldsymbol{X}_k\boldsymbol{R}_k^{-1/2} : \boldsymbol{X}_k \in \mathcal{X}_k^*\}$ where $\{\mathcal{X}_k^*\}_{k=1}^K$ is the solution to (6.5) for the considered uncorrelated fading but with a new power constraint $\frac{1}{n}\sum_{t=1}^n \|\mathsf{X}_k[t]\boldsymbol{R}_k^{-1/2}\|_{\text{F}}^2 \leq P_k T, k \in [K]$.*

[1] The correlation between the antennas of *different* users is not likely since the users are not co-located.

In the next section, we analyze the error probability and derive more specific design criteria.

6.3 Constellation Design Criteria

With $\mathbf{X}$ uniformly distributed in $\mathcal{X}$, $P_e(\mathcal{X})$ can be written as

$$P_e(\mathcal{X}) = \frac{1}{|\mathcal{X}|} \sum_{\boldsymbol{X}\in\mathcal{X}} \mathbb{P}(\Xi(\mathbf{Y}) \neq \boldsymbol{X} | \mathbf{X} = \boldsymbol{X}). \tag{6.6}$$

We denote the pairwise error event as $\{\boldsymbol{X} \to \boldsymbol{X}'\} := \{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\boldsymbol{X}) \leq p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\boldsymbol{X}')|\mathbf{X} = \boldsymbol{X}\}$. For any given $\mathbf{X} = \boldsymbol{X}$, the ML detection error event $\{\Xi(\mathbf{Y}) \neq \boldsymbol{X}|\mathbf{X} = \boldsymbol{X}\}$ is the union of the pairwise error events denoted by $\bigcup_{\boldsymbol{X}'\in\mathcal{X}:\boldsymbol{X}'\neq\boldsymbol{X}}\{\boldsymbol{X} \to \boldsymbol{X}'\}$. After some simple manipulations, we have the following upper and lower bounds on $P_e(\mathcal{X})$

$$\frac{1}{|\mathcal{X}|} \max_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \mathbb{P}(\boldsymbol{X} \to \boldsymbol{X}') \leq P_e(\mathcal{X}) \leq (|\mathcal{X}| - 1) \max_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \mathbb{P}(\boldsymbol{X} \to \boldsymbol{X}'). \tag{6.7}$$

We see that for a given $|\mathcal{X}|$, the symbol detection error $P_e(\mathcal{X})$ vanishes if and only if the *worst-case pairwise error probability (PEP)*, $\max_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \mathbb{P}(\boldsymbol{X} \to \boldsymbol{X}')$, vanishes. Therefore, our goal from now on is to minimize the worst-case PEP.

6.3.1 Pairwise Error Probability Analysis

Following [98], [191], the PEP $\mathbb{P}(\boldsymbol{X} \to \boldsymbol{X}')$ can be derived in closed form as follows.

Proposition 6.2. *Let $\{\lambda_l\}_{l=1}^{L}$ be the distinct non-zero eigenvalues of*

$$\boldsymbol{\Lambda} := (\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})(\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1} - \boldsymbol{I}_T, \tag{6.8}$$

with multiplicities $\{\mu_l\}_{l=1}^{L}$, and let $\{\lambda_l\}_{l=1}^{L_p}$ be positive and $\{\lambda_l\}_{l=L_p+1}^{L}$ negative. The PEP is given by

$$\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}') = \begin{cases} 1 + \sum_{k=1}^{L_p} \xi_k\Big(N \ln \frac{\det(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})}{\det(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})}\Big), & \text{if } \det(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}) \geq \det(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}), \\ -\sum_{k=L_p+1}^{L} \xi_k\Big(N \ln \frac{\det(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})}{\det(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})}\Big), & \text{if } \det(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}) \leq \det(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}), \end{cases} \tag{6.9}$$

with $\xi_k(c) := \mathrm{Res}\left(\frac{e^{sc}}{s\prod_{l=1}^{L}\lambda_l^{\mu_l N}\left(s+1/\lambda_l\right)^{\mu_l N}}, \frac{-1}{\lambda_k}\right)$ *where*

$$\mathrm{Res}(f(s), x) := \frac{1}{(m-1)!} \lim_{s\to x} \frac{\mathrm{d}^{m-1}}{\mathrm{d}s^{m-1}}\left[(s-x)^m f(s)\right]$$

is the residue of a function $f(s)$ in a pole x of multiplicity m.

Proof. The closed-form expression of the PEP follows readily from [98, Proposition 1] by noting that the matrix $\mathbf{C}_{ij}^{\mathrm{NC}}$ therein is equal to $\boldsymbol{\Lambda} \otimes \boldsymbol{I}_N$ in our setting, thus has the same nonzero eigenvalues as $\boldsymbol{\Lambda}$ with multiplicities N. □

Another way to calculate the PEP is as follows. Let us rewrite the PEP as

$$\mathbb{P}(\boldsymbol{X} \to \boldsymbol{X}') = \mathbb{P}\left(\ln \frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\boldsymbol{X})}{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\boldsymbol{X}')} \leq 0\right) = \mathbb{P}(\mathsf{L}(\boldsymbol{X} \to \boldsymbol{X}') \leq 0) \tag{6.10}$$

with the PLLR $\mathsf{L}(\boldsymbol{X} \to \boldsymbol{X}') := \ln \frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\boldsymbol{X})}{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\boldsymbol{X}')}$. Using (6.3), we obtain

$$\mathsf{L}(\boldsymbol{X} \to \boldsymbol{X}') = N \ln \frac{\det(\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})}{\det(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})} - \mathrm{tr}\Big(((\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1} - (\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1})\mathbf{Y}\mathbf{Y}^{\mathsf{H}}\Big). \tag{6.11}$$

Let $\mathbf{Y}_0 := \mathbf{Y}(\boldsymbol{I}_T + \mathbf{X}\mathbf{X}^{\mathsf{H}})^{-1/2}$ be a "whitened" version of $\mathbf{Y}$, then $\mathbf{Y}_0$ is a Gaussian matrix with independent columns following $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{I}_T)$. The PLLR can be expressed as

$$\mathsf{L}(\boldsymbol{X} \to \boldsymbol{X}') = -N \ln \det\Big(\boldsymbol{I}_T + \bar{\boldsymbol{\Lambda}}\Big) + \mathrm{tr}\Big(\bar{\boldsymbol{\Lambda}}\mathbf{Y}_0\mathbf{Y}_0^{\mathsf{H}}\Big), \tag{6.12}$$

where the matrix $\bar{\boldsymbol{\Lambda}} := (\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{\frac{1}{2}}(\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{\frac{1}{2}} - \boldsymbol{I}_T$ has the same eigenvalues as $\boldsymbol{\Lambda}$. The matrix $\mathbf{Y}_0\mathbf{Y}_0^{\mathsf{H}}$ follows the complex Wishart distribution $\mathcal{W}_T(N, \boldsymbol{I})$ (see Appendix 6.A.1.a). If $\bar{\boldsymbol{\Lambda}}$, or equivalently $\boldsymbol{\Lambda}$, is semidefinite, a closed-form expression of the PEP is given as follows.

Proposition 6.3. *If $\boldsymbol{\Lambda} \succeq 0$, i.e. $\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}} \succeq \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$, the PEP is given by $\mathbb{P}(\boldsymbol{X} \to \boldsymbol{X}') = \psi$; if $\boldsymbol{\Lambda} \preceq 0$, i.e. $\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}} \preceq \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$, the PEP is given by $\mathbb{P}(\boldsymbol{X} \to \boldsymbol{X}') = 1 - \psi$, where*

$$\begin{aligned}\psi := {} & \varsigma^{TN} \det(|\boldsymbol{\Lambda}|)^{-N} \\ & \times \sum_{k=1}^{\infty} \frac{\gamma(TN + k, \frac{N}{\varsigma}|\ln\det(\boldsymbol{I}_T + \boldsymbol{\Lambda})|)}{k!(TN + k - 1)!} \sum_{\kappa(T,k)} [N]_{\kappa(T,k)} C_{\kappa(T,k)}(\boldsymbol{I}_T - \varsigma|\boldsymbol{\Lambda}|^{-1}),\end{aligned} \tag{6.13}$$

for arbitrary $\varsigma > 0$. Here, $\kappa(T,k) = \{k_i\}_{i=1}^T$ is a partition of k into T integers $k_1 \geq k_2 \geq \cdots \geq k_T$ such that $\sum_{i=1}^T k_i = k$; $[N]_{\kappa(T,k)} := \prod_{i=1}^T (N - i + 1)_{k_i}$ with $(a)_k := a(a+1)\ldots(a+k-1)$; and $\gamma(s,x) := \int_0^x t^{s-1}e^{-t}dt$ is the lower incomplete Gamma function. For a $T \times T$ matrix $\boldsymbol{\Sigma}$ with eigenvalues $\sigma_1, \ldots, \sigma_T$, $C_{\kappa(T,k)}(\boldsymbol{\Sigma})$ denotes the complex zonal polynomial *corresponding to a partition $\kappa(T,k) = \{k_1, k_2, \ldots, k_T\}$ defined as*

$$C_{\kappa(T,k)}(\boldsymbol{\Sigma}) := k! \frac{\prod_{1 \leq i < j \leq T}(k_i - k_j - i - j)}{\prod_{i=1}^T (k_i + T - i)!} \frac{\det\Big(\{\sigma_i^{k_j + T - j}\}_{i,j}\Big)}{\det\Big(\{\sigma_i^{T-j}\}_{i,j}\Big)}, \tag{6.14}$$

where $\{a(i,j)\}_{i,j}$ denotes a matrix with $a(i,j)$ as the (i,j)-th component.

Proof. From (6.10) and (6.12), the PEP is the CDF (resp. complementary CDF) of the trace of $|\bar{\boldsymbol{\Lambda}}|^{1/2}\mathbf{Y}_0\mathbf{Y}_0^{\mathsf{H}}|\bar{\boldsymbol{\Lambda}}|^{1/2}$ evaluated at $N|\ln\det(\boldsymbol{I}_T + \bar{\boldsymbol{\Lambda}})|$ if $\bar{\boldsymbol{\Lambda}} \succeq 0$ (resp. $\bar{\boldsymbol{\Lambda}} \preceq 0$). Since $\mathbf{Y}_0\mathbf{Y}_0^{\mathsf{H}} \sim \mathcal{W}_T(N, \boldsymbol{I})$, according to [192, Thm. 3.2.5], we have that $|\bar{\boldsymbol{\Lambda}}|^{1/2}\mathbf{Y}_0\mathbf{Y}_0^{\mathsf{H}}|\bar{\boldsymbol{\Lambda}}|^{1/2} \sim \mathcal{W}_T(N, |\bar{\boldsymbol{\Lambda}}|)$. The proof then follows from the CDF of the trace of a complex Wishart matrix given in Lemma 6.2 in Appendix 6.A.1.a, and the fact that $\boldsymbol{\Lambda}$ and $\bar{\boldsymbol{\Lambda}}$ have the same eigenvalues. □

Proposition 6.3 can be seen as a special case of Proposition 6.2 when $L_p \in \{0, L\}$. These closed-form expressions, however, do not bring clear insights to constellation design. A high-SNR asymptotic expression of the PEP was given in [98, Proposition 3], but is also hard to exploit since it involves the number of different individual symbols in $\boldsymbol{X}$ and $\boldsymbol{X}'$, which is highly non-smooth. In the following, we derive some simple but effective design criteria.

6.3.2 Design Criteria

We resort to the following bound on the PEP

$$\mathbb{P}(\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\le 0)\le \frac{\mathsf{Var}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]}{\mathsf{Var}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]+\mathbb{E}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]^2} \tag{6.15}$$

which follows from Cantelli's inequality.[2] Note that the upper bound decreases with $\frac{\mathbb{E}[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')]^2}{\mathsf{Var}[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')]}$. We choose to relax the problem into maximizing the expected PLLR $\mathbb{E}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]$. Although maximizing $\frac{\mathbb{E}[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')]^2}{\mathsf{Var}[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')]}$ and maximizing $\mathbb{E}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]$ are equivalent only when $\mathsf{Var}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]$ is constant over different symbol pairs, the relaxation makes the problem tractable.

We further justify our choice by pointing out the connection of our problem to the following hypothesis testing problem. Let us consider two hypotheses:

$$H_0:\ \{\boldsymbol{y}_i\}_{i=1}^N\sim\mathcal{N}_{\mathbb{C}}(\boldsymbol{0},\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}), \tag{6.16}$$

$$H_1:\ \{\boldsymbol{y}_i\}_{i=1}^N\sim\mathcal{N}_{\mathbb{C}}(\boldsymbol{0},\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}), \tag{6.17}$$

where $\{\boldsymbol{y}_i\}_{i=1}^N$ are realizations of N columns of $\mathbf{Y}$. Then, the PEP $\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}')$ can be seen as the type-1 error probability of the optimal likelihood ratio test. From (6.15) and the fact that $\frac{\mathbb{E}[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')]^2}{\mathsf{Var}[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')]}\to\infty$ as $N\to\infty$, we have that $\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}')\to 0$ as $N\to\infty$ for any constellation satisfying the identifiability condition in Proposition 6.1. (A proof is given in Appendix 6.A.2.) Switching the symbols' roles, we deduce that $\mathbb{P}(\boldsymbol{X}'\to\boldsymbol{X})\le\epsilon\in(0,1/2)$ for N large enough. Then, from the Chernoff-Stein Lemma [21, Theorem 11.8.3], we have that

$$\begin{aligned}\lim_{N\to\infty}\tfrac{1}{N}\ln\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}')&=-D\Big(\mathcal{N}_{\mathbb{C}}(\boldsymbol{0},\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})\Big\|\mathcal{N}_{\mathbb{C}}(\boldsymbol{0},\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})\Big)\\&=-\mathbb{E}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big].\end{aligned} \tag{6.18}$$

Therefore, maximizing $\mathbb{E}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]$ maximizes the pairwise error exponent when $N\to\infty$. The convergence of the pairwise error exponent to the KL divergence was also used in [59], [97].

Therefore, let $e_{\min}(\mathcal{X}):=\frac{1}{N}\min\limits_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}}\mathbb{E}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]$, we consider the following design criterion

$$\mathcal{X}^*=\arg\max_{\mathcal{X}} e_{\min}(\mathcal{X}) \tag{6.19}$$

where it follows from (6.11) and $\mathbb{E}[\mathbf{Y}\mathbf{Y}^{\mathsf{H}}]=N(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})$ that

$$\begin{aligned}\mathbb{E}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]=N\ln\frac{\det(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})}{\det(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})}-N+N\mathrm{tr}\big((\boldsymbol{I}+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\big)\\+N\mathrm{tr}\big((\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\big).\end{aligned} \tag{6.20}$$

Remark 6.3. *A criterion following from treating the joint constellation as a Grassmannian constellation in $G(\mathbb{C}^T,M_{\mathrm{tot}})$ is $\mathcal{X}^*=\arg\max\limits_{\mathcal{X}}\min\limits_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}}\mathrm{tr}\Big(\boldsymbol{I}-\frac{\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}}{\|\boldsymbol{X}\|_{\mathrm{F}}^2\|\boldsymbol{X}'\|_{\mathrm{F}}^2}\Big)$, or equivalently,*

$$\mathcal{X}^*=\arg\min_{\mathcal{X}}\max_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}}\mathrm{tr}\Big(\frac{\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}}{\|\boldsymbol{X}\|_{\mathrm{F}}^2\|\boldsymbol{X}'\|_{\mathrm{F}}^2}\Big). \tag{6.21}$$

[2] Cantelli's inequality [193, Section II.8] states that $\mathbb{P}(\mathsf{x}-\mu\le\lambda)\le\frac{\sigma^2}{\sigma^2+\lambda^2}$ for a real-valued random variable x with mean μ and variance σ^2, and $\lambda<0$. Applying this with $\mathsf{x}=\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')$ and $\lambda=-\mathbb{E}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]$, we obtain (6.15).

Another criterion originally proposed for the single-user channel in [68, Eq.(8)] is

$$\mathcal{X}^* = \arg\min_{\mathcal{X}} \sum_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \det{}^{-N}\Big(\boldsymbol{I} - \frac{\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}}{\|\boldsymbol{X}\|_{\mathrm{F}}^2\|\boldsymbol{X}'\|_{\mathrm{F}}^2}\Big). \tag{6.22}$$

We denote $m_1(\mathcal{X}) := \max\limits_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \mathrm{tr}\Big(\frac{\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}}{\|\boldsymbol{X}\|_{\mathrm{F}}^2\|\boldsymbol{X}'\|_{\mathrm{F}}^2}\Big)$ *and* $m_2(\mathcal{X}) := \ln \sum\limits_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \det{}^{-N}\Big(\boldsymbol{I}-\frac{\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}}{\|\boldsymbol{X}\|_{\mathrm{F}}^2\|\boldsymbol{X}'\|_{\mathrm{F}}^2}\Big)$ *for future reference.*

In the following, we further simplify the design criterion (6.19).

Lemma 6.1. *Let* $\boldsymbol{X}$ *and* $\boldsymbol{X}'$ *be such that* $\|\boldsymbol{X}\|_{\mathrm{F}}^2 = \Theta(P)$ *and* $\|\boldsymbol{X}'\|_{\mathrm{F}}^2 = \Theta(P)$ *as* $P\to\infty$. *We have* $\mathrm{tr}\big((\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\big) = O(1)$; $\ln\frac{\det\left(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\right)}{\det(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})} = O(1)$ *if* $\mathrm{Span}(\boldsymbol{X}) = \mathrm{Span}(\boldsymbol{X}')$ *and* $\Theta(\ln P)$ *otherwise. Furthermore,* $\mathrm{tr}\big((\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\big) = O(1)$ *if* $\mathrm{Span}(\boldsymbol{X}) = \mathrm{Span}(\boldsymbol{X}')$ *and* $\Theta(P)$ *otherwise.*

Proof. The proof is provided in Appendix 6.A.3. □

We see that the only term in (6.20) that can scale up linearly with P is $d(\boldsymbol{X}\to\boldsymbol{X}') := \mathrm{tr}\big((\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\big)$. Let $d_{\min}(\mathcal{X}) := \min\limits_{\boldsymbol{X},\boldsymbol{X}'\in\mathcal{X}:\boldsymbol{X}\neq\boldsymbol{X}'} d(\boldsymbol{X}\to\boldsymbol{X}')$, we have the following design criterion

$$\mathcal{X}^* = \arg\max_{\mathcal{X}} d_{\min}(\mathcal{X}). \tag{6.23}$$

Hereafter, we assume for simplicity that all users have the same number of antennas, i.e. $M_1 = \cdots = M_K = M$, although the general case follows in a straightforward manner.

6.3.3 The Single-User Case

In the single-user case with M transmit antenna, it is known that the high-SNR optimal input signal belongs to the Grassmann manifold [49]. We consider Grassmannian constellation [64] $\mathcal{X}\subset G(\mathbb{C}^T, M)$, thus $\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X} = \frac{PT}{M}\boldsymbol{I}_M, \forall\boldsymbol{X}\in\mathcal{X}$. Using the Woodbury identity $(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1} = \boldsymbol{I}_T - \boldsymbol{X}'(\boldsymbol{I}_M+\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}')^{-1}\boldsymbol{X}'^{\mathsf{H}}$, we have that

$$\begin{aligned} d(\boldsymbol{X}\to\boldsymbol{X}') &= \mathrm{tr}\Big((\boldsymbol{I}_T - \boldsymbol{X}'(\boldsymbol{I}_M+\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}')^{-1}\boldsymbol{X}'^{\mathsf{H}})\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\Big) && (6.24)\\ &= \mathrm{tr}(\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}) - \mathrm{tr}\Big(\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}'(\boldsymbol{I}_M+\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}')^{-1}\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}\Big) && (6.25)\\ &= PT\Big(1 - \alpha_{P,T,M}\frac{\|\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}\|_{\mathrm{F}}^2}{(PT)^2}\Big), && (6.26)\end{aligned}$$

where $\alpha_{P,T,M} := \Big(\frac{1}{PT}+\frac{1}{M}\Big)^{-1}$ and the last equality follows from $\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X} = \boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}' = \frac{PT}{M}\boldsymbol{I}_M$. Therefore, the design criterion (6.23) is equivalent to $\mathcal{X} = \arg\min\limits_{\mathcal{X}} \max\limits_{\boldsymbol{X},\boldsymbol{X}'\in\mathcal{X}:\boldsymbol{X}\neq\boldsymbol{X}'} \|\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}\|_{\mathrm{F}}^2$. This coincides with the common criterion of maximizing the minimum pairwise chordal distance between the symbols [60], [64], [74], [123].

6.3.4 The Two-User Case

In the two-user case, we first develop

$$d(\boldsymbol{X}\to\boldsymbol{X}') = \mathrm{tr}\Big(\boldsymbol{X}_1^{\mathsf{H}}(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}_1\Big) + \mathrm{tr}\Big(\boldsymbol{X}_2^{\mathsf{H}}(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}_2\Big), \tag{6.27}$$

where we recall that $\boldsymbol{X} := [\boldsymbol{X}_1\ \boldsymbol{X}_2]$, $\boldsymbol{X}' := [\boldsymbol{X}_1'\ \boldsymbol{X}_2']$ with $\boldsymbol{X}_k, \boldsymbol{X}_k'\in\mathcal{X}_k$, $k\in\{1,2\}$, and $\boldsymbol{X}'\neq\boldsymbol{X}$. There are two types of error event.

1. Simultaneous detection error, i.e., $\boldsymbol{X}_1 \neq \boldsymbol{X}'_1$, $\boldsymbol{X}_2 \neq \boldsymbol{X}'_2$:

$$d(\boldsymbol{X} \to \boldsymbol{X}') = \mathrm{tr}\Big(\boldsymbol{X}_1^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}'_1\boldsymbol{X}'^{\mathsf{H}}_1 + \boldsymbol{X}'_2\boldsymbol{X}'^{\mathsf{H}}_2)^{-1}\boldsymbol{X}_1\Big) + \mathrm{tr}\Big(\boldsymbol{X}_2^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}'^{\mathsf{H}}_1\boldsymbol{X}'_1 + \boldsymbol{X}'^{\mathsf{H}}_2\boldsymbol{X}'_2)^{-1}\boldsymbol{X}_2\Big). \tag{6.28}$$

2. One sided detection error, i.e., $\boldsymbol{X}_k = \boldsymbol{X}'_k$, $\boldsymbol{X}_l \neq \boldsymbol{X}'_l$, $k \neq l \in \{1, 2\}$:

$$d(\boldsymbol{X} \to \boldsymbol{X}') = \mathrm{tr}\Big(\boldsymbol{X}_k^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}_k\boldsymbol{X}_k^{\mathsf{H}} + \boldsymbol{X}'_l\boldsymbol{X}'^{\mathsf{H}}_l)^{-1}\boldsymbol{X}_k\Big) + \mathrm{tr}\Big(\boldsymbol{X}_l^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}'^{\mathsf{H}}_k\boldsymbol{X}'_k + \boldsymbol{X}'^{\mathsf{H}}_l\boldsymbol{X}'_l)^{-1}\boldsymbol{X}_l\Big). \tag{6.29}$$

Let us define

$$d_1(\mathcal{X}) := \min_{\substack{\boldsymbol{X}_1 \neq \boldsymbol{X}'_1 \in \mathcal{X}_1 \\ \boldsymbol{X}_2 \in \mathcal{X}_2}} \mathrm{tr}(\boldsymbol{X}_1^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}'_1\boldsymbol{X}'^{\mathsf{H}}_1 + \boldsymbol{X}_2\boldsymbol{X}_2^{\mathsf{H}})^{-1}\boldsymbol{X}_1), \tag{6.30}$$

$$d_2(\mathcal{X}) := \min_{\substack{\boldsymbol{X}_2 \neq \boldsymbol{X}'_2 \in \mathcal{X}_2 \\ \boldsymbol{X}_1 \in \mathcal{X}_1}} \mathrm{tr}(\boldsymbol{X}_2^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}_1\boldsymbol{X}_1^{\mathsf{H}} + \boldsymbol{X}'_2\boldsymbol{X}'^{\mathsf{H}}_2)^{-1}\boldsymbol{X}_2). \tag{6.31}$$

Since $0 \leq \mathrm{tr}\Big(\boldsymbol{X}_k^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}_k\boldsymbol{X}_k^{\mathsf{H}} + \boldsymbol{X}'_l\boldsymbol{X}'^{\mathsf{H}}_l)^{-1}\boldsymbol{X}_k\Big) \leq M$, $\forall\, k \neq l$, considering both types of error, we obtain the following bounds

$$\min\{d_1(\mathcal{X}), d_2(\mathcal{X})\} \leq d_{\min}(\mathcal{X}) \leq \min\{d_1(\mathcal{X}), d_2(\mathcal{X})\} + M. \tag{6.32}$$

Therefore, $d_{\min}(\mathcal{X})$ is within a constant gap to $\min\{d_1(\mathcal{X}), d_2(\mathcal{X})\}$, and $d_{\min}(\mathcal{X})$ scales linearly with P when P is large if and only if $\min\{d_1(\mathcal{X}), d_2(\mathcal{X})\}$ does so. Based on this observation, we propose the following design criterion

$$\mathcal{X}^* = \arg\max_{\mathcal{X}} \ \min\{d_1(\mathcal{X}), d_2(\mathcal{X})\}. \tag{6.33}$$

6.3.5 The K-User Case

We now generalize the analysis of the two-user case to the K-user case. We develop

$$d(\boldsymbol{X} \to \boldsymbol{X}') = \sum_{k=1}^{K} \mathrm{tr}\Big(\boldsymbol{X}_k^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}_k\Big), \tag{6.34}$$

where we recall that $\boldsymbol{X} = [\boldsymbol{X}_1 \ldots \boldsymbol{X}_K], \boldsymbol{X}' = [\boldsymbol{X}'_1 \ldots \boldsymbol{X}'_K]$ with $\boldsymbol{X}_k, \boldsymbol{X}'_k \in \mathcal{X}_k, k \in [K]$ and $\boldsymbol{X} \neq \boldsymbol{X}'$. $\boldsymbol{X}$ and $\boldsymbol{X}'$ are regarded as the transmitted and detected joint symbols, respectively. For any $\mathcal{K} \subset [K]$, if all users in $\mathcal{K}$ are wrongly detected, while all users in $\mathcal{L} = [K] \setminus \mathcal{K}$ are correctly detected, then

$$d(\boldsymbol{X} \to \boldsymbol{X}') = \sum_{k \in \mathcal{K}} \mathrm{tr}\Big(\boldsymbol{X}_k^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}'_k\boldsymbol{X}'^{\,\mathsf{H}}_k + \sum_{j \neq k} \boldsymbol{X}'_j\boldsymbol{X}'^{\,\mathsf{H}}_j)^{-1}\boldsymbol{X}_k\Big) + \sum_{l \in \mathcal{L}} \mathrm{tr}\Big(\boldsymbol{X}_l^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}_l\boldsymbol{X}_l^{\mathsf{H}} + \sum_{j \neq l} \boldsymbol{X}'_j\boldsymbol{X}'^{\,\mathsf{H}}_j)^{-1}\boldsymbol{X}_l\Big). \tag{6.35}$$

In this case, the minimal value of $d(\boldsymbol{X} \to \boldsymbol{X}')$ is defined as

$$d^{\mathcal{K}}_{\min}(\mathcal{X}) := \min_{\substack{\boldsymbol{X}_k \neq \boldsymbol{X}'_k \in \mathcal{X}_k, \forall k \in \mathcal{K}, \\ \boldsymbol{X}_l = \boldsymbol{X}'_l \in \mathcal{X}_l, \forall l \in [K] \setminus \mathcal{K}}} d(\boldsymbol{X} \to \boldsymbol{X}')$$

. Then, it is straightforward that $d_{\min}(\mathcal{X})$ is the minimum value of $d_{\min}^{\mathcal{K}}(\mathcal{X})$ over all possible $\mathcal{K} \subset [K]$, i.e., $d_{\min}(\mathcal{X}) = \min_{\mathcal{K}\subset[K]} d_{\min}^{\mathcal{K}}(\mathcal{X})$. Let us define

$$d_k(\mathcal{X}) := \min_{\substack{\boldsymbol{X}_k \neq \boldsymbol{X}'_k \in \mathcal{X}_k \\ \boldsymbol{X}_j \in \mathcal{X}_j, j\neq k}} \operatorname{tr}\Big(\boldsymbol{X}_k^{\mathsf{H}}\Big(\boldsymbol{I}_T + \boldsymbol{X}'_k \boldsymbol{X}'^{\mathsf{H}}_k + \sum_{j\neq k} \boldsymbol{X}_j \boldsymbol{X}_j^{\mathsf{H}}\Big)^{-1} \boldsymbol{X}_k\Big). \tag{6.36}$$

We have that

$$d_{\min}^{\mathcal{K}}(\mathcal{X}) \geq \min_{k\in\mathcal{K}} d_k(\mathcal{X}) \geq \min_{k\in[K]} d_k(\mathcal{X}), \forall \mathcal{K} \subset [K], \tag{6.37}$$

where the first inequality holds since the constraint under the min in $d_k(\mathcal{X})$ subsumes that in $d_{\min}^{\mathcal{K}}(\mathcal{X})$, and the trace in $d_k(\mathcal{X})$ is one of the nonnegative summands in $d_{\min}^{\mathcal{K}}(\mathcal{X})$; the second inequality holds since $\mathcal{K} \subset [K]$. Taking $\mathcal{K}^* = \arg\min_{\mathcal{K}\subset[K]} d_{\min}^{\mathcal{K}}(\mathcal{X})$ yields

$$d_{\min}(\mathcal{X}) = d_{\min}^{\mathcal{K}^*}(\mathcal{X}) \geq \min_{k\in[K]} d_k(\mathcal{X}). \tag{6.38}$$

On the other hand, since $\operatorname{tr}\Big(\boldsymbol{X}_l^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}_l\boldsymbol{X}_l^{\mathsf{H}} + \sum_{j\neq l} \boldsymbol{X}'_j \boldsymbol{X}'^{\mathsf{H}}_j)^{-1}\boldsymbol{X}_l\Big) \leq M$, $\forall\, l \in [K]$, we get that $d_{\min}^{\{k\}}(\mathcal{X}) \leq d_k(\mathcal{X}) + (K-1)M$ for all $k \in [K]$. Let $k^* = \arg\min_{k\in[K]} d_k(\mathcal{X})$, we have that

$$\begin{aligned} d_{\min}(\mathcal{X}) &= \min_{\mathcal{K}\subset[K]} d_{\min}^{\mathcal{K}}(\mathcal{X}) && (6.39)\\ &\leq d_{\min}^{\{k^*\}}(\mathcal{X}) && (6.40)\\ &\leq d_{k^*}(\mathcal{X}) + (K-1)M && (6.41)\\ &\leq \min_{k\in[K]} d_k(\mathcal{X}) + (K-1)M. && (6.42) \end{aligned}$$

From (6.38) and (6.42), we see that $d_{\min}(\mathcal{X})$ is within a constant gap to $\min_{k\in[K]} d_k(\mathcal{X})$, and thus $d_{\min}(\mathcal{X})$ scales linearly with P when P is large if and only if $\min_{k\in[K]} d_k(\mathcal{X})$ does so. Based on this observation, we propose the following design criterion

$$\mathcal{X}^* = \arg\max_{\mathcal{X}} \min_{k\in[K]} d_k(\mathcal{X}). \tag{6.43}$$

6.3.6 Simplifications

6.3.6.a Alternating Optimization

To simplify the constellation optimization, we propose an *alternating optimization* as follows. First $\{\mathcal{X}_k\}_{k=1}^K$ are initialized. Then, for $k = 1, \ldots, K$, we iteratively optimize $\mathcal{X}_k$ by $\mathcal{X}_k^* = \arg\max_{\mathcal{X}_k} m(\mathcal{X})$ for fixed $\{\mathcal{X}_l\}_{l\neq k}$ in a round robin manner, where $m(\mathcal{X})$ is the considered metric. At each iteration, it has fewer variables to optimize than directly solving (6.19), (6.23), or (6.43). However, the solution of alternating optimization is not guaranteed to converge to the exact optimum.

6.3.6.b Solution Space Reduction

In the most general setting, the simplified criteria (6.19), (6.23), (6.43) still have a large solution space. Specifically, $\mathcal{X}$ belongs to a product space of

$$\{\boldsymbol{X}_k^{(1)}, \ldots, \boldsymbol{X}_k^{(|\mathcal{X}_k|)} \in \mathbb{C}^{T\times M_k} \;:\; \frac{1}{|\mathcal{X}_k|}\sum_{i=1}^{|\mathcal{X}_k|} \|\boldsymbol{X}_k^{(i)}\|^2 \leq PT\}$$

, and thus has $\prod_{k=1}^{K}(TM_k)^{|\mathcal{X}_k|}$ free variables to optimize. To reduce the solution space, we make the suboptimal assumption that the individual constellations $\mathcal{X}_k$ are Grassmannian, thus have equal power for each transmitted symbol $\boldsymbol{X}_k$. From a practical perspective, this is desirable since the constellation is oblivious to the presence of the other users and Grassmannian signaling is high-SNR optimal for the single-user channel. Under this assumption, we have that $\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{X}_k = \frac{P_k T}{T}\boldsymbol{I}_M,\ \forall \boldsymbol{X}_k \in \mathcal{X}_k, k \in [K]$. That is, the symbol $\boldsymbol{X}_k$ is a scaled truncated unitary matrix representative of a point in the Grassmann manifold $G(\mathbb{C}^T, M)$. Thus, the solution space is reduced to the Cartesian product of K instances of $G(\mathbb{C}^T, M)$ (for the signal subspace) and K instances of the interval $[0, P]$ (for the signal power). Furthermore, the optimization for signal subspace and power can be done separately, as in the next sections.

6.4 Two Simple Constructions for Given Transmit Power

In this section, inspired by the proposed criteria, we propose two simple constellation constructions for fixed power $P_k, k \in [K]$.

6.4.1 Partitioning Design

In this subsection, we consider the symmetrical power case $P_k = P,\ \forall k \in [K]$. This is a reasonable assumption if the users transmit at symmetric rate $R_1 = \cdots = R_K$. Also, with Grassmannian signaling, we have that $\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{X}_k = \frac{PT}{M}\boldsymbol{I}_M,\ \forall \boldsymbol{X}_k \in \mathcal{X}_k,\ k \in [K]$. Nevertheless, there must be constraints between the symbols of different users. For instance, if the constellations are such that $\boldsymbol{X}_1 = \boldsymbol{X}_2 = \cdots = \boldsymbol{X}_K$ for $\boldsymbol{X}_k \in \mathcal{X}_k$ can occur, then $d_k(\mathcal{X})$ is upper-bounded by a constant for any k. This can be developed in a formal way as follows.

An upper bound can be obtained by removing the terms inside the inverse in $d_k(\mathcal{X})$, namely,

$$d_k(\mathcal{X}) \leq \min\Big\{ \min_{\boldsymbol{X}_k \neq \boldsymbol{X}'_k \in \mathcal{X}_k} \operatorname{tr}(\boldsymbol{X}_k^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}'_k\boldsymbol{X}'^{\mathsf{H}}_k)^{-1}\boldsymbol{X}_k), \\ \min_{\boldsymbol{X}_k \in \mathcal{X}_k, \boldsymbol{X}_l \in \mathcal{X}_l, l \neq k} \operatorname{tr}(\boldsymbol{X}_k^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}_l\boldsymbol{X}_l^{\mathsf{H}})^{-1}\boldsymbol{X}_k)\Big\}. \tag{6.44}$$

For $d_k(\mathcal{X})$ to be large, the upper bounds must be large. This is made precise in the next proposition.

Proposition 6.4 (Necessary condition)**.** *Let $\{\mathcal{X}_k\}_{k=1}^K$ be such that $\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{X}_k = \frac{PT}{M}\boldsymbol{I}_M,\ \forall \boldsymbol{X}_k \in \mathcal{X}_k,\ k \in [K]$. If the following lower bound holds for some $c \in [0, 1/M]$*

$$\min_{k \in [K]} d_k(\mathcal{X}) \geq PT(1 - \alpha_{P,T,M}\, c), \tag{6.45}$$

where $\alpha_{P,T,M} := \left(\frac{1}{PT} + \frac{1}{M}\right)^{-1}$, then we must have

$$\frac{1}{(PT)^2} \max\Big\{ \max_{\boldsymbol{X}_k \neq \boldsymbol{X}'_k \in \mathcal{X}_k, k \in [K]} \|\boldsymbol{X}'^{\mathsf{H}}_k \boldsymbol{X}_k\|_{\mathrm{F}}^2, \max_{\boldsymbol{X}_k \in \mathcal{X}_k, \boldsymbol{X}_l \in \mathcal{X}_l, k \neq l \in [K]} \|\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{X}_l\|_{\mathrm{F}}^2 \Big\} \leq c. \tag{6.46}$$

Proof. The proof follows the same steps as in the single-user case in Section 6.3.3, applying to the upper bound (6.44). □

The above proposition shows that symbol pairs from different users should fulfill similar distance criteria as symbol pairs from the same constellation when it comes to identifiability conditions. However, it is unclear whether (6.46) alone is enough to guarantee a large

value of $\min_{k\in[K]} d_k(\mathcal{X})$. In the following, we shall show that these conditions are indeed sufficient if c is small.

Proposition 6.5 (Sufficient condition)**.** *Let $\{\mathcal{X}_k\}_{k=1}^K$ be such that $\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{X}_k = \frac{PT}{M}\boldsymbol{I}_M$, $\forall\, \boldsymbol{X}_k \in \mathcal{X}_k$, $k \in [K]$. If*

$$\frac{1}{(PT)^2}\max\left\{\max_{\boldsymbol{X}_k\neq\boldsymbol{X}_k'\in\mathcal{X}_k,k\in[K]} \left\|\boldsymbol{X}_k'^{\mathsf{H}}\boldsymbol{X}_k\right\|_{\mathrm{F}}^2, \max_{\boldsymbol{X}_k\in\mathcal{X}_k,\boldsymbol{X}_l\in\mathcal{X}_l,k\neq l\in[K]} \|\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{X}_l\|_{\mathrm{F}}^2\right\} \leq c \tag{6.47}$$

for some $c \in [0, 1/M]$, then we have

$$\min_{k\in[K]} d_k(\mathcal{X}) \geq PT\left(1 - K\left(\frac{1}{PT} + \frac{1}{M} - \sqrt{\frac{K(K-1)c}{2^{\mathbb{1}\{K=2\}}}}\right)^{-1} c\right). \tag{6.48}$$

Proof. The proof is provided in Appendix 6.A.4. □

The two propositions above motivate the following simplified design criterion

$$\mathcal{X}^* = \arg\min_{\mathcal{X}} \max\left\{\max_{\boldsymbol{X}_k\neq\boldsymbol{X}_k'\in\mathcal{X}_k,k\in[K]} \left\|\boldsymbol{X}_k'^{\mathsf{H}}\boldsymbol{X}_k\right\|_{\mathrm{F}}^2, \max_{\boldsymbol{X}_k\in\mathcal{X}_k,\boldsymbol{X}_l\in\mathcal{X}_l,k\neq l\in[K]} \|\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{X}_l\|_{\mathrm{F}}^2\right\}. \tag{6.49}$$

Based on the criterion (6.49), we propose a simple construction as follows. Let $\mathcal{X}_{\mathrm{SU}}$ be a single-user constellation and let $c := \frac{1}{(PT)^2}\max_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}_{\mathrm{SU}}} \|\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}\|_{\mathrm{F}}^2 \in [0, \frac{1}{M}]$. Then, we can generate $\{\mathcal{X}_k\}_{k=1}^K$ by partitioning $\mathcal{X}_{\mathrm{SU}}$ into K disjoint subsets. In this way, from (6.38) and Proposition 6.5, we can guarantee that

$$d_{\min}(\mathcal{X}) \geq PT\left(1 - K\left(\frac{1}{PT} + \frac{1}{M} - \sqrt{\frac{K(K-1)c}{2^{\mathbb{1}\{K=2\}}}}\right)^{-1} c\right). \tag{6.50}$$

With such a construction, the joint constellation design problem becomes essentially an individual constellation design problem. A random partition would be sufficient to guarantee (6.50), although one can smartly partition the set $\mathcal{X}_{\mathrm{SU}}$ to improve over (6.50). The optimal partition problem is equivalent to a min-max graph partitioning problem [194]. Also note that for the right-hand side of (6.50) to scale linearly with P, c must be small enough, which requires the initial single-user constellation $\mathcal{X}_{\mathrm{SU}}$ to be sparse enough in $G(\mathbb{C}^T, M)$ and thus limits the size of $\mathcal{X}_{\mathrm{SU}}$. This is made precise in the following.

Proposition 6.6. *For a joint constellation $\{\mathcal{X}_k\}_{k=1}^K$ generated by partitioning a single-user constellation $\mathcal{X}_{\mathrm{SU}}$, for the lower bound of $d_{\min}(\mathcal{X})$ in (6.50) to scale linearly with P, the minimum pairwise chordal distance between the elements of $\mathcal{X}_{\mathrm{SU}}$, i.e. $\delta_{\min}(\mathcal{X}_{\mathrm{SU}}) := \min_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}_{\mathrm{SU}}} \sqrt{M - \frac{1}{P^2T^2}\|\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}\|_{\mathrm{F}}^2}$, must satisfy*

$$\delta_{\min}(\mathcal{X}_{\mathrm{SU}}) > \sqrt{M - \left[\left(\frac{1}{KPT} + \frac{1}{KM} + \frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}\right)^{1/2} - \sqrt{\frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}}\right]^2}, \tag{6.51}$$

which is possible only if

$$|\mathcal{X}_{\mathrm{SU}}| < c_{T,M}^{-1} 2^{2M(T-M)} \times \left(M - \left[\left(\frac{1}{KPT} + \frac{1}{KM} + \frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}\right)^{\frac{1}{2}} - \sqrt{\frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}}\right]^2\right)^{-M(T-M)}, \tag{6.52}$$

with

$$c_{T,M} = \frac{1}{(M(T-M))!} \prod_{i=1}^{\min\{M,T-M\}} \frac{(T-i)!}{(\min\{M,T-M\}-i)!}. \tag{6.53}$$

Proof. The right-hand side of (6.50) scales linearly with P if

$$1 - K\Big(\frac{1}{PT} + \frac{1}{M} - \sqrt{\frac{K(K-1)c}{2^{\mathbb{1}\{K=2\}}}}\Big)^{-1} c > 0,$$

that is

$$c < \left[\left(\frac{1}{KPT} + \frac{1}{KM} + \frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}\right)^{1/2} - \sqrt{\frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}}\right]^2. \tag{6.54}$$

By definition, we have that $\delta_{\min}(\mathcal{X}_{\mathrm{SU}}) = \sqrt{M-c}$, so (6.54) is equivalent to (6.51). On the other hand, according to Lemma 1.4, the volume of a metric ball $\mathcal{B}(\delta)$ of radius δ (in chordal distance) in $G(\mathbb{C}^T, M)$ with normalized invariant measure $\mu(\cdot)$ is given by $\mu(\mathcal{B}(\delta)) = c_{T,M}\delta^{2M(T-M)}$ with $c_{T,M}$ defined in (6.53). Since $\mathcal{X}_{\mathrm{SU}}$ is a packing on $G(\mathbb{C}^T, M)$ with minimum chordal distance $\delta_{\min}(\mathcal{X}_{\mathrm{SU}})$, we have the Hamming upper bound (1.121) on the size of $\mathcal{X}_{\mathrm{SU}}$ as $|\mathcal{X}_{\mathrm{SU}}| \le \frac{1}{\mu(\mathcal{B}(\delta_{\min}(\mathcal{X}_{\mathrm{SU}})/2))}$. From this and (6.51), we obtain (6.52). □

In the high-SNR regime ($P \to \infty$), the bounds of $\delta_{\min}(\mathcal{X}_{\mathrm{SU}})$ and $|\mathcal{X}_{\mathrm{SU}}|$ in Proposition 6.6 respectively converge to

$$\nu(K,M) := \sqrt{M - \left[\left(\frac{1}{KM} + \frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}\right)^{1/2} - \sqrt{\frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}}\right]^2}, \tag{6.55}$$

$$\begin{aligned}\beta(T,K,M) := {} & c_{T,M}^{-1} 2^{2M(T-M)} \\ & \times \left(M - \left[\left(\frac{1}{KM} + \frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}\right)^{\frac{1}{2}} - \sqrt{\frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}}\right]^2\right)^{-M(T-M)}.\end{aligned} \tag{6.56}$$

Fig. 6.1 shows the values of $\log_2(\beta(T,K,M))$, which is the high-SNR upper bound on the number of bits per symbol $\log_2(|\mathcal{X}_{\mathrm{SU}}|)$ in $\mathcal{X}_{\mathrm{SU}}$, for $K \in \{2,4\}$ and some values of T and M. As can be seen, for a fixed M, the bound monotonically increases in T; for a fixed T, the bound first increases in M then decreases after a peak value and becomes 0 (imposing a zero transmission rate) when $M \approx 0.73T$; and the bound is not sensitive to the value of K.

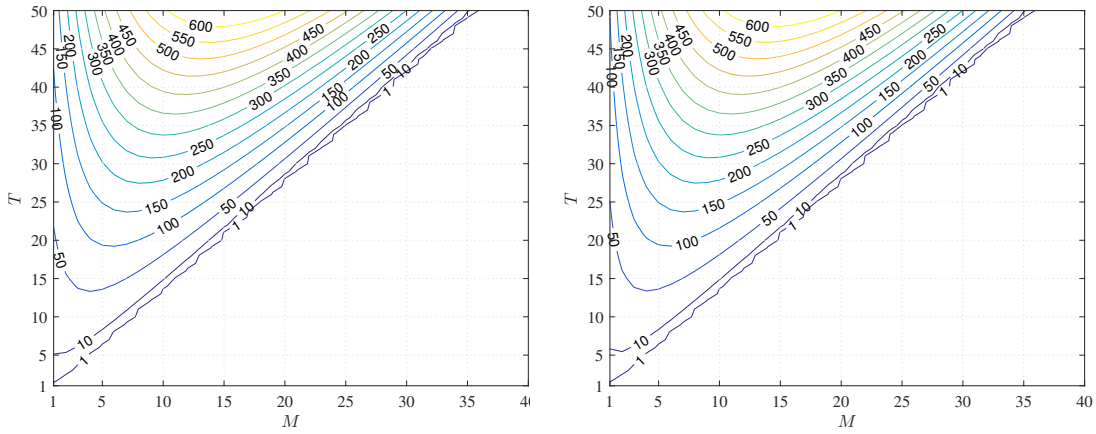
Remark 6.4. *The Grassmann manifold $G(\mathbb{C}^T, M)$ has $2M(T-M)$ real dimensions. From* (6.56)*, an upper bound on the number of bits per real dimension for $\mathcal{X}_{\mathrm{SU}}$ is given by*

$$\frac{\log_2 \beta(T,K,M)}{2M(T-M)} \le \zeta(K,M) \tag{6.57}$$

where

$$\zeta(K,M) := 1 - \frac{1}{2}\log_2\left(1 - \frac{1}{M}\left[\left(\frac{1}{KM} + \frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}\right)^{1/2} - \sqrt{\frac{K-1}{4K2^{\mathbb{1}\{K=2\}}}}\right]^2\right). \tag{6.58}$$

In fact, using Stirling's formula $\sqrt{2\pi}n^{n+1/2}e^{-n} \le n! \le en^{n+1/2}e^{-n}$ [195], we can show that $\frac{\log_2 \beta(T,M)}{2M(T-M)} \uparrow \zeta(K,M)$ as $T \to \infty$, where "$\uparrow$" means "approach from below". After some simple manipulations, we have that $\zeta(K,M) \le 2 - \frac{1}{2}\log_2 3 < \log_2 3$ for any K and M. That is, roughly speaking, one should not pack more than 2 symbols of $\mathcal{X}_{\mathrm{SU}}$ in each real dimension of the manifold on average.



(a) $K = 2$ (b) $K = 4$

FIGURE 6.1: The upper bound $\log_2(\beta(T, K, M))$ on the number of bits per symbol $\log_2(|\mathcal{X}_{\mathrm{SU}}|)$ for the lower bound (6.50) of $d_{\min}(\mathcal{X})$ to scale linearly with P.

6.4.2 Precoding Design

We consider the SIMO case ($M = 1$), where the constellation symbols are vector valued. We propose a constellation design as follows.

6.4.2.a Design Principle

We first define for each user a constellation $\mathcal{C}_k = \{\boldsymbol{c}^{(1)}, \ldots, \boldsymbol{c}^{(|\mathcal{X}_k|)}\}$ in $G(\mathbb{C}^{T-K+1}, 1)$, then construct $\mathcal{X}_k$ as the image of $\mathcal{C}_k$ through a mapping

$$\boldsymbol{f}_k : \begin{cases} \mathbb{C}^{T-K+1} & \to \mathbb{C}^T, \\ \boldsymbol{d}_k^{(i)} & \mapsto \boldsymbol{c}_k^{(i)}. \end{cases} \tag{6.59}$$

That is,

$$\boldsymbol{x}_k^{(i)} = \boldsymbol{f}_k\left(\boldsymbol{c}_k^{(i)}\right), \quad i = 1, \ldots, 2^{|\mathcal{X}_k|}. \tag{6.60}$$

While the constellations $\mathcal{C}_k$ can be identical or different amongst the users, the encoder mapping $\boldsymbol{f}_k$ needs to be unique for each user in order to guarantee the identifiability condition.

In particular, we consider a mapping consisting of a linear transformation followed by a normalization, such that $\boldsymbol{f}_k$ can be defined through a full-rank precoder $\boldsymbol{U}_k \in \mathbb{C}^{T\times(T-K+1)}$ as

$$\boldsymbol{x}_k^{(i)} = \frac{P_k T}{M} \frac{\boldsymbol{U}_k \boldsymbol{c}_k^{(i)}}{\|\boldsymbol{U}_k \boldsymbol{c}_k^{(i)}\|}. \tag{6.61}$$

We refer to this kind of mapping as *normalized linear encoder mapping*. Therefore, each symbol $\boldsymbol{x}_k^{(i)}$ of user k belongs to the column space $\mathcal{U}_k$ of $\boldsymbol{U}_k$. For example, when $T = 3$, $K = 2$, and $|\mathcal{X}_1| = |\mathcal{X}_2| = 4$, a geometric interpretation for the precoders $\boldsymbol{U}_1 = [\boldsymbol{e}_1\ \boldsymbol{e}_3]$ and $\boldsymbol{U}_2 = [\boldsymbol{e}_2\ \boldsymbol{e}_3]$ is provided in Fig. 6.2.

Remark 6.5. *The precoders impose a geometric separation between the individual constellations. This is the discrimination factor of the signals transmitted from different users.*

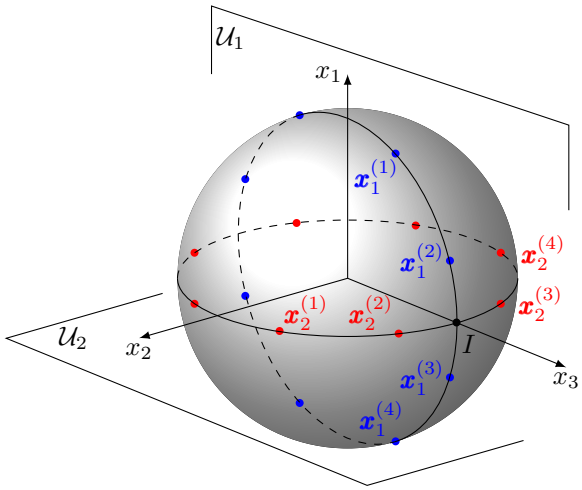


FIGURE 6.2: A geometric interpretation on the real Grassmannian for the precoders $\boldsymbol{U}_1 = [\boldsymbol{e}_1\ \boldsymbol{e}_3]$ and $\boldsymbol{U}_2 = [\boldsymbol{e}_2\ \boldsymbol{e}_3]$ ($\boldsymbol{e}_k$ denotes the k-th column of $\boldsymbol{I}_T$) for $T = 3$, $K = 2$, $|\mathcal{X}_1| = |\mathcal{X}_2| = 4$. The symbols of user 1 and user 2—represented by their projections on the sphere—belong to the column spaces $\mathcal{U}_1$ and $\mathcal{U}_2$ of $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$, respectively. The axis x_t, $t = 1, 2, 3$, corresponds to the t-th component of a symbol.

In Chapter 7, we will exploit this geometric separation to design an efficient detector that decouples the joint detection into multiple single-user detection problems. We recall that we focus on ML detection in the current chapter.

Let $\boldsymbol{V}_k \in \mathbb{C}^{T\times(K-1)}$ denote a basis of the orthogonal subspace of $\mathcal{U}_k$, i.e., $\boldsymbol{V}_k^{\mathsf{H}}\boldsymbol{U}_k = \boldsymbol{0}$, $\boldsymbol{V}_k^{\mathsf{H}}\boldsymbol{V}_k = \boldsymbol{I}_{K-1}$.

Property 6.1 (Distance preservation)**.** *If $\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{U}_k = \boldsymbol{I}_{T-K+1}$, the chordal distance of any pair of symbols in $\mathcal{C}_k$ is preserved by the normalized linear encoder mappings.*

Proof. For any two symbols $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ in $\mathcal{C}_k$, the chordal distance between their precoded versions is

$$d(\boldsymbol{U}_k\boldsymbol{c}_1, \boldsymbol{U}_k\boldsymbol{c}_2) = \sqrt{1 - |\boldsymbol{c}_2^{\mathsf{H}}\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{U}_k\boldsymbol{c}_1|^2} = \sqrt{1 - |\boldsymbol{c}_2^{\mathsf{H}}\boldsymbol{c}_1|^2} = d(\boldsymbol{c}_1, \boldsymbol{c}_2), \tag{6.62}$$

since $\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{U}_k = \boldsymbol{I}$ by assumption. □

This property states that the distance spectrum of the individual constellations $\mathcal{X}_k$ is the same as that of $\mathcal{C}_k$ if $\boldsymbol{U}_k^{\mathsf{H}}\boldsymbol{U}_k = \boldsymbol{I}$.

6.4.2.b Precoder Design

For fixed constellations $\mathcal{C}_k$, the precoders can be optimized according to the criteria in (6.19) and (6.23). To reduce the complexity, we consider a more greedy approach by taking the QR factorization of the precoders:

$$\boldsymbol{U}_k = \boldsymbol{Q}_k\boldsymbol{R}_k, \quad k \in [K], \tag{6.63}$$

where the truncated-unitary matrix $\boldsymbol{Q}_k \in \mathbb{C}^{T\times(T-K+1)}$ determines the *subspace* $\mathcal{U}_k$ which the symbols $\boldsymbol{c}_k^{(i)}$ lie in, and the upper triangular matrix $\boldsymbol{R}_k \in \mathbb{C}^{(T-K+1)\times(T-K+1)}$ partially decides the *orientation* of the symbols in this subspace. Both $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ need to be optimized so as to optimize the metrics $\min_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \mathbb{E}[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')]$ and $d_{\min}(\mathcal{X})$.

From Lemma 6.1, we know that $\mathbb{E}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]$ and $d(\boldsymbol{X}\to\boldsymbol{X}')$ are upper bounded by a constant if $\mathrm{Span}(\boldsymbol{X}) = \mathrm{Span}\,(\boldsymbol{X}')$. In fact, these metrics become larger as $\mathrm{Span}(\boldsymbol{X})$ and $\mathrm{Span}\,(\boldsymbol{X}')$ share less dimensions (see Appendix 6.A.3). Therefore, the precoders should be designed such that $\mathrm{Span}(\boldsymbol{X})$ and $\mathrm{Span}\,(\boldsymbol{X}')$ are maximally orthogonal for any pair $\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}$. Since $\mathrm{Span}(\boldsymbol{X})$ is also the column space of $[\boldsymbol{U}_1\ \ldots\ \boldsymbol{U}_K]$ for any $\boldsymbol{X}$, $\mathrm{Span}(\boldsymbol{X})$ and $\mathrm{Span}(\boldsymbol{X}')$ are both spanned by the shared dimensions of $\mathcal{U}_1,\ldots,\mathcal{U}_K$. Therefore, to separate $\mathrm{Span}(\boldsymbol{X})$ and $\mathrm{Span}(\boldsymbol{X}')$, we minimize the number of shared dimensions in $\mathcal{U}_1,\ldots,\mathcal{U}_K$. In other words, we would like to optimize $\{\boldsymbol{Q}_k\}$ so that the subspaces $\mathcal{U}_k$ are maximally distant from each other. Considering the max-min chordal distance optimization problem, we have

$$\{\boldsymbol{Q}_1,\ldots,\boldsymbol{Q}_K\} = \arg\max_{\boldsymbol{Q}_k:\ \boldsymbol{Q}_k^{\mathsf{H}}\boldsymbol{Q}_k=\boldsymbol{I},k\in[K]}\ \min_{1\le i<j\le K}\|\boldsymbol{Q}_i\boldsymbol{Q}_i^{\mathsf{H}}-\boldsymbol{Q}_j\boldsymbol{Q}_j^{\mathsf{H}}\|_F^2 \tag{6.64}$$

$$= \arg\max_{\boldsymbol{Q}_k:\ \boldsymbol{Q}_k^{\mathsf{H}}\boldsymbol{Q}_k=\boldsymbol{I},k\in[K]}\ \min_{1\le i<j\le K}\|\boldsymbol{V}_i\boldsymbol{V}_i^{\mathsf{H}}-\boldsymbol{V}_j\boldsymbol{V}_j^{\mathsf{H}}\|_F^2, \tag{6.65}$$

since $\boldsymbol{V}_k$ is a basis of the orthogonal complement of $\boldsymbol{Q}_k$, $k\in[K]$. When $T\ge K(K-1)$, the solution is obvious: we set all $\boldsymbol{V}_k$ orthogonal to each other as in the following example.

Example 6.1 (Precoder Type I)**.** *If $T\ge K(K-1)$, for $k\in[K]$,*

$$\boldsymbol{U}_k = [\boldsymbol{e}_1\ \ldots\ \boldsymbol{e}_{(k-1)(K-1)}\ \boldsymbol{e}_{k(K-1)+1}\ \ldots\ \boldsymbol{e}_T], \tag{6.66}$$

$$\boldsymbol{V}_k = [\boldsymbol{e}_{(k-1)(K-1)+1}\ \ldots\ \boldsymbol{e}_{k(K-1)}]. \tag{6.67}$$

For $\boldsymbol{R}_k$, we consider the special case $\boldsymbol{R}_k = \mathrm{diag}(\eta_1^{(k)},\eta_2^{(k)},\ldots,\eta_{T-K+1}^{(k)})$, where the factor $\eta_i^{(k)}$ indicates the weight of a symbol in the dimension of $\mathcal{U}_k$ represented by column i of $\boldsymbol{U}_k$. By adjusting these factors, we control the orientation of the symbols in $\mathcal{U}_k$. When $\mathcal{U}_1,\ldots,\mathcal{U}_K$ overlap, let us denote by $\mathcal{U}_\cap = \bigcap_{k=1}^K\mathcal{U}_k$ the intersection of these subspaces. If there exists a set of symbols $\boldsymbol{x}_k\neq\boldsymbol{x}_k'\in\mathcal{U}_k, k\in[K]$ and $\boldsymbol{x}_l\in\mathcal{U}_l, l\neq k$ all belonging to $\mathcal{U}_\cap$, then it can be shown that $d_k(\mathcal{X})$ in (6.36) is upper bounded by a constant, and so is $\min_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} d(\boldsymbol{X}\to\boldsymbol{X}')$. Therefore, we avoid placing the symbol in the intersection $\mathcal{U}_\cap$. In Fig. 6.2, this can be seen as putting the points representing the symbols further away from the point I representing the intersection of $\mathcal{U}_1$ and $\mathcal{U}_2$. We next present an example of precoders in this spirit.

Example 6.2 (Precoder Type II)**.** *For $k\in[K]$,*

$$\boldsymbol{U}_k = \Big[\sqrt{\eta_1}\boldsymbol{e}_k\quad \sqrt{\eta_2}[\boldsymbol{e}_{K+1}\ \ldots\ \boldsymbol{e}_T]\Big], \tag{6.68}$$

$$\boldsymbol{V}_k = [\boldsymbol{e}_1\ \ldots\ \boldsymbol{e}_{k-1}\ \boldsymbol{e}_{k+1}\ \ldots\ \boldsymbol{e}_K], \tag{6.69}$$

where[3] *$\eta_1 = \frac{K(T-K+1)}{T}$ and $\eta_2 = \frac{T-K+1}{T}$. Note that the factor η_2 indicating the weight of a symbol in the dimensions in the intersection $\mathcal{U}_\cap$ is smaller than η_1 indicating the power in the mutually exclusive part of $\{\mathcal{U}_k\}$. Also, η_1 and η_2 are designed such that in $[\boldsymbol{U}_1\ \ldots\ \boldsymbol{U}_K]$, equal power is allocated to every direction.*

6.5 Power Optimization

When the users transmit at different rates, it might not be optimal to let them transmit at equal power. For example, in the extreme case when only one of the users transmits at

[3]Precoder Type II is of the same spirit as the pilot-based scheme: the users' signals are orthogonal in the first K channel uses.

non-zero rate, all other users should remain silent, i.e., transmit at power zero, to avoid causing interference. Therefore, power optimization also plays a key role. For a fixed constellation $\mathcal{X}$ (possibly generated with equal transmit power), we can further optimize the transmit power so as to maximize the metrics $e_{\min}(\mathcal{X})$ or $d_{\min}(\mathcal{X})$.

Let us focus on the two-user SIMO case and consider a joint Grassmannian constellation $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$ with average symbol power $\{P_1, P_2\}$, that is, $\mathcal{X}_k = \{\sqrt{P_k T}\boldsymbol{x}_k^{(i)} : \|\boldsymbol{x}_k^{(i)}\| = 1, i \in [|\mathcal{X}_k|]\}$, $k = 1, 2$. Let $\bar{\mathcal{X}} = \bar{\mathcal{X}}_1 \times \bar{\mathcal{X}}_2$ where $\bar{\mathcal{X}}_k = \{\boldsymbol{x}_k^{(i)}\}_{i=1}^{[|\mathcal{X}_k|]}$ is the set of the *normalized* symbols of user k. We assume that $\bar{\mathcal{X}}$ is fixed and would like to optimize the transmit powers $\{P_1, P_2\}$. To this end, we define $\theta := P_2/P_1$ and denote $\mathcal{X}$ as $\mathcal{X}^\theta$ for convenience. The parameter θ is the ratio between the transmit powers of the two users that we would like to optimize. From Section **??**, we deduce that $\min\{d_1(\mathcal{X}^\theta), d_2(\mathcal{X}^\theta)\} \le d_{\min}(\mathcal{X}^\theta) \le \min\{d_1(\mathcal{X}^\theta), d_2(\mathcal{X}^\theta)\} + 1$ where $d_1(\mathcal{X}^\theta) = \min\limits_{\boldsymbol{x}_1, \boldsymbol{x}_1', \boldsymbol{x}_2} \delta_1(\theta)$ and $d_2(\mathcal{X}^\theta) = \min\limits_{\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \hat{\boldsymbol{x}}_2'} \delta_2(\theta)$ with

$$\delta_1(\theta) := P_1 T \boldsymbol{x}_1^{\mathsf{H}} (\boldsymbol{I}_T + P_1 T \boldsymbol{x}_1' \boldsymbol{x}_1'^{\mathsf{H}} + \theta P_1 T \boldsymbol{x}_2 \boldsymbol{x}_2^{\mathsf{H}})^{-1} \boldsymbol{x}_1, \tag{6.70}$$

$$\delta_2(\theta) := \theta P_1 T \hat{\boldsymbol{x}}_2^{\mathsf{H}} (\boldsymbol{I}_T + P_1 T \hat{\boldsymbol{x}}_1 \hat{\boldsymbol{x}}_1^{\mathsf{H}} + \theta P_1 T \hat{\boldsymbol{x}}_2' \hat{\boldsymbol{x}}_2'^{\mathsf{H}})^{-1} \hat{\boldsymbol{x}}_2, \tag{6.71}$$

for $\{\boldsymbol{x}_1, \boldsymbol{x}_1', \hat{\boldsymbol{x}}_1\} \subset \bar{\mathcal{X}}_1, \{\boldsymbol{x}_2, \hat{\boldsymbol{x}}_2, \hat{\boldsymbol{x}}_2'\} \subset \bar{\mathcal{X}}_2$ such that $\boldsymbol{x}_1 \neq \boldsymbol{x}_1'$ and $\hat{\boldsymbol{x}}_2 \neq \hat{\boldsymbol{x}}_2'$. The optimal value of θ can be found by analyzing $d_1(\mathcal{X}^\theta)$ and $d_2(\mathcal{X}^\theta)$, as stated in the following proposition.

Proposition 6.7. *In the two-user SIMO case:*

1. $\min\{d_1(\mathcal{X}^\theta), d_2(\mathcal{X}^\theta)\}$ *is maximized at* $\theta = \theta^*$ *such that* $d_1(\mathcal{X}^{\theta^*}) = d_2(\mathcal{X}^{\theta^*})$, *and*

$$d_{\min}(\mathcal{X}^{\theta^*}) \le \max_\theta d_{\min}(\mathcal{X}^\theta) \le d_{\min}(\mathcal{X}^{\theta^*}) + 1; \tag{6.72}$$

2. *the value* θ^* *is given by* $\theta^* = \arg\min\limits_{\theta \in \boldsymbol{\Theta}} \delta(\theta)$ *where* $\boldsymbol{\Theta}$ *is the set of* $\theta > 0$ *such that*

$$\delta_1(\theta) = \delta_2(\theta) =: \delta(\theta) \tag{6.73}$$

for some $\{\boldsymbol{x}_1, \boldsymbol{x}_1', \hat{\boldsymbol{x}}_1\} \subset \bar{\mathcal{X}}_1, \{\boldsymbol{x}_2, \hat{\boldsymbol{x}}_2, \hat{\boldsymbol{x}}_2'\} \subset \bar{\mathcal{X}}_2$ *such that* $\boldsymbol{x}_1 \neq \boldsymbol{x}_1'$ *and* $\hat{\boldsymbol{x}}_2 \neq \hat{\boldsymbol{x}}_2'$; *the solution to* (6.73) *is given by*

$$\theta = \frac{1}{3a}\left[2\sqrt{\Delta}\cos\left(\frac{1}{3}\arccos\left(\frac{9abc - 2b^3 - 27a^2 d}{2\sqrt{\Delta^3}}\right)\right) - b\right] \tag{6.74a}$$

with

$$a := P_1 T\big[1 + P_1 T(1 - |\boldsymbol{x}_1'^{\mathsf{H}} \boldsymbol{x}_2|^2)\big] e_2, \tag{6.74b}$$

$$b := (1 + P_1 T) e_2 + \big[1 + P_1 T(1 - |\boldsymbol{x}_1'^{\mathsf{H}} \boldsymbol{x}_2|^2)\big]\big[1 + P_1 T(1 - |\hat{\boldsymbol{x}}_1^{\mathsf{H}} \hat{\boldsymbol{x}}_2|^2)\big] - \big[P_1 T + P_1^2 T^2 (1 - |\hat{\boldsymbol{x}}_1^{\mathsf{H}} \hat{\boldsymbol{x}}_2'|^2)\big] e_1, \tag{6.74c}$$

$$c := -(1 + P_1 T) e_1 - \big[1 + P_1 T(1 - |\hat{\boldsymbol{x}}_1^{\mathsf{H}} \hat{\boldsymbol{x}}_2'|^2)\big]\big[1 + P_1 T(1 - |\boldsymbol{x}_1^{\mathsf{H}} \boldsymbol{x}'_1|^2)\big] + \big(1 + \tfrac{1}{P_1 T}\big)\big[1 + P_1 T(1 - |\hat{\boldsymbol{x}}_1^{\mathsf{H}} \hat{\boldsymbol{x}}_2|^2)\big], \tag{6.74d}$$

$$d := -\big(1 + \tfrac{1}{P_1 T}\big)\big[1 + P_1 T(1 - |\boldsymbol{x}_1^{\mathsf{H}} \boldsymbol{x}'_1|^2)\big], \tag{6.74e}$$

$$e_1 := 1 - |\boldsymbol{x}_1^{\mathsf{H}} \boldsymbol{x}_2|^2 + P_1 T\big[(1 - |\boldsymbol{x}_1^{\mathsf{H}} \boldsymbol{x}_1'|^2)(1 - |\boldsymbol{x}'^{\mathsf{H}}_1 \boldsymbol{x}_2|^2) - |\boldsymbol{x}_1^{\mathsf{H}} \boldsymbol{x}_1' \boldsymbol{x}'^{\mathsf{H}}_1 \boldsymbol{x}_2 - \boldsymbol{x}_1^{\mathsf{H}} \boldsymbol{x}_2|^2\big], \tag{6.74f}$$

$$e_2 := 1 - |\hat{\boldsymbol{x}}_2^{\mathsf{H}} \hat{\boldsymbol{x}}_2'|^2 + P_1 T\big[(1 - |\hat{\boldsymbol{x}}_2^{\mathsf{H}} \hat{\boldsymbol{x}}_1|^2)(1 - |\hat{\boldsymbol{x}}_1^{\mathsf{H}} \hat{\boldsymbol{x}}_2'|^2) - |\hat{\boldsymbol{x}}_2^{\mathsf{H}} \hat{\boldsymbol{x}}_1 \hat{\boldsymbol{x}}_1^{\mathsf{H}} \hat{\boldsymbol{x}}_2' - \hat{\boldsymbol{x}}_2^{\mathsf{H}} \hat{\boldsymbol{x}}_2'|^2\big], \tag{6.74g}$$

$$\Delta := b^2 - 3ac. \tag{6.74h}$$

Proof. Throughout proof, $\boldsymbol{x}_1, \boldsymbol{x}_1', \hat{\boldsymbol{x}}, \boldsymbol{x}_2, \hat{\boldsymbol{x}}_2, \hat{\boldsymbol{x}}_2'$ implicitly satisfy the conditions mentioned after (6.73), i.e., $\{\boldsymbol{x}_1, \boldsymbol{x}_1', \hat{\boldsymbol{x}}_1\} \subset \bar{\mathcal{X}}_1, \{\boldsymbol{x}_2, \hat{\boldsymbol{x}}_2, \hat{\boldsymbol{x}}_2'\} \subset \bar{\mathcal{X}}_2$ such that $\boldsymbol{x}_1 \neq \boldsymbol{x}_1'$ and $\hat{\boldsymbol{x}}_2 \neq \hat{\boldsymbol{x}}_2'$.

1. From Lemma 6.5 in Appendix 6.A.1.b, we have that $\delta_1(\theta)$ is monotonically decreasing in θ for any $\boldsymbol{x}_1, \boldsymbol{x}_1', \boldsymbol{x}_2$, so $d_1(\mathcal{X}^\theta) = \min\limits_{\boldsymbol{x}_1, \boldsymbol{x}_1', \boldsymbol{x}_2} \delta_1(\theta)$ (see (6.30)) is also monotonically decreasing in θ. Also from Lemma 6.5, $\delta_2(\theta)$ is strictly increasing in θ for any $\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \hat{\boldsymbol{x}}_2'$, and so is $d_2(\mathcal{X}^\theta) = \min\limits_{\hat{\boldsymbol{x}}_2, \hat{\boldsymbol{x}}_2', \hat{\boldsymbol{x}}_1} \delta_2(\theta)$ (see (6.31)). Furthermore, $\delta_1(0) = P_1 T - \frac{P_1^2 T^2 |\boldsymbol{x}_1^{\mathsf{H}} \boldsymbol{x}_1'|^2}{1+P_1 T} > \frac{P_1 T}{1+P_1 T} > 0 = \delta_2(0)$ for any $\boldsymbol{x}_1, \boldsymbol{x}_1', \hat{\boldsymbol{x}}, \boldsymbol{x}_2, \hat{\boldsymbol{x}}_2, \hat{\boldsymbol{x}}_2'$, so $d_1(\mathcal{X}^\theta) > d_2(\mathcal{X}^\theta)$ at $\theta = 0$. Therefore, there exists a *unique* $\theta^* > 0$ such that $d_1(\mathcal{X}^{\theta^*}) = d_2(\mathcal{X}^{\theta^*})$, and thus θ^* maximizes $\min\{d_1(\mathcal{X}^\theta), d_2(\mathcal{X}^\theta)\}$.

   Let $\tilde{\theta} = \arg\max_\theta d_{\min}(\mathcal{X}^\theta)$. Then $d_{\min}(\mathcal{X}^{\theta^*}) \le d_{\min}(\mathcal{X}^{\tilde{\theta}})$. Also, we have that $d_{\min}(\mathcal{X}^{\tilde{\theta}}) \le \min\{d_1(\mathcal{X}^{\tilde{\theta}}), d_2(\mathcal{X}^{\tilde{\theta}})\} + 1 \le \min\{d_1(\mathcal{X}^{\theta^*}), d_2(\mathcal{X}^{\theta^*})\} + 1 \le d_{\min}(\mathcal{X}^{\theta^*}) + 1$, where the first and the third inequalities follow from (6.32), and the second inequality holds because θ^* maximizes $\min\{d_1(\mathcal{X}^\theta), d_2(\mathcal{X}^\theta)\}$. Therefore, $d_{\min}(\mathcal{X}^{\theta^*}) \le \max_\theta d_{\min}(\mathcal{X}^\theta) \le d_{\min}(\mathcal{X}^{\theta^*}) + 1$, implying that θ^* is approximately the solution to $\max_\theta d_{\min}(\mathcal{X}^\theta)$.

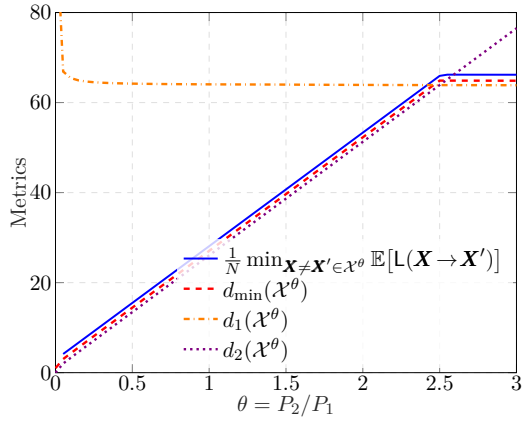
2. Since $d_1(\mathcal{X}^{\theta^*}) = d_2(\mathcal{X}^{\theta^*}) =: \delta(\theta^*)$, it is straightforward that $\theta^* \in \boldsymbol{\Theta}$. Let $\hat{\theta} := \arg\min\limits_{\theta \in \boldsymbol{\Theta}} \delta(\theta)$ and $\hat{\delta}_k(\theta)$ be the function $\delta_k(\theta)$ with $\boldsymbol{x}_1, \boldsymbol{x}_1', \hat{\boldsymbol{x}}, \boldsymbol{x}_2, \hat{\boldsymbol{x}}_2, \hat{\boldsymbol{x}}_2'$ satisfying (6.73), i.e., $\hat{\delta}_1(\hat{\theta}) = \hat{\delta}_2(\hat{\theta}) = \delta(\hat{\theta})$. We have $\hat{\delta}_1(\theta^*) \ge d_1(\mathcal{X}^{\theta^*}) = \delta(\theta^*) \ge \delta(\hat{\theta}) = \hat{\delta}_1(\hat{\theta})$ where the first equality follows from the min in $d_1(\mathcal{X}^{\theta^*})$ and the second inequality holds because $\theta^* \in \boldsymbol{\Theta}$. As a consequence, $\theta^* \le \hat{\theta}$ because $\hat{\delta}_1(\theta)$ is decreasing in θ. Similarly, we have that $\hat{\delta}_2(\theta^*) \ge d_2(\mathcal{X}^{\theta^*}) = \delta(\theta^*) \ge \delta(\hat{\theta}) = \hat{\delta}_2(\hat{\theta})$, so $\theta^* \ge \hat{\theta}$ because $\hat{\delta}_1(\theta)$ is increasing in θ. We conclude that $\theta^* = \hat{\theta}$. Finally, (6.73) can be written as a cubic equation $a\theta^3 + b\theta^2 + c\theta + d = 0$ (note that $a > 0$). Then, (6.74) follows by solving this equation for a positive root. □

The first part of Proposition 6.7 says that there exists a unique θ^* that maximizes $\min\{d_1(\mathcal{X}^\theta), d_2(\mathcal{X}^\theta)\}$, and this θ^* is also approximately the value of θ maximizing $d_{\min}(\mathcal{X}^\theta)$. The second part states that θ^* can be found by enumerating the closed-form expression (6.74) over the given set of normalized symbols $\bar{\mathcal{X}}_k$, $k = 1, 2$. In Fig. 6.3, we numerically verify Proposition 6.7 by plotting the values of $e_{\min}(\mathcal{X}^\theta)$, $d_{\min}(\mathcal{X}^\theta)$, $d_1(\mathcal{X}^\theta)$ and $d_2(\mathcal{X}^\theta)$ as a function of θ for $P_1 = 20$ dB and different $\bar{\mathcal{X}}$ with $T = 4$, $B_1 = 6$, and $B_2 = 2$. We see that $d_{\min}(\mathcal{X}^\theta)$ is within a constant gap from the minimum of $d_1(\mathcal{X}^\theta)$, which decreases with θ, and $d_2(\mathcal{X}^\theta)$, which increases with θ. $d_{\min}(\mathcal{X}^\theta)$ is maximized approximately at θ^* such that $d_1(\mathcal{X}^{\theta^*}) = d_2(\mathcal{X}^{\theta^*})$. These observations agree with Proposition 6.7. Furthermore, θ^* is also near the value of θ that maximizes the metric $e_{\min}(\mathcal{X}^\theta)$.

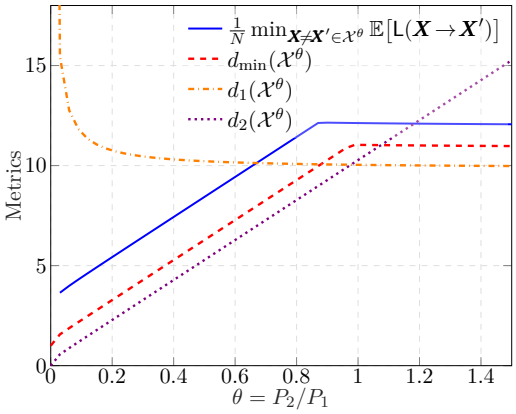
Therefore, given a joint normalized constellation $\bar{\mathcal{X}}$, one can further optimize the power so as to maximize the design metrics. Taking the power constraint P into account, the optimal power can be obtained as follows. First, we let $P_1 = P, P_2 = \theta P$, i.e., user 1 transmits at full power, and find the optimal θ in $[0, 1]$ by solving

$$\tilde{\theta} = \arg\max_{\theta \in [0,1]} e_{\min}(\mathcal{X}^\theta), \quad \text{or} \quad \tilde{\theta} = \arg\max_{\theta \in [0,1]} d_{\min}(\mathcal{X}^\theta). \tag{6.75}$$

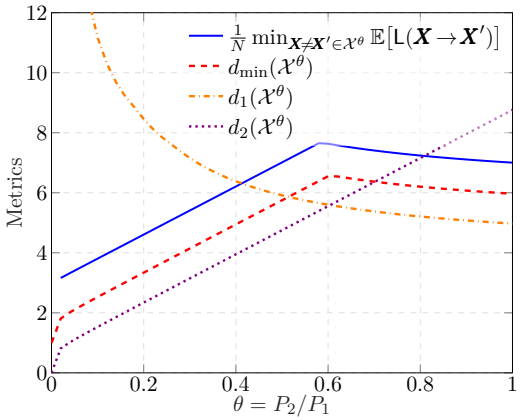
Following Proposition 6.7, $\tilde{\theta}$ can be approximated by θ^* such that $d_1(\mathcal{X}^{\theta^*}) = d_2(\mathcal{X}^{\theta^*})$, which can be found in closed form. Then, we let $P_2 = P, P_1 = \theta P_2$, i.e., user 2 transmits at full power, and solve (6.75) with the users' roles swapped to find the optimal $\breve{\theta}$ in $[0, 1]$. Finally, the optimal power is given by choosing between $\{P_1 = P, P_2 = \tilde{\theta} P\}$ and $\{P_2 = P, P_1 = \breve{\theta} P_2\}$ the option with higher metric value. In the numerical result in the next section, we shall see that it is favorable to let the user with higher transmission rate transmit at full power P.



(a) $\bar{\mathcal{X}}$ obtained by maximizing $d_{\min}(\mathcal{X})$ (6.23) at 30 dB for both users



(b) $\bar{\mathcal{X}}$ obtained by precoding with Precoder Type II



(c) $\bar{\mathcal{X}}$ obtained by partitioning

FIGURE 6.3: The values of the metrics $e_{\min}(\mathcal{X}^\theta)$, $d_{\min}(\mathcal{X}^\theta)$, $d_1(\mathcal{X}^\theta)$ and $d_2(\mathcal{X}^\theta)$ as a function of θ for $P_1 = 20$ dB, $P_2 = \theta P_1$, $T = 4$, $B_1 = 6$, $B_2 = 2$, and different given normalized constellation $\bar{\mathcal{X}}$.

In Table 6.1, we summarize the existing/proposed design criteria/constructions (not including the proposed criteria (6.43), (6.49), and the alternating optimization).

6.6 Numerical Results

In this section, we consider the single transmit antenna case ($M_1 = \cdots = M_K = 1$) and assume Grassmannian signaling, i.e., $\mathcal{X}_k = \{\sqrt{P_k T}\boldsymbol{x}_{ki}\}_{i=1}^{|\mathcal{X}_k|}$ with each $\boldsymbol{x}_{ki}$ being a unit-norm vector representative of a point in $G(\mathbb{C}^T, 1)$, $k \in [K]$, and a joint symbol $\boldsymbol{X} \in \mathcal{X}$ is formed as $\boldsymbol{X} = [\sqrt{P_1 T}\boldsymbol{x}_1 \ \ldots \ \sqrt{P_K T}\boldsymbol{x}_K]$ for $\boldsymbol{x}_k \in \mathcal{X}_k$. For the precoding design, we consider Precoder Type II and numerically optimized $\mathcal{C}_k$. We will compare our design to the constellations optimized with the criteria Min-m_1 (6.21) and Min-m_2 (6.22) in terms of joint ML symbol error rate (6.6). We also consider the joint constellation in which the symbols in $\mathcal{X}_k$ are generated as $\boldsymbol{x}_k = \sqrt{P_k}\left[\sqrt{K}\tilde{\boldsymbol{e}}_k^{\mathsf{T}} \ \sqrt{\frac{T-K}{P_{\text{avg}}}}\tilde{\boldsymbol{x}}_k^{\mathsf{T}}\right]^{\mathsf{T}}$ where $\tilde{\boldsymbol{e}}_k$ is the k-th column of $\boldsymbol{I}_K$ and $\tilde{\boldsymbol{x}}_k$ is a vector of scalar symbols in a QAM constellation with average power P_{avg}. This corresponds to the scenario where K users transmit mutually orthogonal pilot sequences, followed by spatially multiplexed parallel data transmission. With this *pilot-based constellation*, the receiver uses either an ML detector (6.4) or a linear MMSE detector consisting of MMSE channel estimation, MMSE equalization, and component-wise QAM demapper.

TABLE 6.1: The existing/proposed joint constellation design criteria/constructions

	Criterion/Construction	Shorthand	Motivation
Existing	$\mathcal{X}^* = \arg\min_{\mathcal{X}} m_1(\mathcal{X})$ (6.21)	Min-m_1	Treating $\mathcal{X}$ as a Grassmannian constellation in $G(\mathbb{C}^T, M_{\text{tot}})$
	$\mathcal{X}^* = \arg\min_{\mathcal{X}} m_2(\mathcal{X})$ (6.22) (from [68, Eq.(8)])	Min-m_2	
Proposed	$\mathcal{X}^* = \arg\max_{\mathcal{X}} e_{\min}(\mathcal{X})$ (6.19)	Max-$e_{\min}$	Minimizing the PEP
	$\mathcal{X}^* = \arg\max_{\mathcal{X}} d_{\min}(\mathcal{X})$ (6.23)	Max-$d_{\min}$	Maximizing the high-SNR dominant term in $e_{\min}(\mathcal{X})$
	Partitioning a single-user constellation $\mathcal{X}_{\text{SU}}$ (Section 6.4.1)	Partitioning	$d_{\min}(\mathcal{X})$ is large if $\mathcal{X}_{\text{SU}}$ is sparse enough
	Precoding single-user constellations in $G(\mathbb{C}^{T-K+1}, 1)$ (Section 6.4.2, for the SIMO MAC)	Precoding	Imposing a geometric separation between individual constellations

6.6.1 Numerical Optimization

We solve numerically Max-$e_{\min}$ (6.19), Max-$d_{\min}$ (6.23) and the alternating optimization for given powers $\{P_k\}$. In general, we want to solve the manifold-constrained optimization $\max_{\mathcal{X}=\mathcal{X}_1\times\cdots\times\mathcal{X}_K} \min_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} f(\boldsymbol{X},\boldsymbol{X}')$, where $f(\boldsymbol{X},\boldsymbol{X}')$ is customized according to the considered criterion. Note that the objective function is not smooth because of the min. To smoothen it, we use the approximation $\max_i x_i \approx \epsilon \ln \sum_i \exp(x_i/\epsilon)$ with a small ϵ and obtain

$$\min_{\mathcal{X}=\mathcal{X}_1\times\cdots\times\mathcal{X}_K} \epsilon \ln \sum_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \exp\left(-\frac{f(\boldsymbol{X},\boldsymbol{X}')}{\epsilon}\right). \tag{6.76}$$

This smooth optimization is, however, jointly over multiple points on the Grassmannian of lines. To tackle this, we construct the matrix $\boldsymbol{C} := [\boldsymbol{x}_{11}\ldots\boldsymbol{x}_{1|\mathcal{X}_1|}\ \ldots\ \boldsymbol{x}_{K1}\ldots\boldsymbol{x}_{K|\mathcal{X}_2|}] \in \mathbb{C}^{T\times\sum_{k=1}^K|\mathcal{X}_k|}$, then $\boldsymbol{C}$ belongs to the oblique manifold $\mathcal{OB}(T,\sum_{k=1}^K|\mathcal{X}_k|)$ (see Example 1.1). Consequently, the optimization problem (6.76) can be reformulated as a single-variable optimization in this oblique manifold as

$$\min_{\boldsymbol{C}\in\mathcal{OB}\left(T,\sum_{k=1}^K|\mathcal{X}_k|\right)} \underbrace{\epsilon \ln \sum_{\substack{\boldsymbol{X}=[\sqrt{P_1T}\boldsymbol{x}_{1i_1}\ \cdots\ \sqrt{P_KT}\boldsymbol{x}_{Ki_K}] \\ \neq\boldsymbol{X}'=[\sqrt{P_1T}\boldsymbol{x}_{1j_1}\cdots\sqrt{P_KT}\boldsymbol{x}_{Kj_K}]}} \exp\left(-\frac{f(\boldsymbol{X},\boldsymbol{X}')}{\epsilon}\right)}_{=:g(\boldsymbol{C})}. \tag{6.77}$$

In Appendix 6.A.5, we compute the Riemannian gradient of $g(\boldsymbol{C})$. Finally, we resort to the Manopt toolbox [141] to solve the optimization by conjugate gradient descent on the manifold. Recall that this toolbox provides several search direction construction rules and a restart strategy in order to ensure a convergence to a local minimum. Furthermore, we initialize the optimization with either the precoding-based, partitioning-based, or pilot-based constellation (which are easily generated), and choose the best optimized constellation among these options.

The other criteria (6.21) and (6.22) are similarly solved numerically with the Manopt toolbox. Hereafter, in all figures, the legends representing our proposed schemes are in bold face.

6.6.2 The Symmetrical Rate and Equal Power Case

We first consider the symmetrical rate setting $R_1 = \cdots = R_K = B/T$ with equal power $P_1 = \cdots = P_K = P$ for all users. We optimize the joint constellation at $P = 30$ dB (even when the performance of the resulting constellations is benchmarked at other SNR values).

6.6.2.a The Two-User Case

For the two-user ($K = 2$) case, in Fig. 6.4, we plot the joint symbol error rate (SER) (6.6) of the considered schemes for $T = 5$, $B = 5$, and $N = 4$. We observe that the constellations optimized with our metrics $e_{\min}(\mathcal{X})$ (6.19) and $d_{\min}(\mathcal{X})$ (6.23) achieve similar performance and are the best among the schemes with the same rate pair. The performance of the alternatively optimized constellation is slightly inferior to that and better than the pilot-based scheme. The partitioning design (with random partition) and the precoding design have similar performance. The constellations optimized with the Min-m_1 and Min-m_2 criteria, especially the latter, perform worse than that with our criteria.

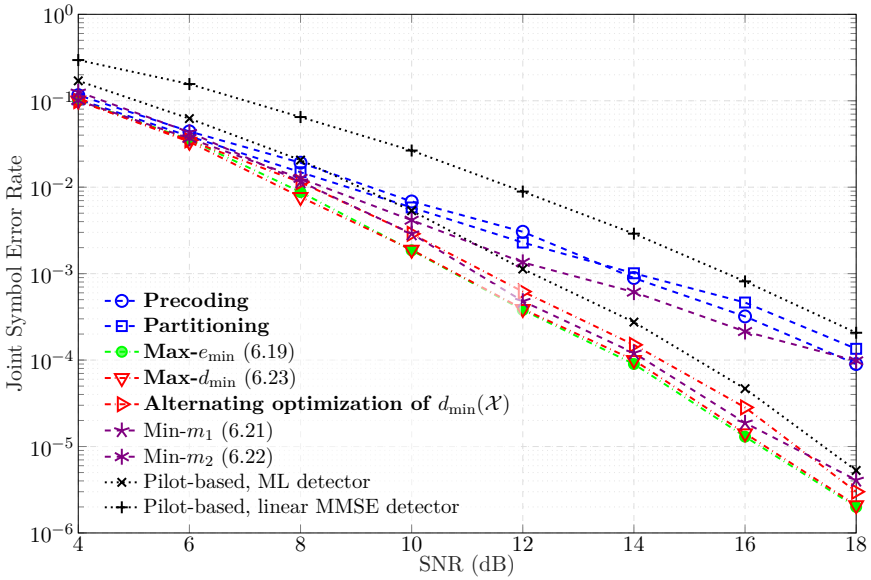
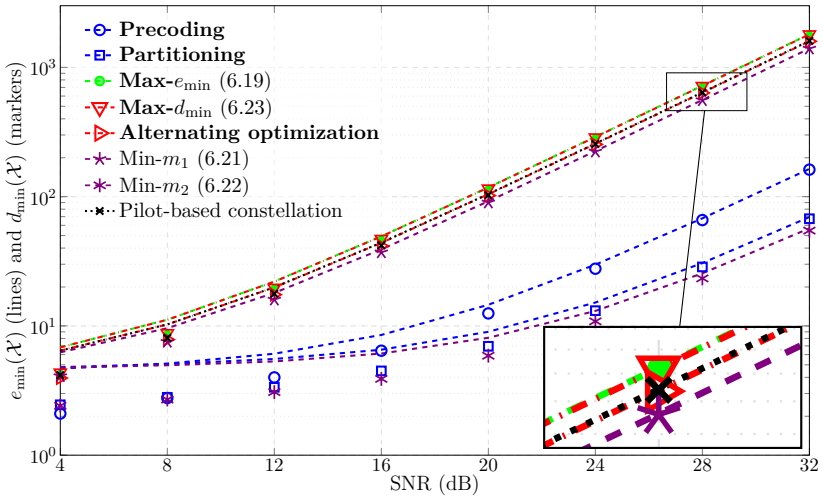


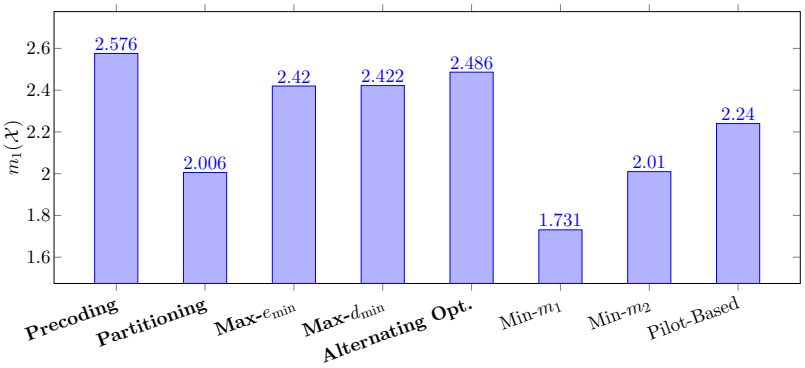
FIGURE 6.4: The joint SER of the proposed constellations compared to the baselines for $T = 5$, $K = 2$, $B = 5$, and $N = 4$.

To assess the effectiveness of the design metrics, we show the values of our metrics $d_{\min}(\mathcal{X})$ and $e_{\min}(\mathcal{X})$, and the other metrics $m_1(\mathcal{X})$ and $m_2(\mathcal{X})$ for these constellations in the same setting in Fig. 6.5. In Fig. 6.5(a), $d_{\min}(\mathcal{X})$ is very close to $e_{\min}(\mathcal{X})$ for SNR ≥ 20 dB. The constellations with low joint ML SER in Fig. 6.4 exhibit a large value of these metrics. This confirms that our proposed metrics are meaningful for constellation design and evaluation. In Fig. 6.5(b) and Fig. 6.5(c), we see that the relative order of the constellations in terms of the value of the metrics in (6.21) and (6.22) is unrevealing about the SER performance in Fig. 6.4. In particular, although the constellation optimized with the metric $m_1(\mathcal{X})$ (6.21) also achieves a low joint ML SER (and a high value of our metrics), this is not true for other constellations, such as the partitioning design.

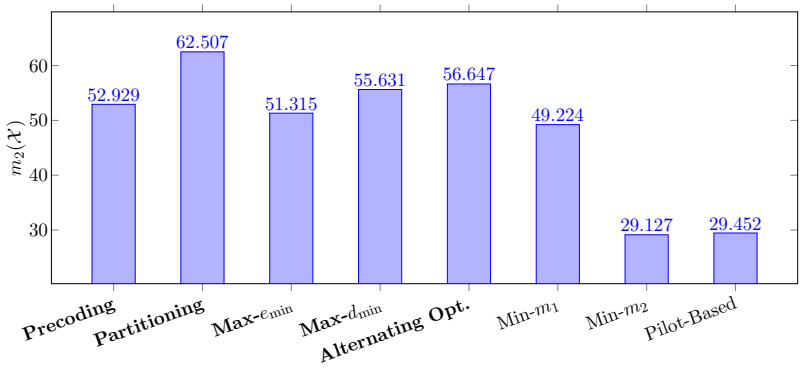
In Fig. 6.6, we consider larger constellations ($B = 8$) for which numerical optimization of the joint metrics becomes cumbersome. However, the partitioning and precoding constructions, which are based on our metrics, achieve good performance and outperform the pilot-based constellation.



(a) The metrics $e_{\min}(\mathcal{X})$ (lines) and $d_{\min}(\mathcal{X})$ (markers)



(b) The metric $m_1(\mathcal{X})$ in (6.21)

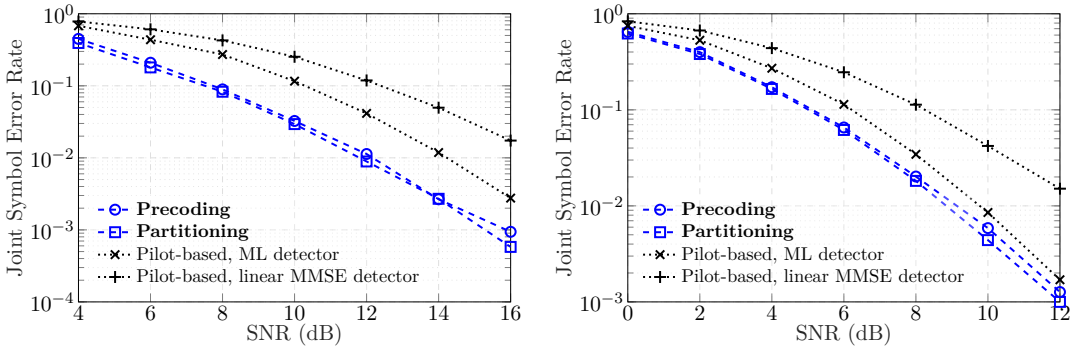


(c) The metric $m_2(\mathcal{X})$ in (6.22) with $N=4$

FIGURE 6.5: The value of the metrics $e_{\min}(\mathcal{X})$, $d_{\min}(\mathcal{X})$, $m_1(\mathcal{X})$ and $m_2(\mathcal{X})$ for the considered constellations for $T = 5$ and $B = 5$.

6.6.2.b The Three-User Case

In the three-user ($K = 3$) case, we consider $T = 4$, $B = 3$, $N = 4$, and plot the joint SER of various constellations in Fig. 6.7. We observe again that maximizing $d_{\min}(\mathcal{X})$ results in a good constellation that outperforms all others. The constellations obtained from Min-m_1 and Min-m_2 are outperformed by the pilot-based and precoding-based constellations. The SER of the Min-m_2 constellation and the partitioning-based constellation decreases slower with the SNR than the other schemes.



(a) $T = 5, B = 8, N = 4$ (b) $T = 6, B = 8, N = 4$

FIGURE 6.6: The joint SER of the partitioning design in comparison with the precoding design and a pilot-based scheme for $T \in \{5, 6\}$, $K = 2$, $B = 8$, and $N = 4$.

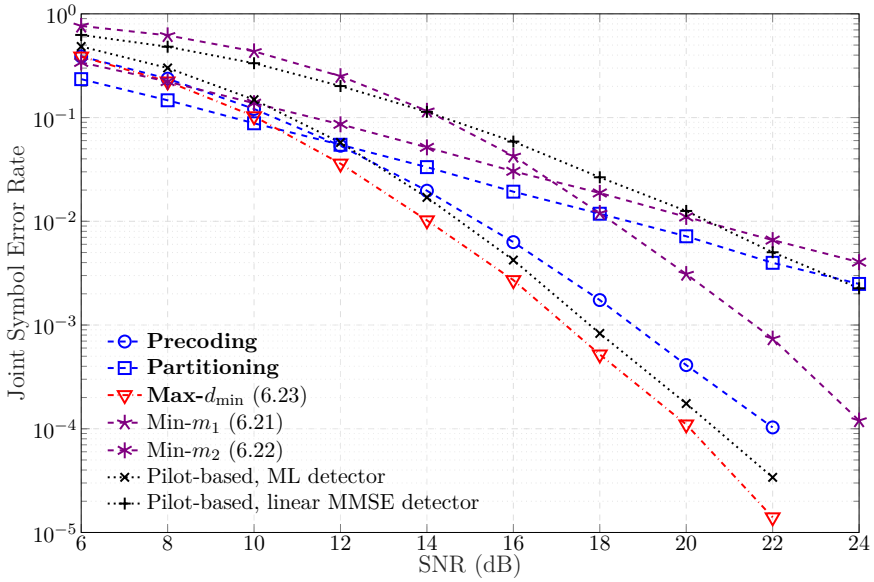
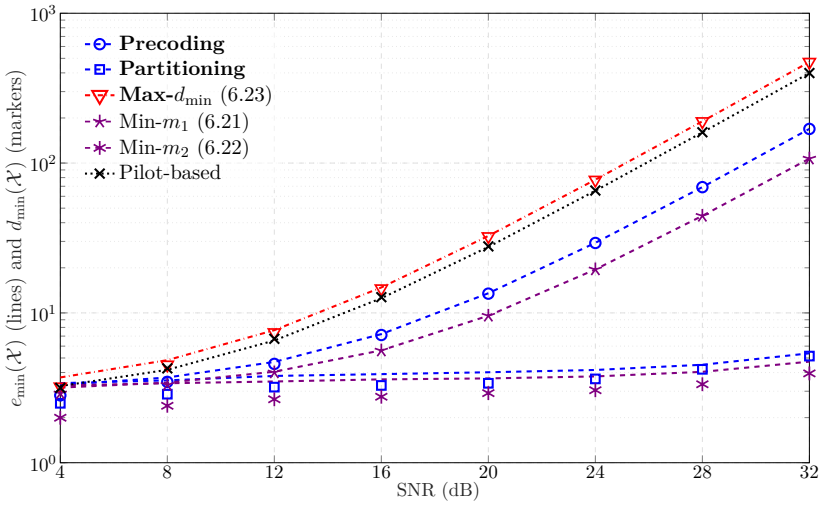


FIGURE 6.7: The joint SER of the proposed constellations compared to the baselines for $T = 4$, $K = 3$, $B = 3$, and $N = 4$.

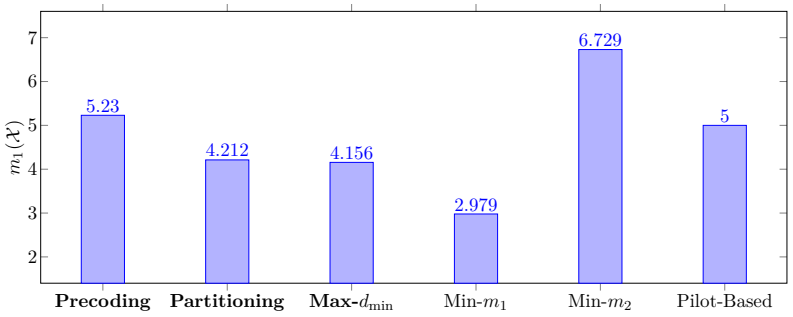
Fig. 6.8 depicts the values of our metrics $e_{\min}(\mathcal{X})$ and $d_{\min}(\mathcal{X})$, and the other metrics $m_1(\mathcal{X})$ and $m_2(\mathcal{X})$ for these constellations. In Fig. 6.8(a), the values of $e_{\min}(\mathcal{X})$ and $d_{\min}(\mathcal{X})$ are very close, and the relative order of these values for different constellations agree with the order of the SER performance of these constellations at high SNR in Fig. 6.7. For the Min-m_2 constellation and the partitioning-based constellation, the metrics $e_{\min}(\mathcal{X})$ and $d_{\min}(\mathcal{X})$ do not grow linearly in P, which is consistent with their SER performance. In fact, we can verify that these constellations violate (6.51) in Proposition 6.6 when P is large: the minimum pairwise chordal distance $\delta_{\min}(\mathcal{X}_{\text{SU}})$ of the concatenation $\mathcal{X}_{\text{SU}}$ of their individual constellations are about 0.2958 and 0.8264 respectively, both smaller than $\nu(K = 3, M = 1) \approx 0.9543$. From Fig. 6.8(b) and Fig. 6.8(c), we further observe that the metrics $m_1(\mathcal{X})$ and $m_2(\mathcal{X})$ are not meaningful for constellation design and evaluation.

6.6.3 The Asymmetrical Rate Case with Power Optimization

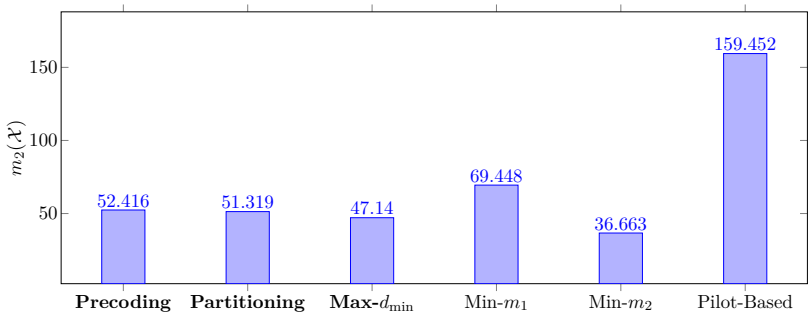
We now consider the asymmetrical rate case and focus on the two-user case. We set $T = 4$, $B_1 = 6$, and $B_2 = 2$ (as in Fig. 6.3). In Fig. 6.9, we plot the joint SER of the constellations generated by Max-$d_{\min}$, precoding, or partitioning and compare with



(a) The metrics $e_{\min}(\mathcal{X})$ (lines) and $d_{\min}(\mathcal{X})$ (markers)



(b) The metric $m_1(\mathcal{X})$ in (6.21)



(c) The metric $m_2(\mathcal{X})$ in (6.22) with $N = 4$

FIGURE 6.8: The value of the metrics $e_{\min}(\mathcal{X})$, and $d_{\min}(\mathcal{X})$, $m_1(\mathcal{X})$ and $m_2(\mathcal{X})$ for the considered constellations for $T = 4$, $K = 3$, and $B = 3$.

a pilot-based constellation with the same transmit rate for each user. Furthermore, we consider equal and full transmit power $P_1 = P_2 = P$, or optimized power as in Section 6.5. The constellations obtained by Max-$d_{\min}$ significantly outperforms other schemes. For this constellation, the optimal power coincides with full power $P_1 = P_2 = P$ for all $P > 4$ dB. For the precoding and partitioning designs, the optimal power is to let user 1 (which has higher transmission rate) transmit at full power $P_1 = P$ and user 2 at lower power $P_2 = \tilde{\theta}P$ with $\tilde{\theta}$ obtained from optimizing $d_{\min}(\mathcal{X})$ as in (6.75). The SER with optimized power is only slightly lower than the SER with full power. This is because the values of the metrics with optimized power is not significantly higher than that with full power, as seen in Fig. 6.3. However, using optimized power helps reduce the transmit power of user 2. In Fig. 6.10, we plot the optimized power fraction $\tilde{\theta}$ for user 2 obtained from (6.75). For the

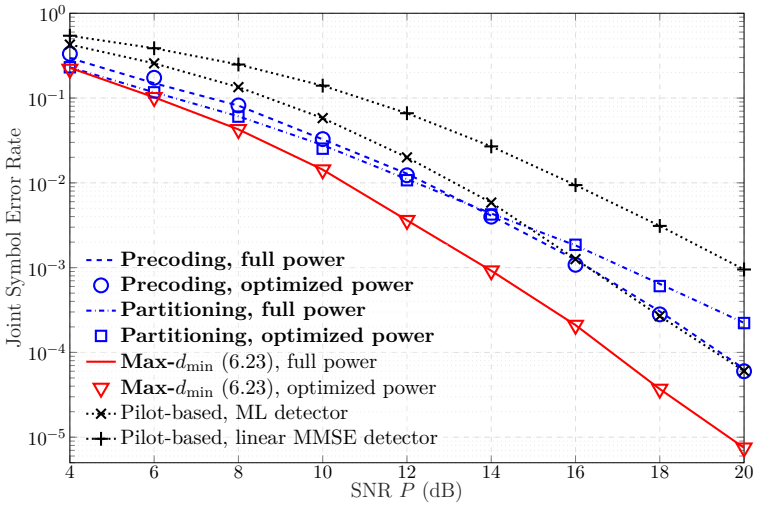


FIGURE 6.9: The joint SER of the proposed constellations with full transmit power $P_1 = P_2 = P$ or optimized transmit power as in Section 6.5, compared to a pilot-based constellation for $T = 4$, $K = 2$, $B_1 = 6$, $B_2 = 2$, and $N = 4$.

precoding design, as power constraint P grows, $\tilde{\theta}$ increases, i.e., user 2 should use more power. Whereas for the partitioning design, user 2 should use less power as P grows. We note that this behavior might not hold for all constellations of the kind.

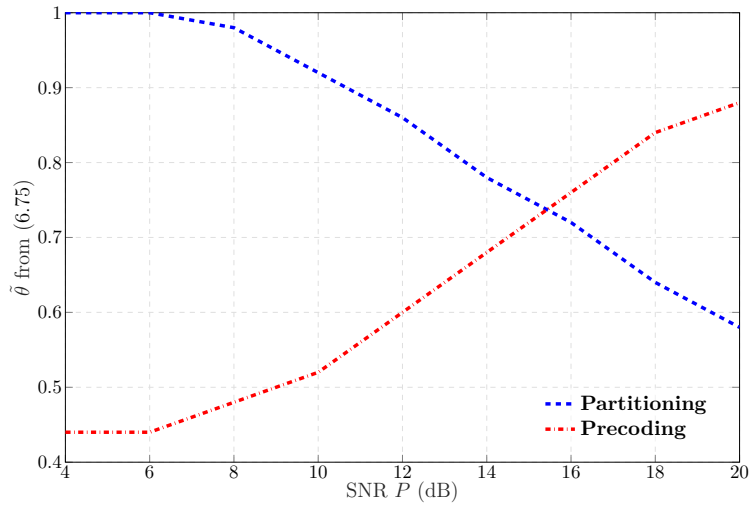


FIGURE 6.10: The optimal power fraction $\tilde{\theta}$ for user 2 by (6.75) for the precoding and partitioning designs for $T = 4$, $B_1 = 6$, $B_2 = 2$.

6.7 Closing Remarks

This chapter is an attempt of joint constellation design for non-coherent MIMO MAC in Rayleigh block fading. By analyzing the joint detection error, we have derived some closed-form metrics which turned out to be effective for designing joint constellations that achieve a low symbol error rate.

6.A Appendices

6.A.1 Mathematical Preliminaries

6.A.1.a Complex Wishart Matrix and the Distribution of the Trace

Definition 6.1 (Complex Wishart distribution)**.** *An $m \times m$ random matrix $\mathbf{A}$ is said to have a* complex Wishart distribution *with n degrees of freedom and covariance matrix $\mathbf{\Sigma}$, denoted by $\mathbf{A} \sim \mathcal{W}_m(n, \mathbf{\Sigma})$, if $\mathbf{A} = \mathbf{B}\mathbf{B}^{\mathsf{H}}$, where the columns of $\mathbf{B} \in \mathbb{C}^{m \times n}$ are independent and follow $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{\Sigma})$.*

The support of $\mathbf{A}$ is the space $\mathcal{S}_m$ of $m \times m$ positive definite Hermitian matrices. When $m \leq n$, the PDF of $\mathbf{A}$ is given by [196]

$$f(\mathbf{A}) = \frac{\det(\mathbf{A})^{n-m}}{\Gamma_m(n)\det(\mathbf{\Sigma})^n} \exp\Big(-\mathrm{tr}\Big(\mathbf{\Sigma}^{-1}\mathbf{A}\Big)\Big), \ \mathbf{A} \in \mathcal{S}_m, \tag{6.78}$$

where $\Gamma_m(n) := \pi^{m(m-1)/2} \prod_{i=1}^{m} \Gamma(n - i + 1)$ is the complex multivariate Gamma function. When $m > n$, the matrix $\mathbf{A}$ is singular and does not admit a PDF on $\mathcal{S}_m$. In this case, the distribution is referred to as the *singular Wishart*, and $\mathbf{A}$ admits a PDF on the space $\mathcal{S}_{m,n}$ of $m \times m$ positive semidefinite Hermitian matrices of rank n given by [197, Theorem 3]

$$f(\mathbf{A}) = \frac{\pi^{n(n-m)}\det(\mathbf{\Gamma})^{n-m}}{\Gamma_n(n)\det(\mathbf{\Sigma})^n} \exp\Big(-\mathrm{tr}\Big(\mathbf{\Sigma}^{-1}\mathbf{A}\Big)\Big), \ \mathbf{A} \in \mathcal{S}_{m,n}, \tag{6.79}$$

where $\mathbf{\Gamma}$ is the diagonal matrix containing the n non-zero eigenvalues of $\mathbf{A}$.

Lemma 6.2 (Distribution of the trace of a complex Wishart matrix)**.** *Let $\mathbf{A} \sim \mathcal{W}_m(n, \mathbf{\Sigma})$ and $\mathsf{u} := \mathrm{tr}(\mathbf{A})$. The moment generation function (MGF) of u is given by*

$$\phi_{\mathsf{u}}(t) = \det(\boldsymbol{I}_m - t\mathbf{\Sigma})^{-n}. \tag{6.80}$$

The CDF of u is given by

$$F_{\mathsf{u}}(u) = \varsigma^{mn}\det(\mathbf{\Sigma})^{-n} \sum_{k=0}^{\infty} \frac{\gamma(mn + k, u/\varsigma)}{k!(mn + k - 1)!} \sum_{\kappa(n,k)} [n]_{\kappa(n,k)} C_{\kappa(n,k)}(\boldsymbol{I}_m - \varsigma\mathbf{\Sigma}^{-1}) \tag{6.81}$$

for arbitrary $\varsigma > 0$, where the lower incomplete Gama function $\gamma(\cdot, \cdot)$, the partition $\kappa(n, k)$, the function $[n]_{\kappa(n,k)}$, and the complex zonal polynomial $C_{\kappa(n,k)}(\cdot)$ were defined in Proposition 6.3.

Proof. The proof follows by generalizing [192, Theorem 8.3.4] to the complex case and to include the non-singular case. Let

$$\tilde{\Gamma}(m, n) := \Gamma_m(n)\mathbb{1}\{m \leq n\} + \frac{\Gamma_n(n)}{\pi^n(n - m)}\mathbb{1}\{m > n\}, \tag{6.82}$$

$$\tilde{\mathcal{S}}(m, n) := \mathcal{S}_m\mathbb{1}\{m \leq n\} \cup \mathcal{S}_{m,n}\mathbb{1}\{m > n\}, \tag{6.83}$$

and $\widetilde{\det}(\mathbf{A})$ be the product of non-zero eigenvalues of $\mathbf{A}$. From the Wishart PDF given in (6.78) and (6.79), the MGF of u is derived as

$$\begin{aligned}
\phi_{\mathsf{u}}(t) &= \mathbb{E}[\exp(t\mathrm{tr}(\mathbf{A}))] && (6.84)\\
&= \int_{\mathbf{A}\in\tilde{\mathcal{S}}(m,n)} \exp(t\mathrm{tr}(\mathbf{A}))f(\mathbf{A})\,\mathrm{d}\mathbf{A} && (6.85)\\
&= \frac{\det(\boldsymbol{\Sigma})^{-n}}{\tilde{\Gamma}(m,n)} \int_{\mathbf{A}\in\tilde{\mathcal{S}}(m,n)} \exp\Big(-\mathrm{tr}\Big((\boldsymbol{\Sigma}^{-1}-t\boldsymbol{I}_m)\Big)\mathbf{A}\Big)\widetilde{\det}(\mathbf{A})^{n-m}\,\mathrm{d}\mathbf{A} && (6.86)\\
&= \det(\boldsymbol{\Sigma})^{-n}\det\Big(\boldsymbol{\Sigma}^{-1}-t\boldsymbol{I}_m\Big)^{-n} && (6.87)\\
&= \det(\boldsymbol{I}_m - t\boldsymbol{\Sigma})^{-n}. && (6.88)
\end{aligned}$$

In order to invert $\phi_{\mathsf{u}}(t)$ to find the PDF of u, we first write $\phi_{\mathsf{u}}(t)$ in terms of more familiar MGFs. Specifically, we write

$$\begin{aligned}
\phi_{\mathsf{u}}(t) &= (1-t\varsigma)^{-mn}\det\Big(\varsigma^{-1}\boldsymbol{\Sigma}\Big)^{-n}\det\Big(\boldsymbol{I}_m - \frac{1}{1-t\varsigma}(\boldsymbol{I}_m-\varsigma\boldsymbol{\Sigma}^{-1})\Big)^{-n} && (6.89)\\
&= (1-t\varsigma)^{-mn}\det\Big(\varsigma^{-1}\boldsymbol{\Sigma}\Big)^{-n}\sum_{k=0}^{\infty}\sum_{\kappa(n,k)}\frac{[n]_{\kappa(n,k)}C_{\kappa(n,k)}(\boldsymbol{I}_m-\varsigma\boldsymbol{\Sigma}^{-1})}{k!(1-t\varsigma)^k} && (6.90)\\
&= \varsigma^{mn}\det(\boldsymbol{\Sigma})^{-n}\sum_{k=0}^{\infty}\frac{1}{k!}(1-t\varsigma)^{-mn-k}\sum_{\kappa(n,k)}[n]_{\kappa(n,k)}C_{\kappa(n,k)}(\boldsymbol{I}_m-\varsigma\boldsymbol{\Sigma}^{-1}) && (6.91)
\end{aligned}$$

for arbitrary $\varsigma > 0$. Here, (6.90) follows from the fact that

$$\sum_{k=1}^{\infty}\sum_{\kappa(n,k)}\frac{1}{k!}[n]_{\kappa(n,k)}C_{\kappa(n,k)}(\boldsymbol{X}) = {}_1F_0(n;\boldsymbol{X}) = \det(\boldsymbol{I}_m-\boldsymbol{X})^{-n},$$

where ${}_1F_0(n;\boldsymbol{X})$ is a special case of the hypergeometric function with matrix argument [198, Section 2.1.2] and the infinite sum converges if $\|\boldsymbol{X}\|_2 < 1$ with $\|\cdot\|_2$ denoting the spectral norm of a matrix. Note that the convergence condition for the infinite sum in (6.90) reads $\|\boldsymbol{I}_m - \varsigma\boldsymbol{\Sigma}^{-1}\|_2 < |1-t\varsigma|$. By recognizing that $(1-t\varsigma)^{-mn-k}$ is the MGF of a Gamma distribution with shape $mn+k$ and scale ς, i.e. Gamma$(mn+k,\varsigma)$, we can invert $\phi_{\mathsf{u}}(t)$ term by term to get the PDF of u as

$$f_{\mathsf{u}}(u) = \varsigma^{mn}\det(\boldsymbol{\Sigma})^{-n}\sum_{k=0}^{\infty}\frac{1}{k!}g_{mn+k,\varsigma}(u)\sum_{\kappa(n,k)}[n]_{\kappa(n,k)}C_{\kappa(n,k)}(\boldsymbol{I}_m-\varsigma\boldsymbol{\Sigma}^{-1}), \tag{6.92}$$

where $g_{k,\theta}(x) := \frac{x^{k-1}e^{-x/\theta}}{(k-1)!\theta^k}$ denotes the PDF of the Gamma(k,θ) distribution. By integrating w.r.t. u, we get the CDF of u given in (6.81). □

6.A.1.b Eigenvalues of Matrices, and Other Preliminaries

For an $n\times n$ matrix $\boldsymbol{M}$, we denote by $\sigma_1(\boldsymbol{M}),\sigma_2(\boldsymbol{M}),\ldots,\sigma_n(\boldsymbol{M})$ its n eigenvalues sorted in decreasing order.

Lemma 6.3 (Eigenvalue perturbation of Hermitian matrix)**.** *For two $T\times T$ Hermitian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, it holds that*

$$|\sigma_i(\boldsymbol{A}+\boldsymbol{B}) - \sigma_i(\boldsymbol{A})| \le \|\boldsymbol{B}\|_{\mathrm{F}}, \quad \forall i\in[T]. \tag{6.93}$$

Proof. From [199, Corollary 8.1.6], $|\sigma_i(\boldsymbol{A}+\boldsymbol{B}) - \sigma_i(\boldsymbol{A})|$ is upper bounded by the spectral norm of $\boldsymbol{B}$. Then, (6.93) follows since the spectral norm is upper bounded by the Frobenius norm. □

Corollary 6.1. *For two $T \times T$ Hermitian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ whose entries depend on a parameter P, if $\|\boldsymbol{B}\|_{\mathrm{F}} = O(1)$ as $P \to \infty$, then*

$$\sigma_i(\boldsymbol{A}+\boldsymbol{B}) = \sigma_i(\boldsymbol{A}) + O(1), \quad \forall i \in [T], P \to \infty. \tag{6.94}$$

Lemma 6.4. *Let $\boldsymbol{Q} := \left[\begin{smallmatrix} \boldsymbol{I}_m & \boldsymbol{A}_{m\times n} \\ \boldsymbol{A}_{m\times n}^{\mathsf{H}} & \boldsymbol{I}_n \end{smallmatrix}\right]$ be positive semidefinite. Then, the $m+n$ eigenvalues of $\boldsymbol{Q}$ are*

$$1+\sigma_1(\boldsymbol{A}), \ldots, 1+\sigma_{\min\{m,n\}}(\boldsymbol{A}), 1, \ldots, 1, 1-\sigma_{\min\{m,n\}}(\boldsymbol{A}), \ldots, 1-\sigma_1(\boldsymbol{A}). \tag{6.95}$$

Proof. The singular value decomposition of $\boldsymbol{A}$ leads to a block diagonalization of $\boldsymbol{Q}$ with 2×2 blocks. The result then follows immediately. □

Lemma 6.5. *Consider three* distinct *T-dimensional unit vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, a parameter $\rho > 0$, a variable $\theta \geq 0$, and two functions $\delta_1(\theta) := \rho \boldsymbol{a}^{\mathsf{H}}(\boldsymbol{I}_T + \rho \boldsymbol{b}\boldsymbol{b}^{\mathsf{H}} + \rho\theta \boldsymbol{c}\boldsymbol{c}^{\mathsf{H}})^{-1}\boldsymbol{a}$ and $\delta_2(\theta) := \rho\theta \boldsymbol{a}^{\mathsf{H}}(\boldsymbol{I}_T + \rho \boldsymbol{b}\boldsymbol{b}^{\mathsf{H}} + \rho\theta \boldsymbol{c}\boldsymbol{c}^{\mathsf{H}})^{-1}\boldsymbol{a}$. Then, $\delta_1(\theta)$ is monotonically decreasing in θ while $\delta_2(\theta)$ is strictly increasing in θ.*

Proof. After some simple manipulations, we obtain

$$\frac{\partial \delta_1}{\partial \theta} = -\frac{\rho^2 |\rho \boldsymbol{a}^{\mathsf{H}}\boldsymbol{b}\boldsymbol{b}^{\mathsf{H}}\boldsymbol{c} - (1+\rho)\boldsymbol{a}^{\mathsf{H}}\boldsymbol{c}|^2}{(1+\rho+\rho(1+\rho(1-|\boldsymbol{b}^{\mathsf{H}}\boldsymbol{c}|^2))\theta)^2}, \tag{6.96}$$

$$\frac{\partial \delta_2}{\partial \theta} = \frac{1}{\rho + \rho^2(1-|\boldsymbol{b}^{\mathsf{H}}\boldsymbol{c}|^2)} \Bigg[\rho^2(1-|\boldsymbol{a}^{\mathsf{H}}\boldsymbol{c}|^2) + \rho^3 \big[(1-|\boldsymbol{a}^{\mathsf{H}}\boldsymbol{b}|^2)(1-|\boldsymbol{b}^{\mathsf{H}}\boldsymbol{c}|^2) - |\boldsymbol{a}^{\mathsf{H}}\boldsymbol{b}\boldsymbol{b}^{\mathsf{H}}\boldsymbol{c} - \boldsymbol{a}^{\mathsf{H}}\boldsymbol{c}|^2\big] + \frac{(1+\rho)|\rho \boldsymbol{a}^{\mathsf{H}}\boldsymbol{b}\boldsymbol{b}^{\mathsf{H}}\boldsymbol{c} - (1+\rho)\boldsymbol{a}^{\mathsf{H}}\boldsymbol{c}|^2}{(1+\rho+\rho(1+\rho(1-|\boldsymbol{b}^{\mathsf{H}}\boldsymbol{c}|^2))\theta)^2} \Bigg]. \tag{6.97}$$

It is obvious that $\frac{\partial \delta_1}{\partial \theta} \leq 0, \forall \theta \geq 0$. Let $\{\bar{\boldsymbol{b}}_i\}_{i=1}^{T-1}$ be a basis of the orthogonal complement of $\mathrm{Span}(\boldsymbol{b})$, i.e., $\boldsymbol{b}\boldsymbol{b}^{\mathsf{H}} + \sum_{i=1}^{T-1} \bar{\boldsymbol{b}}_i \bar{\boldsymbol{b}}_i^{\mathsf{H}} = \boldsymbol{I}_T$, we have that

$$\begin{aligned}
&(1-|\boldsymbol{a}^{\mathsf{H}}\boldsymbol{b}|^2)(1-|\boldsymbol{b}^{\mathsf{H}}\boldsymbol{c}|^2) - |\boldsymbol{a}^{\mathsf{H}}\boldsymbol{b}\boldsymbol{b}^{\mathsf{H}}\boldsymbol{c} - \boldsymbol{a}^{\mathsf{H}}\boldsymbol{c}|^2 \\
&= \boldsymbol{a}^{\mathsf{H}}(\boldsymbol{I}_T - \boldsymbol{b}\boldsymbol{b}^{\mathsf{H}})\boldsymbol{a}\boldsymbol{c}^{\mathsf{H}}(\boldsymbol{I}_T - \boldsymbol{b}\boldsymbol{b}^{\mathsf{H}})\boldsymbol{c} - |\boldsymbol{a}^{\mathsf{H}}(\boldsymbol{I}_T - \boldsymbol{b}\boldsymbol{b}^{\mathsf{H}})\boldsymbol{c}|^2 && (6.98)\\
&= \left(\sum_{i=1}^{T-1} |\boldsymbol{a}^{\mathsf{H}}\bar{\boldsymbol{b}}_i|^2\right)\left(\sum_{i=1}^{T-1} |\boldsymbol{c}^{\mathsf{H}}\bar{\boldsymbol{b}}_i|^2\right) - \left|\sum_{i=1}^{T-1} \boldsymbol{a}^{\mathsf{H}}\bar{\boldsymbol{b}}_i\bar{\boldsymbol{b}}_i^{\mathsf{H}}\boldsymbol{c}\right|^2 && (6.99)\\
&\geq 0, && (6.100)
\end{aligned}$$

where the last inequality is due to the Cauchy–Schwarz inequality. This and $|\boldsymbol{a}^{\mathsf{H}}\boldsymbol{c}|^2 < 1$ (since $\boldsymbol{a} \neq \boldsymbol{c}$) imply that $\frac{\partial \delta_2}{\partial \theta} > 0, \forall \theta \geq 0$. □

6.A.2 Proof that $\lim_{N\to\infty} \mathbb{P}(\boldsymbol{X} \to \boldsymbol{X}') = 0$ for Any Pair of Distinct Symbols $\boldsymbol{X}$ and $\boldsymbol{X}'$ of an Identifiable Joint Constellation

From (6.12), we develop $\mathsf{L}(\boldsymbol{X} \to \boldsymbol{X}')$ as $\mathsf{L}(\boldsymbol{X} \to \boldsymbol{X}') = -N \ln\det(\boldsymbol{I}_T + \bar{\boldsymbol{\Lambda}}) + \sum_{i=1}^{T} \lambda_i \mathsf{g}_i$, where $\lambda_1, \ldots, \lambda_T$ are T eigenvalues of $\bar{\boldsymbol{\Lambda}} = (\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{\frac{1}{2}}(\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{\frac{1}{2}} - \boldsymbol{I}_T$, and $\mathsf{g}_1, \ldots, \mathsf{g}_T$ are independent Gamma random variables with shape N and scale 1. It follows that

$$\mathbb{E}\big[\mathsf{L}(\boldsymbol{X} \to \boldsymbol{X}')\big] = -N \ln\det(\boldsymbol{I}_T + \bar{\boldsymbol{\Lambda}}) + \textstyle\sum_{i=1}^{T} \lambda_i = -N \ln\det(\boldsymbol{I}_T + \bar{\boldsymbol{\Lambda}}) + N\mathrm{tr}(\bar{\boldsymbol{\Lambda}}), \tag{6.101}$$

$$\mathsf{Var}\big[\mathsf{L}(\boldsymbol{X} \to \boldsymbol{X}')\big] = N \textstyle\sum_{i=1}^{T} \lambda_i^2 = N\mathrm{tr}(\bar{\boldsymbol{\Lambda}}^2). \tag{6.102}$$

For any joint constellation satisfying the identifiability condition in Proposition 6.1, we have $\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}} \neq \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$, thus $\bar{\boldsymbol{\Lambda}} \neq \boldsymbol{0}$. Therefore, $\operatorname{tr}(\bar{\boldsymbol{\Lambda}}) - \ln\det(\boldsymbol{I}_T + \bar{\boldsymbol{\Lambda}})$ is strictly positive. We have that

$$\frac{\mathbb{E}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]^2}{\mathsf{Var}\big[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\big]} = N\frac{\big(\operatorname{tr}(\bar{\boldsymbol{\Lambda}}) - \ln\det(\boldsymbol{I}_T + \bar{\boldsymbol{\Lambda}})\big)^2}{\operatorname{tr}\big(\bar{\boldsymbol{\Lambda}}^2\big)} \to \infty, \quad \text{as } N\to\infty. \tag{6.103}$$

From this and (6.15), we conclude that $\lim_{N\to\infty} \mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}') = 0$ for any pair of distinct symbols $\boldsymbol{X}$ and $\boldsymbol{X}'$ of a joint constellation satisfying the identifiability condition.

6.A.3 Proof of Lemma 6.1

It is straightforward that $\operatorname{tr}\big((\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\big) = O(1)$ because the eigenvalues of $(\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}$ are all smaller than 1.

Using polar decomposition, the input matrix $\boldsymbol{X}$ can be decomposed into an orthonormal matrix $\boldsymbol{W} \in \mathbb{C}^{T\times M_{\text{tot}}}$ whose columns span the column space of $\boldsymbol{X}$ and a full-rank spanning matrix $\boldsymbol{D}$. That is $\boldsymbol{X} = \boldsymbol{W}\boldsymbol{D}^{\mathsf{H}}$, where $\|\boldsymbol{D}\|_F^2 = \Theta(P)$. Similarly, $\boldsymbol{X}' = \boldsymbol{W}'\boldsymbol{D}'^{\mathsf{H}}$, for some orthonormal matrix $\boldsymbol{W}' \in \mathbb{C}^{T\times M_{\text{tot}}}$ and some full-rank spanning matrices $\boldsymbol{D}'$ such that $\|\boldsymbol{D}'\|_F^2 = \Theta(P)$. We assume w.l.o.g. that the column subspaces of $\boldsymbol{X}$ and $\boldsymbol{X}'$ share $r \le M_{\text{tot}}$ eigenmodes and thus express $\boldsymbol{W}$ and $\boldsymbol{W}'$ as

$$\boldsymbol{W} = [\boldsymbol{U}\ \boldsymbol{V}], \quad \text{and} \quad \boldsymbol{W}' = [\boldsymbol{U}\ \boldsymbol{V}'], \tag{6.104}$$

with $\boldsymbol{U} \in \mathbb{C}^{T\times r}$, $\boldsymbol{V} \in \mathbb{C}^{T\times(M_{\text{tot}}-r)}$, and $\boldsymbol{V}' \in \mathbb{C}^{T\times(M_{\text{tot}}-r)}$ such that $\boldsymbol{U}^{\mathsf{H}}\boldsymbol{U} = \boldsymbol{I}$, $\boldsymbol{V}^{\mathsf{H}}\boldsymbol{V} = \boldsymbol{I}$, $\boldsymbol{V}'^{\mathsf{H}}\boldsymbol{V}' = \boldsymbol{I}$, $\boldsymbol{U}^{\mathsf{H}}\boldsymbol{V} = \boldsymbol{0}$, $\boldsymbol{U}^{\mathsf{H}}\boldsymbol{V}' = \boldsymbol{0}$, and $\boldsymbol{V}^{\mathsf{H}}\boldsymbol{V}' = \boldsymbol{0}$.

6.A.3.a Proof that $\ln\frac{\det(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})}{\det(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})} = O(1)$ if $\operatorname{Span}(\boldsymbol{X}) = \operatorname{Span}(\boldsymbol{X}')$ and $\Theta(\ln P)$ otherwise

Let

$$\boldsymbol{\Gamma} := (\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1} \tag{6.105}$$

$$= \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1} + (\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}. \tag{6.106}$$

Applying Corollary 6.1 with $\boldsymbol{A} = \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}$ and $\boldsymbol{B} = (\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}$, we have that

$$\sigma_i(\boldsymbol{\Gamma}) = \sigma_i(\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}) + O(1) \tag{6.107}$$

$$= \begin{cases} \sigma_i(\boldsymbol{X}'^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}\boldsymbol{X}') + O(1), & i \le M_{\text{tot}}, \\ O(1), & i > M_{\text{tot}}. \end{cases} \tag{6.108}$$

Recalling the decomposition $\boldsymbol{X} = \boldsymbol{W}\boldsymbol{D}^{\mathsf{H}}$, we have that $\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}} = \boldsymbol{W}\boldsymbol{\Sigma}\boldsymbol{W}^{\mathsf{H}}$ with $\boldsymbol{\Sigma} := \boldsymbol{D}^{\mathsf{H}}\boldsymbol{D}$. Let $\boldsymbol{W}_\perp$ be the orthonormal complement of $\boldsymbol{W}$, i.e., $[\boldsymbol{W}\ \boldsymbol{W}_\perp]$ is unitary. Thus, $\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}} = \boldsymbol{W}(\boldsymbol{I}_T + \boldsymbol{\Sigma})\boldsymbol{W}^{\mathsf{H}} + \boldsymbol{W}_\perp\boldsymbol{W}_\perp^{\mathsf{H}}$, $(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1} = \boldsymbol{W}(\boldsymbol{I}_T + \boldsymbol{\Sigma})^{-1}\boldsymbol{W}^{\mathsf{H}} + \boldsymbol{W}_\perp\boldsymbol{W}_\perp^{\mathsf{H}}$, and we can expand

$$\boldsymbol{X}'^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X})^{-1}\boldsymbol{X}'^{\mathsf{H}} = \boldsymbol{X}'^{\mathsf{H}}\boldsymbol{W}(\boldsymbol{I}_T + \boldsymbol{\Sigma})^{-1}\boldsymbol{W}^{\mathsf{H}}\boldsymbol{X}' + \boldsymbol{X}'^{\mathsf{H}}\boldsymbol{W}_\perp\boldsymbol{W}_\perp^{\mathsf{H}}\boldsymbol{X}'. \tag{6.109}$$

Recalling that $\boldsymbol{X}' = \boldsymbol{W}'\boldsymbol{D}'^{\mathsf{H}}$ and using (6.104), we have that

$$\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{W}(\boldsymbol{I}_T + \boldsymbol{\Sigma})^{-1}\boldsymbol{W}^{\mathsf{H}}\boldsymbol{X}' = \boldsymbol{D}'^{\mathsf{H}}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}(\boldsymbol{I}_T + \boldsymbol{\Sigma})^{-1}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{D}' \tag{6.110}$$

$$= \boldsymbol{D}_1'^{\mathsf{H}}(\boldsymbol{I}_r + \boldsymbol{\Sigma}_1)^{-1}\boldsymbol{D}_1', \tag{6.111}$$

where $\boldsymbol{D}'_1$ contains the first r columns of $\boldsymbol{D}'$ and $\boldsymbol{\Sigma}_1$ denotes the top-left $r \times r$ block of $\boldsymbol{\Sigma}$, respectively. Therefore, $\|\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{W}(\boldsymbol{I}_T+\boldsymbol{\Sigma})^{-1}\boldsymbol{W}^{\mathsf{H}}\boldsymbol{X}'\|_F^2 \leq \frac{\|\boldsymbol{D}'_1\|_{\mathrm{F}}^2}{1+\sigma_{\min}(\boldsymbol{\Sigma}_1)} = O(1)$, where $\sigma_{\min}(\boldsymbol{\Sigma}_1)$ is the smallest eigenvalue of $\boldsymbol{\Sigma}_1$. With this, we apply Corollary 6.1 to (6.109) and obtain $\sigma_i(\boldsymbol{X}'^{\mathsf{H}}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}\boldsymbol{X}') = \sigma_i(\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{W}_\perp\boldsymbol{W}_\perp^{\mathsf{H}}\boldsymbol{X}') + O(1)$. Plugging this in (6.108), we get that

$$\sigma_i(\boldsymbol{\Gamma}) = \begin{cases} \sigma_i(\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{W}_\perp\boldsymbol{W}_\perp^{\mathsf{H}}\boldsymbol{X}') + O(1), & i \leq M_{\mathrm{tot}}, \\ O(1), & i > M_{\mathrm{tot}}, \end{cases} \tag{6.112}$$

$$= \begin{cases} \Theta(P), & i \leq M_{\mathrm{tot}} - r, \\ O(1), & i > M_{\mathrm{tot}} - r. \end{cases} \tag{6.113}$$

If $\mathrm{Span}(\boldsymbol{X}) = \mathrm{Span}(\boldsymbol{X}')$, then $r = M_{\mathrm{tot}}$, so $\sigma_i(\boldsymbol{\Gamma}) = O(1)$ for all $i \in [T]$. Thus $\ln\det(\boldsymbol{\Gamma}) = O(1)$. Otherwise, $\sigma_i(\boldsymbol{\Gamma}) = \Theta(P)$ for some i, so $\ln\det(\boldsymbol{I}_T+\boldsymbol{\Gamma}) = \sum_{i=1}^T \ln\sigma_i = \Theta(\ln P)$.

6.A.3.b Proof that $\mathrm{tr}((\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}) = O(1)$ if $\mathrm{Span}(\boldsymbol{X}) = \mathrm{Span}(\boldsymbol{X}')$ and $\Theta(P)$ otherwise

We expand

$$\boldsymbol{X}^{\mathsf{H}}(\boldsymbol{I}+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X} = \boldsymbol{X}^{\mathsf{H}}\Big(\boldsymbol{I}-\boldsymbol{X}'(\boldsymbol{I}+\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}')^{-1}\boldsymbol{X}'^{\mathsf{H}}\Big)\boldsymbol{X} \tag{6.114}$$

$$= \boldsymbol{D}\boldsymbol{W}^{\mathsf{H}}\Big(\boldsymbol{I}_T-\boldsymbol{W}'\boldsymbol{D}'^{\mathsf{H}}(\boldsymbol{I}_T+\boldsymbol{D}'\boldsymbol{W}'^{\mathsf{H}}\boldsymbol{W}'\boldsymbol{D}'^{\mathsf{H}})^{-1}\boldsymbol{D}'\boldsymbol{W}'^{\mathsf{H}}\Big)\boldsymbol{W}\boldsymbol{D}^{\mathsf{H}} \tag{6.115}$$

$$= \boldsymbol{D}\boldsymbol{D}^{\mathsf{H}} - \boldsymbol{D}\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{D}'^{\mathsf{H}}(\boldsymbol{I}_T+\boldsymbol{D}'\boldsymbol{D}'^{\mathsf{H}})^{-1}\boldsymbol{D}'\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{D}^{\mathsf{H}} \tag{6.116}$$

$$= \boldsymbol{D}\boldsymbol{D}^{\mathsf{H}} - \begin{bmatrix}P^2T^2\boldsymbol{D}_1\boldsymbol{D}'^{\mathsf{H}}_1(\boldsymbol{I}_r+\boldsymbol{D}'_1\boldsymbol{D}'^{\mathsf{H}}_1)^{-1}\boldsymbol{D}'_1\boldsymbol{D}_1^{\mathsf{H}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix} \tag{6.117}$$

where $\boldsymbol{D}_1$ and $\boldsymbol{D}'_1$ contain the first r columns of $\boldsymbol{D}$ and $\boldsymbol{D}'$, respectively. Thus,

$$\mathrm{tr}\Big((\boldsymbol{I}+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\Big) = \mathrm{tr}(\boldsymbol{D}_2\boldsymbol{D}_2^{\mathsf{H}}) + \mathrm{tr}\Big(\boldsymbol{D}_1\boldsymbol{D}_1^{\mathsf{H}} - \boldsymbol{D}_1\boldsymbol{D}'^{\mathsf{H}}_1(\boldsymbol{I}_r+\boldsymbol{D}'_1\boldsymbol{D}'^{\mathsf{H}}_1)^{-1}\boldsymbol{D}'_1\boldsymbol{D}_1^{\mathsf{H}}\Big) \tag{6.118}$$

$$= \|\boldsymbol{D}_2\|_{\mathrm{F}}^2 + \mathrm{tr}\Big((\boldsymbol{I}_r+\boldsymbol{D}'_1\boldsymbol{D}'^{\mathsf{H}}_1)^{-1}\boldsymbol{D}_1^{\mathsf{H}}\boldsymbol{D}_1\Big), \tag{6.119}$$

where $\boldsymbol{D}_2$ contains the last $M_{\mathrm{tot}}-r$ columns of $\boldsymbol{D}$. Since $(\boldsymbol{I}_r+\boldsymbol{D}'_1\boldsymbol{D}'^{\mathsf{H}}_1)^{-1} \preceq (1+\sigma_{\min}(\boldsymbol{D}'^{\mathsf{H}}_1\boldsymbol{D}'_1))^{-1}\boldsymbol{I}$ where $\sigma_{\min}(\boldsymbol{D}'^{\mathsf{H}}_1\boldsymbol{D}'_1)$ is the smallest eigenvalue of $\boldsymbol{D}'^{\mathsf{H}}_1\boldsymbol{D}'_1$, we have that

$$\|\boldsymbol{D}_2\|_{\mathrm{F}}^2 \leq \mathrm{tr}\Big((\boldsymbol{I}+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\Big) \leq \|\boldsymbol{D}_2\|_{\mathrm{F}}^2 + \frac{\|\boldsymbol{D}_1\|_{\mathrm{F}}^2}{1+\sigma_{\min}(\boldsymbol{D}'^{\mathsf{H}}_1\boldsymbol{D}'_1)}. \tag{6.120}$$

If $\mathrm{Span}(\boldsymbol{X}) = \mathrm{Span}\,(\boldsymbol{X}')$, we have $r = M_{\mathrm{tot}}$ and thus $\boldsymbol{D}_2$ is an empty matrix. Therefore, $\mathrm{tr}\Big((\boldsymbol{I}+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\Big) \leq \frac{\|\boldsymbol{D}_1\|_{\mathrm{F}}^2}{1+\sigma_{\min}(\boldsymbol{D}'^{\mathsf{H}}_1\boldsymbol{D}'_1)} = O(1)$. Otherwise, $r < M_{\mathrm{tot}}$ and $\|\boldsymbol{D}_2\|_{\mathrm{F}}^2 = \Theta(P)$, thus $\mathrm{tr}\Big((\boldsymbol{I}+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\Big) = \Theta(P)$.

6.A.4 Proof of Proposition 6.5

Let us rewrite $\boldsymbol{X}'_k\boldsymbol{X}'^{\mathsf{H}}_k + \sum_{l\neq k}\boldsymbol{X}_l\boldsymbol{X}_l^{\mathsf{H}} = \bar{\boldsymbol{X}}\bar{\boldsymbol{X}}^{\mathsf{H}}$ where $\bar{\boldsymbol{X}} := [\boldsymbol{X}_1\ \ldots\ \boldsymbol{X}_{k-1}\ \boldsymbol{X}'_k\ \boldsymbol{X}_{k+1}\ \ldots\ \boldsymbol{X}_K] \in \mathcal{X}$. Then, the trace in (6.36) becomes

$$\mathrm{tr}\Big(\boldsymbol{X}_k^{\mathsf{H}}(\boldsymbol{I}_T+\bar{\boldsymbol{X}}\bar{\boldsymbol{X}}^{\mathsf{H}})^{-1}\boldsymbol{X}_k\Big) = \mathrm{tr}(\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{X}_k) - \mathrm{tr}\Big(\boldsymbol{X}_1^{\mathsf{H}}\bar{\boldsymbol{X}}(\boldsymbol{I}_T+\bar{\boldsymbol{X}}^{\mathsf{H}}\bar{\boldsymbol{X}})^{-1}\bar{\boldsymbol{X}}^{\mathsf{H}}\boldsymbol{X}_k\Big) \tag{6.121}$$

$$= PT - \mathrm{tr}\Big(\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{U}\boldsymbol{\Sigma}(\boldsymbol{I}_T+\boldsymbol{\Sigma}^2)^{-1}\boldsymbol{\Sigma}\boldsymbol{U}^{\mathsf{H}}\boldsymbol{X}_k\Big), \tag{6.122}$$

where $\bar{\boldsymbol{X}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathsf{H}}$ with $\boldsymbol{U} \in \mathbb{C}^{r\times T}$, $\boldsymbol{V} \in \mathbb{C}^{KM\times r}$ being orthogonal matrices, and r being the rank of $\bar{\boldsymbol{X}}$; $\boldsymbol{\Sigma}$ contains r singular values of $\bar{\boldsymbol{X}}$ in decreasing order. Then, since $(\boldsymbol{I}_T+\boldsymbol{\Sigma}^2)^{-1} \preceq (1+\sigma^2_{\min}(\bar{\boldsymbol{X}}))^{-1}\boldsymbol{I}$ with $\sigma^2_{\min}(\bar{\boldsymbol{X}})$ being the minimum non-zero singular value of $\bar{\boldsymbol{X}}$, we have

$$\begin{aligned}
&\operatorname{tr}\Big(\boldsymbol{X}_k^{\mathsf{H}}(\boldsymbol{I}_T+\bar{\boldsymbol{X}}\bar{\boldsymbol{X}}^{\mathsf{H}})^{-1}\boldsymbol{X}_k\Big)\\
&\geq PT-(1+\sigma^2_{\min}(\bar{\boldsymbol{X}}))^{-1}\operatorname{tr}(\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}\boldsymbol{U}^{\mathsf{H}}\boldsymbol{X}_k) &(6.123)\\
&= PT-(1+\sigma^2_{\min}(\bar{\boldsymbol{X}}))^{-1}\|\bar{\boldsymbol{X}}^{\mathsf{H}}\boldsymbol{X}_k\|_{\mathrm{F}}^2 &(6.124)\\
&= PT-(1+\sigma^2_{\min}(\bar{\boldsymbol{X}}))^{-1}\left(\|\boldsymbol{X}_k'^{\mathsf{H}}\boldsymbol{X}_k\|_{\mathrm{F}}^2+\sum_{l\neq k}\|\boldsymbol{X}_l^{\mathsf{H}}\boldsymbol{X}_k\|_{\mathrm{F}}^2\right). &(6.125)
\end{aligned}$$

From (6.125), the key is to find a lower bound on the non-zero singular value $\sigma_{\min}(\bar{\boldsymbol{X}})$.

- For any K, applying Lemma 6.3 with $\boldsymbol{A}=\boldsymbol{I}_{KM}$ and $\boldsymbol{B}=\bar{\boldsymbol{X}}^{\mathsf{H}}\bar{\boldsymbol{X}}-\frac{PT}{M}\boldsymbol{I}_{KM}$, we have that $|\sigma_{\min}(\bar{\boldsymbol{X}}^{\mathsf{H}}\bar{\boldsymbol{X}})-\frac{PT}{M}| \leq \|\bar{\boldsymbol{X}}^{\mathsf{H}}\bar{\boldsymbol{X}}-\frac{PT}{M}\boldsymbol{I}_{KM}\|_{\mathrm{F}} = \sqrt{\sum_{k\neq l\in[K]}\|\boldsymbol{X}_l^{\mathsf{H}}\boldsymbol{X}_k\|_{\mathrm{F}}^2} \leq PT\sqrt{K(K-1)c}$, thus $\sigma^2_{\min}(\bar{\boldsymbol{X}}) \geq PT(\frac{1}{M}-\sqrt{K(K-1)c})$.
- For $K=2$, the bound can be tightened. For $k\neq l\in\{1,2\}$, applying Lemma 6.4 to the matrix $\boldsymbol{Q}=\frac{M}{PT}\bar{\boldsymbol{X}}^{\mathsf{H}}\bar{\boldsymbol{X}}$ with $\boldsymbol{A}=\frac{M}{PT}\boldsymbol{X}_k'^{\mathsf{H}}\boldsymbol{X}_l$, we see that the minimum non-zero eigenvalues of $\boldsymbol{Q}$ is $1-\sigma^*(\frac{M}{PT}\boldsymbol{X}_k'^{\mathsf{H}}\boldsymbol{X}_l)$ if there exists at least one singular value of $\frac{M}{PT}\boldsymbol{X}_k'^{\mathsf{H}}\boldsymbol{X}_l$ strictly smaller than 1 and $\sigma^*(\frac{M}{PT}\boldsymbol{X}_k'^{\mathsf{H}}\boldsymbol{X}_l)$ is the largest among such values. Otherwise, if all singular values of $\frac{M}{PT}\boldsymbol{X}_k'^{\mathsf{H}}\boldsymbol{X}_l$ are 1, the minimum non-zero eigenvalue of $\boldsymbol{Q}$ is two. In any case, the minimum non-zero eigenvalue of $\boldsymbol{Q}$ is lower bounded by $1-\|\frac{M}{PT}\boldsymbol{X}_k'\boldsymbol{X}_l^{\mathsf{H}}\|_{\mathrm{F}} \geq 1-M\sqrt{c}$. Hence, $\sigma^2_{\min}(\bar{\boldsymbol{X}}) \geq PT(\frac{1}{M}-\sqrt{c})$.

Plugging the bound of $\sigma^2_{\min}(\bar{\boldsymbol{X}})$ into (6.125) yields (6.48).

6.A.5 The Riemannian Gradient of $g(\boldsymbol{C})$

According to [140, Section 3.6] (see (1.126)), the Riemannian gradient can be computed by projection as $\nabla_{\mathrm{R}}g(\boldsymbol{C}) = (\boldsymbol{I}_T-\boldsymbol{C}\boldsymbol{C}^{\mathsf{H}})\nabla_{\mathrm{E}}g(\boldsymbol{C})$ where

$$\nabla_{\mathrm{E}}g(\boldsymbol{C}) = \left[\frac{\partial g(\boldsymbol{C})}{\partial \boldsymbol{x}_{11}} \;\cdots\; \frac{\partial g(\boldsymbol{C})}{\partial \boldsymbol{x}_{1|\mathcal{X}_1|}} \;\cdots\; \frac{\partial g(\boldsymbol{C})}{\partial \boldsymbol{x}_{K1}} \;\cdots\; \frac{\partial g(\boldsymbol{C})}{\partial \boldsymbol{x}_{K|\mathcal{X}_K|}}\right]$$

is the Euclidean gradient of $g(\boldsymbol{C})$ with

$$\begin{aligned}
\frac{\partial g(\boldsymbol{C})}{\partial \boldsymbol{x}_{kn}} = -&\left(\sum_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}}\exp\left(-\frac{f(\boldsymbol{X},\boldsymbol{X}')}{\epsilon}\right)\right)^{-1}\\
&\times\sum_{\substack{\boldsymbol{X}=[\sqrt{P_1T}\boldsymbol{x}_{1i_1}\;\cdots\;\sqrt{P_KT}\boldsymbol{x}_{Ki_K}]\\ \neq\boldsymbol{X}'=[\sqrt{P_1T}\boldsymbol{x}_{1j_1}\cdots\sqrt{P_KT}\boldsymbol{x}_{1j_K}],\\ \{1i_1,\ldots,Ki_k,1j_1,\ldots,Kj_K\}\ni kn}}\exp\left(-\frac{f(\boldsymbol{X},\boldsymbol{X}')}{\epsilon}\right)\frac{\partial f(\boldsymbol{X},\boldsymbol{X}')}{\partial \boldsymbol{x}_{kn}}. \quad (6.126)
\end{aligned}$$

In our proposed criteria, $f(\boldsymbol{X},\boldsymbol{X}')$ is given by $\frac{1}{N}\mathbb{E}[\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')]$ for the criterion (6.19) and $\operatorname{tr}((\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})$, for the criterion (6.23). Essentially, we would like to compute the derivative of $d(\boldsymbol{X}\to\boldsymbol{X}') = \operatorname{tr}((\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})$ (the derivative of $\operatorname{tr}((\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1})$

is similar) and $\psi(\boldsymbol{X},\boldsymbol{X}') := \ln \frac{\det(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})}{\det(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})}$. With $\boldsymbol{X} = [\sqrt{P_1T}\boldsymbol{x}_{1i_1} \dots \sqrt{P_KT}\boldsymbol{x}_{Ki_K}]$ and $\boldsymbol{X}' = [\sqrt{P_1T}\boldsymbol{x}_{1j_1} \dots \sqrt{P_KT}\boldsymbol{x}_{1j_K}]$, after some manipulations, we get

$$\frac{\partial d(\boldsymbol{X}\to\boldsymbol{X}')}{\partial \boldsymbol{x}_{kn}} = \begin{cases} 2P_kT\Big(\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\Big)^{-1}\boldsymbol{x}_{kn}, \text{ if } n = i_k, \\ 2P_kT\Big(\frac{\boldsymbol{x}_{1n}^{\mathsf{H}}\boldsymbol{A}^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\boldsymbol{A}^{-1}\boldsymbol{x}_{1n}(\boldsymbol{I}_T+P_kT\boldsymbol{A}^{-1})}{(1+P_kT\boldsymbol{x}_{kn}^{\mathsf{H}}\boldsymbol{A}^{-1}\boldsymbol{x}_{kn})^2} - \frac{\boldsymbol{A}^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\boldsymbol{A}^{-1}}{1+P_kT\boldsymbol{x}_{kn}^{\mathsf{H}}\boldsymbol{A}^{-1}\boldsymbol{x}_{kn}}\Big)\boldsymbol{x}_{kn}, \\ \qquad \text{with } \boldsymbol{A} := \boldsymbol{I}_T + \sum_{l\neq k}\boldsymbol{x}_{lj_l}\boldsymbol{x}_{lj_l}^{\mathsf{H}}, \text{ if } n = j_k. \end{cases} \tag{6.127}$$

In the two-user case, we have $\psi(\boldsymbol{X},\boldsymbol{X}') = \ln \frac{Q-|\boldsymbol{x}_{1i_1}^{\mathsf{H}}\boldsymbol{x}_{2i_2}|^2}{Q-|\boldsymbol{x}_{1j_1}^{\mathsf{H}}\boldsymbol{x}_{2j_2}|^2}$ with $Q := (1+\frac{1}{P_1T})(1+\frac{1}{P_2T})$, so

$$\frac{\partial \psi(\boldsymbol{X},\boldsymbol{X}')}{\partial \boldsymbol{x}_{kn_k}} = \frac{2\boldsymbol{x}_{\bar{k}n_{\bar{k}}}\boldsymbol{x}_{\bar{k}n_{\bar{k}}}^{\mathsf{H}}}{Q-|\boldsymbol{x}_{kn_k}^{\mathsf{H}}\boldsymbol{x}_{\bar{k}n_{\bar{k}}}|^2}\boldsymbol{x}_{kn_k}, \quad \text{for } \ k\in\{1,2\}, n\in\{i,j\}, \bar{k} = 2-k. \tag{6.128}$$

(The gradients of the metrics in (6.21) and (6.22) are computed similarly.)

Chapter 7

A Multi-User Detection Scheme Based on Expectation Propagation

We consider the non-coherent SIMO multiple access channel under spatially correlated Rayleigh block fading. We propose a novel soft-output multi-user detector that computes an approximate marginal posterior of each transmitted signal. Our detector is based on EP approximate inference and has polynomial complexity in the number of users, number of receive antennas and channel coherence interval. We also propose two simplifications of this detector with reduced complexity. The proposed detectors can be used for general signaling with vector-valued symbols transmitted over each coherence block. In particular, with the precoding-based constellation in the previous chapter, the proposed detectors outperform a non-coherent detector with projection-based interference mitigation that exploits the precoder structure. With pilot-assisted signaling, the EP detector outperforms, in terms of symbol error rate, some conventional coherent pilot-based detectors, namely, a linear MMSE decoder, a sphere decoder and a joint channel estimation–data detection scheme. Our EP-based detectors produce accurate approximates of the true posterior leading to high achievable sum-rates. The gains of these detectors are further observed in terms of the bit error rate when using their soft outputs for a turbo channel decoder.

7.1 Overview

We consider a SIMO MAC in Rayleigh flat and block fading with coherence interval $T \geq 2$. We assume that the communication signals are independently transmitted from K single-antenna non-cooperating users. We design the detector without assuming any specific structure of the signal transmitted over a coherence block. Thus, the designed detector can be used for both pilot-assisted and pilot-free transmission. We consider the case where the receiver is interested not only in the hard detection of the symbols but also in their posterior marginal PMFs. This "soft" information is needed, for example, when computing the bit-wise LLRs required for soft-input soft-output channel decoding. Computing an exact marginal PMF would require enumerating all possible combinations of other-user signals, which is infeasible with many users, many antennas, or large constellations. Thus, we seek sub-optimal schemes with practical complexity.

In contrast to soft coherent MIMO detection, for which many schemes have been proposed (e.g., [103]–[105]), the soft non-coherent MIMO detection under general signaling, and Grassmannian signaling in particular, has not been well investigated. The list-based soft demapper in [80] reduces the number of terms considered in posterior marginalization by including only those symbols at a certain distance from a reference point. However, it was designed for the single-user case only and has no obvious generalization to the MAC. The semi-blind joint/iterative channel estimation and data detection approaches [99]–[102] for the MIMO P2P channel can be extended to the MAC. However, these schemes are restricted to transmitted signals with pilots.

In this chapter, we propose message-passing algorithms for posterior marginal inference of non-coherent multi-user MIMO transmissions over spatially correlated Rayleigh block fading channels. Our algorithms are based on EP approximate inference [200], [201]. EP provides an iterative framework for approximating posterior beliefs by parametric distributions in the exponential family [202, Section 1.6]. Although there are many possible ways to apply EP to our non-coherent multi-user detection problem, we do so by choosing as variable nodes the indices of the transmitted symbols and the noiseless received signal from each user. The EP algorithm passes messages between the corresponding variable nodes and factor nodes on a bipartite factor graph. In doing so, the approximate posteriors of these variables are iteratively refined. We also address numerical implementation issues of the EP algorithm.

To measure the accuracy of the approximate posterior generated by the soft detectors, we compute the mismatched sum-rate of the system that uses the approximate posterior as the decoding metric. This mismatched sum-rate approaches the achievable rate of the system as the approximate posterior gets close to the true posterior. We also evaluate the symbol error rate when using the proposed schemes for hard detection, and the bit error rate when using these schemes for turbo equalization with a standard turbo code.

The contributions of this chapter are summarized as follows:

1. We propose soft and hard multi-user detectors for the non-coherent SIMO MAC using EP approximate inference, and methods to stabilize the EP updates. The proposed detectors work for general vector-valued transmitted symbols within each channel coherence block, i.e., it is general enough to include both the pilot-assisted and pilot-free signaling cases.

2. We propose two simplifications of the EP detector with reduced complexity. The first one, so-called EPAK, is based on approximating the EP messages with Kronecker products. The second one can be interpreted as soft MMSE estimation and successive interference approximation (successive interference approximation (SIA)).

3. We design some efficient greedy detection schemes for the precoding-based constellation in Chapter 6, called POCIS.

4. We analyze the complexity and numerically evaluate the convergence, running time, and performance of the proposed EP, EPAK, and MMSE-SIA, and POCIS detectors, the optimal ML detector, a genie-aided detector, the state-of-the-art detector from [125], and some conventional coherent pilot-based schemes. Our results suggest that the proposed EP-based detectors offer significantly improved mismatched sum-rate, symbol error rate, and coded bit error rate w.r.t. some existing sub-optimal schemes, while having lower complexity than the ML detector.

To the best of our knowledge, our proposed approach is the first message-passing scheme for non-coherent multi-user MIMO detection in block fading with general constellations.

The remainder of this chapter is organized as follows. The system model is presented in Section 7.2. A brief review of EP is presented in Section 7.3, and the EP approach to non-coherent detection is presented in Section 7.4. In Section 7.5, two simplifications (MMSE-SIA and EPAK) of the EP detector are presented. Implementation aspects of EP, MMSE-SIA, and EPAK are discussed in Section 7.6. In Section 7.7, we present greedy detectors for the precoding-based constellations (from Chapter 6), which later serve as baseline for the EP-based detectors. Numerical results and conclusions are presented in Section 7.8 and Section 7.9, respectively. The mathematical preliminaries and proofs are provided in the appendices.

7.2 System Model

7.2.1 Channel Model

We consider a SIMO MAC in which K single-antenna users transmit to an N-antenna receiver. We assume that the channel is flat and block fading with an equal-length and synchronous (across the users) coherence block of T channel uses. That is, the channel vectors $\mathsf{h}_k \in \mathbb{C}^{N\times 1}$, which contain the fading coefficients between the transmit antenna of user $k \in [K]$ and the N receive antennas, remain constant within each coherence block of T channel uses and change independently between blocks. Furthermore, the *distribution* of h_k is assumed to be known, but its *realizations* are unknown to both ends of the channel. Since the users are not co-located, we assume that the h_k are independent across users. We consider Rayleigh fading with receiver-side correlation, i.e., $\mathsf{h}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{R}_k)$, where $\boldsymbol{R}_k \in \mathbb{C}^{N\times N}$ is the spatial correlation matrix. We assume that $\frac{1}{N}\mathrm{tr}(\boldsymbol{R}_k) =: \xi_k$ where ξ_k is the large-scale average channel gain from one of the receive antennas to user k. We assume that $T > K$ and $N \geq K$.

Within a coherence block, each transmitter k sends a signal vector $\mathsf{x}_k \in \mathbb{C}^T$, and the receiver receives a realization $\boldsymbol{Y}$ of the random matrix

$$\mathsf{Y} = \sum_{k=1}^{K} \mathsf{x}_k \mathsf{h}_k^{\mathsf{T}} + \mathsf{Z} = \mathsf{X}\mathsf{H}^{\mathsf{T}} + \mathsf{Z}, \tag{7.1}$$

where $\mathsf{X} = [\mathsf{x}_1 \ \ldots \ \mathsf{x}_K] \in \mathbb{C}^{T\times K}$ and $\mathsf{H} = [\mathsf{h}_1 \ \ldots \ \mathsf{h}_K] \in \mathbb{C}^{N\times K}$ concatenate the transmitted signals and channel vectors, respectively, $\mathsf{Z} \in \mathbb{C}^{T\times N}$ is the Gaussian noise with IID $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ entries independent of H, and the block index is omitted for simplicity. We assume that the transmitted signals have average unit norm, i.e., $\mathbb{E}[\|\mathsf{x}_k\|^2] = 1, k \in [K]$. Under this normalization, the SNR of the transmitted signal from user k at each receive antenna is $\mathrm{SNR}_k = \xi_k/(T\sigma^2)$.

We assume that the transmitted signals belong to *disjoint* finite discrete individual constellations with vector-valued symbols. That is, $\mathsf{x}_k \in \mathcal{X}_k := \left\{\boldsymbol{x}_k^{(1)}, \ldots, \boldsymbol{x}_k^{(|\mathcal{X}_k|)}\right\}$, $k \in [K]$. In particular, $\mathcal{X}_k$ can be a Grassmannian constellation on $G(\mathbb{C}^T, 1)$, i.e., each constellation symbol $\boldsymbol{x}_k^{(i)}$ is a unit-norm vector representative of a point in $G(\mathbb{C}^T, 1)$. Another example is when the constellation symbols contain pilots and scalar data symbols.[1] Each symbol in $\mathcal{X}_k$ is labeled with a binary sequence of length $B_k := \log_2 |\mathcal{X}_k|$.

[1] In this case, the constellations are disjoint thanks to the fact that pilot sequences are user-specific.

7.2.2 Multi-User Detection Problem

Given $\mathsf{X} = \boldsymbol{X} = [\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \ldots,\ \boldsymbol{x}_K]$, the received signal Y has zero mean complex jointly Gaussian entries with the correlation characterized by

$$\begin{aligned}\mathbb{E}\Big[[\mathsf{Y}]_{t_1,n_1}[\mathsf{Y}]^*_{t_2,n_2}\ \Big|\ \mathsf{X}=\boldsymbol{X}\Big] &= \sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\mathbb{E}\Big[[\mathsf{H}]_{k_1,n_1}[\mathsf{H}]^*_{k_2,n_2}\Big][S]_{t_1,k_1}[S]^*_{t_2,k_2}+\mathbb{E}\Big[[\mathsf{Z}]_{t_1,n_1}[\mathsf{Z}]^*_{t_2,n_2}\Big]\\ &= \sum_{k=1}^{K}[\boldsymbol{R}_k]_{n_1,n_2}[\boldsymbol{x}_k\boldsymbol{x}_k^{\mathsf{H}}]_{t_1,t_2}+\sigma^2\mathbb{1}\{t_1=t_2,n_1=n_2\}. \end{aligned}\tag{7.2}$$

Therefore, the conditional probability density $p_{\mathbf{Y}|\mathbf{X}}$, also known as likelihood function, is derived similar to [46, Eq.(9)] as

$$p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X}) = \frac{\exp\big(-\mathrm{vec}(\boldsymbol{Y}^{\mathsf{T}})^{\mathsf{H}}(\sigma^2\boldsymbol{I}_{NT}+\sum_{k=1}^{K}\boldsymbol{x}_k\boldsymbol{x}_k^{\mathsf{H}}\otimes\boldsymbol{R}_k)^{-1}\mathrm{vec}(\boldsymbol{Y}^{\mathsf{T}})\big)}{\pi^{NT}\det(\sigma^2\boldsymbol{I}_{NT}+\sum_{k=1}^{K}\boldsymbol{x}_k\boldsymbol{x}_k^{\mathsf{H}}\otimes\boldsymbol{R}_k)}.\tag{7.3}$$

Given the received signal $\mathsf{Y}=\boldsymbol{Y}$, the joint multi-user ML symbol decoder is then

$$\begin{aligned}\hat{\boldsymbol{X}} = \arg\min_{\boldsymbol{X}\in\prod_{k=1}^{K}\mathcal{X}_k}\Bigg(&\mathrm{vec}(\boldsymbol{Y}^{\mathsf{T}})^{\mathsf{H}}\Big(\sigma^2\boldsymbol{I}_{NT}+\sum_{k=1}^{K}\boldsymbol{x}_k\boldsymbol{x}_k^{\mathsf{H}}\otimes\boldsymbol{R}_k\Big)^{-1}\mathrm{vec}(\boldsymbol{Y}^{\mathsf{T}})\\ &+\log\det\Big(\sigma^2\boldsymbol{I}_{NT}+\sum_{k=1}^{K}\boldsymbol{x}_k\boldsymbol{x}_k^{\mathsf{H}}\otimes\boldsymbol{R}_k\Big)\Bigg).\end{aligned}\tag{7.4}$$

Since the ML decoding metric depends on $\boldsymbol{X}$ only through $\sum_{k=1}^{K}\boldsymbol{x}_k\boldsymbol{x}_k^{\mathsf{H}}\otimes\boldsymbol{R}_k$, for identifiability, it must hold that $\sum_{k=1}^{K}\boldsymbol{x}_k\boldsymbol{x}_k^{\mathsf{H}}\otimes\boldsymbol{R}_k \neq \sum_{k=1}^{K}\boldsymbol{x}'_k\boldsymbol{x}'^{\,\mathsf{H}}_k\otimes\boldsymbol{R}_k$ for any pair of distinct joint symbols $\boldsymbol{X}=[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_K]$ and $\boldsymbol{X}'=[\boldsymbol{x}'_1,\ldots,\boldsymbol{x}'_K]$ in $\prod_{k=1}^{K}\mathcal{X}_k$.

When a channel code is used, most channel decoders require the LLRs of the bits. The LLR of the j-th bit of user k, denoted by $\mathsf{b}_{k,j}$, given the observation $\mathsf{Y}=\boldsymbol{Y}$ is defined as

$$\mathrm{LLR}_{k,j}(\boldsymbol{Y}) := \log\frac{p_{\mathbf{Y}|\mathsf{b}_{k,j}}(\boldsymbol{Y}|1)}{p_{\mathbf{Y}|\mathsf{b}_{k,j}}(\boldsymbol{Y}|0)} = \log\frac{\sum_{\boldsymbol{\alpha}\in\mathcal{X}_{k,j}^{(1)}}p_{\mathbf{Y}|\mathbf{x}_k}(\boldsymbol{Y}|\boldsymbol{\alpha})}{\sum_{\boldsymbol{\beta}\in\mathcal{X}_{k,j}^{(0)}}p_{\mathbf{Y}|\mathbf{x}_k}(\boldsymbol{Y}|\boldsymbol{\beta})} = \log\frac{\sum_{\boldsymbol{\alpha}\in\mathcal{X}_{k,j}^{(1)}}p_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{\alpha}|\boldsymbol{Y})}{\sum_{\boldsymbol{\beta}\in\mathcal{X}_{k,j}^{(0)}}p_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{\beta}|\boldsymbol{Y})}\tag{7.5}$$

where $\mathcal{X}_{k,j}^{(b)}$ denotes the set of all possible symbols in $\mathcal{X}_k$ with the j-th bit being equal to b for $j\in[B_k]$ and $b\in\{0,1\}$. To compute (7.5), the posteriors $p_{\mathbf{x}_k|\mathbf{Y}}$, $k\in[K]$, are marginalized from

$$p_{\mathbf{X}|\mathbf{Y}}(\boldsymbol{X}|\boldsymbol{Y}) = \frac{p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X})p_{\mathbf{X}}(\boldsymbol{X})}{p_{\mathbf{Y}}(\boldsymbol{Y})} \propto p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X})p_{\mathbf{X}}(\boldsymbol{X}).\tag{7.6}$$

Assuming that the transmitted signals are independent and uniformly distributed over the respective constellations, the prior $p_{\mathbf{X}}$ factorizes as

$$\Pr(\mathsf{X}=[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_K]) = \prod_{k=1}^{K}\frac{1}{|\mathcal{X}_k|}\mathbb{1}\{\boldsymbol{x}_k\in\mathcal{X}_k\}.\tag{7.7}$$

On the other hand, the likelihood function $p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_K])$ involves all $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_K$ in such a manner that it does not straightforwardly factorize. Exact marginalization of $p_{\mathbf{X}|\mathbf{Y}}$ requires computing

$$p_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{x}_k|\boldsymbol{Y}) = \sum_{\boldsymbol{x}_l\in\mathcal{X}_l,\forall l\neq k}p_{\mathbf{X}|\mathbf{Y}}([\boldsymbol{x}_1,\ldots,\boldsymbol{x}_K]|\boldsymbol{Y})\quad\text{for}\quad k\in[K].\tag{7.8}$$

That is, it requires computing $p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X})$ (which requires the inversion of an $NT \times NT$ matrix) for all $\boldsymbol{X} \in \prod_{k=1}^{K} \mathcal{X}_k$. Thus, the total complexity of exact marginalization is $O(K^6 2^{KB})$.[2] This is formidable for many users or large constellations. Thus, we seek alternative approaches to estimate

$$p_{\mathbf{X}|\mathbf{Y}}([\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K]|\boldsymbol{Y}) \approx \hat{p}_{\mathbf{X}|\mathbf{Y}}([\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K]|\boldsymbol{Y}) = \prod_{k=1}^{K} \hat{p}_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{x}_k|\boldsymbol{Y}). \tag{7.9}$$

7.2.3 Achievable Rate

According to [25, Section II], the highest sum-rate reliably achievable with a given decoding metric $\hat{p}_{\mathbf{X}|\mathbf{Y}}$, so-called the mismatched sum-rate, is lower bounded by the generalized mutual information (GMI) given by

$$R_{\mathrm{GMI}} = \frac{1}{T} \sup_{s \geq 0} \mathbb{E}\left[\log_2 \frac{\hat{p}_{\mathbf{X}|\mathbf{Y}}(\mathbf{X}|\mathbf{Y})^s}{\sum_{\boldsymbol{X}' \in \prod_{k=1}^{K} \mathcal{X}_k} \Pr(\mathbf{X} = \boldsymbol{X}') \hat{p}_{\mathbf{X}|\mathbf{Y}}(\boldsymbol{X}'|\mathbf{Y})^s} \right] \tag{7.10}$$

$$= \frac{1}{T} \sup_{s \geq 0} \mathbb{E}\left[\sum_{k=1}^{K} B_k - \log_2 \frac{\sum_{\boldsymbol{X}' \in \prod_{k=1}^{K} \mathcal{X}_k} \hat{p}_{\mathbf{X}|\mathbf{Y}}(\boldsymbol{X}'|\boldsymbol{Y})^s}{\hat{p}_{\mathbf{X}|\mathbf{Y}}(\mathbf{X}|\mathbf{Y})^s} \right] \tag{7.11}$$

$$= \frac{1}{T} \sum_{k=1}^{K} B_k - \frac{1}{T} \inf_{s \geq 0} \mathbb{E}\left[\sum_{k=1}^{K} \log_2 \frac{\sum_{\boldsymbol{x}'_k \in \mathcal{X}_k} \hat{p}_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{x}'_k|\mathbf{Y})^s}{\hat{p}_{\mathbf{x}_k|\mathbf{Y}}(\mathbf{x}_k|\mathbf{Y})^s} \right] \tag{7.12}$$

bits/channel use, where the expectation is over the joint distribution of $\mathbf{X}$ and $\mathbf{Y}$, i.e., $p_{\mathbf{Y}|\mathbf{X}} p_{\mathbf{X}}$. Here, (7.11) holds because the transmitted symbols are independent and have uniform prior distribution, and (7.12) follows from the factorization of $\hat{p}_{\mathbf{X}|\mathbf{Y}}$ in (7.9). The generalized mutual information R_{GMI} is upper bounded by the sum-rate achieved with the optimal decoding metric $p_{\mathbf{X}|\mathbf{Y}}$ given by

$$R = \frac{1}{T} I(\mathbf{X}; \mathbf{Y}) = \frac{1}{T} h(\mathbf{X}) - \frac{1}{T} h(\mathbf{X}|\mathbf{Y}) \tag{7.13}$$

$$= \frac{1}{T} \sum_{k=1}^{K} B_k - \frac{1}{T} \mathbb{E}\left[\log_2 \frac{1}{p_{\mathbf{X}|\mathbf{Y}}(\mathbf{X}|\mathbf{Y})} \right] \tag{7.14}$$

$$= \frac{1}{T} \sum_{k=1}^{K} B_k - \frac{1}{T} \mathbb{E}\left[\log_2 \frac{\sum_{\boldsymbol{X}' \in \prod_{k=1}^{K} \mathcal{X}_k} p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\boldsymbol{X}')}{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X})} \right] \tag{7.15}$$

bits/channel use, where (7.15) follows from the Bayes' law and the uniformity of the prior distribution. R_{GMI} approaches R as $\hat{p}_{\mathbf{X}|\mathbf{Y}}$ gets close to $p_{\mathbf{X}|\mathbf{Y}}$. Note that if we fix $s = 1$ in place of the infimum in (7.12), it holds that

$$R - R_{\mathrm{GMI}}(s = 1) = \frac{1}{T} \mathbb{E}\left[\log_2 \frac{p_{\mathbf{X}|\mathbf{Y}}(\mathbf{X}|\mathbf{Y})}{\hat{p}_{\mathbf{X}|\mathbf{Y}}(\mathbf{X}|\mathbf{Y})} \right] = \frac{1}{T} \mathbb{E}_{\mathbf{Y}}\left[D(p_{\mathbf{X}|\mathbf{Y}} \| \hat{p}_{\mathbf{X}|\mathbf{Y}}) \right], \tag{7.16}$$

which converges to zero when the KL divergence between $\hat{p}_{\mathbf{X}|\mathbf{Y}}$ and $p_{\mathbf{X}|\mathbf{Y}}$ vanishes.

The expectations in (7.12) and (7.15) cannot be derived in closed form in general. Alternatively, we can evaluate R and R_{GMI} (and also $\mathbb{E}_{\mathbf{Y}}[D(p_{\mathbf{X}|\mathbf{Y}} \| \hat{p}_{\mathbf{X}|\mathbf{Y}})]$) numerically with

[2] Throughout the chapter, as far as the complexity analysis is concerned, we assume for notational simplicity that $T = O(K)$, $N = O(K)$, and $|\mathcal{X}_k| = O(2^B)$, $\forall k \in [K]$. If the channels are uncorrelated ($\boldsymbol{R}_k = \boldsymbol{I}_N$), the likelihood function can be simplified as $p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X}) = \frac{\exp\left(-\mathrm{tr}\left\{\boldsymbol{Y}^{\mathsf{H}}(\sigma^2 \boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}\boldsymbol{Y}\right\}\right)}{\pi^{NT} \det^N(\sigma^2 \boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})}$. Thus, the complexity of exact marginalization is reduced to $O(K^3 2^{KB})$.

the Monte Carlo method. Note that when K or B_k is large, even a numerical evaluation of R and $\mathbb{E}_{\mathbf{Y}}[D(p_{\mathbf{X}|\mathbf{Y}}\|\hat{p}_{\mathbf{X}|\mathbf{Y}})]$ is not possible. Therefore, we choose to use the mismatched sum-rate lower bound R_{GMI} as an information-theoretic metric to evaluate how close $\hat{p}_{\mathbf{X}|\mathbf{Y}}$ is to $p_{\mathbf{X}|\mathbf{Y}}$.

In what follows, we design a posterior marginal estimation scheme based on EP. We start by providing a brief review of EP in the next section.

7.3 Expectation Propagation

The EP algorithm was first proposed in [200] and summarized in, e.g., [201] for approximate inference in probabilistic graphical models. EP is an iterative framework for approximating posterior beliefs by parametric distributions in the exponential family [202, Section 1.6]. Let us consider a set of unknown variables represented by a random vector $\mathbf{s}$ with posterior of the form

$$p_{\mathbf{s}}(\boldsymbol{s}) \propto \prod_{\alpha} \psi_\alpha(\boldsymbol{s}_\alpha), \tag{7.17}$$

where $\boldsymbol{s}_\alpha$ is the subset of variables involved in the factor ψ_α corresponding to a partition $\{\mathbf{s}_\alpha\}$ of $\mathbf{s}$. Furthermore, let us partition the components of $\mathbf{s}$ into some sets $\{\mathbf{s}_\beta\}$, where no $\mathbf{s}_\beta$ is split across factors (i.e., $\forall\alpha,\beta$ either $\mathbf{s}_\beta \subset \mathbf{s}_\alpha$ or $\mathbf{s}_\beta \cap \mathbf{s}_\alpha = \emptyset$). The partition $\{\mathbf{s}_\alpha\}$ represents the local dependency of the variables given by the intrinsic factorization (7.17), while the partition $\{\mathbf{s}_\beta\}$ groups the variables that always occur together in a factor. We are interested in the posterior marginals w.r.t. the partition $\{\mathbf{s}_\beta\}$. In the following, we omit $\mathbf{s}$ in the subscripts since it is obvious.

EP approximates the true posterior p from (7.17) by a distribution $\hat{p}$ that can be expressed in two ways. First, it can be expressed w.r.t. the "target" partition $\{\mathbf{s}_\beta\}$ as

$$\hat{p}(\boldsymbol{s}) = \prod_{\beta} \hat{p}_\beta(\boldsymbol{s}_\beta), \tag{7.18}$$

where $\hat{p}_\beta$ are constrained to be in the exponential family (see Appendix 7.A.1), such that

$$\hat{p}_\beta(\boldsymbol{s}_\beta) = \exp\left(\boldsymbol{\gamma}_\beta^{\mathsf{T}}\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta) - A_\beta(\boldsymbol{\gamma}_\beta)\right), \tag{7.19}$$

for sufficient statistics $\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta)$, parameters $\boldsymbol{\gamma}_\beta$, and log-partition function

$$A_\beta(\boldsymbol{\gamma}) := \ln \int e^{\boldsymbol{\gamma}^{\mathsf{T}}\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta)}\, \mathrm{d}\boldsymbol{s}_\beta.$$

Second, $\hat{p}$ can also be expressed w.r.t. the partition $\{\mathbf{s}_\alpha\}$ as

$$\hat{p}(\boldsymbol{s}) \propto \prod_{\alpha} m_\alpha(\boldsymbol{s}_\alpha), \tag{7.20}$$

in accordance with (7.17). For (7.18) and (7.20) to be consistent, the terms m_α should also factorize over β, i.e., there exist factors $m_{\alpha,\beta}$ of the form $m_{\alpha,\beta}(\boldsymbol{s}_\beta) = \exp\left(\boldsymbol{\gamma}_{\alpha,\beta}^{\mathsf{T}}\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta)\right)$ such that

$$m_\alpha(\boldsymbol{s}_\alpha) = \prod_{\beta\in\mathfrak{N}_\alpha} m_{\alpha,\beta}(\boldsymbol{s}_\beta) = \exp\left(\sum_{\beta\in\mathfrak{N}_\alpha} \boldsymbol{\gamma}_{\alpha,\beta}^{\mathsf{T}}\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta)\right), \tag{7.21}$$

$$\hat{p}_\beta(\boldsymbol{s}_\beta) \propto \prod_{\alpha\in\mathfrak{N}_\beta} m_{\alpha,\beta}(\boldsymbol{s}_\beta) = \exp\left(\sum_{\alpha\in\mathfrak{N}_\beta} \boldsymbol{\gamma}_{\alpha,\beta}^{\mathsf{T}}\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta)\right), \tag{7.22}$$

where $\mathfrak{N}_\alpha$ collects the indices β for which $\mathbf{s}_\beta \subset \mathbf{s}_\alpha$, and $\mathfrak{N}_\beta$ collects the indices α for which $\mathbf{s}_\beta \subset \mathbf{s}_\alpha$. It turns out that $m_{\alpha,\beta}$ can be interpreted as a message from the factor node α to the variable node β on a bipartite factor graph [203]. In this case, $m_\alpha(\boldsymbol{s}_\alpha)$ is the product of all messages emanating from factor node α, and $\hat{p}_\beta(\boldsymbol{s}_\beta)$ is proportional to the product of all messages impinging on variable node β.

EP works by first initializing all $m_\alpha(\boldsymbol{s}_\alpha)$ and $\hat{p}_\beta(\boldsymbol{s}_\beta)$ (typically by the respective priors, which are assumed to also belong to the considered exponential family), then iteratively updating each approximation factor m_α in turn. Let us fix a factor index α. According to [200], the "tilted" distribution q_α is constructed by swapping the true potential ψ_α for its approximate m_α in $\hat{p}(\boldsymbol{s})$ as

$$q_\alpha(\boldsymbol{s}) = \frac{\hat{p}(\boldsymbol{s})\psi_\alpha(\boldsymbol{s}_\alpha)}{m_\alpha(\boldsymbol{s}_\alpha)}, \tag{7.23}$$

where it is assumed that $\int q_\alpha(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} < \infty$. This tilted distribution is projected back onto the exponential family by minimizing the KL divergence:

$$\hat{p}_\alpha^{\text{new}}(\boldsymbol{s}) = \arg\min_{\underline{p}\in\mathcal{P}} D\big(q_\alpha(\boldsymbol{s})\big\|\underline{p}(\boldsymbol{s})\big), \tag{7.24}$$

where $\mathcal{P}$ is the set of distributions of the form of $\hat{p}$ in (7.18), i.e., $\underline{p}(\boldsymbol{s}) = \prod_\beta \underline{p}_\beta(\boldsymbol{s}_\beta) = \prod_\beta \exp\left(\underline{\boldsymbol{\gamma}}_\beta^{\mathsf{T}}\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta) - A_\beta(\underline{\boldsymbol{\gamma}}_\beta)\right)$ for some $\{\underline{\boldsymbol{\gamma}}_\beta\}$. Following [200], [201], the solution to (7.24) is given as follows.

Proposition 7.1. *The solution to* (7.24) *is given by* $\hat{p}_\alpha^{\text{new}}(\boldsymbol{s}) = \prod_\beta \hat{p}_{\alpha,\beta}^{\text{new}}(\boldsymbol{s}_\beta)$ *with* $\hat{p}_{\alpha,\beta}^{\text{new}}(\boldsymbol{s}_\beta) = \hat{p}_\beta(\boldsymbol{s}_\beta)$, $\forall\beta \notin \mathfrak{N}_\alpha$, *and* $\hat{p}_{\alpha,\beta}^{\text{new}}(\boldsymbol{s}_\beta) = \exp\left(\underline{\boldsymbol{\gamma}}_\beta^{\mathsf{T}}\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta) - A_\beta(\underline{\boldsymbol{\gamma}}_\beta)\right)$ *with* $\underline{\boldsymbol{\gamma}}_\beta$ *such that* $\mathbb{E}_{\hat{p}_{\alpha,\beta}^{\text{new}}}[\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta)] = \mathbb{E}_{q_\alpha}[\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta)]$, $\forall\beta \in \mathfrak{N}_\alpha$, *whenever the expectation* $\mathbb{E}_{q_\alpha}[\cdot]$ *exists.*

Proof. The proof is provided in Appendix 7.A.3. □

The factor m_α is then updated via

$$\begin{aligned}
m_\alpha^{\text{new}}(\boldsymbol{s}_\alpha) &= \frac{\hat{p}_\alpha^{\text{new}}(\boldsymbol{s})m_\alpha(\boldsymbol{s}_\alpha)}{\hat{p}(\boldsymbol{s})} && (7.25)\\
&= \Big[\prod_{\beta\in\mathfrak{N}_\alpha} m_{\alpha,\beta}(\boldsymbol{s}_\beta)\Big]\frac{\prod_{\beta\in\mathfrak{N}_\alpha}\hat{p}_{\alpha,\beta}^{\text{new}}(\boldsymbol{s}_\beta)}{\prod_{\beta\in\mathfrak{N}_\alpha}\hat{p}_\beta(\boldsymbol{s}_\beta)} && (7.26)\\
&\propto \Big[\prod_{\beta\in\mathfrak{N}_\alpha} m_{\alpha,\beta}(\boldsymbol{s}_\beta)\Big]\frac{\prod_{\beta\in\mathfrak{N}_\alpha}\hat{p}_{\alpha,\beta}^{\text{new}}(\boldsymbol{s}_\beta)}{\prod_{\beta\in\mathfrak{N}_\alpha}\left[m_{\alpha,\beta}(\boldsymbol{s}_\beta)\prod_{\alpha'\in\mathfrak{N}_\beta\setminus\alpha} m_{\alpha',\beta}(\boldsymbol{s}_\beta)\right]} && (7.27)\\
&= \prod_{\beta\in\mathfrak{N}_\alpha} m_{\alpha,\beta}^{\text{new}}(\boldsymbol{s}_\beta), && (7.28)
\end{aligned}$$

with

$$m_{\alpha,\beta}^{\text{new}}(\boldsymbol{s}_\beta) := \frac{\hat{p}_{\alpha,\beta}^{\text{new}}(\boldsymbol{s}_\beta)}{\prod_{\alpha'\in\mathfrak{N}_\beta\setminus\alpha} m_{\alpha',\beta}(\boldsymbol{s}_\beta)}. \tag{7.29}$$

Note that, on the right-hand side of (7.25), all terms dependent on $\{\boldsymbol{s}_\beta\}_{\beta\notin\mathfrak{N}_\alpha}$ cancel, leaving the dependence only on $\{\boldsymbol{s}_\beta\}_{\beta\in\mathfrak{N}_\alpha}$. Thus, the update of m_α only affects the approximate posterior of nodes β in the neighborhood of node α. After that, the process is repeated with the next α.

A message-passing view of Proposition 7.1 can be seen by expanding $q_\alpha(\boldsymbol{s})$ as

$$q_\alpha(\boldsymbol{s}) = \frac{\psi_\alpha(\boldsymbol{s}_\alpha)}{m_\alpha(\boldsymbol{s}_\alpha)}\Big[\prod_{\beta\in\mathfrak{N}_\alpha}\prod_{\alpha'\in\mathfrak{N}_\beta} m_{\alpha',\beta}(\boldsymbol{s}_\beta)\Big]\Big[\prod_{\beta\notin\mathfrak{N}_\alpha}\hat{p}_\beta(\boldsymbol{s}_\beta)\Big] \tag{7.30}$$

$$= \psi_\alpha(\boldsymbol{s}_\alpha)\Big[\prod_{\beta\in\mathfrak{N}_\alpha}\prod_{\alpha'\in\mathfrak{N}_\beta\setminus\alpha} m_{\alpha',\beta}(\boldsymbol{s}_\beta)\Big]\Big[\prod_{\beta\notin\mathfrak{N}_\alpha}\hat{p}_\beta(\boldsymbol{s}_\beta)\Big], \tag{7.31}$$

then, using the natural logarithm for the KL divergence, it follows that

$$\begin{aligned}
&D(q_\alpha(\boldsymbol{s})\|\underline{p}(\boldsymbol{s})) \\
&= \int q_\alpha(\boldsymbol{s}) \ln\frac{q_\alpha(\boldsymbol{s})}{\underline{p}(\boldsymbol{s})}\,\mathrm{d}\boldsymbol{s} && (7.32)\\
&= \int \psi_\alpha(\boldsymbol{s}_\alpha)\Big[\prod_{\beta\in\mathfrak{N}_\alpha}\prod_{\alpha'\in\mathfrak{N}_\beta\setminus\alpha} m_{\alpha',\beta}(\boldsymbol{s}_\beta)\Big]\Big[\prod_{\beta\notin\mathfrak{N}_\alpha}\hat{p}_\beta(\boldsymbol{s}_\beta)\Big] \\
&\quad\times \ln\Bigg(\frac{\psi_\alpha(\boldsymbol{s}_\alpha)\prod_{\beta\in\mathfrak{N}_\alpha}\prod_{\alpha'\in\mathfrak{N}_\beta\setminus\alpha} m_{\alpha',\beta}(\boldsymbol{s}_\beta)}{\prod_{\beta\in\mathfrak{N}_\alpha}\underline{p}_\beta(\boldsymbol{s}_\beta)}\frac{\prod_{\beta\notin\mathfrak{N}_\alpha}\hat{p}_\beta(\boldsymbol{s}_\beta)}{\prod_{\beta\notin\mathfrak{N}_\alpha}\underline{p}_\beta(\boldsymbol{s}_\beta)}\Bigg)\,\mathrm{d}\boldsymbol{s} && (7.33)\\
&= \int \psi_\alpha(\boldsymbol{s}_\alpha)\Big[\prod_{\beta\in\mathfrak{N}_\alpha}\prod_{\alpha'\in\mathfrak{N}_\beta\setminus\alpha} m_{\alpha',\beta}(\boldsymbol{s}_\beta)\Big]\ln\frac{\psi_\alpha(\boldsymbol{s}_\alpha)\prod_{\beta\in\mathfrak{N}_\alpha}\prod_{\alpha'\in\mathfrak{N}_\beta\setminus\alpha} m_{\alpha',\beta}(\boldsymbol{s}_\beta)}{\prod_{\beta\in\mathfrak{N}_\alpha}\underline{p}_\beta(\boldsymbol{s}_\beta)}\,\mathrm{d}\boldsymbol{s}_\alpha \\
&\quad+\sum_{\beta\notin\mathfrak{N}_\alpha}\int \hat{p}_\beta(\boldsymbol{s}_\beta)\ln\frac{\hat{p}_\beta(\boldsymbol{s}_\beta)}{\underline{p}_\beta(\boldsymbol{s}_\beta)}\,\mathrm{d}\boldsymbol{s}_\beta && (7.34)\\
&= \sum_{\beta\in\mathfrak{N}_\alpha}\int q_{\alpha,\beta}(\boldsymbol{s}_\beta)\ln\frac{q_{\alpha,\beta}(\boldsymbol{s}_\beta)}{\underline{p}_\beta(\boldsymbol{s}_\beta)}\,\mathrm{d}\boldsymbol{s}_\beta + \sum_{\beta\notin\mathfrak{N}_\alpha} D(\hat{p}_\beta\|\underline{p}_\beta) + c_0 && (7.35)\\
&= \sum_{\beta\in\mathfrak{N}_\alpha} D(q_{\alpha,\beta}\|\underline{p}_\beta) + \sum_{\beta\notin\mathfrak{N}_\alpha} D(\hat{p}_\beta\|\underline{p}_\beta) + c_0, && (7.36)
\end{aligned}$$

where

$$q_{\alpha,\beta}(\boldsymbol{s}_\beta) := \int \psi_\alpha(\boldsymbol{s}_\alpha)\Big[\prod_{\beta\in\mathfrak{N}_\alpha}\prod_{\alpha'\in\mathfrak{N}_\beta\setminus\alpha} m_{\alpha',\beta}(\boldsymbol{s}_\beta)\Big]\,\mathrm{d}\boldsymbol{s}_{\alpha\setminus\beta} \tag{7.37}$$

and c_0 represents a constant w.r.t. the distribution $\underline{p}$ (which we optimize) whose value is irrelevant and may change at each occurrence. Equation (7.36) says that, for each β in the neighborhood of node α, the optimal $\underline{p}_\beta$ (i.e., $\hat{p}^{\text{new}}_{\alpha,\beta}$) is uniquely identified as the moment match of $q_{\alpha,\beta}$ in the exponential family with sufficient statistics $\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta)$, where $q_{\alpha,\beta}$ is formed by taking the product of the true factor ψ_α and all the messages impinging on that factor, and then integrating out all variables except $\boldsymbol{s}_\beta$. Furthermore, (7.29) says that the new message $m^{\text{new}}_{\alpha,\beta}$ passed from α to $\beta\in\mathfrak{N}_\alpha$ equals $\hat{p}^{\text{new}}_{\alpha,\beta}$ divided by the product of messages $\{m_{\alpha',\beta}\}_{\alpha'\in\mathfrak{N}_\beta\setminus\alpha}$, i.e., previous messages to β from all directions except α. An illustrative example is shown in Fig. 7.1.

7.4 Application of EP to Non-Coherent Detection

In order to apply EP to the non-coherent detection problem described in Section 7.2, we express the transmitted signal as $\mathbf{x}_k = \boldsymbol{x}_k^{(\mathsf{i}_k)}$, where $\mathsf{i}_1,\ldots,\mathsf{i}_K$ are independent random indices.[3] With the assumption that the constellation symbols are transmitted with equal

[3]The application of EP to non-coherent multi-user detection is non-trivial. Many choices can be made to model and partition the unknowns, but may not result in tractable derivation. Our choice is carefully made to enable closed-form message updates.

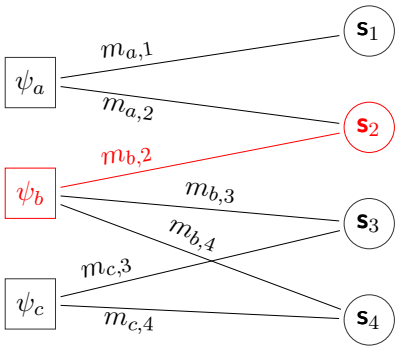


FIGURE 7.1: An example of the factor graph representation of EP for $\alpha \in \{a, b, c\}$ and $\beta \in \{1, 2, 3, 4\}$. For $\alpha = b$ and $\beta = 2$, according to (7.37) and (7.29), $q_{b,2}(\mathbf{s}_2) = \int \psi_b(\mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4) m_{a,2}(\mathbf{s}_2) m_{c,3}(\mathbf{s}_3) m_{c,4}(\mathbf{s}_4)\, \mathrm{d}\mathbf{s}_3\, \mathrm{d}\mathbf{s}_4$ and $m_{b,2}^{\text{new}}(\mathbf{s}_2) = \frac{\hat{p}_{b,2}^{\text{new}}(\mathbf{s}_2)}{m_{a,2}(\mathbf{s}_2)}$, respectively.

probability, i_k are uniformly distributed over $[|\mathcal{X}_k|]$, $k \in [K]$. We rewrite the received signal (7.1) in vector form as

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{w}_k + \mathbf{z}, \tag{7.38}$$

where $\mathbf{y} := \mathrm{vec}(\mathbf{Y}^\mathsf{T})$, $\mathbf{w}_k := (\boldsymbol{x}_k^{(\mathsf{i}_k)} \otimes \boldsymbol{I}_N)\mathbf{h}_k$, and $\mathbf{z} := \mathrm{vec}(\mathbf{Z}^\mathsf{T}) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2 \boldsymbol{I}_{NT})$. The problem of estimating $p_{\mathbf{x}_k|\mathbf{Y}}$ is equivalent to estimating $p_{\mathsf{i}_k|\mathbf{Y}}$ since they admit the same PMF.

With $\mathbf{w} := [\mathbf{w}_1^\mathsf{T}, \ldots, \mathbf{w}_K^\mathsf{T}]^\mathsf{T}$ and $\mathbf{i} := [\mathsf{i}_1, \ldots, \mathsf{i}_K]^\mathsf{T}$, we can write

$$p_{\mathbf{i},\mathbf{w}|\mathbf{y}}(\boldsymbol{i}, \boldsymbol{w}|\boldsymbol{y}) \propto p_{\mathbf{i},\mathbf{w},\mathbf{y}}(\boldsymbol{i}, \boldsymbol{w}, \boldsymbol{y}) = p_{\mathbf{y}|\mathbf{w}}(\boldsymbol{y}|\boldsymbol{w}) p_{\mathbf{w}|\mathbf{i}}(\boldsymbol{w}|\boldsymbol{i}) p_{\mathbf{i}}(\boldsymbol{i}) \tag{7.39}$$

$$= \psi_0(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K) \Big[\prod_{k=1}^{K} \psi_{k1}(\boldsymbol{w}_k, i_k) \Big] \Big[\prod_{k=1}^{K} \psi_{k2}(i_k) \Big], \tag{7.40}$$

corresponding to (7.17), where

$$\psi_0(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K) := p_{\mathbf{y}|\mathbf{w}}(\boldsymbol{y}|\boldsymbol{w}) = \mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{y}; \sum_{k=1}^{K} \boldsymbol{w}_k, \sigma^2 \boldsymbol{I}_{NT}\Big), \tag{7.41}$$

$$\psi_{k1}(\boldsymbol{w}_k, i_k) := p_{\mathbf{w}_k|\mathsf{i}_k}(\boldsymbol{w}_k) = \mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{w}_k; \mathbf{0}, (\boldsymbol{x}_k^{(i_k)} \boldsymbol{x}_k^{(i_k)\mathsf{H}}) \otimes \boldsymbol{R}_k\Big), \tag{7.42}$$

$$\psi_{k2}(i_k) := p_{\mathsf{i}_k}(i_k) = \frac{1}{|\mathcal{X}_k|} \quad \text{for} \quad i_k \in [|\mathcal{X}_k|]. \tag{7.43}$$

In the following, we consider a realization $\boldsymbol{y}$ of $\mathbf{y}$ and use EP to infer the posterior of the indices $\{\mathsf{i}_k\}$ and, as a by-product, the posterior of $\mathbf{w}_k$, $k \in [K]$. To do so, we choose the partition $\mathbf{s} = \{\mathbf{w}_k, \mathsf{i}_k\}_{k=1}^K$ and illustrate the interaction between these variables and the factors $\psi_0, \psi_{k1}, \psi_{k2}$ on the bipartite factor graph in Fig. 7.2. This graph is a tree with a root $\mathbf{y}$ and K leaves $\{\psi_{k2}\}_{k=1}^K$.

We write the EP approximation according to (7.18) as

$$\hat{p}_{\mathbf{s}|\mathbf{y}}(\boldsymbol{s}|\boldsymbol{y}) = \hat{p}_{\mathbf{i},\mathbf{w}|\mathbf{y}}(\boldsymbol{i}, \boldsymbol{w}|\boldsymbol{y}) = \prod_{k=1}^{K} \hat{p}_{\mathbf{w}_k}(\boldsymbol{w}_k) \hat{p}_{\mathsf{i}_k}(i_k), \tag{7.44}$$

where $\hat{p}_{\mathbf{w}_k}(\boldsymbol{w}_k)$ and $\hat{p}_{\mathsf{i}_k}(i_k)$ are implicitly conditioned on $\mathbf{y} = \boldsymbol{y}$ and constrained to be a Gaussian vector distribution and a discrete distribution with support $[|\mathcal{X}|]$ (both belong to

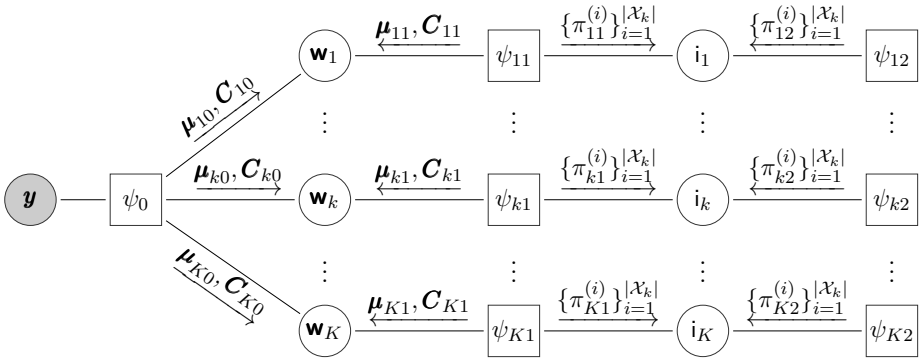


FIGURE 7.2: A factor graph representation of the non-coherent multi-user detection problem. Each factor node represents the conditional PDF of the variable on the left given the variable(s) on the right. The messages are depicted with under-arrows showing their direction from a factor node to a variable node.

the exponential family), respectively. Specifically, they are parameterized as

$$\hat{p}_{\mathbf{w}_k}(\boldsymbol{w}_k) = \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \hat{\boldsymbol{w}}_k, \boldsymbol{\Sigma}_k) \;\text{ such that }\; \boldsymbol{\Sigma}_k \text{ is positive definite}, \tag{7.45}$$

$$\hat{p}_{\mathsf{i}_k}(i_k) = \hat{\pi}_k^{(i_k)} \;\text{ for }\; i_k \in [|\mathcal{X}_k|] \;\text{ such that }\; \textstyle\sum_{i=1}^{|\mathcal{X}_k|} \hat{\pi}_k^{(i)} = 1. \tag{7.46}$$

We also write the EP approximation according to (7.20) as

$$\hat{p}_{\mathbf{s}|\mathbf{y}}(\boldsymbol{s}|\boldsymbol{y}) = \hat{p}_{\mathbf{i},\mathbf{w}|\mathbf{y}}(\boldsymbol{i}, \boldsymbol{w}|\boldsymbol{y}) \propto m_0(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K) \Big[\prod_{k=1}^{K} m_{k1}(\boldsymbol{w}_k, i_k) \Big] \Big[\prod_{k=1}^{K} m_{k2}(i_k) \Big], \tag{7.47}$$

where we define

$$m_0(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K) \propto \prod_{k=1}^{K} \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k0}, \boldsymbol{C}_{k0}), \tag{7.48}$$

$$m_{k1}(\boldsymbol{w}_k, i_k) \propto \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1}) \pi_{k1}^{(i_k)}, \tag{7.49}$$

$$m_{k2}(i_k) = \pi_{k2}^{(i_k)} \;\text{ for }\; i_k \in [|\mathcal{X}_k|]. \tag{7.50}$$

On the factor graph in Fig. 7.2, we can interpret $(\boldsymbol{\mu}_{k0}, \boldsymbol{C}_{k0})$ as the message from factor node ψ_0 to variable node $\mathbf{w}_k$, $(\boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1})$ as the message from factor node ψ_{k1} to variable node $\mathbf{w}_k$, $\{\pi_{k1}^{(\mathsf{i}_k)}\}_{\mathsf{i}_k=1}^{|\mathcal{X}_k|}$ as the message from factor node ψ_{k1} to variable node i_k, and $\{\pi_{k2}^{(\mathsf{i}_k)}\}_{\mathsf{i}_k=1}^{|\mathcal{X}_k|}$ as the message from factor node ψ_{k2} to variable node i_k.

Remark 7.1. *Our choice of Gaussian distribution (within the exponential family) in* (7.45) *is motivated by the fact that when the noise and channel are Gaussian, the symbol posterior takes the form of a Gaussian mixture. It also allows a tractable derivation (using the Gaussian PDF multiplication rule) and closed-form update expressions, as will be shown in the next subsection. If a general (possibly non-Gaussian) channel model is considered, the factor* $\psi_{k1}(\boldsymbol{w}_k, i_k)$ *in* (7.42) *may be different, but the factor graph in Fig. 7.2 remains unchanged.*

7.4.1 The EP Message Updates

In the following, we derive the message updates from each of the factor nodes ψ_0, ψ_{k1}, and ψ_{k2}, $k \in [K]$, to the corresponding variable nodes. To do so, for each $\alpha \in \{k1, k2, 0\}$, we compute the projected density $\hat{p}_\alpha^{\text{new}} = \prod_{k=1}^{K} \hat{p}_{\alpha,\mathbf{w}_k}^{\text{new}}(\boldsymbol{w}_k) \hat{p}_{\alpha,\mathsf{i}_k}^{\text{new}}(i_k)$ according to (7.44) and Proposition 7.1, and then update the factor m_α according to (7.28).

7.4.1.a Message $\{\pi_{k2}^{(i_k)}\}_{i_k=1}^{|\mathcal{X}_k|}$ from Factor Node ψ_{k2} to Variable Node i_k

First, we compute $\hat{p}_{k2,\mathsf{i}_k}^{\text{new}}$ and then the EP message $\left\{\pi_{k2}^{(i_k)}\right\}_{i_k=1}^{|\mathcal{X}_k|}$ from node ψ_{k2} to node i_k. From (7.36) and (7.46), we know that $\hat{p}_{k2,\mathsf{i}_k}^{\text{new}}$ is the discrete distribution with PMF $\left\{\hat{\pi}_{k2}^{(i)}\right\}_{i=1}^{|\mathcal{X}_k|}$ proportional to $\psi_{k2}(i_k)\pi_{k1}^{(i_k)}$, and so

$$\hat{\pi}_{k2}^{(i_k)} = \frac{\psi_{k2}(i_k)\pi_{k1}^{(i_k)}}{\sum_{i=1}^{|\mathcal{X}_k|}\psi_{k2}(i)\pi_{k1}^{(i)}} = \frac{\pi_{k1}^{(i_k)}}{\sum_{i=1}^{|\mathcal{X}_k|}\pi_{k1}^{(i)}} \quad \text{for } i_k \in [|\mathcal{X}_k|], \tag{7.51}$$

since $\psi_{k2}(i_k)$ is constant over these i_k. With $\hat{p}_{k2,\mathsf{i}_k}^{\text{new}}$ computed, (7.28) implies that the message from node ψ_{k2} to node i_k is the PMF proportional to

$$\frac{\hat{p}_{k2,\mathsf{i}_k}^{\text{new}}(i_k)}{\pi_{k1}^{(i_k)}} = \frac{\hat{\pi}_{k2}^{(i_k)}}{\pi_{k1}^{(i_k)}} = \frac{1}{\sum_{i=1}^{|\mathcal{X}_k|}\pi_{k1}^{(i)}} = c_0 \quad \text{for } i_k \in [|\mathcal{X}_k|], \tag{7.52}$$

and thus $\pi_{k2}^{(i_k)} = \frac{1}{|\mathcal{X}_k|}$ for $i_k \in [|\mathcal{X}_k|]$.

7.4.1.b Messages from Factor Node ψ_{k1} to Variable Nodes $\mathbf{w}_k$ and i_k

Next, we compute $\hat{p}_{k1}^{\text{new}} = \prod_{k=1}^{K}\hat{p}_{k1,\mathbf{w}_k}^{\text{new}}(\boldsymbol{w}_k)\hat{p}_{k1,\mathsf{i}_k}^{\text{new}}(i_k)$ and the messages $\{\pi_{k1}^{(i_k)}\}_{i_k=1}^{|\mathcal{X}_k|}$ and $(\boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1})$ from node ψ_{k1} to nodes i_k and $\mathbf{w}_k$, respectively.

Message $\left\{\pi_{k1}^{(i_k)}\right\}_{i_k=1}^{|\mathcal{X}_k|}$ *from Node* ψ_{k1} *to Node* i_k*:*

We first compute $\hat{p}_{k1,\mathsf{i}_k}^{\text{new}}(i_k)$. From (7.36) and (7.46), we know that $\hat{p}_{k1,\mathsf{i}_k}^{\text{new}}(i_k)$ is the discrete distribution with support $[|\mathcal{X}_k|]$ and PMF $\hat{\pi}_{k1}^{(i_k)}$ proportional to

$$\begin{aligned}
&\int \psi_{k1}(\boldsymbol{w}_k, i_k)\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k0}, \boldsymbol{C}_{k0})\pi_{k2}^{(i_k)}\,\mathrm{d}\boldsymbol{w}_k \\
&= \frac{1}{|\mathcal{X}_k|}\int \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{0}, (\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}}) \otimes \boldsymbol{R}_k)\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k0}, \boldsymbol{C}_{k0})\,\mathrm{d}\boldsymbol{w}_k && (7.53)\\
&= \frac{1}{|\mathcal{X}_k|}\int \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \hat{\boldsymbol{w}}_{ki_k}, \boldsymbol{\Sigma}_{ki_k})\mathcal{N}_{\mathbb{C}}(\boldsymbol{0}; \boldsymbol{\mu}_{k0}, (\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}}) \otimes \boldsymbol{R}_k + \boldsymbol{C}_{k0})\,\mathrm{d}\boldsymbol{w}_k && (7.54)\\
&= \frac{1}{|\mathcal{X}_k|}\mathcal{N}_{\mathbb{C}}(\boldsymbol{0}; \boldsymbol{\mu}_{k0}, (\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}}) \otimes \boldsymbol{R}_k + \boldsymbol{C}_{k0}), && (7.55)
\end{aligned}$$

where the second equality follows from the Gaussian PDF multiplication rule in Lemma 7.1 (see Appendix 7.A.2) with

$$\begin{aligned}
\boldsymbol{\Sigma}_{ki} &= \left([(\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}) \otimes \boldsymbol{R}_k]^{-1} + \boldsymbol{C}_{k0}^{-1}\right)^{-1} && (7.56)\\
&= \left[(\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}) \otimes \boldsymbol{R}_k\right]\left((\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}) \otimes \boldsymbol{R}_k + \boldsymbol{C}_{k0}\right)^{-1}\boldsymbol{C}_{k0}, && (7.57)\\
\hat{\boldsymbol{w}}_{ki} &= \boldsymbol{\Sigma}_{ki}\boldsymbol{C}_{k0}^{-1}\boldsymbol{\mu}_{k0} && (7.58)\\
&= \left[(\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}) \otimes \boldsymbol{R}_k\right]\left((\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}) \otimes \boldsymbol{R}_k + \boldsymbol{C}_{k0}\right)^{-1}\boldsymbol{\mu}_{k0}. && (7.59)
\end{aligned}$$

Thus

$$\hat{\pi}_{k1}^{(i_k)} = \frac{\mathcal{N}_{\mathbb{C}}(\boldsymbol{0}; \boldsymbol{\mu}_{k0}, (\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}}) \otimes \boldsymbol{R}_k + \boldsymbol{C}_{k0})}{\sum_{i=1}^{|\mathcal{X}_k|}\mathcal{N}_{\mathbb{C}}(\boldsymbol{0}; \boldsymbol{\mu}_{k0}, (\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}) \otimes \boldsymbol{R}_k + \boldsymbol{C}_{k0})} \quad \text{for } i_k \in [|\mathcal{X}_k|]. \tag{7.60}$$

With $\hat{p}^{\text{new}}_{k1,\mathsf{i}_k}(i_k)$ computed, (7.28) implies that the message $\pi^{(i_k)}_{k1}$ from node ψ_{k1} to node i_k is the PMF proportional to $\frac{\hat{p}^{\text{new}}_{k1,\mathsf{i}_k}(i_k)}{\pi^{(i_k)}_{k2}} = |\mathcal{X}_k|\hat{\pi}^{(i_k)}_{k1}$ for $i_k \in [|\mathcal{X}_k|]$, and thus

$$\pi^{(i_k)}_{k1} = \frac{|\mathcal{X}_k|\hat{\pi}^{(i_k)}_{k1}}{\sum_{i=1}^{|\mathcal{X}_k|}|\mathcal{X}_k|\hat{\pi}^{(i)}_{k1}} = \hat{\pi}^{(i_k)}_{k1} \quad \text{for} \quad i_k \in [|\mathcal{X}_k|]. \tag{7.61}$$

Message $(\boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1})$ from Node ψ_{k1} to Nodes $\mathbf{w}_k$:

We next compute $\hat{p}^{\text{new}}_{k1,\mathbf{w}_k}(\boldsymbol{w}_k)$. From (7.36) and (7.45), we know that $\hat{p}^{\text{new}}_{k1,\mathbf{w}_k}(\boldsymbol{w}_k)$ is the Gaussian distribution with mean $\hat{\boldsymbol{w}}_k$ and covariance $\boldsymbol{\Sigma}_k$ matched to that of the PDF proportional to

$$\begin{aligned}
&\sum_{i_k=1}^{|\mathcal{X}_k|}\psi_{k1}(\boldsymbol{w}_k, i_k)\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k0},\boldsymbol{C}_{k0})\pi^{(i_k)}_{k2} \\
&= \frac{1}{|\mathcal{X}_k|}\sum_{i=1}^{|\mathcal{X}_k|}\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\mathbf{0},(\boldsymbol{x}^{(i)}_k\boldsymbol{x}^{(i)\mathsf{H}}_k)\otimes\boldsymbol{R}_k)\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k0},\boldsymbol{C}_{k0}) &(7.62)\\
&= \frac{1}{|\mathcal{X}_k|}\sum_{i=1}^{|\mathcal{X}_k|}\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\hat{\boldsymbol{w}}_{ki},\boldsymbol{\Sigma}_{ki})\mathcal{N}_{\mathbb{C}}(\mathbf{0};\boldsymbol{\mu}_{k0},(\boldsymbol{x}^{(i)}_k\boldsymbol{x}^{(i)\mathsf{H}}_k)\otimes\boldsymbol{R}_k+\boldsymbol{C}_{k0}) &(7.63)\\
&\propto \sum_{i=1}^{|\mathcal{X}_k|}\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\hat{\boldsymbol{w}}_{ki},\boldsymbol{\Sigma}_{ki})\hat{\pi}^{(i)}_{k1}, &(7.64)
\end{aligned}$$

where the second equality follows from the Gaussian PDF multiplication rule in Lemma 7.1 with $\boldsymbol{\Sigma}_{ki}$ and $\hat{\boldsymbol{w}}_{ki}$ defined in (7.57) and (7.59), respectively. Thus, from (7.61), we have

$$\hat{\boldsymbol{w}}_k = \sum_{i=1}^{|\mathcal{X}_k|}\pi^{(i)}_{k1}\hat{\boldsymbol{w}}_{ki}, \tag{7.65}$$

$$\boldsymbol{\Sigma}_k = \sum_{i=1}^{|\mathcal{X}_k|}\pi^{(i)}_{k1}(\hat{\boldsymbol{w}}_{ki}\hat{\boldsymbol{w}}^{\mathsf{H}}_{ki}+\boldsymbol{\Sigma}_{ki}) - \hat{\boldsymbol{w}}_k\hat{\boldsymbol{w}}^{\mathsf{H}}_k. \tag{7.66}$$

With $\hat{p}^{\text{new}}_{k1,\mathbf{w}_k}(\boldsymbol{w}_k)$ computed, (7.28) implies that the message from node ψ_{k1} to node $\mathbf{w}_k$ is proportional to

$$\frac{\hat{p}^{\text{new}}_{k1,\mathbf{w}_k}(\boldsymbol{w}_k)}{\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k0},\boldsymbol{C}_{k0})} = \frac{\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\hat{\boldsymbol{w}}_k,\boldsymbol{\Sigma}_k)}{\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k0},\boldsymbol{C}_{k0})} \propto \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k1},\boldsymbol{C}_{k1}), \tag{7.67}$$

with

$$\boldsymbol{C}_{k1} = (\boldsymbol{\Sigma}^{-1}_k - \boldsymbol{C}^{-1}_{k0})^{-1}, \tag{7.68}$$

$$\boldsymbol{\mu}_{k1} = \boldsymbol{C}_{k1}(\boldsymbol{\Sigma}^{-1}_k\hat{\boldsymbol{w}}_k - \boldsymbol{C}^{-1}_{k0}\boldsymbol{\mu}_{k0}). \tag{7.69}$$

Equations (7.68) and (7.69) can be verified using

$$\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\hat{\boldsymbol{w}}_k,\boldsymbol{\Sigma}_k) \propto \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k1},\boldsymbol{C}_{k1})\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k0},\boldsymbol{C}_{k0}),$$

which follows from (7.22) and the Gaussian PDF multiplication rule in Lemma 7.1.

7.4.1.c Message $(\boldsymbol{\mu}_{k0}, \boldsymbol{C}_{k0})$ from Factor Node ψ_0 to Variable Node $\mathbf{w}_k$

Finally, we compute $\hat{p}^{\text{new}}_{0,\mathbf{w}_k}$ and the EP message $(\boldsymbol{\mu}_{k0}, \boldsymbol{C}_{k0})$ from node ψ_0 to node $\mathbf{w}_k$ for each $k \in [K]$. From (7.36) and (7.45), we know that $\hat{p}^{\text{new}}_{0,\mathbf{w}_k}$ is the Gaussian distribution with mean $\hat{\boldsymbol{w}}_{k0}$ and covariance $\boldsymbol{\Sigma}_{k0}$ matched to that of the PDF proportional to

$$\begin{aligned}
&\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1}) \int \psi_0(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K) \Big[\prod_{j \neq k} \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_j; \boldsymbol{\mu}_{j1}, \boldsymbol{C}_{j1}) \, \mathrm{d}\boldsymbol{w}_j \Big] \\
&= \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1}) \int \mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{y}; \boldsymbol{w}_k + \sum_{j \neq k} \boldsymbol{w}_j, \sigma^2 \boldsymbol{I}_{NT}\Big) \Big[\prod_{j \neq k} \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_j; \boldsymbol{\mu}_{j1}, \boldsymbol{C}_{j1}) \, \mathrm{d}\boldsymbol{w}_j \Big] \qquad (7.70) \\
&= \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1}) \mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{w}_k; \boldsymbol{y} - \sum_{j \neq k} \boldsymbol{\mu}_{j1}, \sigma^2 \boldsymbol{I}_{NT} + \sum_{j \neq k} \boldsymbol{C}_{j1}\Big), \qquad (7.71)
\end{aligned}$$

where (7.71) follows by applying repeatedly Lemma 7.1. Applying the Gaussian PDF multiplication rule to (7.71), we obtain

$$\boldsymbol{\Sigma}_{k0} = \Big(\boldsymbol{C}_{k1}^{-1} + \Big[\sigma^2 \boldsymbol{I}_{NT} + \sum_{j \neq k} \boldsymbol{C}_{j1}\Big]^{-1}\Big)^{-1}, \qquad (7.72)$$

$$\hat{\boldsymbol{w}}_{k0} = \boldsymbol{\Sigma}_{k0}\Big(\boldsymbol{C}_{k1}^{-1}\boldsymbol{\mu}_{k1} + \Big[\sigma^2 \boldsymbol{I}_{NT} + \sum_{j \neq k} \boldsymbol{C}_{j1}\Big]^{-1}\Big[\boldsymbol{y} - \sum_{j \neq k} \boldsymbol{\mu}_{j1}\Big]\Big). \qquad (7.73)$$

Given $\hat{p}^{\text{new}}_{0,\mathbf{w}_k}(\boldsymbol{w}_k) = \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \hat{\boldsymbol{w}}_{k0}, \boldsymbol{\Sigma}_{k0})$, (7.28) implies that the message from node ψ_0 to node $\mathbf{w}_k$ is proportional to

$$\frac{\hat{p}^{\text{new}}_{0,\mathbf{w}_k}(\boldsymbol{w}_k)}{\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1})} = \frac{\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \hat{\boldsymbol{w}}_{k0}, \boldsymbol{\Sigma}_{k0})}{\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1})} \propto \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k0}, \boldsymbol{C}_{k0}), \qquad (7.74)$$

with $\boldsymbol{C}_{k0} = (\boldsymbol{\Sigma}_{k0}^{-1} - \boldsymbol{C}_{k1}^{-1})^{-1}$ and $\boldsymbol{\mu}_{k0} = \boldsymbol{C}_{k0}(\boldsymbol{\Sigma}_{k0}^{-1}\boldsymbol{w}_{k0} - \boldsymbol{C}_{k1}^{-1}\boldsymbol{\mu}_{k1})$. This is verified using $\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \hat{\boldsymbol{w}}_{k0}, \boldsymbol{\Sigma}_{k0}) \propto \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1})\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k0}, \boldsymbol{C}_{k0})$, which follows from (7.22), and the Gaussian PDF multiplication rule in Lemma 7.1. Plugging in the expressions for $\boldsymbol{\Sigma}_{k0}^{-1}$ and $\hat{\boldsymbol{w}}_{k0}$ from (7.72) and (7.73) yields

$$\boldsymbol{C}_{k0} = \sigma^2 \boldsymbol{I}_{NT} + \sum_{j \neq k} \boldsymbol{C}_{j1}, \qquad (7.75)$$

$$\boldsymbol{\mu}_{k0} = \boldsymbol{y} - \sum_{j \neq k} \boldsymbol{\mu}_{j1}. \qquad (7.76)$$

This concludes the derivation of the EP message updates.

7.4.2 Initialization of the EP Messages

We initialize the EP messages as follows. First, we choose the non-informative initialization $\boldsymbol{C}_{k0}^{-1} = \mathbf{0}$ and $\boldsymbol{\mu}_{k0} = \mathbf{0}$, so that, from (7.60), the initial message from node ψ_{k1} to node i_k coincides with the uniform prior $\pi_{k1}^{(i_k)} = \hat{\pi}_{k1}^{(i_k)} = \frac{1}{|\mathcal{X}_k|}$ for $i_k \in [|\mathcal{X}_k|]$, and, from (7.57) and (7.59), the initial parameters $\boldsymbol{\Sigma}_{ki} = (\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}) \otimes \boldsymbol{R}_k$ and $\boldsymbol{w}_{ki} = \mathbf{0}$, respectively, for $k \in [K]$ and $i \in [|\mathcal{X}_k|]$. This leads to the initial parameters of $\hat{p}_k(\boldsymbol{w}_k)$ from (7.65) and (7.66) as $\hat{\boldsymbol{w}}_k = \mathbf{0}$ and $\boldsymbol{\Sigma}_k = \frac{1}{|\mathcal{X}_k|}\sum_{i=1}^{|\mathcal{X}_k|}(\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}) \otimes \boldsymbol{R}_k$, and the initial message from node ψ_{k1} to node $\mathbf{w}_k$ given in (7.68) and (7.69) as $\boldsymbol{C}_{k1} = \boldsymbol{\Sigma}_k = \frac{1}{|\mathcal{X}_k|}\sum_{i=1}^{|\mathcal{X}_k|}(\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}) \otimes \boldsymbol{R}_k$, and $\boldsymbol{\mu}_{k1} = \hat{\boldsymbol{w}}_k = \mathbf{0}$. Finally, the initial messages from node ψ_0 to node $\mathbf{w}_k$ follows from (7.75) and (7.76) as $\boldsymbol{C}_{k0} = \sigma^2 \boldsymbol{I}_{NT} + \sum_{j \neq k} \frac{1}{|\mathcal{X}_j|}\sum_{i=1}^{|\mathcal{X}_j|}(\boldsymbol{x}_j^{(i)}\boldsymbol{x}_j^{(i)\mathsf{H}}) \otimes \boldsymbol{R}_k$, and $\boldsymbol{\mu}_{k0} = \boldsymbol{y}$.

7.4.3 The Algorithm

We summarize the proposed EP scheme for probabilistic non-coherent detection in Algorithm 1. In the end, according to (7.22) and (7.46), the estimated PMF $\hat{p}_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{x}_k^{(i_k)}|\boldsymbol{Y})$ is given by $\hat{p}_k(i_k) = \hat{\pi}_k^{(i_k)} \propto \pi_{k1}^{(i_k)}\pi_{k2}^{(i_k)}$, that is $\hat{p}_k(i_k) = \pi_{k1}^{(i_k)}$ since $\pi_{k2}^{(i_k)}$ is constant. The algorithm goes through the branches of the tree graph in Fig. 7.2 in a round-robin manner. In each branch, the factor nodes are visited from leaf to root. We note that other message passing schedules can be implemented.

Algorithm 1: EP for probabilistic non-coherent detection

Input: the observation $\boldsymbol{Y}$; the constellations $\mathcal{X}_1, \ldots, \mathcal{X}_K$;

set the maximal number of iterations $t_{\max}$;

initialize the messages $\{\pi_{k1}^{(i_k)}\}_{i_k=1}^{|\mathcal{X}_k|}, \boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1}, \boldsymbol{\mu}_{k0}, \boldsymbol{C}_{k0}$, for $k \in [K]$;

$t \longleftarrow 0$;

repeat

  $t \longleftarrow t + 1$;

  for $k \leftarrow 1$ **to** K **do**

    update $\{\pi_{k1}^{(i_k)}\}_{i_k=1}^{|\mathcal{X}_k|}$ according to (7.61) and (7.60) ;

    compute $\{\hat{\boldsymbol{w}}_{ki}\}_{i=1}^{|\mathcal{X}_k|}$ and $\{\boldsymbol{\Sigma}_{ki}\}_{i=1}^{|\mathcal{X}_k|}$ according to (7.59) and (7.57), respectively ;

    compute $\hat{\boldsymbol{w}}_k$ and $\boldsymbol{\Sigma}_k$ according to (7.65) and (7.66), respectively ;

    update $\boldsymbol{\mu}_{k1}$ and $\boldsymbol{C}_{k1}$ according to (7.69) and (7.68), respectively ;

    update $\{\boldsymbol{\mu}_{j0}\}_{j\neq k}$ and $\{\boldsymbol{C}_{j0}\}_{j\neq k}$ according to (7.76) and (7.75), respectively ;

  end

until convergence or $t = t_{\max}$;

return the PMF $\left\{\pi_{k1}^{(i_k)}\right\}_{i_k=1}^{|\mathcal{X}_k|}$ of $\hat{p}_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{x}_k^{(i_k)}|\boldsymbol{Y})$ for $k \in [K]$;

In the EP algorithm, the dominant operation is the update of $\pi_{k1}^{(i_k)}$, $\boldsymbol{\Sigma}_{ki}$, and $\hat{\boldsymbol{w}}_{ki}$, which involves the inverse of the $NT \times NT$ matrix $\left(\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}}\right) \otimes \boldsymbol{R}_k + \boldsymbol{C}_{k0}$ (with complexity $O(K^6)$) *for all* $k \in [K]$ and $i_k \in [|\mathcal{X}_k|]$. The complexity of computing $\hat{\boldsymbol{w}}_k$, $\boldsymbol{\Sigma}_k$, $\boldsymbol{\mu}_{k1}$, $\boldsymbol{C}_{k1}$, $\{\boldsymbol{\mu}_{j0}\}_{j\neq k}$, and $\{\boldsymbol{C}_{j0}\}_{j\neq k}$ are all of lower order. Therefore, the complexity per iteration is given by $O(K^7 2^B)$. In order to reduce this complexity, we derive two simplifications of the EP scheme in the next section.

7.5 Simplifications of the EP Detector

In this section, we attempt to simplify EP by avoiding the inverse of $NT \times NT$ matrices.

7.5.1 EP with Approximate Kronecker Products (EPAK)

We observe that if $\boldsymbol{C}_{k0}$ could be expressed as a Kronecker product $\bar{\boldsymbol{C}}_{k0} \otimes \boldsymbol{R}_k$ with $\bar{\boldsymbol{C}}_{k0} \in \mathbb{C}^{T\times T}$, we could rewrite $\pi_{k1}^{(i_k)}$ in (7.60) as

$$\pi_{k1}^{(i_k)} = \frac{\mathcal{N}_{\mathbb{C}}(\boldsymbol{0}; \boldsymbol{\mu}_{k0}, (\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}} + \bar{\boldsymbol{C}}_{k0}) \otimes \boldsymbol{R}_k)}{\sum_{i=1}^{|\mathcal{X}_k|} \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}; \boldsymbol{\mu}_{k0}, (\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}} + \bar{\boldsymbol{C}}_{k0}) \otimes \boldsymbol{R}_k)}. \tag{7.77}$$

Let $\boldsymbol{M}_{k0} \in \mathbb{C}^{T\times N}$ such that $\boldsymbol{\mu}_{k0} = \mathrm{vec}(\boldsymbol{M}_{k0}^{\mathsf{T}})$, (7.77) could be computed efficiently using

$$\mathcal{N}_{\mathbb{C}}\Big(\mathbf{0}; \boldsymbol{\mu}_{k0}, (\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}} + \bar{\boldsymbol{C}}_{k0}) \otimes \boldsymbol{R}_k\Big) \propto (1 + \boldsymbol{x}_k^{(i_k)\mathsf{H}}\bar{\boldsymbol{C}}_{k0}^{-1}\boldsymbol{x}_k^{(i_k)})^{-N} \exp\left(\frac{\mathrm{tr}\{\bar{\boldsymbol{C}}_{k0}^{-1}\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}}\boldsymbol{M}_{k0}(\boldsymbol{R}_k^{-1})^{\mathsf{T}}\boldsymbol{M}_{k0}^{\mathsf{H}}\}}{1 + \boldsymbol{x}_k^{(i_k)\mathsf{H}}\bar{\boldsymbol{C}}_{k0}^{-1}\boldsymbol{x}_k^{(i_k)}}\right) \tag{7.78}$$

since only the $T \times T$ matrix $\bar{\boldsymbol{C}}_{k0}$ needs to be inverted (the inverse of $\boldsymbol{R}_k$ can be precomputed and stored). In general, $\boldsymbol{C}_{k0}$ does not have a Kronecker structure. Thus we propose to fit $\boldsymbol{C}_{k0}$ to the form of a Kronecker product by solving the least squares problem

$$\min_{\bar{\boldsymbol{C}}_{k0}\in\mathbb{C}^{T\times T}} \|\boldsymbol{C}_{k0} - \bar{\boldsymbol{C}}_{k0} \otimes \boldsymbol{R}\|_F^2 \tag{7.79}$$

as formulated in [204, Section 4]. Let $\boldsymbol{C}_{k0}\{i,j\}$ be the $N \times N$ sub-matrix containing the elements in rows from $(i-1)N+1$ to iN and columns from $(j-1)N+1$ to jN of $\boldsymbol{C}_{k0}$. Let $\bar{c}_{ij}$ be the element in row i and column j of $\bar{\boldsymbol{C}}_{k0}$. It follows that

$$\begin{aligned}
&\|\boldsymbol{C}_{k0} - \bar{\boldsymbol{C}}_{k0} \otimes \boldsymbol{R}_k\|_F^2 \\
&= \sum_{i=1}^{T}\sum_{j=1}^{T} \|\boldsymbol{C}_{k0}\{i,j\} - \bar{c}_{ij}\boldsymbol{R}_k\|_F^2 &(7.80)\\
&= \sum_{i=1}^{T}\sum_{j=1}^{T} \|\boldsymbol{C}_{k0}\{i,j\}\|_F^2 - \bar{c}_{ij}\mathrm{tr}(\boldsymbol{C}_{k0}\{i,j\}^{\mathsf{H}}\boldsymbol{R}_k) - \bar{c}_{ij}^{*}\mathrm{tr}(\boldsymbol{R}_k\boldsymbol{C}_{k0}\{i,j\}) + |\bar{c}_{ij}|^2\mathrm{tr}\Big(\boldsymbol{R}_k^2\Big). &(7.81)
\end{aligned}$$

Observe that $\|\boldsymbol{C}_{k0} - \bar{\boldsymbol{C}}_{k0} \otimes \boldsymbol{R}_k\|_F^2$ is the sum of convex quadratic functions of $\bar{c}_{ij}$. Setting the partials $\frac{\partial\|\boldsymbol{C}_{k0}-\bar{\boldsymbol{C}}_{k0}\otimes\boldsymbol{R}_k\|_F^2}{\partial\bar{c}_{ij}}$ to zeros, the optimal $\bar{\boldsymbol{C}}_{k0}$ is given by

$$\bar{c}_{ij} = \frac{\mathrm{tr}(\boldsymbol{C}_{k0}\{i,j\}\boldsymbol{R}_k)}{\mathrm{tr}\Big(\boldsymbol{R}_k^2\Big)}. \tag{7.82}$$

With the approximation $\boldsymbol{C}_{k0} \approx \bar{\boldsymbol{C}}_{k0} \otimes \boldsymbol{R}_k$, we can approximate $\pi_{k1}^{(i_k)}$ by the right-hand side of (7.77). Also, it follows from (7.57) and (7.59) that

$$\boldsymbol{\Sigma}_{ki} \approx [(\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}})(\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}} + \bar{\boldsymbol{C}}_{k0})^{-1}\bar{\boldsymbol{C}}_{k0}] \otimes \boldsymbol{R}_k, \tag{7.83}$$

$$\hat{\boldsymbol{w}}_{ki} \approx \mathrm{vec}\big([\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}}(\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}} + \bar{\boldsymbol{C}}_{k0})^{-1}\boldsymbol{M}_{k0}]^{\mathsf{T}}\big). \tag{7.84}$$

To compute $\boldsymbol{C}_{k1}$ and $\boldsymbol{\mu}_{k1}$ in (7.69) and (7.68), the inversion of $\boldsymbol{C}_{k0}$ can be simplified as $\boldsymbol{C}_{k0}^{-1} \approx \bar{\boldsymbol{C}}_{k0}^{-1} \otimes \boldsymbol{R}_k^{-1}$, but the inverse of $NT \times NT$ matrices involving $\boldsymbol{\Sigma}_k$ is still required.

To keep an accurate message update at early iterations[4], let us fix a threshold $t_0 \in [t_{\max}]$ and modify Algorithm 1 as follows. At iteration t, if $t \le t_0$, the messages are updated as in lines 7–11; if $t > t_0$, in line 7, (7.60) is replaced by (7.77) for the update of $\pi_{k1}^{(i_k)}$, and in line 8, (7.59) and (7.57) are replaced by (7.84) and (7.83) for the update of $\boldsymbol{\Sigma}_{ki}$ and $\hat{\boldsymbol{w}}_{ki}$, respectively. We refer to this scheme as expectation propagation with approximate Kronecker (EPAK). It coincides with EP if $t_0 = t_{\max}$. At iteration $t > t_0$, the dominant operations in EPAK are the inverse of $\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}} + \bar{\boldsymbol{C}}_{k0}$ (with complexity $O(K^3)$) in (7.83) and (7.84) for each $k \in [K]$ and $i \in [|\mathcal{X}_k|]$, and the inverse of $NT \times NT$ matrices (with complexity $O(K^6)$) to compute $\boldsymbol{C}_{k1}$ and $\boldsymbol{\mu}_{k1}$ for each $k \in [K]$. Thus the complexity at iteration t of EPAK is $O(K^7 2^B)$ if $t \le t_0$ and $O(K^4 2^B + K^7)$ if $t > t_0$.

[4]In the uncorrelated fading case, i.e., $\boldsymbol{R}_k = \boldsymbol{I}_N$, the approximation of $\boldsymbol{C}_{k0}$ with Kronecker products becomes more accurate when $\hat{\pi}_{k1}$ is closer to a Kronecker-delta distribution, i.e., we have high confidence in one of the symbols. This is likely the case at high SNR after some EP iterations. At early iterations, however, the approximation $\boldsymbol{C}_{k0} \approx \bar{\boldsymbol{C}}_{k0} \otimes \boldsymbol{R}$ can be inaccurate.

7.5.2 Minimum Mean Square Error—Successive Interference Approximation (MMSE-SIA)

Another method to simplify EP is as follows. In the EP scheme, as in (7.64) and (7.67), the message $\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k1},\boldsymbol{C}_{k1})$ from node ψ_{k1} to node $\mathbf{w}_k$ is derived by first projecting $\hat{p}^{\text{new}}_{k1,\mathbf{w}_k}(\boldsymbol{w}_k) \propto \sum_{i=1}^{|\mathcal{X}_k|}\pi_{k1}^{(i)}\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\hat{\boldsymbol{w}}_{ki},\boldsymbol{\Sigma}_{ki})$ onto the Gaussian family, then dividing the projected Gaussian by $\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k0},\boldsymbol{C}_{k0})$. If we skip the projection of $\hat{p}^{\text{new}}_{k1}(\boldsymbol{w}_k)$ onto the Gaussian family, i.e., we derive $\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k1},\boldsymbol{C}_{k1})$ by dividing directly $\hat{p}^{\text{new}}_{k1,\mathbf{w}_k}(\boldsymbol{w}_k)$ to $\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k0},\boldsymbol{C}_{k0})$, then the mean $\boldsymbol{\mu}_{k1}$ and covariance matrix $\boldsymbol{C}_{k1}$ are matched to that of the PDF proportional to

$$\frac{\hat{p}^{\text{new}}_{k1,\mathbf{w}_k}(\boldsymbol{w}_k)}{\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k0},\boldsymbol{C}_{k0})} = \sum_{i=1}^{|\mathcal{X}_k|}\pi_{k1}^{(i)}\frac{\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\hat{\boldsymbol{w}}_{ki},\boldsymbol{\Sigma}_{ki})}{\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k0},\boldsymbol{C}_{k0})} \tag{7.85}$$

$$\propto \sum_{i=1}^{|\mathcal{X}_k|}\pi_{k1}^{(i)}\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{0},(\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}})\otimes\boldsymbol{R}_k) \tag{7.86}$$

$$= \mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{0},\boldsymbol{Q}_k\otimes\boldsymbol{R}_k). \tag{7.87}$$

where $\boldsymbol{Q}_k := \sum_{i=1}^{|\mathcal{X}_k|}\pi_{k1}^{(i)}\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}$. (7.86) can be verified using

$$\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\hat{\boldsymbol{w}}_{ki},\boldsymbol{\Sigma}_{ki}) \propto \mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{w}_k;\boldsymbol{0},\Big(\boldsymbol{x}_k^{(i)}\boldsymbol{x}_k^{(i)\mathsf{H}}\Big)\otimes\boldsymbol{R}_k\Big)\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k;\boldsymbol{\mu}_{k0},\boldsymbol{C}_{k0}),$$

which follows from the Gaussian PDF multiplication rule with $\hat{\boldsymbol{w}}_{ki}$ and $\boldsymbol{\Sigma}_{ki}$ given in (7.59) and (7.57), respectively. It follows that $\boldsymbol{\mu}_{k1}=\boldsymbol{0}$ and $\boldsymbol{C}_{k1}=\boldsymbol{Q}_k\otimes\boldsymbol{R}_k$. As a consequence (see (7.76) and (7.75)), $\boldsymbol{\mu}_{k0}=\boldsymbol{y}$ and $\boldsymbol{C}_{k0}=\sigma^2\boldsymbol{I}_{NT}+\sum_{l\neq k}\boldsymbol{Q}_l\otimes\boldsymbol{R}_k$.

This scheme can be alternatively interpreted as follows. We expand $\mathbf{y}$ in (7.38) as

$$\mathbf{y} = (\mathbf{x}_k\otimes\boldsymbol{I}_N)\mathbf{h}_k + \sum_{l\neq k}(\mathbf{x}_l\otimes\boldsymbol{I}_N)\mathbf{h}_l + \mathbf{z}. \tag{7.88}$$

The second term $\mathbf{t}_k := \sum_{l\neq k}(\mathbf{x}_l\otimes\boldsymbol{I}_N)\mathbf{h}_l$ is the interference from other users while decoding the signal of user k. Since the signals $\mathbf{x}_l$ are independent of the channels $\mathbf{h}_l$ and the channels $\mathbf{h}_l$ have zero mean, we have that $\mathbb{E}[\mathbf{t}_k]=\boldsymbol{0}$. The covariance matrix of $\mathbf{t}_k$ is $\mathbb{E}[\mathbf{t}_k\mathbf{t}_k^{\mathsf{H}}] = \sum_{l\neq k}\mathbb{E}[\mathbf{x}_l\mathbf{x}_l^{\mathsf{H}}]\otimes\boldsymbol{R}_k = \sum_{l\neq k}\boldsymbol{Q}_l\otimes\boldsymbol{R}_k$. If we treat the interference term $\mathbf{t}_k$ as a Gaussian vector with the same mean and covariance matrix[5], then $\mathbf{t}_k+\mathbf{z}\sim\mathcal{N}_{\mathbb{C}}(\boldsymbol{0},\sum_{l\neq k}\boldsymbol{Q}_l\otimes\boldsymbol{R}_k+\sigma^2\boldsymbol{I}_{NT})$. The single-user likelihood under this approximation is

$$\hat{p}_{\mathbf{y}|\mathbf{x}_k}(\boldsymbol{y}|\boldsymbol{x}_k) = \mathcal{N}_{\mathbb{C}}\Big(\boldsymbol{y};\boldsymbol{0},\boldsymbol{x}_k\boldsymbol{x}_k^{\mathsf{H}}\otimes\boldsymbol{R}_k + \sum_{l\neq k}\boldsymbol{Q}_l\otimes\boldsymbol{R}_l + \sigma^2\boldsymbol{I}_{NT}\Big). \tag{7.89}$$

With this and Lemma 7.1, the update of the approximate posterior $\hat{p}_{\mathbf{x}_k|\mathbf{y}}(\boldsymbol{x}_k|\boldsymbol{y}) \propto \hat{p}_{\mathbf{y}|\mathbf{x}_k}(\boldsymbol{y}|\boldsymbol{x}_k)$ coincides with (7.60) for $\boldsymbol{\mu}_{k0}=\boldsymbol{y}$ and $\boldsymbol{C}_{k0}=\sigma^2\boldsymbol{I}_{NT}+\sum_{l\neq k}\boldsymbol{Q}_l\otimes\boldsymbol{R}_k$. The matrix $\boldsymbol{Q}_k$ is then recalculated with the updated value of $\hat{p}_{\mathbf{x}_k|\mathbf{y}}(\boldsymbol{x}_k^{(i_k)}|\boldsymbol{y})$, $i_k\in[|\mathcal{X}_k|]$. The matrices $\boldsymbol{C}_{l0}$ are updated accordingly, and then used to update $\hat{p}_{\mathbf{x}_l|\mathbf{y}}(\boldsymbol{x}_l^{(i_l)}|\boldsymbol{y})$, $i_l\in[|\mathcal{X}_l|]$, $l\neq k$.

In short, the derived simplification of the EP scheme above iteratively MMSE-estimates the signal $\mathbf{w}_k$ of one user at a time while treating the interference as Gaussian. At each iteration, the Gaussian approximation of the interference for each user is successively improved using the estimates of the signals of other users. We refer to this scheme as MMSE-SIA and summarize it in Algorithm 2. In particular, as for the EP scheme, we can start with the non-informative initialization $\hat{p}_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{x}|\boldsymbol{Y}) = \frac{1}{|\mathcal{X}_k|}\mathbb{1}\{\boldsymbol{x}\in\mathcal{X}_k\}$.

[5]Another choice is to treat each $\mathbf{x}_l$, $l\neq k$, as a Gaussian. With this choice, however, the interference term $\mathbf{t}_k$ is a product of Gaussians which makes the approximate single-user likelihood difficult to evaluate.

Algorithm 2: MMSE-SIA for probabilistic non-coherent detection

Input: the observation $\boldsymbol{Y}$; the constellations $\mathcal{X}_1, \ldots, \mathcal{X}_K$;

set the maximal number of iterations $t_{\max}$;

initialize the posteriors $\hat{p}_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{x}_k|\boldsymbol{Y})$ for $\boldsymbol{x}_k \in \mathcal{X}_k$, and $\boldsymbol{Q}_k = \mathbb{E}_{\hat{p}_{\mathbf{x}_k|\mathbf{Y}}}[\mathbf{x}_k\mathbf{x}_k^{\mathsf{H}}]$ for $k \in [K]$;

$t \longleftarrow 0$;

repeat

  $t \longleftarrow t+1$;

  for $k \leftarrow 1$ **to** K **do**

    compute $\boldsymbol{C}_{k0} = \sigma^2\boldsymbol{I}_{NT} + \sum_{l\neq k}\boldsymbol{Q}_l \otimes \boldsymbol{R}_k$;

    update $\hat{p}_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{x}_k|\boldsymbol{Y})$, $\boldsymbol{x}_k \in \mathcal{X}_k$, according to (7.60) with $\boldsymbol{\mu}_{k0} = \boldsymbol{y}$ and $\boldsymbol{C}_{k0}$ computed ;

    update $\boldsymbol{Q}_k = \mathbb{E}_{\hat{p}_{\mathbf{x}_k|\mathbf{Y}}}[\mathbf{x}_k\mathbf{x}_k^{\mathsf{H}}]$;

  end

until convergence or $t = t_{\max}$;

return $\hat{p}_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{x}_k|\boldsymbol{Y})$ for $\boldsymbol{x}_k \in \mathcal{X}_k$, $k \in [K]$;

The complexity order of Algorithm 2 is the same as EP due to the $NT \times NT$ matrix inversion in (7.60). However, MMSE-SIA still has a complexity advantage over EP since no other matrix inversion is required, and there is no need to compute $\{\hat{\boldsymbol{w}}_{ki}\}$, $\{\boldsymbol{\Sigma}_{ki}\}$, $\hat{\boldsymbol{w}}_k$, $\boldsymbol{\Sigma}_k$, or update $\boldsymbol{\mu}_{k1}$. If the channel is uncorrelated ($\boldsymbol{R}_k = \boldsymbol{I}_N$), the complexity order of MMSE-SIA can be reduced. In this case, $\boldsymbol{C}_{k0}$ is the Kronecker product $\bar{\boldsymbol{C}}_{k0} \otimes \boldsymbol{I}_N$ with $\bar{\boldsymbol{C}}_{k0} := \sum_{l=1,l\neq k}^{K}\boldsymbol{Q}_l + \sigma^2\boldsymbol{I}_T$, and thus in (7.60),

$$\begin{aligned}
&\mathcal{N}_{\mathbb{C}}(\mathbf{0}; \boldsymbol{\mu}_{k0}, (\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}}) \otimes \boldsymbol{R}_k + \boldsymbol{C}_{k0}) \\
&= \mathcal{N}_{\mathbb{C}}\Big(\mathbf{0}; \mathbf{y}, (\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}} + \bar{\boldsymbol{C}}_{k0}) \otimes \boldsymbol{I}_N\Big) \qquad (7.90) \\
&\propto \big(1 + \boldsymbol{x}_k^{(i_k)\mathsf{H}}\bar{\boldsymbol{C}}_{k0}^{-1}\boldsymbol{x}_k^{(i_k)}\big)^{-N} \exp\left(\frac{\|\mathbf{Y}^{\mathsf{H}}\bar{\boldsymbol{C}}_{k0}^{-1}\boldsymbol{x}_k^{(i_k)}\|^2}{1 + \boldsymbol{x}_k^{(i_k)\mathsf{H}}\bar{\boldsymbol{C}}_{k0}^{-1}\boldsymbol{x}_k^{(i_k)}}\right). \qquad (7.91)
\end{aligned}$$

Then, only the inverse of $\bar{\boldsymbol{C}}_{k0}$ is computed, which requires $O(K^3)$ operations. Given $\bar{\boldsymbol{C}}_{k0}^{-1}$, the complexity of computing the right-hand side of (7.91) is then $O(K^2)$ for each $i_k \in [|\mathcal{X}_k|]$. Therefore, the complexity of computing $\hat{p}_{\mathbf{x}_k|\mathbf{Y}}(\boldsymbol{x}_k|\boldsymbol{Y})$ is $O(K^3 + K^2 2^B)$ for $k \in [K]$. Finally, the complexity per iteration of the MMSE-SIA algorithm for uncorrelated fading is given by $O(K^4 + K^3 2^B)$.

7.6 Implementation Aspects

7.6.1 Complexity

We summarize the computational complexity of the considered schemes in Table 7.1.

7.6.2 Stabilization

We discuss some possible numerical problems in the EP algorithm and our solutions.

7.6.2.a Singularity of $\boldsymbol{\Sigma}_k$

First, in (7.66), since the $NT \times NT$ matrix $\boldsymbol{\Sigma}_k$ is the weighted sum of the terms of rank less than NT, it can be close to singular if at a certain iteration, only few of the weights $\pi_{k1}^{(i)}$

TABLE 7.1: Complexity order of different non-coherent multi-user detectors with $T = O(K), N = O(K)$, and $|\mathcal{X}_k| = O(2^B)$, $k \in [K]$

Detector	**Complexity order**	
	Correlated fading	Uncorrelated fading $\boldsymbol{R}_k = \boldsymbol{I}_N, \forall k$
Optimal (exact marginalization)	$O(K^6 2^{BK})$	$O(K^3 2^{BK})$
EP	$O(K^7 2^B t_{\max})$	
EPAK	$O\big(K^7 2^B t_0 + (K^4 2^B + K^7)(t_{\max} - t_0)\big)$	
MMSE-SIA	$O(K^7 2^B t_{\max})$	$O(K^4 t_{\max} + K^3 2^B t_{\max})$

$t_{\max}$ denotes the number of iterations. $t_0 \in [t_{\max}]$.

are sufficiently larger than zero. The singularity of $\boldsymbol{\Sigma}_k$ can also arise from the constellation structure. For example, the constellations proposed in [125] are precoded versions of a constellation in $G(\mathbb{C}^{T-K+1}, 1)$ and the maximal rank of $\boldsymbol{\Sigma}_k$ is $N(T-K+1) \leq NT$. To avoid the inverse of $\boldsymbol{\Sigma}_k$, we express $\boldsymbol{C}_{k1}$ in (7.68) and $\boldsymbol{\mu}_{k1}$ in (7.69) respectively as

$$\boldsymbol{C}_{k1} = -\boldsymbol{C}_{k0}(\boldsymbol{\Sigma}_k - \boldsymbol{C}_{k0})^{-1}\boldsymbol{\Sigma}_k, \tag{7.92}$$

$$\boldsymbol{\mu}_{k1} = \boldsymbol{C}_{k0}(\boldsymbol{\Sigma}_k - \boldsymbol{C}_{k0})^{-1}\Big(\boldsymbol{\Sigma}_k - \sum_{i=1}^{|\mathcal{X}_k|} \pi_{k1}^{(i)} \boldsymbol{\Sigma}_{ki}\Big)\boldsymbol{C}_{k0}^{-1}\boldsymbol{\mu}_{k0}. \tag{7.93}$$

7.6.2.b "Negative variance"

Another problem is that $\boldsymbol{C}_{k1}$ is not guaranteed to be positive definite even if both $\boldsymbol{C}_{k0}$ and $\boldsymbol{\Sigma}_k$ are. When $\boldsymbol{C}_{k1}$ is not positive definite, from (7.75), $\boldsymbol{C}_{k0}$ can have negative eigenvalues, which, through (7.60), can make $\hat{\pi}_{k1}^{(i_k)}$ become close to a Kronecker-delta distribution (even at low SNR) where the position of the mode can be arbitrary, and the algorithm may diverge. Note that this "negative variance" problem is common in EP (see, e.g., [200, Section 3.2.1], [205, Section 5.3]). There has been no generally accepted solution and one normally resorts to various heuristics adapted to each problem. In our problem, to control the eigenvalues of $\boldsymbol{C}_{k1}$, we modify (7.92) by first computing the eigendecomposition $-\boldsymbol{C}_{k0}(\boldsymbol{\Sigma}_k - \boldsymbol{C}_{k0})^{-1}\boldsymbol{\Sigma}_k = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}$, then computing $\boldsymbol{C}_{k1}$ as $\boldsymbol{C}_{k1} = \boldsymbol{V}|\boldsymbol{\Lambda}|\boldsymbol{V}^{-1}$, where $|\boldsymbol{\Lambda}|$ is the element-wise absolute value of $\boldsymbol{\Lambda}$. This manipulation of replacing the variance parameters by their absolute values was also used in [206].

7.6.2.c Overconfidence at early iterations

Finally, due to the nature of the message passing between continuous and discrete distribution, it can happen that all the mass of the PMF $\hat{\pi}_{k1}^{(i_k)}$ is concentrated on a small region of a potentially large constellation $\mathcal{X}_k$. For example, if $\pi_{k1}^{(i_k)}$ is close to a Kronecker-delta distribution with a single mode at i_0, then (7.59) and (7.57) implies that $\boldsymbol{\Sigma}_k$ is approximately $\boldsymbol{\Sigma}_{ki_0}$, and then from (7.68), $\boldsymbol{C}_{k1} \approx (\boldsymbol{x}_k^{(i_0)}\boldsymbol{x}_k^{(i_0)\mathsf{H}}) \otimes \boldsymbol{R}_k$. In this case, almost absolute certainty is placed on the symbol $\boldsymbol{x}_k^{(i_0)}$, and the algorithm will not be able significantly update its belief in the subsequent iterations. This can be problematic when the mode of $\pi_{k1}^{(i_k)}$ is placed on the wrong symbol at early iterations. To smooth the updates, we apply damping on the update of the parameters of the continuous distributions $\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k1}, \boldsymbol{C}_{k1})$ and $\mathcal{N}_{\mathbb{C}}(\boldsymbol{w}_k; \boldsymbol{\mu}_{k0}, \boldsymbol{C}_{k0})$. That is, with a damping factor $\eta \in [0; 1]$, at iteration t and for each

user k, we update

$$\boldsymbol{C}_{k1}(t) = \eta \boldsymbol{V}(t)|\boldsymbol{\Lambda}(t)|\boldsymbol{V}^{-1}(t) + (1-\eta)\boldsymbol{C}_{k1}(t-1), \tag{7.94}$$

$$\begin{aligned}\boldsymbol{\mu}_{k1}(t) = {} & \eta \boldsymbol{C}_{k0}(t-1)\left(\boldsymbol{\Sigma}_k(t) - \boldsymbol{C}_{k0}(t-1)\right)^{-1} \\ & \times \left(\boldsymbol{\Sigma}_k(t) - \sum_{i=1}^{|\mathcal{X}_k|} \pi_{k1}^{(i)}(t)\boldsymbol{\Sigma}_{ki}(t)\right)\boldsymbol{C}_{k0}^{-1}(t-1)\boldsymbol{\mu}_{k0}(t-1) \\ & + (1-\eta)\boldsymbol{\mu}_{k1}(t-1),\end{aligned} \tag{7.95}$$

$$\boldsymbol{C}_{l0}(t) = \eta\Big(\sigma^2 \boldsymbol{I}_{NT} + \sum_{j\neq l} \boldsymbol{C}_{j1}(t)\Big) + (1-\eta)\boldsymbol{C}_{l0}(t-1), \quad \forall l \neq k, \tag{7.96}$$

$$\boldsymbol{\mu}_{l0}(t) = \eta\Big(\boldsymbol{y} - \sum_{j\neq l} \boldsymbol{\mu}_{j1}(t)\Big) + (1-\eta)\boldsymbol{\mu}_{l0}(t-1), \quad \forall l \neq k. \tag{7.97}$$

In short, we stabilize the EP message updates by replacing (7.94), (7.95), (7.96), and (7.97) for (7.68), (7.69), (7.75), and (7.76), respectively. This technique also applies to EPAK. For MMSE-SIA, we damp the update of $\bar{\boldsymbol{C}}_{k0}$ and $\boldsymbol{Q}_k$ in a similar manner as $\bar{\boldsymbol{C}}_{k0}(t) = \eta(\sum_{l\neq k}\boldsymbol{Q}_l(t-1)+\sigma^2\boldsymbol{I}_T)+(1-\eta)\bar{\boldsymbol{C}}_{k0}(t-1)$ and $\boldsymbol{Q}_k(t) = \eta\sum_{i_k=1}^{|\mathcal{X}_k|}\pi_{k1}^{(i_k)}(t)\boldsymbol{x}_k^{(i_k)}\boldsymbol{x}_k^{(i_k)\mathsf{H}} + (1-\eta)\boldsymbol{Q}_k(t-1)$. Note that damping does not change the complexity order of these schemes. The approaches described in this subsection were implemented for the numerical results in the next section.

7.7 Greedy Detectors for the Precoding-Based Constellation

In this section, we present a greedy approach for detection of signals transmitted with the precoding-based constellation in Chapter 6. Recall that in this constellation, the symbols are constructed as

$$\boldsymbol{x}_k^{(i)} = \frac{\boldsymbol{U}_k \boldsymbol{c}_k^{(i)}}{\left\|\boldsymbol{U}_k \boldsymbol{c}_k^{(i)}\right\|}, \quad i \in [|\mathcal{X}_k|], k \in [K], \tag{7.98}$$

where $\left\{\boldsymbol{c}_k^{(i)}, \ldots, \boldsymbol{c}_k^{(|\mathcal{X}_k|)}\right\} =: \mathcal{C}_k$ is a single-user Grassmannian constellation in $G(\mathbb{C}^{T-K+1}, 1)$ and $\boldsymbol{U}_k \in \mathbb{C}^{T\times T-K+1}$ is the precoder matrix defined uniquely for user k. The received signal is given in (7.1) with x_k drawn from $\mathcal{X}_k$. Recall that $\boldsymbol{V}_k \in \mathbb{C}^{T\times(K-1)}$ is a basis of the orthogonal complement of Span($\boldsymbol{U}_k$), i.e., $\boldsymbol{V}_k^{\mathsf{H}}\boldsymbol{U}_k = \boldsymbol{0}$, $\boldsymbol{V}_k^{\mathsf{H}}\boldsymbol{V}_k = \boldsymbol{I}_{K-1}$. We have the following property.

Property 7.1. *With normalized linear encoder mappings* (7.98)*, the transmitted signals can be separated and detected error-free in the noiseless case if* $\operatorname{rank}(\mathsf{H}) = K$ *and*

$$\boldsymbol{E}_k := \boldsymbol{V}_k^{\mathsf{H}}[\boldsymbol{U}_1\tilde{\mathsf{x}}_1 \ldots \boldsymbol{U}_{k-1}\tilde{\mathsf{x}}_{k-1}\ \boldsymbol{U}_{k+1}\tilde{\mathsf{x}}_{k+1} \ldots \boldsymbol{U}_K\tilde{\mathsf{x}}_K] \tag{7.99}$$

is full rank for any $k \in [K]$ *and any possible symbol association*

$$\{\tilde{\mathsf{x}}_1, \ldots, \tilde{\mathsf{x}}_{k-1}, \tilde{\mathsf{x}}_{k+1}, \ldots, \tilde{\mathsf{x}}_K\} \in \prod_{l\neq k} \mathcal{C}_l. \tag{7.100}$$

Proof. For any realization of H such that $\operatorname{rank}(\mathsf{H}) = K$, there exists a matrix $\mathsf{A}_k \in \mathbb{C}^{N\times(N-K+1)}$ with orthogonal columns such that $\mathsf{H}_{-k}^{\mathsf{T}}\mathsf{A}_k = \boldsymbol{0}$, with

$$\mathsf{H}_{-k} := [\mathsf{h}_1\ \ldots\ \mathsf{h}_{k-1}\ \mathsf{h}_{k+1}\ \ldots\ \mathsf{h}_K]. \tag{7.101}$$

That is, the columns of $\mathbf{A}_k$ span the orthogonal complement of the column space of $\mathbf{H}_{-k}^*$. With $\mathbf{X}$ constructed using the mapping (7.98), in the noiseless case $\mathbf{Y} = \mathbf{X}\mathbf{H}^{\mathsf{T}}$, we have

$$\boldsymbol{U}_k^{\dagger}\mathbf{Y}\mathbf{A}_k = \boldsymbol{U}_k^{\dagger}\mathbf{X}\mathbf{H}^{\mathsf{T}}\mathbf{A}_k = \tilde{\mathbf{x}}_k\mathbf{h}_k^{\mathsf{T}}\mathbf{A}_k. \tag{7.102}$$

Each column of $\tilde{\mathbf{x}}_k\mathbf{h}_k^{\mathsf{T}}\mathbf{A}_k$ is a scaled version of the Grassmannian symbol $\tilde{\mathbf{x}}_k \in G(\mathbb{C}^{T-K+1}, 1)$, and thus is *identical* to $\tilde{\mathbf{x}}_k$ in $G(\mathbb{C}^{T-K+1}, 1)$. Therefore, $\tilde{\mathbf{x}}_k$ can be detected *without error* from $\mathbf{Y}$ if $\mathbf{A}_k$ is known. Note that in the noiseless case, $\mathbf{A}_k$ spans the null space of $\boldsymbol{V}_k^{\mathsf{H}}\mathbf{Y}$ since $\boldsymbol{V}_k^{\mathsf{H}}\mathbf{Y}\mathbf{A}_k = \boldsymbol{E}_k\mathbf{H}_{-k}^{\mathsf{T}}\mathbf{A}_k = \mathbf{0}$, and thus $\mathbf{A}_k$ can be identified from $\boldsymbol{V}_k^{\mathsf{H}}\mathbf{Y}$ *without* the knowledge of $\mathbf{H}$ as long as $\boldsymbol{E}_k$ is full rank. □

We now specify the condition in Property 7.1 for the two precoder designs in Chapter 6. This condition can be equivalently stated as

$$\sum_{l=1, l\neq k}^{K} \alpha_l \boldsymbol{U}_l\tilde{\mathbf{x}}_l \notin \mathrm{Span}(\boldsymbol{U}_k), \tag{7.103}$$

for all $k \in [K]$ and any possible codeword association $\{\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_{k-1}, \tilde{\mathbf{x}}_{k+1}, \ldots, \tilde{\mathbf{x}}_K\} \in \prod_{l\neq k}\mathcal{C}_l$, and any complex weights $\{\alpha_l\}$ such that $\sum_{l=1,l\neq k}^{K}|\alpha_l|^2 > 0$.

Proposition 7.2. *Precoder Type I fulfills the condition in Property 7.1 if*

$$\det(\mathbf{X}_{-k,(k-1)(K-1)+1:k(K-1)}) \neq 0, \tag{7.104}$$

for any $k \in [K]$ and any $\mathbf{X}_{-k} \in \prod_{l\neq k}\mathcal{X}_l$, where $\mathbf{X}_{-k,(k-1)(K-1)+1:k(K-1)}$ denotes the sub-matrix containing the rows from $(k-1)(K-1)+1$ to $k(K-1)$ of $\mathbf{X}_{-k}$.

Proof. For Precoder Type I, the condition (7.103) is violated when

$$\left(\sum_{l=1,l\neq k}^{K}\alpha_l\boldsymbol{U}_l\tilde{\mathbf{x}}_l\right)_{(k-1)(K-1)+1:k(K-1)} = \mathbf{0}, \tag{7.105}$$

which can be written as

$$\mathbf{X}_{-k,(k-1)(K-1):k(K-1)+1}\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_{k-1}\\ \alpha_{k+1}\\ \vdots\\ \alpha_K\end{bmatrix} = \mathbf{0}, \tag{7.106}$$

where $\mathbf{X}_{-k,(k-1)(K-1)+1:k(K-1)}$ denotes the sub-matrix containing the rows from $(k-1)(K-1)+1$ to $k(K-1)$ of $\mathbf{X}_{-k} = [\mathbf{x}_1 \ldots \mathbf{x}_{k-1}\ \mathbf{x}_{k+1} \ldots \mathbf{x}_K]$. There exists non-zero solution $\{\alpha_l\}$ to the homogeneous system (7.104) if $\det(\mathbf{X}_{-k,(k-1)(K-1)+1:k(K-1)}) = 0$. Therefore, the condition for Precoder Type I is finally stated as in (7.104). □

Remark 7.2. As soon as $K > 3$, a necessary (but not sufficient) condition is that the constellations $\mathcal{C}_k$, $k \in [K]$, are all different, since otherwise we have $\det(\mathbf{X}_{-k,(k-1)(K-1)+1:k(K-1)}) = 0$ whenever $\tilde{\mathbf{x}}_{k+1} = \tilde{\mathbf{x}}_{k+2}$ or $\tilde{\mathbf{x}}_{k-1} = \tilde{\mathbf{x}}_{k-2}$.

Proposition 7.3. *Precoder Type II fulfills the condition in Property 7.1 if*

$$\prod_{k=1}^{K}\tilde{\mathbf{x}}_k^{\mathsf{T}}\boldsymbol{e}_1 \neq 0, \quad \forall\{\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_K\} \in \prod_{k=1}^{K}\mathcal{D}_k. \tag{7.107}$$

Proof. Let $\tilde{\mathsf{x}}_{k,l} = \tilde{\mathbf{x}}_k^{\mathsf{T}} \boldsymbol{e}_l$ denote the l-th entry of $\tilde{\mathbf{x}}_k$. We can derive

$$\sum_{l=1,l\neq k}^{K} \alpha_l \boldsymbol{U}_l \tilde{\mathbf{x}}_l = \begin{bmatrix} \alpha_1\sqrt{\eta_1}\tilde{\mathsf{x}}_{1,1} \\ \vdots \\ \alpha_{k-1}\sqrt{\eta_1}\tilde{\mathsf{x}}_{k-1,1} \\ 0 \\ \alpha_{k+1}\sqrt{\eta_1}\tilde{\mathsf{x}}_{k+1,1} \\ \vdots \\ \alpha_K\sqrt{\eta_1}\tilde{\mathsf{x}}_{K,1} \\ \sum_{l\neq k}\alpha_l\sqrt{\eta_2}\tilde{\mathsf{x}}_{l,2} \\ \vdots \\ \sum_{l\neq k}\alpha_l\sqrt{\eta_2}\tilde{\mathsf{x}}_{l,T-K+1} \end{bmatrix}. \tag{7.108}$$

Thus the condition (7.103) is fulfilled if

$$\begin{bmatrix} \alpha_1\sqrt{\eta_1}\tilde{\mathsf{x}}_{1,1} \\ \vdots \\ \alpha_{k-1}\sqrt{\eta_1}\tilde{\mathsf{x}}_{k-1,1} \\ \alpha_{k+1}\sqrt{\eta_1}\tilde{\mathsf{x}}_{k+1,1} \\ \vdots \\ \alpha_K\sqrt{\eta_1}\tilde{\mathsf{x}}_{K,1} \end{bmatrix} \neq \mathbf{0}, \tag{7.109}$$

for any k and any $\{\alpha_l\}$ such that $\sum_{l=1,l\neq k}^{K} |\alpha_l|^2 > 0$. This is true when $\eta_1 = \frac{K(T-K+1)}{T} > 0$ and $\tilde{\mathsf{x}}_{1,1}, \ldots, \tilde{\mathsf{x}}_{k-1,1}, \tilde{\mathsf{x}}_{k+1,1}, \ldots, \tilde{\mathsf{x}}_{K,1}$ are all non-zero for any $k \in [K]$. Since $T > K-1$, this means that (7.107) holds. □

The condition (7.107) for Precoder Type II is much easier to be fulfilled than the condition (7.104) for Precoder Type I. We note that given a fixed $\{\mathcal{C}_k\}$ whose design does not necessarily take the conditions in Propositions 7.2 and 7.3 into account, we can apply a random rotation to have $\tilde{\mathcal{C}}_k = \mathbf{F}_k\mathcal{C}_k$, which has the same distance spectrum as $\mathcal{C}_k$, where $\mathbf{F}_k \in \mathbb{C}^{(T-K+1)\times(T-K+1)}$ is a random unitary matrix. Then, the condition in Property 7.1 is guaranteed with high probability.

At the receiver, the greedy approach exploits Property 7.1 to separate the signals of different users and detect each signal as in the single-user case. As a consequence, the multi-user ML detection is decoupled into K single-user detection problems. The details are as follows.

7.7.1 Separation-First Detector

Inspired by Property 7.1, we directly separate the signal of user k using a matrix $\mathbf{A}_k \in \mathbb{C}^{N\times(N-K+1)}$, $\mathbf{A}_k^{\mathsf{H}}\mathbf{A}_k = \boldsymbol{I}_{N-K+1}$, computed from $\mathbf{Y}$ such that $\boldsymbol{V}_k^{\mathsf{H}}\mathbf{Y}\mathbf{A}_k = \mathbf{0}$. The matrix $\mathbf{A}_k$ spans the null-space of $\boldsymbol{V}_k^{\mathsf{H}}\mathbf{Y} = \boldsymbol{E}_k\mathbf{H}_{-k} + \boldsymbol{V}_k^{\mathsf{H}}\mathbf{Z}$ with

$$\boldsymbol{E}_k := \boldsymbol{V}_k^{\mathsf{H}}[\boldsymbol{U}_1\tilde{\mathbf{x}}_1 \ \ldots \ \boldsymbol{U}_{k-1}\tilde{\mathbf{x}}_{k-1} \ \boldsymbol{U}_{k+1}\tilde{\mathbf{x}}_{k+1} \ \ldots \ \boldsymbol{U}_K\tilde{\mathbf{x}}_K]. \tag{7.110}$$

Assuming that $\boldsymbol{E}_k$ is full rank, $\mathbf{A}_k$ serves as an estimate of the basis of the subspace orthogonal to the row space of $\mathbf{H}_{-k}^*$. Then,

$$\tilde{\mathbf{Y}}_k = \boldsymbol{U}_k^{\dagger}\mathbf{Y}\mathbf{A}_k = \tilde{\mathbf{x}}_k\mathbf{h}_k^{\mathsf{T}}\mathbf{A}_k + \boldsymbol{U}_k^{\dagger}\left(\sum_{l=1,l\neq k}^{K}\boldsymbol{U}_l\tilde{\mathbf{x}}_l\mathbf{h}_l^{\mathsf{T}} + \mathbf{Z}\right)\mathbf{A}_k$$

is equivalent to a SIMO P2P channel output with channel vector $\tilde{\mathbf{h}}_k = \mathbf{A}_k^{\mathsf{T}}\mathbf{h}_k \in \mathbb{C}^{N-K+1}$ and equivalent noise $\tilde{\mathbf{Z}}_k = \boldsymbol{U}_k^{\dagger}\Big(\sum_{l=1,l\neq k}^{K}\boldsymbol{U}_l\tilde{\mathbf{x}}_l\mathbf{h}_l^{\mathsf{T}} + \mathbf{Z}\Big)\mathbf{A}_k$. We next decode $\tilde{\mathbf{x}}_k$ from $\tilde{\mathbf{Y}}_k$ using a single-user detector.

By construction, $\mathbf{A}_k$ is a truncated unitary matrix independent of $\mathbf{h}_k$. Thus, in the case of IID Rayleigh fading $\mathbf{h}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{I}_N)$, we get $\tilde{\mathbf{h}}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{I}_{N-K+1})$ which allows to approximate the ML detector for $\tilde{\mathbf{x}}_k$, considering the equivalent noise as a white Gaussian vector, by

$$\hat{\tilde{\mathbf{x}}}_k = \arg\max_{\tilde{\mathbf{x}}\in\mathcal{C}_k} \big\|\tilde{\mathbf{Y}}_k^{\mathsf{H}}\tilde{\mathbf{x}}\big\|^2. \tag{7.111}$$

For correlated Rayleigh fading $\mathbf{h}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{R}_k)$, we get $\tilde{\mathbf{h}}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{A}_k^{\mathsf{T}}\boldsymbol{R}_k\mathbf{A}_k)$, and the approximate ML detector is given (similar to (7.4)) by

$$\hat{\tilde{\mathbf{x}}}_k = \arg\min_{\tilde{\mathbf{x}}\in\mathcal{C}_k}\Big(\mathrm{vec}(\tilde{\mathbf{Y}}^{\mathsf{T}})^{\mathsf{H}}\big(\sigma^2\boldsymbol{I}+\tilde{\mathbf{x}}_k\tilde{\mathbf{x}}_k^{\mathsf{H}}\otimes\mathbf{A}_k^{\mathsf{T}}\boldsymbol{R}_k\mathbf{A}_k\big)^{-1}\mathrm{vec}(\tilde{\mathbf{Y}}^{\mathsf{T}}) + \log\det\big(\sigma^2\boldsymbol{I}+\tilde{\mathbf{x}}_k\tilde{\mathbf{x}}_k^{\mathsf{H}}\otimes\mathbf{A}_k^{\mathsf{T}}\boldsymbol{R}_k\mathbf{A}_k\big)\Big). \tag{7.112}$$

The marginal posteriors can be computed accordingly. In general, the approximate ML detector has complexity $O(K^2 2^B)$ per user for uncorrelated fading and $O(K^6 2^B)$ per user for correlated fading. (Note that this complexity can be further reduced if the single-user constellation $\mathcal{C}_k$ has a structure, such as the cube-split constellation in Chapter 5.) Since the computation of $\mathbf{A}_k$ and $\tilde{\mathbf{Y}}_k$ both have complexity order $O(K^3)$, $k \in [K]$, the total complexity order of the separation-first detector is $O(K^4 + K^3 2^B)$ for uncorrelated fading and $O(K^7 2^B)$ for correlated fading, equivalent to that of one MMSE-SIA iteration.

7.7.2 Denoising-First Detector

This detector is similar to the separation-first detector except that before the separation step (find $\mathbf{A}_k$ and compute $\tilde{\mathbf{Y}}_k$), we perform a denoising step by solving

$$\Big\{\hat{\mathbf{X}}, \mathbf{H}\Big\} = \arg\max_{\mathbf{H}\in\mathbb{C}^{N\times K},\ \mathbf{X}\in\mathbb{C}^{T\times K}:\,\mathbf{X}^{\mathsf{H}}\mathbf{X}=\boldsymbol{I}_K} \|\mathbf{Y} - \mathbf{X}\mathbf{H}^{\mathsf{T}}\|_{\mathrm{F}}^2. \tag{7.113}$$

The solution of this optimization is $\mathbf{H} = \mathbf{Y}^{\mathsf{T}}\hat{\mathbf{X}}^*\Big(\hat{\mathbf{X}}^{\mathsf{T}}\hat{\mathbf{X}}^*\Big)^{-1}$ with

$$\hat{\mathbf{X}} = \arg\max_{\mathbf{X}\in\mathbb{C}^{T\times K}:\,\mathbf{X}^{\mathsf{H}}\mathbf{X}=\boldsymbol{I}} \|\mathbf{Y}^{\mathsf{H}}\mathbf{X}\|_{\mathrm{F}}^2. \tag{7.114}$$

Thus $\hat{\mathbf{X}}$ is the collection of the K singular vectors of $\mathbf{Y}$ associated to the K strongest singular values. Then we find the unique $\hat{\tilde{\mathbf{x}}}_1, \ldots, \hat{\tilde{\mathbf{x}}}_K$ such that $[\boldsymbol{U}_1\hat{\tilde{\mathbf{x}}}_1 \ \ldots \ \boldsymbol{U}_K\hat{\tilde{\mathbf{x}}}_K]$ spans the column space of $\hat{\mathbf{X}}$, seeing that each $\boldsymbol{U}_k\hat{\tilde{\mathbf{x}}}_k$ is in the column space of $\boldsymbol{U}_k$. In order to do so, we now apply the separation step in the separation-first detector to $\hat{\mathbf{X}}$. That is, we first find a unit vector $\mathbf{a}_k \in \mathbb{C}^K$ such that $\boldsymbol{V}_k^{\mathsf{H}}\hat{\mathbf{X}}\mathbf{a}_k = \mathbf{0}$, then consider $\tilde{\mathbf{y}}_k = \boldsymbol{U}_k^{\dagger}\hat{\mathbf{X}}\mathbf{a}_k$ as the input of a single-user detector. In doing so, the multi-user detection is decoupled into K single-user SISO detection problems. The total complexity order is $O(K^4 + K^2 2^B)$.

7.7.3 POCIS: Interference Mitigation

In the above detectors, the interference $\sum_{l=1,l\neq k}^{K}\boldsymbol{U}_l\tilde{\mathbf{x}}_l\mathbf{h}_l^{\mathsf{T}}$ is partially mitigated using an estimation ($\mathbf{A}_k$) of its *row space*. Now, assuming that we have the estimates $\hat{\tilde{\mathbf{x}}}_1, \ldots, \hat{\tilde{\mathbf{x}}}_K$ of $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_K$ from the separation-first or denoising-first detector, we can try to further cancel the residual interference. Let $\mathbf{O}_k$ be an orthonormal basis of the column space

of $[\boldsymbol{U}_1\hat{\tilde{\mathbf{x}}}_1 \dots \boldsymbol{U}_{k-1}\hat{\tilde{\mathbf{x}}}_{k-1}\ \boldsymbol{U}_{k+1}\hat{\tilde{\mathbf{x}}}_{k+1} \dots \boldsymbol{U}_K\hat{\tilde{\mathbf{x}}}_K]$, which is approximately, up to the noise perturbation, the *column space* of the interference. We project the received signal onto the orthogonal of this approximate interference space

$$\tilde{\mathbf{Y}}_k = \boldsymbol{U}_k^{\dagger}(\boldsymbol{I}_T - \mathbf{O}_k\mathbf{O}_k^{\mathsf{H}})\mathbf{Y}\mathbf{A}_k \tag{7.115}$$

$$= \boldsymbol{U}_k^{\dagger}(\boldsymbol{I}_T - \mathbf{O}_k\mathbf{O}_k^{\mathsf{H}})\boldsymbol{U}_k\tilde{\mathbf{x}}_k\tilde{\mathbf{h}}_k^{\mathsf{T}} + \boldsymbol{U}_k^{\dagger}(\boldsymbol{I}_T - \mathbf{O}_k\mathbf{O}_k^{\mathsf{H}})\left(\sum_{l=1,l\neq k}^{K} \boldsymbol{U}_l\tilde{\mathbf{x}}_l\mathbf{h}_l^{\mathsf{T}} + \mathbf{Z}\right)\mathbf{A}_k. \tag{7.116}$$

We then decode for $\tilde{\mathbf{x}}_k$ or its posterior from $\tilde{\mathbf{Y}}_k$ using a single-user detector. With these new estimates $\hat{\tilde{\mathbf{x}}}_1, \dots, \hat{\tilde{\mathbf{x}}}_K$, we can repeat the interference mitigation, making it an iterative process. Each iteration is equivalent in complexity to the separation-first or denoising-first detector. We refer to this iterative interference mitigation detector as Projection onto the Orthogonal Complement of the Interference Subspace (POCIS). Using the separation-first detector for each iteration, POCIS has the same complexity order as MMSE-SIA.

7.8 Performance Evaluation

In this section, we evaluate the performance of our proposed schemes for a given set of individual constellations. We assume that $B_1 = \dots B_K =: B$. We consider the local scattering model [207, Section 2.6] for the correlation matrices $\boldsymbol{R}_k$. Specifically, the (l, m)-th element of $\boldsymbol{R}_k$ is generated as

$$[\boldsymbol{R}_k]_{l,m} = \xi_k \mathbb{E}_{\delta_k}[\exp(2\pi d_H(l-m)\sin(\varphi_k + \delta_k))], \tag{7.117}$$

where d_H is the antenna spacing in the receiver array (measured in number of wavelengths), φ_k is a deterministic nominal angle, and δ_k is a random deviation. We consider $d_H = \frac{1}{2}$, φ_k generated uniformly in $[-\pi, \pi]$, and δ_k uniformly distributed in $[-\sqrt{3}\sigma_\varphi, \sqrt{3}\sigma_\varphi]$ with angular standard deviation $\sigma_\varphi = 10^\circ$. We also consider $\xi_k = 1, \forall k$. We set a damping factor $\eta = 0.9$ for EP, EPAK, and MMSE-SIA.

7.8.1 Test Constellations, State-of-the-Art Detectors, and Benchmarks

7.8.1.a Precoding-Based Grassmannian Constellations

We consider the precoding-based constellation design in Chapter 6. Recall that this design imposes a geometric separation between the individual constellations through a set of precoders $\boldsymbol{U}_k$, $k \in [K]$. Specifically, starting with a Grassmannian constellation[6] $\mathcal{C} = \{\boldsymbol{c}^{(1)}, \dots, \boldsymbol{c}^{(2^B)}\}$ in $G(\mathbb{C}^{T-K+1}, 1)$, the individual constellation $\mathcal{X}_k$ is generated as $\boldsymbol{x}_k^{(i)} = \frac{\boldsymbol{U}_k\boldsymbol{c}^{(i)}}{\|\boldsymbol{U}_k\boldsymbol{c}^{(i)}\|}, i \in [2^B]$. We consider the Precoder Type II and two candidates for $\mathcal{C}$:

- A numerically optimized constellation generated by solving the max-min distance criteria

$$\max_{\boldsymbol{c}^{(i)} \in G(\mathbb{C}^{T-K+1},1), i=1,\dots,2^B} \ \min_{1 \leq i < j \leq 2^B} d(\boldsymbol{c}^{(i)}, \boldsymbol{c}^{(j)}), \tag{7.118}$$

  where $d(\boldsymbol{c}^{(i)}, \boldsymbol{c}^{(j)}) := \sqrt{1 - |\boldsymbol{c}^{(i)\mathsf{H}}\boldsymbol{c}^{(j)}|^2}$ is the chordal distance between two Grassmannian points represented by $\boldsymbol{c}^{(i)}$ and $\boldsymbol{c}^{(j)}$. A constellation with maximal minimum pairwise distance leads to low symbol error rate in the absence of the interference. In our simulation, we approximate (7.118) by $\min_{\mathcal{C}} \log \sum_{1 \leq i < j \leq 2^B} \exp\left(\frac{|\boldsymbol{c}^{(i)\mathsf{H}}\boldsymbol{c}^{(j)}|}{\epsilon}\right)$ with a small ϵ for smoothness, then solve it using gradient descent on the Grassmann manifold using the Manopt toolbox [141]. Please see Example 1.1 for details.

[6] The constellation $\mathcal{C}$ can be different for different users. Here, we consider a common constellation.

- The cube-split constellation introduced in Chapter 5. This structured constellation has good distance properties and allows for low-complexity single-user decoding and a simple yet effective binary labeling scheme.

We consider the POCIS detector in Section 7.7 as baseline. Note that only the indices of the estimated symbols are passed in POCIS, as opposed to the soft information on the symbols as in EP, MMSE-SIA, and EPAK.

7.8.1.b Pilot-Based Constellations

We also consider the pilot-based constellations in which the symbols are generated as $\boldsymbol{x}_k^{(i)} = \left[\sqrt{\frac{K}{T}}\boldsymbol{e}_k^{\mathsf{T}} \ \sqrt{\frac{T-K}{TP_{\text{avg}}}}\tilde{\boldsymbol{x}}_k^{(i)\mathsf{T}}\right]^{\mathsf{T}}$ where $\boldsymbol{e}_k$ is the k-th column of $\boldsymbol{I}_K$, $\tilde{\boldsymbol{x}}_k^{(i)}$ is a vector of data symbols taken from a scalar constellation, such as QAM, and P_{avg} is the average symbol power of the considered scalar constellation. Note that this corresponds to the scenario where the K users transmit mutually orthogonal pilot sequences, followed by spatially multiplexed parallel data transmission. Many MIMO detectors have been proposed specifically for these constellations. We consider some representatives as follows.

- The receiver MMSE-estimates the channel based on the first K rows of $\mathbf{Y}$, then MMSE-equalizes the received data symbols in the remaining $T-K$ rows of $\mathbf{Y}$, and performs a scalar demapper on the equalized symbols. This is referred to as a linear MMSE detector.
- The receiver MMSE-estimates the channel, then decodes the data symbols using the sphere decoder by Schnorr and Euchner [208], referred to as Schnorr-Euchner sphere decoder (SESD).
- The receiver performs the semi-blind joint ML channel estimation and data detection scheme in [100] with repeated weighted boosting search (RWBS) for channel estimation and SESD for data detection, referred to as RWBS-SESD.

We note that the sphere decoder has near optimal performance given the channel knowledge, but its complexity is non-deterministic and can be exponential in the channel dimension if the channel matrix is ill-conditioned.

7.8.1.c Benchmarks

We consider the optimal ML detector, whenever it is feasible, as a benchmark. When the optimal detector is computationally infeasible, we resort to another benchmark consisting in giving the receiver, while it decodes the signal $\mathbf{x}_k$ of user k, the knowledge of the signals $\mathbf{x}_l$ (but not the channel $\mathbf{h}_l$) of all the interfering users $l \neq k$. With this genie-aided information, optimal ML decoding (7.4) can be performed by keeping $\mathbf{x}_l$ fixed for all $l \neq k$ and searching for the best $\mathbf{x}_k$ in $\mathcal{X}_k$, thus reducing the total search space size from 2^{BK} to $K2^B$. The posterior marginals are computed separately for each user accordingly. This genie-aided detector gives an upper bound on the performance of EP, MMSE-SIA, EPAK, and POCIS.

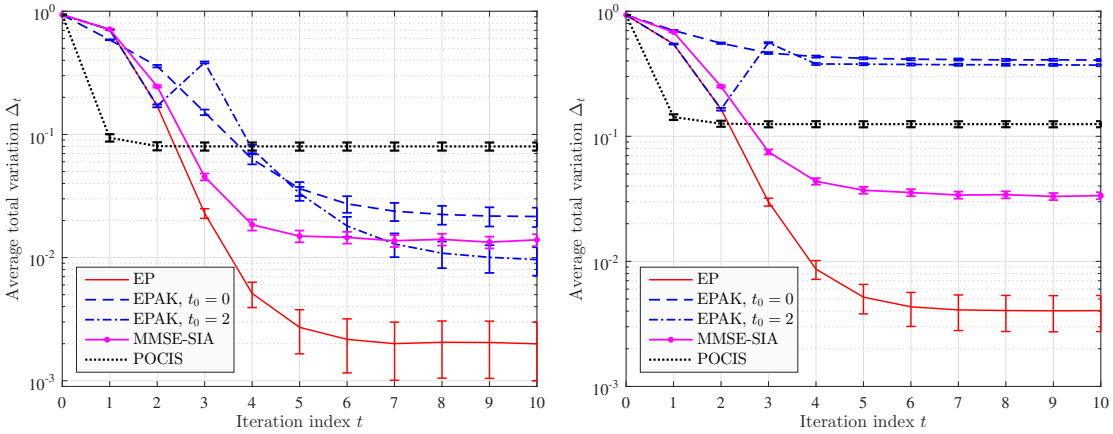
7.8.2 Convergence and Running Time

To assess the convergence of the algorithms, we evaluate the total variation distance between the estimated marginal posteriors $\hat{p}_{\mathbf{x}_k|\mathbf{Y}}$ at each iteration and the exact marginal posteriors $p_{\mathbf{x}_k|\mathbf{Y}}$ when exact marginalization (7.8) is possible. The total variation distance between two probability measures P and Q on $\mathcal{S}$ is defined as $\mathcal{TV}(P,Q) := \frac{1}{2}\sum_{x\in\mathcal{S}}|P(x)-Q(x)|$. At iteration t where the estimated posteriors are $\hat{p}^{(t)}_{\mathbf{x}_k|\mathbf{Y}}$, $k \in [K]$, we evaluate the average

total variation distance as

$$\Delta_t = \frac{1}{K}\sum_{k=1}^{K} \mathbb{E}_{\mathbf{Y}}\Big[\mathcal{TV}\Big(\hat{p}^{(t)}_{\mathbf{x}_k|\mathbf{Y}}, p_{\mathbf{x}_k|\mathbf{Y}}\Big)\Big]. \tag{7.119}$$

We consider the precoding-based Grassmannian constellations. In Fig. 7.3, we show the empirical average total variation Δ_t for $T = 6$, $K = 3$, $N = 4$, and $B = 4$ at SNR $= 8$ dB. As can be seen, at convergence, EP provides the most accurate estimates of the marginal posteriors in average although it is less stable than other schemes. EP converges after 6 iterations while MMSE-SIA converges after 5 iterations. For uncorrelated fading, EPAK with $t_0 = 2$ can be eventually better than MMSE-SIA, but converges slower. For correlated fading, EPAK totally fails because of the inaccuracy of the approximation with Kronecker products. POCIS converges very quickly after 2 iterations but achieves a relatively low accuracy of the posterior estimation.



(a) Uncorrelated fading $\boldsymbol{R}_k = \boldsymbol{I}_N$ (b) Correlated Fading

FIGURE 7.3: The empirical average total variation Δ_t over 1000 realizations of the transmitted signal, channel, and noise versus iteration for different non-coherent soft detection schemes for $T = 6$, $K = 3$, $B = 4$, and $N = 4$ at SNR $= 8$ dB. The error bars show the standard error, which is the standard deviation normalized by the square root of the number of samples. For correlated fading, these figures are further averaged over 10 realizations of the correlation matrices.

Fig. 7.4 depicts the average running time (on a local server) of exact marginalization in comparison with 6 iterations of EP, EPAK, MMSE-SIA, and POCIS at SNR $= 8$ dB. These schemes have significantly lower computation time than exact marginalization. The running time saving of EPAK w.r.t. EP is not significant, even with $t_0 = 0$. For uncorrelated fading, MMSE-SIA has much shorter running time than all other schemes.

From these convergence behaviors, hereafter, we fix the number of iterations of EP, MMSE-SIA, and EPAK as 6 and of POCIS as 3. Furthermore, we consider EPAK only for uncorrelated fading. For correlated fading, we generate the correlation matrices once and fix them over the simulation.

7.8.3 Achievable Rate

We first plot the achievable mismatched sum-rate R_{GMI} of the system calculated as in (7.12) for $T = 6$, $K = 3$, $N = 4$ and $B \in \{4, 8\}$ in Fig. 7.5. We consider the precoding-based Grassmannian constellations. For $\mathcal{C}$, we use the numerically optimized constellation if

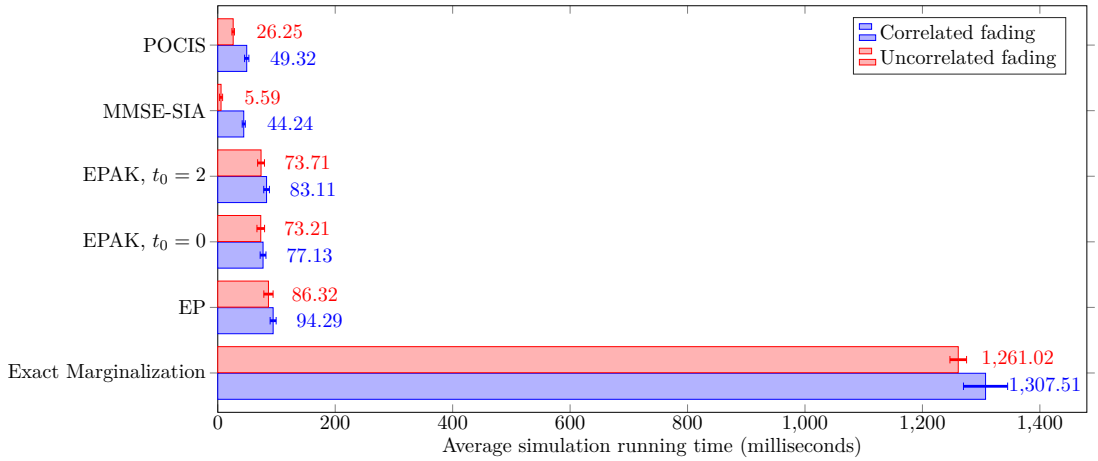


FIGURE 7.4: The average running time over 1000 realizations of the transmitted signal, channel, and noise of exact marginalization in comparison with 6 iterations of the considered detection schemes for $T = 6$, $K = 3$, $B = 4$, and $N = 4$ at SNR $= 8$ dB. The error bars show the standard deviation. For correlated fading, the running time is further averaged over 10 realizations of the correlation matrices.

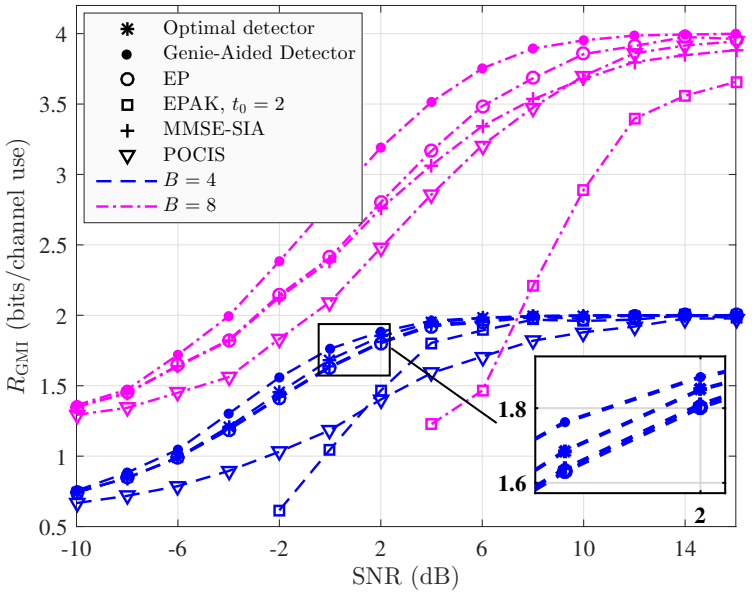
$B = 4$ and the cube-split constellation if $B = 8$. For uncorrelated fading (Fig. 7.5(a), the rates achieved with EP and MMSE-SIA detectors are very close to the achievable rate of the system (with the optimal detector) and not far from that of the genie-aided detector. EPAK (with $t_0 = 2$) achieves a very low rate, especially in the low SNR regime where the Kronecker approximation is not accurate. For correlated fading, (Fig. 7.5(b)), the rates achieved with EP and MMSE-SIA are only marginally lower than that of the optimal detector and genie-aided detector. In both cases, the rate achieved with POCIS is lower than that of EP and MMSE-SIA in the lower SNR regime and converges slowly with SNR to the limit $\frac{BK}{T}$ bits/channel use.

7.8.4 Symbol Error Rates of Hard Detection

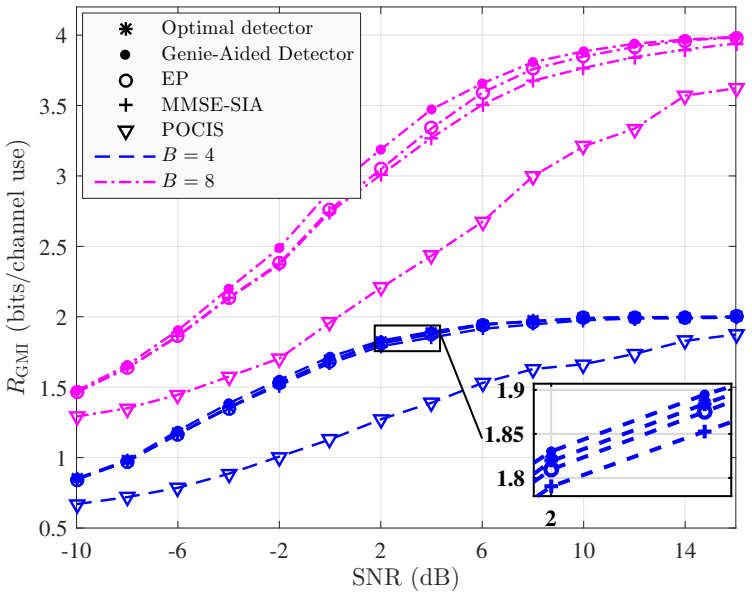
Next, we use the outputs of EP, EPAK, MMSE-SIA and POCIS for a maximum-a-posteriori (MAP) hard detection. We evaluate the performance in terms of SER.

In Fig. 7.6, we consider the precoding-based constellations with $T = 6$, $K = 3$, $N \in \{4, 8\}$, and $B = 4$, for which the optimal ML detector (7.4) is computationally feasible. We observe that the SER of the EP and MMSE-SIA detectors are not much higher than that of the optimal detector, especially in the lower SNR regime. The SER of EPAK is significantly higher than that of EP and MMSE-SIA for $t_0 = 0$. This is greatly improved by setting $t_0 = 2$, i.e., keeping the first two iterations of EP. The gain of EP w.r.t. EPAK and MMSE-SIA is more pronounced when the SNR increases. For correlated fading, EP performs almost as good as the optimal detector, whose SER performance is closely approximated by the genie-aided detector.

In Fig. 7.7, we consider $T = 6$, $K = 3$, $N = 8$, and $B = 9$ and use the genie-aided detector as a benchmark. In Fig. 7.7(a), we consider uncorrelated fading and use the pilot-based constellations with 8-QAM data symbols. The performance of EP is very close to that of the genie-aided detector. The performance of MMSE-SIA is close to EP in the low SNR regime (SNR ≤ 8 dB). We also depict the SER of the three pilot-based detectors in Section 7.8.1.b, namely, 1) linear MMSE detector, 2) SESD, and 3) RWBS-SESD. These three schemes are outperformed by the EP detectors. In Fig. 7.7(b), we consider correlated fading and use the precoding-based Grassmannian constellations



(a) Uncorrelated fading



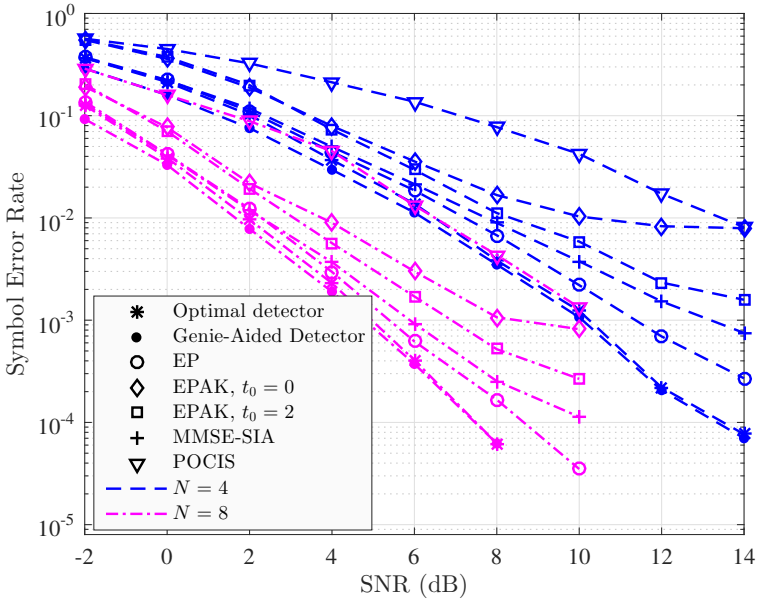
(b) Correlated fading

FIGURE 7.5: The mismatched rate of the system with EP, EPAK (with $t_0 = 2$), MMSE-SIA, and POCIS detectors in comparison with the optimal detector and/or the genie-aided detector for $T = 6$, $K = 3$, $N = 4$, and $B \in \{4, 8\}$.

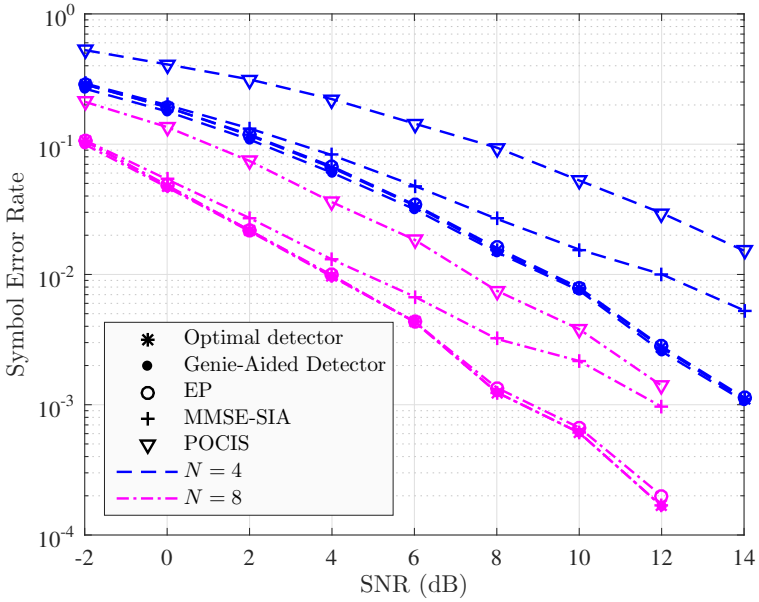
with $\mathcal{C}$ numerically optimized. We observe again that EP achieves almost the same SER performance as the genie-aided detector.

7.8.5 Bit Error Rates with a Channel Code

In this subsection, we use the output of the soft detectors for channel decoding. We consider the precoding-based Grassmannian constellations with the cube-split constellation for $\mathcal{C}$



(a) Uncorrelated fading

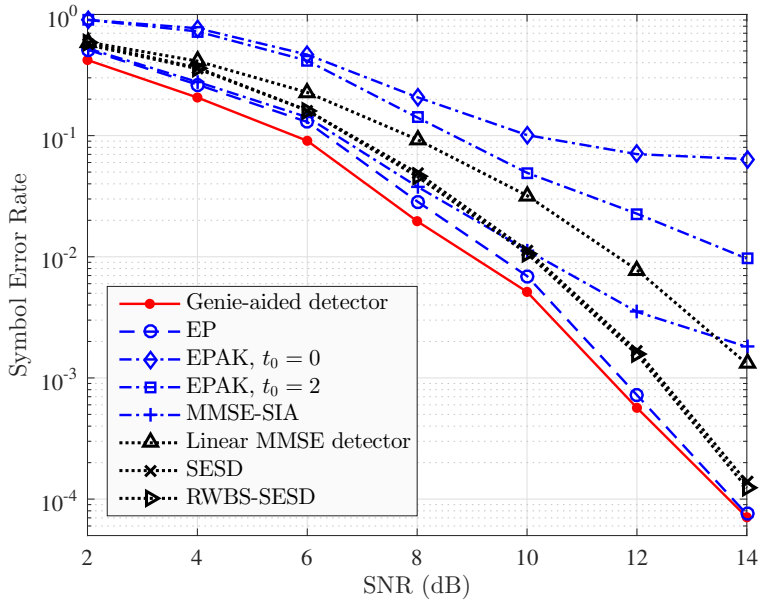


(b) Correlated fading

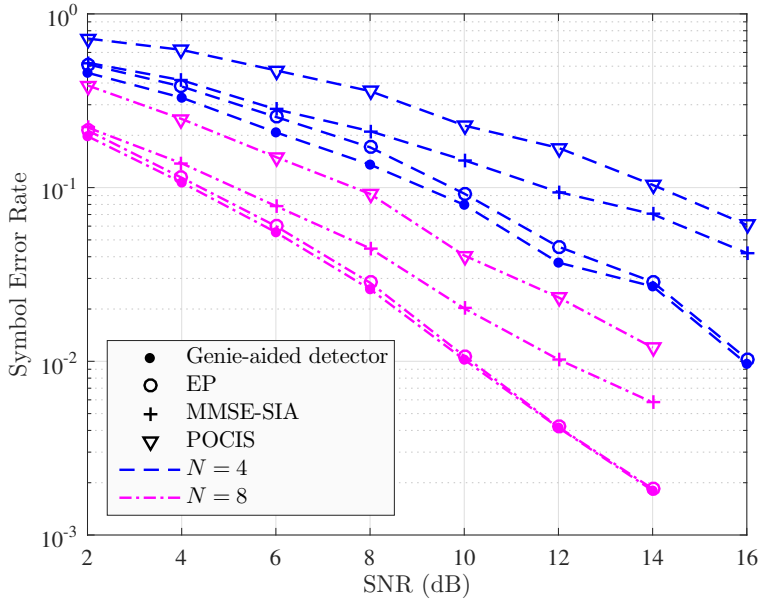
FIGURE 7.6: The symbol error rate of the system with EP, EPAK (with $t_0 \in \{0, 2\}$), MMSE-SIA, and POCIS detectors in comparison with the optimal detector and the genie-aided detector for $T = 6$, $K = 3$, $N \in \{4, 8\}$ and $B = 4$.

since it admits an effective and simple binary labeling [123]. We take the binary labels of the symbols in $\mathcal{C}$ for the corresponding symbols in $\mathcal{X}_k$. We integrate a standard symmetric parallel concatenated rate-1/3 turbo code [184]. The turbo encoder accepts packets of 1008 bits; the turbo decoder computes the bit-wise LLR from the soft outputs of the detection scheme as in (7.5) and performs 10 decoding iterations for each packet.

In Fig. 7.8, we show the BER with this turbo code using $B = 8$ bits/symbol and different values of T and $K = N$. EP achieves the closest performance to the genie-aided



(a) Uncorrelated fading, pilot-based constellations

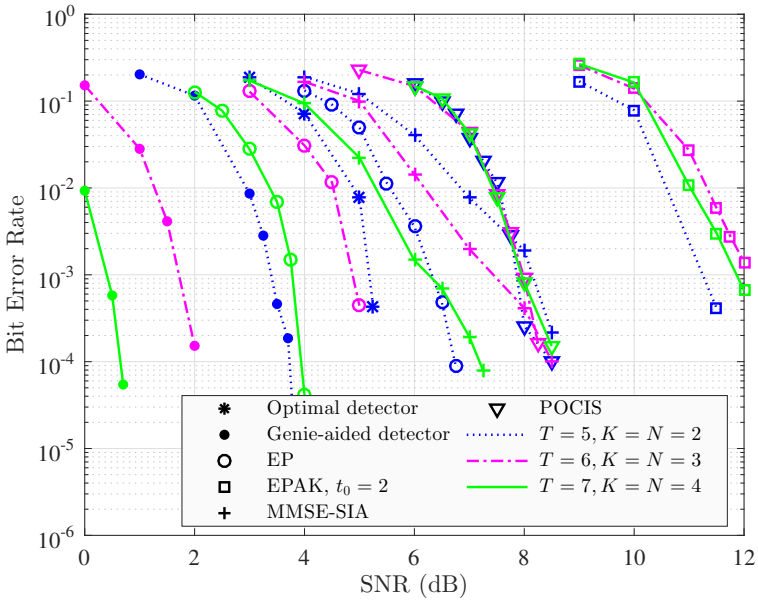


(b) Correlated fading, precoding-based constellations

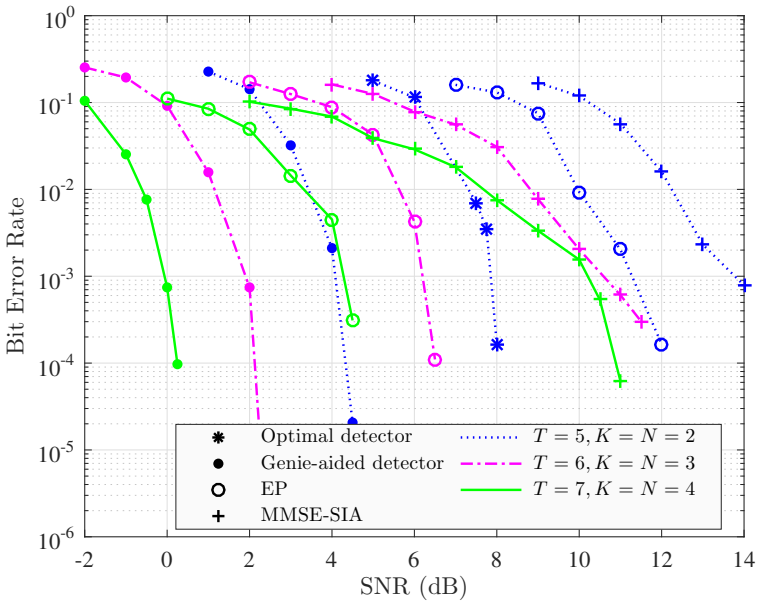
FIGURE 7.7: The symbol error rate of the system with EP, EPAK ($t_0 \in \{0, 2\}$), MMSE-SIA, POCIS vs. the genie-aided detector for $T = 6$, $K = 3$, $N = 8$, and $B = 9$. For uncorrelated fading, these schemes are compared with three pilot-based detectors using respectively MMSE equalizer, sphere decoding [208], and joint channel estimation–data detection [100].

detector and the optimal detector (7.8). The BER of MMSE-SIA vanishes slower with the SNR than the other schemes, and becomes better than POCIS as K and N increase. The BER of EPAK with $t_0 = 2$ is higher than all other schemes. Under uncorrelated fading, for $T = 7$ and $K = N = 4$, the power gain of EP w.r.t. MMSE-SIA, POCIS, and EPAK for the same BER of 10^{-3} is about 3 dB, 4 dB, and 8 dB, respectively. We also observe that the genie-aided detector gives very optimistic BER performance results compared to the

optimal detector.



(a) Uncorrelated fading

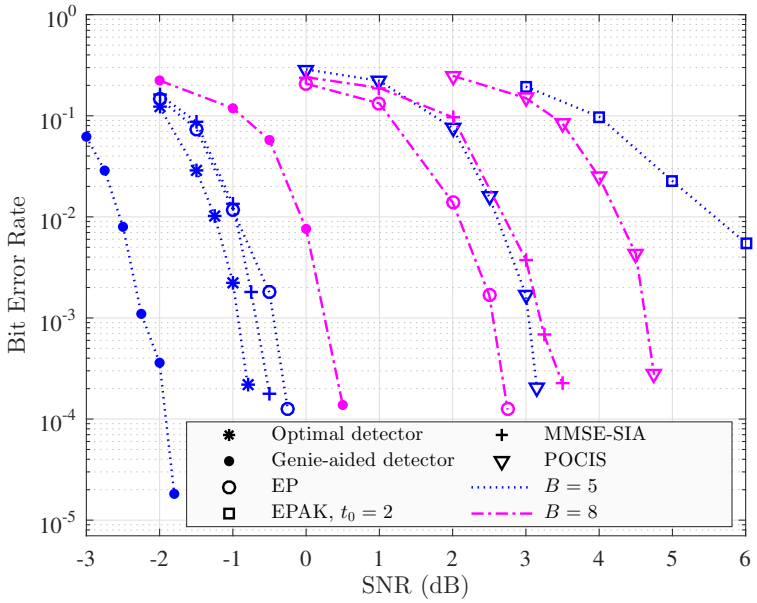


(b) Correlated fading

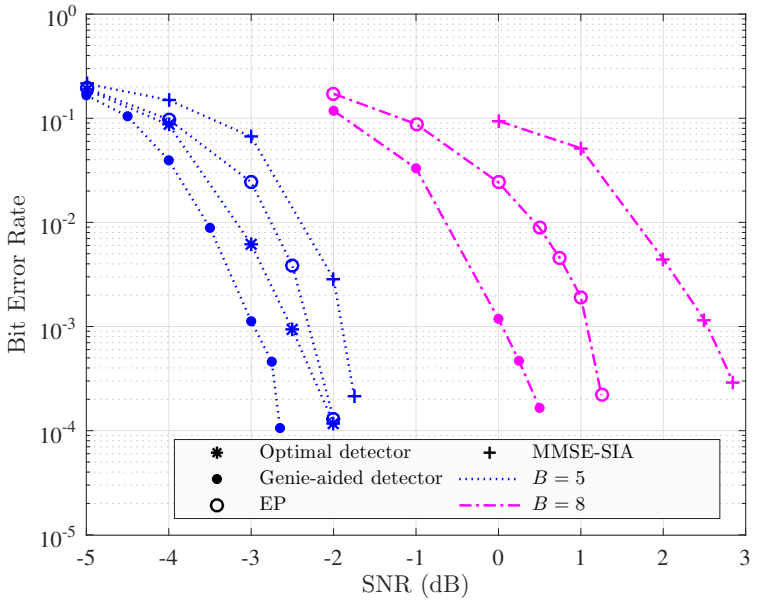
FIGURE 7.8: The bit error rate with turbo codes of EP, EPAK (with $t_0 = 2$), MMSE-SIA, POCIS, and the optimal/genie-aided detector for $B = 8$ bits/symbol and $K = N$.

Finally, in Fig. 7.9, we consider $T = 6$, $K = 3$, $N = 4$, and compare the BER with the same turbo code for different B. For $B = 5$, both EP and MMSE-SIA have performance close to the optimal detector. Under uncorrelated fading, MMSE-SIA can be slightly better than EP. This is due to the residual effect (after damping) of the phenomenon that all the mass of $\pi_{k1}^{(i_k)}$ is concentrated on a possibly wrong symbol at early iterations, and EP may not be able to refine significantly the PMF in the subsequent iterations if the constellation

is sparse. This situation is not observed for $B = 8$, i.e., larger constellations. Also, as compared to the case $T = 6, K = 3, B = 8$ in Fig. 7.8, the performance of MMSE-SIA is significantly improved as the number of receive antennas increases from $N = 3$ to $N = 4$. As in the previous case, EPAK does not perform well.



(a) Uncorrelated fading



(b) Correlated fading

FIGURE 7.9: The bit error rate with turbo codes of EP, EPAK (with $t_0 = 2$), MMSE-SIA, POCIS, and the optimal/genie-aided detector for $T = 6$, $K = 3$, and $N = 4$.

7.9 Closing Remarks

We proposed an EP based scheme and two simplifications (EPAK and MMSE-SIA) of this scheme for multi-user detection in non-coherent SIMO multiple access channel with spatially correlated Rayleigh fading. EP and MMSE-SIA are shown to achieve good performance in terms of mismatched sum-rate, symbol error rate when they are used for hard detection, and bit error rate when they are used for soft-input soft-output channel decoding. EPAK has acceptable performance with uncorrelated fading. It performs well for hard symbol detection but inadequately for soft-output detection. While MMSE-SIA and EPAK have lower complexity than EP, the performance gain of EP with respect to MMSE-SIA and EPAK is more significant when the number of users and/or the constellation size increase.

7.A Appendices

7.A.1 Exponential Family

The exponential family [202, Section 1.6] is a common statistical model for inference. The exponential family is a set of probability distributions whose PDF (or PMF) can be parameterized as

$$p_{\mathbf{s}}(\boldsymbol{s}|\boldsymbol{\gamma}) = h(\boldsymbol{s}) \exp\left(\boldsymbol{\gamma}^{\mathsf{T}} \boldsymbol{\phi}(\boldsymbol{s}) - A(\boldsymbol{\gamma})\right), \tag{7.120}$$

where $\mathbf{s} = [\mathsf{s}_1 \ \ldots \ \mathsf{s}_n]^{\mathsf{T}}$ is a n-dimensional complex random vector with realization $\boldsymbol{s} \in \mathbb{C}^n$, $\boldsymbol{\gamma} \in \mathbb{C}^m$ is a parameter vector (m can be different from n), $h(\boldsymbol{s})$ is a scalar function and $\boldsymbol{\phi}(\boldsymbol{s}) = \{\phi_i(\boldsymbol{s})\}_{i=1}^m$ is a vector of functions evaluated at $\boldsymbol{s}$, and we define the log-partition function

$$A(\boldsymbol{\gamma}) := \ln \int_{\mathbb{C}^n} h(\boldsymbol{s}) \exp\left(\boldsymbol{\gamma}^{\mathsf{T}} \boldsymbol{\phi}(\boldsymbol{s})\right) \mathrm{d}\boldsymbol{s}. \tag{7.121}$$

Here, $\boldsymbol{\phi}(\mathbf{s})$ is called the sufficient statistics of $\mathbf{s}$.

An example of a distribution in the exponential family is the discrete distribution

$$p_{\mathbf{s}}(\boldsymbol{s}) = \prod_{i=1}^{n} \pi_i(s_i) \mathbb{1}\{s_i \in \mathcal{S}_i\}, \tag{7.122}$$

for $\boldsymbol{s} = [s_1, \ldots, s_n]^{\mathsf{T}}$, where $\mathcal{S}_i$ is the discrete domain of the component s_i and $\pi_i(.)$ such that (such that) $\sum_{s \in \mathcal{S}_i} \pi_i(s) = 1$ defines a marginal PMF of s_i. This can be deduced from (7.120) by taking $h(\boldsymbol{s}) = \prod_{i=1}^n \mathbb{1}\{s_i \in \mathcal{S}_i\}$, $\phi_i(\boldsymbol{s}) = \ln(\pi_i(s_i))$, $i \in [n]$, $\boldsymbol{\gamma} = \mathbf{1}_n$, and $A(\boldsymbol{\gamma}) = 0$.

Another example is the Gaussian vector distribution denoted by $\mathbf{s} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu}$ is the mean vector and $\boldsymbol{\Sigma}$ is the covariance matrix. It has PDF

$$\mathcal{N}_{\mathbb{C}}(\boldsymbol{s}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) := \frac{1}{\pi^n \det(\boldsymbol{\Sigma})} \exp\left(-(\boldsymbol{s}-\boldsymbol{\mu})^{\mathsf{H}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{s}-\boldsymbol{\mu})\right), \ \boldsymbol{s} \in \mathbb{C}^n. \tag{7.123}$$

This distribution is generated from (7.120) with $h(\boldsymbol{s}) = 1$ for all $\boldsymbol{s} \in \mathbb{C}^n$ and

$$\boldsymbol{\phi}(\boldsymbol{s}) = \left[\{\mathsf{s}_i\}_{i=1}^n, \{\mathsf{s}_i^*\}_{i=1}^n, \{\mathsf{s}_i^* \mathsf{s}_j\}_{1 \le i \le j \le n}\right]^{\mathsf{T}}, \tag{7.124}$$

$$\boldsymbol{\gamma} = \left[\left\{\sum_{j=1}^n \sigma_{ij} \mu_j^*\right\}_{i=1}^n, \left\{\sum_{j=1}^n \sigma_{ji} \mu_j\right\}_{i=1}^n, \{2\sigma_{ij}\}_{1 \le i \le j \le n}\right]^{\mathsf{T}}, \tag{7.125}$$

$$A(\boldsymbol{\gamma}) = \ln \det(\boldsymbol{\Sigma}) + \boldsymbol{\mu}^{\mathsf{H}} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + n \ln \pi, \tag{7.126}$$

where $\boldsymbol{\Sigma}^{-1} = \{\sigma_{i,j}\}_{i,j=1,\ldots,n}$.

7.A.2 Properties of the Gaussian PDF

Lemma 7.1. *Let* $\mathbf{s}$ *be an* n*-dimensional complex Gaussian vector. It holds that*

1. $\mathcal{N}_{\mathbb{C}}(\boldsymbol{s};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \mathcal{N}_{\mathbb{C}}(\boldsymbol{s}+\boldsymbol{y};\boldsymbol{\mu}-\boldsymbol{y},\boldsymbol{\Sigma})$ *for* $\boldsymbol{y}\in\mathbb{C}^n$*;*
2. *Gaussian PDF multiplication rule:*

$$\mathcal{N}_{\mathbb{C}}(\boldsymbol{s};\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)\mathcal{N}_{\mathbb{C}}(\boldsymbol{s};\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2) = \mathcal{N}_{\mathbb{C}}(\boldsymbol{s};\boldsymbol{\mu}_{\text{new}},\boldsymbol{\Sigma}_{\text{new}})\mathcal{N}_{\mathbb{C}}(\mathbf{0};\boldsymbol{\mu}_1-\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2), \quad (7.127)$$

where $\boldsymbol{\Sigma}_{\text{new}} := \left(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_2^{-1}\right)^{-1}$ *and* $\boldsymbol{\mu}_{\text{new}} := \boldsymbol{\Sigma}_{\text{new}}(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2)$.

Proof. The first part follows readily from the definition of $\mathcal{N}_{\mathbb{C}}(\boldsymbol{s};\boldsymbol{\mu},\boldsymbol{\Sigma})$ in (7.123). The complex Gaussian PDF multiplication rule is a straightforward generalization of the real counterpart [209]. Specifically, we express the Gaussian PDF as

$$\mathcal{N}_{\mathbb{C}}(\boldsymbol{s};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \frac{1}{\pi^n \det(\boldsymbol{\Sigma})}\exp(-\boldsymbol{s}^{\mathsf{H}}\boldsymbol{\Sigma}^{-1}\boldsymbol{s}+\boldsymbol{s}^{\mathsf{H}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}+\boldsymbol{\mu}^{\mathsf{H}}\boldsymbol{\Sigma}^{-1}\boldsymbol{s}-\boldsymbol{\mu}^{\mathsf{H}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}). \quad (7.128)$$

Then

$$\begin{aligned}
&\mathcal{N}_{\mathbb{C}}(\boldsymbol{s};\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)\mathcal{N}_{\mathbb{C}}(\boldsymbol{s};\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)\\
&= \frac{1}{\pi^{2n}\det(\boldsymbol{\Sigma}_1)\det(\boldsymbol{\Sigma}_2)}\exp\big(-\boldsymbol{s}^{\mathsf{H}}(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_2^{-1})\boldsymbol{s}+\boldsymbol{s}^{\mathsf{H}}(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2)\\
&\quad+(\boldsymbol{\mu}_1^{\mathsf{H}}\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\mu}_2^{\mathsf{H}}\boldsymbol{\Sigma}_2^{-1})\boldsymbol{s}-\boldsymbol{\mu}_1^{\mathsf{H}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1-\boldsymbol{\mu}_2^{\mathsf{H}}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2\big) \quad (7.129)\\
&= \frac{\exp\big(-\boldsymbol{s}^{\mathsf{H}}\boldsymbol{\Sigma}_{\text{new}}^{-1}\boldsymbol{s}+\boldsymbol{s}^{\mathsf{H}}\boldsymbol{\Sigma}_{\text{new}}^{-1}\boldsymbol{\mu}_{\text{new}}+\boldsymbol{\mu}_{\text{new}}^{\mathsf{H}}\boldsymbol{\Sigma}_{\text{new}}^{-1}\boldsymbol{s}-\boldsymbol{\mu}_1^{\mathsf{H}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1-\boldsymbol{\mu}_2^{\mathsf{H}}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2\big)}{\pi^{2n}\det(\boldsymbol{\Sigma}_1)\det(\boldsymbol{\Sigma}_2)} \quad (7.130)\\
&= C(\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2)\frac{\exp\big(-(\boldsymbol{s}-\boldsymbol{\mu}_{\text{new}})^{\mathsf{H}}\boldsymbol{\Sigma}_{\text{new}}^{-1}(\boldsymbol{s}-\boldsymbol{\mu}_{\text{new}})\big)}{\pi^n\det(\boldsymbol{\Sigma}_{\text{new}})} \quad (7.131)\\
&= C(\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2)\mathcal{N}_{\mathbb{C}}(\boldsymbol{s};\boldsymbol{\mu}_{\text{new}},\boldsymbol{\Sigma}_{\text{new}}), \quad (7.132)
\end{aligned}$$

where $\boldsymbol{\mu}_{\text{new}}$ and $\boldsymbol{\Sigma}_{\text{new}}$ were defined above; in the second equality, we used the fact that $\boldsymbol{\Sigma}_{\text{new}}^{-1} = \boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_2^{-1}$ and $\boldsymbol{\Sigma}_{\text{new}}^{-1}\boldsymbol{\mu}_{\text{new}} = \boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2$; and the scaling factor is defined as

$$C(\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2) := \frac{\exp\left(\boldsymbol{\mu}_{\text{new}}^{\mathsf{H}}\boldsymbol{\Sigma}_{\text{new}}^{-1}\boldsymbol{\mu}_{\text{new}}-\boldsymbol{\mu}_1^{\mathsf{H}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1-\boldsymbol{\mu}_2^{\mathsf{H}}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2\right)}{\pi^n\det\left(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_{\text{new}}^{-1}\boldsymbol{\Sigma}_2\right)}. \quad (7.133)$$

After some manipulations, we deduce that $C(\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2) = \mathcal{N}_{\mathbb{C}}(\mathbf{0};\boldsymbol{\mu}_1-\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)$. □

7.A.3 Proof of Proposition 7.1

Using the natural logarithm for the KL divergence, we can derive that

$$
\begin{aligned}
D(q_\alpha(\boldsymbol{s})\|\underline{p}(\boldsymbol{s})) &= \int q_\alpha(\boldsymbol{s}) \ln \frac{q_\alpha(\boldsymbol{s})}{\prod_\beta \underline{p}_\beta(\boldsymbol{s}_\beta)} \,\mathrm{d}\boldsymbol{s} && (7.134)\\
&= \sum_\beta \int q_\alpha(\boldsymbol{s}) \ln \frac{1}{\underline{p}_\beta(\boldsymbol{s}_\beta)} \,\mathrm{d}\boldsymbol{s} + c_0 && (7.135)\\
&= \sum_{\beta\in\mathfrak{N}_\alpha} \int q_\alpha(\boldsymbol{s}) \ln \frac{1}{\underline{p}_\beta(\boldsymbol{s}_\beta)} \,\mathrm{d}\boldsymbol{s} + \sum_{\beta\notin\mathfrak{N}_\alpha} \int q_\alpha(\boldsymbol{s}) \ln \frac{1}{\underline{p}_\beta(\boldsymbol{s}_\beta)} \,\mathrm{d}\boldsymbol{s} + c_0 && (7.136)\\
&= \sum_{\beta\in\mathfrak{N}_\alpha} \int q_\alpha(\boldsymbol{s}) \ln \frac{1}{\underline{p}_\beta(\boldsymbol{s}_\beta)} \,\mathrm{d}\boldsymbol{s} + \sum_{\beta\notin\mathfrak{N}_\alpha} \int \hat{p}_\beta(\boldsymbol{s}_\beta) \ln \frac{1}{\underline{p}_\beta(\boldsymbol{s}_\beta)} \,\mathrm{d}\boldsymbol{s}_\beta + c_0 && (7.137)\\
&= -\sum_{\beta\in\mathfrak{N}_\alpha} \int q_\alpha(\boldsymbol{s}) \Big[\underline{\boldsymbol{\gamma}}_\beta^{\mathsf{T}} \boldsymbol{\phi}(\boldsymbol{s}_\beta) - A_\beta(\underline{\boldsymbol{\gamma}}_\beta)\Big] \,\mathrm{d}\boldsymbol{s} + \sum_{\beta\notin\mathfrak{N}_\alpha} D(\hat{p}_\beta\|\underline{p}_\beta) + c_0 && (7.138)\\
&= \sum_{\beta\in\mathfrak{N}_\alpha} \Big[A_\beta(\underline{\boldsymbol{\gamma}}_\beta) - \underline{\boldsymbol{\gamma}}_\beta^{\mathsf{T}} \mathbb{E}_{q_\alpha}[\boldsymbol{\phi}(\boldsymbol{s}_\beta)]\Big] + \sum_{\beta\notin\mathfrak{N}_\alpha} D(\hat{p}_\beta\|\underline{p}_\beta) + c_0, && (7.139)
\end{aligned}
$$

where (7.137) follows from $q_\alpha(\boldsymbol{s}) = \frac{\psi_\alpha(\boldsymbol{s}_\alpha)}{m_\alpha(\boldsymbol{s}_\alpha)} \Big[\prod_{\beta\in\mathfrak{N}_\alpha} \hat{p}_\beta(\boldsymbol{s}_\beta)\Big]\Big[\prod_{\beta\notin\mathfrak{N}_\alpha} \hat{p}_\beta(\boldsymbol{s}_\beta)\Big]$, and (7.138) follows from (7.19). From (7.139), we can see that the optimization (7.24) of $\underline{p}$ decouples over $\underline{p}_\beta$, and the optimal distribution can be expressed as $\hat{p}_\alpha^{\text{new}}(\boldsymbol{s}) = \prod_\beta \hat{p}_{\alpha,\beta}^{\text{new}}(\boldsymbol{s}_\beta)$.

For $\beta \notin \mathfrak{N}_\alpha$, the minimum of $D(\hat{p}_\beta\|\underline{p}_\beta)$ is simply 0 and achieved with $\hat{p}_{\alpha,\beta}^{\text{new}}(\boldsymbol{s}_\beta) = \hat{p}_\beta(\boldsymbol{s}_\beta)$.

For $\beta \in \mathfrak{N}_\alpha$, since the log-partition function $A_\beta(\underline{\boldsymbol{\gamma}}_\beta)$ is convex in $\underline{\boldsymbol{\gamma}}_\beta$ (see, e.g., [210, Lemma 1]), the minimum of $A_\beta(\underline{\boldsymbol{\gamma}}_\beta) - \underline{\boldsymbol{\gamma}}_\beta^{\mathsf{T}} \mathbb{E}_{q_\alpha}[\boldsymbol{\phi}(\boldsymbol{s}_\beta)]$ is achieved at the value of $\underline{\boldsymbol{\gamma}}_\beta$ where its gradient is zero. Using the well-known property of the log-partition function, $\nabla_{\boldsymbol{\gamma}} A_\beta(\boldsymbol{\gamma}) = \mathbb{E}_{\hat{p}_\beta}[\boldsymbol{\phi}_\beta(\boldsymbol{\gamma})]$, we get that the zero-gradient equation is equivalent to the moment matching criterion $\mathbb{E}_{\hat{p}_{\alpha,\beta}^{\text{new}}}[\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta)] = \mathbb{E}_{q_\alpha}[\boldsymbol{\phi}_\beta(\boldsymbol{s}_\beta)]$.

Chapter 8

Conclusions and Outlook

8.1 Conclusions

Non-coherent wireless communications assume that the communicating nodes do not know the instantaneous channel state information in advance, and only know the channel distribution. This thesis contributes to the understanding of fundamental limits of non-coherent communications, as well as to the design of non-coherent communication systems in block fading, with a focus on the P2P channel, the MAC, and the BC. The main results and corresponding possible extensions of the thesis are summarized as follows.

Fundamental Limits of Non-Coherent Communications

- In Chapter 2, we found the optimal DoF for the MIMO P2P channel under generic block fading. Possible extensions include analyzing the finite-SNR achievable rate or capacity, in particular, quantifying the constants after the logarithmic term. To this end, the duality approach can be still useful, but one would need to refine the choice of auxiliary input distribution. Furthermore, as far as the constant term is concerned, one would need to rely on the escape-to-infinity property of the optimal input.
- In Chapter 3, we found the optimal DoF region for the two-user SIMO MAC under generic block fading, where the corner points can be achieved by simple pilot-based schemes. It would be interesting to extend to the MIMO and/or more-than-two-user case. Finite-SNR characterization of the capacity region would be challenging, but one can attempt to analyze the sum-rate capacity.
- In Chapter 4, we derive achievable rate and DoF regions for the MIMO BC under spatially correlated generic block fading. In doing so, we propose novel transmission schemes based on rate splitting, product superposition, and a combination of them in order to effectively exploit the transmit correlation diversity. Although these achievable regions improve significantly upon TDMA, one needs to characterize the outer bounds to see how far these regions are from optimality. One approach is to use the cooperative bound by letting users cooperate and considering them as one macro user. However, the existing results for the P2P channel, such as [117, Theorem 1], do not apply immediately because the channel columns are not statistically equivalent (due to the difference in their correlation matrices). Our attempts using

the duality approach show that it is challenging to choose an adequate auxiliary output distribution.

Transceiver Design for Non-Coherent Communications

- In Chapter 5, we propose a structured Grassmannian constellation for the SIMO P2P channel under Rayleigh block fading. This so-called cube-split constellation has a good distance properties and is suitable for practical uses because it is simple to generate—thus available for large constellation size and large symbol length and does not require to be stored, it admits a simple yet effective binary labeling scheme, and its structure enables efficient hard and soft detection. A future work is to extend it to the MIMO case where the main challenge is to find an efficient way to partition the Grassmannian.
- In Chapter 6, we propose simple and effective metrics for joint constellation design for the MIMO MAC under Rayleigh block fading so as to minimize the detection error. Our metrics can be used to generate well-performing joint constellations through numerical optimization, where the individual constellations can be Grassmannian. They can also be used as benchmark to design joint constellations with a structure that facilitates efficient detection as in the single-user case. We have demonstrated the latter use with a precoding-based design. In the future, it is desirable to design more effective structured joint constellations.
- In Chapter 7, we propose efficient multi-user soft detectors for the SIMO MAC under spatially correlated Rayleigh block fading. Our detectors are based on expectation propagation, a message passing approximate inference framework in probabilistic graphical models. They achieve good performance in terms of mismatched sum-rate, symbol error rate when they are used for hard detection, and bit error rate when they are used for soft-input soft-output channel decoding. Possible extensions include considering more complicated fading models and analyzing theoretically the performance of EP for non-coherent detection.

8.2 Outlook

In this section, we provide a further outlook of non-coherent communications in the context of other topics in wireless communications.

- **Non-coherent communications in continuously varying channel:** In block fading, it is assumed that the channel coefficients are fully correlated (i.e., equal) within a coherence block and uncorrelated across the blocks. In practice, the channel coefficients are highly correlated but can vary within a block. Furthermore, the coefficients separated by a coherence interval are weakly correlated, but two adjacent coefficients across two adjacent blocks can be highly correlated. This can be observed in both the time domain and frequency domain (with OFDM). In this scenario, the DoF was studied in [211], [212]. In addition, Grassmannian signaling becomes sub-optimal, and the block-fading ML detection becomes a mismatched detection. In would be interesting to identify how "mismatched" Grassmannian constellation and detection are, as has been investigated in [81] in terms of the error rate. This mismatch can also be analyzed in terms of the mismatch data rate. The per-block ML detection can be modified by taking into account the correlation of channel coefficients within a block. On the other hand, we can try to adapt Grassmannian constellations to continuously varying channel. For example, leveraging the fact that

a Grassmannian symbol is invariant to scaling by a complex scalar, one can encode some additional information bits in the phase shift of the symbols across coherence blocks. We have reported this idea in [124].

- **Using Grassmannian constellations for more sophisticated system design**: The design of Grassmannian constellations focus on the bit-to-symbol mapping and symbol-to-bit demapping of a digital communication system. More general system design using Grassmannian constellations as an ingredient can be done. For example, in [213]–[215], multi-layer coding schemes are proposed in which Grassmannian constellations are used in the lower-resolution layer that does not have access to CSI. The integration of Grassmannian constellations into LTE networks was investigated with a link-level performance study in [216]. Bit-interleaved coded modulation systems using Grassmannian constellations have been designed in [80], [177].

- **Non-coherent low-latency communications:** In this thesis, we assume that an information-carrying message, represented by a channel codeword, spans a large number of coherence blocks. Therefore, we employ asymptotic metrics such as the Shannon capacity and the DoF. In ultra-reliable low-latency communication (URLLC), information is carried in small packets that should be received with low latency. In this regime, non-coherent communication is suitable since the size of the metadata (e.g., control information, pilots) is comparable to the size of data [217]. However, the classic asymptotic information-theoretic benchmarks are not applicable, and one needs to switch to a framework that provides understanding of short-packet transmissions. Finite-blocklength information theory [218] is appropriate for that purpose, where the metric of interest is the *maximum coding rate* $R^*(n, \epsilon)$ at finite blocklength[1] n and finite packet error probability ϵ. Non-asymptotic bounds on the non-coherent maximal coding rate have been proposed for the Rayleigh quasi-static fading channel (a channel codeword spans a single coherence block) for the SISO case in [219]. In block fading channel (a channel codeword spans multiple, but a finite number of, coherence blocks), non-asymptotic bounds for the non-coherent maximal coding rate was presented for the SISO case in [220] (Rayleigh fading) and [187] (Rician fading), and the MIMO case in [186]. A non-coherent communication scheme in which information is modulated onto the zeros of the transmitted baseband signal's z-transform was proposed for short-packet communication in [221].

- **Non-coherent machine-type communication and massive random access**: In machine-type communication and massive random access, the system has to support a very large number of users and only a random subset of them are active at a time. The set of active users might not be known in advance, and it is impossible to pre-assign mutually orthogonal pilots to every user. Therefore, non-coherent pilot-free communication can be a suitable strategy. A non-coherent random access protocol based on Gabor frames, which are also Grassmannian frames, was proposed in [222]. In [223], Senel *et al.* proposed a non-coherent transmission scheme for massive machine-type communication and developed a modified approximate message passing algorithm to exploit the structured sparsity of this scheme. Compared with coherent transmission, this non-coherent scheme performs significantly better, especially when the number of bits is small. Recently, our cube-split constellation has been used to construct a tensor-based unsourced massive random access scheme in [224]. Furthermore, in the context of machine-type communication systems, one

[1]The term "block" in *finite blocklength* refers to a channel codeword and is different from that in *coherence block*: in finite-blocklength information theory, a codeword can span multiple coherence blocks.

may be interested in other performance metrics than capacity or error probability, such as, e.g., information freshness.

- **Machine learning for non-coherent communications:** The use of machine learning, especially deep learning, in wireless communications is growing rapidly and profoundly in many aspects of physical-layer communications, from using learning for a specific task such as channel estimation, symbol equalizer, signal detection, channel coding/decoding, to a complete end-to-end learning-based transmission [225]–[227]. Non-coherent communications can certainly benefit from machine learning. A data-driven machine learning method was proposed to design non-coherent transceiver for short-packet transmission in [228]. Learning methods on the Grassmann manifold can be used for non-coherent communication with Grassmannian signaling [229]. For example, Du *et al.* employed data clustering algorithms on the Grassmann manifold to design an automatic modulation recognition scheme for Grassmannian constellations in [230]. In end-to-end learning of communication systems through neural network based autoencoders, channel estimation can be performed even with an unknown channel model [231]. When the channel model is unknown, one can use data-driven learning to fit the empirical input-output conditional probability to a parametric model. We proposed a framework using generalized Gaussian model for such purpose in [130]. An autoencoder for non-coherent energy-based communication in a multi-carrier multi-user SIMO system was proposed in [232]. The paper [233] presents waveform design for non-coherent multi-user MIMO systems through deep learning.

Appendix A

Resumé en Français

A.1 Communications Sans Fil

A.1.1 L'Histoire et l'Évolution des Communications Sans Fil

La communication sans fil [2]–[5] est communément considérée comme le transfert *électromagnétique* d'informations entre des points qui ne sont pas connectés par un conducteur électrique [6]. Le transfert électromagnétique d'informations a commencé lorsque James C. Maxwell a postulé la transmission des ondes électromagnétiques en 1864, puis Heinrich Hertz l'a vérifiée et démontrée en 1880 et 1887, respectivement. Cet effet permet d'utiliser les ondes électromagnétiques comme supports d'information à l'aide d'appareils électroniques au niveau de l'émetteur (source) et du récepteur (destination). Marconi a mis en place un télégraphe sans fil et breveté un système sans fil complet en 1897. Avec le développement des circuits intégrés, la communication sans fil électromagnétique s'est développée rapidement à mesure que la radiodiffusion et la télévision se répandaient dans le monde entier. Les systèmes sans fil sont passés de la transmission de signaux analogiques à la transmission de signaux numériques composés de bits, ancrée dans le travail fondateur de Claude Shannon en 1948 [7] et déployée dans les années 1980.

Depuis les années 1980, l'évolution des systèmes sans fil mobiles progresse d'une génération à l'autre tous les dix ans environ. Chaque génération présente une réglementation, des services et des innovations différents [8], [9]. La cinquième génération (5G) [10], [11] est en cours de développement depuis le début des années 2010. Les réseaux 5G sont censés prendre en charge un grand nombre et une grande hétérogénéité des terminaux, c'est-à-dire l'Internet des objets (IoT) [12]. Les trois principaux cas d'utilisation visés par la 5G sont les communications mobiles à large bande améliorées (eMBB), les communications ultra-fiables à faible latence (URLLC) et les communications massives de type machine (mMTC) [13]. Les réseaux 5G ont été testés dans de nombreux pays et en sont maintenant aux premiers stades du déploiement commercial [14]. Entre-temps, les activités de recherche vers la sixième génération (6G) ont été lancées. Bien que l'on ne sache pas exactement ce que sera la 6G, de nombreuses visions (spéculatives) pour la 6G ont été fournies sous différents angles, par exemple dans [15]–[20].

A.1.2 Propagation Sans Fil

Les ondes électromagnétiques sont émises par une antenne au niveau de l'émetteur, se propagent à travers un canal sans fil et sont interceptées par une antenne au niveau du récepteur. En principe, on pourrait résoudre les équations du champ électromagnétique

de Maxwell pour trouver les ondes reçues à l'antenne de réception. Cependant, comme c'est trop complexe, des modèles plus simples sont utilisés pour approximer la propagation du signal. Une onde transmise est réfléchie, diffusée et diffractée lorsqu'elle interagit avec des objets de l'environnement lorsqu'elle se propage vers le destinataire. Par conséquent, le récepteur observe plusieurs copies de cette onde à différents délais, chacune subissant une atténuation spécifique. En d'autres termes, l'onde parcourt plusieurs trajets, s'atténue dans chaque trajet et arrive à l'antenne de réception avec différents retards. De plus, en raison du mouvement relatif de l'émetteur, du récepteur et des objets dans l'environnement, les facteurs d'atténuation, c'est-à-dire les gains de canal, varient dans le temps. Cet effet est appelé *évanouissement*. Soit $a_n(t)$ le gain de canal complexe du trajet n au temps t. Considérons une transmission d'un signal $x(t)$ au temps t et à la fréquence porteuse f_c, le signal reçu peut être exprimé comme [4, Eq.(3.5)]

$$y(t) = \left(\int_{-\infty}^{\infty} h(\tau, t) x(t-\tau) \, \mathrm{d}\tau \right) e^{\jmath 2\pi f_c t}, \tag{A.1}$$

où $n = 0$ correspond au trajet en visibilité directe, $N_p(t)$ est le nombre de composants à trajets multiples résolvables, $\tau_n(t)$ est le retard du n-ème trajet au temps t, et

$$h(\tau, t) := \sum_{n=0}^{N_p(t)} e^{-\jmath 2\pi f_c \tau_n(t)} a_n(t) \delta(\tau - \tau_n(t)) \tag{A.2}$$

est la réponse impulsionnelle du canal au temps t et au retard τ. De cette façon, les équations de Maxwell sont remplacées par une relation entrée–sortie d'un système linéaire variant dans le temps.

En général, en raison de la nature aléatoire de l'environnement, le gain d'atténuation $a_n(t)$, le retard $\tau_n(t)$ et le nombre de trajets $N_p(t)$ sont aléatoires. Par conséquent, la réponse impulsionnelle de canal $h(\tau, t)$ est modélisée comme la réalisation d'une variable aléatoire $\mathsf{h}(\tau, t)$. On suppose généralement que $\mathsf{h}(\tau, t)$ est un processus Gaussien, lorsque le nombre de trajets $N_p(t)$ est grand, en évoquant le théorème de la limite centrale. Une autre hypothèse courante est que la phase de chaque composant à trajets multiples est uniformément distribuée. Si les statistiques conjointes de $\mathsf{h}(\tau_1, t_1)$ et $\mathsf{h}(\tau_2, t_2)$ ne dépendent que de la différence de temps $\Delta t = t_2 - t_1$, le canal est dit stationnaire au sens large (WSS). De plus, si la réponse de canal d'une composante à trajets multiples donnée à différents retards n'est pas corrélée, le canal n'a pas de corrélation diffusion (US).

Dans le domaine fréquentiel, la réponse en fréquence du canal est donnée par la transformée de Fourier de $\mathsf{h}(\tau, t)$ par rapport à (w.r.t.) τ [4, Eq.(3.57)]

$$\mathsf{H}(f, t) := \int_{-\infty}^{\infty} \mathsf{h}(t, \tau) e^{-\jmath 2\pi f t} \, \mathrm{d}\tau. \tag{A.3}$$

Étant donné que $\mathsf{h}(t, \tau)$ est un processus Gaussien WSSUS (c'est-à-dire WSS et US), $\mathsf{H}(f, t)$ est également un processus Gaussien WSSUS avec fonction d'autocorrélation

$$A_{\mathsf{H}}(\Delta f; \Delta t) := \mathbb{E}[\mathsf{H}^*(f, t) \mathsf{H}(f + \Delta f, t + \Delta t)], \tag{A.4}$$

ce qui ne dépend que de la différence de temps Δt et de la différence de fréquence Δ_f. Les réponses en fréquence à deux canaux au même instant, c'est-à-dire $\Delta t = 0$, et la séparation de fréquence Δf sont approximativement indépendantes si $A_{\mathsf{H}}(\Delta f; 0) \approx 0$. La fréquence B_c où $A_{\mathsf{H}}(\Delta f; 0) \approx 0$ pour tous $\Delta f > B_c$ est appelée *largeur de bande de cohérence* du canal. En général, si le signal transmis a une bande passante étroite $B \ll B_c$, alors la réponse du canal est à peu près constante sur l'ensemble du signal bande passante. C'est

ce qu'on appelle la *l'évanouissement plate*. Au contraire, si $B \gg B_c$, alors la réponse du canal varie largement à travers la bande passante du signal. C'est ce que l'on appelle *l'évanouissement sélectif en fréquence*. En revanche, l'autocorrélation $A_{\mathsf{H}}(\Delta f; \Delta t)$ pour $\Delta f = 0$ caractérise la décorrélation des réponses des canaux dans le temps. Si $B \gg B_c$, alors le canal mesuré à des instants temporels séparés par t est approximativement non corrélé et donc indépendant. La durée T_c où $A_{\mathsf{H}}(0; \Delta t) \approx 0$ pour tout $\Delta t > T_c$ est appelée *temps de cohérence* du canal. Elle est inversement proportionnelle à la propagation Doppler. Un bloc de largeur de bande de cohérence B_c et de temps de cohérence T_c est appelé *bloc de cohérence*, et la longueur totale $T = T_c B_c$ d'un bloc de cohérence est appelée *intervalle de cohérence*.

Par un décalage dans le domaine fréquentiel (c'est-à-dire une conversion vers le bas) du signal reçu (A.1), nous avons une représentation en bande de base équivalente. De plus, en échantillonnant et en considérant un bruit additif, nous obtenons un modèle de bande de base à temps discret donné par [5, Eq.(2.39)]

$$\mathsf{y}[m] = \sum_l \mathsf{h}_l[m]\mathsf{x}[m-l] + \mathsf{z}[m], \tag{A.5}$$

où $\mathsf{y}[m]$ et $\mathsf{x}[m]$ sont respectivement les échantillons du signal reçu et transmis en bande de base, $\mathsf{z}[\mathsf{m}]$ est le bruit filtré passe-bas, et $\mathsf{h}_l[m]$ est le l-ème prise de filtre de canal complexe. Ici, $\mathsf{h}_l[m]$ et $\mathsf{z}[m]$ sont aussi normalement supposés être un processus Gaussien (discret).

A.1.3 Le Canal MIMO

Nous supposons maintenant que l'émetteur est équipé de M antennes et le récepteur de N antennes. Nous supposons un canal plat en fréquence et utilisons un modèle de canal statistique comme dans la sous-section précédente avec une représentation en bande de base en temps discret à une seule pression. Le signal en bande de base reçu est

$$\mathbf{y}[m] = \mathbf{H}[m]\mathbf{x}[m] + \mathbf{z}[m], \tag{A.6}$$

où $\mathbf{H}[m] \in \mathbb{C}^{N \times M}$ est la matrice de canal contenant les coefficients d'évanouissement des M antennes d'émission aux N antennes de réception avec une distribution arbitraire, et $\mathbf{z}[m]$ est le bruit Gaussien blanc additif (AWGN) suivant la distribution $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{I})$. Ce canal est appelé canal point à point (P2P) à entrées multiples et sorties multiples (MIMO).

A.1.3.a Limites Fondamentales du Canal MIMO

En raison de la réglementation et des limites matérielles, le signal transmis est soumis à une contrainte de puissance

$$\frac{1}{N_s} \sum_{i=1}^{N_s} \|\mathbf{x}[i]\|^2 \leq P \tag{A.7}$$

où N_s est la longueur d'un mot de code de canal, qui représente un message. On serait intéressé à trouver le débit maximal auquel les informations peuvent être transmises et reçues de manière fiable au niveau du récepteur. Cette quantité est appelée la capacité (Shannon) du canal et défini comme suit.

Definition A.1 (Capacité du canal)**.** *La capacité C du canal est le débit de données maximal, c'est-à-dire le nombre maximal d'unités d'informations nécessaires pour représenter un message normalisé par le temps de communication, de sorte qu'il existe un schéma de codage de canal réalisant une probabilité d'erreur arbitrairement faible lorsque la longueur de code N_s va à l'infini.*

Ainsi, une communication fiable n'est pas possible, même avec une longueur de code infinie, si l'on émet à un débit supérieur à la capacité du canal.

Étant donné que le canal considéré est sans mémoire, la capacité du canal est donnée par [21], [22]

$$C(P) = \max_{p_{\mathbf{x}}:\, \mathbb{E}[\|\mathbf{x}\|^2] \leq P} I(\mathbf{x};\mathbf{y}). \tag{A.8}$$

Une *distribution d'entrée qui atteint la capacité* est une solution à la maximisation ci-dessus. Si l'on transmet avec une distribution d'entrée $p_{\mathbf{x}}$ satisfaisant à la contrainte de puissance mais pas nécessairement à la capacité atteinte, alors l'information mutuelle

$$R(P) = I(\mathbf{x};\mathbf{y}) \tag{A.9}$$

est un *débit réalisable* du canal. Dans de nombreuses situations, il est très difficile de calculer la capacité du canal ou le débit réalisable, et on peut recourir à une représentation asymptotique grossière de ces quantités donnée par les *degrés de liberté* (DoF). Un DoF réalisable et le DoF optimal du canal sont respectivement définis par

$$d_{\text{réalisable}} := \lim_{P\to\infty} \frac{R(P)}{\log_2(P)} \quad \text{and} \quad d_{\text{optimal}} := \lim_{P\to\infty} \frac{C(P)}{\log_2(P)}. \tag{A.10}$$

Avec cela, le débit et la capacité réalisables se comportent dans le régime de SNR élevé comme $R(P) = d_{\text{réalisable}} \log_2(P) + o(\log_2 P)$ et $C(P) = d_{\text{optimal}} \log_2(P) + o(\log_2 P)$. Par conséquent, le DoF est également appelé le facteur de *pré-log* du débit / capacité. En gros, le DoF est le nombre de bits supplémentaires qui peuvent être transmis de manière fiable lorsque la puissance du signal est doublée.

A.1.3.b Conception Pratique pour un Canal MIMO

Conception de la constellation: Bien que la distribution d'entrée permettant d'atteindre la capacité soit souvent continue, dans la pratique, le signal transmis $\mathbf{x}$ est normalement tiré d'un ensemble discret fini $\mathcal{X}$ pour réduire la complexité. Cet ensemble est appelé la *constellation* et chaque élément est appelé *un point / symbole de constellation*. La constellation doit être conçue de manière à atteindre un débit de données élevé et / ou une faible erreur de détection.

Conception de détection efficace: Le récepteur souhaite détecter le symbole transmis. En raison du domaine discret des symboles, la détection du maximum de vraisemblance (ML) est souvent NP-difficile [24]. Pour le résoudre, il faut énumérer toute la constellation, ce qui est trop complexe si la taille de la constellation est grande. Par conséquent, d'un point de vue pratique, il est favorable d'utiliser une détection sous-optimale à faible complexité. Cela peut être fait en remplaçant la métrique de la fonction de vraisemblance par une métrique de détection sous-optimale qui peut être plus facile à calculer. Si la constellation a une structure, on peut également exploiter cette structure pour concevoir une détection efficace.

A.1.4 Communications Cohérentes: le Rôle des Informations sur l'État des Canaux

La valeur instantanée de $\mathbf{H}$ est appelée *information d'état de canal* (CSI). Si l'émetteur et / ou le récepteur dispose de ces informations, la communication est dite *cohérente*. Si le canal est bien conditionné, alors la capacité cohérente est donnée par [26], [27]

$$C(P) = \min\{M, N\} \log_2 P + O(1), \tag{A.11}$$

ainsi le DoF optimal est $\min\{M, N\}$. En fait, connaissant les réalisations de $\mathbf{H}$, le canal MIMO peut être converti en $\min\{M, N\}$ canaux parallèles, à entrée unique et à sortie unique (SISO). Par conséquent, avec CSI, la capacité évolue linéairement avec le nombre d'antennes.

Dans les communications cohérentes, on utilise normalement une constellation scalaire pour chaque composante de $\mathbf{x}$. Les constellations scalaires les plus courantes sont la modulation d'amplitude d'impulsion (PAM), la modulation par déphasage (PSK) et la modulation d'amplitude en quadrature (QAM), où les points de constellation diffèrent respectivement dans leur amplitude, phase et à la fois amplitude et phase.

Étant donné la matrice d'évanouissement de Rayleigh indépendante et identiquement distribué (IID) $\mathbf{H} = \boldsymbol{H}$, la détection ML donnée $\mathbf{y} = \boldsymbol{y}$ est équivalente au problème des moindres carrés (LS)

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}\in\mathcal{X}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2. \tag{A.12}$$

Ce problème de détection cohérent a été étudié de manière approfondie dans la littérature. Puisqu'il est NP-difficile [24], de nombreux schémas sous-optimaux ont été proposés pour réduire la complexité, notamment:

- détecteurs linéaires consistant en un démappeur par composants de $\hat{\boldsymbol{x}} = \boldsymbol{T}\boldsymbol{y}$, tel qu'un filtre adapté $\boldsymbol{T} = \boldsymbol{H}^{\mathsf{H}}$, un forçage nul (ZF) $\boldsymbol{T} = (\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathsf{H}} =: \boldsymbol{H}^{\dagger}$ (en supposant $N \geq M$) et une moyenne minimale erreur carrée (MMSE) $\boldsymbol{T} = (\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H} + \boldsymbol{I}_N)^{-1}\boldsymbol{H}^{\mathsf{H}}$;
- détecteurs aidés à l'annulation des interférences, tels que l'annulation des interférences successives / parallèles / multi-étapes / à retour de décision;
- détecteurs basés sur la recherche dans les arbres, tels que les décodeurs de sphères;
- détecteurs assistés à réduction de lattice.

En bref, CSI permet d'adapter la transmission et la réception à l'état d'évanouissement actuel. Des enquêtes complètes sur la transmission et la détection cohérentes de MIMO peuvent être trouvées respectivement dans [29] et [30].

A.1.5 Le Coût d'Acquisition de CSI

Bien que la communication cohérente exploite efficacement les ressources spatiales supplémentaires d'un canal MIMO pour augmenter l'efficacité spectrale, elle repose sur la disponibilité de CSI. En pratique, comme la matrice de canaux est aléatoire et s'estompe dans le temps et la fréquence, sa valeur n'est pas donnée *a priori* et doit être estimée. Typiquement, l'estimation de canal est effectuée en envoyant des symboles de référence, appelés *pilotes*, connus du récepteur dans certaines utilisations de canal d'un bloc de cohérence. Le récepteur estime le canal dans ces utilisations de canal en utilisant les pilotes connus, puis les interpole / extrapole pour déduire les gains de canal des utilisations de canal restantes dans le bloc de cohérence. En traitant l'estimation de canal comme le canal connu, une communication cohérente peut être effectuée dans ces utilisations de canal restantes. C'est ce qu'on appelle *un schéma / une approche pilote* [31]. Pour analyser correctement les performances du système, il faut prendre en compte le coût de l'estimation de canal et l'erreur d'estimation de canal.

Les symboles pilotes étant connus du récepteur, ils ne portent aucune information. En revanche, ils occupent une fraction des ressources temps / fréquence de communication. Dans un canal MIMO avec M antennes d'émission, il faudrait envoyer au moins M symboles

pilotes pour le récepteur afin de déterminer M vecteurs de canal correspondant aux M antennes [31]. Soit $T = T_c B_c$ l'intervalle de cohérence, alors la fraction de ressource dépensée pour l'estimation de canal est $\frac{M}{T}$ et il reste une fraction $1 - \frac{M}{T}$ du bloc de cohérence pour une transmission de données cohérente. Dans un environnement très mobile où l'état du canal change rapidement, l'intervalle de cohérence T est court et la fraction de transmission pilote peut être disproportionnée par rapport à la transmission de données, surtout si le nombre d'antennes est grande.

L'erreur d'estimation de canal a un impact négatif sur le débit du canal et les performances d'erreur. D'une part, si l'on traite l'estimation de canal comme le vrai canal et sans tenir compte de toute inexactitude, le détecteur optimal dans cette hypothèse est un détecteur qui ne correspond pas pour le canal avec une erreur d'estimation de canal, et le débit du canal est déterminé par le débit dépareillé. En revanche, même si les statistiques de l'erreur d'estimation de canal sont prises en compte, cette erreur résiduelle impose un bruit secondaire qui augmente la puissance de bruit totale et réduit la capacité du canal pour une puissance de signal donnée. L'effet d'une connaissance de canal imparfaite sur la capacité du canal a été étudié, par exemple, [32], [33] et sur l'erreur de détection, par exemple, [34], [35].

Dans certains scénarios, l'estimation de canal basée sur pilote devient difficile, voire impossible. Par exemple, dans la liaison montante d'un système multi-utilisateur, des séquences pilotes sont attribuées par utilisateur et orthogonalement entre utilisateurs. Si le nombre total d'utilisateurs est supérieur à l'intervalle de cohérence (mais probablement seul un nombre aléatoire d'utilisateurs est actif à la fois), il n'est pas possible de pré-attribuer des séquences pilotes mutuellement orthogonales à chaque utilisateur présent dans le système. On peut envisager des pilotes non orthogonaux, mais l'acquisition précise de CSI reste difficile.

La communication sans CSI a priori est dite *non-cohérente*. Ce cadre fournit un point de vue plus réaliste à l'analyse et à la conception du système. La communication non-cohérente est également un cadre plus général puisque, comme mentionné ci-dessus, on peut estimer le canal, puis effectuer une communication cohérente. Néanmoins, l'estimation des canaux et la détection de données cohérentes de manière séquentielle / conjointe pourraient ne pas être optimales. Lorsque le coût de l'estimation de canal est important, que l'erreur d'estimation de canal est grave ou que l'estimation de canal basée sur le pilote devient gênante / impossible, il peut être avantageux de s'abstenir de le faire en utilisant un schéma de communication qui ne repose pas sur la connaissance de CSI. La communication non-cohérente explique un tel schéma.

Dans cette thèse, nous nous concentrons sur les communications non-cohérentes.

A.2 Communications Sans Fil Non-Cohérentes

Dans cette section, nous fournissons un aperçu de l'état de l'art des communications sans fil non-cohérentes, ainsi que les questions auxquelles nous voulons répondre dans cette thèse.

A.2.1 Hypothèses

- **Pas de CSI:** la distribution des coefficients de canal est supposée connue, mais la valeur instantanée des coefficients de canal, c'est-à-dire CSI, n'est pas connue.

- **Évanouissement par bloc:** l'hypothèse d'évanouissement par bloc se réfère au cas où les coefficients de canal dans un bloc de cohérence de longueur $T = T_c B_c$ sont fortement corrélés de sorte qu'ils peuvent être considérés comme inchangés dans le bloc. De plus, les coefficients de canal dans différents blocs sont supposés être indépendants et identiquement distribués. Lorsque $T = 1$, cela devient un évanouissement rapide stationnaire.
- **Distribution des canaux:** le canal est censé être un évanouissement Rayleigh IID si les composants sont indépendantes et distribuées de manière Gaussienne identiquement. Plus généralement, le canal est un évanouissement *générique* si la matrice du canal a une entropie différentielle finie et un second moment fini. Le canal est corrélé spatialement si les coefficients de canal entre différentes antennes sont corrélés, c'est-à-dire que la norme et la direction de canal ne sont pas indépendants.
- **Distribution du bruit:** Nous supposons tout au long que le bruit est distribué de manière Gaussienne IID.

A.2.2 Le Canal Point à Point

Considérons un canal P2P MIMO avec M antennes d'émission et N antennes de réception, et contrainte de puissance P.

A.2.2.a Limites Fondamentales et Signal Qui Atteint la Capacité

La capacité du canal P2P non-cohérent a été étudiée dans un grand nombre de travaux. En général, la capacité explicite est inconnue et seules des approximations dans le régime du rapport signal / bruit (SNR) extrême sont connues dans certains contextes.

$T = 1$ **(évanouissement rapide stationnaire)**

Sous l'hypothèse d'évanouissement Rayleigh IID pour le canal SISO dans le fading rapide ($T = 1$), Richters a conjecturé en 1967 que, bien que le canal soit continu, la distribution des entrées de capacité est discrète [43]. Plus tard, Abou-Faycal *et al.* a prouvé cette conjecture et a montré que la distribution des intrants générateurs de capacité a un nombre fini de points de masse avec l'un d'eux situé à l'origine [44]. Dans le même cadre, Taricco et Elia ont montré que la capacité s'échelonne en tant que $\alpha \log\log P \leq C \leq \log\log P + O(1)$ pour $\alpha \in (0, 1)$ lorsque $P \to \infty$ [45].

La mise à l'échelle logarithmique double de la capacité w.r.t. le SNR a été prouvé rigoureusement dans un cadre plus général par Lapidoth et Moser en [36]. En supposant d'évanouissement générique, pour limiter la information mutuelle, ils ont utilisé une double expression et remplacé la maximisation sur la distribution d'entrée du canal par une minimisation sur la distribution de sortie du canal. Avec cette approche de dualité, ils ont prouvé que les capacités évolue en

$$C = \log\log P + \chi(\mathbf{H}) + o(1), \tag{A.13}$$

où $\chi(\mathbf{H})$ est ce que l'on appelle le *nombre d'évanouissement* du canal. Par conséquent, en cas d'évanouissement rapide, la communication à des débits nettement supérieurs au nombre d'évanouissements est extrêmement inefficace en termes de puissance.

$T \geq 2$ **(évanouissement par bloc)**

Lorsque le canal reste constant pendant au moins deux utilisations de canal, l'une des premières études pour aborder la capacité a été réalisée par Marzetta et Hochward [47].

En supposant d'évanouissement Rayleigh IID, ils ont montré deux résultats importants comme suit.

- La capacité obtenue avec $M > T$ est la même que la capacité obtenue avec $M = T$ antennes d'émission pour tout T et N et SNR arbitraire. Cela contraste avec la croissance linéaire illimitée de la capacité avec $\min\{M, N\}$ lorsque CSI est disponible.
- Le signal qui atteint la capacité peut être représenté comme

$$\mathbf{X} = \mathbf{D}\mathbf{\Phi}, \tag{A.14}$$

où $\mathbf{D} \in \mathbb{R}^{M\times M}$ est une matrice diagonale, non négative, et $\mathbf{\Phi} \in \mathbb{C}^{M\times T}$ est une matrice unitaire tronquée et distribuée de façon isotrope. Autrement dit, la densité de probabilité de $\mathbf{\Phi}$ est inchangée lorsqu'elle est postmultipliée par une matrice unitaire $T \times T$.

Ces résultats ont ensuite été généralisés et affinés par Zheng et Tse dans [49] pour le cas $T \geq \min\{M, N\} + N$ et par Yang *et al.* dans [50] pour le cas $T \leq M + N, M \leq \min\{N, \lfloor T/2 \rfloor\}$. Ils ont collectivement montré que, sous l'évanouissement Rayleigh IID, la capacité dans le régime à SNR élevé est

$$C = M^*\left(1 - \frac{M^*}{T}\right)\log P + O(1), \tag{A.15}$$

où $M^* := \min\{M, N, \lfloor T/2 \rfloor\}$. Ainsi, le DoF optimal est $M^*\left(1 - \frac{M^*}{T}\right)$. La distribution d'entrée qui atteint la capacité au terme de second ordre à SNR élevé hérite de la structure en (A.14) avec les distributions spécifiées de $\mathbf{D}$ et $\mathbf{\Phi}$. En général, $\mathbf{\Phi}$ est une matrice unitaire tronquée et distribuée isotrope. Alors que la distribution de $\mathbf{D}$ doit être adaptée en fonction de la relation de l'intervalle de cohérence et du nombre d'antennes.

- Si $T \geq M' + N$ où $M' := \min\{M, N\}$, il est optimal de laisser les premiers M' éléments diagonaux de $\mathbf{D}$ égaux à $\sqrt{\frac{PT}{M'}}$ et les éléments diagonaux restants égaux à 0 avec la probabilité 1 [49]. C'est-à-dire que seules M' antennes d'émission sont utilisées pour émettre les premières M' lignes de puissance égale par antenne. La distribution d'entrée résultante est référencée à *modulation spatio-temporelle unitaire* (USTM) [53].
- Si $T < M + N$ et $M \leq \min\{N, \lfloor T/2 \rfloor\}$, les éléments diagonaux carrés du $\mathbf{D}$ optimal ont la même distribution conjointe que les valeurs propres ordonnées d'une matrice $M \times M$ aléatoire bêta distribuée et positive-définie [50]. La distribution d'entrée résultante est appelée *modulation spatio-temporelle à variation beta* (BSTM). Notez que dans ce cas, le débit réalisable avec l'entrée USTM est à un écart constant inférieur à la capacité atteinte avec l'entrée BSTM.

Étant donné que la plupart des travaux susmentionnés dans le cas $T \geq 2$ reposent sur l'hypothèse d'évanouissement Rayleigh IID, la question suivante est toujours ouverte:

Question A.1. *Quelle est la limite de capacité du canal P2P MIMO non-cohérent en évanouissement* générique *par bloc?*

A.2.2.b Conception de Constellation

Dans cette sous-section, nous nous concentrons sur la conception de la constellation pour le cas $T \geq 2$ et, motivés par le DoF optimal, supposons que $M \leq \min\{N, \lfloor T/2 \rfloor\}$. Nous

considérons l'entrée USTM

$$\mathbf{X} = \sqrt{\frac{PT}{M}}\mathbf{\Phi}, \quad \text{where } \mathbf{\Phi}\mathbf{\Phi}^{\mathsf{H}} = \boldsymbol{I}_M. \tag{A.16}$$

Avec USTM, le signal d'entrée est invariant par rapport à la rotation à partir de la droite et les informations sont intégrées dans l'espace de ligne de la matrice $\mathbf{\Phi}$. L'intuition derrière l'optimalité de USTM est que la matrice de canaux $\mathbf{H}$ ne fait que mettre à l'échelle et faire pivoter les bases de la matrice de signaux transmis $\mathbf{X}$ sans changer son sous-espace car les espaces de lignes de $\mathbf{X}$ et l'observation sans bruit $\mathbf{HX}$ sont les mêmes. L'information est transportée à la position de l'espace de rangées de dans la variété Grassmann $G(\mathbb{C}^T, M)$, qui est l'espace des sous-espaces à dimensions M de $\mathbb{C}^T$. Par définition, $\mathbf{X}$ et $\mathbf{HX}$ représentent le même élément de $G(\mathbb{C}^T, M)$. À SNR élevé, le bruit additif a un faible impact sur le sous-espace du signal de sortie, et le sous-espace de $\mathbf{X}$ peut être récupéré avec précision à partir du sous-espace de la sortie bruyante.

D'après l'observation ci-dessus, une constellation pour une communication MIMO non-cohérente sur un évanouissement par bloc peut être conçue comme un ensemble de représentants de sous-espaces à dimension M en $\mathbb{C}^T$. De manière équivalente, ces constellations représentent un ensemble de points sur la variété Grassmann $G(\mathbb{C}^T, M)$. Nous appelons les constellations de ce type les *constellations Grassmanniennes*. Compte tenu de la taille de la constellation, la construction de la constellation Grassmannienne peut être interprétée comme un emballage de points dans la variété Grassmann. Le critère d'emballage ultime est de minimiser l'erreur de détection sous observation bruyante. Cela revient généralement à maximiser la distance par paire minimale entre les points de constellation. Un certain nombre de constellations grassmanniennes ont été proposées avec différents critères, génération de constellation et méthodes de détection. Ils suivent deux approches principales.

- La première approche utilise des outils d'optimisation numérique pour résoudre le problème d'emballage de sphères sur le Grassmannien [63]–[69]. Il en résulte des constellations avec un bon spectre de distance mais sans structure particulière. En raison du manque de structure, ce type de constellation doit être stocké à la fois à l'émetteur et au récepteur, et décodé avec le décodeur ML de haute complexité, ce qui limite l'utilisation pratique à seulement petites constellations.
- La seconde approche impose une structure particulière à la constellation [70]–[77]. L'entrée structurée de données–pilotes d'un schéma basé sur pilote peut également être considérée comme un code non-cohérent [78]. La structure de la constellation facilite la modulation et, probablement, la démodulation de faible complexité.

Dans cette thèse, nous souhaitons apporter notre réponse à la question suivante.

Question A.2. *Comment concevoir une constellation Grassmannienne qui a une efficacité d'emballage élevée tout en étant simple à générer (donc disponible pour une grande taille de constellation et une grande longueur de symbole, et ne nécessitant pas d'être stockée), en admettant un schéma d'étiquetage binaire simple et efficace, et en permettant une schéma de détection efficace?*

A.2.3 Le Canal d'Accès Multiple

Nous considérons un canal d'accès multiple (MAC) MIMO avec K émetteurs émettant vers un récepteur. L'utilisateur k est équipé de M_k antennes, $k \in [K]$, et le récepteur a N antennes.

A.2.3.a Limites Fondamentales

Shamai et Marzetta ont étudié la capacité du MAC à entrée unique et sorties multiples (SIMO) ($M_k = 1, k \in [K]$) dans l'évanouissement Rayleigh IID dans [84]. Pour l'évanouissement rapide ($T = 1$), ils ont montré que la capacité totale pour $K > 1$ utilisateurs est égale à la capacité pour $K = 1$ utilisateur, donc l'accès multiple par répartition dans le temps (TDMA) est optimal. Pour l'évanouissement par bloc ($T > 1$), ils ont supposé que la capacité de somme maximale ne peut être atteinte que par plus de $K = T$ utilisateurs, ce qui est soutenu par une analyse asymptotique. Dans le même contexte avec $T > 1$, Gopalan *et al.* a dérivé une borne inférieure constructive de la capacité de somme d'un schéma de décodage successif [85]. Étant donné que la capacité de somme du MAC peut être supérieure par la capacité du canal P2P en permettant la coopération de l'utilisateur, la capacité de somme du MAC sous l'évanouissement rapide peut être montrée à l'échelle comme un double logarithme du SNR plus un nombre d'évanouissement. Le nombre décroissant de le MAC à entrées multiples sortie unique (MISO) dans l'évanouissement Ricienne a été dérivé par Lin et Moser dans [86]. Devassy *et al.* fourni des limites supérieures et inférieures non asymptotiques sur la capacité de somme du MAC MIMO sous l'évanouissement Rayleigh dans [87]. Dans ce cas, la limite inférieure dérivée avec une signalisation Grassmannienne indépendante de chaque utilisateur a un petit écart avec la limite supérieure même à un SNR modéré.

Les travaux susmentionnés portent sur la capacité totale du MAC non-cohérent. La région de pleine capacité est inconnue et seules quelques régions réalisables du DoF ont été proposées. Une région DoF réalisable pour le MAC MIMO à deux utilisateurs sous l'évanouissement Rayleigh IID a été proposée dans [88] en utilisant une approche géométrique. Une généralisation de cette région DoF réalisable au cas de K-user est présentée dans [37, Theorem 5]. Ces régions réalisables peuvent être atteintes par un schéma pilote simple. Ces régions DoF réalisables sont somme-DoF optimales, mais l'optimalité globale n'est pas connue. Une limite extérieure coopérative pour la région optimale de DoF a également été donnée dans [37, Section VI-B].

Dans cette thèse, nous étudierons la région DoF optimale du MAC non-cohérent et répondrons à la question ouverte suivante.

Question A.3. *Quelle est la région DoF optimale pour le MAC MIMO non-cohérent dans l'évanouissement générique par bloc?*

A.2.3.b Constellation Design

Comme dans le cas d'un utilisateur unique (P2P), les signaux transmis $\mathbf{X}_k$ de l'utilisateur k sont normalement tirés d'une constellation discrète finie $\mathcal{X}_k$, donc $\mathcal{X} := \{[\boldsymbol{X}_1\ \boldsymbol{X}_2\ \dots\ \boldsymbol{X}_K] : \boldsymbol{X}_k \in \mathcal{X}_k, k \in [K]\}$ est la constellation conjointe du MAC.

Une extension simple du schéma pilote pour le cas d'un utilisateur unique consiste à diviser le bloc cohérent en deux parties: 1) la partie d'apprentissage dans laquelle des séquences pilotes mutuellement orthogonales sont envoyées pour estimer le CSI pour chaque utilisateur, et 2) le partie de transmission de données dans laquelle différents utilisateurs communiquent simultanément [89] en utilisant une constellation scalaire (par exemple, PAM, QAM, PSK). L'optimalité de cette approche en termes de taux réalisable et d'erreur de détection reste incertaine. De plus, il n'est pas toujours possible d'affecter des pilotes orthogonaux à tous les utilisateurs du système avant la transmission, comme indiqué au Section A.1.5.

Un schéma de codage basé sur l'amplitude a été proposé dans [90], [91], mais le détecteur d'énergie qui l'accompagne repose sur un grand nombre d'antennes de réception de sorte que la puissance reçue moyenne à travers toutes les antennes se concentre. Également avec un réseau d'antennes de réception massif, certains schémas de codage différentiel ont été étudiés sur la base de PSK [92], [93] ou QAM [94]. Une autre ligne de travail est basée sur les constellations PSK qui sont absolument additivement décomposables de manière unique, c'est-à-dire que chaque symbole PSK individuel peut être décodé de manière unique à partir de n'importe quelle combinaison linéaire de deux points de constellation PSK avec des poids positifs [95], [96]. Dans ce schéma, la décodabilité unique de la matrice du signal repose sur l'orthogonalité asymptotique entre les canaux des utilisateurs dans le régime MIMO massif. Une propriété similaire uniquement décomposable a également été exploitée pour la modulation spatio-temporelle multi-utilisateurs basée sur QAM [97]. En bref, ces schémas supposent un grand nombre d'antennes de réception.

Contrairement au cas d'un utilisateur unique où un bon critère de conception de constellation est connu pour être le remplissage de sphère dans le Grassmannien, pour le MAC MIMO non-cohérent en évanouissement par bloc, un critère de conception de constellation simple et efficace reste incertain. En général, la constellation conjointe $\mathcal{X}$ pour le MAC doit être conçue de manière à minimiser l'erreur de détection de symbole. Si les utilisateurs pouvaient coopérer, le système pourrait être considéré comme un canal non-cohérent P2P MIMO $\left(\sum_{k=1}^{K} M_k\right) \times N$, pour lequel l'entrée optimale à SNR élevé est uniformément distribué sur le Grassmannien $G(\mathbb{C}^T, \sum_{k=1}^{K} M_k)$ [49]. Inspirée de cela, la constellation conjointe pour le MAC peut être traitée comme une constellation Grassmannienne en $G(\mathbb{C}^T, \sum_{k=1}^{K} M_k)$, ce qui conduit à un critère de conception imitant le remplissage des sphères dans ce Grassmannien en maximisant la distance cordale par paire minimale. Brehler et Varanasi ont dérivé la probabilité d'erreur du détecteur ML pour le MAC MIMO dans [98] et ont analysé les asymptotiques à SNR élevé. Cependant, un critère explicite de conception de constellation n'a pas été fourni et l'analyse a conduit à une conception pour le MAC uniquement avec les utilisateurs coopérants. Avec les utilisateurs coopérants, le critère de conception est similaire à celui d'un canal à utilisateur unique proposé dans [68, Eq.(8)] par les mêmes auteurs, qui est différent du critère de max-min distance cordale par paire. Ce critère peut être utilisé pour le cas des utilisateurs non coopérants en modifiant l'espace d'optimisation. On peut montrer que l'exposant d'erreur par paire converge vers la divergence de Kullback-Leibler (KL) entre les distributions de sortie conditionnées sur l'un ou l'autre des symboles transmis [59]. Sur cette base, un critère consistant à maximiser la divergence KL minimale a été proposé pour le MAC dans [97]. Cependant, ce travail se concentre sur la modulation spatio-temporelle basée sur QAM et utilise uniquement le critère pour optimiser les puissances de transmission et l'attribution de la sous-constellation.

Dans cette thèse, nous souhaitons répondre davantage à la question suivante:

Question A.4. *Comment concevoir une constellation conjointe efficace pour le MAC MIMO non-cohérent afin d'obtenir un faible taux d'erreur de symbole (SER)?*

A.2.3.c Détection Multi-Utilisateurs

Nous déplaçons maintenant le focus sur le récepteur, dont la tâche est de détecter les symboles transmis (ou plutôt les bits sous-jacents) sur la base de l'observation bruyante. Étant donné que le problème de détection cohérente ML optimal est NP-difficile, le récepteur peut utiliser un détecteur sous-optimal pour réduire la complexité.

Si les signaux transmis contiennent des pilotes, le récepteur peut estimer (normalement imparfaitement) le canal sur la base des symboles pilotes, puis effectuer une détection cohérente sur la base de l'estimation du canal. Cette approche s'appuie sur une riche littérature de détection cohérente, dans laquelle de nombreux algorithmes sous-optimaux ont été proposés, comme indiqué dans la Section A.1.4. L'estimation de canal et la détection de données peuvent également être effectuées de manière itérative [99], [100] ou conjointement sur la base d'une recherche d'arbre [101], [102]. Ces schémas nécessitent une transmission pilote pour une estimation initiale du canal ou pour garantir l'identifiabilité des symboles de données. En revanche, avec la transmission sans pilote, le détecteur non-cohérent n'effectue pas d'estimation de canal explicite et exploite uniquement la connaissance statistique de la matrice de canal.

Le récepteur pourrait être intéressé non seulement par la détection dure des symboles, mais aussi par leur fonction de masse de probabilité marginale postérieure (PMF). Ces informations "douce" sont nécessaires, par exemple, lors du calcul des rapports de vraisemblance logarithmique (LLR) au niveau du bit requis pour le décodage de canal. Le calcul d'une PMF marginale exacte nécessiterait d'énumérer toutes les combinaisons possibles de signaux provenant d'autres utilisateurs, ce qui est impossible avec de nombreux utilisateurs, de nombreuses antennes ou de grandes constellations. Ainsi, comme pour la détection dure, un schéma sous-optimal est nécessaire. Contrairement à la détection probabiliste MIMO cohérente, pour laquelle de nombreux schémas ont été proposés (par exemple, [103]–[105]), la détection probabiliste non-cohérente MIMO sous signalisation générale, et la signalisation Grassmannienne en particulier, n'a pas été bien étudiée. Par conséquent, dans cette thèse, nous aimerions répondre à la question suivante:

Question A.5. *Comment détecter efficacement les symboles et estimer leurs postérieurs marginaux dans le MAC MIMO non-cohérent avec des constellations générales?*

A.2.4 Le Canal de Diffusion

Prenons un canal de diffusion (BC) MIMO dans lequel une station de base avec M antennes transmet à K utilisateurs. L'utilisateur k est équipé de N_k antennes, $k \in [K]$.

La plupart des études sur les limites de capacité du BC supposent des informations parfaites sur l'état des canaux au niveau des récepteurs (CSIR) et différents niveaux d'informations sur l'état des canaux au niveau de l'émetteur (CSIT): parfait CSIT [106], [107] (avec codage sur papier sale (DPC)), CSIT imparfait (partiel) [108]–[110] et aucun CSIT [111], [112]. Le BC MISO non-cohérent (pas de CSIT, pas de CSIR) avec canal distribué isotrope a été brièvement mentionné par Jafar et Goldsmith dans [111, Section VII-C]. Ils ont montré que *la limite supérieure scalaire* proposée sur la région de capacité, qui a été appliquée avec succès au cas parfait du CSIR, devient lâche pour le BC non-cohérent car elle ne tient pas compte de la perte de DoF en raison de l'absence de CSIR. Dans le cadre de l'évanouissement Rayleigh IID par bloc, Fadel et Nosratinia ont trouvé la région DoF optimale du MIMO BC non-cohérent donnée dans [37, Theorem 1], qui est réalisable avec un schéma TDMA.

Les résultats susmentionnés supposent l'indépendance statistique entre chaque paire d'antennes d'émission et de réception. En pratique, cependant, les canaux entre différentes antennes sont souvent corrélés parce que l'environnement de propagation provoque souvent des gains de signal reçu plus forts dans certaines directions spatiales, et également en raison des modèles spatialement dépendants des antennes. Comme les utilisateurs ne sont pas colocalisés, ils peuvent avoir des matrices de corrélation non identiques. Dans ce cas, un

outil utile tirant parti de la différence entre les corrélations spatiales observées par différents utilisateurs est la diversité de corrélation de transmission. Pour les matrices de corrélation spatiale de transmission qui ont des espaces propres mutuellement exclusifs, la *diversité de corrélation de transmission* peut être récoltée en utilisant un schéma de transmission de multiplexage spatial joint (JSDM) [113], [114] qui réduit la surcharge nécessaire pour l'estimation de canal. L'idée principale de JSDM est de partitionner les utilisateurs en groupes avec approximativement le même espace de corrélation de canal. Un autre schéma de transmission en *superposition de produits* proposé par Li et Nosratinia dans [115], [116] pour le BC avec des utilisateurs mixtes statiques (avec CSIR) et dynamiques (sans CSIR), puis appliqué à des BC non-cohérents spatialement corrélés avec des espaces propres de corrélation qui se chevauchent complètement [117].

Notez que le schéma de transmission JSDM tente d'exploiter les parties non chevauchantes (mutuellement exclusives) des espaces de corrélation, tandis que la superposition de produits exploite les parties chevauchantes. Dans cette thèse, nous souhaitons exploiter les deux parties en répondant à la question suivante.

Question A.6. *Quelles sont les limites fondamentales de la BC non-cohérente spatialement corrélée dans l'évanouissement générique par bloc avec des espaces de corrélation se chevauchant partiellement et comment exploiter efficacement la diversité de corrélation de transmission dans cette condition?*

A.3 Description de la Thèse

Cette thèse est une contribution aux communications sans fil non-cohérentes. Nous supposons tout au long que la distribution des canaux est connue, mais le CSI n'est connu par aucun nœud communiquant, et les coefficients des canaux sont en fondu avec l'intervalle de cohérence T. L'hypothèse sur la distribution des canaux (Rayleigh / générique, corrélée / non corrélée) est spécifiée dans chaque chapitre. Nous étudions le canal P2P, le MAC et le BC en termes de limites fondamentales (capacité, débit réalisable, DoF) et de conception d'émetteur-récepteur (conception de constellation, conception de détection efficace). La principale contribution de la thèse est la réponse ou l'extension des réponses existantes aux questions posées dans la section précédente. Le plan détaillé et les contributions de la thèse sont présentés ci-dessous.

A.3.1 Aperçu et Contributions de la Thèse

Cette thèse comprend deux parties principales traitant respectivement des limites fondamentales et des aspects de conception d'émetteur-récepteur des communications non-cohérentes, suivies de la conclusion et des perspectives.

Partie I: Limites Fondamentales des Communications Non-Cohérentes

Dans *Chapitre 2*, nous répondons en partie à *Question 1.1* en étudiant la DoF optimale du canal P2P MIMO non-cohérent dans l'évanouissement générique par bloc. Nous montrons que le DoF optimal pour le canal P2P MIMO sous l'évanouissement Rayleigh IID (trouvé dans [49], [50]) tient également sous l'évanouissement générique. Nous introduisons une nouvelle technique de preuve inverse basée sur une liaison assistée par le génie et l'approche de la dualité [36]. Les extensions possibles incluent l'analyse du débit / capacité réalisable de SNR fini. Le résultat de ce chapitre a été présenté dans [118], [119].

Dans *Chapitre 3*, nous répondons à *Question 1.3*, c'est-à-dire la région DoF optimale pour le MAC non-cohérent dans l'évanouissement générique par bloc, dans le cas SIMO. La preuve inverse suit en montrant que la contrainte de puissance moyenne peut être remplacée par une contrainte de puissance de crête sans changer la région DoF optimale, avec une borne assistée par le génie similaire et l'approche de la dualité comme au Chapitre 2. Il serait intéressant d'étendre à le cas MIMO et / ou plus de deux utilisateurs. La caractérisation par SNR fini de la région de capacité serait difficile, mais on peut tenter d'analyser la capacité à somme. Le résultat de ce chapitre a été présenté dans [118], [119].

Chapitre 4 présente notre réponse à *Question 1.6*. En considérant un BC MIMO spatialement corrélé avec des espaces de corrélation se chevauchant partiellement, nous exploitons la diversité de corrélation de transmission en concevant soigneusement des schémas de transmission basés sur le fractionnement de débit, la superposition de produits, et une version hybride de ceux-ci. Ce faisant, nous trouvons des régions de débit et de DoF réalisables pour le BC dans le cas de deux utilisateurs, et certaines régions de DoF réalisables dans le cas de K utilisateurs. Les régions DoF réalisables s'améliorent significativement par rapport à TDMA, qui s'est avéré être optimal pour le DoF pour la décoloration non corrélée [37]. En tant que sous-produit, nous dérivons également un débit réalisable pour le canal P2P MIMO avec un évanouissement spatialement corrélé. Bien que ces régions réalisables s'améliorent considérablement par rapport à la TDMA, il faut caractériser les limites extérieures pour voir dans quelle mesure ces régions sont éloignées de l'optimalité. Une approche consiste à utiliser la coopérative liée en laissant les utilisateurs coopérer et en les considérant comme un seul macro-utilisateur. Cependant, les résultats existants pour le canal P2P, tels que [117, Theorem 1], ne s'appliquent pas immédiatement car les colonnes de canal ne sont pas statistiquement équivalentes (en raison de la différence dans leurs matrices de corrélation). Nos tentatives utilisant l'approche de la dualité montrent qu'il est difficile de choisir une distribution de sortie auxiliaire adéquate. Le résultat de ce chapitre a été publié dans [120] et [121].

Partie II: Conception de Système pour les Communications Non-Cohérentes

Chapitre 5 donne notre réponse à *Question 1.2* dans le cas SIMO. Nous proposons une constellation structurée dans le Grassmannien de lignes dites *constellation de cube-split*. Il est généré en partitionnant le Grassmannien de lignes en une collection d'hypercubes courbés et en définissant un mappage sur chacun de ces hypercubes courbés de telle sorte que les symboles résultants sont répartis approximativement uniformément sur le Grassmannien. Cette constellation remplit toutes les caractéristiques souhaitées à la *Question 1.2*: elle a une efficacité de tassement élevée représentée par la distance cordale par paire minimale tout en étant simple à générer (donc disponible pour une grande taille de constellation et une grande longueur de symbole, et ne nécessitant pas d'être stockée); il admet un schéma d'étiquetage binaire simple et efficace; et il permet une détection dure et douce efficace. Ces avantages par rapport au schéma pilote et à d'autres constellations Grassmanniennes structurées sont plus prononcés dans le régime d'intervalle de cohérence court et de grande taille de constellation, comme le montrent les résultats numériques. Un travail futur est de l'étendre au cas MIMO où le principal défi est de trouver un moyen efficace de partitionner le Grassmannien. Le résultat de ce chapitre a été publié dans [122], [123], and [124].

Dans *Chapitre 6*, nous présentons une conception de constellation conjointe pour le MAC MIMO non-cohérent dans l'évanouissement par bloc de Rayleigh IID, ainsi donnons notre réponse à *Question 1.4*. Nous analysons l'erreur de détection ML et introduisons de nouveaux critères de conception afin de minimiser la probabilité d'erreur. Nous simplifions

davantage les métriques par une analyse SNR élevée. Nos métriques peuvent être utilisées pour la construction de constellations conjointes en les optimisant (numériquement) sur l'ensemble des matrices de signaux. De plus, sur la base de ces métriques, nous proposons deux constructions simples consistant respectivement en partitionnant une constellation mono-utilisateur ou en précodant des constellations mono-utilisateur de dimension inférieure. Nous étudions la possibilité de construire chaque constellation individuelle sous la forme d'une constellation Grassmannienne mise à l'échelle par la puissance de transmission respective. Les résultats numériques montrent que nos métriques proposées sont significatives et que les constellations résultantes fonctionnent mieux, pour le même débit de transmission et la même puissance, qu'un schéma basé sur un pilote et les constellations optimisées avec les métriques existantes. À l'avenir, il est souhaitable d'utiliser nos critères comme lignes directrices pour concevoir des constellations articulaires structurées plus efficaces. Les résultats de ce chapitre sont présentés collectivement dans [125], [124], [126], and [127].

Dans *Chapitre 7*, nous nous concentrons sur le côté récepteur d'un MAC et répondons à *Question 1.5*. Nous proposons un schéma de détection douce multi-utilisateur non-cohérent pour le MAC SIMO sous évanouissement par bloc de Rayleigh corrélé spatialement. Notre détecteur est basé sur la propagation d'espérance (EP) inférence approximative et présente une complexité polynomiale dans le nombre d'utilisateurs, le nombre d'antennes de réception et l'intervalle de cohérence. Nous proposons également deux simplifications de ce détecteur à complexité réduite. Les détecteurs proposés peuvent être utilisés pour la signalisation générale avec des symboles à valeur vectorielle transmis sur chaque bloc de cohérence. Dans ce chapitre, nous proposons également un schéma de détection efficace pour la constellation basée sur le précodage dans Chapitre 6, qui a une complexité plus faible mais fonctionne moins bien que le détecteur EP. Avec la signalisation assistée par pilote, le détecteur EP surpasse, en termes de taux d'erreur de symboles, certains détecteurs pilotes cohérents conventionnels, à savoir un décodeur MMSE linéaire, un décodeur sphérique et un schéma joint d'estimation de canal et de détection de données. Nos détecteurs basés sur l'EP produisent des approximations précises de la vraie postérieure conduisant à un débit de somme réalisables élevé. Les gains de ces détecteurs sont également observés en termes de taux d'erreur binaire lors de l'utilisation de leurs sorties pour un décodeur de canal turbo. Les extensions possibles incluent la prise en compte de modèles d'évanouissement plus complexes et l'analyse théorique des performances de l'EP pour une détection non-cohérente. Les résultats de ce chapitre ont été publiés dans [128], [125] and [129].

Nous résumons les principaux résultats de cette thèse dans Tableau 1.1.

TABLE A.1: Les principaux résultats de cette thèse

Canal	**Partie I: Limites fondamentales**		**Partie II: Conception du système**	
	Débit / DoF réalisable	DoF optimal	Conception de constellation	Conception de détection
P2P	Chapitre 4	Chapitre 2	Chapitre 5	
MAC	—	Chapitre 3	Chapitre 6	Chapitre 7
BC	Chapitre 4	—	—	—

Conclusion et Perspectives

Nous concluons la thèse du *Chapitre 8* qui résume les principaux résultats et donne des perspectives pour les travaux futurs. Nous fournissons également une perspective

supplémentaire des communications non-cohérentes dans le contexte d'autres sujets dans les communications sans fil dans ce qui suit.

- **Communications non-cohérentes dans un canal à variation continue**: Dans le cas d'un évanouissement par bloc, on suppose que les coefficients du canal sont entièrement corrélés (c'est-à-dire égaux) dans un bloc de cohérence et non corrélés entre les blocs. En pratique, les coefficients de canal sont fortement corrélés mais peuvent varier au sein d'un bloc. De plus, les coefficients séparés par un intervalle de cohérence sont faiblement corrélés, mais deux coefficients adjacents à travers deux blocs adjacents peuvent être fortement corrélés. Cela peut être observé à la fois dans le domaine temporel et le domaine fréquentiel (avec OFDM). Dans ce scénario, le DoF a été étudié dans [211], [212]. De plus, la signalisation Grassmannienne devient sous-optimale, et la détection ML à évanouissement par bloc devient une détection ne correspondent pas. Il serait intéressant de déterminer dans quelle mesure la constellation et la détection Grassmanniennes ne correspondent pas, comme cela a été étudié dans [81] en termes de taux d'erreur. Cette non-concordance peut également être analysée en termes de débit de données. La détection ML par bloc peut être modifiée en tenant compte de la corrélation des coefficients de canal au sein d'un bloc. D'un autre côté, nous pouvons essayer d'adapter les constellations Grassmanniennes à des canaux variant en continu. Par exemple, en exploitant le fait qu'un symbole Grassmannien est invariant à la mise à l'échelle par un scalaire complexe, on peut coder des bits d'information supplémentaires dans le déphasage des symboles à travers des blocs de cohérence. Nous avons rapporté cette idée dans [124].
- **Utilisation des constellations grassmanniennes pour une conception de système plus sophistiquée**: La conception des constellations Grassmanniennes se concentre sur le mappage bit à symbole et le démappage symbole à bit d'un système de communication numérique. Une conception plus générale du système utilisant des constellations Grassmanniennes comme ingrédient peut être effectuée. Par exemple, dans [213]–[215], des schémas de codage multicouches sont proposés dans lesquels des constellations Grassmanniennes sont utilisées dans la couche de résolution inférieure qui n'a pas accès à CSI. L'intégration des constellations Grassmanniennes dans les réseaux d'évolution à long terme (LTE) a été étudiée avec une étude de performance au niveau de la liaison dans [216]. Des systèmes de modulation codés entrelacés par bits utilisant des constellations Grassmanniennes ont été conçus dans [80], [177].
- **Communications non-cohérentes à faible latence**: Dans cette thèse, nous supposons qu'un message porteur d'informations, représenté par un mot de code de canal, s'étend sur un grand nombre de blocs de cohérence. Par conséquent, nous utilisons des métriques asymptotiques telles que la capacité de Shannon et le DoF. Dans la communication ultra-fiable à faible latence (URLLC), les informations sont transportées dans de petits paquets qui doivent être reçus avec une faible latence. Dans ce régime, la communication non-cohérente convient car la taille des métadonnées (par exemple, informations de contrôle, pilotes) est comparable à la taille des données [217]. Cependant, les repères classiques de la théorie de l'information asymptotique ne sont pas applicables, et il faut passer à un cadre qui permet de comprendre les transmissions de paquets courts. La théorie de l'information sur la longueur de bloc finie[1] [218] est appropriée à cette fin, où la métrique d'intérêt est *le débit de codage maximal* $R^*(n, \epsilon)$ à la longueur de bloc finie n et la probabilité d'erreur de paquet fini

[1]Le terme "bloc" en *longueur de bloc fini* fait référence à un mot de code de canal et est différent de celui de bloc de cohérence: dans la théorie de l'information à longueur de bloc finie, un mot de code peut s'étendre sur plusieurs blocs de cohérence.

ϵ. Des limites non asymptotiques sur le taux de codage maximal non-cohérent ont été proposées pour le canal d'évanouissement quasi-statique de Rayleigh (un mot de code de canal s'étend sur un seul bloc de cohérence) pour le cas SISO dans [219]. Dans le canal à évanouissement par blocs (un mot de code de canal s'étend sur plusieurs, mais un nombre fini, de blocs de cohérence), des limites non asymptotiques pour le débit de codage maximal non-cohérent ont été présentées pour le cas SISO dans [187], [220], et le cas MIMO dans [186]. Un schéma de communication non-cohérent dans lequel les informations sont modulées sur les zéros de la transformée z du signal de bande de base transmis a été proposé pour la communication par paquets courts dans [221].

- **Communication de type machine non-cohérente et accès aléatoire massif**: Dans la communication de type machine et l'accès aléatoire massif, le système doit prendre en charge un très grand nombre d'utilisateurs et seul un sous-ensemble aléatoire d'entre eux est actif à la fois. L'ensemble des utilisateurs actifs peut ne pas être connu à l'avance et il est impossible de pré-affecter des pilotes mutuellement orthogonaux à chaque utilisateur. Par conséquent, une communication sans pilote non-cohérente peut être une stratégie appropriée. Un protocole d'accès aléatoire non-cohérent basé sur des trames de Gabor, qui sont également des trames de Grassmannian, a été proposé dans [222]. Dans [223], Senel *et al.* a proposé un schéma de transmission non-cohérent pour les communications massives de type machine et a développé un algorithme de passage de message approximatif modifié pour exploiter la rareté structurée de ce schéma. Comparé à une transmission cohérente, ce schéma non-cohérent fonctionne nettement mieux, surtout lorsque le nombre de bits est petit. Récemment, notre constellation cube-split a été utilisée pour construire un schéma d'accès aléatoire massif basé sur le tenseur dans [224].

- **Apprentissage automatique pour les communications non-cohérentes**: L'utilisation de l'apprentissage automatique, en particulier l'apprentissage en profondeur, dans les communications sans fil se développe rapidement et profondément dans de nombreux aspects des communications de la couche physique, de l'utilisation de l'apprentissage pour une tâche spécifique telle que l'estimation de canal, l'égaliseur de symboles, détection de signal, codage / décodage de canal, vers une transmission complète basée sur l'apprentissage de bout en bout [225]–[227]. Les communications non-cohérentes peuvent certainement bénéficier de l'apprentissage automatique. Une méthode d'apprentissage automatique pilotée par les données a été proposée pour concevoir un émetteur-récepteur non-cohérent pour la transmission de paquets courts dans [228]. Les méthodes d'apprentissage sur la Grassmannien peuvent être utilisées pour une communication non-cohérente avec la signalisation Grassmannienne [229]. Par exemple, Du *et al.* utilisé des algorithmes de regroupement de données sur la variété Grassmann pour concevoir un schéma de reconnaissance de modulation automatique pour les constellations Grassmanniennes dans [230]. Dans l'apprentissage de bout en bout des systèmes de communication via des auto-encodeurs basés sur un réseau de neurones, l'estimation de canal peut être effectuée même avec un modèle de canal inconnu [231]. Lorsque le modèle de canal est inconnu, on peut utiliser l'apprentissage basé sur les données pour ajuster la probabilité conditionnelle d'entrée-sortie empirique à un modèle paramétrique. Nous avons proposé un cadre utilisant un modèle gaussien généralisé à cette fin dans [130]. Un autoencodeur pour une communication basée sur l'énergie non-cohérente dans un système SIMO multi-porteuses multi-utilisateurs a été proposé dans [232]. L'article [233] présente la conception de formes d'ondes pour les systèmes MIMO multi-utilisateurs non-cohérents grâce à l'apprentissage en profondeur.

A.3.2 Liste des Publications

Les publications incluses dans le résultat principal de cette thèse sont listées ci-dessous.

Articles de Journaux

[119] K.-H. Ngo, S. Yang, and M. Guillaud, "The optimal degrees of freedom for the point-to-point and multiple-access channels in generic block fading," *in preparation to submit to IEEE Trans. Inf. Theory*, 2020.

[127] K.-H. Ngo, S. Yang, M. Guillaud, and A. Decurninge, "Joint constellation design for the noncoherent MIMO multiple-access channel," *in preparation to submit to IEEE Trans. Wireless Commun.*, 2020.

[121] F. Zhang, K.-H. Ngo, S. Yang, and A. Nosratinia, "Transmit correlation diversity: Generalization, new techniques, and improved bounds," *submitted to IEEE Trans. Inf. Theory*, 2020.

[129] K.-H. Ngo, M. Guillaud, A. Decurninge, S. Yang, and P. Schniter, "Multi-user detection based on expectation propagation for the noncoherent SIMO multiple access channel," *IEEE Trans. Wireless Commun.*, 2020, (arXiv preprint arXiv:1905.11152).

[123] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Cube-split: A structured Grassmannian constellation for noncoherent SIMO communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1948–1964, Mar. 2020.

Préimpression

[126] K.-H. Ngo, S. Yang, M. Guillaud, and A. Decurninge, "Joint constellation design for the two-user noncoherent multiple-access channel," arXiv preprint arXiv:2001.04970, 2020.

Conférence

[128] K.-H. Ngo, M. Guillaud, A. Decurninge, S. Yang, S. Sarkar, and P. Schniter, "Noncoherent multi-user detection based on expectation propagation," in *53rd Asilomar Conference on Signals, Systems, and Computers*, CA, USA, Nov. 2019.

[125] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "A multiple access scheme for noncoherent SIMO communications," in *52nd Asilomar Conference on Signals, Systems, and Computers*, CA, USA, Oct. 2018, pp. 1846–1850.

[118] K.-H. Ngo, S. Yang, and M. Guillaud, "The optimal DoF region for the two-user noncoherent SIMO multiple-access channel," in *IEEE Information Theory Workshop (ITW)*, arXiv preprint arXiv:1806.04102, Guangzhou, China, Nov. 2018.

[120] K.-H. Ngo, S. Yang, and M. Guillaud, "An achievable DoF region for the two-user noncoherent MIMO broadcast channel with statistical CSI," in *IEEE Information Theory Workshop (ITW)*, Nov. 2017, pp. 604–608.

[122] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Design and analysis of a practical codebook for noncoherent communications," in *51st Asilomar Conference on Signals, Systems, and Computers*, CA, USA, Oct. 2017, pp. 1237–1241.

Brevet

[124] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Transmitter and receiver communication apparatus for noncoherent communication," English, European patent, Application number 6 860 390, Filed on 30 Oct. 2018.

A.3.3 Contributions en Dehors du Champ de la Thèse

Au cours de la thèse, nous avons également publié quelques autres contributions qui ne sont pas incluses dans le résultat principal de cette thèse:

[130] K.-H. Ngo, S. Yang, and M. Guillaud, "Generalized Gaussian model for data-driven learning in communications," in *International Zurich Seminar on Information and Communication (IZS)*, poster, Zurich, Switzerland, Feb. 2020.

[131] K.-H. Ngo, S. Yang, and M. Kobayashi, "Scalable content delivery with coded caching in multi-antenna fading channels," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 548–562, Jan. 2018.

[132] A. Ghorbel, K.-H. Ngo, R. Combes, M. Kobayashi, and S. Yang, "Opportunistic content delivery in fading broadcast channels," in *IEEE Global Communications Conference (GLOBECOM)*, Singapore, Dec. 2017.

[133] T. T. Q. Tran, V.-L. Nguyen, K.-H. Ngo, L.-T. Nguyen, Q.-T. Nguyen, E. Bastug, S. Azarian, M. Debbah, and P. Duhamel, "Network coding and information security in industry 4.0," in *1st ASEAN IVO Workshop on Cybersecurity and Information Security in Industry 4.0*, poster, Hanoi, Vietnam, Mar. 2019.

Bibliography

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. National Institute of Standards and Technology, 1964.

[2] W. C. Jakes and D. C. Cox, Eds., *Microwave Mobile Communications*. Wiley-IEEE Press, 1994, ISBN: 0780310691.

[3] A. F. Molisch, *Wireless Communications*, 2nd ed. John Wiley & Sons, Jan. 2011.

[4] A. Goldsmith, *Wireless Communications*. Cambridge University Press, Aug. 2005.

[5] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005, ISBN: 0-5218-4527-0.

[6] R. Stair and G. Reynolds, *Principles of Information Systems*. Cengage Learning, 2015.

[7] C. E. Shannon, "A mathematical theory of communication," *Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[8] J. De Vriendt, P. Laine, C. Lerouge, and Xiaofeng Xu, "Mobile network evolution: A revolution on the move," *IEEE Communications Magazine*, vol. 40, no. 4, pp. 104–111, Apr. 2002, ISSN: 1558-1896. DOI: 10.1109/35.995858.

[9] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Communications Surveys Tutorials*, vol. 20, no. 2, pp. 1124–1148, 2018, ISSN: 1553-877X. DOI: 10.1109/COMST.2017.2785181.

[10] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[11] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[12] M. R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, and L. Ladid, "Internet of things in the 5G era: Enablers, architecture, and business models," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 510–527, 2016.

[13] M Series, "IMT vision–Framework and overall objectives of the future development of IMT for 2020 and beyond," *Recommendation ITU*, pp. 2083–0, 2015.

[14] M. Patzold, "5G is coming around the corner," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 4–10, 2019.

[15] K. David and H. Berndt, "6G vision and requirements: Is there any need for beyond 5G?" *IEEE Veh. Technol. Mag.*, vol. 13, no. 3, pp. 72–80, 2018.

[16] V. Raghavan and J. Li, "Evolution of physical-layer communications research in the post-5G era," *IEEE Access*, vol. 7, pp. 10 392–10 401, 2019.

[17] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, pp. 1–9, 2019.

[18] E. Calvanese Strinati, S. Barbarossa, J. L. Gonzalez-Jimenez, D. Ktenas, N. Cassiau, L. Maret, and C. Dehos, "6G: The next frontier: From holographic messaging to artificial intelligence using subterahertz and visible light communication," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 42–50, 2019.

[19] F. Tariq, M. Khandaker, K.-K. Wong, M. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *arXiv preprint arXiv:1902.06700*, 2019.

[20] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?" *Nature Electronics*, vol. 3, no. 1, pp. 20–29, 2020.

[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd. New York, NY, USA: John Wiley & Sons, 2006.

[22] A. El Gamal and Y.-H. Kim, *Network Information Theory*. New York, NY, USA: Cambridge University Press, 2011, ISBN: 1107008735, 9781107008731.

[23] R. F. Fischer, *Precoding and Signal Shaping for Digital Transmission*. John Wiley & Sons, 2005.

[24] S. Verdú, "Computational complexity of optimum multiuser detection," *Algorithmica*, vol. 4, no. 1, pp. 303–312, 1989, ISSN: 1432-0541. DOI: 10.1007/BF01553893.

[25] A. Ganti, A. Lapidoth, and I. E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, 2000, ISSN: 0018-9448. DOI: 10.1109/18.887846.

[26] I. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, pp. 585–595, 1999.

[27] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless personal communications*, vol. 6, no. 3, pp. 311–335, 1998.

[28] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996, ISSN: 1538-7305. DOI: 10.1002/bltj.2015.

[29] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, "A survey on MIMO transmission with finite input signals: Technical challenges, advances, and future trends," *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1779–1833, 2018.

[30] S. Yang and L. Hanzo, "Fifty years of MIMO detection: The road to large-scale MIMOs," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 1941–1988, 2015, ISSN: 1553-877X. DOI: 10.1109/COMST.2015.2475242.

[31] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003, ISSN: 0018-9448. DOI: 10.1109/TIT.2003.809594.

[32] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000, ISSN: 0018-9448. DOI: 10.1109/18.841172.

[33] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203–2214, 2006.

[34] X. Tang, M.-. Alouini, and A. J. Goldsmith, "Effect of channel estimation error on M-QAM BER performance in Rayleigh fading," *IEEE Trans. Commun.*, vol. 47, no. 12, pp. 1856–1864, 1999.

[35] B. Xia and J. Wang, "Effect of channel-estimation error on QAM systems with antenna diversity," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 481–488, 2005.

[36] A. Lapidoth and S. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003, ISSN: 0018-9448. DOI: 10.1109/TIT.2003.817449.

[37] M. Fadel and A. Nosratinia, "Coherence disparity in broadcast and multiple access channels," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7383–7401, Dec. 2016, ISSN: 0018-9448. DOI: 10.1109/TIT.2016.2616126.

[38] S. M. Moser, "Duality-based bounds on channel capacity," Ph.D. dissertation, ETH Zürich, Switzerland, Jan. 2005.

[39] J. Sebastian and S. Diggavi, "On the generalized degrees of freedom of noncoherent interference channel," *arXiv preprint arXiv:1812.03579*, 2018.

[40] T. Koch and G. Kramer, "On noncoherent fading relay channels at high signal-to-noise ratio," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2221–2241, 2013.

[41] Z. Utkovski and T. Eftimov, "A pre-log region for the non-coherent MIMO two-way relaying channel," in *21st European Signal Processing Conference (EUSIPCO)*, 2013, pp. 1–5.

[42] J. Sebastian and S. Diggavi, "Generalized degrees of freedom of noncoherent diamond networks," *IEEE Trans. Inf. Theory*, 2020, Early Access.

[43] J. S. Richters, "Communication over fading dispersive channels," MIT Res. Lab. Electronics, Cambridge, MA, USA, Tech. Rep. 464, Nov. 1967.

[44] I. C. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001, ISSN: 1557-9654. DOI: 10.1109/18.923716.

[45] G. Taricco and M. Elia, "Capacity of fading channel with no side information," *Electronics Letters*, vol. 33, no. 16, pp. 1368–1370, Jul. 1997, ISSN: 0013-5194. DOI: 10.1049/el:19970918.

[46] S. A. Jafar and A. Goldsmith, "Multiple-antenna capacity in correlated Rayleigh fading with channel covariance information," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 990–997, May 2005, ISSN: 1536-1276. DOI: 10.1109/TWC.2005.847029.

[47] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999, ISSN: 0018-9448. DOI: 10.1109/18.746779.

[48] B. Hassibi and T. L. Marzetta, "Multiple-antennas and isotropically random unitary inputs: The received signal density in closed form," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1473–1484, Jun. 2002, ISSN: 1557-9654. DOI: 10.1109/TIT.2002.1003835.

[49] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002, ISSN: 0018-9448. DOI: 10.1109/18.978730.

[50] W. Yang, G. Durisi, and E. Riegler, "On the capacity of large-MIMO block-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 117–132, Feb. 2013, ISSN: 0733-8716. DOI: 10.1109/JSAC.2013.130202.

[51] F. Rusek, A. Lozano, and N. Jindal, "Mutual information of IID complex Gaussian signals on block Rayleigh-faded channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 331–340, Jan. 2012, ISSN: 1557-9654. DOI: 10.1109/TIT.2011.2171520.

[52] G. Alfano, C. Chiasserini, A. Nordio, and S. Zhou, "Closed-form output statistics of MIMO block-fading channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7782–7797, Dec. 2014, ISSN: 1557-9654. DOI: 10.1109/TIT.2014.2363838.

[53] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000, ISSN: 0018-9448. DOI: 10.1109/18.825818.

[54] K. Takeuchi, R. R. Müller, M. Vehkaperä, and T. Tanaka, "On an achievable rate of large Rayleigh block-fading MIMO channels with no CSI," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6517–6541, 2013.

[55] J. Sebastian and S. Diggavi, "Generalized degrees freedom of noncoherent MIMO channels with asymmetric link strengths," *arXiv preprint arXiv:1705.07355v6*, 2019.

[56] R. H. Gohary and H. Yanikomeroglu, "The ergodic high SNR capacity of the spatially-correlated non-coherent MIMO channel within an SNR-independent gap," in *Proc. IEEE Information Theory Workshop - Fall (ITW)*, 2015, pp. 234–238.

[57] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, Apr. 1999, ISSN: 0895-4798. DOI: 10.1137/S0895479895290954. [Online]. Available: https://doi.org/10.1137/S0895479895290954.

[58] A. Barg and D. Y. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Inf. Theory*, vol. 48, pp. 2450–2454, 2002.

[59] M. J. Borran, A. Sabharwal, and B. Aazhang, "On design criteria and construction of noncoherent space-time constellations," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2332–2351, 2003, ISSN: 0018-9448. DOI: 10.1109/TIT.2003.817431.

[60] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experiment. Math.*, vol. 5, no. 2, pp. 139–159, 1996. [Online]. Available: https://projecteuclid.org:443/euclid.em/1047565645.

[61] O. Henkel, "Sphere-packing bounds in the Grassmann and Stiefel manifolds," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3445–3456, Oct. 2005, ISSN: 0018-9448. DOI: 10.1109/TIT.2005.855594.

[62] W. Dai, Y. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds and applications to MIMO communications," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1108–1123, 2008, ISSN: 0018-9448. DOI: 10.1109/TIT.2007.915691.

[63] D. Agrawal, T. J. Richardson, and R. L. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2618–2626, Sep. 2001, ISSN: 0018-9448. DOI: 10.1109/18.945279.

[64] R. H. Gohary and T. N. Davidson, "Noncoherent MIMO communication: Grassmannian constellations and efficient detection," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1176–1205, Mar. 2009, ISSN: 0018-9448. DOI: 10.1109/TIT.2008.2011512.

[65] M. Beko, J. Xavier, and V. A. N. Barroso, "Noncoherent communication in multiple-antenna systems: Receiver design and codebook construction," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5703–5715, 2007, ISSN: 1053-587X. DOI: 10.1109/TSP.2007.901151.

[66] B. Tahir, S. Schwarz, and M. Rupp, "Constructing Grassmannian frames by an iterative collision-based packing," *IEEE Signal Process. Lett.*, vol. 26, no. 7, pp. 1056–1060, 2019, ISSN: 1070-9908. DOI: 10.1109/LSP.2019.2919391.

[67] L. Peng, D. Hu, L. Zhang, and Z. Qin, "Grassmannian constellation based on antipodal points and orthogonal design and its simplified detecting algorithm," *Journal of Computer Networks and Communications*, vol. 2017, 2017.

[68] M. L. McCloud, M. Brehler, and M. K. Varanasi, "Signal design and convolutional coding for noncoherent space-time communication on the block-Rayleigh-fading channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1186–1194, 2002, ISSN: 0018-9448. DOI: 10.1109/18.995648.

[69] Y. Wu, K. Ruotsalainen, and M. Juntti, "Unitary space-time constellation design based on the Chernoff bound of the pairwise error probability," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3842–3850, 2008, ISSN: 0018-9448. DOI: 10.1109/TIT.2008.926310.

[70] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 1962–1973, Sep. 2000, ISSN: 0018-9448. DOI: 10.1109/18.868472.

[71] V. Tarokh and I.-M. Kim, "Existence and construction of noncoherent unitary space-time codes," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3112–3117, 2002, ISSN: 0018-9448. DOI: 10.1109/TIT.2002.805075.

[72] W. Zhao, G. Leus, and G. B. Giannakis, "Orthogonal design of unitary constellations for uncoded and trellis-coded noncoherent space-time systems," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1319–1327, 2004, ISSN: 0018-9448. DOI: 10.1109/TIT.2004.828154.

[73] I. Kammoun and J. Belfiore, "A new family of Grassmann space-time codes for non-coherent MIMO systems," *IEEE Communications Letters*, vol. 7, no. 11, pp. 528–530, Nov. 2003, ISSN: 2373-7891. DOI: 10.1109/LCOMM.2003.820081.

[74] I. Kammoun, A. M. Cipriano, and J. C. Belfiore, "Non-coherent codes over the Grassmannian," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3657–3667, Oct. 2007, ISSN: 1536-1276. DOI: 10.1109/TWC.2007.06059.

[75] Y. Jing and B. Hassibi, "Unitary space-time modulation via Cayley transform," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2891–2904, 2003, ISSN: 1053-587X. DOI: 10.1109/TSP.2003.818202.

[76] J. Zhang, F. Huang, and S. Ma, "Full diversity blind space-time block codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6109–6133, 2011, ISSN: 0018-9448. DOI: 10.1109/TIT.2011.2161918.

[77] K. M. Attiah, K. Seddik, R. H. Gohary, and H. Yanikomeroglu, "A systematic design approach for non-coherent Grassmannian constellations," in *IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 2948–2952. DOI: 10.1109/ISIT.2016.7541839.

[78] P. Dayal, M. Brehler, and M. K. Varanasi, "Leveraging coherent space-time codes for noncoherent communication via training," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2058–2080, 2004, ISSN: 0018-9448. DOI: 10.1109/TIT.2004.833341.

[79] S. G. Srinivasan and M. K. Varanasi, "Constellation design for the noncoherent MIMO Rayleigh fading channel at general SNR," in *IEEE International Symposium on Information Theory (ISIT)*, Jul. 2006, pp. 1798–1802. DOI: 10.1109/ISIT.2006.261744.

[80] M. A. El-Azizy, R. H. Gohary, and T. N. Davidson, "A BICM-IDD scheme for non-coherent MIMO communication," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 541–546, Feb. 2009, ISSN: 1536-1276. DOI: 10.1109/TWC.2009.070887.

[81] J. Cabrejas, S. Roger, D. Calabuig, Y. M. M. Fouad, R. H. Gohary, J. F. Monserrat, and H. Yanikomeroglu, "Non-coherent open-loop MIMO communications over temporally-correlated channels," *IEEE Access*, vol. 4, pp. 6161–6170, 2016, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2016.2580680.

[82] J. Fanjul, J. Ibáñez, I. Santamaria, and C. Loucera, "Experimental evaluation of non-coherent mimo grassmannian signaling schemes," in *International Conference on Ad-Hoc Networks and Wireless*, Springer, 2017, pp. 221–230.

[83] R. H. Gohary and H. Yanikomeroglu, "Grassmannian signalling achieves tight bounds on the ergodic high-SNR capacity of the noncoherent MIMO full-duplex relay channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2480–2494, 2014.

[84] S. Shamai and T. L. Marzetta, "Multiuser capacity in block fading with no channel state information," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, pp. 938–942, Apr. 2002, ISSN: 0018-9448. DOI: 10.1109/18.992788.

[85] R. Gopalan, K. Padmanabhan, S. Ranganathan, and O. M. Collins, "Calculating and achieving capacity on the unknown fading MIMO channel," in *Proc. IEEE Int. Symp. Information Theory*, Jul. 2006, pp. 1579–1583. DOI: 10.1109/ISIT.2006.261542.

[86] G. Lin and S. M. Moser, "The fading number of a multiple-access Rician fading channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4983–4991, Aug. 2011, ISSN: 1557-9654. DOI: 10.1109/TIT.2011.2158470.

[87] R. Devassy, G. Durisi, J. Östman, W. Yang, T. Eftimov, and Z. Utkovski, "Finite-SNR bounds on the sum-rate capacity of Rayleigh block-fading multiple-access channels with no a priori CSI," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3621–3632, Oct. 2015, ISSN: 1558-0857. DOI: 10.1109/TCOMM.2015.2466536.

[88] Z. Utkovski, D. Ilik, and L. Kocarev, "An achievable pre-log region for the non-coherent block fading MIMO multiple access channel," in *The Tenth Int. Symp. Wireless Communication Systems (ISWCS)*, Aug. 2013, pp. 1–5.

[89] S. Murugesan, E. Uysal-Biyikoglu, and P. Schniter, "Optimization of training and scheduling in the non-coherent SIMO multiple access channel," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1446–1456, Sep. 2007, ISSN: 0733-8716. DOI: 10.1109/JSAC.2007.070917.

[90] M. Chowdhury, A. Manolakos, and A. Goldsmith, "Scaling laws for noncoherent energy-based communications in the SIMO MAC," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1980–1992, Apr. 2016, ISSN: 1557-9654. DOI: 10.1109/TIT.2016.2527679.

[91] A. Manolakos, M. Chowdhury, and A. Goldsmith, "Energy-based modulation for noncoherent massive SIMO systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7831–7846, 2016, ISSN: 1536-1276. DOI: 10.1109/TWC.2016.2608883.

[92] A. Schenk and R. F. H. Fischer, "Noncoherent detection in massive MIMO systems," in *17th International ITG Workshop on Smart Antennas (WSA)*, Mar. 2013, pp. 1–8.

[93] V. M. Baeza, A. G. Armada, W. Zhang, M. El-Hajjar, and L. Hanzo, "A noncoherent multiuser large-scale SIMO system relying on M-ary DPSK and BICM-ID," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1809–1814, 2018, ISSN: 0018-9545. DOI: 10.1109/TVT.2017.2750114.

[94] D. Kong, X. Xia, and T. Jiang, "A differential QAM detection in uplink massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6371–6383, 2016, ISSN: 1558-2248. DOI: 10.1109/TWC.2016.2583428.

[95] S. Li, J. Zhang, and X. Mu, "Noncoherent massive space-time block codes for uplink network communications," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 5013–5027, 2018, ISSN: 1939-9359. DOI: 10.1109/TVT.2018.2815981.

[96] Y. Yu, P. Chen, and J. Zhang, "Design of optimal finite alphabet NOMA scheme for uplink noncoherent massive MIMO channels," in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, 2019, pp. 1–6.

[97] H. Chen, Z. Dong, J.-K. Zhang, and B. Vucetic, "Design of non-orthogonal and noncoherent massive MIMO for scalable URLLC beyond 5G," *arXiv preprint arXiv:2001.10728*, 2020.

[98] M. Brehler and M. K. Varanasi, "Noncoherent multiuser space-time communications: Optimum receivers and signal design," in *Proceedings of the 35th Annual Conference Information Science System*, The Johns Hopkins University, Mar. 2001, pp. 379–383.

[99] S. Buzzi, M. Lops, and S. Sardellitti, "Performance of iterative data detection and channel estimation for single-antenna and multiple-antennas wireless communications," *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1085–1104, Jul. 2004, ISSN: 1939-9359. DOI: 10.1109/TVT.2004.830144.

[100] M. Abuthinien, S. Chen, and L. Hanzo, "Semi-blind joint maximum likelihood channel estimation and data detection for MIMO systems," *IEEE Signal Processing Letters*, vol. 15, pp. 202–205, 2008, ISSN: 1558-2361. DOI: 10.1109/LSP.2007.911758.

[101] Weiyu Xu, M. Stojnic, and B. Hassibi, "On exact maximum-likelihood detection for non-coherent MIMO wireless systems: A branch-estimate-bound optimization framework," in *IEEE International Symposium on Information Theory (ISIT)*, Jul. 2008, pp. 2017–2021. DOI: 10.1109/ISIT.2008.4595343.

[102] H. A. J. Alshamary and W. Xu, "Efficient optimal joint channel estimation and data detection for massive MIMO systems," in *IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 875–879. DOI: 10.1109/ISIT.2016.7541424.

[103] G. Caire, R. R. Muller, and T. Tanaka, "Iterative multiuser joint decoding: Optimal power allocation and low-complexity implementation," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1950–1973, 2004, ISSN: 0018-9448. DOI: 10.1109/TIT.2004.833351.

[104] J. Goldberger and A. Leshem, "MIMO detection for high-order QAM based on a Gaussian tree approximation," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4973–4982, 2011, ISSN: 0018-9448. DOI: 10.1109/TIT.2011.2159037.

[105] J. Céspedes, P. M. Olmos, M. Sánchez-Fernández, and F. Perez-Cruz, "Probabilistic MIMO symbol detection with expectation consistency approximate inference," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3481–3494, Apr. 2018, ISSN: 0018-9545. DOI: 10.1109/TVT.2017.2786638.

[106] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003, ISSN: 1557-9654. DOI: 10.1109/TIT.2003.813523.

[107] H. Weingarten, Y. Steinberg, and S. S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006, ISSN: 1557-9654. DOI: 10.1109/TIT.2006.880064.

[108] A. Bennatan and D. Burshtein, "On the fading-paper achievable region of the fading MIMO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 100–115, Jan. 2008, ISSN: 1557-9654. DOI: 10.1109/TIT.2007.911268.

[109] A. Gholami Davoodi and S. A. Jafar, "Transmitter cooperation under finite precision CSIT: A GDoF perspective," *IEEE Trans. Inf. Theory*, vol. 63, no. 9, pp. 6020–6030, Sep. 2017, ISSN: 1557-9654. DOI: 10.1109/TIT.2016.2619899.

[110] 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.

[111] 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, Mar. 2005, ISSN: 1557-9654. DOI: 10.1109/TIT.2004.842621.

[112] C. Huang, S. A. 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, ISSN: 1557-9654. DOI: 10.1109/TIT.2011.2173615.

[113] J. Nam, A. Adhikary, J. Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 876–890, Oct. 2014.

[114] J. Nam, G. Caire, and J. Ha, "On the role of transmit correlation diversity in multiuser MIMO systems," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 336–354, Jan. 2017.

[115] Y. Li and A. Nosratinia, "Product superposition for MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6839–6852, Nov. 2012, ISSN: 0018-9448. DOI: 10.1109/TIT.2012.2209862.

[116] Y. Li and A. Nosratinia, "Coherent product superposition for downlink multiuser MIMO," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1746–1754, Mar. 2015, ISSN: 1558-2248. DOI: 10.1109/TWC.2014.012314.121592.

[117] F. Zhang, M. Fadel, and A. Nosratinia, "Spatially correlated MIMO broadcast channel: Analysis of overlapping correlation eigenspaces," in *IEEE International Symposium on Information Theory (ISIT)*, Jun. 2017, pp. 1097–1101. DOI: 10.1109/ISIT.2017.8006698.

[118] K.-H. Ngo, S. Yang, and M. Guillaud, "The optimal DoF region for the two-user non-coherent SIMO multiple-access channel," in *IEEE Information Theory Workshop (ITW)*, arXiv preprint arXiv:1806.04102, Guangzhou, China, Nov. 2018.

[119] ——, "The optimal degrees of freedom for the point-to-point and multiple-access channels in generic block fading," *in preparation to submit to IEEE Trans. Inf. Theory*, 2020.

[120] K.-H. Ngo, S. Yang, and M. Guillaud, "An achievable DoF region for the two-user non-coherent MIMO broadcast channel with statistical CSI," in *IEEE Information Theory Workshop (ITW)*, Nov. 2017, pp. 604–608. DOI: 10.1109/ITW.2017.8277972.

[121] F. Zhang, K.-H. Ngo, S. Yang, and A. Nosratinia, "Transmit correlation diversity: Generalization, new techniques, and improved bounds," *submitted to IEEE Trans. Inf. Theory*, 2020.

[122] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Design and analysis of a practical codebook for non-coherent communications," in *51st Asilomar Conference on Signals, Systems, and Computers*, CA, USA, Oct. 2017, pp. 1237–1241. DOI: 10.1109/ACSSC.2017.8335549.

[123] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Cube-split: A structured Grassmannian constellation for non-coherent SIMO communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1948–1964, Mar. 2020, ISSN: 1558-2248. DOI: 10.1109/TWC.2019.2959781.

[124] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Transmitter and receiver communication apparatus for non-coherent communication," English, European pat. 6 860 390, 2018.

[125] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "A multiple access scheme for non-coherent SIMO communications," in *52nd Asilomar Conference on Signals, Systems, and Computers*, CA, USA, Oct. 2018, pp. 1846–1850. DOI: 10.1109/ACSSC.2018.8645403.

[126] K.-H. Ngo, S. Yang, M. Guillaud, and A. Decurninge, "Joint constellation design for the two-user non-coherent multiple-access channel," *arXiv preprint arXiv:2001.04970*, 2020.

[127] ——, "Joint constellation design for the non-coherent MIMO multiple-access channel," *in preparation to submit to a journal*, 2020.

[128] K.-H. Ngo, M. Guillaud, A. Decurninge, S. Yang, S. Sarkar, and P. Schniter, "Non-coherent multi-user detection based on expectation propagation," in *53rd Asilomar Conference on Signals, Systems, and Computers*, CA, USA, Nov. 2019.

[129] K.-H. Ngo, M. Guillaud, A. Decurninge, S. Yang, and P. Schniter, "Multi-user detection based on expectation propagation for the non-coherent SIMO multiple access channel," *IEEE Trans. Wireless Commun.*, 2019, (under revision, arXiv preprint arXiv:1905.11152).

[130] K.-H. Ngo, S. Yang, and M. Guillaud, "Generalized Gaussian model for data-driven learning in communications," in *International Zurich Seminar on Information and Communication (IZS)*, poster, Zurich, Switzerland, Feb. 2020.

[131] K.-H. Ngo, S. Yang, and M. Kobayashi, "Scalable content delivery with coded caching in multi-antenna fading channels," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 548–562, Jan. 2018, ISSN: 1536-1276. DOI: 10.1109/TWC.2017.2768361.

[132] A. Ghorbel, K.-H. Ngo, R. Combes, M. Kobayashi, and S. Yang, "Opportunistic content delivery in fading broadcast channels," in *IEEE Global Communications Conference (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6. DOI: 10.1109/GLOCOM.2017.8254966. [Online]. Available: https://arxiv.org/pdf/1702.02179.pdf.

[133] T. T. Q. Tran, V.-L. Nguyen, K.-H. Ngo, L.-T. Nguyen, Q.-T. Nguyen, E. Bastug, S. Azarian, M. Debbah, and P. Duhamel, "Network coding and information security in industry 4.0," in *1st ASEAN IVO Workshop on Cybersecurity and Information Security in Industry 4.0*, poster, Hanoi, Vietnam, Mar. 2019.

[134] W. Rudin, *Real and Complex Analysis*. Tata McGraw-hill education, 2006.

[135] I. Csiszar and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.

[136] I. S. Dhillon, J. R. Heath, T. Strohmer, and J. A. Tropp, "Constructing packings in Grassmannian manifolds via alternating projection," *Experimental mathematics*, vol. 17, no. 1, pp. 9–35, 2008.

[137] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, 2008, ISSN: 0733-8716. DOI: 10.1109/JSAC.2008.081002.

[138] M. H. Conde and O. Loffeld, "Fast approximate construction of best complex antipodal spherical codes," *arXiv preprint arXiv:1705.03280*, 2017.

[139] P. Shor and N. J. A. Sloane, "A family of optimal packings in Grassmannian manifolds," *Journal of Algebraic Combinatorics*, vol. 7, no. 2, pp. 157–163, 1998.

[140] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ: Princeton University Press, 2008, pp. xvi+224, ISBN: 978-0-691-13298-3.

[141] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a Matlab toolbox for optimization on manifolds," *Journal of Machine Learning Research*, vol. 15, pp. 1455–1459, 2014. [Online]. Available: http://www.manopt.org.

[142] W. W. Hager and H. Zhang, "A survey of nonlinear conjugate gradient methods," *Pacific Journal of Optimization*, vol. 2, no. 1, pp. 35–58, 2006.

[143] M. J. D. Powell, "Restart procedures for the conjugate gradient method," *Mathematical programming*, vol. 12, no. 1, pp. 241–254, 1977.

[144] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, 2nd. San Diego, CA, USA: Academic press, 1986, vol. 120.

[145] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd. New York, NY, USA: Cambridge University Press, 2012, ISBN: 0521548233, 9780521548236.

[146] S. Yang and S. Shamai (Shitz), "On the multiplexing gain of discrete-time MIMO phase noise channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 2394–2408, Apr. 2017, ISSN: 0018-9448. DOI: 10.1109/TIT.2017.2653718.

[147] J. Chen, S. Yang, and P. Elia, "On the fundamental feedback-vs-performance tradeoff over the MISO-BC with imperfect and delayed CSIT," in *IEEE International Symposium on Information Theory (ISIT)*, 2013, pp. 997–1001. [Online]. Available: https://arxiv.org/pdf/1302.0806.pdf.

[148] B. Wang and F. Zhang, "Some inequalities for the eigenvalues of the product of positive semidefinite Hermitian matrices," *Linear algebra and its applications*, vol. 160, pp. 113–118, 1992.

[149] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002, ISSN: 1558-0008. DOI: 10.1109/JSAC.2002.801223.

[150] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, P. Karlsson, and M. Beach, "Modeling of wide-band MIMO radio channels based on NLoS indoor measurements," *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 655–665, May 2004, ISSN: 1939-9359. DOI: 10.1109/TVT.2004.827164.

[151] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Pearson Education, 2015.

[152] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016. DOI: 10.1017/CBO9781316799895.

[153] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," vol. 48, no. 3, pp. 502–513, Mar. 2000.

[154] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," vol. 3, no. 4, pp. 1165–1175, Jul. 2004.

[155] E. A. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," vol. 3, no. 5, pp. 1543–1553, Sep. 2004, ISSN: 1536-1276.

[156] A. M. Tulino, A. Lozano, and S. Verdu, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.

[157] W. Chang, S. Chung, and Y. H. Lee, "Diversity-multiplexing tradeoff in rank-deficient and spatially correlated MIMO channels," in *IEEE International Symposium on Information Theory*, Jul. 2006, pp. 1144–1148. DOI: 10.1109/ISIT.2006.261963.

[158] E. Dall'Anese, A. Assalini, and S. Pupolin, "On the effect of imperfect channel estimation upon the capacity of correlated MIMO fading channels," in *IEEE Vehicular Technology Conference*, Apr. 2009, pp. 1–5.

[159] A. Soysal, "Tightness of capacity bounds in correlated MIMO systems with channel estimation error," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2010, pp. 667–671.

[160] T. Al-naffouri, M. Sharif, and B. Hassibi, "How much does transmit correlation affect the sum-rate scaling of MIMO Gaussian broadcast channels?," vol. 57, no. 2, pp. 562–572, Feb. 2009.

[161] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 550–560, Apr. 2002, ISSN: 1558-0008. DOI: 10.1109/49.995514.

[162] J.-. Lee, J.-. Ko, and Y.-. Lee, "Effect of transmit correlation on the sum-rate capacity of two-user broadcast channels," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2597–2599, Sep. 2009, ISSN: 1558-0857. DOI: 10.1109/TCOMM.2009.09.070513.

[163] J. W. Lee, H. N. Cho, H. J. Park, and Y. H. Lee, "Sum-rate capacity of correlated multi-user MIMO channels," in *Information Theory and Applications Workshop (ITA)*, Jan. 2010, pp. 1–5.

[164] A. Adhikary and G. Caire, "JSDM and multi-cell networks: Handling inter-cell interference through long-term antenna statistics," in *48th Asilomar Conference on Signals, Systems and Computers*, Nov. 2014, pp. 649–655.

[165] A. Adhikary, H. S. Dhillon, and G. Caire, "Massive-MIMO meets HetNet: Interference coordination through spatial blanking," vol. 33, no. 6, pp. 1171–1186, Jun. 2015.

[166] A. Adhikary, E. A. Safadi, and G. Caire, "Massive MIMO and inter-tier interference coordination," in *Information Theory and Applications Workshop (ITA)*, Feb. 2014, pp. 1–10.

[167] E. M. Luks, F. Rákóczi, and C. R. Wright, "Some algorithms for nilpotent permutation groups," *J. Symb. Comput.*, vol. 23, no. 4, pp. 335–354, Apr. 1997, ISSN: 0747-7171. DOI: 10.1006/jsco.1996.0092. [Online]. Available: http://dx.doi.org/10.1006/jsco.1996.0092.

[168] M. Guillaud, M. Debbah, and A. L. Moustakas, "A maximum entropy characterization of spatially correlated MIMO wireless channels," in *IEEE Wireless Communications and Networking Conference*, 2007, pp. 1039–1044. DOI: 10.1109/WCNC.2007.197.

[169] M. Debbah and R. R. Muller, "MIMO channel modeling and the principle of maximum entropy," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1667–1690, 2005. DOI: 10.1109/TIT.2005.846388.

[170] R. Diestel, *Graph Theory: 5th edition*, ser. Springer Graduate Texts in Mathematics. Springer-Verlag, © Reinhard Diestel, 2017, ISBN: 9783961340057.

[171] M. R. Garey and D. S. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co., 1990, ISBN: 0716710455.

[172] A. Decurninge and M. Guillaud, "Cube-split: Structured quantizers on the Grassmannian of lines," in *IEEE Wireless Communications and Networking Conference (WCNC)*, San Francisco, CA, USA, Mar. 2017.

[173] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1361–1391, Jul. 1999. DOI: 10.1109/18.771140.

[174] J. Kim, K. Cheun, and S. Choi, "Unitary space-time constellations based on quasi-orthogonal sequences," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 35–39, Jan. 2010, ISSN: 0090-6778. DOI: 10.1109/TCOMM.2010.01.080123.

[175] E. Agrell, J. Lassing, E. G. Strom, and T. Ottosson, "On the optimality of the binary reflected Gray code," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3170–3182, 2004, ISSN: 0018-9448. DOI: 10.1109/TIT.2004.838367.

[176] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, 1998, ISSN: 0018-9448. DOI: 10.1109/18.669123.

[177] N. H. Tran, H. H. Nguyen, and T. Le-Ngoc, "Coded unitary space-time modulation with iterative decoding: Error performance and mapping design," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 703–716, Apr. 2007, ISSN: 0090-6778. DOI: 10.1109/TCOMM.2007.894092.

[178] S. Baluja and M. Covell, "Neighborhood preserving codes for assigning point labels: Applications to stochastic search," in *International Conference on Computational Science*, 2013, pp. 956–965.

[179] G. W. K. Colman, R. H. Gohary, M. A. El-Azizy, T. J. Willink, and T. N. Davidson, "Quasi-Gray labelling for Grassmannian constellations," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, pp. 626–636, Feb. 2011, ISSN: 1536-1276. DOI: 10.1109/TWC.2011.121410.100404.

[180] J. Marcum, "A statistical theory of target detection by pulsed radar: Mathematical appendix," *IRE Trans. Inf. Theory*, vol. 6, no. 2, pp. 59–267, Apr. 1960, ISSN: 2168-2712. DOI: 10.1109/TIT.1960.1057560.

[181] S. Ten Brink, "Designing iterative decoding schemes with the extrinsic information transfer chart," *AEU Int. J. Electron. Commun*, vol. 54, no. 6, pp. 389–398, 2000.

[182] M. Tüchler, R. Koetter, and A. C. Singer, "Turbo equalization: Principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, May 2002. DOI: 10.1109/TCOMM.2002.1006557.

[183] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, Seventh. Elsevier/Academic Press, Amsterdam, 2007, pp. xlviii+1171, ISBN: 978-0-12-373637-6; 0-12-373637-4.

[184] 3rd Generation Partnership Project (3GPP), *Multiplexing and channel coding*, Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA);

[185] A. Lancho, J. Östman, G. Durisi, T. Koch, and G. Vazquez-Vilar, "Saddlepoint approximations for short-packet wireless communications," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2020.

[186] G. Durisi, T. Koch, J. Östman, Y. Polyanskiy, and W. Yang, "Short-packet communications over multiple-antenna Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 618–629, 2016.

[187] J. Östman, G. Durisi, E. G. Ström, M. C. Coskun, and G. Liva, "Short packets over block-memoryless fading channels: Pilot-assisted or noncoherent transmission?" *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1521–1536, 2019.

[188] C. Auderset, C. Mazza, and E. A. Ruh, "Angular Gaussian and Cauchy estimation," *J. Multivar. Anal.*, vol. 93, no. 1, pp. 180–197, Mar. 2005, ISSN: 0047-259X. DOI: 10.1016/j.jmva.2004.01.007. [Online]. Available: http://dx.doi.org/10.1016/j.jmva.2004.01.007.

[189] R. J. Baxley, B. T. Walkenhorst, and G. Acosta-Marum, "Complex Gaussian ratio distribution with applications for error rate calculation in fading channels with imperfect CSI," in *IEEE Global Telecommunications Conference (GLOBECOM)*, 2010. DOI: 10.1109/GLOCOM.2010.5683407.

[190] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, ser. Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley & Sons, 1995, vol. 2, ISBN: 9780471584940.

[191] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2383–2399, 2001, ISSN: 1557-9654. DOI: 10.1109/18.945253.

[192] R. J. Muirhead, *Aspects of multivariate statistical theory*. John Wiley & Sons, 2009, vol. 197.

[193] I. R. Savage, "Probability inequalities of the Tchebycheff type," *Journal of Research of the National Bureau of Standards-B. Mathematics and Mathematical Physics*, vol. 65B, no. 3, pp. 211–222, 1961.

[194] A. Buluç, H. Meyerhenke, I. Safro, P. Sanders, and C. Schulz, "Recent advances in graph partitioning," in *Algorithm Engineering*, Springer, 2016, pp. 117–158.

[195] H. Robbins, "A remark on Stirling's formula," *The American mathematical monthly*, vol. 62, no. 1, pp. 26–29, 1955.

[196] J. Wishart, "The generalised product moment distribution in samples from a normal multivariate population," *Biometrika*, vol. 20, no. 1-2, pp. 32–52, Dec. 1928.

[197] T. Ratnarajah and R. Vaillancourt, "Complex singular Wishart matrices and applications," *Computers & Mathematics with Applications*, vol. 50, no. 3-4, pp. 399–411, 2005.

[198] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. New York, NY, USA: Cambridge University Press, 2011, ISBN: 1107011639, 9781107011632.

[199] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Third. The Johns Hopkins University Press, 1996.

[200] T. P. Minka, "A family of algorithms for approximate Bayesian inference," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, Jan. 2001.

[201] T. Heskes, M. Opper, W. Wiegerinck, O. Winther, and O. Zoeter, "Approximate inference techniques with expectation constraints," *J. Stat. Mech: Theory Exp.*, p. 11 015, 2005.

[202] P. J. Bickel and K. A. Doksum, *Mathematical statistics : Basic ideas and selected topics*, English. Holden-Day San Francisco, 1977, ix, 492 p. : ISBN: 0816207844.

[203] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001, ISSN: 0018-9448. DOI: 10.1109/18.910572.

[204] C. F. Van Loan and N. Pitsianis, "Approximation with Kronecker products," in *Linear algebra for large scale and real-time applications*, Springer, 1993, pp. 293–314.

[205] A. Vehtari, A. Gelman, T. Sivula, P. Jylänki, D. Tran, S. Sahai, P. Blomstedt, J. P. Cunningham, D. Schiminovich, and C. Robert, "Expectation propagation as a way of life: A framework for bayesian inference on partitioned data," *arXiv preprint arXiv:1412.4869*, 2014.

[206] P. Sun, C. Zhang, Z. Wang, C. N. Manchón, and B. H. Fleury, "Iterative receiver design for ISI channels using combined belief- and expectation-propagation," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1733–1737, Oct. 2015, ISSN: 1070-9908. DOI: 10.1109/LSP.2015.2404822.

[207] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017, ISSN: 1932-8346. DOI: 10.1561/2000000093.

[208] C.-P. Schnorr and M. Euchner, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," *Mathematical programming*, vol. 66, no. 1-3, pp. 181–199, 1994.

[209] P. Bromiley, "Products and convolutions of Gaussian probability density functions," *Tina-Vision Memo*, vol. 3, no. 4, p. 1, 2003.

[210] M. J. Wainwright, T. S. Jaakkola, and A. S. Willsky, "A new class of upper bounds on the log partition function," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2313–2335, Jul. 2005, ISSN: 0018-9448. DOI: 10.1109/TIT.2005.850091.

[211] V. I. Morgenshtern, E. Riegler, W. Yang, G. Durisi, S. Lin, B. Sturmfels, and H. Bölcskei, "Capacity pre-log of noncoherent SIMO channels via Hironaka's theorem," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4213–4229, 2013.

[212] G. Koliander, E. Riegler, G. Durisi, and F. Hlawatsch, "Degrees of freedom of generic block-fading MIMO channels without a priori channel state information," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7760–7781, 2014.

[213] M. T. Hussien, K. G. Seddik, R. H. Gohary, M. Shaqfeh, H. Alnuweiri, and H. Yanikomeroglu, "Multi-resolution broadcasting over the Grassmann and Stiefel manifolds," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, 2014, pp. 1907–1911.

[214] K. G. Seddik, R. H. Gohary, M. T. Hussien, M. Shaqfeh, H. Alnuweiri, and H. Yanikomeroglu, "Multi-resolution multicasting over the Grassmann and Stiefel

manifolds," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5296–5310, 2017.

[215] M. A. ElMossallamy, K. G. Seddik, and R. H. Gohary, "Multi-resolution multicasting using Grassmannian codes and space shift keying," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 988–992, 2019.

[216] J. Cabrejas, D. Martín-Sacristán, S. Roger, D. Calabuig, and J. F. Monserrat, "On the integration of Grassmannian constellations into LTE networks: A link-level performance study," in *13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2017, pp. 1824–1829.

[217] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1711–1726, 2016.

[218] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[219] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Diversity versus channel knowledge at finite block-length," in *Proc. IEEE Information Theory Workshop (ITW)*, 2012, pp. 572–576.

[220] A. Lancho, T. Koch, and G. Durisi, "On single-antenna Rayleigh block-fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 496–519, 2020.

[221] P. Walk, P. Jung, and B. Hassibi, "MOCZ for blind short-packet communication: Basic principles," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5080–5097, 2019.

[222] Z. Utkovski, T. Eftimov, and P. Popovski, "Random access protocols with collision resolution in a noncoherent setting," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 445–448, 2015.

[223] K. Senel and E. G. Larsson, "Grant-free massive MTC-enabled massive MIMO: A compressive sensing approach," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6164–6175, 2018.

[224] A. Decurninge, I. Land, and M. Guillaud, *Tensor-based modulation for unsourced massive random access*, 2020. arXiv: 2006.06797 [cs.IT].

[225] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. on Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, 2017.

[226] S. Dörner, S. Cammerer, J. Hoydis, and S. t. Brink, "Deep learning based communication over the air," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 132–143, 2018.

[227] Z. Qin, H. Ye, G. Y. Li, and B. F. Juang, "Deep learning in physical layer communications," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 93–99, 2019.

[228] J. Huang, M. Lan, H. Zhang, C. Huang, W. Zhang, and S. Cui, "Design of noncoherent communications: From statistical method to machine learning," *IEEE Wireless Commun.*, vol. 27, no. 1, pp. 76–83, 2020.

[229] J. Zhang, G. Zhu, R. W. Heath Jr, and K. Huang, "Grassmannian learning: Embedding geometry awareness in shallow and deep learning," *arXiv preprint arXiv:1808.02229*, 2018.

[230] Y. Du, G. Zhu, J. Zhang, and K. Huang, "Automatic recognition of space-time constellations by learning on the Grassmann manifold," *IEEE Trans. Signal Process.*, vol. 66, no. 22, pp. 6031–6046, 2018.

[231] F. A. Aoudia and J. Hoydis, "Model-free training of end-to-end communication systems," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2503–2516, 2019.

[232] T. Van Luong, Y. Ko, N. A. Vien, M. Matthaiou, and H. Q. Ngo, "Deep energy autoencoder for noncoherent multicarrier MU-SIMO systems," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2020.

[233] S. Xue, Y. Ma, N. Yi, and R. Tafazolli, "On deep learning solutions for joint transmitter and noncoherent receiver design in MU-MIMO systems," *arXiv preprint arXiv:2004.06599*, 2020.

About the Author

Khac-Hoang NGO
Dept. of Telecommunications, CentraleSupélec, University of Paris-Saclay
3 rue Joliot Curie, 91192 Gif-sur-Yvette, France
Email: khachoang.ngo@supelec.fr Homepage: sites.google.com/site/khachoangngo

Education

2017 - 2020	Ph.D., Wireless Communications, CentraleSupélec, Univ. of Paris-Saclay, France
2015 - 2016	M.Sc., Wireless Communications, CentraleSupélec, Univ. of Paris-Saclay, France
2010 - 2014	B.E., Electronics and Telecommunications, Univ. of Engineering and Technology (UET), Vietnam National Univ., Hanoi (VNU), Vietnam

Experience

2016 - 2020	Research Engineer, Mathematical and Algorithmic Sciences Lab., Paris Research Center, Huawei Technologies France
2016	Research Assistant, L2S - CentraleSupélec, France
2012 - 2015	Research Assistant, UET, VNU, Vietnam
2013 - 2014	Intern, Vietnam Posts and Telecommunications Group, Vietnam
2012	Intern, Dept. of Electrical and Computer Engineering, National Univ. of Singapore

Professional Activities

Nov. 2019	Participant, 2nd Global Young Vietnamese Scholars Forum, Vietnam
Sep. 2019	Participant, 7th Heidelberg Laureate Forum, Germany
Feb. 2019	Scientific Committee Member, Junior Conference on Wireless and Optical Communications, Univ. of Paris-Saclay
Nov. 2015	Executive Committee Member, Young Engineers and Scientists Forum, Tokyo, Japan
2017 - now	Reviewer for IEEE Journals and Conferences
2019 - now	Copyeditor, ICT Research Journal, Vietnam Ministry of Information and Communications

Honors

2016	Graduate with first-class honors (master level), CentraleSupélec
2015	Univ. Paris-Saclay scholarship for international master students
2014	Graduate with first-class honors (bachelor level), UET, VNU
2013	Honda Young Engineers and Scientists Award in Vietnam, Honda Foundation, Japan

Personal

Born in 1992, Bac Giang, Vietnam

université PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre: Communications Sans Fil Noncohérentes: Limites Fondamentales et Conception du Système

Mots clés: communications noncohérentes, évanouissement par bloc, information sur l'état du canaux, degrés de liberté, constellations Grassmanniennes, détection multi-utilisateurs

Résumé: Dans les communications sans fil sur des canaux à évanouissement, en particulier à antennes multiples, la connaissance instantanée des coefficients de canal, appelés informations d'état de canal (CSI), est essentielle car elle permet d'adapter la transmission et la réception aux conditions actuelles du canal. La communication avec CSI *a priori* au niveau du récepteur est dite *cohérente*. En pratique, cependant, le CSI n'est pas disponible avant la communication et doit être estimé à un coût qui ne doit pas être ignoré, en particulier dans un environnement hautement mobile. Ainsi, la communication sans CSI *a priori*, également appelée communication *noncohérente*, est un cadre plus pratique et général. Cette thèse contribue à la compréhension des limites théoriques des communications noncohérentes, ainsi qu'à la conception d'un système de communication pratique noncohérent à évanouissement par bloc. Nous considérons trois scénarios: le canal point à point (P2P), le canal à accès multiple (MAC) et le canal de diffusion (BC).
Dans la première partie, nous étudions les limites fondamentales des communications noncohérentes en termes de débit de données et de degrés de liberté (DoF) réalisables. Nous considérons un évanouissement par bloc générique dans lequel le canal a une entropie différentielle finie et un second moment fini. Nous établissons d'abord le DoF optimal pour le canal P2P noncohérent à entrées multiples et sorties multiples (MIMO) en utilisant l'approche de la dualité pour borner les informations mutuelles. Deuxièmement, en utilisant une approche de dualité similaire, nous dérivons la région de DoF optimale pour le MAC SIMO à deux utilisateurs, qui peut être obtenue par partage de temps entre des schémas à pilotes simples. Troisième, nous dérivons les régions débit et DoF réalisable pour le BC MIMO noncohérent avec un évanouissement spatialement corrélé en exploitant la diversité de corrélation de transmission, qui est la différence entre la corrélation subie par différents utilisateurs. Ce faisant, nous concevons soigneusement des schémas de transmission basés sur des pilotes et sur le partage de débit, la superposition de produits et une combinaison de ceux-ci pour transmettre efficacement des signaux dans les parties communes et mutuellement exclusives des sous-espaces de corrélation.
Dans la deuxième partie, nous concevons la constellation et les schémas de détection efficaces pour les communications noncohérentes sur le canal d'évanouissement de type Rayleigh par bloc. Premièrement, nous proposons une constellation Grassmannienne structurée pour le canal P2P SIMO qui est simple à générer, a une efficacité d'empilement élevée, admet un étiquetage binaire simple mais efficace et permet une détection efficace douce et dure. Deuxièmement, nous étudions la conception de constellation conjointe pour le MAC MIMO. Nous introduisons des critères de conception simples et efficaces afin de minimiser l'erreur de détection conjointe et proposons quelques constructions de constellation simples. Troisièmement, nous proposons un schéma de détection souple multi-utilisateurs noncohérent pour le MAC SIMO à l'évanouissement de Rayleigh corrélé spatialement basé sur l'inférence approximative par propagation d'espérance. Ce schéma présente une complexité polynomiale dans la dimension du canal tout en produisant des marginaux postérieurs approximatifs par utilisateur précis conduisant à des performances d'erreur quasi-optimales.

Title: Noncoherent Wireless Communications: Fundamental Limits and System Design

Keywords: noncoherent communications, block fading, channel state information, degrees of freedom, Grassmannian constellations, multi-user detection

Abstract: In wireless communication over fading channels, especially multiple-antenna communication, the instantaneous knowledge of channel coefficients, so-called channel state information (CSI), is critical because it enables to adapt the transmission and reception to current channel conditions. The communication with *a priori* CSI at the receiver is said to be *coherent*. In practice, however, CSI is not granted for free prior to communication and needs to be estimated at a cost that should not be ignored, especially in a highly mobile environment. Thus, communication without *a priori* CSI, also known as *noncoherent* communication, is a more practical and general framework. This thesis contributes to the understanding of the theoretical limits of noncoherent communications, as well as the design of a practical noncoherent communication system in block fading. We consider three scenarios: the point-to-point (P2P) channel, the multiple-access channel (MAC), and the broadcast channel (BC).
In the first part, we study the fundamental limits of noncoherent communications in terms of achievable data rate and degrees of freedom (DoF). We consider generic block fading in which the channel has finite differential entropy and finite second moment. First, we derive the optimal DoF for the noncoherent multiple-input multiple-output (MIMO) P2P channel by using the duality approach to bound the input-output mutual information. Second, using a similar duality approach, we derive the optimal DoF region for the two-user noncoherent single-input multiple-output (SIMO) MAC, which can be achieved by time sharing between simple pilot-based schemes. Third, we derive achievable rate and DoF regions for the noncoherent MIMO BC with spatially correlated fading by exploiting the transmit correlation diversity, which is the difference between the correlation experienced by different users. In doing so, we carefully design pilot-based transmission schemes based on rate splitting, product superposition, and a combination of them to effectively transmit signals in both the common and mutually exclusive parts of the correlation subspaces.
In the second part, we design the constellation and efficient detection schemes for noncoherent communications over Rayleigh block fading channel. First, we propose a structured Grassmannian constellation for the SIMO P2P channel that is simple to generate, has high packing efficiency, admits a simple yet effective binary labeling, and allows for efficient soft and hard detection. Second, we investigate joint constellation design for the MIMO MAC. We introduce some simple and effective design criteria so as to minimize the joint detection error, and propose some simple constellation constructions. Third, we propose a noncoherent multi-user soft detection scheme for the SIMO MAC in spatially correlated Rayleigh fading based on expectation propagation approximate inference. This scheme has polynomial complexity in the channel dimension while producing accurate approximate per-user posterior marginals leading to near-optimal error performance.

Université Paris-Saclay

Espace Technologique / Immeuble Discovery

Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France