

Research Reports from the Communications Research Laboratory
at Ilmenau University of Technology

Bin Song

**Linear Transmit-Receive Strategies
for Multi-user MIMO
Wireless Communications**



ILMENAU UNIVERSITY OF TECHNOLOGY



Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität Ilmenau

LINEAR TRANSMIT-RECEIVE STRATEGIES
FOR MULTI-USER MIMO
WIRELESS COMMUNICATIONS

February 17, 2014

Bin Song
geboren am 01.12.1977 in China

Dissertation zur Erlangung des
akademischen Grades Doktor-Ingenieur (Dr.-Ing)

Anfertigung im: Fachgebiet Nachrichtentechnik
Institut für Informationstechnik
Fakultät für Elektrotechnik und Informationstechnik

Gutachter: Univ.-Prof. Dr.-Ing. Martin Haardt
Univ.-Prof. Dr.-Ing. Gerald Schuller
Prof. Rodrigo C. de Lamare

Vorgelegt am: 05.12.2013

Verteidigt am: 29.01.2014

urn:nbn:de:gbv:ilm1-2014000027

Acknowledgments

This thesis is a result of my research works during my Ph.D. study. It would not have been possible without the help and support from the kind people around me.

Firstly, I would like to express my greatest gratitude to Prof. Martin Haardt for his supervision and guidance. His constant encouragement, valuable suggestions, and fruitful discussions help me to be productive and have significantly influenced the contents of this thesis. I was very lucky and really enjoyed to work with him in a free and open-minded working environment.

I am grateful to Prof. Dr. Rodrigo de Lamare and Prof. Gerald Schuller for taking the time to be the reviewers of my thesis. The comments and discussions have effectively improved the quality of the thesis and are also very valuable for my future works.

In addition, a deep appreciation to Prof. Arie Yeredor for our fruitful collaboration on closed-form coordinated beamforming design presented in Chapter 4 of this thesis. Without his perceptive comments, professional ideas, and great patience, our collaboration work would not have been successful.

Many thanks to all of my colleagues for building lovely working environments. I appreciate every help from you and every talk we had. Special thanks goes to Veljko Stankovic and Florian Roemer, who introduced me linear precoding techniques when I started my work at Communications Research Laboratory, for sharing their bright ideas, skills and knowledge, for helping me to solve difficult technical problems when help was really needed. I would also like to thank Martin Fuchs-Lautensack, Martin Weis, Dominik Schulz, Yao Cheng, Nuan Song, Jens Steinwandt, Bilal Zafar, Jianshu Zhang, Mike Wolf, Jianhui Li, Peng Li and Keke Zu for fruitful technical discussions and personal talks and for being good friends. Additionally, special thanks to Wolfgang Erdtmann and our secretary Ms. Christina Patotschka for being always available for the problems of computer and organization.

I wish to thank my parents for their endless love and support throughout my life, for sharing my sadness and happiness, and for helping me out of the setbacks. Finally, I would like to give my deepest appreciation to my husband and my little boy. Every moment of care, hug, and smile is the source of my energy. Thank you for coming into my life and making it colorful and meaningful.

Abstract

The ever growing demand for reliable high data rates, enlarged coverage, and spectral efficiency in the existing third generation (3G) and fourth generation (4G) of mobile communication systems and future systems, has inspired intensive research efforts in the field of multi-user multiple-input multiple-output (MIMO) communications. In order to combat interference and exploit large multiplexing gains of the multi-antenna systems, a particular interest in spatial division multiple access (SDMA) techniques has emerged. Linear precoding techniques, as one of the SDMA strategies, have obtained more attention due to the fact that an increasing number of users and antennas involved into the existing and future mobile communication systems requires a simplification of the precoding design. Therefore, this thesis contributes to the design of linear transmit and receive strategies for multi-user MIMO broadcast channels in a single cell and clustered multiple cells. We efficiently solve some open problems in the existing linear transmission strategies and further enhance some existing algorithms.

First, we present a throughput approximation framework for multi-user MIMO broadcast channels employing regularized block diagonalization (RBD) linear precoding. Comparing dirty paper coding (DPC, which can achieve the capacity of multi-user MIMO broadcast channels) and linear precoding algorithms (e.g., zero forcing (ZF) and block diagonalization (BD)), we further quantify lower and upper bounds of the rate and power offset between them as a function of the system parameters such as the number of users and antennas. These analytical results are useful from the system design perspective.

Next, we develop a novel closed-form coordinated beamforming (CBF) algorithm (i.e., sequentially drilled joint congruence (SeDJoCo) transformation based closed-form CBF) to solve the existing open problem of CBF. To the best of our knowledge, with the exception of our new algorithm, so far only one algorithm has been proposed as a closed-form CBF which can only support a MIMO system with two users and two transmit antennas. In contrast, our new algorithm solves this problem and can support a MIMO system with an arbitrary number of users and transmit antennas. Moreover, the application of our new algorithm is not only for CBF, but also for blind source separation (BSS), since the same mathematical model has been used in BSS application.

Then, we further propose a new iterative CBF algorithm (i.e., flexible coordinated beamforming (FlexCoBF)) for multi-user MIMO broadcast channels. Compared to the existing iterative CBF algorithms, the most promising advantage of our new algorithm is that it pro-

vides freedom in the choice of the linear transmit and receive beamforming strategies, i.e., any existing linear precoding method can be chosen as the transmit strategy and the receive beamforming strategy can be flexibly chosen from maximum ratio combining (MRC) or minimum mean square error (MMSE) receivers. In other words, FlexCoBF algorithm does not require special designs of the transmit-receive beamforming vectors in contrast to the existing iterative CBF algorithms. Considering clustered multiple cell scenarios, we extend the FlexCoBF algorithm further and introduce the concept of the coordinated multipoint (CoMP) transmission which is one hot topic in fourth generation (4G) and fifth generation (5G) system design. Our objective is to efficiently reduce the inter-cluster and intra-cluster interference and enhance the performance of the cluster edge users in terms of users' throughput.

Finally, we present three strategies for channel state information (CSI) acquisition regarding various channel conditions and channel estimation strategies such as the time-varying correlated channel model, the finite rate feedback channel, and blind channel estimation techniques. The CSI knowledge is required at the base station in order to implement SDMA techniques. The quality of the obtained CSI heavily affects the system performance. The performance enhancement achieved by our new strategies has been demonstrated by numerical simulation results in terms of the system sum rate and the bit error rate.

Zusammenfassung

Die stetig wachsende Nachfrage nach zuverlässig hohen Durchsätzen, größeren Reichweiten und verbesserten spektralen Effizienzen in Mobilkommunikationssystemen der dritten und vierten Generation war der Auslöser für starke Forschungsanstrengungen im Bereich von Mehrnutzer-Systemen mit „Multiple-Input Multiple-Output“-Technik (MIMO-Technik). Die Notwendigkeit zur Unterdrückung von Interferenzen auf der einen Seite und zur Ausnutzung der durch Mehrfachzugriffsverfahren erzielbaren Gewinne auf der anderen Seite rückte die räumlichen Mehrfachzugriffsverfahren (Space Division Multiple Access, SDMA) in den Fokus der Forschung. Ein Vertreter der räumlichen Mehrfachzugriffsverfahren, die lineare Vorkodierung, fand aufgrund steigender Anzahl an Nutzern und Antennen in heutigen und zukünftigen Mobilkommunikationssystemen besondere Beachtung, da diese Verfahren das Design von Algorithmen zur Vorkodierung vereinfachen. Aus diesem Grund leistet diese Dissertation einen Beitrag zur Entwicklung linearer Sende- und Empfangstechniken für MIMO-Technologie mit mehreren Nutzern. Die vorliegende Arbeit konzentriert sich dabei auf den Fall der Broadcast-Kanäle in einzelnen Mobilkommunikationszellen oder in Clustern derselben. In diesem Zusammenhang stellen wir effiziente Ansätze vor, die offene Probleme in aktuellen linearen Übertragungsverfahren lösen. Weiterhin zeigen wir, wie aktuelle Algorithmen verbessert werden können.

Zunächst stellen wir ein Framework zur Approximation des Datendurchsatzes in Broadcast-MIMO-Kanälen mit mehreren Nutzern vor. In diesem Framework nehmen wir das lineare Vorkodierverfahren regularisierte Blockdiagonalisierung (RBD) an. Durch den Vergleich von Dirty Paper Coding (DPC, erreicht die Kapazität eines Broadcast-MIMO-Kanals mit mehreren Nutzern) und linearen Vorkodieralgorithmen (z.B. Zero Forcing (ZF) und Blockdiagonalisierung (BD)) ist es uns möglich, untere und obere Schranken für den Unterschied bezüglich Datenraten und bezüglich Leistung zwischen beiden anzugeben. Hierbei sind die Schranken Funktionen der Systemparameter - wie die Anzahl der Nutzer und Antennen. Aus Sicht des Systemdesigns sind dies nützliche analytischen Ergebnisse.

Im Weiteren entwickeln wir einen Algorithmus für koordiniertes Beamforming (Coordinated Beamforming, CBF), dessen Lösung sich in geschlossener Form angeben lässt. Dieser CBF-Algorithmus basiert auf der SeDJoCo-Transformation „Sequentially Drilled Joint Congruence“-Transformation) und löst bisher vorhandene Probleme im Bereich CBF. Nach Wissen der Autoren gibt es - mit Ausnahme unseres neuen Algorithmus - nur ein einziges

CBF-Verfahren mit geschlossener Lösung. Dieses unterstützt jedoch nur MIMO-Systeme mit zwei Nutzern und zwei Sendeantennen. Demgegenüber lässt unser neues Verfahren eine beliebiger Anzahl von Nutzern und Sendeantennen zu. Außerdem ist unser neues Verfahren nicht auf CBF als Anwendung beschränkt, da dasselbe mathematische Modell auch in der blinden Trennung von Quellen (Blind Source Separation, BSS) auftritt.

Im Anschluss schlagen wir einen iterativen CBF-Algorithmus namens FlexCoBF (flexible coordinated beamforming) für MIMO-Broadcast-Kanäle mit mehreren Nutzern vor. Im Vergleich mit bis dato existierenden iterativen CBF-Algorithmen kann als vielversprechendster Vorteil die freie Wahl der linearen Sende- und Empfangsstrategie herausgestellt werden. Das heißt, jede existierende Methode der linearen Vorkodierung kann als Sendestrategie genutzt werden, während die Strategie zum Empfangsbeamforming frei aus Maximum Ratio Combining (MRC) oder der Methode der kleinsten, mittleren Fehlerquadrate (Minimum Mean Square Error, MMSE) gewählt werden darf. Mit anderen Worten: FlexCoBF setzt keine spezielle Form des Sende-Empfangs-Beamformingvektors voraus, wie es bei bisherigen iterativen CBF-Algorithmen der Fall ist. Im Hinblick auf Szenarien, in denen Mobilfunkzellen in Clustern zusammengefasst sind, erweitern wir FlexCoBF noch weiter. Hier wurde das in Mobilfunknetzen der dritten und vierten Generation viel diskutierte Konzept der koordinierten Mehrpunktverbindung (Coordinated Multipoint (CoMP) transmission) integriert. Unser Ziel ist es, die Kommunikation innerhalb und zwischen den Clustern auf möglichst effiziente Weise zu verringern und gleichzeitig den Datendurchsatz für Nutzer an den Zellrändern zu verbessern.

Zuletzt stellen wir drei Möglichkeiten vor, Kanalzustandsinformationen (Channel State Information, CSI) unter verschiedenen Kanalumständen zu erlangen. Mögliche Umstände sind der zeitvariante, korrelierte Kanal, der Kanal mit endlichem Durchsatz im Rückkanal sowie der Einsatz blinder Kanalschätzalgorithmen. Kanalzustandsinformationen an der Basisstation sind Voraussetzung, um SDMA-Techniken einsetzen zu können. Die Qualität der Kanalzustandsinformationen hat einen starken Einfluss auf die Güte des Übertragungssystems. Die durch unsere neuen Algorithmen erzielten Verbesserungen haben wir mittels numerischer Simulationen von Summenraten und Bitfehlerraten belegt.

Contents

Acknowledgments	5
Abstract	7
Zusammenfassung	9
Contents	11
List of Figures	15
List of Tables	18
1. Introduction	1
1.1. Background and Motivations	1
1.2. Scope of the thesis and contributions	8
2. Fundamentals of MIMO wireless communication	15
2.1. Introduction and Motivation	15
2.2. Benefits of MIMO technology	16
2.2.1. Spatial diversity gain	16
2.2.2. Spatial multiplexing gain	17
2.2.3. Array gain	17
2.2.4. Interference reduction and avoidance	18
2.3. MIMO channel models	18
2.3.1. Physical channel models	19
2.3.2. Analytical channel models	21
2.4. Capacity region of systems with MIMO	23
2.4.1. Single-User MIMO Capacity	25
2.4.2. Multi-User MIMO Capacity	28
2.4.3. Multi-cell MIMO capacity	32
2.5. Summary and conclusions	35
3. Linear precoding techniques for multi-user MIMO systems	37

3.1. Introduction and Motivation	37
3.2. Previous work on linear precoding techniques	38
3.2.1. Zero Forcing precoding	39
3.2.2. Block Diagonalization Precoding	40
3.2.3. Regularized Block Diagonalization	42
3.3. Achievable throughput approximation at high SNRs	44
3.3.1. System Model	45
3.3.2. Capacity Approximation Framework	46
3.3.3. Throughput approximation for DPC	47
3.3.4. Throughput approximation for ZF	48
3.3.5. Throughput approximation for BD	50
3.3.6. Throughput approximation for RBD under the condition $M_T \geq M_R$. . .	52
3.3.7. Throughput approximation for RBD under the condition $M_T < M_R$. . .	56
3.4. Summary and Conclusions	60
4. Coordinated beamforming techniques	61
4.1. Introduction and Motivation	61
4.2. Previous work on coordinated beamforming techniques	62
4.2.1. Iterative coordinated beamforming	63
4.2.2. Closed-form coordinated beamforming	65
4.3. SeDJoCo transformation based closed-form CBF	66
4.3.1. Motivation in ML Blind or Semi-Blind Source Separation	69
4.3.2. Equivalent formulations and existence of a solution	71
4.3.3. Solutions of SeDJoCo	75
4.3.4. Convergence behavior	85
4.3.5. Achievable sum rate of SeDJoCo-based CBF	88
4.4. Flexible coordinated beamforming (FlexCoBF)	91
4.4.1. FlexCoBF in a single cell	93
4.4.2. FlexCoBF in clustered cellular MIMO network	102
4.5. Summary and Conclusions	120
5. Channel State Information Acquisition	125
5.1. Introduction and Motivation	125
5.2. Linear precoding with long-term channel state information	129
5.2.1. System and data model	129
5.2.2. Previous long-term CSI method	130

5.2.3. ROLT-CSI	131
5.2.4. Simulation Results	132
5.3. Quantized finite rate feedback for multi-user MIMO broadcast channels	137
5.3.1. Quantization scheme for ZF based system	138
5.3.2. Quantization scheme for BD based system	143
5.3.3. Quantization scheme for RBD based system	147
5.4. Blind channel estimation using a tensor-based subspace method	155
5.4.1. Notations and operations of tensor	155
5.4.2. Higher-Order SVD decomposition	157
5.4.3. Tensor-based blind estimation of SIMO channels	166
5.4.4. Tensor-based blind estimation of MIMO channels	181
5.5. Summary and Conclusions	190
6. Conclusions and Outlook	195
6.1. Conclusions	195
6.2. Future Works	197
Appendix A. Glossary of Acronyms, Symbols and Notation	200
A.1. Acronyms	200
A.2. Symbols and Notation	201
Appendix B. Proofs and derivations for Chapter 2	204
B.1. Proof of equation (2.14)	204
B.2. Schur-convex and Schur-concave function	204
Appendix C. Proofs and derivations for Chapter 3	206
C.1. Proof of equations (3.24) to (3.25)	206
C.2. Notes on several fundamental functions and distributions	207
C.2.1. Gamma function	207
C.2.2. Digamma function	208
C.2.3. Wishart distribution	208
C.2.4. Chi-square distribution	209
C.3. Proof of Theorem 3.3.4	209
C.4. Proof of Theorem 3.3.5	210
Appendix D. Proofs and derivations for Chapter 4	212
D.1. Proof of Theorem (4.2.1)	212

D.2. Notes on inverse discrete time Fourier transform (IDTFT)	212
D.3. Notes on Lie group	213
Appendix E. Proofs and derivations for Chapter 5	214
E.1. Notes on Jensen's inequality	214
E.2. Notes on Euclidean space and Grassmannian space	214
E.2.1. Euclidean space	214
E.2.2. Grassmannian space	214
E.3. Proof of Theorem 5.3.7	215
Bibliography	217

List of Figures

1.1. Generations of mobile communications and their keywords.	1
1.2. Block diagram of multi-user MIMO BC.	5
1.3. Block diagram of multi-user MIMO MAC.	6
2.1. The block diagram of a equivalent baseband MIMO communication system. . .	15
3.1. Achievable throughput for a system with configuration $\{1, 1, 1, 1\} \times 4$	44
3.2. Achievable throughput for a system with configuration $\{2, 2\} \times 4$	45
3.3. An example of the capacity approximation at high SNRs.	47
3.4. Achievable throughputs approximation for DPC, BD and ZF at high SNRs . .	52
3.5. Achievable throughputs approximation for RBD at high SNRs with $M_r = 2$. . .	57
3.6. Achievable throughputs approximation for RBD at high SNRs with $M_r = 2$. . .	57
3.7. Achievable throughputs approximation for RBD at high SNRs with $M_r = 3$. . .	58
3.8. Achievable throughputs approximation for RBD at high SNRs with $M_T < M_R$.	59
4.1. Block diagram of multi-user MIMO BC employing CBF.	63
4.2. An example for the five users case with zero MUI constraint.	68
4.3. RMS error for IR, NCG, and STJOCO for $N = n = 3$	86
4.4. RMS error for IR, NCG, and STJOCO for $N = n = 10$	87
4.5. Achievable sum rate comparison for $M_T = M_{R_i} = K = 2$	89
4.6. Achievable sum rate comparison for $M_T = M_{R_i} = K = 3$	90
4.7. Achievable sum rate comparison for $M_T = M_{R_i} = K = 4$	90
4.8. The tracked residual MUI after 50 iterations for FlexCoBF with BD precoding	96
4.9. Achievable sum rate for $M_T = K = 3$ and $M_{R_i} = 2$	98
4.10. CCDF of required number of iterations with stopping criterion I	99
4.11. CCDF of required number of iterations with stopping criterion I and II	99
4.12. Achievable sum rate for $M_T = K = 3$ and $M_{R_i} = 2$ at SNR = 5 dB vs. ρ	101
4.13. Achievable sum rate for $M_T = K = 3$ and $M_{R_i} = 2$ at SNR = 25 dB vs. ρ	101
4.14. Achievable sum rate for $M_T = 6$, $K = M_{R_i} = 3$ and $\rho = 0$	102
4.15. Achievable sum rate for $M_T = 6$ and $K = M_{R_i} = 3$ at SNR = 5 dB v.s. ρ	103
4.16. Achievable sum rate for $M_T = 6$ and $K = M_{R_i} = 3$ at SNR = 25 dB v.s. ρ	103
4.17. Clustered cellular scenario	104

4.18. Two adjacent clusters with limited cooperation	105
4.19. Block diagram of extended FlexCoBF for two adjacent clusters.	107
4.20. Achievable sum rate for two-cluster scenario with $M_T = 3$, $M_{R_i} = 2$, and $\rho = 0$. . .	113
4.21. CCDF of required number of iterations in two-cluster scenario	114
4.22. A clustered cellular network with 7 adjacent clusters	115
4.23. Utility function $U(\xi)$ for different ξ	116
4.24. Sum rates of the home cluster with different cluster size.	117
4.25. Sum rates of the home cluster with different cooperation strategies.	119
4.26. An example of the throughput of user 3 with different cooperation strategies. . .	119
4.27. Sum rates of the cluster edge users with different cooperation strategies.	120
5.1. Reciprocity for channel state information (CSI) acquisition.	126
5.2. channel state information (CSI) acquisition using feedback.	127
5.3. One chunk structure in time and frequency domain.	130
5.4. The geometrical representation of the simulation scenario.	133
5.5. CCDF of the sum rates with BD and RBD precoding based on long-term CSI. . .	134
5.6. CCDF of the individual user throughput with BD based on long-term CSI. . . .	135
5.7. CCDF of the individual user throughput with RBD based on long-term CSI. . .	135
5.8. CCDF of the sum rates with BD precoding based on long-term CSI at the BS. .	136
5.9. CCDF of the sum rates with RBD precoding based on long-term CSI at the BS. .	137
5.10. Limited feedback system model.	138
5.11. Achievable throughputs for ZF $M_T = K = 5$ with fixed feedback bits per user. .	142
5.12. Achievable throughputs for ZF $M_T = K = 5$ with increased feedback bits per user. .	143
5.13. Achievable throughputs for BD $M_T = 4$, $K = 2$, and $M_r = 2$	147
5.14. Achievable throughputs for BD $M_T = 6$, $K = 3$, and $M_r = 2$	148
5.15. Achievable throughputs for RBD $M_T = 4$, $K = 2$, and $M_r = 2$ with increased B . .	152
5.16. Achievable throughputs for RBD $M_T = 4$, $K = 2$, and $M_r = 2$ with fixed B	155
5.17. Achievable throughputs for RBD with DE-LBG codebooks.	156
5.18. Unfolding of the tensor $\mathcal{A} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$	158
5.19. Visualization of the multiplication of a tensor with matrices	158
5.20. Block diagram of the tensor based data model in equation (5.77)	170
5.21. Block diagram of the tensor based data model with oversampling.	174
5.22. The smoothing window with the smoothing parameter $\eta = W$	176
5.23. RMSE for blind estimation with $M_R = 4$ and $r_1 = L + 1 = 3$	176
5.24. RMSE for blind estimation with $M_R = 6$ and $r_1 = L + 1 = 3$	177
5.25. RMSE for blind estimation with $M_R = 4$, $L + 1 = 5$	178

5.26. The smoothing window with the smoothing parameter $1 \leq \eta < W$	178
5.27. RMSE for blind estimation of SIMO channels with varied η	179
5.28. BER for semi-blind estimation of SIMO channels	180
5.29. BER performance comparison for SBCE and training-based channel estimation.	180
5.30. Block diagram of the tensor based data model.	183
5.31. Construction of the slices of \mathbf{H}_T	184
5.32. Block diagram of the tensor based data model with oversampling.	186
5.33. RMSE for MIMO channels: $M_T = 2$, $L + 1 = 2$, $r_1 = 4$, SNR = 20 dB.	189
5.34. RMSE for MIMO channels: $M_R = 5$, $M_T = 2$, $L + 1 = 4$, SNR = 20 dB.	190
5.35. RMSE for blind estimation of MIMO channels with varied η	191
5.36. BER performance comparison for SBCE and training-based channel estimation.	191
C.1. Example of $\Gamma(x)$ and $\frac{1}{\Gamma(x)}$ for real valued number, where $x_0 = 1.46163$	207

List of Tables

4.1. Convergence investigation for $M_T = K = 3$, $M_{R_i} = 2$, and SNR = 10 dB	97
4.2. User grouping and the involved help clusters	118
5.1. OFDM Parameters	133
5.2. The number of feedback bits B for the case of BD with increased feedback bits	147
5.3. The number of feedback bits B for the case of RBD with increased feedback bits	151
5.4. Variations of the matrix SVD	160
5.5. Variations of the tensor HOSVD	161
5.6. Comparison of the computational complexities	172

1. Introduction

1.1. Background and Motivations

During the past several decades, mobile communications have experienced a promising evolution from analog systems of the first-generation (1G) to digital systems of the fourth-generation (4G) as shown in Figure 1.1. 1G and the second-generation (2G) cellular systems have been

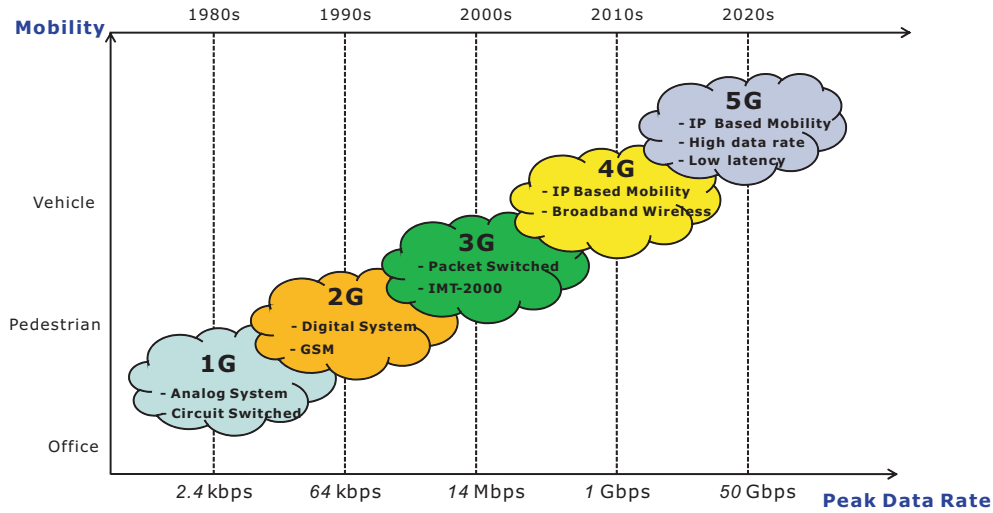


Figure 1.1.: Generations of mobile communications and their keywords.

mainly used for voice applications and supporting circuit-switched type services. The data rates of the users on the air interface are limited to less than several tens of kilobits per second (kbps). The International Mobile Telecommunications-2000 (IMT-2000) has been introduced in the beginning of the 21st century as third-generation (3G) cellular systems, which can provide 2 megabits per second (Mbps) and 384 kbps minimum data rates in indoor and vehicular environments, respectively. The applications of 3G systems have been found in wireless voice telephony, mobile internet access, fixed wireless internet access, video calls, and mobile TV. The first 4G system has been deployed since 2006 in South Korea. 4G systems have been specified to support all-internet protocol (IP) based communication which means all data and signaling will be transferred via IP on the network layer. The conceivable applications of 4G systems are found in IP telephony, mobile web access, gaming services, high-definition mobile TV, video conferencing and so on. The minimum data rate target ranges from 2 to 20 Mbps.

The fifth-generation (5G) systems do not exist yet, but this generation is expected to allow higher data rates than 4G systems. The minimum data rate target ranges from 20 to 100 Mbps. It is obvious that the growth of the number of subscribers and the involved communication terminals presents huge challenges for the future generations. For example, 4G systems are expected to fulfill the following requirements [OYN00, FZJJ06]:

1. Peak data rates for high mobility (e.g., mobile access) must approximate 100 Mbps and for low mobility (e.g., local wireless access) must approximate 1 Gbps.
2. Possibility to dynamically share and utilize the network resources to support more simultaneous users per cell
3. Wide coverage area and seamless roaming among different systems
4. High capacity, peak link spectral efficiency of up to 15 bit/s/Hz on the downlink and 6.75 bit/s/Hz on the uplink
5. The cost per bit has to be kept low.
6. Wireless quality of service (QoS) control (e.g., latency, jitter and packet loss) to offer high quality of service and support various applications

Regarding the demand for reliable high data rates, the data rate growth in previous wireless communication systems (e.g., a single-input single-output (SISO) system) has been primarily achieved by using more base stations (BS) and spectrum. However, this growth is constrained by the limited transmit power and the limited frequency bandwidth. In 4G systems, the data rate growth is achieved by strongly increasing the spectral efficiency per BS, which is supported by the coordinated multipoint (CoMP) concept or multiple-input multiple-output (MIMO) techniques [MRS⁺12].

CoMP approaches were firstly proposed in [BMWT00, SZ01, And05] where multiple BSs jointly transmit to multiple terminals and the intercell interference is effectively avoided to obtain large gains in spectral efficiency. These examples are cases of multi-cell joint signal processing. However, CoMP can also refer to the schemes with limited cooperation between BSs, for example coordinated beamforming where the intercell interferences at the edge users (i.e., users allocated around the cell border) are avoided rather than accepting them as noise via limited cooperation between BSs. The details can be found in Section 4.4. Compared to other options which can support the increasing data demand, such as using more spectrum, increasing the degree of sectorization, and using more base stations or introducing relays and

micro/femto cells, CoMP approaches require a fairly small change of infrastructure, and may lead to a more homogeneous quality of service (QoS) distribution over the area [MKF06]. For this reason, CoMP has been identified as a key technology of LTE-Advanced [PDF⁺08]. CoMP technology has opened a new path for multi-cell wireless technology development. Meanwhile, it faces several challenges regarding clustering, synchronization, channel knowledge, efficient and robust algorithm implementation, and Backhaul [MF11].

MIMO techniques have been proposed as a breakthrough in wireless communication system design. Pioneering works by van Etti [Ett76], Winters [Win87] and Telatar [Tel95] have predicted remarkable spectral efficiency and link reliability of the multiple-input multiple-output (MIMO) technology, which inspired an explosion of research activities associated with MIMO channels [KS01] and extended multi-user MIMO systems.

Compared to SISO technology, MIMO technology exploits the spatial domain and provides a number of advantages [BCC⁺07].

- * The sensitivity to fading is reduced due to the spatial diversity gain provided by multiple independent spatial paths, thereby improving the reliability of the received signals. For a MIMO channel with M_T transmit antennas and M_R receive antennas, $M_T \cdot M_R$ independently fading links can be potentially realized. Therefore, a spatial diversity order of $M_T \cdot M_R$ is offered.
- * The channel capacity is significantly increased due to the spatial multiplexing gain which is promised by transmitting multiple, independent data streams under a rich scattering environment. The number of data streams is decided by the minimum of the number of transmit antennas and the number of receive antennas in a MIMO channel, i.e., $\min(M_T, M_R)$. Therefore, capacity scales linearly with $\min(M_T, M_R)$ relative to a SISO channel.
- * The spatial dimension exploited in MIMO systems can help to mitigate the interference between different data streams or users. For instance, we can use transmit/receive beam-patterns to direct the signal power towards the desired user and minimize the interference to other users. Interference reduction improves the quality and coverage of a wireless communication system.

In general, we cannot achieve all the benefits described above simultaneously [BCC⁺07]. It is observed that maximizing one type of gain comes at the price of sacrificing the other. For example, the coding structure from orthogonal designs [TJC99] achieves the full diversity gain,

while reducing the achievable spatial multiplexing gain. In order to support maximum data rate or diversity order, the appropriate linear transmit and receive strategies are necessary.

Many optimal and sub-optimal beamforming algorithms have been proposed for MIMO broadcast channels. Different optimization criteria are addressed, such as maximum signal-to-noise ratio (SNR) [LJS03], minimum MSE [SSP01], and maximum information rate [JG01, JB04]. The channel state information (CSI) is assumed to be known at the transmitter and helps to increase the data rate, enhance coverage, and reduce the receiver complexity. In some applications such as Time Division Duplex (TDD) systems, the transmitter can obtain CSI directly due to the reciprocity principle. In Frequency Division Duplex (FDD) systems, CSI can be acquired at the transmitter by using feedback from the receiver. When the CSI is known perfectly at both transmitter and receiver, the MIMO channel (ergodic) capacity is the maximum mutual information averaged over all channel states.

Multi-user MIMO systems are considered as an enhancement of MIMO system when the number of users is greater than one. In contrast to single-user MIMO, multi-user MIMO allows a base station (BS) or an access point (AP) to communicate to multiple users in the same band simultaneously with the benefit of space division multiple access (SDMA). It is achieved at the cost of additional hardware like filters and antennas, but does not come at the expense of additional bandwidth. Using additional processing and the available CSI, the BS/AP can mitigate or completely eliminate the multi-user interference (MUI).

Multi-user MIMO techniques have been investigated intensely due to several key advantages over single-user MIMO [GKH⁺07, SSH04, SPSH04]. For example, because of an overall multiplexing gain given by the minimum number of base station antennas, MU-MIMO systems obtain a direct gain in multiple access capacity. Furthermore, this achievable multiplexing gain can be obtained without the need for multiple antenna terminals, which therefore facilitates the development of small and cheap terminals.

Multi-user MIMO can be classified into two categories: MIMO multiple access channels (MIMO MAC) and MIMO broadcast channels (MIMO BC). The MIMO MAC represents multi-user MIMO uplink channels where multiple terminals transmit to one receiver. The receiver performs much of the processing. For example, the joint interference cancellation and SDMA-based uplink user scheduling. The channel state information at the receiver (CSIR) has to be available at the receiver for the advanced processing, which requires a significant level of uplink capacity to transmit the dedicated pilots from each terminal. Conversely, the MIMO BC represents multi-user MIMO downlink channels where a single transmitter transmits to multiple terminals. The transmit processing is required at the transmitter side such as precoding and SDMA-based downlink user scheduling. The implementation of the

transmit processing requires that the transmitter has to know the channel state information at the transmitter (CSIT) in order to properly serve the spatially multiplexed terminals.

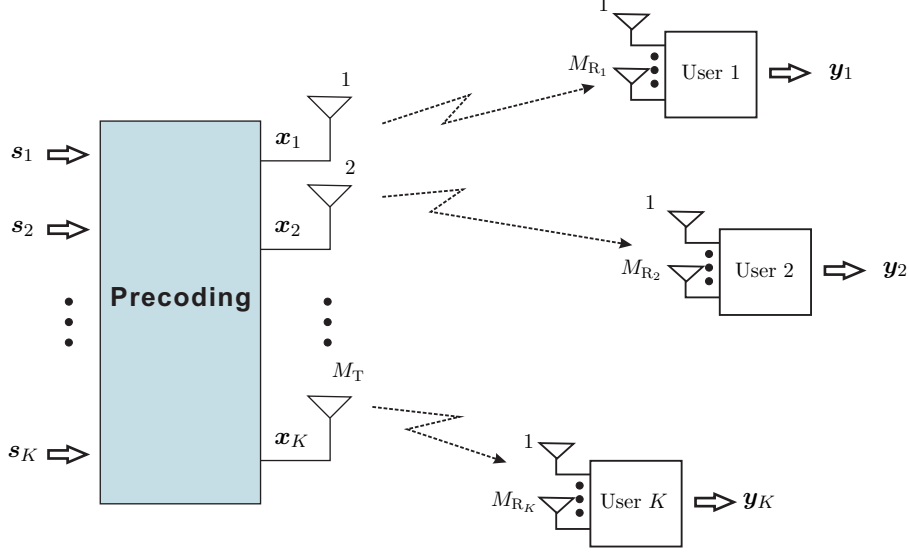


Figure 1.2.: Block diagram of multi-user MIMO BC.

We consider a MIMO system with K -users where the base station has M_T antennas and the i th user is equipped with M_{R_i} antennas. Figures 1.2 and 1.3 show the multi-user MIMO downlink channels and multi-user MIMO uplink channels, respectively. The capacity becomes a K -dimensional rate region, where each point is a vector of achievable rates by all the K users simultaneously. From [CV93, YRBC01, YRC01], the capacity region of a general MAC has been known for both constant channels and fading channels. Although the characterization of the general broadcast capacity region is a long standing problem, substantial progress has been made for the Gaussian MIMO BC channel by the work of Caire and Shamai [CS00] and the subsequent research by Yu and Cioffi [YC01]. Considering full CSIT, the sum capacity of the Gaussian MIMO BC by using the idea of dirty paper coding (DPC) [Cos83] has been found and demonstrated to equal the achievable region of the maximum sum rate of DPC.

In [VJG02], a promising result by Vishwanath shows that the rate region of MIMO BC obtained by using DPC with the power constraint P is equal to the capacity region of the dual MIMO MAC where the sum of all individual power constraints is set to P . This MAC-BC duality is very useful, since it substitutes the non-concave functions of the covariances in the DPC rate region with the concave functions of the covariance matrices in the dual MAC rate region.

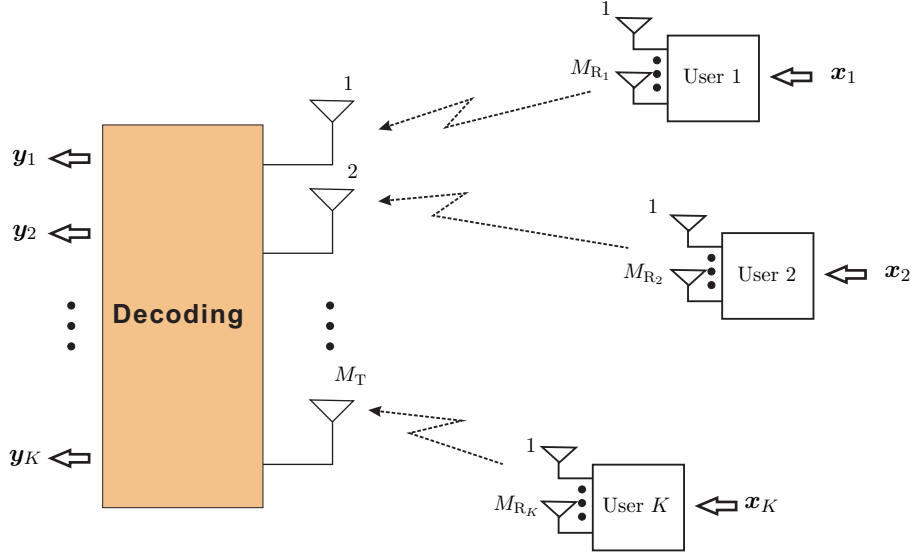


Figure 1.3.: Block diagram of multi-user MIMO MAC.

DPC is a technique to pre-subtract the interference at the transmitter to achieve the maximum sum rate of the system. The existing DPC schemes are hard to implement in practice due to the very high complexity at both transmitter and receiver, especially when the number of users becomes large. For example, general nested lattices [ZSE02, EZ04] are the most efficient encoding strategies. However, the construction of such lattices for a reasonably high dimension is not systematic because of the very high coding and decoding complexity. In addition, performing DPC with imperfect CSI is still challenging. As a simplified version of DPC, Tomlinson-Harashima precoding (THP) has been proposed. THP is a non-linear precoding technique originally developed for SISO multipath channels, which can be interpreted as moving the feedback part of the decision-feedback equalization (DFE) to the transmitter. It is also applied for the pre-equalization of MUI in MIMO systems [CC00], where it performs spatial pre-equalization instead of temporal pre-equalization for inter-symbol interference (ISI) channels. Some subsequent works have appeared [FWLH02, JBU04, SH05c], which have proposed different precoding schemes based on THP with improved power efficiency, increased achievable throughput, or reduced complexity at the receiver. However, the complexity at the base station is still very high. The main drawbacks of these THP-based non-linear techniques have been pointed out in [SHGJ06]. They are a higher computational complexity, the required signaling overhead and their sensitivity to channel estimation errors.

Linear precoding techniques have been proposed for the multi-user MIMO BC by considering the trade-off between performance and complexity. In general, linear precoding tech-

niques achieve a reasonable performance with a much lower complexity, and they also have remarkable flexibility for practical implementation due to the adaptation to various degrees of CSI, such as instantaneous CSI, long-term statistic of CSI (definition in Section 5.2), or limited CSI (definition in Section 5.3). Linear precoding techniques can be classified by the amount of the MUI they allow as zero or non-zero MUI techniques. For example, zero-forcing (ZF) precoding [VJ98] aims at nulling the MUI with the expense of losing some signal gain. Minimum mean-square-error (MMSE) precoding [VJ98] has been designed to maximize the signal-to-interference-plus-noise ratio (SINR) at each user and allows an amount of MUI. ZF precoding is proposed to support single data stream transmission to each user where each user is equipped with only one receive antenna. The subsequent work by Spencer et al. [SSH04] extends it to the case of multiple data streams per user where each user can have multiple receive antennas. This algorithm is known as block diagonalization (BD) precoding. MMSE precoding improves the system performance by introducing a certain amount of interference especially for users equipped with a single antenna. However, it suffers a performance loss when it attempts to mitigate the interference between two closely spaced antennas when the user terminal is equipped with more than one receive antenna. The work by Stankovic et al. [SH04b] has proposed a new algorithm that deals with this problem by successively calculating the columns of the precoding matrix for each of the receive antennas separately. This new algorithm is known as successive MMSE (SMMSE) precoding. Although linear precoding techniques cannot achieve the sum capacity of the MIMO BC in general, with some special conditions they can achieve the same performance as DPC. The work of Yoo and Goldsmith in [YG06] has shown that zero-forcing (ZF) precoding can achieve the same asymptotic sum capacity as that of DPC, when the number of users goes to infinity.

Unfortunately, the application of these linear precoding strategies is constrained by the dimensionality restriction which states that the total number of receive antennas must be smaller than or equal to the number of transmit antennas. This condition is not fulfilled in many scenarios, especially when a large number of users is present. Then, additionally an efficient user scheduling algorithm [FGH07] or receiver antenna selection [WZZ⁺05] is needed. Regularized block diagonalization (RBD) linear precoding [SH08] has been proposed to release the dimensionality restriction, while achieving an improved sum rate as well as diversity order compared to BD and ZF. However, the performance of RBD degrades heavily with an increasing aggregate number of receive antennas [SH09a]. Coordinated beamforming (CBF) [SSH04, ZHV08, SH08, CMIH08a, CMJH08, SH09b, SRH10b] has been proposed as a powerful solution to overloaded scenarios (i.e., the total number of receive antennas is larger than the number of transmit antennas), which jointly optimizes the beamforming vectors at the trans-

mitter and receiver and allows to transmit a number of data streams that is equal to the number of transmit antennas. The sum rate performance has been demonstrated to reach the sum capacity of the MIMO BC. Most of the existing CBF algorithms employ iterative operations to jointly update the transmit-receive beamforming vectors, but without the consideration of the convergence behavior. In [SSH04, ZHV08, SH08], the number of iterations is set directly by hand. To the best of our knowledge, closed-form CBF has only be addressed in [CMJH08] and our proposal [YSRH12]. The closed-form CBF in [CMJH08] is restricted to a system with two transmit antennas and two users. In contrast, our method in [YSRH12], named "sequentially drilled "joint congruence (SeDJoCo) transformation based (SeDJoCo-based) CBF, is valid for an arbitrary number of users and transmit antennas. However, in both of them a single data stream transmission per user is demonstrated only.

1.2. Scope of the thesis and contributions

The prime focus of the thesis is on the design of a linear transmit-receive strategy. We provide new schemes that can give proper solutions to some open problems (shown in introduction and motivation of each chapter) in the existing linear transmission techniques or help to enhance the existing algorithms. Therefore, we consider multi-user MIMO down-link channels in a single cell and multiple cells. The whole thesis consists of 5 chapters. The following part provides a brief motivation for the different chapters, the open problems of the existing solutions, and summarizes the major contributions and the possible applications. A more detailed introduction for each chapter is found at the beginning of each chapter. Our publications and patent that are relevant to the thesis can be found in [SH09a, SH09b, SRH10b, YSRH12, LCZ⁺12, SRH10c, SHMK09, KTS⁺10, SRH08, SH09c, SRH10a, RSS⁺11, SRH13a, SRH13b, CLZ⁺13, ZSHdL14a, ZSHdL14b, DCL⁺13]

Chapter 2: this chapter presents an overview of the fundamentals of MIMO technology, which helps to understand the motivations, developments, and applications of the major contributions in this thesis. It starts with major achievable benefits of MIMO techniques. Although these benefits cannot be exploited simultaneously, some combinations of the benefits across a wireless system can lead to an improved capacity, enhanced reliability, and enlarged coverage. Then, the existing MIMO channel models have been introduced from a physical perspective and an analytical perspective. Several widely used MIMO channel models are reviewed in details such as Rayleigh fading channels, spatially correlated Rayleigh fading channels, Ricean fading channels, etc. Furthermore, the capacity results for single-user MIMO, multi-user MIMO, and multi-cell MIMO have

been reviewed. These results indicate that the capacity gain obtained from multiple antennas heavily depends on the available channel information at either transmitter or receiver [FG98, JG05a], the channel signal-to-noise ratio (SNR), and the spatial channel correlations [JG05a]. Although there are still many open problems in terms of capacity of systems with MIMO (shown in Sections 2.4.1.6 and 2.4.2.4), the existing results do provide an insight into the capacity limits of systems and the implications of the limits for practical system design. The proofs and notes of Chapter 2 are found in Appendix B.

Chapter 3: this chapter is focused on the existing well-known linear precoding schemes (e.g., ZF, BD, and RBD). An overview of them is provided firstly. Since the linear precoding techniques have a low complexity at the transmitter and the receiver by sacrificing an amount of throughput, a throughput loss between the linear precoding and DPC does exist. It is interesting to quantify the throughput loss with respect to the system parameters such as the number of receive antennas, the number of transmit antennas and the number of users. The research work associated with this loss has been done in [JG05b, SCA⁺06, LJ07]. The authors of [JG05b, SCA⁺06] have analyzed the ratio between the achievable sum rates of DPC and BD. A more practical metric has been proposed in [LJ07], where the absolute rate and power offsets between DPC, BD, and ZF are studied at high SNRs. However, none of them has mentioned RBD.

RBD was firstly proposed by [SH08] in 2008. It is designed to relax the dimensionality constraint which restricts the applications of ZF and BD. In other words, RBD can still be utilized in the situation when the multi-user MIMO system has less transmit antennas (i.e., M_T) than the aggregate number of receive antennas (i.e., M_R , here, $M_T < M_R$), although the performance is degraded compared to the application of RBD in the situation that the number of transmit antennas is larger than or equal to the total number of the receive antenna in a multi-user MIMO system (i.e., $M_T \geq M_R$). Furthermore, RBD has an improved throughput and diversity order compared to BD.

Inspired by the previous work in [LJ07], we analyze the performance of the throughput loss between RBD, DPC, and BD at high SNRs for two cases. In the first case (i.e., $M_T \geq M_R$), the bounds of the average rate and power offsets between these schemes have been derived as a function of the system parameters. The achievable multiplexing gain of RBD is the same as DPC at high SNRs. In the case $M_T < M_R$, the approximated throughput of RBD shows that the achievable multiplexing gain drops to 1 at high SNRs. The application of RBD is only recommended in the low or medium SNR regime. This

contribution has been published in [SH09a]. The proofs and notes of Chapter 3 are found in Appendix C.

Chapter 4: this chapter is devoted to coordinated beamforming techniques. We present new CBF algorithms which can provide a better sum rate relative to the existing CBF algorithms with comparable complexity. There are two basic categories for the existing CBF algorithms (i.e., iterative CBF and closed-form CBF). The closed-form solution of CBF is more difficult to obtain than the iterative one. Thus, most of the existing CBF algorithms are iterative. To the best of our knowledge, only the authors in [CMJH08] have provided a closed-form CBF but only for a two-user two-transmit-antenna system. It indicates in [CMJH08] that the transmit beamforming vectors \mathbf{f}_1 and \mathbf{f}_2 satisfy zero inter-user interference conditions if they are chosen as the generalized eigenvector pair of the correlation matrices of the two users (i.e., \mathbf{R}_1 and \mathbf{R}_2). In other words, the matrix $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2]$ can simultaneously diagonalize the correlation matrices \mathbf{R}_1 and \mathbf{R}_2 such that $\mathbf{F}^H \mathbf{R}_1 \mathbf{F}$ and $\mathbf{F}^H \mathbf{R}_2 \mathbf{F}$ are diagonal matrices.

Let us consider the case of K users. There is a matrix $\mathbf{F} \in \mathbb{C}^{M_T \times K}$ in which the i th column corresponds to the transmit beamforming vector of user i and a set of matrices $\{\mathbf{R}_i\}_{i=1}^K \in \mathbb{C}^{M_T \times M_T}$ which are the correlation matrices of all users. If the matrix \mathbf{F} can jointly minimize the magnitude of the off-diagonal elements on the i th row and the i th column of the set of correlation matrices $\{\mathbf{R}_i\}_{i=1}^K$, then, the columns of the matrix \mathbf{F} are chosen as the transmit beamforming vectors of K users which minimize the multi-user interference. To find such a matrix \mathbf{F} , a particular transformation algorithm is needed. Fortunately, the problems of blind source separation (BSS) and independent component analysis (ICA) [Yer10] give a hint to find \mathbf{F} . Since the generic algorithmic tool to solve BSS and ICA problems is an approximated joint diagonalization (AJD), the general framework of AJD considers a set of N square, symmetric, real-valued $n \times n$ matrices $\{\mathbf{C}_i\}_{i=1}^N$. The goal of AJD is to find a single matrix \mathbf{B} which best jointly diagonalizes the target matrices. Inspired by AJD, we have derived a "sequentially drilled" joint congruence (SeDJoCo) transformation which can be applied to successfully find the matrix \mathbf{F} [SRH10c, YSRH12, SH09b]. The solutions of SeDJoCo have been proven to exist, but may not be unique. Note that the SeDJoCo transformation can be used to solve not only the problem of the closed-form CBF but also the problems of BSS and ICA.

Except for the closed-form CBF, another important contribution is devoted to the iterative CBF. A new iterative CBF [SRH10b, SRH13a] (namely flexible coordinated beam-

forming (FlexCoBF)) has been designed with a high flexibility on the transmit-receive beamforming vector design, a sum rate close to the sum capacity of the MIMO broadcast channel, and efficient stopping criteria to achieve a good convergence. The aims of FlexCoBF are to efficiently utilize well-known transmit strategies (e.g., ZF, BD, and RBD) and receive strategies (e.g., MMSE and MRC), combat some drawbacks in the existing iterative CBF (e.g., only one data stream per user [CMIH08a] and the manually defined number of iterations [SSH04, ZHV08, SH08]), and give an insight into the convergence behavior. We have successfully obtained a patent [DCL⁺13] which is based on FlexCoBF.

Furthermore, FlexCoBF has been extended to the cellular scenario [SRH13a]. Unlike some existing beamforming methods which assume the K -user MIMO interference channel and only consider the cell edge users [CHHT13, CKH09, GMK10, TGR09], the extended FlexCoBF considers both cell interior users and cell edge users. By introducing the coordinated multipoint (CoMP) concept which is one of the hot topics for 4G and beyond system design [PDF⁺08], we consider a limited cooperation between base stations to support the extended FlexCoBF. As a result, both intra-cell and inter-cell interferences can be efficiently suppressed.

A short summary is given at the end of this chapter. It summarizes the advantages of the new designed CBF algorithms compared to the existing CBF and also the shortcomings which should be considered in future work. The contributions associated with this chapter are published in [SH09b, SRH10b, SRH10c, LCZ⁺12, YSRH12, SRH13a, CLZ⁺13]. The proofs and notes of Chapter 4 are found in Appendix D.

Chapter 5: since channel state information (CSI) acquisition at the base station is an important issue for the linear transmit-receive strategy design in multi-user MIMO broadcast channels, we discuss several CSI acquisition strategies for three different channel conditions.

Firstly, we consider a time-varying correlated channel model. In this case, the channel varies too fast to capture the instantaneous CSI (namely short-term CSI), the long-term CSI based on second-order channel statistics is considered alternatively. The authors in [SH05b, RFH08] have introduced a method to exploit the long-term CSI for multi-user linear precoding. However, this method is only efficient for the case that the channels are weakly correlated or entirely uncorrelated. Therefore, we have proposed a rank-one approximated long-term CSI (ROLT-CSI) approach [SHMK09] which captures the channel information by exploiting the knowledge of the estimated spatial correlation per receive antenna and transmitting then along the dominant eigenmode of the exploited

spatial correlations. Compared to the previous long-term method, ROLT-CSI is more efficient if the channels are highly spatially correlated (e.g., LOS channels). But even for uncorrelated channels, the presented ROLT-CSI still achieves some performance gain relative to the previous long-term CSI method.

Then, we consider a system employing a limited feedback channel to inform the base station about the CSI. In order to reduce the feedback overhead, the channel matrix is quantized first at the user side before it is fed back. There are several ways to quantize the channel matrix. A straightforward way is to encode every complex number of the channel matrix individually, but it is inefficient, since the number of feedback bits is still high. Alternatively, we quantize every individual channel matrix by looking up a predefined codebook. Then, only the index of the chosen codeword needs to be fed back. Most channel quantization schemes for the multi-user MIMO downlink employing precoding only consider the directions of the channel [Jin06, RJ07], or they quantize the channel directions and magnitudes separately [KZH08]. We propose a quantization scheme [SRH08] which stacks the vectors of the channel matrix to maintain the relative magnitude information for the columns of the channels and further quantizes the stacked vector. It is well suited for the multi-user MIMO downlink employing RBD precoding or employing other linear precoding algorithms (e.g., ZF and BD) with power allocation. Moreover, we have developed an analytic performance analysis for the multi-user MIMO downlink employing our quantization scheme and RBD precoding. The quality of the codebook significantly affects the quality of the quantized CSI. Except for the study of the random vector quantization (RVQ) codebook, we also investigate a new codebook design (i.e., the dominant eigenvector Linde-Buzo-Gray (DE-LBG) vector quantization codebook) [SRH08], which is more efficient than the RVQ codebook by introducing optimality criteria and a sequence of channel samples (details are found in Section 5.3.3.4). As a result, the complexity of the DE-LBG codebook is higher than the RVQ codebook.

Finally, we consider blind channel estimation in order to reduce the channel estimation overhead. Blind channel estimation has already been discussed for several decades and several blind methods of channel estimation have been proposed for various channel contexts. The existing methods are classified into the moment-based and the maximum likelihood (ML) methods. The moment-based methods can be further divided into the higher order statistical and the second order statistical approaches. Many applications of the higher order statistical approaches [HN89, PF91, Tug95] do not consider multichannel models. In such a case it may be necessary to exploit higher order statistics. Since the publication of [TXK91], second order moment techniques have received considerable

attention. These methods exploit the subspace structures of the observation [TXK91, MDCM95, Nef04, ZMG02, VS08, MdCDB99, YP03]. Thus, they are often referred to as subspace-based algorithms. One attractive property of the subspace-based blind channel estimation methods is that the channel estimates can often be obtained in a closed form from optimizing a quadratic cost function (as shown in equation (5.76)). On the other hand, subspace methods may not be robust against modeling errors, especially when the channel matrix is close to being singular, as they rely on the property that the channel lies in a unique subspace. Unlike subspace-based methods, the ML methods usually cannot be obtained in closed form. Their implementations are further complicated [BM86, Hua96, TVP94]. However, ML methods can be very effective by including the subspace methods as initialization procedures [TP98].

In the presented subspace-based approaches to blind channel estimation, the measurement data is generally stored in one highly structured vector by a stacking operation. The structure inherent in the measurement data is thus not considered in the subspace estimation step. Exploiting the inherent structure of the measurement data often results in a benefit. Therefore, we introduce the concept of *tensors*. Tensor-based signal processing techniques have become increasingly popular in many different areas of signal processing due to several fundamental advantages over their matrix-based counterparts [KB09, dLdMV00a].

* **Identifiability:**

There is a major difference between matrices and higher-order tensors when rank properties are considered. For a matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$, the rank of the matrix r is constrained to $r \leq \min(M, N)$. In contrast, the tensor rank can largely exceed its dimensions which leads to the advantage that more sources than sensors can be identified [KB09].

* **Uniqueness:**

The tensor decompositions are essentially unique up to permutation and scaling, whereas matrix decompositions are not. For example, the singular value decomposition (SVD) of a matrix is unique only because of the additional orthogonality constraints. This property of tensors allows to separate more components compared to the matrix approach and is very useful for applications such as blind source separation (BSS) [WJG⁺10].

* **Multilinear rank reduction:**

Since the tensor notation preserves the structure of the measurement data, more

efficient denoising is achieved by low-rank approximations of tensors. This advantage can be used for many applications such as chemometrics [AB03], psychometrics [CC70], computer vision [VT03], watermarking [AHB07], data mining [STF06], array processing [SBG00], and ICA [LC08].

* **Improved subspace estimate:**

By using the multilinear rank reduction, improved signal subspace estimates can be achieved. Therefore, multidimensional subspace-based parameter estimation schemes can lead to a higher accuracy [HRD08]. Some example applications include R -D harmonic retrieval [Boy08, HRD08], channel modeling [MHS04], surveillance RADAR [JLL09], blind interference cancellation [dLHSN08, dL11], and high-resolution parameter estimation [dLSNH07, HN95, HTR04].

Considering these advantages, we have designed tensor-based blind channel estimation algorithms of SIMO and MIMO channels to achieve an enhanced accuracy [SRH10a, SRH13b]. The contributions presented in this chapter are published in [SHMK09, KTS⁺10, SRH08, SH09c, SRH10a, SRH13b, RSS⁺11]. The proofs and notes of Chapter 5 are found in Appendix E.

Chapter 6: in this final chapter we collect all the contributions from the thesis again and summarize the future research directions related to the thesis.

There are five appendices to the thesis. Appendix A summarizes the list of acronyms and the mathematical notation used throughout the thesis. Appendices B, C, D, and E contain proofs, derivations, and notes of Chapter 2, 3, 4, and 5, respectively. The bibliography is split into two parts: one part with our own publications and a second part with all other references.

2. Fundamentals of MIMO wireless communication

2.1. Introduction and Motivation

Fading and interference are the two key challenges faced by wireless communication systems. Fading influences the reliability of any point-to-point wireless communication (e.g., between a base station and a user) and interference limits the achievable throughput. MIMO technology, known as the use of multiple antennas at the transmitter and the receiver in wireless systems, emerges as a breakthrough for the future wireless communication. It offers a number of benefits which help to meet the existing challenges such as the potentials for high spectral efficiency, increased diversity, and interference reduction.

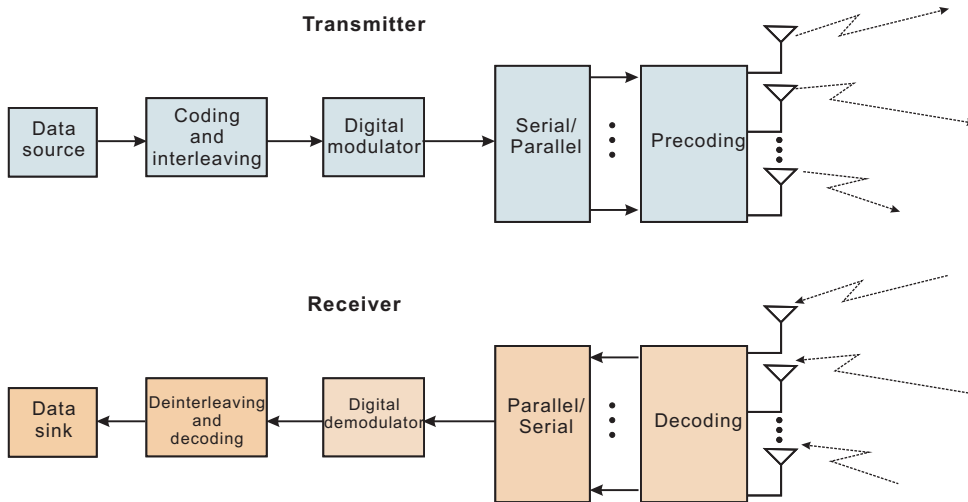


Figure 2.1.: The block diagram of a equivalent baseband MIMO communication system.

Figure 2.1 shows an example of a MIMO communication system. At the transmitter, the information bits are generated, encoded, and interleaved. The interleaved codeword is mapped to data symbols by the digital modulation. These data symbols are thus split into more spatial data streams. The spatial data streams are mapped to the transmit antennas by the precoding block. The launched signals propagate through the channel and are finally collected by the receiver. The receiver reverses the transmit operations to decode the data. Notice that the placement, the functionality, and the interactions of the blocks can vary due to different system designs.

MIMO technology has been integrated into Third Generation (3G) cellular systems, 4G standards (e.g., LTE-Advanced), and broadband fixed/mobile wireless access networks (IEEE 802.16e, also known as WiMax). It will also be included into the future IEEE 802.11n releases. In this chapter, we review several fundamental concepts in MIMO wireless communications. This knowledge will help to understand the motivations, developments, and applications of the major contributions in this thesis. This chapter starts with the overview of the achievable benefits of MIMO technology in Section 2.2. In Section 2.3 the construction of the wireless MIMO channels is described. The capacity regions of the single-user MIMO, the multi-user MIMO, and the multi-cell MIMO systems are discussed in Section 2.4. Finally, a summary is provided in Section 2.5.

2.2. Benefits of MIMO technology

MIMO technology provides a number of advantages over the conventional single-input single-output (SISO) technology, which have been traditionally defined as spatial diversity gain, spatial multiplexing gain, array gain, and interference reduction and avoidance [PNG03].

2.2.1. Spatial diversity gain

Diversity techniques have been proposed to combat the impact of fading. The basic idea of diversity is to provide the receiver multiple copies of the same transmitted signal in time, frequency, or space. With these independent copies, the probability that all links experience a deep fade reduces dramatically. Therefore, the quality and reliability of the reception has been improved. In general, time and frequency diversity techniques lead to a loss in time and bandwidth due to the introduction of redundancy. In contrast, spatial diversity is provided by employing multiple antennas at the transmitter and/or the receiver, thereby no sacrifice of time and bandwidth occurs.

The diversity gain is defined as the negative slope of the log-log plot of the average error probability \bar{P} versus SNR ρ [OC10]

$$g_d \triangleq -\frac{\log(\bar{P})}{\log(\rho)}. \quad (2.1)$$

Note that the spatial diversity gain is usually taken as the asymptotic slope (i.e., $\rho \rightarrow \infty$). For example, in a slow Rayleigh-fading environment with one transmit antenna and n receive antennas, the transmitted signal is passed through n different paths. If these n paths experience independent fading, a maximal diversity gain of n can be achieved. In general, for a MIMO

channel with M_T transmit antennas and M_R receive antennas, a spatial diversity order of $M_T M_R$ is potentially offered.

Compared to a single antenna system, a spatial diversity scheme inherently requires additional hardware (i.e., antennas) and additional signal processing at the transmitter and/or the receiver. However, regarding the fact that signal reliability is a key requirement of mobile wireless communication systems, employing multiple antennas is an effective method to decrease the number of failed connections and improve the reception quality.

2.2.2. Spatial multiplexing gain

In MIMO communication systems, multiple independent data streams can be simultaneously transmitted from the transmitter employing multiple antennas. By introducing proper SDMA strategies (e.g., linear precoding) at the transmitter, the receiver can reliably separate these data streams under favorable channel conditions such as a rich scattering environment. Thereby, the capacity of the MIMO system is enhanced by a multiplicative factor equal to the number of data streams. In general, the maximum number of data streams is equal to the minimum of the number of the transmit antennas and the number of the receive antennas. If a MIMO system is equipped with M_T transmit antennas and M_R receive antennas, the achievable spatial multiplexing gain is $\min(M_T, M_R)$.

Let us denote the transmission rate as $R(\rho)$. Then, the spatial multiplexing gain is defined as [OC10]

$$g_s \triangleq \lim_{\rho \rightarrow \infty} \frac{R(\rho)}{\log_2(\rho)}. \quad (2.2)$$

In practice the spatial multiplexing gain can be limited by spatial correlation.

2.2.3. Array gain

Array gain reflects the increase in the receive SNR ρ_{out} by using multiple antenna at the transmitter and/or receiver with respect to the receive SNR ρ_{SISO} of the SISO case, which is defined as [OC10]

$$g_a \triangleq \frac{\rho_{out}}{\rho_{SISO}}, \quad (2.3)$$

Array gain can be realized through spatial processing at the transmit antenna array and/or at the receiver antenna array (e.g., maximal ratio combining (MRC)). With the improved receive SNR, the coverage and range of a wireless system can therefore be enlarged.

2.2.4. Interference reduction and avoidance

In wireless communication system, interference is caused by sharing time and frequency resources between multiple users. To utilize MIMO technology, interference can be avoided by exploiting the additional spatial dimension for user separation. For instance, beamforming techniques direct signal energy towards the intended user and minimize interference to other users. Therefore, the achievable system throughput is improved and the coverage of a wireless system is enlarged as well. Furthermore, the interference power can be effectively reduced by the array gain, which results in an increased signal to noise plus interference ratio (SINR).

Unfortunately, it is impossible to simultaneously exploit all the benefits described above for a MIMO system due to the conflicting requirements on the spatial degrees of freedom. For example, spatial diversity is used to combat fading. However, multiple independent fading events increase the degrees of freedom available for communications. If the individual transmit-receive antenna pairs fade independently, the channel matrix can be well conditioned (i.e., very low or no spatial correlations) with high probability. In this case, independent data streams can be supported. Thus, spatial multiplexing can be used. In other word, the spatial multiplexing gain comes at the price of sacrificing diversity. A fundamental trade-off has been derived between the spatial multiplexing gain and the spatial diversity gain that any multiple antenna scheme can achieve in the Rayleigh fading channel [ZT03], which is useful for evaluating and comparing the existing multiple antenna schemes. Despite all that, using some combinations of the benefits in a wireless system can lead to an improved capacity, an enhanced reliability, and an enlarged coverage.

2.3. MIMO channel models

Note that radio propagation has a significant impact on the performance of wireless communication systems. Its impact on future broadband systems is even more important due to demand for the increased data rates, bandwidth, mobility, adaptivity, etc. Therefore, versatile and accurate MIMO channel models are required. Furthermore, a profound understanding of MIMO channels is crucial for an efficient design of algorithms and a selection of proper signaling strategies in wireless MIMO systems.

The wireless channel is a challenging environment due to the existence of multipaths with different time-varying delays, attenuations caused by fading, different phases, and directions of departure and arrival. Furthermore, high mobility results in rapid variations across the time-dimension and angular spread causes significant variations in the spatial channel responses. Therefore, a versatile MIMO channel model must accurately track all dimensions of the channel

responses (e.g., time, frequency, and space) and capture all the basic channel characteristics which include

- * Temporal characteristics (e.g., delay spread and distribution, power delay profile)
- * Frequency-domain characteristics (e.g., Doppler shifts, Doppler power spectrum)
- * Spatial characteristics (e.g., angle of departure spread and distribution, angle of arrival spread and distribution, spatial correlations, and power azimuth spectrum)

Considering a MIMO system with M_T transmit antennas and M_R receive antennas, the MIMO channel is given by the $M_R \times M_T$ matrix \mathbf{H} with

$$\mathbf{H}(\tau, t) = \begin{bmatrix} h_{1,1}(\tau, t) & h_{1,2}(\tau, t) & \cdots & h_{1,M_T}(\tau, t) \\ h_{2,1}(\tau, t) & h_{2,2}(\tau, t) & \cdots & h_{2,M_T}(\tau, t) \\ \vdots & \vdots & \ddots & \vdots \\ h_{M_R,1}(\tau, t) & h_{M_R,2}(\tau, t) & \cdots & h_{M_R,M_T}(\tau, t) \end{bmatrix}, \quad (2.4)$$

where $h_{m,n}(\tau, t)$ is the channel impulse response between n th transmit antenna for $n = 1, \dots, M_T$ and m th receive antenna for $m = 1, \dots, M_R$, which depends on several parameters such as time, delay, and directions of departure and arrival.

In general, the MIMO channel models can be classified into two different categories: physical channel models and analytical channel models. The difference and properties of these two models are briefly described as follows [OC10].

2.3.1. Physical channel models

Physical channel models are mostly used to evaluate the system performance. They specify the geometrical parameters (e.g., the locations of scatterers and obstacles) and the array configuration. The MIMO channels are thus generated to provide values of the basic characteristics of the channels (e.g., spatial correlations, multi-path fading, Doppler spread, etc.) as a function of the defined geometrical parameters and array configuration.

A large number of physical models exists [OC10]. For example, geometry-based models, using a simplified ray-based approach, are stochastic models of the channel. One-ring models, two-ring models, combined elliptical-ring models, and elliptical and circular models are popular geometry-based models as example. In all cases, the channel between any transmit and receive antennas is generally obtained as [Gol05]

$$h_{m,n}(\tau, t) = \sum_{\ell=1}^L c_{\ell} e^{j2\pi f_{D,\ell} t} \delta(\tau - \tau_{\ell}) \mathbf{a}_R(\Theta_{R,\ell}) \mathbf{a}_T(\Theta_{T,\ell}), \quad (2.5)$$

where

- L is the total number of paths.
- c_ℓ is the complex scalar amplitude gains of the ℓ th path.
- $\Theta_{R,\ell}$ is the direction of arrival of the ℓ th path.
- $\Theta_{T,\ell}$ is the direction of departure of the ℓ th path.
- $\mathbf{a}_R(\Theta_{R,\ell})$ is the array steering vector of the received array as the function of the direction of arrival (DOA).
- $\mathbf{a}_T(\Theta_{T,\ell})$ is the array steering vector of the transmitted array as the function of the direction of departure (DOD).
- τ_ℓ indicates the time delay of the ℓ th path.
- $f_{D,\ell}$ denotes the Doppler frequency of the ℓ th path.

There are several popular physical channel models for comparing the performance of different system implementations such as

- * **3GPP/3GPP2 spatial channel models (SCM)**: this set of models were released in September 2003 [rGPPG03] and originally established for the 2 GHz range and 5 MHz bandwidth 3G networks in urban and suburban macrocells as well as in urban microcells. The modeling structure is similar to the COST 259 directional channel model [Cor01], but the major difference is that SCM is defined as a discretized model instead of the continuous one in COST 259. For example, continuous, large-scale movements of the mobile terminal is not permitted for SCM. Instead, SCM considers different possible positions of the mobile terminal within a cell.
- * **WINNER channel model**: this set of models was developed by the Wireless World Initiative New Radio (WINNER) project. The WINNER channel model-phase I (WIM1) was firstly described in the deliverable D 5.4 in 2005 [WIN05]. WIM2 referring to phase II was available in the deliverable D 1.1.2 in 2007 [WIN07]. WIM1 and WIM2 were proposed for the 5 GHz range and 100 MHz bandwidth in different environments such as indoor, urban microcells, indoor hotspot, stationary feeder sub-scenarios, urban and rural macrocells, indoor-to-outdoor, outdoor-to-indoor, moving networks, and bad urban. The comparisons of SCM and WIM have been evaluated in [NST⁺07]. The WINNER channel model has been considered in our publication [KTS⁺10].

- * **IlmProp channel model:** it is capable of modeling time variant frequency selective multi-user MIMO channels [DHS03]. The channel impulse response (CIR) is generated as a sum of propagation rays. These rays are defined by a three dimensional geometry which can be defined manually or be retrieved from measurements. We have utilized the IlmProp in our publication [SHMK09].

2.3.2. Analytical channel models

On the other hand, analytical models are mostly used for the explicit algorithm design. They provide a mathematical, more abstract representation of the MIMO channel matrix. To this end, the MIMO channel matrix is expressed as a function of a random fading matrix and various channel correlations.

2.3.2.1. Rayleigh fading channels

The Rayleigh fading assumption is often used by MIMO system designers due to the fact that it is realistic when the environment has rich scatterers. The narrowband transmission between a transmit-receive antenna pair is modeled as a sum of a large number of contributions which have random and statistically independent phase, directions of departure (DODs), and directions of arrival (DOAs). If the wide-sense stationary uncorrelated scattering homogeneous (WSSUSH) Rayleigh fading channels are assumed, each element of the channel matrix \mathbf{H} is thus a zero-mean complex circularly symmetric Gaussian variable. We have

$$\begin{aligned} \mathbb{E}\{h_{m,n}\} &= 0, \\ \mathbb{E}\{|h_{m,n}|^2\} &= 1, \\ \mathbb{E}\{h_{m,n}h_{i,j}^*\} &= 0, \text{ for } i \neq m, j \neq n. \end{aligned} \tag{2.6}$$

In this case, the channel matrix \mathbf{H} is denoted as \mathbf{H}_w [PNG03]. This so-called independent identically distributed (i.i.d.) Rayleigh assumption has been extensively used for the designs of space-time coding, linear precoding, coordinated beamforming, distributed beamforming, etc. However, it is noticed that the real-world channel sometimes significantly deviates from this ideal channel model \mathbf{H}_w due to a number of reasons:

- * Insufficient spacing between antenna elements and/or the placement of scatterers leads to spatial fading correlation. Channels are not independent anymore.
- * Gain imbalance between the various elements of the channel matrix can be created by the use of multiple polarizations. Channels are no longer identically distributed.

- * The Rayleigh fading channels may become Ricean fading channels due to the presence of a fixed component (e.g., line-of-sight) in the channel.

2.3.2.2. Spatial correlated Rayleigh fading channels

When the antenna spacings or the angular spreading of the energy at the transmit and/or the receiver are not large enough, spatial correlations may result. The full spatial correlation matrix \mathbf{R} can be modeled as [PNG03]

$$\begin{aligned} \mathbf{R} &= \mathbb{E} \{ \text{vec}(\mathbf{H}) \text{vec}(\mathbf{H})^H \} \\ &= \begin{bmatrix} 1 & \rho_{1,2}^* & \rho_{1,3}^* & \cdots & \rho_{1,M_T \cdot M_R}^* \\ \rho_{1,2} & 1 & \rho_{2,3}^* & \cdots & \rho_{2,M_T \cdot M_R}^* \\ \rho_{1,3} & \rho_{2,3} & 1 & \cdots & \rho_{3,M_T \cdot M_R}^* \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{1,M_T \cdot M_R} & \rho_{2,M_T \cdot M_R} & \rho_{3,M_T \cdot M_R} & \cdots & 1 \end{bmatrix} \in \mathbb{C}^{M_T \cdot M_R \times M_T \cdot M_R}, \end{aligned} \quad (2.7)$$

where the correlation coefficients $0 \leq |\rho_{k,l}| \leq 1$ for $k, l = 1, \dots, M_T \cdot M_R$ and $k \neq l$. The channel matrix \mathbf{H} represents the spatially correlated Rayleigh fading channel which can be constructed by

$$\text{vec}(\mathbf{H}) = \mathbf{R}^{\frac{1}{2}} \text{vec}(\mathbf{H}_w). \quad (2.8)$$

Here, the operation $\text{vec}(\mathbf{H})$ stacks all elements of the matrix \mathbf{H} column by column into a column vector.

In many applications, a simpler and less general model, known as Kronecker model, is given by [PNG03]

$$\mathbf{H} = \mathbf{R}_r^{1/2} \mathbf{H}_w (\mathbf{R}_t^{1/2})^H, \quad (2.9)$$

where $\mathbf{R}_r \in \mathbb{C}^{M_R \times M_R}$ is the receive correlation matrix and $\mathbf{R}_t \in \mathbb{C}^{M_T \times M_T}$ is the transmit correlation matrix. Both \mathbf{R}_r and \mathbf{R}_t are positive semi-definite matrices. Compared to the model defined in equation (2.8), the Kronecker model has less degrees of freedom, since the model defined in equation (2.8) is able to capture any correlation effects between the elements of \mathbf{H} .

2.3.2.3. Ricean fading channels

The Rayleigh assumption typically holds in mobile scenarios. However, there are situations where the Rayleigh assumption is not suitable due to the fact that a strong coherent component may exist in real-world cellular networks. This component does not experience any fading

over time and corresponds to a line-of-sight (LOS) path. In the presence of a LOS component between the transmitter and the receiver, the MIMO channel may be modeled as a sum of a fixed component and a fading component [PNG03]

$$\mathbf{H} = \sqrt{\frac{K}{1+K}} \bar{\mathbf{H}} + \sqrt{\frac{1}{1+K}} \mathbf{H}_w, \quad (2.10)$$

where the first term on the right-hand-side (RHS) is the LOS component of the channel ($\sqrt{\frac{K}{1+K}} \bar{\mathbf{H}} = \mathbb{E}\{\mathbf{H}\}$) and the second term on the RHS is the fading component assuming uncorrelated fading. The factor $K \geq 0$ is the Ricean K -factor of the channel which is defined as the ratio of the power in the LOS component of the channel to the power in the fading component. If $K = 0$ the channel \mathbf{H} exhibits pure Rayleigh fading. At the other extreme $K = \infty$ corresponds to a non-fading channel.

Except for the above introduced analytical channel models, many other analytical channel models have been established for different applications [OC10] such as dual-polarized channel models and double-Rayleigh fading models for keyhole channels. In this chapter, we do not overview them in detail.

2.4. Capacity region of systems with MIMO

MIMO systems provide tremendous capacity gains and high spectral efficiencies [BCC⁺07], which has inspired significant activities to develop transmitter and receiver techniques that realize the capacity benefits and exploit diversity-multiplexing trade-offs. In this section, we focus on capacities of the single-user MIMO system, the multi-user MIMO system, and the multi-cell MIMO system in the Shannon theoretic sense. The Shannon capacity of a single-user time-invariant channel is defined as the maximum data rate at which reliable communication can be performed under a specific power constraint and without any constraint on the transmitter and receiver complexity. If the channel is time variant, we can define ergodic capacity and outage capacity. These different definitions are related to what is known about the channel state information (CSI) or channel distribution information (CDI) at the transmitter and/or the receiver. When the instantaneous CSI is known perfectly at both transmitter and receiver, the transmitter can adapt its transmission strategy according to the instantaneous CSI. In this case, the ergodic capacity is the maximum mutual information averaged over all channel states. The ergodic capacity is an appropriate capacity metric for fast varying channels and typically achieved by using an adaptive rate and power policy relative to the channel state variations. Alternatively, outage capacity is an appropriate capacity metric for slowly varying

channels and requires a fixed data rate in all non-outage channel states. Consequently, the average rate associated with the outage capacity is typically smaller than the ergodic capacity. When the channel varies too fast to capture the instantaneous CSI at the transmitter, only the receiver has perfect CSI. The transmission strategy is based on the CDI instead of the CSI. In this case, the transmitter must maintain a fixed-rate transmission strategy optimized with respect to its CDI. Then, the ergodic capacity defines the rate that can be achieved based on averaging over all channel states. Alternatively, the transmitter can send at a rate which cannot be supported by all channel states. In the poor channel states the receiver declares an outage and the transmitted data are lost. In this case, a percentage outage capacity p is defined to be the transmission rate that can be supported $(100 - p)\%$ of the time.

The calculation of the for general a channel distribution is a hard problem. Almost all references in this area are focused on three special distribution models described as follows [GJJV03, BCC⁺07]. In these three cases, the channel coefficients are modeled as complex jointly Gaussian random variables.

- * **Zero-mean spatially white (ZMSW) model:** This model typically captures the long-term average distribution of the channel coefficients. In this model, the channel mean is zero and the channel covariance is white.
- * **Channel mean information (CMI) model:** This model is applicable for a system where the feedback delay leads to an imperfect estimate at the transmitter. In this model, the mean of the channel distribution is non-zero while the covariance is modeled as being white with a constant scaling factor. The channel mean reflects the outdated channel measurement and the constant factor indicates the estimation error.
- * **Channel covariance information (CCI) model:** This model deals with the channel which varies too fast to track its mean. Therefore, the mean of the channel is set to be zero in this model and the information regarding the relative geometry of the propagation paths is captured by a non-white covariance matrix.

In single user MIMO (SU MIMO) systems, the achievable capacity is determined and identical for the uplink and downlink by assuming that the channel impulse responses are the same on the downlink and the uplink (e.g., TDD system), the channel is known at the transmitter and the receiver, and the downlink and uplink have the same transmit power. Without CSI at the transmitter, SU MIMO suffers only a small penalty (i.e., loss) on capacity. In multi-user MIMO (MU MIMO) systems, the achievable capacities are characterized by capacity regions of multiple access channels (MAC) on the uplink and broadcast channels (BC) on the downlink. The downlink and uplink channels are duals of each other under the conditions that the

channel impulse responses of each user are the same on the downlink and the uplink, the noise statistics of each receiver on the downlink are the same as those of the receiver on the uplink, and the sum of the individual power constraints on the uplink is equal to the power constraint on the downlink. The duality permits us to compute the capacity region of the MIMO BC much easier from the capacity region of the MIMO MAC, since it is very difficult to compute the MIMO BC capacity region directly. Without CSI at the transmitter, MU MIMO has a much larger penalty on the downlink than SU MIMO. For multi-cell MIMO, the analysis of the capacity is inevitably a hard problem. Some results do exist, but only for simplified interference models (e.g., treating interference as Gaussian noise) [BCC⁺07]. Such results cannot be extended to more general models which do not treat interference as Gaussian noise.

2.4.1. Single-User MIMO Capacity

In this part, we consider the single-user MIMO capacity first, since it is much easier to derive for a single user than for multiple users and the results have been known for many cases. Although there are still several open problems in obtaining the SU MIMO capacity under general assumptions of CSI and CDI which are shown later (see Section 2.4.1.6), for several interesting cases the solutions are known. In this section we will overview several cases with special CSI and CDI models. It can be found that the statistical properties of the channel and the correlation between the antenna elements influence the capacity significantly [BCC⁺07].

2.4.1.1. System Model

Considering a MIMO system where the transmitter has M_T transmit antennas and the receiver is equipped with M_R receive antennas, the channel is represented by a $M_R \times M_T$ matrix \mathbf{H} with the entries h_{ij} denoting the channel gain between the j th transmit antenna and the i th receive antenna. The received signal $\mathbf{y} \in \mathbb{C}^{M_R \times 1}$ can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (2.11)$$

where \mathbf{x} is the $M_T \times 1$ transmitted vector and \mathbf{n} indicates the $M_R \times 1$ additive white circularly symmetric complex Gaussian noise vector. The transmitted signal \mathbf{x} is subject to a transmit power constraint P_T , i.e., $\text{tr}(\mathbb{E}\{\mathbf{x}\mathbf{x}^H\}) \leq P_T$. The covariance matrix of the noise is normalized to be the identity matrix (i.e., $\sigma_n^2 = 1$). Furthermore, let us define $\tilde{P}_T = \frac{P_T}{\sigma_n^2}$ and the signal-to-noise ratio (SNR) equals $10 \log_{10} \tilde{P}_T$ (i.e., $\text{SNR} = 10 \log_{10} \tilde{P}_T$).

2.4.1.2. Constant MIMO Channel Capacity with Perfect CSIT and CSIR

When the channel is constant, it is reasonable to assume that the channel matrix \mathbf{H} is known perfectly at the transmitter and the receiver. Thus, the MIMO capacity can be calculated as [BCC⁺07]

$$C_{\text{SU}} = \max_{\mathbf{Q}: \text{tr}(\mathbf{Q})=P_{\text{T}}} \log_2 \det(\mathbf{I}_{M_{\text{R}}} + \mathbf{H}\mathbf{Q}\mathbf{H}^{\text{H}}), \quad (2.12)$$

where the matrix $\mathbf{Q} \in \mathbb{C}^{M_{\text{T}} \times M_{\text{T}}}$ is the positive semi-definite input covariance matrix. The total transmit power is denoted by P_{T} .

It has been shown in [Tel95] that the capacity is achieved by converting the MIMO channel into parallel, non-interfering SISO channels through a singular value decomposition (SVD) of the channel matrix \mathbf{H} and by applying water-filling as described below. The channel matrix \mathbf{H} is rewritten as $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\text{H}}$ by SVD, where the matrix $\mathbf{U} \in \mathbb{C}^{M_{\text{R}} \times M_{\text{R}}}$ is unitary and contains the left singular vectors of \mathbf{H} , the unitary matrix $\mathbf{V} \in \mathbb{C}^{M_{\text{T}} \times M_{\text{T}}}$ contains the right singular vectors of \mathbf{H} , and $\mathbf{\Sigma}$ is a matrix of size $M_{\text{R}} \times M_{\text{T}}$, and the (i, i) th element of $\mathbf{\Sigma}$ is the singular values σ_i of \mathbf{H} . Here, we use (i, i) to indicate the diagonal element on the i th row and the i th column. If the matrix \mathbf{H} has exactly r positive singular values ($r \leq \min(M_{\text{R}}, M_{\text{T}})$), the channel is decomposed into r subchannels by premultiplying the input by $\mathbf{V}_s \in \mathbb{C}^{M_{\text{T}} \times r}$ (i.e., $\mathbf{V} = [\mathbf{V}_s \ \mathbf{V}_n]$) and post-multiplying the output by the matrix $\mathbf{U}_s^{\text{H}} \in \mathbb{C}^{r \times M_{\text{R}}}$ (i.e., $\mathbf{U} = [\mathbf{U}_s \ \mathbf{U}_n]$). This linear transmit-receive strategy is widely used in practical systems because of its low complexity and significant performance benefit.

The water-filling algorithm is used as the optimal solution of the power loading over the parallel channels, which leads to [BCC⁺07]

$$P_i = \left(\mu - \frac{1}{\sigma_i^2} \right)_+, i = 1, \dots, r. \quad (2.13)$$

The waterfill level μ is chosen such that $\sum_{i=1}^r P_i = P_{\text{T}}$. Here, $(x)_+$ indicates $\max(x, 0)$. Then, the capacity is achieved by defining the covariance matrix $\mathbf{Q} = \mathbf{V}_s \mathbf{P} \mathbf{V}_s^{\text{H}}$ where the matrix $\mathbf{P} \in \mathbb{C}^{r \times r}$ is equal to $\text{diag}(P_1, \dots, P_r)$. The resulting capacity is given by

$$C_{\text{SU}} = \sum_{i=1}^r \left(\log_2(\mu \sigma_i^2) \right)_+. \quad (2.14)$$

Proof: see Appendix B.1

2.4.1.3. Spatially White Fading MIMO Channel Capacity with Perfect CSIT and CSIR

With fading, the ergodic capacity is simply defined as an average of the capacities achieved with each channel realization. We have [BCC⁺07]

$$C_{SU} = \mathbb{E}_{\mathbf{H}} \left\{ \max_{\mathbf{Q}: \text{tr}(\mathbf{Q})=P_T} \log_2 \det(\mathbf{I}_{M_R} + \mathbf{H}\mathbf{Q}\mathbf{H}^H) \right\}. \quad (2.15)$$

Here, the input covariance matrix \mathbf{Q} is changed with each channel realization. If the channel changes slowly enough to allow the reliable estimation of the channel at the receiver and the estimated CSI can be promptly fed back to the transmitter, we can assume perfect CSI at both transmitter and receiver. Then, the capacity for each channel realization can be calculated by equation (2.14) introduced in the subsection above. Correspondingly, the covariance matrix \mathbf{Q} for each channel realization is chosen according to the water-filling procedure.

2.4.1.4. Spatially White Fading MIMO Channel Capacity with Perfect CSIR and CDIT

Obtaining CSIT can be rather difficult in time-varying channels, as it generally requires either high-rate feedback from the receiver, or time-division duplex (TDD) operation on a sufficiently fast scale. Assume that we have perfect CSI at the receiver and a ZMSW channel distribution information at the transmitter (CDIT). When the transmitter does not know the instantaneous CSI and only has the knowledge of the fading distribution, it is impossible to align the input covariance matrix with the eigenmodes of the channel \mathbf{H} . In this case, the optimal transmit strategy is to equally allocate the transmit power in each direction [FG98, Tel95]. The optimum input covariance matrix is the scaled identity matrix, i.e., $\mathbf{Q} = \frac{P_T}{M_T} \mathbf{I}_{M_T}$. Thus the ergodic capacity is given as

$$C_{SU} = \mathbb{E}_{\mathbf{H}} \left\{ \log_2 \det(\mathbf{I}_{M_R} + \frac{\tilde{P}_T}{M_T} \mathbf{H}\mathbf{H}^H) \right\}. \quad (2.16)$$

2.4.1.5. Spatially correlated Fading MIMO Capacity

Although the capacity of spatially correlated MIMO under the CCI model by considering a general spatial correlation model is still an open problem, several research publications have studied capacities of spatially correlated fading MIMO under the CCI model by considering the Kronecker model (shown in equation (2.9)). Here, we like to point out that the work by Ozcelik et al. [OHW⁺03] has shown that the Kronecker model may not render the multipath structure correctly, leading to pessimistic capacity estimates in some cases.

Boche and Jorswieck have shown in [BJ03b, BJ03a] that the ergodic capacity of the correlated MISO systems with perfect CSI and without CSI at the transmitter is Schur-concave

(definition in Appendix B.2), regarding perfect CSI at the receiver. When only covariance knowledge of the channels is known at the transmitter (i.e., the CCI model) and the receiver has perfect CSI, the ergodic capacity of the correlated MISO system is Schur-convex (definition in Appendix B.2) with respect to the correlation properties. Additionally, in the presence of perfect CSIR, it has been shown in [CTKV02] that the capacity of spatially correlated MIMO channels still increases linearly with $\min(M_T, M_R)$, but the rate of growth is 10 – 20% smaller compared to the spatially white fading case. It is noticed that the analysis in [CTKV02] further considers several assumptions on the transmit and receive correlation matrices in the Kronecker model. For example, the correlation between the fading at two antennas depends only on the relative and not the absolute positions of the antennas. Now, we come to the question: does the correlated fading always decrease the growth of capacity? Actually the results exploited in [JG05a, BJ03b, BJ03a] have indicated that in the absence of CSIT correlations can improve capacity, especially for the fast fading channels.

2.4.1.6. Open Problems in Single-User MIMO

The results summarized in the existing publications (e.g., [GJJV03, BCC⁺07]) give the basis of the understanding of the SU MIMO channel capacity under special CSI and CDI assumptions. However, the knowledge of the MIMO capacity with CDI only is still far from complete. In the following, some of the many open problems have been pointed out as examples [GJJV03, BCC⁺07].

- CCI or CMI: Capacity is not known under the CCI model for completely general spatial correlations, or under the CMI model for an arbitrary channel mean matrix.
- Channel distribution information at the receiver (CDIR): Capacity for almost all cases with only CDIR are open problems.
- Outage capacity: Due to the fact that the outage capacity is less analytically tractable compared to the ergodic capacity, most existing results with partial CSI are only for the ergodic capacity. An abundance of open problems is left for the outage capacity.

2.4.2. Multi-User MIMO Capacity

Instead of a single real number for the channel capacity of the single-user MIMO system, the capacity of the multi-user MIMO system is a region which is defined as the set of simultaneously achieved rates. This region denotes the fundamental limit of reliable communications under certain channel characteristics. The capacity benefit of the multi-user MIMO system is even

greater than the single user MIMO case due to the increased number of antennas and users. In this part, we introduce the results for the capacity of the MIMO MAC and MIMO BC [GJJV03, BCC⁺07].

2.4.2.1. System Model

We consider a single cellular system where the base station has M_T antennas and the i th user has M_{R_i} antennas. There are K users in this system. We use $\mathbf{H}_i \in \mathbb{C}^{M_{R_i} \times M_T}$ to indicate the downlink channel of user i . Accordingly, the expression $\mathbf{H}_i^T \in \mathbb{C}^{M_T \times M_{R_i}}$ denotes the uplink channel of user i by assuming that the same channel is used for the uplink and downlink.

Considering the MIMO MAC as shown in Figure 1.3, the i th user transmits signal vector $\mathbf{x}_i \in \mathbb{C}^{M_{R_i} \times 1}$ to the base station. The received signal at the base station is denoted by $\mathbf{y} \in \mathbb{C}^{M_T \times 1}$ and given as

$$\mathbf{y} = \sum_{i=1}^K \mathbf{H}_i^T \mathbf{x}_i + \mathbf{n}_i, \quad (2.17)$$

where the noise vector $\mathbf{n}_i \in \mathbb{C}^{M_T \times 1}$ indicates circularly symmetric complex Gaussian noise with an identity covariance matrix (i.e., $\sigma_n^2 = 1$). Each user is subject to an individual power constraint P_i , i.e., $\text{tr}(\mathbb{E}\{\mathbf{x}_i \mathbf{x}_i^H\}) \leq P_i$. The total transmit power of all users is assumed to be $\sum_{i=1}^K P_i = P_T$.

Considering the MIMO BC as shown in Figure 1.2, the transmitted signal from the base station is denoted as $\mathbf{x} \in \mathbb{C}^{M_T \times 1}$. The received signal at the i th user $\mathbf{y}_i \in \mathbb{C}^{M_{R_i} \times 1}$ is given as

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{x} + \mathbf{n}_i. \quad (2.18)$$

Here, the noise $\mathbf{n}_i \in \mathbb{C}^{M_{R_i} \times 1}$ is assumed to be circularly symmetric complex Gaussian noise with an identity covariance matrix (i.e., $\sigma_n^2 = 1$). The base station is subject to a transmit power constraint P_T , i.e., $\text{tr}(\mathbb{E}\{\mathbf{x} \mathbf{x}^H\}) \leq P_T$. Let us define $\tilde{P}_T = \frac{P_T}{\sigma_n^2}$ and the signal-to-noise ratio (SNR) equals $10 \log_{10} \tilde{P}_T$ (i.e., $\text{SNR} = 10 \log_{10} \tilde{P}_T$).

2.4.2.2. Capacity Region of the MIMO MAC

In the MIMO MAC, K independent data rates are present. Therefore, the capacity region is a K -dimensional region. With individual power constraints $P_i, i = 1, \dots, K$ and successive decoding, the capacity region is given as [SXLK98]

$$\sum_{i=1}^K R_i \leq \max_{\text{tr}(\mathbf{Q}_i \mathbf{Q}_i^H) \leq P_i} \log_2 \det(\mathbf{I}_{M_T} + \sum_{i=1}^K \mathbf{H}_i^T \mathbf{Q}_i \mathbf{H}_i^*). \quad (2.19)$$

The boundary of the MIMO MAC capacity region can be achieved by finding optimal covariance matrices \mathbf{Q}_i . Since the MAC capacity region is convex, efficient convex optimization tools can be employed to solve it [BV04]. The iterative water-filling algorithm proposed in [YRBC04] is a more efficient numerical algorithm to compute the optimal covariance matrices that maximize the sum capacity, because it decomposes the multi-user problem into a sequence of single user problems. This method suggests that each user's covariance matrix is a water-filling covariance by regarding the interference generated by other users as noise.

For fading MIMO multiple-access channels, the ergodic capacity region with perfect CSIR and CSIT is equal to the time average of the capacity obtained at each fading instant with a constant transmit policy. If users' channel matrices are ZMSW and each user has the same power constraint, then the optimal covariance matrices are scaled version of the identity matrix, i.e., $\mathbf{Q}_i = \frac{P_i}{M_R} \mathbf{I}_{M_{R_i}}$ [Tel95]. In this case, the sum rate capacity of the MAC is expressed by

$$C_{\text{MAC}} = \mathbb{E}_{\mathbf{H}} \left\{ \log_2 \det \left(\mathbf{I}_{M_T} + \sum_{i=1}^K \frac{P_i}{M_R \sigma_n^2} \mathbf{H}_i^T \mathbf{H}_i^* \right) \right\}. \quad (2.20)$$

Note that this expression is exactly the ergodic capacity of the point-to-point MIMO fading channels with $\sum_{i=1}^K M_{R_i}$ transmit antennas and M_T receive antennas as given in equation (2.16). This implies that the lack of cooperation between the K transmitting users does not reduce the capacity under this fading model. The sum rate capacity of the MAC grows as $\min(M_T, \sum_{i=1}^K M_{R_i}) \log_2(\tilde{P}_T)$. Thus, for a system with large number of users, the capacity can be linearly increased by increasing the number of receive antennas M_T at the base station. This is a key benefit of MIMO in multi-user systems.

2.4.2.3. Capacity Region of the MIMO BC

Unlike the MIMO multiple-access channels, the general expression for the capacity region of the MIMO BC is still an open question due to the lack of a general theory on the capacity of non-degraded broadcast channels. However, dirty paper coding (DPC) [CS00, Cos83, YC01] has been shown as a capacity-achieving tool for MIMO BC and achieves the sum rate capacity of the MIMO BC.

DPC allows the multi-user interference to be pre-subtracted at the transmitter, but in such a way that the transmit power is not increased. The transmission strategy under dirty paper coding is exploited in the code design. The transmitter first chooses a codeword for receiver 1. Then, the codeword for receiver 2 is picked with full knowledge of the codeword for receiver 1. Therefore, the receiver 2 does not see the interference caused by the codeword intended for

receiver 1. Similarly, the transmitter chooses codeword for receiver 3 such that receiver 3 does not see the interference caused by the signals for receivers 1 and 2. To continue this process for all K receivers, an achievable set of rates is given as [BCC⁺07]

$$R_{\pi(i)} = \log_2 \frac{\det(\mathbf{I}_{M_R} + \mathbf{H}_{\pi(i)}(\sum_{j \geq i} \mathbf{Q}_{\pi(j)})\mathbf{H}_{\pi(i)}^H)}{\det(\mathbf{I}_{M_R} + \mathbf{H}_{\pi(i)}(\sum_{j > i} \mathbf{Q}_{\pi(j)})\mathbf{H}_{\pi(i)}^H)}, \quad i = 1, 2, \dots, K, \quad (2.21)$$

where $\pi(i)$ is a permutation of the user indices depending on the encoding order. The dirty paper region is the convex hull of the union of all such rates vectors over all covariance matrices $\mathbf{Q}_1, \dots, \mathbf{Q}_K$ with sum power constraint P_T . Notice that computing such a set of achievable rates is extremely complex, especially for a large number of antennas at either the transmitter or the receiver.

2.4.2.3.1. Duality between the MIMO BC and the MIMO MAC

The channel capacities of downlink and uplink may be different due to fundamental differences. For instance, downlink has a single power constraint associated with the base station, whereas the uplink has different power constraints associated with each user. Another difference is that on the downlink there is an additive noise term related to each user, while there is only one additive noise term related to the base station on the uplink. However, the fact that the downlink and the uplink channels looks like mirror images of each other implies a duality between them.

The duality between the scalar Gaussian BC and Gaussian MAC¹ was first shown in [JVG01], which indicates that the capacity region of a scalar Gaussian BC with power P_T is equal to the union of capacity regions of the dual MAC with power (P_1, \dots, P_K) such that $\sum_{i=1}^K P_i = P_T$. One key point is that to achieve the same rate vector in the BC and MAC, the decoding order must in general be reversed. A multiple-antenna extension has been proposed in [VJG03], which establishes the duality between the DPC region of the MIMO BC and the capacity region of the MIMO MAC. This duality is very useful from a numerical point of view, because the rate equations (2.21) are neither a convex nor a concave function of the covariance matrices, whereas the boundary of the dual MIMO MAC capacity region can be taken as a convex optimization problem. By applying the duality, the solved optimal MAC covariances can be transformed to the corresponding optimal BC covariances using the MAC-BC transformations given in [VJG03].

A simple expression of the capacity region of the MIMO BC can be observed in terms of

¹The scalar Gaussian BC and MAC channels defined in the paper are subject to a flat Rayleigh fading

the sum power MIMO MAC by the duality

$$\begin{aligned}
 C_{\text{BC}}(P_T; \{\mathbf{H}_i\}_{i=1}^K) &= C_{\text{MAC}}(P_1, \dots, P_K; \{\mathbf{H}_i^T\}_{i=1}^K) \\
 &= \bigcup_{\substack{\text{tr}(\mathbf{Q}_i \mathbf{Q}_i^H) \geq 0, \\ \sum_{i=1}^K \text{tr}(\mathbf{Q}_i \mathbf{Q}_i^H) \leq P_T}} \left(\sum_{i=1}^K R_i \leq \log_2 \det(\mathbf{I}_{M_T} + \sum_{i=1}^K \mathbf{H}_i^T \mathbf{Q}_i \mathbf{H}_i^*) \right),
 \end{aligned} \tag{2.22}$$

where the $\{\mathbf{Q}_i\}_{i=1}^K$ are the dual MIMO MAC covariance matrices and \cup is a union notation. Thus, the MIMO BC capacity region can be numerically computed by power convex optimization algorithms.

Similar to the MIMO MAC, the sum rate capacity of the MIMO BC with perfect CSIR and CSIT grows approximately as $\min(M_T, \sum_{i=1}^K M_{R_i}) \log_2(\tilde{P}_T)$ when \tilde{P}_T goes to infinity. In other words, this MIMO BC has a multiplexing gain of $\min(M_T, \sum_{i=1}^K M_{R_i})$.

2.4.2.4. Open Problems in Multi-User MIMO

A large number of open problems in multi-user MIMO cases exist. Some of them are listed as follows [GJJV03, BCC⁺07].

- BC with CSIR: The BC capacity is only known when both the transmitter and the receivers have perfect CSI.
- CDIT and CDIR: Due to the fact that perfect CSI is rarely possible, a study of capacity with CDI at both the transmitters and receivers for both MAC and BC is practical.
- Non-DPC techniques for BC: DPC is a very powerful capacity-achieving scheme, however, it has prohibitively high complexity for implementation in practice. Thus, non-DPC multi-user transmissions schemes for BC are more practical.

2.4.3. Multi-cell MIMO capacity

Transmissions in one cell are not limited within this cell due to the fundamental feature of wireless propagation. Thus, inter-cell interference is introduced between the cells where the resources (e.g., time, frequency, and space) are shared. In fact the majority of current systems are interference limited rather than noise limited. As a result, multi-cell environments must be explicitly considered in order to accurately assess the benefits of MIMO technology.

In the presence of multiple cells, multiple users, multiple antennas, and possibilities of cooperation between base stations, the analysis of the information-theoretic capacity of the cellular

network is inevitably a hard problem. A thorough understanding of it regarding fading and path loss effects has not been obtained yet. However, such an analysis is required because it can provide a common benchmark to assess the efficiency of any practical scheme. Some capacity results do exist, but only for simplified interference models. Such results provide intuition for the general performance behavior, whereas they are difficult to extend to general channel models. In this part, we review such analysis results from two categories. One assumes that the base stations cannot cooperate. The other considers full cooperation between base stations. Note that the cooperation, considered in this thesis, implies cooperative transmission and reception which takes place at the physical layer. It is different from the existing cooperation in the cellular systems such as GSM which takes place at the networking layer (e.g., hand-off).

2.4.3.1. Multi-cell MIMO without cooperation between base stations

If no cooperation is introduced, the channel becomes an interference channel and the system is interference limited. Unfortunately, the Shannon capacity of the interference channel is still an open problem in information theory. Even for the interference channel with two transmit-receive pairs and single transmit and receive antenna, the capacity region is not fully known [CG87].

In order to get insight into this problem, a Gaussian assumption has been proposed to treat the inter-cell interference as additive white Gaussian noise. This assumption can be viewed as a worst-case about the interference. If the structure of the interference is known, it can presumably help in the decoding of the desired signals. However, with this Gaussian assumption, the capacities of both uplink and downlink in a cellular environment can be determined by the analysis of the single-cell case. A lot of efforts have been spent on it and a lot of promising results have been achieved. For example, the capacity of a single-antenna cellular system uplink with the Gaussian assumption was obtained in [SW97]. The capacity of multiple transmit and receive antenna array in cellular system has been studied in [LT02, MK09, AV11, CDG00]. These capacity results show that an orthogonal multiple access method (e.g., TDMA, FDMA) is optimal in one cell, if the inter-cell interference is non-negligible.

2.4.3.2. Multi-cell MIMO with full cooperation between base stations

The full cooperation between base stations can be established via unrestricted backhaul links (error free and unlimited capacity) to a central processor. If such cooperation is assumed, the entire cellular system can be viewed as a single cell with a distributed antenna array at the

base station. It is typically referred as Wyner-type model [Wyn94]. Consequently, capacity results of a single-cell case can be applied. In fact, the geographical separations of the base station and the users have impacts on the channel gains in the composite network.

The analysis of the uplink capacity in cellular systems with cooperation was started in the early works of [Wyn94, HW93]. In both cases, the channels between the users and the base station are assumed to be AWGN channels and interference takes place only between adjacent cells. The Wyner model of [Wyn94, HW93] considers cells arranged on a line or in a more conventional hexagonal model. The per user capacity is derived in both cases. It is shown that the uplink capacity is achieved by using orthogonal multiple access methods (e.g., TDMA, FDMA) in each cell.

The downlink channel in cellular systems can be modeled as a MIMO broadcast channel by considering full cooperation between the base stations. The work in [SZ01] is viewed as a pioneer in the application of dirty paper coding (DPC) to a cellular system with cooperation between base stations. A single antenna at each user and each base station is considered in [SZ01]. Its results show that the capacity of the cellular downlink is enhanced by the relatively simple application of DPC, and this DPC scheme is asymptotically optimal at high SNRs. Inspired by [SZ01], a number of other works have also studied the capacity of cellular downlink channels for either finite or asymptotic regimes [HV04, GHW⁺11]. Especially, the authors in [GHW⁺11] did not utilize the Gaussian assumption, a new analytical co-channel interference model for MIMO cellular networks has been proposed by considering a Poisson spatial distribution of interfering transmitters and taking into account fading and shadowing effects in wireless channels. The downlink average capacity of MIMO cellular networks is derived based on this new interference model. The duality of the MAC and BC discussed in Section 2.4.2.3.1 can still be applied to the uplink and downlink of the cellular composite channels.

Research results show that full cooperating cellular systems can achieve a significantly increased capacity relative to the cellular systems without cooperation. However, the equivalence between multi-cell MIMO systems and MIMO systems only holds when an ideal backhaul is assumed. This assumption is very challenging from a practical perspective. Therefore, practical cooperation schemes must operate with a limited backhaul (i.e., finite capacity and finite latency). For future mobile wireless communications, the major goal is to find good signal processing and coding techniques that can achieve large cooperation gains with the information about the local channel state and local user data. It is referred as limited cooperation between base stations. Furthermore, full cooperating MIMO cellular systems involve a large number of antennas and users, which requires a reduction of the complexity of the current precoding

and decoding schemes (e.g., optimal precoding and optimal joint decoding).

2.5. Summary and conclusions

In this section, we have provided an overview of the fundamentals of the single user, the multi-user and the multi-cell MIMO technologies, which helps to understand the motivations, developments, and applications of the major contributions in this thesis. We have emphasized the major achievable benefits of MIMO techniques and reviewed the MIMO channel models from two groups: physical MIMO channel models and analytical MIMO channel models. The capacity results have been summarized for single-user MIMO channels, multi-user MIMO channels, and multi-cell MIMO channels. Significant capacity gains can be predicted for these systems under some assumptions (e.g., perfect CSI). In fact, the capacity gains are highly dependent on the nature of the CSI, the channel SNR, and the spatial correlation of the channels. An interesting insight is that with perfect CSI the spatial correlations are found to increase capacity at low SNRs and decrease capacity at high SNRs [GJJV03, BCC⁺07].

There are still many open problems in this area. Most capacity regions associated with multi-user MIMO channels remain unsolved such as the ergodic capacity for the MIMO BC under perfect receive CSI only. The capacity of MIMO cellular systems is a relatively open area. One reason is that the Shannon capacity of a cellular system is not well defined and heavily depends on some conditions (e.g., frequency assumption, propagation models, etc.). Another reason comes from the unsolved single-cell problems. We can expect that MIMO technology is likely to remain important and attractive for many years.

3. Linear precoding techniques for multi-user MIMO systems

3.1. Introduction and Motivation

Compared to time division multiple access (TDMA) systems, the multi-user MIMO systems employing space division multiple access (SDMA) can simultaneously serve a group of users and achieve a linear increase of the sum rate. Some information theoretic results in [CS00, Cos83, VJG02] have shown that dirty paper coding (DPC) achieves the capacity region of the Gaussian MIMO broadcast channel and provides the maximum diversity order. The capacity of the multi-user MIMO multiple access channels can be achieved via an MMSE receiver with successive interference cancellation [SXLK98]. A more exciting result has been shown by Jindal and Goldsmith [JG05b] that the sum rate of a multi-user MIMO broadcast channel employing DPC and a multi-user MIMO uplink utilizing successive interference cancellation are at most $\min(M_T, K)$ times larger than the maximum achievable sum rate of a system using TDMA, where M_T is the number of antennas at the base station (BS) and K is the number of users. This result is valid at all SNR regimes and independent of the number of receive antennas and the channel gain matrix.

As we have shown in Chapter 2, channel state information at the transmitter (CSIT) helps to increase the transmission rate and reduce the receiver complexity. In the multi-user MIMO BC, the base station is able to use the available CSI to reduce or completely eliminate the multi-user interference (MUI) via linear or non-linear precoding techniques (e.g., DPC or Tomlinson-Harashima precoding (THP)), which results in a significant capacity gain and a high spectral efficiency. In the multi-user MIMO MAC, each user terminal has only its own CSI, and it is impossible to acquire the CSIs of other users. In this case, space-time coding techniques can be employed at the user terminal to enhance reliability through diversity. The BS has the possibility to successfully mitigate the MUI by employing successive interference cancellation.

In this chapter we focus on the multi-user MIMO broadcast channels, since most of solutions for the downlink can be applied to the uplink in a straightforward way. Moreover, the cost of the hardware at the base station can be reduced by using the same multi-user MIMO processing techniques on both the downlink and the uplink. The precoding design for the multi-user MIMO broadcast channels has been one of the hot topics in the last decade. As

we have mentioned before, DPC is proposed as a capacity-optimal SDMA strategy to achieve the sum capacity of a multi-user MIMO downlink system. However, deploying DPC in a real system is very impractical due to the prohibitively high complexity at both the BS and the user terminals. Considering the demand for low complexity and low power consumption at the base station and the users, several sub-optimal linear precoding algorithms have been proposed such as zero-forcing (ZF) [VJ98], block diagonalization (BD) [SSH04], and regularized block diagonalization (RBD) [SH08]. Compared to DPC, they have significantly reduced complexities at the BS, while achieving the same multiplexing gain as DPC. Furthermore, they have the ability to adapt to various degrees of CSI. However, this low complexity comes at the price of a throughput loss relative to DPC. Therefore, a profound analytical study about how much loss does incur is really interesting and useful for the system design. A remarkable work has been done by Lee and Jindal in [LJ07], where the absolute rate and power offsets between ZF, BD, and DPC have been studied at high SNRs. In this chapter, we further consider RBD precoding which has an improved sum rate and diversity order relative to BD [SH08]. Moreover, RBD has the advantage that it is not constrained by the dimensionality condition that the aggregate number of receive antennas is not larger than the number of transmit antennas. We approximate the achievable throughput of an RBD based system [SH09a]. Compared to DPC and BD based systems, the bounds of the average rate and power offsets among these strategies are derived as a function of the system parameters (e.g., the number of users and receive antennas).

This chapter starts with an overview of several sub-optimal linear precoding algorithms (i.e., ZF, BD, and RBD) in Section 3.2. Then, in Section 3.3 the performance analysis of the throughput loss between these linear precoding algorithms and DPC at high SNRs is addressed. Finally, a summary is provided in Section 3.4.

3.2. Previous work on linear precoding techniques

In a multi-user MIMO broadcast channel employing linear precoding technique, the transmit signal vector $\mathbf{x} \in \mathbb{C}^{M_T \times 1}$ in equation (2.18) can be rewritten as

$$\mathbf{x} = \sum_{i=1}^K \mathbf{F}_i \mathbf{s}_i, \quad (3.1)$$

where $\mathbf{s}_i \in \mathbb{C}^{r_i \times 1}$ contains the data symbols for user i and r_i represents the number of data streams intended for the i th user. The matrix $\mathbf{F}_i \in \mathbb{C}^{M_T \times r_i}$ denotes the linear precoding matrix for user i . There are K users in this system. Then, the received signal $\mathbf{y}_i \in \mathbb{C}^{M_{R_i} \times 1}$ for user i

is given by

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{F}_i \mathbf{s}_i + \sum_{j=1, j \neq i}^K \mathbf{H}_i \mathbf{F}_j \mathbf{s}_j + \mathbf{n}_i. \quad (3.2)$$

Here, the matrix $\mathbf{H}_i \in \mathbb{C}^{M_{R_i} \times M_T}$ is the channel gain matrix for user i and the vector $\mathbf{n}_i \in \mathbb{C}^{M_{R_i} \times 1}$ represents additive white Gaussian noise (AWGN) with unit variance (i.e., $\sigma_n^2 = 1$). The first term on the right-hand-side (RHS) of equation (3.2) is the desired signals for the i th user. The term $\sum_{j=1, j \neq i}^K \mathbf{H}_i \mathbf{F}_j \mathbf{s}_j$ indicates the interference caused by the signals intended for the other users, which should be removed by the linear precoding.

Let us define the combined channel matrix and precoding matrix for all users as \mathbf{H} and \mathbf{F} , respectively. We have

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1^T & \mathbf{H}_2^T & \dots & \mathbf{H}_K^T \end{bmatrix}^T \in \mathbb{C}^{M_R \times M_T} \quad (3.3)$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \dots & \mathbf{F}_K \end{bmatrix} \in \mathbb{C}^{M_T \times r}, \quad (3.4)$$

where the term M_R indicates the total number of receive antennas (i.e., $M_R = \sum_{i=1}^K M_{R_i}$) and $r \leq \min(M_R, M_T)$ indicates the total number of transmitted data streams (i.e., $r = \sum_{i=1}^K r_i$).

The equivalent combined channel matrix of all users after the precoding is equal to

$$\mathbf{H}\mathbf{F} = \begin{bmatrix} \mathbf{H}_1 \mathbf{F}_1 & \mathbf{H}_1 \mathbf{F}_2 & \dots & \mathbf{H}_1 \mathbf{F}_K \\ \mathbf{H}_2 \mathbf{F}_1 & \mathbf{H}_2 \mathbf{F}_2 & \dots & \mathbf{H}_2 \mathbf{F}_K \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_K \mathbf{F}_1 & \mathbf{H}_K \mathbf{F}_2 & \dots & \mathbf{H}_K \mathbf{F}_K \end{bmatrix}, \quad (3.5)$$

where the effective channel of the user i is given by $\mathbf{H}_i \mathbf{F}_i$ and the off-diagonal elements $\mathbf{H}_i \mathbf{F}_j$ ($j = 1, \dots, K, j \neq i$) determine the interference caused by the signals intended for the other users. The linear precoding matrix \mathbf{F} can be viewed as attempting to block diagonalize the product $\mathbf{H}\mathbf{F}$. Although the optimal solution is not necessarily perfectly block diagonal, it will generally be near block diagonal.

3.2.1. Zero Forcing precoding

ZF precoding was investigated extensively in the literatures [PNG03, VJ98], where each user is equipped with one receive antenna. To this end, we have $K = M_R$.

The precoding matrix \mathbf{F} and a scaling factor β can be obtained from the following opti-

mization [PNG03, VJ98]

$$\{\mathbf{F}, \beta\} = \arg \min_{\mathbf{F}, \beta} \mathbb{E} \left\{ \|\mathbf{y} - \mathbf{s}\|_2^2 \right\}, \text{ s.t. : } \mathbf{y}|_{\mathbf{n}_i=0, \forall i} = \mathbf{s}, \mathbf{H}\mathbf{F} = \mathbf{I}_K, \quad (3.6)$$

where the vectors $\mathbf{s} = [s_1, s_2, \dots, s_K]^T \in \mathbb{C}^{K \times 1}$ and $\mathbf{y} = [y_1, y_2, \dots, y_K]^T \in \mathbb{C}^{K \times 1}$ denote the transmit signals and the received signals of all users, respectively. We assume that the complex data symbols s_i are independent and identically distributed (i.i.d.) random variables. The average power of the complex data symbols is set to one (i.e., $\mathbb{E} \{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_K$). The parameter β is chosen according to the transmit power constraint (i.e., $\beta^2 \text{tr}(\tilde{\mathbf{F}} \mathbb{E} \{\mathbf{s}\mathbf{s}^H\} \tilde{\mathbf{F}}^H) \leq P_T$). Then, the solution of this optimization problem is a pseudo-inverse of the combined channel matrix \mathbf{H} :

$$\mathbf{F} = \beta \tilde{\mathbf{F}}, \quad (3.7)$$

where

$$\tilde{\mathbf{F}} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1}, \quad \beta = \sqrt{\frac{P_T}{\|\tilde{\mathbf{F}}\|_{\mathcal{F}}^2}}. \quad (3.8)$$

With ZF precoding, the equivalent combined channel of all users (i.e., $\mathbf{H}\mathbf{F}$) is perfectly diagonalized under the condition $M_T \geq K$. In this case, the MUI can be entirely removed. However, ZF precoding has a serious drawback. The design of ZF disregards the noise term and focuses only on perfectly removing the interference term. It results in a noise enhancement problem [Gol05] which implies that the noise is amplified when the channel has a high attenuation. In this case, the transmit power is required to be increased in order to maintain the received SNR.

Compared to DPC, ZF is sub-optimal. A significant performance degradation is introduced as a penalty. The diversity order and array gain of each data stream is proportional to $M_T - M_R + 1$ [PNG03].

3.2.2. Block Diagonalization Precoding

Block diagonalization (BD) precoding was first proposed in [SSH04], which is designed to solve either the problem of maximizing the sum rate under a transmit power constraint or to minimize the total transmit power subject to achieving a desired arbitrary rate for each user. Compared to ZF precoding, BD approaches the optimal solution at high SNRs and allows multiple receive antennas at each user. However, BD is also limited by the dimensionality constraint that the number of transmit antennas has to be equal to or larger than the total number of receive antennas (i.e., $M_T \geq M_R$).

To eliminate all multi-user interference (MUI), the precoding matrix \mathbf{F}_i of BD is found to lie in the null space of the other users' channel matrices. Therefore, a multi-user MIMO broadcast channel is decomposed into multiple parallel independent single-user MIMO channels.

Let us define $\widetilde{\mathbf{H}}_i$ as

$$\widetilde{\mathbf{H}}_i = \begin{bmatrix} \mathbf{H}_1^T & \cdots & \mathbf{H}_{i-1}^T & \mathbf{H}_{i+1}^T & \cdots & \mathbf{H}_K^T \end{bmatrix}^T \in \mathbb{C}^{(M_R - M_{R_i}) \times M_T}. \quad (3.9)$$

The zero MUI constraint forces the matrix \mathbf{F}_i to lie in the null space of $\widetilde{\mathbf{H}}_i$. This definition introduces the dimensionality constraint (i.e., $M_T \geq M_R$) to guarantee that all users can be accommodated under the zero MUI constraint. Let us assume that the matrix $\widetilde{\mathbf{H}}_i$ has the rank \widetilde{L}_i and compute the singular value decomposition (SVD) of $\widetilde{\mathbf{H}}_i$. We have

$$\widetilde{\mathbf{H}}_i = \widetilde{\mathbf{U}}_i \widetilde{\Sigma}_i \begin{bmatrix} \widetilde{\mathbf{V}}_i^{(1)} & \widetilde{\mathbf{V}}_i^{(0)} \end{bmatrix}^H. \quad (3.10)$$

Here, $\widetilde{\mathbf{V}}_i^{(1)} \in \mathbb{C}^{M_T \times \widetilde{L}_i}$ holds the first \widetilde{L}_i right singular vectors. $\widetilde{\mathbf{V}}_i^{(0)} \in \mathbb{C}^{M_T \times (M_T - \widetilde{L}_i)}$ holds the last $M_T - \widetilde{L}_i$ right singular vectors which form an orthogonal basis for the null space of $\widetilde{\mathbf{H}}_i$. The columns of $\widetilde{\mathbf{V}}_i^{(0)}$ are candidates for the precoding matrix \mathbf{F}_i . Then, the effective channel of user i after the MUI elimination is defined as $\mathbf{H}_i \widetilde{\mathbf{V}}_i^{(0)}$, which has the dimension $M_{R_i} \times (M_T - \widetilde{L}_i)$. This effective channel is equivalent to a conventional single-user MIMO channel with $M_T - \widetilde{L}_i$ transmit antennas and M_{R_i} receive antennas. Let us compute SVD of $\mathbf{H}_i \widetilde{\mathbf{V}}_i^{(0)}$

$$\mathbf{H}_i \widetilde{\mathbf{V}}_i^{(0)} = \mathbf{U}_i \Sigma_i \begin{bmatrix} \mathbf{V}_i^{(1)} & \mathbf{V}_i^{(0)} \end{bmatrix} \in \mathbb{C}^{M_{R_i} \times (M_T - \widetilde{L}_i)} \quad (3.11)$$

and define the rank of the matrix $\mathbf{H}_i \widetilde{\mathbf{V}}_i^{(0)}$ be L_i . The product of the first L_i right singular vectors $\mathbf{V}_i^{(1)} \in \mathbb{C}^{(M_T - \widetilde{L}_i) \times L_i}$ and $\widetilde{\mathbf{V}}_i^{(0)}$ produces an orthogonal basis of dimension L_i and represents the transmission vectors that maximize the information rate for user i subject to the zero MUI constraint. Thus, the precoding matrix \mathbf{F} for all users can be defined as [SSH04]

$$\mathbf{F} = \begin{bmatrix} \widetilde{\mathbf{V}}_1^{(0)} \mathbf{V}_1^{(1)} & \widetilde{\mathbf{V}}_2^{(0)} \mathbf{V}_2^{(1)} & \cdots & \widetilde{\mathbf{V}}_K^{(0)} \mathbf{V}_K^{(1)} \end{bmatrix} \mathbf{\Lambda}^{\frac{1}{2}}, \quad (3.12)$$

where $\mathbf{\Lambda} \in \mathbb{C}^{r \times r}$ is a diagonal matrix whose elements λ_i scale the power transmitted into each of the columns of \mathbf{F} . The optimal power loading coefficients in $\mathbf{\Lambda}$ are found by performing water-filling on the singular values Σ_i from all users collected together, while assuming a total power constraint. The receive beamforming vectors of user i are chosen as $\mathbf{D}_i = \mathbf{U}_i^H$.

3.2.3. Regularized Block Diagonalization

Precoder design such as ZF and BD assumes zero MUI which imposes a dimensionality constraint regarding to the total number of antennas at the BS and the users. When the total number of receive antennas is larger than the number of antennas at the BS, the solutions that overcome this dimensionality constraint use either only a subset of antennas or a subset of eigenmodes [SSH04, TUBN05]. However, an additional control overhead is usually required to inform the users about the selection. Regularized block diagonalization (RBD) [SH08] was proposed to release this dimensionality constraint, while using as much as possible of the available spatial degrees of freedom and minimizing the interference between different users at the same time.

The RBD precoding matrix is described as [SH08]

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \cdots & \mathbf{F}_K \end{bmatrix} = \beta \mathbf{F}_a \mathbf{F}_b, \quad (3.13)$$

where

$$\mathbf{F}_a = \begin{bmatrix} \mathbf{F}_{a_1} & \mathbf{F}_{a_2} & \cdots & \mathbf{F}_{a_K} \end{bmatrix} \in \mathbb{C}^{M_T \times M_R}$$

and

$$\mathbf{F}_b = \begin{bmatrix} \mathbf{F}_{b_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{b_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{F}_{b_K} \end{bmatrix} \in \mathbb{C}^{M_R \times r}.$$

Here, $r \leq \min(M_R, M_T)$ is the total number of transmitted data streams. The parameter β is chosen to set the total transmit power to P_T .

RBD precoding is performed in two steps. In the first step the precoding matrix \mathbf{F}_a is designed to suppress the MUI by reducing the overlap of the row spaces spanned by the effective channel matrices of different users (i.e., $\mathbf{H}_i \mathbf{F}_{a_i}$, for $i = 1, \dots, K$). The equivalent combined channel matrix of all users after the precoding matrix \mathbf{F}_a is equal to

$$\mathbf{H} \mathbf{F}_a = \begin{bmatrix} \mathbf{H}_1 \mathbf{F}_{a_1} & \mathbf{H}_1 \mathbf{F}_{a_2} & \cdots & \mathbf{H}_1 \mathbf{F}_{a_K} \\ \mathbf{H}_2 \mathbf{F}_{a_1} & \mathbf{H}_2 \mathbf{F}_{a_2} & \cdots & \mathbf{H}_2 \mathbf{F}_{a_K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_K \mathbf{F}_{a_1} & \mathbf{H}_K \mathbf{F}_{a_2} & \cdots & \mathbf{H}_K \mathbf{F}_{a_K} \end{bmatrix}, \quad (3.14)$$

where the i th user's effective channel is given by $\mathbf{H}_i \mathbf{F}_{a_i}$ and the interference generated to the other users is determined by $\widetilde{\mathbf{H}}_i \mathbf{F}_{a_i}$. The matrix $\widetilde{\mathbf{H}}_i$ has the same definition as in equation

(3.9). The Frobenius norm of the matrix $\widetilde{\mathbf{H}}_i \mathbf{F}_{a_i}$ is related to the level of the overlap of the row spaces of the effective channels of different users $\mathbf{H}_i \mathbf{F}_{a_i}$, $i = 1, \dots, K$. Then, the precoding matrix \mathbf{F}_a results from the following optimization [SH08]

$$\mathbf{F}_a = \min_{\mathbf{F}_a} \mathbb{E} \left\{ \sum_{i=1}^K \|\widetilde{\mathbf{H}}_i \mathbf{F}_{a_i}\|_{\mathcal{F}}^2 + \frac{\|\mathbf{n}\|_{\mathcal{F}}^2}{\beta^2} \right\}. \quad (3.15)$$

Here, the vector $\mathbf{n} \in \mathbb{C}^{M_R \times 1}$ contains the samples of a zero mean additive white Gaussian noise at the receive antennas. The noise is assumed to be uncorrelated with the same variance σ_n^2 . The solution of the minimization of equation (3.15) results in

$$\mathbf{F}_{a_i} = \widetilde{\mathbf{V}}_i \left(\widetilde{\boldsymbol{\Sigma}}_i^2 + \frac{M_R \sigma_n^2}{P_T} \mathbf{I}_{M_T} \right)^{-1/2} \in \mathbb{C}^{M_T \times M_T}, \quad (3.16)$$

where $\widetilde{\mathbf{V}}_i$ and $\widetilde{\boldsymbol{\Sigma}}_i$ are the matrices of the right singular vectors and the diagonal matrix of the matrix $\widetilde{\mathbf{H}}_i$ (i.e., $\widetilde{\mathbf{H}}_i = \widetilde{\mathbf{U}}_i \widetilde{\boldsymbol{\Sigma}}_i \widetilde{\mathbf{V}}_i^H$), respectively. The precoding matrix \mathbf{F}_a forces each user to transmit on the eigenmodes of the combined channel matrix of all other users with the power $\frac{M_R \sigma_n^2}{P_T}$. At high SNRs, each user transmits only in the nullspace of all other users, which indicates that the off-diagonal block matrices in equation (3.14) converge to zero with increasing SNR.

In step two the precoding matrix \mathbf{F}_b is designed to optimize the system performance by any specific optimization criterion assuming a set of parallel single user MIMO channels. The matrix $\mathbf{F}_{b_i} \in \mathbb{C}^{M_T \times r_i}$ has the form

$$\mathbf{F}_{b_i} = \mathbf{V}_i \mathbf{D}_i. \quad (3.17)$$

Here, the matrix $\mathbf{V}_i \in \mathbb{C}^{M_T \times M_T}$ is the right singular vector matrix of $\mathbf{H}_i \mathbf{F}_{a_i}$ (i.e., $\mathbf{H}_i \mathbf{F}_{a_i} = \mathbf{U}_i \boldsymbol{\Sigma}_i \mathbf{V}_i^H$) and the matrix $\mathbf{D}_i \in \mathbb{C}^{M_T \times r_i}$ is the power loading according to the optimization criteria proposed in [SH08]. After the generation of the matrices \mathbf{F}_a and \mathbf{F}_b , the parameter β is used to set the total transmit power to P_T , i.e., $\beta^2 = P_T / \|\mathbf{F}_a \mathbf{F}_b\|_{\mathcal{F}}^2$ due to the assumption $\mathbb{E} \{ \mathbf{s}_i \mathbf{s}_i^H \} = \mathbf{I}_{r_i}$ applying in the RBD design. The receive beamforming vectors of user i are chosen as $\mathbf{U}_i^H \in \mathbb{C}^{M_T \times M_{R_i}}$, which is the left singular vector matrix of $\mathbf{H}_i \mathbf{F}_{a_i}$.

In Figures 3.1 and 3.2 we compare the achievable throughput of multi-user MIMO systems employing ZF, BD and RBD precoding. We assume flat fading channels with $M_T = 4$ transmit antennas. Moreover, 4 users are considered in Figure 3.1, where each user has only one antenna. Dominant eigenmode transmission is performed in this case. In Figure 3.2, two users are considered for BD and RBD precoding, where each user is equipped with two receive antennas and receives two different data streams from the base station. Water-filling is employed to

allocate the power to the different data streams for all users. It is shown that a throughput loss relative to DPC exists. Furthermore, it is shown that BD and RBD precoding provide obviously higher achievable throughput than ZF precoding. RBD precoding outperforms BD and ZF at low SNRs.

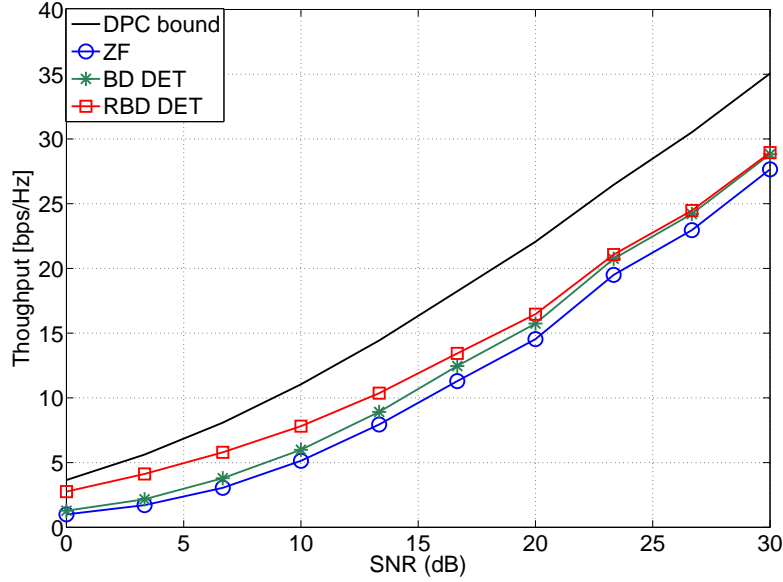


Figure 3.1.: Achievable throughput for a system with configuration $\{1, 1, 1, 1\} \times 4$.

3.3. Achievable throughput approximation at high SNRs

This section is devoted to the achievable throughput analysis for the optimal strategy of dirty paper coding (DPC) and sub-optimal lower complexity linear precoding techniques (e.g., ZF, BD and RBD). As we have shown in Figures 3.1 and 3.2, linear precodings such as ZF, BD, and RBD can achieve the same multiplexing gain as DPC, but do incur a throughput loss compared to DPC. The authors in [JG05b, SCA⁺06] have analyzed the ratio between the achievable sum rates of DPC and BD precoding. A more practical metric has been proposed in [LJ07] to study the absolute rate and power offsets between the DPC, BD and ZF algorithms at high SNRs. In [LJ07], the achievable throughput of DPC, ZF, and BD is approximated for high SNRs by introducing a capacity approximation framework which was firstly proposed in [SV01] for the context of code-division multiple access (CDMA) with random spreading. The details of this capacity approximation framework are described in the following Section 3.3.2. The rate and

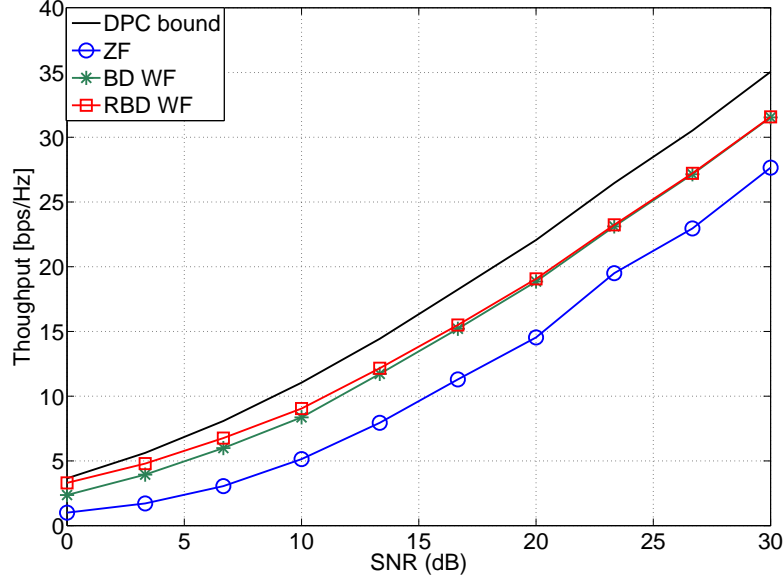


Figure 3.2.: Achievable throughput for a system with configuration $\{2, 2\} \times 4$, ZF system with configuration $\{1, 1, 1, 1\} \times 4$.

power offsets between them are derived as a function of the number of users and the number of the transmit and receive antennas. Inspired by it, we further approximate the achievable throughput of an RBD based multi-user MIMO downlink channel at high SNRs in this section and obtain the bounds of the average rate and power offsets between RBD, BD, and DCP as a function of the system parameters (e.g., the number of users and the number of the transmit and receive antennas).

3.3.1. System Model

We consider a multi-user MIMO system with a single base station (BS) and K users, where the BS is equipped with M_T transmit antennas. For notational simplicity, each user has M_r receive antennas (i.e., $M_{R_i} = M_r$ for $\forall i$). The aggregate number of receive antennas, denoted by M_R , is equal to $K \cdot M_r$. The propagation channel between the BS and each user is assumed to be a spatially uncorrelated Rayleigh fading channel. The received signal of the i th user is expressed as

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{x} + \mathbf{n}_i \quad (3.18)$$

where $\mathbf{x} \in \mathbb{C}^{M_T \times 1}$ is the transmit signal vector and has the same definition in equation (3.1). With an average total power limitation P_T at the BS, we require that $\text{tr}(\mathbb{E}\{\mathbf{x}\mathbf{x}^H\}) \leq P_T$. The matrix $\mathbf{H}_i \in \mathbb{C}^{M_r \times M_T}$ is the channel gain matrix for user i and the vector $\mathbf{n}_i \in \mathbb{C}^{M_r \times 1}$ represents additive white Gaussian noise (AWGN) with unit variance (i.e., $\sigma_n^2 = 1$). Here, we still use $\mathbf{H} \in \mathbb{C}^{M_R \times M_T}$ to indicate the combined channel of all users (defined in equation (3.3)). Furthermore, we assume that each user has perfect knowledge of its own channel and the BS has perfect knowledge of all users' channels. In order to be consistent with the definitions in Chapter 2, let us define $\tilde{P}_T = \frac{P_T}{\sigma_n^2}$ and the signal-to-noise ratio (SNR) in dB equals $10 \log_{10} \tilde{P}_T$ (i.e., $\text{SNR} = 10 \log_{10} \tilde{P}_T$).

3.3.2. Capacity Approximation Framework

The capacity approximation framework used in this section is the same as the framework in [LJ07]. This capacity approximation framework was firstly proposed in [SV01] for the context of code-division multiple access (CDMA) with random spreading. Sequentially, this framework has been utilized to quantify the capacities of single user MIMO at high SNRs for independent and identically distributed (i.i.d) Rayleigh fading channels and correlated Rayleigh fading channels in [LTV05]. Furthermore, the authors of [JSO08] have also introduced this framework to analyze the achievable throughput of TDMA based opportunistic beamforming at high SNRs.

This capacity approximation framework enables the channel capacity $C(\tilde{P}_T)$ to be well approximated at high SNRs as [LJ07]

$$C(\tilde{P}_T) = \mathcal{S}_\infty \cdot (\log_2 \tilde{P}_T - \mathcal{L}_\infty) + o(1) \quad (3.19)$$

$$= \mathcal{S}_\infty \cdot \left(\frac{\text{SNR}}{3 \text{ dB}} - \mathcal{L}_\infty \right) + o(1) ,$$

$$\text{where } \mathcal{S}_\infty = \lim_{\tilde{P}_T \rightarrow \infty} \ln 2 \cdot \tilde{P}_T \dot{C}(\tilde{P}_T) \quad (3.20)$$

$$\text{and } \mathcal{L}_\infty = \lim_{\tilde{P}_T \rightarrow \infty} \left(\log_2 \tilde{P}_T - \frac{C(\tilde{P}_T)}{\mathcal{S}_\infty} \right) . \quad (3.21)$$

Here, \mathcal{S}_∞ represents the multiplexing gain (i.e., the asymptotic slope of the spectral efficiency in bps/Hz per 3 dB) and \mathcal{L}_∞ refers to the power offset in 3 dB units as shown in Figure 3.3. We use $\dot{C}(P_T)$ to denote the first derivative of the capacity with respect to P_T . Furthermore, $O(1)$ is a notation which describes the limiting behavior of the capacity. When $\tilde{P}_T \rightarrow \infty$, the term $O(1)$ vanishes. For either multi-user MIMO broadcast channels or point-to-point MIMO channels, the multiplexing gain \mathcal{S}_∞ is found to be the minimum of the aggregate number of

receive antennas and the number of transmit antennas under the assumption of uncorrelated Rayleigh fading (i.e., $\mathcal{S}_\infty = \min(M_T, M_R)$). The rate offset \mathcal{L}_∞ depends on the fading statistics and the transmission strategies. In the following sections, we start with an overview of the throughput approximations for DPC, ZF, and BD.

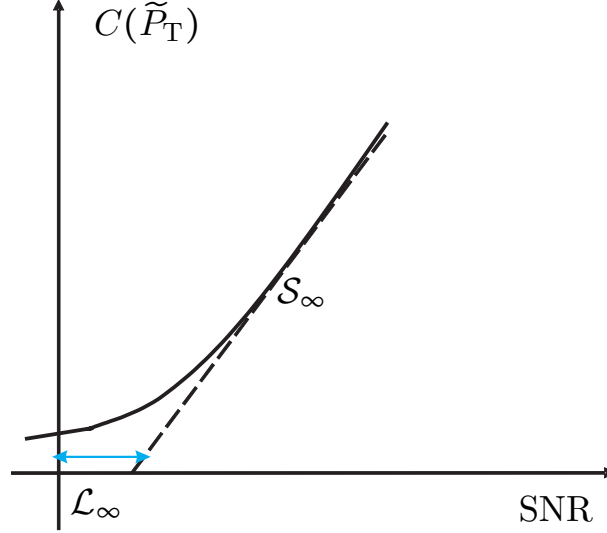


Figure 3.3.: An example of the capacity approximation at high SNRs.

3.3.3. Throughput approximation for DPC

The achievable sum rate of DPC can be expressed from the MIMO BC-MAC duality as¹ [VJG03]

$$C_{\text{DPC}}(\mathbf{H}, \tilde{P}_T) = \max_{\sum_{i=1}^K \text{tr}(\mathbf{Q}_i) \leq P_T} \log_2 \left| \mathbf{I}_{M_T} + \sum_{i=1}^K \mathbf{H}_i^H \mathbf{Q}_i \mathbf{H}_i \right|. \quad (3.22)$$

There is no known closed-form solution for this capacity. But it has been shown in [Jin05] that $C_{\text{DPC}}(\mathbf{H}, \tilde{P}_T)$ converges to the capacity of a point-to-point MIMO channel with the matrix $\mathbf{H} \in \mathbb{C}^{M_R \times M_T}$ under the condition $M_T \geq M_R$, thus

$$\lim_{\tilde{P}_T \rightarrow \infty} \left(C_{\text{DPC}}(\mathbf{H}, \tilde{P}_T) - \log_2 \left| \mathbf{I}_{M_R} + \frac{\tilde{P}_T}{M_R} \mathbf{H} \mathbf{H}^H \right| \right) = 0, \quad (3.23)$$

which corresponds to the fact that choosing each of the covariance matrices as $\mathbf{Q}_i = \frac{P_T}{M_R} \mathbf{I}_{M_R}$ in (3.22) is asymptotically optimal at high SNRs. As a result, the approximation of the DPC

¹In order to simplify the equations, we use $|\cdot|$ instead of $\det(\cdot)$ to indicate the determinant of a matrix.

throughput at high SNRs can be found as [LJ07]

$$C_{\text{DPC}}(\mathbf{H}, \tilde{P}_T) \cong M_R \log_2 \tilde{P}_T - M_R \log_2 M_R + \log_2 |\mathbf{H}\mathbf{H}^H|. \quad (3.24)$$

Here, \cong refers to equivalence in the limit (i.e., the difference between both sides converges to zero as $P_T \rightarrow \infty$). Applying the approximation framework we get

$$\mathcal{S}_\infty = M_R, \quad (3.25)$$

$$\mathcal{L}_\infty = \log_2 M_R - \frac{1}{\mathcal{S}_\infty} \log_2 |\mathbf{H}\mathbf{H}^H|. \quad (3.26)$$

The proof can be found in Appendix C.1.

Note that with DPC and the equal power allocation, the multi-user MIMO broadcast channel is equivalent to the $M_T \times M_R$ point-to-point MIMO channel where the CSI is known at the transmitter and $M_T \geq M_R$.

3.3.4. Throughput approximation for ZF

Since zero-forcing (ZF) precoding can totally eliminate the multi-user interference (MUI), the received signal of user i is given by

$$y_i = \mathbf{h}_i^T \mathbf{f}_i s_i + n_i, \quad (3.27)$$

where the i th user's channel matrix \mathbf{h}_i has the dimension $M_T \times 1$ and the precoding matrix \mathbf{f}_i is a column vector of the dimension $M_T \times 1$. Thus, the ZF based system is converted into M_R parallel channels with the effective channel $g_i = \mathbf{h}_i^T \mathbf{f}_i$. The maximum achievable throughput is obtained by optimizing the power allocation across these parallel channels

$$C_{\text{ZF}}(\mathbf{H}, \tilde{P}_T) = \max_{\sum_{i=1}^K P_i \leq \tilde{P}_T} \sum_{i=1}^K \log_2 (1 + P_i |g_i|^2). \quad (3.28)$$

Since the optimal power allocation policy converges to equal power allocation at high SNRs, we have [LJ07]

$$C_{\text{ZF}}(\mathbf{H}, \tilde{P}_T) \cong M_R \log_2 \tilde{P}_T - M_R \log_2 M_R + \log_2 \prod_{i=1}^K |g_i|^2. \quad (3.29)$$

Here, the effective channel norm $|g_i|^2$ obeys a chi-square distribution (definition shown in Appendix C.2) with $2(M_T - K + 1)$ degrees of freedom due to the fact that the ZF precoding vector for each user is chosen orthogonal to the other users' channels. Therefore, $K - 1$ degrees

of freedom are consumed at the BS regarding this orthogonality constraint. Then, the degree of freedom for the i th user is $M_T - K + 1$ which leads to the chi-squared distributed $|g_i|^2$ with $2(M_T - K + 1)$ degrees of freedom. It can be also interpreted as that a ZF-based system is equivalent to K parallel SISO channels.

Compared to equation (3.24), the approximation of ZF is identical to that for DPC except for the final term. Let us define the rate offset between the achievable throughput of DPC and ZF at high SNRs as

$$\Delta_{\text{DPC-ZF}}(\mathbf{H}) = \lim_{\tilde{P}_T \rightarrow \infty} [C_{\text{DPC}}(\mathbf{H}, \tilde{P}_T) - C_{\text{ZF}}(\mathbf{H}, \tilde{P}_T)]. \quad (3.30)$$

Considering equations (3.24) and (3.29), the rate loss incurred by ZF for one channel realization is

$$\Delta_{\text{DPC-ZF}}(\mathbf{H}) = \log_2 \frac{|\mathbf{H}\mathbf{H}^H|}{\prod_{i=1}^K |g_i|^2}. \quad (3.31)$$

By averaging over the fading distribution, the average rate offset can be expressed as

$$\bar{\Delta}_{\text{DPC-ZF}} = \mathbb{E} \{ \Delta_{\text{DPC-ZF}}(\mathbf{H}) \}, \quad (3.32)$$

which allows a comparison of the average throughput. Note that under the i.i.d. Rayleigh fading assumption, the matrix $\mathbf{H}\mathbf{H}^H$ is Wishart distributed (definition shown in Appendix C.2) with M_T degrees of freedom and $|g_i|^2$ has a chi-square distribution with $2(M_T - K + 1)$ degree of freedom. Utilizing the expression for the expected log-determinant of Wishart matrices and chi-squared variables in terms of Euler's digamma function (definition shown in Appendix C.2), e.g.,

$$\mathbb{E} \{ \log_e \det \mathbf{W} \} = \sum_{\ell=0}^{m-1} \varphi(n - \ell) \quad (3.33)$$

where the matrix $\mathbf{W} \in \mathbb{C}^{m \times m}$ is a complex Wishart matrix with the degrees of freedom n ($n \geq m$), the average rate offset can be computed in closed form as follows.

Theorem 3.3.1. *The average rate offset between DPC and ZF in the i.i.d. Rayleigh fading at high SNRs is given by [LJ07]*

$$\bar{\Delta}_{\text{DPC-ZF}} = (\log_2 e) \sum_{i=1}^{K-1} \frac{i}{M_T - i} \text{ (bps/Hz)}. \quad (3.34)$$

Proof: see [LJ07].

Since the capacity curve has a slope of $\frac{\mathcal{S}_\infty}{3}$ in units of bps/Hz/dB, the average rate offset

can easily be translated into an average power offset. Let us define the average power offset by $\Delta \bar{P}_{\text{DPC-ZF}}$ and get

$$\Delta \bar{P}_{\text{DPC-ZF}} = \frac{3\bar{\Delta}_{\text{DPC-ZF}}}{\mathcal{S}_{\infty}} = \frac{3\bar{\Delta}_{\text{DPC-ZF}}}{M_{\text{R}}}. \quad (3.35)$$

3.3.5. Throughput approximation for BD

Block diagonalization (BD) is an extension of the zero-forcing precoding for the case that the users have multiple receive antennas as explained in Section 3.2.2. With BD, the i th user's precoding matrix $\tilde{\mathbf{F}}_i$ (i.e., $\tilde{\mathbf{F}}_i = \tilde{\mathbf{V}}_i^{(0)}$) lies in the null space of all other users' channels (i.e., $\tilde{\mathbf{V}}_i^{(0)}$ in equation (3.10)). Thus, the system is converted into K parallel MIMO channels with effective channel matrices $\mathbf{G}_i^{(0)} = \mathbf{H}_i \tilde{\mathbf{F}}_i = \mathbf{H}_i \tilde{\mathbf{V}}_i^{(0)}$, $i = 1, \dots, K$. There is no MUI at each user. The received signal for user i is given by

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{F}_i \mathbf{s}_i + \mathbf{n}_i, \quad (3.36)$$

The achievable throughput of BD-based systems is given by

$$C_{\text{BD}}(\mathbf{H}, \tilde{P}_{\text{T}}) = \max_{\sum_{i=1}^K \text{tr}\{\mathbf{Q}_i\} \leq P_{\text{T}}} \sum_{i=1}^K \log_2 \left| \mathbf{I}_{M_{\text{R}_i}} + \mathbf{G}_i^{(0)} \mathbf{Q}_i \mathbf{G}_i^{(0)H} \right|. \quad (3.37)$$

Since the optimal power allocation policy converges to equal power allocation (i.e., $\mathbf{Q}_i = \frac{P_{\text{T}}}{M_{\text{R}}} \mathbf{I}_{M_{\text{R}_i}}$) at high SNRs, the throughput approximation of BD is performed as [LJ07]

$$C_{\text{BD}}(\mathbf{H}, \tilde{P}_{\text{T}}) \cong M_{\text{R}} \log_2 \tilde{P}_{\text{T}} - M_{\text{R}} \log_2 M_{\text{R}} + \log_2 \prod_{i=1}^K \left| \mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H} \right|. \quad (3.38)$$

Here, the orthogonality constraint consumes $(K-1) \cdot M_r$ degrees of freedom. Therefore, the degrees of freedom for the i th user are $M_{\text{T}} - (K-1) \cdot M_r$, which implies that the $M_r \times M_r$ matrix $\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H}$ has a Wishart distribution with $M_{\text{T}} - M_{\text{R}} + M_r$ degrees of freedom.

Similar to the analysis for ZF, let us define the rate loss relative to DPC for one channel realization as

$$\Delta_{\text{DPC-BD}}(\mathbf{H}) = \lim_{\tilde{P}_{\text{T}} \rightarrow \infty} [C_{\text{DPC}}(\mathbf{H}, \tilde{P}_{\text{T}}) - C_{\text{BD}}(\mathbf{H}, \tilde{P}_{\text{T}})] \quad (3.39)$$

and the average rate offset over Rayleigh fading as

$$\bar{\Delta}_{\text{DPC-BD}} = \text{E} \{ \Delta_{\text{DPC-BD}}(\mathbf{H}) \}, \quad (3.40)$$

respectively. We reach to the following theorem.

Theorem 3.3.2. *The average rate offset between DPC and BD in the i.i.d. Rayleigh fading at high SNRs is given by [LJ07]*

$$\bar{\Delta}_{\text{DPC-BD}} = (\log_2 e) \sum_{k=0}^{K-1} \sum_{n=0}^{M_r-1} \sum_{i=kM_r+1}^{(K-1)M_r} \frac{1}{M_T - n - i} \text{ (bps/Hz)}. \quad (3.41)$$

Proof: see [LJ07].

Consequently, the average power offset can be obtained as

$$\Delta \bar{P}_{\text{DPC-BD}} = \frac{3\bar{\Delta}_{\text{DPC-BD}}}{\mathcal{S}_\infty} = \frac{3\bar{\Delta}_{\text{DPC-BD}}}{M_R}. \quad (3.42)$$

With equations (3.34) and (3.41), it is possible to gain more intuition by considering the rate and power offsets between BD and ZF, where we consider K users with M_r receive antennas each for BD and KM_r users with one antenna each for ZF in order to ensure the same number of data streams for the both cases. Let us define $M_T = \alpha KM_r$ where α is an integer and greater than zero.

Theorem 3.3.3. *If $\alpha \geq 1$ and $M_r > 1$, the average rate offset $\bar{\Delta}_{\text{BD-ZF}}$ and power offset $\Delta \bar{P}_{\text{BD-ZF}}$ between BD and ZF in the i.i.d. Rayleigh fading at high SNRs are given by [LJ07]*

$$\begin{aligned} \bar{\Delta}_{\text{BD-ZF}} &= \bar{\Delta}_{\text{DPC-ZF}} - \bar{\Delta}_{\text{DPC-BD}}, \\ &= (\log_2 e) K \sum_{i=1}^{M_r-1} \frac{M_r - i}{(\alpha - 1)KM_r + i} \text{ (bps/Hz)} \end{aligned} \quad (3.43)$$

$$(3.44)$$

and

$$\Delta \bar{P}_{\text{BD-ZF}} = \frac{3(\log_2 e)}{M_r} \sum_{i=1}^{M_r-1} \frac{M_r - i}{(\alpha - 1)KM_r + i} \text{ (dB)}, \quad (3.45)$$

respectively.

Proof: see [LJ07].

When $\alpha = 1$, from equation (3.45) we get

$$\Delta \bar{P}_{\text{BD-ZF}} = \frac{3(\log_2 e)}{M_r} \sum_{i=1}^{M_r-1} \frac{M_r - i}{i} \text{ (dB)}. \quad (3.46)$$

This expression shows us an important feature that the average power offset $\Delta \bar{P}_{\text{BD-ZF}}$ only depends on the number of receive antennas M_r and is independent of the number of transmit

antennas M_T . For example, equation (3.46) indicates that the power advantage of using BD in the $M_r = 2$ system is $\Delta \bar{P}_{\text{BD-ZF}} = 2.1640$ (dB) relative to performing ZF. This power advantage is the same for $M_T = 4$, $K = 4$, and $M_r = 1$ (i.e., ZF-based system) vs. $M_T = 4$, $K = 2$, and $M_r = 2$ (i.e., BD-based system) as well as for $M_T = 6$ and $K = 6$, $M_r = 1$ (i.e., ZF-based system) vs. $K = 3$, $M_r = 2$ (i.e., BD-based system).

Figure 3.4 is shown as an example. The solid lines and dashed lines correspond to the simulated results and the approximated results, respectively. The achievable throughput of DPC is simulated by applying an algorithm proposed in [JRV⁺05]. Utilizing $C_i = \mathbb{E} \{\log_2(1 + \text{SINR}_i)\}$ for each user with Gaussian inputs, the achievable throughput of ZF and BD are simulated. For the $M_T = K M_r$ case, the average rate offset between BD and ZF increases with the increasing number of receive antennas.

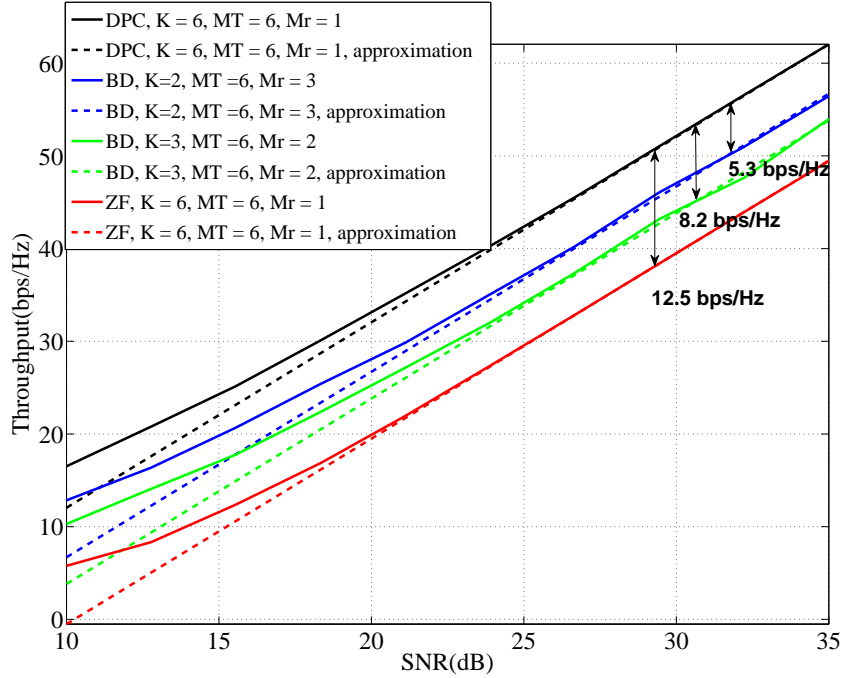


Figure 3.4.: Achievable throughputs approximation for DPC, BD and ZF at high SNRs

3.3.6. Throughput approximation for RBD under the condition $M_T \geq M_R$

In this subsection, we further study the throughput approximation for regularized block diagonalization (RBD) based multi-user MIMO broadcast channels which has been discussed in our publication [SH09a]. The RBD precoding has been proposed to relax the dimensionality

constraint that the number of transmit antennas has to be equal to or larger than the total number of receive antennas, while achieving an improved throughput and an improved diversity order relative to BD [SH08]. Since RBD allows some multi-user interference (MUI), the received signal $\mathbf{y}_i \in \mathbb{C}^{M_{R_i} \times 1}$ for user i is given by

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{F}_i \mathbf{s}_i + \sum_{j=1, j \neq i}^K \mathbf{H}_i \mathbf{F}_j \mathbf{s}_j + \mathbf{n}_i . \quad (3.47)$$

With RBD the channel \mathbf{H}_i is converted to an equivalent channel $\bar{\mathbf{H}}_i = \mathbf{H}_i \mathbf{F}_{a_i}$ which has less overlap with the other users' channels. The achievable throughput of RBD based systems is expressed as

$$C_{\text{RBD}}(\mathbf{H}, \tilde{P}_T) = \max_{\sum_{i=1}^K \text{tr}(\mathbf{Q}_i) \leq P_T} \sum_{i=1}^K \log_2 \left| \mathbf{I}_{M_{R_i}} + \bar{\mathbf{H}}_i \mathbf{Q}_i \bar{\mathbf{H}}_i^H \right| , \quad (3.48)$$

where the covariance matrix \mathbf{Q}_i is set to be $\frac{P_T}{M_R} \mathbf{I}_{M_T}$ due to the fact that the optimal power allocation converges to equal power allocation at high SNRs [LJ07]. Thus, we get

$$C_{\text{RBD}}(\mathbf{H}, \tilde{P}_T) \cong M_R \log_2 \tilde{P}_T - M_R \log_2 M_R + \log_2 \prod_{i=1}^K \left| \bar{\mathbf{H}}_i \bar{\mathbf{H}}_i^H \right| \quad (3.49)$$

and

$$\mathcal{S}_\infty = M_R , \quad (3.50)$$

$$\mathcal{L}_\infty = \log_2 M_R - \frac{1}{\mathcal{S}_\infty} \log_2 \prod_{i=1}^K \left| \bar{\mathbf{H}}_i \bar{\mathbf{H}}_i^H \right| . \quad (3.51)$$

It is found that RBD can maintain the same multiplexing gain as DPC, but has a different power offset.

The power offset of RBD can be further quantified. Note that in [SH08] the expression of \mathbf{F}_{a_i} is derived under the condition $\mathbb{E} \{ \mathbf{s}_i \mathbf{s}_i^H \} = \mathbf{I}_{M_T}$. Taking into account $\mathbb{E} \{ \mathbf{s}_i \mathbf{s}_i^H \} = \frac{P_T}{M_R} \mathbf{I}_{M_T}$ in our throughput approximation, the expressions (22) and (26) in [SH08] can be easily modified as

$$\mathbf{F}_a = \min_{\mathbf{F}_a} \mathbb{E} \left\{ \sum_{i=1}^K \left\| \bar{\mathbf{H}}_i \mathbf{F}_{a_i} \mathbf{s}_i \right\|_F^2 + \frac{\|\mathbf{n}_i\|_F^2}{\beta^2} \right\} \quad \text{and} \quad (3.52)$$

$$\beta^2 = \frac{P_T}{\sum_{i=1}^K \text{tr}(\mathbf{F}_{a_i} \mathbf{s}_i \mathbf{s}_i^H \mathbf{F}_{a_i}^H)} = \frac{M_R}{\sum_{i=1}^K \text{tr}(\mathbf{F}_{a_i} \mathbf{F}_{a_i}^H)} . \quad (3.53)$$

Then using a similar derivation as in [SH08] we get a new expression for \mathbf{F}_{a_i} as

$$\mathbf{F}_{a_i} = \tilde{\mathbf{V}}_i \left(\frac{\tilde{P}_T}{M_R} \tilde{\mathbf{\Sigma}}_i^T \tilde{\mathbf{\Sigma}}_i + \mathbf{I}_{M_T} \right)^{-1/2}. \quad (3.54)$$

With the help of the above expression, the approximation of $|\bar{\mathbf{H}}_i \bar{\mathbf{H}}_i^H|$ can be obtained as follows

$$\begin{aligned} |\bar{\mathbf{H}}_i \bar{\mathbf{H}}_i^H| &= \left| \mathbf{H}_i \tilde{\mathbf{V}}_i \left(\frac{\tilde{P}_T}{M_R} \tilde{\mathbf{\Sigma}}_i^T \tilde{\mathbf{\Sigma}}_i + \mathbf{I}_{M_T} \right)^{-1} \tilde{\mathbf{V}}_i^H \mathbf{H}_i^H \right| \\ &\stackrel{(1)}{=} \left| \mathbf{H}_i \begin{bmatrix} \tilde{\mathbf{V}}_i^{(1)} & \tilde{\mathbf{V}}_i^{(0)} \end{bmatrix} \left(\frac{\tilde{P}_T}{M_R} \tilde{\mathbf{\Sigma}}_i^T \tilde{\mathbf{\Sigma}}_i + \mathbf{I}_{M_T} \right)^{-1} \begin{bmatrix} \tilde{\mathbf{V}}_i^{(1)} & \tilde{\mathbf{V}}_i^{(0)} \end{bmatrix}^H \mathbf{H}_i^H \right| \\ &\stackrel{(2)}{=} \left| \mathbf{H}_i \tilde{\mathbf{V}}_i^{(1)} \left(\frac{\tilde{P}_T}{M_R} \tilde{\mathbf{\Lambda}}_i^2 + \mathbf{I}_{(M_R-M_r)} \right)^{-1} \tilde{\mathbf{V}}_i^{(1)H} \mathbf{H}_i^H + \mathbf{H}_i \tilde{\mathbf{V}}_i^{(0)} \tilde{\mathbf{V}}_i^{(0)H} \mathbf{H}_i^H \right| \\ &\stackrel{(3)}{\approx} \left| \mathbf{H}_i \tilde{\mathbf{V}}_i^{(1)} \left(\frac{\tilde{P}_T}{M_R} \tilde{\mathbf{\Lambda}}_i^2 \right)^{-1} \tilde{\mathbf{V}}_i^{(1)H} \mathbf{H}_i^H + \mathbf{H}_i \tilde{\mathbf{V}}_i^{(0)} \tilde{\mathbf{V}}_i^{(0)H} \mathbf{H}_i^H \right| \\ &\stackrel{(4)}{=} \left| \frac{M_R}{\tilde{P}_T} \mathbf{G}_i^{(1)} \mathbf{G}_i^{(1)H} + \mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H} \right| \\ &\stackrel{(5)}{\approx} \left| \mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H} \right| \left(1 + \underbrace{\frac{M_R}{\tilde{P}_T} \text{tr} \left[(\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H})^{-1} \mathbf{G}_i^{(1)} \mathbf{G}_i^{(1)H} \right]}_{\mu_i} \right) \\ &= \left| \mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H} \right| \left(1 + \frac{\mu_i M_R}{\tilde{P}_T} \right), \end{aligned} \quad (3.55)$$

where the definitions of $\tilde{\mathbf{V}}_i$ and $\tilde{\mathbf{\Sigma}}_i$ can be found in equation 3.10. At step (1) $\tilde{\mathbf{V}}_i$ is separated into $\tilde{\mathbf{V}}_i^{(1)} \in \mathbb{C}^{M_T \times (M_R - M_r)}$ and $\tilde{\mathbf{V}}_i^{(0)} \in \mathbb{C}^{M_T \times (M_T - M_R + M_r)}$, which refer to the right singular vectors corresponding to non-zero singular values and the right singular vectors corresponding to zero singular values, respectively. At step (2) we use $\tilde{\mathbf{\Lambda}}_i \in \mathbb{C}^{(M_R - M_r) \times (M_R - M_r)}$ to represent the economy-size version of $\tilde{\mathbf{\Sigma}}_i$. For large P_T , we neglect $\mathbf{I}_{(M_R - M_r)}$ and reach step (3). Then we replace $\mathbf{H}_i \tilde{\mathbf{V}}_i^{(1)} \tilde{\mathbf{\Lambda}}_i^{-1}$ and $\mathbf{H}_i \tilde{\mathbf{V}}_i^{(0)}$ by $\mathbf{G}_i^{(1)}$ and $\mathbf{G}_i^{(0)}$, respectively. At step (4) the following property of matrix determinants is utilized

$$\det(\mathbf{A} + \epsilon \mathbf{X}) = \det(\mathbf{A}) \left(1 + \text{tr}(\mathbf{A}^{-1} \mathbf{X}) \epsilon \right), \quad (3.56)$$

where \mathbf{A} and \mathbf{X} are square matrices and ϵ is very small number, which leads to step (5). Due to the fact that

$$0 \leq \text{tr}(\mathbf{A}\mathbf{B})^n \leq \text{tr}(\mathbf{A})^n \text{tr}(\mathbf{B})^n \quad (3.57)$$

if \mathbf{A} and \mathbf{B} are positive semi-definite matrices of the same order, the μ_i defined in equation

(3.55) can be bounded as

$$0 \leq \mu_i \leq \text{tr} \left((\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H})^{-1} \right) \text{tr} \left(\mathbf{G}_i^{(1)} \mathbf{G}_i^{(1)H} \right). \quad (3.58)$$

Substituting equation (3.55) into equation (3.51), \mathcal{L}_∞ can be rewritten as

$$\begin{aligned} \mathcal{L}_\infty = \log_2 M_R - \frac{1}{M_R} \log_2 \prod_{i=1}^K \left| \mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H} \right| \\ - \frac{1}{M_R} \log_2 \prod_{i=1}^K \left(1 + \frac{\mu_i M_R}{\tilde{P}_T} \right). \end{aligned} \quad (3.59)$$

As a result, the throughput approximation of RBD at high SNRs can be further expressed as

$$C_{\text{RBD}}(\mathbf{H}, \tilde{P}_T) \cong M_R \log_2 \tilde{P}_T - M_R \log_2 M_R + \log_2 \prod_{i=1}^K \left| \mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H} \right| + \log_2 \prod_{i=1}^K \left(1 + \frac{\mu_i M_R}{\tilde{P}_T} \right). \quad (3.60)$$

3.3.6.1. RBD vs. BD

The rate loss between RBD and BD at high SNRs for one channel realization is defined as

$$\Delta_{\text{RBD-BD}} = \lim_{\tilde{P}_T \rightarrow \infty} [C_{\text{RBD}}(\tilde{P}_T) - C_{\text{BD}}(\tilde{P}_T)]. \quad (3.61)$$

Averaging over the fading distribution, the average rate offset is calculated as

$$\bar{\Delta}_{\text{RBD-BD}} = \mathbb{E} \{ \Delta_{\text{RBD-BD}} \}. \quad (3.62)$$

Theorem 3.3.4. *The average rate offset between RBD and BD in the i.i.d. Rayleigh fading at high SNRs is upper bounded by*

$$\bar{\Delta}_{\text{RBD-BD}} \leq K \log_2 \left(1 + \frac{\mu M_R}{\tilde{P}_T} \right), \quad (3.63)$$

where $\mu = \mathbb{E} \{ \mu_i \}$, $i = 1, \dots, K$, and

$$\begin{cases} 0 \leq \mu \leq \frac{M_r^2(M_R - M_r)}{M_T - M_R} & , \text{ for } M_T > M_R \\ 0 \leq \mu \leq \frac{M_r}{\xi} - \frac{1}{\xi(2M_r - \xi)} + 1 & , \text{ for } M_T = M_R \end{cases} \quad (3.64)$$

Here, ξ is a small positive number which should be equal to the average of the largest eigenvalue of $\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H}$ in Rayleigh fading.

Proof: see C.3.

3.3.6.2. DPC vs. RBD

The rate loss relative to DPC for one channel realization is also defined as

$$\Delta_{\text{DPC-RBD}} = \lim_{\tilde{P}_T \rightarrow \infty} [C_{\text{DPC}}(\tilde{P}_T) - C_{\text{RBD}}(\tilde{P}_T)]. \quad (3.65)$$

Averaging over the fading distribution, the average rate offset is calculated as

$$\bar{\Delta}_{\text{DPC-RBD}} = \mathbb{E} \{ \Delta_{\text{DPC-RBD}} \}. \quad (3.66)$$

Theorem 3.3.5. *The average rate offset between DPC and RBD in the i.i.d Rayleigh fading at high SNRs is lower bounded by*

$$\bar{\Delta}_{\text{DPC-RBD}} \geq \log_2 e \sum_{m=0}^{M_R-1} \varphi(M_T - m) - K \log_2 \left(1 + \frac{\mu M_R}{\tilde{P}_T} \right) - K \cdot \log_2 e \sum_{n=0}^{M_r-1} \varphi(M_T - M_R + M_r - n), \quad (3.67)$$

where $\varphi(\cdot)$ denotes the digamma function.

Proof: see C.4.

Using a similar expression as in equation (3.42), the average power offset between DPC and RBD is calculated by

$$\bar{P}_{\text{DPC-RBD}} = \frac{3\bar{\Delta}_{\text{DPC-RBD}}}{\mathcal{S}_\infty} = \frac{3\bar{\Delta}_{\text{DPC-RBD}}}{M_R}. \quad (3.68)$$

In Figures 3.5, 3.6, and 3.7, we show the simulation results and approximation results for RBD. We assume a quasi-static i.i.d. block Rayleigh fading channel. In each block, the channel is assumed to be constant. The parameter ξ only depends on the equivalent channel $\mathbf{G}_i^{(0)}$ of each user. For instance, ξ is calculated as 3.53 by numerical experiments for the case $M_T = 6$ and $M_r = 2$. Then, the parameter μ is bounded by $0 \leq \mu \leq 0.96$. Consequently, the upper bound of the average rate offset between RBD and BD can be calculated according to the SNRs. For 30 dB, this upper bound is calculated as 0.026 bps/Hz for $\mu = 0.96$ as an example. It implies that RBD achieves almost the same throughput as BD at high SNRs. Additionally, μ can be arbitrarily chosen within its range (e.g., $0 \leq \mu \leq 0.96$). In Figures 3.5 and 3.6 we choose $\mu = 0.25$ and $\mu = 0.96$, respectively. There is no obvious difference between these two cases, since μ is really small relative to the observed high SNRs.

3.3.7. Throughput approximation for RBD under the condition $M_T < M_R$

Under the condition $M_T < M_R$ each user's precoding matrix cannot lie in the null space of all other users' channels anymore, since there is no null space left. As a result, ZF and BD cannot

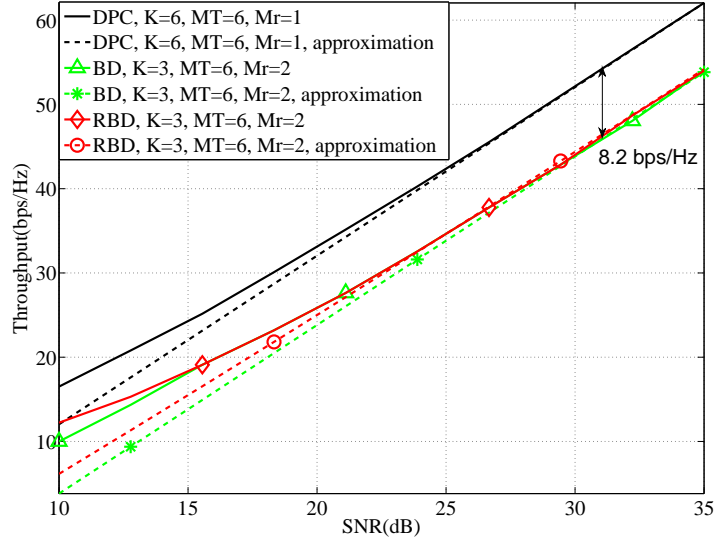


Figure 3.5.: Achievable throughputs approximation for RBD at high SNRs with $M_T = 6$, $K = 3$, $M_r = 2$ and $\mu = 0.25$

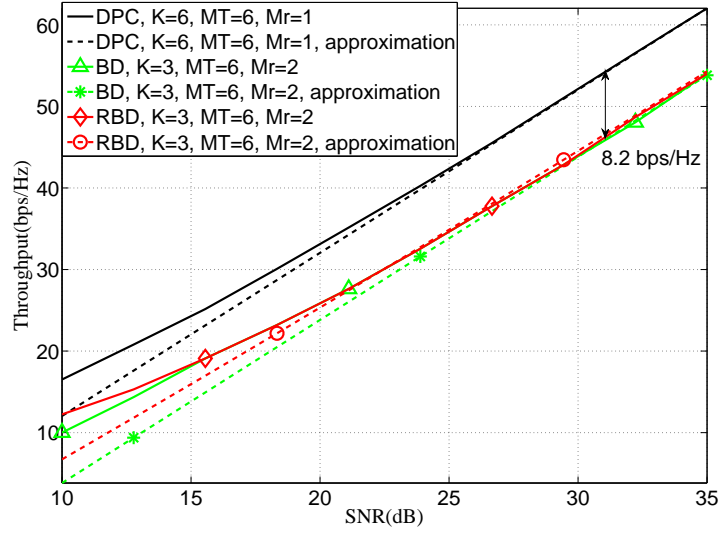


Figure 3.6.: Achievable throughputs approximation for RBD at high SNRs with $M_T = 6$, $K = 3$, $M_r = 2$ and $\mu = 0.96$

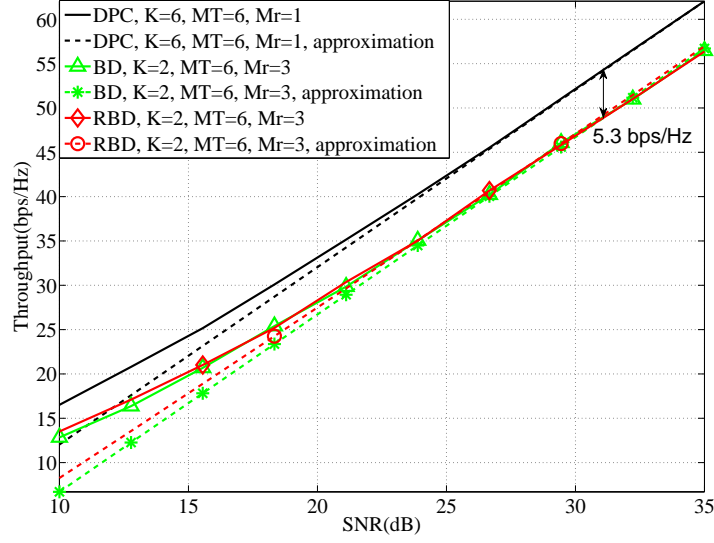


Figure 3.7.: Achievable throughputs approximation for RBD at high SNRs with $M_T = 6$, $K = 2$, $M_r = 3$ and $\mu = 0.25$

be performed in this case. However, the RBD precoding matrix \mathbf{F}_{a_i} ($i = 1, \dots, K$) does not only lie in the null space of all other users' channels (i.e., $\tilde{\mathbf{V}}_i^{(0)}$ in equation (3.10)), but also in the space $\tilde{\mathbf{V}}_i^{(1)}$ in equation (3.10) with a power that is inversely proportional to the singular values of the all other users' channels. In this section we study the achievable throughput of RBD precoding for the case $M_T < M_R$ and approximate this throughput at high SNRs.

First let us use some results derived in Section 3.3.6, but rewrite the expression for $|\bar{\mathbf{H}}_i \bar{\mathbf{H}}_i^H|$ as

$$\begin{aligned}
 |\bar{\mathbf{H}}_i \bar{\mathbf{H}}_i^H| &= \left| \mathbf{H}_i \tilde{\mathbf{V}}_i \left(\frac{\tilde{P}_T}{M_R} \tilde{\Sigma}_i^T \tilde{\Sigma}_i + \mathbf{I}_{M_T} \right)^{-1} \tilde{\mathbf{V}}_i^H \mathbf{H}_i^H \right| \\
 &\stackrel{(1)}{\approx} \left| \frac{M_R}{\tilde{P}_T} \mathbf{H}_i \tilde{\mathbf{V}}_i \left(\tilde{\Sigma}_i^T \tilde{\Sigma}_i \right)^{-1} \tilde{\mathbf{V}}_i^H \mathbf{H}_i^H \right| \\
 &\stackrel{(2)}{=} \left| \frac{M_R}{\tilde{P}_T} \mathbf{H}_i \left(\tilde{\mathbf{H}}_i^H \tilde{\mathbf{H}}_i \right)^{-1} \mathbf{H}_i^H \right|.
 \end{aligned} \tag{3.69}$$

In step (1) we neglect \mathbf{I}_{M_T} by considering the high SNR regime. Considering the SVD of the matrix $\tilde{\mathbf{H}}_i \in \mathbb{C}^{(K-1)M_r \times M_T}$ (i.e., $\tilde{\mathbf{H}}_i = \tilde{\mathbf{U}}_i \tilde{\Sigma}_i \tilde{\mathbf{V}}_i^H$), we reach step (2).

Substituting equation (3.69) into equation (3.49), the achievable throughput of RBD at high

SNRs can be approximated as

$$\begin{aligned}
 C_{\text{RBD}}(\mathbf{H}, \tilde{P}_T) &\cong M_R \log_2 \tilde{P}_T - M_R \log_2 M_R + \log_2 \prod_{i=1}^K \left| \frac{M_R}{\tilde{P}_T} \mathbf{H}_i (\tilde{\mathbf{H}}_i^H \tilde{\mathbf{H}}_i)^{-1} \mathbf{H}_i^H \right| \\
 &= M_R \log_2 \tilde{P}_T - M_R \log_2 M_R + K \mathbb{E} \left\{ \log_2 \left| \frac{M_R}{\tilde{P}_T} \mathbf{H}_i (\tilde{\mathbf{H}}_i^H \tilde{\mathbf{H}}_i)^{-1} \mathbf{H}_i^H \right| \right\} \\
 &= M_R \log_2 \tilde{P}_T - M_R \log_2 M_R + K \mathbb{E} \left\{ \log_2 \left(\frac{M_R}{\tilde{P}_T} \right)^{M_r} \cdot |\mathbf{H}_i (\tilde{\mathbf{H}}_i^H \tilde{\mathbf{H}}_i)^{-1} \mathbf{H}_i^H| \right\} \\
 &= M_R \log_2 \tilde{P}_T - M_R \log_2 M_R + K \log_2 \left(\frac{M_R}{\tilde{P}_T} \right)^{M_r} + K \mathbb{E} \left\{ \log_2 |\mathbf{H}_i (\tilde{\mathbf{H}}_i^H \tilde{\mathbf{H}}_i)^{-1} \mathbf{H}_i^H| \right\} \\
 &= K \mathbb{E} \left\{ \log_2 |\mathbf{H}_i (\tilde{\mathbf{H}}_i^H \tilde{\mathbf{H}}_i)^{-1} \mathbf{H}_i^H| \right\} \\
 &\stackrel{(1)}{=} K \log_2 e \sum_{n=0}^{M_r-1} \left(\varphi(M_T - n) - \varphi(M_R - M_T - n) \right) \tag{3.70}
 \end{aligned}$$

At step 1 $\mathbb{E} \left\{ \log_2 |\mathbf{H}_i (\tilde{\mathbf{H}}_i^H \tilde{\mathbf{H}}_i)^{-1} \mathbf{H}_i^H| \right\}$ is rewritten by applying the property of Wishart matrix found in Appendix C.2 equation (C.10).

From equation (3.70) and Figure 3.8 we can see that under the condition $M_T < M_R$, the throughput of RBD stays almost constant for high SNRs, which shows that the benefit of spatial multiplexing is completely lost, i.e., $\mathcal{S}_\infty = 1$. In this case ($M_T < M_R$), we recommend that RBD should be performed only for low or medium SNRs.

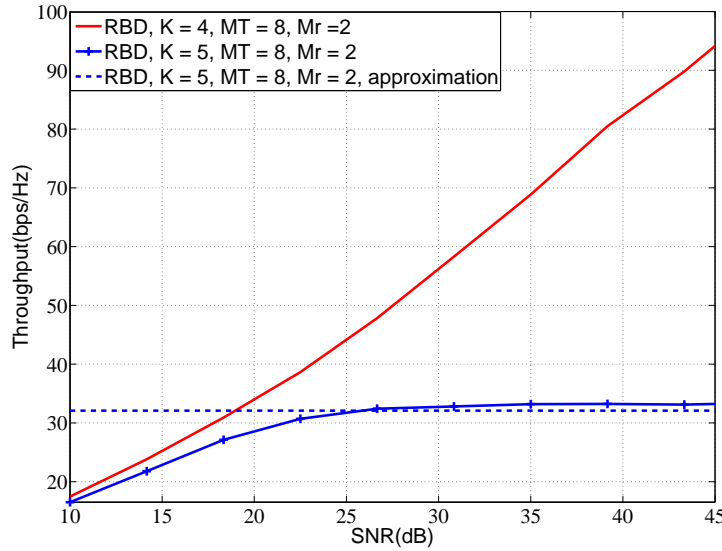


Figure 3.8.: Achievable throughputs approximation for RBD at high SNRs with $M_T < M_R$

3.4. Summary and Conclusions

In this chapter, we have studied the throughput loss between the throughput achieved by DPC and the throughputs achieved with linear precoding strategies (i.e., ZF, BD, and RBD) by utilizing an affine approximation at high SNRs. Under the condition that the aggregate number of receive antennas is not larger than the number of transmit antennas, the average rate and power offsets between DPC, ZF, and BD in a spatially white Rayleigh fading environment have been derived in a closed form which depends on the number of transmit and receive antennas. The average rate and power offsets between DPC, RBD, and BD are bounded by a simple function of the system parameters (e.g., the number of users and receive antennas). Sequentially, we consider the condition that the aggregate number of receive antennas is larger than the number of transmit antennas, where only RBD can be performed. The achievable throughput of RBD at high SNRs is approximated, which shows that the benefit of spatial multiplexing is completely lost. In this case, we suggest to utilize RBD only for the low to medium SNR range. Therefore, we can further conclude that the RBD precoding is not a good choice for the multi-user MIMO downlink with the antenna configuration $M_T < M_R$. Compared to it, coordinated beamforming (CBF) techniques can be viewed as a good solution for this case, which will be discussed in detail in the following chapter .

4. Coordinated beamforming techniques

4.1. Introduction and Motivation

For multi-user MIMO broadcast channels various transmit strategies have been proposed [Cos83, CC00, VJ98, SSH04, SH08, CMIH08a, CMJH08, SH09b, SRH10b, YSRH12, LCZ⁺12]. They offer different trade-offs between the sum rate performance and complexity. As we have shown in Chapter 3, linear precoding techniques represent a promising transmit strategy because of a lower complexity while being able to achieve the same multiplexing gain as DPC. Unfortunately, their applications are constrained by the dimensionality restriction which states that the total number of receive antennas must be smaller than or equal to the number of transmit antennas. This condition is not fulfilled in many scenarios that have been studied recently. For example, the users across cell borders have to be considered jointly by base stations (BSs) for coordinated multi-point (CoMP) transmission [PDF⁺08]. Furthermore, when each user is equipped with multiple antennas, the BS simultaneously serves as many users as possible, which corresponds to a large number of receive antennas. Regularized block diagonalization (RBD) linear precoding has been proposed to relax the dimensionality restriction and has an improved sum rate as well as diversity order compared to BD and ZF. However, it has been shown in Chapter 3 that the performance of RBD degrades heavily with an increasing aggregate number of receive antennas. It is possible to apply and improve ZF, BD, and RBD through eigenmode selection [CHA07] or receive antenna selection [CHM06]. In both cases, though, the transmitter and the receiver are not jointly optimized. Additionally, some signaling techniques are required to indicate the selected eigenmodes or receive antennas.

Coordinated beamforming (CBF) algorithms have been proposed to transmit a number of data streams that is smaller than the total number of receive antennas [SSH04, ZHV08, SH08, CMIH08a, CMJH08]. The methods in [SSH04, ZHV08, SH08, CMIH08a] employ iterative operations to jointly update the transmit-receive beamforming vectors. However, the convergence behavior of the iterations is not considered in [SSH04, ZHV08, SH08]. The number of iterations is set by hand. The coordinated transmission strategy in [CMIH08a] has a lower complexity, and its sum rate performance is closest to the sum capacity of the MIMO broadcast channel compared to the other CBF algorithms [SSH04, ZHV08, SH08]. But, only a single data stream to each user is considered in [CMIH08a] and the receive beamforming strategy is fixed to maximum ratio combining (MRC) matched filtering. Therefore, it is more attractive

if an iterative CBF can support both single and multiple data stream transmissions, achieve a sum rate performance which is not worse than [CMIH08a], and take the convergence behavior into account. Inspired by it, we propose a new iterative CBF named flexible coordinated beamforming (FlexCoBF) [SRH10b] which possesses several advantages compared to the existing iterative CBF algorithms (shown in Section 4.4). FlexCoBF has originally been designed for the multi-user MIMO downlink channel. In our subsequent works, we have extended it to multi-user multi-cell MIMO systems [SRH13a] and multi-carrier multi-user MIMO systems [LCZ⁺12, CLZ⁺13].

Closed-form expressions for CBF have been proposed in [CMJH08] in order to avoid iteratively updating between the transmit and receive beamforming vectors, while achieving the same sum rate performance as the iterative CBF in [CMIH08a]. However, the algorithm proposed in [CMJH08] is restricted to a system with two transmit antennas and two users. The transmit beamformers are designed as the generalized eigenvectors of the channel correlation matrices of the two users when a MRC matched filter is used at each user side. It is an open problem for closed-form CBF where the number of users is greater than 2. Fortunately, we have found a solution to solve this problem. Our new algorithm is named SeDJoCo-based closed-form CBF where the transmit beamforming vectors of an arbitrary number of users are calculated by a sequentially drilled joint congruence (SeDJoCo) transformation (details are found in Section 4.3).

In this chapter, we begin by reviewing the existing coordinated beamforming algorithms in Section 4.2. In Section 4.3, SeDJoCo-based closed-form coordinated beamforming is proposed for multi-user MIMO broadcast channels with an arbitrary number of users. The new iterative coordinated beamforming algorithm (i.e., FlexCoBF) for a single cell and its extension to a cellular scenario are discussed in Section 4.4. Finally, a summary is provided in Section 4.5.

4.2. Previous work on coordinated beamforming techniques

Coordinated beamforming algorithms allow fewer data streams than the number of receive antennas by jointly optimizing the transmit and receive beamformers [SSH04, ZHV08, SH08, CMIH08a, CMJH08]. These approaches perform close to the sum capacity of the MIMO broadcast channel but most of them require an iterative computation for the transmit and receive beamformers. The convergence of these iterative algorithms cannot be guaranteed.

In this section, we consider a multi-user MIMO broadcast channel with K users as shown in Figure 4.1. The i th user has M_{R_i} receive antennas and the base station (BS) is equipped with M_T transmit antennas. The total number of receive antennas is indicated by M_R (i.e.,

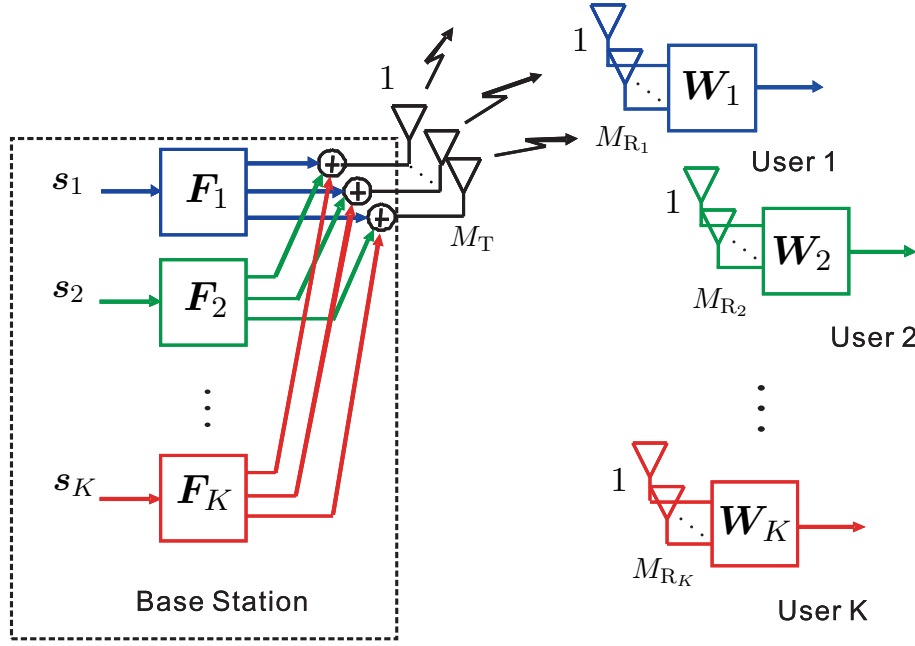


Figure 4.1.: Block diagram of multi-user MIMO BC employing CBF.

$M_R = \sum_{i=1}^K M_{R_i}$) and we have $M_R > M_T$. The channel between the BS and the i th user is denoted by $\mathbf{H}_i \in \mathbb{C}^{M_{R_i} \times M_T}$ which is assumed to be quasi-static block-fading. In general, the received signal of the i th user by applying CBF can be expressed as

$$\mathbf{y}_i = \mathbf{W}_i^H \mathbf{H}_i \mathbf{F}_i \mathbf{s}_i + \mathbf{W}_i^H \mathbf{H}_i \sum_{\ell=1, \ell \neq i}^K \mathbf{F}_\ell \mathbf{s}_\ell + \mathbf{W}_i^H \mathbf{n}_i. \quad (4.1)$$

Let $\mathbf{s}_i \in \mathbb{C}^{r_i}$ denote the transmitted signal for the i th user. The matrices $\mathbf{F}_i \in \mathbb{C}^{M_T \times r_i}$ and $\mathbf{W}_i \in \mathbb{C}^{M_{R_i} \times r_i}$ indicate the transmit and receive beamforming matrix of user i , respectively. The variable r_i denotes the number of data streams to user i . The total number of data streams for all users is indicated as r (i.e., $r = \sum_{i=1}^K r_i < M_R$). Obviously, with the CBF algorithms which can only support one data stream per user such as [CMIH08a], we have $r_i = 1$, the transmit beamforming vector $\mathbf{f}_i \in \mathbb{C}^{M_T}$, the receive beamforming vector $\mathbf{w}_i \in \mathbb{C}^{M_{R_i}}$, and $r = K$.

4.2.1. Iterative coordinated beamforming

The coordinated beamforming algorithm in [SSH04], i.e., coordinated BD, was proposed as a pioneering work. The coordinated BD is an iterative method that uses a reasonable initialization for the receive beamforming matrices (i.e., \mathbf{W}_i , for $i = 1, \dots, K$) followed by the

application of BD. The number of the iteration is set to two. More than one data stream per user is supported by coordinated BD.

The equivalent multi-user channel \mathbf{H}_e is defined as

$$\mathbf{H}_e = \begin{bmatrix} \mathbf{W}_1^H \mathbf{H}_1 \\ \mathbf{W}_2^H \mathbf{H}_2 \\ \vdots \\ \mathbf{W}_K^H \mathbf{H}_K \end{bmatrix} \in \mathbb{C}^{r \times M_T}, \quad (4.2)$$

The matrix \mathbf{H}_e has the dimensions that are compatible with the BD algorithm when $r \leq M_T$. The problem then becomes one of choosing r_i and the beamformers \mathbf{W}_i for each user.

Obviously, the number of data streams r_i allocated to each user should be 1 when $K = M_T$, assuming that all users are to be accommodated. When $K < M_T$ the question is more difficult. In such a case, the BS can still send only one data stream to each user, but with an increased gain, or allocate the additional subchannels to some or all users. If the sum rate is the primary concern, the optimal solution is to give extra subchannels to the stronger users. If power control is the goal, the more beneficial solution is to give the extra subchannels to the users with weaker channels.

When the values of r_i have been determined, the approach to determine the matrix \mathbf{W}_i is to use the r_i dominant left singular vectors of \mathbf{H}_i . The coordinated BD algorithm is summarized as follows

1. For $i = 1, \dots, K$: compute the SVD of \mathbf{H}_i (i.e., $\mathbf{H}_i = \mathbf{U}_i \mathbf{\Sigma}_i \mathbf{V}_i^H$).
2. Determine r_i , which is the number of data streams assigned to each user.
3. For $i = 1, \dots, K$, let \mathbf{W}_i contain the first r_i columns of \mathbf{U}_i . Then, calculate \mathbf{H}_e .
4. Apply the BD algorithm on the matrix \mathbf{H}_e to calculate the precoding matrix \mathbf{F} . Then, the procedure has ended.

Inspired by the coordinated BD, the authors in [CMIH08a] have proposed a low complexity iterative coordinated beamforming algorithm which supports one data stream for each user and uses maximal ratio combining filters as the receive combining vectors. Thus, the i th user's receive combining vector $\mathbf{w}_i \in \mathbb{C}^{M_{R_i} \times 1}$ is given as $\mathbf{w}_i = \mathbf{H}_i \mathbf{f}_i$ where $\mathbf{f}_i \in \mathbb{C}^{M_T \times 1}$ is the transmit beamforming vector of the user i . In this case, the equivalent multi-user channel matrix \mathbf{H}_e is

calculated as

$$\mathbf{H}_e = \begin{bmatrix} \mathbf{w}_1^H \mathbf{H}_1 \\ \mathbf{w}_2^H \mathbf{H}_2 \\ \vdots \\ \mathbf{w}_K^H \mathbf{H}_K \end{bmatrix} \in \mathbb{C}^{K \times M_T}. \quad (4.3)$$

The receive combining vectors are initialized to some random vectors. Then, with increasing iteration index p the following two steps are repeated until a stopping criterion is satisfied.

1. Compute the equivalent multi-user channel matrix $\mathbf{H}_e^{(p)}$.

$$\mathbf{H}_e^{(p)} = \begin{bmatrix} \mathbf{w}_1^{(p-1)H} \mathbf{H}_1 \\ \mathbf{w}_2^{(p-1)H} \mathbf{H}_2 \\ \vdots \\ \mathbf{w}_K^{(p-1)H} \mathbf{H}_K \end{bmatrix}$$

2. Compute the transmit beamforming vectors for all users

$$\mathbf{F}^{(p)} = \mathbf{H}_e^{(p)+},$$

where

$$\mathbf{F}^{(p)} = [\mathbf{f}_1^{(p)}, \dots, \mathbf{f}_K^{(p)}].$$

The notation $(\cdot)^+$ indicates the Moore-Penrose pseudo-inverse. The BS repeats this procedure until the changes in the transmit beamformers are sufficiently small (i.e., $\|\mathbf{f}_i^{(p+1)} - \mathbf{f}_i^{(p)}\| < \epsilon$, where ϵ is an arbitrary small number). It is noted that although the iterative algorithm seems to converge in most cases, this cannot be guaranteed [CMIH08a].

4.2.2. Closed-form coordinated beamforming

Since the convergence of the iterative coordinated beamforming algorithms cannot be guaranteed. The authors in [CMJH08] have proposed a closed-form coordinated beamforming algorithm which supports a system with 2 users and allows one data stream to each user.

We describe this algorithm for the $M_T = 2$ case. There are two users in the system, where each user has at least two receive antennas. Utilizing maximal ratio combining matched filters as the receive beamforming vectors (i.e., $\mathbf{w}_i = \mathbf{H}_i \mathbf{f}_i$), the received signals at user 1 and user 2 are given by

$$y_1 = \mathbf{w}_1^H \mathbf{H}_1 \mathbf{f}_1 s_1 + \mathbf{w}_1^H \mathbf{H}_2 \mathbf{f}_2 s_2 + \mathbf{w}_1^H \mathbf{n}_1, \quad (4.4)$$

$$y_2 = \mathbf{w}_2^H \mathbf{H}_2 \mathbf{f}_2 s_2 + \mathbf{w}_2^H \mathbf{H}_2 \mathbf{f}_1 s_1 + \mathbf{w}_2^H \mathbf{n}_2, \quad (4.5)$$

Let us define the sample correlation matrix of user i by $\mathbf{R}_i = \mathbf{H}_i^H \mathbf{H}_i$. Then, the above equations can be rewritten as

$$y_1 = \mathbf{f}_1^H \mathbf{R}_1 \mathbf{f}_1 s_1 + \mathbf{f}_1^H \mathbf{R}_1 \mathbf{f}_2 s_2 + \mathbf{w}_1^H \mathbf{n}_1, \quad (4.6)$$

$$y_2 = \mathbf{f}_2^H \mathbf{R}_2 \mathbf{f}_2 s_2 + \mathbf{f}_2^H \mathbf{R}_2 \mathbf{f}_1 s_1 + \mathbf{w}_2^H \mathbf{n}_2, \quad (4.7)$$

Theorem 4.2.1. [CMJH08] *If $M_T = 2$, $M_{R_i} \geq 2$ for $i = 1, 2$, and \mathbf{R}_1 and \mathbf{R}_2 are both invertible, then the following claim holds. If the transmit beamforming vectors \mathbf{f}_1 and \mathbf{f}_2 satisfy the zero inter-user interference conditions, i.e.,*

$$\mathbf{f}_2^H \mathbf{R}_1 \mathbf{f}_1 = 0 \quad (4.8)$$

$$\mathbf{f}_2^H \mathbf{R}_2 \mathbf{f}_1 = 0 \quad (4.9)$$

then $\mathbf{f}_1, \mathbf{f}_2$ are generalized eigenvectors of $(\mathbf{R}_1, \mathbf{R}_2)$.

Proof: see Appendix D.1.

Theorem 4.2.1 indicates that for $M_T = 2$, the solutions achieving zero inter-user interference are the generalized eigenvectors of \mathbf{R}_1 and \mathbf{R}_2 . For the case where the BS and each user have more than two antennas and the system supports two users, the transmit and receive beamformers are calculated as follows. First the BS finds all generalized eigenvectors of the sample correlation matrices \mathbf{R}_1 and \mathbf{R}_2 . Let \mathbf{V} be the set of generalized eigenvectors of \mathbf{R}_1 and \mathbf{R}_2 . Then, the eigenvector pair that maximizes the sum rate of the system is chosen as the transmit beamformers, i.e.,

$$\{\mathbf{f}_1, \mathbf{f}_2\} = \arg \max_{\mathbf{v}_j, \mathbf{v}_k \in \{\mathbf{V}\}, \mathbf{v}_j \neq \mathbf{v}_k} \left\{ \log_2 \left(1 + \frac{P}{2\sigma^2} |\mathbf{H}_1 \mathbf{v}_j|^2 \right) + \log_2 \left(1 + \frac{P}{2\sigma^2} |\mathbf{H}_2 \mathbf{v}_k|^2 \right) \right\} \quad (4.10)$$

where P is the total transmit power. In general, if there are n generalized eigenvectors, $n(n-1)$ computations are required to find the transmit beamformers.

4.3. SeDJoCo transformation based closed-form CBF

As it has been shown in Section 4.2, the closed-form coordinated beamforming proposed in [CMJH08] is only valid for a system supporting two users. In this section we design a closed-form coordinated beamforming algorithm for an arbitrary number of users and transmit antennas.

We consider a multi-user MIMO system with a single base station (BS) and K users, where the BS is equipped with M_T transmit antennas and user i has M_{R_i} receive antennas. The aggregate number of receive antennas is denoted by M_R (i.e., $M_R = \sum_{i=1}^K M_{R_i}$). The number of transmit antennas is assumed to be equal to the number of users, but smaller than the aggregate number of receive antennas (i.e., $M_T = K < M_R$). An additional discussion on the parameter setting for M_T and K is found in Section 4.3.5. The propagation channel between the BS and each user is considered as a quasi-static block fading MIMO channel. The matrix $\mathbf{H}_i \in \mathbb{C}^{M_{R_i} \times M_T}$ represents the channel between the BS and the i th user and is constant in each fading block. Let x_i denote the transmit signal for the user i and $\mathbf{f}_i \in \mathbb{C}^{M_T \times 1}$ indicate the unit-norm transmit beamformer of user i . Denoting the receive beamforming vector for user i by $\mathbf{w}_i \in \mathbb{C}^{M_{R_i} \times 1}$, and restricting our attention to one data stream per user, the received signal of the i th user is given by

$$y_i = \mathbf{w}_i^H \mathbf{H}_i \mathbf{f}_i s_i + \mathbf{w}_i^H \mathbf{H}_i \sum_{\substack{\ell=1 \\ \ell \neq i}}^K \mathbf{f}_\ell s_\ell + \mathbf{w}_i^H \mathbf{n}_i, \quad (4.11)$$

where $\mathbf{n}_i \in \mathbb{C}^{M_{R_i} \times 1}$ denotes the additive, zero-mean complex-valued white noise vector present at the i th receiver with unit variance (i.e., $\sigma_n^2 = 1$).

We assume maximum ratio combining (i.e., matched filters) $\mathbf{w}_i = \mathbf{H}_i \mathbf{f}_i$ at the receivers. The coordinated transmission strategies choose the transmit and receive beamforming vectors such that each user experiences zero multi-user interference (MUI). This implies that for the i th user $\mathbf{w}_i^H \mathbf{H}_i \mathbf{f}_\ell = 0$ for all $\ell \neq i$, which is equivalent to $\mathbf{f}_i^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{f}_\ell = 0$ ($\forall \ell \neq i$).

If $\mathbf{F} \in \mathbb{C}^{M_T \times K}$ denotes the combined transmit beamformers for all users and $\mathbf{R}_i \in \mathbb{C}^{M_T \times M_T}$ denotes the sample correlation matrix of the user i , we have

$$\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_K], \quad (4.12)$$

$$\mathbf{R}_i = \mathbf{H}_i^H \mathbf{H}_i, \quad (4.13)$$

and $\mathbf{F}^H \mathbf{R}_i \mathbf{F}$ can be calculated as

$$\mathbf{F}^H \mathbf{R}_i \mathbf{F} = \begin{bmatrix} \mathbf{f}_1^H \mathbf{R}_i \mathbf{f}_1 & \cdots & \mathbf{f}_1^H \mathbf{R}_i \mathbf{f}_i & \cdots & \mathbf{f}_1^H \mathbf{R}_i \mathbf{f}_K \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{f}_i^H \mathbf{R}_i \mathbf{f}_1 & \cdots & \mathbf{f}_i^H \mathbf{R}_i \mathbf{f}_i & \cdots & \mathbf{f}_i^H \mathbf{R}_i \mathbf{f}_K \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{f}_K^H \mathbf{R}_i \mathbf{f}_1 & \cdots & \mathbf{f}_K^H \mathbf{R}_i \mathbf{f}_i & \cdots & \mathbf{f}_K^H \mathbf{R}_i \mathbf{f}_K \end{bmatrix} \in \mathbb{C}^{K \times K}. \quad (4.14)$$

Considering the zero MUI constraint for the i th user, we find that the off-diagonal elements on the i th row and the i th column of $\mathbf{F}^H \mathbf{R}_i \mathbf{F}$ must be zero. This indicates that the combined transmit beamformer \mathbf{F} is the matrix which can jointly minimize the magnitude of the off-diagonal elements on the i th row and the i th column of a set of matrices $\{\mathbf{R}_i\}_{i=1}^K$. Figure 4.2 gives an example for the five users case where we depict the zero MUI constraint. Obviously,

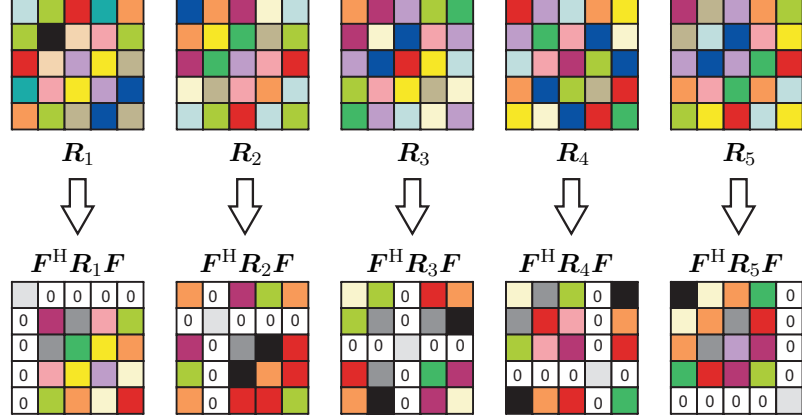


Figure 4.2.: An example for the five users case with zero MUI constraint.

for the two transmit antennas and two users case the matrix \mathbf{F} is just the diagonalizer that can simultaneously diagonalize the matrices \mathbf{R}_1 and \mathbf{R}_2 . In other word, the transmit beamformers \mathbf{f}_1 and \mathbf{f}_2 in the matrix \mathbf{F} are the generalized eigenvectors of $(\mathbf{R}_1, \mathbf{R}_2)$. The generalized eigenvectors can be found by calculating the eigenvalue decomposition of $\mathbf{R}_1^{-1} \mathbf{R}_2$ or $\mathbf{R}_2^{-1} \mathbf{R}_1$. Similarly, the combined transmit beamformer \mathbf{F} for more than two users can be found by designing a particular transformation which we have first proposed in [YSRH12].

Regarding the problems of blind source separation (BSS) and independent component analysis (ICA), approximate joint diagonalization (AJD) algorithms are widely used as a generic tool to solve them. The general framework of AJD considers a set of N (typically more than two) square, symmetric, real-valued $n \times n$ matrices denoted as $\{\mathbf{C}_i\}_{i=1}^N$. The goal of AJD is to find a single matrix \mathbf{B} (or its inverse \mathbf{A}) which best jointly diagonalizes the target matrices in some sense. Inspired by AJD, in [YSRH12] we have derived a “sequentially drilled” joint congruence (SeDJoCo) transformation which can be applied to calculate the matrix \mathbf{F} directly. The SeDJoCo transformation is a particular form of the classical approximate joint diagonalization (AJD) problem. However, unlike the problem of general AJD, the basic form of SeDJoCo considers exactly $N = n$ target matrices $\mathbf{C}_1, \dots, \mathbf{C}_N$ (namely, the number of matrices is equal to their dimension, the relationship between N and the number of users K discussed in Section 4.3.5), and seeks a matrix \mathbf{B} , such that the i th row and the i th column of

the transformed i th target matrix $\mathbf{B}\mathbf{C}_i\mathbf{B}^H$ would be all-zeros, except for the diagonal (i,i) th element. For each matrix \mathbf{C}_i , this structure resembles a square that has been “drilled” along the i th row and column (considering the elements which have been zeroed-out as “empty”). Since the index of the “drilled” row and column progresses sequentially with the matrix-index, we call this congruence transformation “sequentially-drilled” - hence the term SeDJoCo.

4.3.1. Motivation in ML Blind or Semi-Blind Source Separation

Seeking the SeDJoCo transformation is motivated by the closed-form coordinated beamforming as we have shown above. In fact, another important motivation is encountered in the context of maximum likelihood (ML) blind (or semi-blind) source separation which we give a short introduction in the sequel.

Consider the problem of blind (or semi-blind) source separation, in which N statistically independent, zero-mean wide-sense stationary (and real-valued) source signals $\mathbf{s}[t] \triangleq [s_1[t], \dots, s_N[t]]^T$ (with different spectra) are mixed by an unknown, square invertible (real-valued) mixing-matrix \mathbf{A} , yielding the N mixture signals $\mathbf{x}[t] \triangleq [x_1[t], \dots, x_N[t]]^T$,

$$\mathbf{x}[t] = \mathbf{A}\mathbf{s}[t] \quad , \quad t = 1, 2, \dots, T. \quad (4.15)$$

When the power spectral densities (PSDs) of the sources $h_1(\nu), \dots, h_N(\nu)$ (resp.) are known, the scenario is called “semi-blind”. When the PSDs are unknown, the scenario is “fully blind” (see, e.g., [Yer10]). In either case, consider some presumed PSDs (either the true PSDs in a semi-blind scenario or some “educated guess” in a fully blind scenario) $\hat{h}_1(\nu), \dots, \hat{h}_N(\nu)$, and denote by $\hat{\phi}_n[t]$ the Inverse Discrete-Time Fourier Transform (IDTFT, more details are found in Appendix D.2) of $\hat{h}_n^{-1}(\nu)$, namely

$$\hat{\phi}_n[t] \triangleq \int_{-1/2}^{1/2} \frac{1}{\hat{h}_n(\nu)} \cdot e^{j2\pi\nu t} d\nu \quad , \quad n \in \{1, 2, \dots, N\}. \quad (4.16)$$

It is shown in [PG97] (see also [CJ10], Ch.7) that for ML (in the semi-blind scenario, assuming Gaussian sources) or Quasi-ML (QML) (in the fully-blind scenario) separation, the likelihood equations (often also called “estimating equations” in this context) for estimation of \mathbf{A} from $\mathbf{x}[1], \dots, \mathbf{x}[T]$ take the form

$$\sum_{\tau=1-T}^{T-1} \hat{\phi}_n[\tau] \mathbf{e}_m^T \hat{\mathbf{A}}^{-1} \hat{\mathbf{R}}[\tau] \hat{\mathbf{A}}^{-T} \mathbf{e}_n = 0 \quad \forall m \neq n, \quad m, n \in \{1, 2, \dots, N\}, \quad (4.17)$$

where the pinning vector \mathbf{e}_n denotes the n -th column of the $N \times N$ identity matrix \mathbf{I}_N , and

where $\widehat{\mathbf{R}}[\tau]$ denotes the observations' empirical (biased) correlation matrix estimate at lag τ ,

$$\widehat{\mathbf{R}}[\tau] \triangleq \frac{1}{T} \sum_{t=\max(1, 1-\tau)}^{\min(T, T-\tau)} \mathbf{x}[t] \mathbf{x}^T[t + \tau]. \quad (4.18)$$

Thus, defining the set of N matrices

$$\mathbf{Q}_n \triangleq \sum_{\tau=1-T}^{T-1} \hat{\phi}_n[\tau] \widehat{\mathbf{R}}[\tau], \quad n = 1, 2, \dots, N \quad (4.19)$$

and denoting $\widehat{\mathbf{B}} \triangleq \widehat{\mathbf{A}}^{-1}$ (the ML or QML estimate of the demixing matrix), we observe that the likelihood equations (4.17) can also take the form

$$\mathbf{e}_m^T (\widehat{\mathbf{B}} \mathbf{Q}_n \widehat{\mathbf{B}}^T) \mathbf{e}_n = 0 \quad \forall m \neq n, \quad m, n \in \{1, 2, \dots, N\}, \quad (4.20)$$

which implies that for each $n \in \{1, 2, \dots, N\}$, all off-diagonal elements in the n -th column of the transformed matrix $\widehat{\mathbf{B}} \mathbf{Q}_n \widehat{\mathbf{B}}^T$ should be zeros. It is straightforward to show that a “symmetrized” version of $\widehat{\mathbf{R}}[\tau]$ (a result of averaging $\widehat{\mathbf{R}}[\tau]$ with $\widehat{\mathbf{R}}^T[\tau]$) can also be used in (4.17), in which case the resulting matrices \mathbf{Q}_n would also be symmetric, and the form (4.20) would imply that all off-diagonal elements in both the n -th column and n -th row of $\widehat{\mathbf{B}} \mathbf{Q}_n \widehat{\mathbf{B}}^T$ must be all-zeros (for each $n \in \{1, 2, \dots, N\}$).

It is also shown in [PG97, CJ10] that an additional likelihood equation (related to the scaling of the reconstructed sources) requires that the expressions in (4.17) and (4.20) equal 1 for $n = m$. Consequently, the respective diagonal (n, n) -th element of $\widehat{\mathbf{B}} \mathbf{Q}_n \widehat{\mathbf{B}}^T$ should equal 1 as well - but this is merely a scaling condition, which may be substituted with other scaling constraints if desired. For example, this scaling constraint is used in the context of BSS, but is not applicable in other contexts, such as our CBF application. Note that all other elements (in columns and rows other than the n -th) of $\widehat{\mathbf{B}} \mathbf{Q}_n \widehat{\mathbf{B}}^T$ are irrelevant to the ML (or QML) solution, namely, the resulting structure of each $\widehat{\mathbf{B}} \mathbf{Q}_n \widehat{\mathbf{B}}^T$ may generally be far from diagonality, as long as its n -th row and column are *exactly* of the form expected in a diagonal matrix.

A similar form of estimating equations is encountered in a somewhat more specific context of Gaussian Auto-Regressive (AR) sources in [vdV01] (see also [CJ10], Ch.7), and in a more general context (of Gaussian source signals which are not necessarily stationary, but have general temporal-covariance patterns) in [Yer10]. The complex-valued version would also be encountered in these contexts (with complex-valued sources), but only when all signals in question are *circular* complex-valued random processes [Yer12].

General AJD is basically an *ad-hoc* tool which attempts to “best fit” a prescribed model to

the set of target matrices, with no claim of optimality in any significant sense. As shown in [Yer00, TY09], in some particular cases general AJD can be made asymptotically optimal by the introduction of proper weighting. However, the same asymptotic optimality appears in a much more “natural” and computationally simpler way in SeDJoCo (with particular choices of target-matrices), since SeDJoCo can directly attain the ML estimate of \mathbf{A} or \mathbf{B} in such cases. In fact, following [Yer10], it can be concluded that (asymptotically) optimal separation of independent Gaussian sources with any kind of time/frequency diversity (whether stationary, non-stationary, partly stationary and partly non-stationary, etc.) can always be attained via the solution of a SeDJoCo problem.

4.3.2. Equivalent formulations and existence of a solution

The SeDJoCo problem formulation can take several alternative, equivalent forms, each shedding a somewhat different light on the basic aspects of this problem. The three alternative formulations presented below apply both to the real-valued and complex-valued cases.

As already mentioned, in the basic SeDJoCo formulation the number of matrices N must equal the matrices’ dimensions, namely $N = n$. Thus, consider N symmetric (in the real-valued case) or Hermitian symmetric (in the complex-valued case) target-matrices $\mathbf{C}_1, \dots, \mathbf{C}_N$, each of dimensions $N \times N$. The SeDJoCo problem can be stated as:

Proposition 4.3.1. *Given N target-matrices $\mathbf{C}_1, \dots, \mathbf{C}_N$, find an $N \times N$ matrix $\mathbf{B} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_N]^H$, such that*

$$\mathbf{b}_j^H \mathbf{C}_i \mathbf{b}_i = \delta_{ji} \quad \forall j, i \in \{1, 2, \dots, N\}, \quad (4.21)$$

where δ_{ji} denotes Kronecker’s delta function (which is 1 if $j = i$ and 0 otherwise).

Equivalently, the same problem can be stated as:

Proposition 4.3.2. *Given N target-matrices $\mathbf{C}_1, \dots, \mathbf{C}_N$, find an $N \times N$ matrix \mathbf{B} , such that*

$$\mathbf{B} \mathbf{C}_i \mathbf{B}^H \mathbf{e}_i = \mathbf{e}_i \quad \forall i \in \{1, 2, \dots, N\}. \quad (4.22)$$

In other words, each transformed matrix $\mathbf{B} \mathbf{C}_i \mathbf{B}^H$ should be exactly “diagonal” in its i th column (and, since it is symmetric/Hermitian, also in its i th row), in the sense that all off-diagonal elements in these row and column must be exactly zero. All other elements may take arbitrary (nonzero) values. In addition, with the problem formulations above we also require that the diagonal (i, i) th element of $\mathbf{B} \mathbf{C}_i \mathbf{B}^H$ be 1 - but this is merely a scaling constraint on the rows of \mathbf{B} - once any matrix \mathbf{B} satisfying the exact off-diagonal zero constraint is found,

it is straightforward to simply rescale each of its rows such that $\mathbf{b}_i^H \mathbf{C}_i \mathbf{b}_i = 1$, without any effect on the “ i -wise diagonality” property. As we shall see in the sequel, this scaling constraint is only used in one of the SeDJoCo solutions.

Multiplying both sides of equation (4.22) by $\mathbf{A} = \mathbf{B}^{-1}$ on the left we obtain

$$\mathbf{C}_i \mathbf{B}^H \mathbf{e}_i = \mathbf{A} \mathbf{e}_i \Rightarrow \mathbf{a}_i = \mathbf{C}_i \mathbf{b}_i \quad \forall i \in \{1, 2, \dots, N\}, \quad (4.23)$$

where \mathbf{a}_i denotes the i th column of $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_N]$. In other words, the same problem can be stated as follows:

Proposition 4.3.3. *Given N target-matrices $\mathbf{C}_1, \dots, \mathbf{C}_N$, find two reciprocal $N \times N$ matrices \mathbf{B} and $\mathbf{A} = \mathbf{B}^{-1}$, such that the i th column of \mathbf{A} is given by $\mathbf{C}_i \mathbf{b}_i$, with \mathbf{b}_i^H denoting the i th row of \mathbf{B} , $i = 1, 2, \dots, N$.*

Assuming that all target-matrices are invertible, we may also swap the roles between \mathbf{B} and \mathbf{A} , obtaining that the i th column \mathbf{b}_i of \mathbf{B}^H should be given by $\mathbf{C}_i^{-1} \mathbf{a}_i$, where \mathbf{a}_i^H denotes the i th row of \mathbf{A}^H , $i = 1, 2, \dots, N$. This means that the same problem formulations Proposition 4.3.1 and Proposition 4.3.2 above may be cast in terms of \mathbf{A}^H (instead of \mathbf{B}) with the inverses of the target-matrices substituting the target matrices. This implies that the “direct” and “indirect” formulations of SeDJoCo coincide: If \mathbf{B} is the SeDJoCo diagonalizer of $\mathbf{C}_1, \dots, \mathbf{C}_N$, then its (conjugate) transposed inverse \mathbf{A}^H is the SeDJoCo diagonalizer of the inverse set $\mathbf{C}_1^{-1}, \dots, \mathbf{C}_N^{-1}$. It is important to note that this desirable “self-reciprocity” property, is generally *not* shared by other non-orthogonal AJD algorithms. In fact, it is easy to show that this property is satisfied in non-orthogonal AJD when (and only when) the target matrices are *exactly* jointly diagonalizable. In general, however, the target matrices are not exactly jointly diagonalizable.

Obviously, for real-valued and complex-valued matrices cases, SeDJoCo requires the solution of N^2 real-valued equation in N^2 real-valued unknowns and $(2N)^2$ real-valued equations in $(2N)^2$ real-valued unknowns, respectively. Since these equations are nonlinear, the real-valued solutions may not definitely exist and be unique. However, we will show that a solution must exist if all the N target-matrices are positive definite (PD), but there is not an explicit condition for uniqueness.

Let us consider the real-valued case first. Let $\mathbf{C}_1, \dots, \mathbf{C}_N$ denote a set of symmetric, real-valued PD target matrices, and let $\lambda_i > 0$ denote the smallest eigenvalue of \mathbf{C}_i , $i = 1, \dots, N$. Consider the likelihood function for estimation of \mathbf{B} from \mathbf{C}_i which has been derived in [Yer10] on equation (24)

$$C(\mathbf{B}) \triangleq \log |\det \mathbf{B}| - \frac{1}{2} \sum_{i=1}^N \mathbf{e}_i^T \mathbf{B} \mathbf{C}_i \mathbf{B}^T \mathbf{e}_i. \quad (4.24)$$

For all nonsingular \mathbf{B} , $C(\mathbf{B})$ is obviously a continuous and differentiable function of all elements of \mathbf{B} . In addition, $C(\mathbf{B})$ is bounded from above:

$$\begin{aligned}
 C(\mathbf{B}) &= \log |\det \mathbf{B}| - \frac{1}{2} \sum_{i=1}^N \mathbf{b}_i^T \mathbf{C}_i \mathbf{b}_i \\
 &\leq \log \prod_{i=1}^N \|\mathbf{b}_i\| - \frac{1}{2} \sum_{i=1}^N \lambda_i \mathbf{b}_i^T \mathbf{b}_i \\
 &= \frac{1}{2} \sum_{i=1}^N \{ \log \|\mathbf{b}_i\|^2 - \lambda_i \|\mathbf{b}_i\|^2 \} \\
 &\leq \frac{1}{2} \sum_{i=1}^N \{ -\log \lambda_i - 1 \},
 \end{aligned} \tag{4.25}$$

where $\|\mathbf{b}_i\|^2 \triangleq \mathbf{b}_i^T \mathbf{b}_i$ denotes the squared 2 norm of \mathbf{b}_i , and where we have used the properties

1. $|\det \mathbf{B}| \leq \prod_{i=1}^N \|\mathbf{b}_i\|$ (Hadamard's inequality);
2. $\mathbf{b}_i^T \mathbf{C}_i \mathbf{b}_i \geq \lambda_i \|\mathbf{b}_i\|^2$; and
3. $\log x - \lambda x \leq -\log \lambda - 1$ for all $x > 0$.

Note also that $C(\mathbf{B})$ tends to $-\infty$ when \mathbf{B} approaches any singular matrix, and $C(\mathbf{B})$ has additionally the property

$$C(\alpha \cdot \mathbf{B}) \xrightarrow{\alpha \rightarrow \infty} -\infty \quad \forall \mathbf{B} \tag{4.26}$$

Consequently, $C(\mathbf{B})$ must attain a maximum for some nonsingular \mathbf{B} . Being a smooth function of \mathbf{B} for all nonsingular \mathbf{B} , its derivative with respect to (w.r.t.) \mathbf{B} at the maximum point must vanish.

Indeed, differentiating $C(\mathbf{B})$ w.r.t. $B_{(i,j)}$ (the (i,j) -th element of \mathbf{B}) and equating it to zero we get (for all $i, j \in \{1, 2, \dots, N\}$)

$$\begin{aligned}
 \frac{\partial C(\mathbf{B})}{\partial B_{(i,j)}} &= A_{(j,i)} - \frac{1}{2} \sum_{k=1}^N 2 \mathbf{e}_k^T \mathbf{E}_{ij} \mathbf{C}_k \mathbf{B}^T \mathbf{e}_k \\
 &= A_{(j,i)} - \sum_{k=1}^N \delta_{ki} \mathbf{e}_j^T \mathbf{C}_k \mathbf{B}^T \mathbf{e}_k \\
 &= A_{(j,i)} - \mathbf{e}_j^T \mathbf{C}_i \mathbf{b}_i = 0
 \end{aligned} \tag{4.27}$$

where we have used the relation

$$\frac{\partial \log |\det \mathbf{B}|}{\partial B_{(i,j)}} = A_{(j,i)}, \tag{4.28}$$

since $\frac{\partial \log |\det \mathbf{X}|}{\partial \mathbf{X}} = (\mathbf{X}^{-1})^T = (\mathbf{X}^T)^{-1}$ for a square matrix \mathbf{X} [Dyr04].

The first equality of 4.27 holds for all nonsingular real-valued \mathbf{B} , and $\mathbf{E}_{ij} \triangleq \mathbf{e}_i \mathbf{e}_j^T$ denotes an all-zeros matrix with an only 1 at the (i, j) -th location. By concatenating these equations for $j = 1, 2, \dots, N$ into a vector we get $\mathbf{a}_i = \mathbf{C}_i \mathbf{b}_i$, which has to be satisfied for each $i = 1, 2, \dots, N$. This means that the solution of SeDJoCo can be expressed as the maximizer of $C(\mathbf{B})$, which always exists when the target matrices are all PD.

Naturally, this derivation is closely related to the fact that SeDJoCo yields the ML (or QML) estimate of the demixing matrix in some specific BSS contexts (e.g., [PG97, DZ04]) with some specific target-matrices. However, we obtained here a more general result, which holds for *any* set of PD target matrices, and not only for the specific matrices used for ML or QML estimation in [PG97, DZ04].

We now consider the complex-valued case. The main formal difficulty in applying the same proof to the complex-valued case stems from the fact that $C(\mathbf{B})$ as defined above would be a real-valued function of a complex-valued matrix, and as such would not be differentiable w.r.t. \mathbf{B} . To mitigate this difficulty, we take the well-known approach of Brandwood [Bra83] (or van den Bos [vdB94]), reformulating $C(\mathbf{B})$ as $\hat{C}(\mathbf{B}, \mathbf{B}^*)$, such that \mathbf{B} and \mathbf{B}^* are considered independent variables. The “*complex-gradient*” w.r.t. \mathbf{B} is then defined as the partial derivative of $\hat{C}(\mathbf{B}, \mathbf{B}^*)$ w.r.t. \mathbf{B} , considering \mathbf{B}^* to be constant (and this gradient equals the complex-conjugate of the similarly-defined complex-gradient w.r.t. \mathbf{B}^*). At a maximum point, the complex-gradients of $\hat{C}(\mathbf{B}, \mathbf{B}^*)$ w.r.t. both \mathbf{B} and \mathbf{B}^* must vanish.

Indeed, define

$$\hat{C}(\mathbf{B}, \mathbf{B}^*) \triangleq \log \det \mathbf{B} + \log \det \mathbf{B}^* - \sum_{i=1}^N \mathbf{e}_i^T \mathbf{B} \mathbf{C}_i (\mathbf{B}^*)^T \mathbf{e}_i, \quad (4.29)$$

and assume that the target-matrices \mathbf{C}_i are all Hermitian and PD, denoting the smallest eigenvalue of \mathbf{C}_i as $\lambda_i > 0$ ($i = 1, \dots, N$). Using the complex-valued version of the same arguments used above in support of equation (4.25), we have

$$\hat{C}(\mathbf{B}, \mathbf{B}^*) \leq \sum_{i=1}^N \{-\log \lambda_i - 1\}. \quad (4.30)$$

Here, the upper bound of $\hat{C}(\mathbf{B}, \mathbf{B}^*)$ is given. This upper bound can be achieved by some nonsingular \mathbf{B} , such that its complex-gradient w.r.t. \mathbf{B} (and to \mathbf{B}^*) at the maximum point must vanish.

Differentiating w.r.t. $B_{(i,j)}$ we obtain

$$\begin{aligned}
 \frac{\partial \hat{C}(\mathbf{B}, \mathbf{B}^*)}{\partial B_{(i,j)}} &= A_{(j,i)} - 0 - \sum_{k=1}^N \mathbf{e}_k^T \mathbf{E}_{ij} \mathbf{C}_k(\mathbf{B}^*)^T \mathbf{e}_k \\
 &= A_{(j,i)} - 0 - \sum_{k=1}^N \delta_{ki} \mathbf{e}_j^T \mathbf{C}_k(\mathbf{B}^*)^T \mathbf{e}_k \\
 &= A_{(j,i)} - \mathbf{e}_j^T \mathbf{C}_i \mathbf{b}_i = 0
 \end{aligned} \tag{4.31}$$

where \mathbf{b}_i is the i th column of \mathbf{B}^* , namely \mathbf{b}_i^H is the i th row of \mathbf{B} . Differentiation w.r.t. \mathbf{B}^* would simply yield the complex-conjugate version of the same equation. Once again, by concatenating these equations for $j = 1, 2, \dots, N$ into a vector we get $\mathbf{a}_i = \mathbf{C}_i \mathbf{b}_i$.

We have shown the existence of the solution, but not the uniqueness. In fact, with arbitrary (positive-definite) target-matrices the SeDJoCo solution might not be unique.

4.3.3. Solutions of SeDJoCo

Unlike classical AJD, the SeDJoCo problem and its solutions have rarely been addressed in the literature. To the best of our knowledge, with the exception of our journal paper [YSRH12] and two conference papers [SRH10c, Yer09], so far only two different iterative algorithms have been proposed (both in the context of ML or QML BSS): One by Pham and Garat [PG97], which is based on multiplicative updates of \mathbf{B} , and the other by Dégerine and Zaïdi [DZ04], which is based on alternating oblique projections w.r.t. the columns of \mathbf{B} . Both algorithms were developed for the real-valued case only, but can also be extended to the complex-valued case.

In this section we describe two new solutions proposed in our journal paper [YSRH12]. One is based on Newton's method and employs a conjugate gradient solution of the intermediate sets of sparse linear equations. The other is based on a modification of an existing LU-based non-orthogonal AJD algorithm [Afs06]. Both algorithms will be presented for both the real-valued and complex-valued versions of the problem.

4.3.3.1. Solution by Newton's method with conjugate gradient (NCG)

4.3.3.1.1. Real-valued target matrices

Beginning with the real-valued version, we propose to apply Newton's method for the maximization of $C(\mathbf{B})$ in order to solve the nonlinear equations (4.27). To this end, let us define the $N^2 \times 1$ gradient vector \mathbf{g} and the $N^2 \times N^2$ Hessian matrix \mathbf{H} as follows. First, we define the

indexing function $Ix(j, i) \triangleq (j-1)N + i$, which determines the location of $B_{(j,i)}$ in $\text{vec}(\mathbf{B}^T)$ (the concatenation of the columns of \mathbf{B}^T into an $N^2 \times 1$ vector). Then, as we have already seen in (4.27), the elements of the gradient vector \mathbf{g} are given (for $j, i \in \{1, 2, \dots, N\}$) by

$$g_{Ix(j,i)} \triangleq \frac{\partial C(\mathbf{B})}{\partial B_{(j,i)}} = A_{(i,j)} - \mathbf{e}_j^T \mathbf{C}_i \mathbf{b}_i. \quad (4.32)$$

. The vector \mathbf{g} can also be seen as a vectorized version $\mathbf{g} = \text{vec}(\mathbf{G}^T)$ of the gradient matrix

$$\mathbf{G} \triangleq \mathbf{A}^T - [\mathbf{C}_1 \mathbf{b}_1 \dots \mathbf{C}_N \mathbf{b}_N]^T. \quad (4.33)$$

Differentiating (4.27) once again w.r.t. $B_{(p,q)}$ we get the elements of the Hessian \mathbf{H} (for all $j, i, p, q \in \{1, 2, \dots, N\}$),

$$\begin{aligned} H_{(Ix(j,i), Ix(p,q))} &\triangleq \frac{\partial^2 C(\mathbf{B})}{\partial B_{(j,i)} \partial B_{(p,q)}} \\ &= \frac{\partial}{\partial B_{(p,q)}} \{A_{(i,j)} - \mathbf{e}_i^T \mathbf{C}_j \mathbf{b}_j\} \\ &\stackrel{(1)}{=} -\mathbf{e}_i^T \mathbf{A} \mathbf{E}_{pq} \mathbf{A} \mathbf{e}_j - \mathbf{e}_i^T \mathbf{C}_j \mathbf{e}_q \cdot \delta_{jp} \\ &= -A_{(i,p)} A_{(q,j)} - C_{j(i,q)} \cdot \delta_{jp}. \end{aligned} \quad (4.34)$$

At step (1) we have used the relation $\partial \mathbf{A} = -\mathbf{A} \cdot \partial \mathbf{B} \cdot \mathbf{A}$. The key observation here, is that if we differentiate at $\mathbf{B} = \mathbf{I}_N$, then \mathbf{H} becomes considerably sparse, since at $\mathbf{B} = \mathbf{I}_N$ we also have $\mathbf{A} = \mathbf{I}_N$, thus $A_{(n,p)} A_{(q,m)} = \delta_{np} \delta_{qm}$. The computation of the associated $N^2 \times 1$ update vector $-\mathbf{H}^{-1} \mathbf{g}$ can then be attained with relative computational simplicity using the conjugate gradient method (which exploits this sparsity). Note that with $\mathbf{B} = \mathbf{I}_N$ we have

$$\mathbf{H} = -\mathbf{P} - \text{Bdiag}(\mathbf{C}_1, \dots, \mathbf{C}_N), \quad (4.35)$$

where the $\text{Bdiag}(\cdot)$ operator creates a block-diagonal matrix from its matrix arguments, and where \mathbf{P} is merely a permutation matrix transforming the $\text{vec}(\cdot)$ of a matrix into the $\text{vec}(\cdot)$ of its transpose, namely for any $N \times N$ matrix \mathbf{Y} , we have $\mathbf{P} \cdot \text{vec}(\mathbf{Y}) = \text{vec}(\mathbf{Y}^T)$ (note also that $\mathbf{P} = \mathbf{P}^T = \mathbf{P}^{-1}$).

Therefore, the operation of \mathbf{H} on any vectorized $N \times N$ matrix \mathbf{Y}^T can be easily expressed as:

$$\mathbf{H} \cdot \text{vec}(\mathbf{Y}^T) = -\text{vec}([\mathbf{C}_1 \mathbf{y}_1, \dots, \mathbf{C}_N \mathbf{y}_N]^T + \mathbf{Y}), \quad (4.36)$$

where $\mathbf{y}_1, \dots, \mathbf{y}_N$ denote the columns of \mathbf{Y}^T (rows of \mathbf{Y}). This relatively simple relation, requiring N^3 rather than N^4 multiplications, can be conveniently exploited in a conjugate-gradient-based computation of $\mathbf{H}^{-1} \text{vec}(\mathbf{G}^T)$.

Luckily, the joint congruence structure of the SeDJoCo problem enables us to always work in the vicinity of $\mathbf{B} = \mathbf{I}_N$, as each update of \mathbf{B} can be translated into transformation of the target matrices, defining a “new” problem in terms of the transformed matrices. In other words, suppose that a set of target matrices $\mathbf{C}_1, \dots, \mathbf{C}_N$ is given, and an initial guess for \mathbf{B} is $\mathbf{B}^{(0)} = \mathbf{I}_N$. Following a single iteration of the Newton algorithm at $\mathbf{B}^{(0)} = \mathbf{I}_N$, a correction matrix $\mathbf{\Delta}$ is found and used for updating $\mathbf{B}^{(1)} = \mathbf{B}^{(0)} + \mathbf{\Delta} = \mathbf{I}_N + \mathbf{\Delta}$. Apparently, the next step is to apply the next iteration of the Newton algorithm by calculating the correction matrix at $\mathbf{B}^{(1)}$, but this would no longer be computationally appealing, since at $\mathbf{B}^{(1)} \neq \mathbf{I}_N$ the structure of the Hessian severely departs from equation (4.36) and becomes cumbersome and nonsparse. Fortunately, an attractive alternative exists in SeDJoCo. Rather than computing the next update at $\mathbf{B}^{(1)}$ with the original target matrices, we can transform these matrices into a new set of target matrices, using the congruence transformation implied by $\mathbf{B}^{(1)}$, namely obtain $\tilde{\mathbf{C}}_i = \mathbf{B}^{(1)} \mathbf{C}_i \mathbf{B}^{(1)T}$ for $i = 1, 2, \dots, N$. This transformation fully accounts for the update in \mathbf{B} , so that with the new set $\tilde{\mathbf{C}}_1, \tilde{\mathbf{C}}_2, \dots, \tilde{\mathbf{C}}_N$, $\mathbf{B}^{(0)} = \mathbf{I}_N$ can be used again as an “initial guess”, which leads to a convenient calculation of the next update. The process proceeds by retransforming the new target matrices at each step, and accumulating the updates by applying the respective left-multiplicative updates of \mathbf{B} .

With the given target matrices and some initial guess of \mathbf{B} , the NCG algorithm is summarized as follows.

1. Update the transformed target matrices

$$\tilde{\mathbf{C}}_i \leftarrow \mathbf{B} \mathbf{C}_i \mathbf{B}^T \quad i = 1, 2, \dots, N.$$

2. Using (4.33), construct the gradient matrix \mathbf{G} at $\mathbf{B} = \mathbf{I}_N$,

$$\mathbf{G} = \mathbf{I}_N - [\tilde{\mathbf{C}}_1 \mathbf{e}_1, \dots, \tilde{\mathbf{C}}_N \mathbf{e}_N]^T.$$

3. Find the correction matrix $\mathbf{\Delta}$, given by

$$\text{vec}(\mathbf{\Delta}^T) = -\mathbf{H}^{-1} \cdot \text{vec}(\mathbf{G}^T).$$

Note: a key observation is that the associated system of linear equations $\mathbf{H} \text{vec}(\mathbf{\Delta}^T) = -\text{vec}(\mathbf{G}^T)$ may be conveniently solved by using the conjugate-gradient method or the

conjugate-gradient-squared method (used in the simulations). However, as \mathbf{H} has the dimension $N^2 \times N^2$, a direct solution is computationally too expensive for large values of N . In such a case, the sparsity of \mathbf{H} in (4.35) calls for employing the conjugate-gradient method¹ to obtain an iterative solution with guaranteed convergence in a finite number of steps. The method does not involve an explicit inversion of \mathbf{H} , but merely requires a computation of products of the form $\mathbf{H}\mathbf{y}$ in each iteration. As shown in equation (4.36) above, such products are computed with N^3 (rather than N^4) multiplications by exploiting the special sparse structure of \mathbf{H} .

4. Apply and accumulate the correction $\mathbf{B} \leftarrow (\mathbf{I}_N + \Delta)\mathbf{B}$

These 4 steps are repeated until convergence. This algorithm is somewhat similar in structure to Pham’s multiplicative updates algorithm [PG97]. However, unlike the Pham’s multiplicative updates algorithm based on the direct solution of (4.21), it is based on an iterative solution of (4.27), which conveniently lends itself to the use of a conjugate gradient algorithm in each Newton iteration by exploiting the sparsity of \mathbf{H} . Furthermore, we note that the multiplicative updates algorithm in [PG97] assumes that at the vicinity of a solution the transformed matrices $\widehat{\mathbf{B}}\mathbf{C}_i\widehat{\mathbf{B}}^T$ are all nearly-diagonal for further simplification. This assumption may be reasonable in the context of BSS (since near separation the empirical correlation matrices are all nearly diagonal if the observation length T is sufficiently long). However, it excludes non-BSS applications (such as our proposed CBF), in which there is no reason for the transformed matrices to exhibit any diagonality on top of the attained “ i -wise diagonality”.

4.3.3.1.2. Complex-valued target matrices

Since the gradient and the Hessian of the real-valued $C(\mathbf{B})$ w.r.t. a complex-valued \mathbf{B} are undefined, we must therefore resort again to van den Bos’ “complex-gradient” and “complex-Hessian” [vdB94], and apply Newton’s approach to the maximization of (4.29). To this end, we need:

- * The gradient of $\hat{C}(\mathbf{B}, \mathbf{B}^*)$ w.r.t. \mathbf{B} , which we shall denote in vector form as the $N \times 1$ vector \mathbf{g}° ;
- * The gradient w.r.t. \mathbf{B}^* , which we shall denote \mathbf{g}^* ;
- * The Hessian w.r.t. \mathbf{B} and \mathbf{B}^* , which we shall denote by the $N^2 \times N^2$ matrix $\mathbf{H}^{\circ\circ}$;

¹In Matlab, the conjugate-gradient-squared method is easily implemented by applying the command $\mathbf{X} = \text{cgs}(\mathbf{A}, \mathbf{B})$ which attempts to solve the system of linear equations $\mathbf{A}\mathbf{X} = \mathbf{B}$ for $\mathbf{X} \in \mathbb{C}^{N \times M}$. The dimension of \mathbf{A} is $N \times N$ and \mathbf{B} has the dimension of $N \times M$

* The Hessian w.r.t. \mathbf{B} and \mathbf{B}^* , which we shall denote $\mathbf{H}^{\circ*}$;

* The Hessian w.r.t. \mathbf{B}^* and \mathbf{B} , which we shall denote $\mathbf{H}^{*\circ}$;

* The Hessian w.r.t. \mathbf{B}^* and \mathbf{B}^* , which we shall denote \mathbf{H}^{**} ;

Evidently,

$$g_{Ix(j,i)}^\circ \triangleq \frac{\partial \hat{C}(\mathbf{B}, \mathbf{B}^*)}{\partial B_{(j,i)}} = A_{(i,j)} - \mathbf{e}_i^\top \mathbf{C}_j \mathbf{b}_j^*, \quad (4.37)$$

which is a vectorized version $\mathbf{g}^\circ = \text{vec}((\mathbf{G}^\circ)^\top)$ of

$$\mathbf{G}^\circ \triangleq \mathbf{A}^\top - [\mathbf{C}_1 \mathbf{b}_1^* \dots \mathbf{C}_N \mathbf{b}_N^*]^\top. \quad (4.38)$$

Differentiating once again w.r.t. $B_{(p,q)}$,

$$\begin{aligned} H_{(Ix(j,i), Ix(p,q))}^{\circ\circ} &\triangleq \frac{\partial^2 \hat{C}(\mathbf{B}, \mathbf{B}^*)}{\partial B_{(j,i)} \partial B_{(p,q)}} \\ &= \frac{\partial}{\partial B_{(p,q)}} \{A_{(i,j)} - \mathbf{e}_i^\top \mathbf{C}_j (\mathbf{B}^*)^\top \mathbf{e}_j\} \\ &= -\mathbf{e}_i^\top \mathbf{A} \mathbf{E}_{pq} \mathbf{A} \mathbf{e}_j - 0 = -A_{(i,p)} A_{(q,j)}. \end{aligned} \quad (4.39)$$

Conversely, differentiating w.r.t. $B_{(p,q)}^*$,

$$\begin{aligned} H_{(Ix(m,n), Ix(p,q))}^{\circ*} &\triangleq \frac{\partial^2 \hat{C}(\mathbf{B}, \mathbf{B}^*)}{\partial B_{(j,i)} \partial B_{(p,q)}^*} \\ &= \frac{\partial}{\partial B_{(p,q)}^*} \{A_{(i,j)} - \mathbf{e}_i^\top \mathbf{C}_j (\mathbf{B}^*)^\top \mathbf{e}_j\} \\ &= 0 - \mathbf{e}_i^\top \mathbf{C}_j \mathbf{E}_{pq}^\top \mathbf{e}_j = -\mathbf{e}_i^\top \mathbf{C}_j \mathbf{e}_q \delta_{jp}. \end{aligned} \quad (4.40)$$

Naturally, we also have $\mathbf{g}^* = (\mathbf{g}^\circ)^*$, $\mathbf{H}^{**} = (\mathbf{H}^{\circ\circ})^*$ and $\mathbf{H}^{*\circ} = (\mathbf{H}^{\circ*})^*$.

The vectorized update matrix is then given by

$$\begin{bmatrix} \delta \\ \delta^* \end{bmatrix} = - \begin{bmatrix} \mathbf{H}^{\circ\circ} & \mathbf{H}^{\circ*} \\ \mathbf{H}^{*\circ} & \mathbf{H}^{**} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{g}^\circ \\ \mathbf{g}^* \end{bmatrix}, \quad (4.41)$$

where $\delta = \text{vec}(\Delta^T)$. Obviously, it is sufficient to solve for the first half only. Using the four-blocks matrix inversion relation, the solution for δ is also given by

$$\delta = -[H^{\circ\circ} - H^{\circ*}(H^{\circ\circ})^{-1}H^{*\circ}]^{-1}[g^\circ - H^{\circ*}(H^{\circ\circ})^{-1}g^*]. \quad (4.42)$$

In order to simplify, we take advantage of the ability to work at $B = I_N$. Substituting $B = I_N$ (and $A = I_N$) for the Hessian matrices (4.39), (4.40), we get

$$\begin{aligned} H^{\circ\circ} &= H^{**} = -P \\ H^{\circ*} &= (H^{*\circ})^* = -\text{Bdiag}(C_1, \dots, C_N). \end{aligned} \quad (4.43)$$

Thus, equation (4.42) reduces into

$$\delta = [P - \Lambda P^T \Lambda^*]^{-1}[g^\circ - \Lambda^* P^T g^*], \quad (4.44)$$

where $\Lambda \triangleq \text{Bdiag}\{C_1, \dots, C_N\}$ is used as a shorthand notation.

The conjugate gradient method can be used by exploiting the sparsity of the complete Hessian matrix (at $B = I_N$), since for any $N \times N$ complex-valued matrix Y , the product

$$\begin{bmatrix} -P & -\Lambda \\ -\Lambda^* & -P \end{bmatrix} \begin{bmatrix} \text{vec}(Y^T) \\ \text{vec}(Y^H) \end{bmatrix} = - \begin{bmatrix} \text{vec}(Y + [C_1 y_1, \dots, C_N y_N]^T) \\ \text{vec}(Y^* + [C_1 y_1, \dots, C_N y_N]^H) \end{bmatrix} \quad (4.45)$$

can be computed in N^3 instead of $(2N)^4$ complex-valued multiplications.

With the given target matrices C_i , $i = 1, 2, \dots, N$ and an initial guess of B , the complex-values version of the NCG algorithm is summarized. We repeat the following until convergence.

1. Update the transformed target matrices

$$\tilde{C}_i \leftarrow B C_i B^H \quad i = 1, 2, \dots, N.$$

2. Using (4.38), construct the gradient matrix G° at $B = I_N$,

$$G^\circ = I_N - [\tilde{C}_1 e_1, \dots, \tilde{C}_N e_N]^T,$$

and denote $g^\circ = \text{vec}((G^\circ)^T)$.

3. Find the correction matrix Δ , given by

$$\text{vec}(\Delta^T) = [P - \Lambda P^T \Lambda^*]^{-1}[g^\circ - \Lambda^* P^T (g^\circ)^*],$$

with $\mathbf{\Lambda} = \text{Bdiag}(\tilde{\mathbf{C}}_1, \dots, \tilde{\mathbf{C}}_N)$. To alleviate the computational load, the conjugate-gradient method can be used for this part, exploiting the sparsity of the complex Hessian by the use of (4.45) (with each \mathbf{C}_i substituted by $\tilde{\mathbf{C}}_i$).

4. Apply the correction $\mathbf{B} \leftarrow (\mathbf{I}_N + \mathbf{\Delta})\mathbf{B}$

We emphasize in passing that although the scaling equations $\mathbf{b}_i^T \mathbf{C}_i \mathbf{b}_i = 1$ (or $\mathbf{b}_i^H \mathbf{C}_i \mathbf{b}_i = 1$) are inherently built into the NCG algorithm above, they are sometimes irrelevant. For example, in the coordinated beamforming application this scaling constraint does not apply, since a large value of $\mathbf{b}_i^H \mathbf{C}_i \mathbf{b}_i$ leads to a large power of the desired signal at the i th user. Clearly, if a matrix \mathbf{B} solves the SeDJoCo problem with any scaling equations, then for any diagonal matrix \mathbf{D} , $\mathbf{D} \cdot \mathbf{B}$ also solves SeDJoCo, but with possibly different scaling equations. Therefore, the NCG solution can be used with different scaling constraints, simply by re-normalizing the rows of \mathbf{B} as desired.

4.3.3.2. Solution by structured joint congruence (STJOCO) transformation

The structured joint congruence (STJOCO) transformation is derived by modifying Afsari's LU-based non-orthogonal matrix joint diagonalization (NOJD) [Afs06]. However, unlike the algorithm in [Afs06] which was designed to simultaneously diagonalize a set of real-valued symmetric matrices $\{\mathbf{C}_i\}_{i=1}^N$ of dimension $n \times n$, the STJOCO transformation aims to solve SeDJoCo problem. Since the STJOCO transformation is an extension of the LU-based NOJD algorithm, we first review this algorithm

4.3.3.2.1. LU based NOJD algorithm

In [Afs06], an LU-based non-orthogonal joint diagonalization (NOJD) algorithm was proposed as a class of NOJD methods using triangular Jacobi matrices which are based on the LU factorization of the sought diagonalizer. A scaling-invariant cost function is used for the LU-based NOJD which has the form [Afs06]

$$J_1(\mathbf{B}) = \sum_{i=1}^N \left\| \mathbf{C}_i - \mathbf{B}^{-1} \text{diag}(\mathbf{B} \mathbf{C}_i \mathbf{B}^T) \mathbf{B}^{-T} \right\|_F^2 \quad (4.46)$$

Here, scaling-invariant means that the cost function J_1 is invariant under scaling by a non-singular diagonal matrix $\mathbf{\Lambda}$, e.g., $J_1(\mathbf{\Lambda} \mathbf{B}) = J_1(\mathbf{B})$.

The diagonalizer \mathbf{B} is updated iteratively in the form

$$\mathbf{B}_{q+1} = (\mathbf{I}_N + \mathbf{\Delta}_q) \mathbf{B}_q \quad (4.47)$$

where \mathbf{I}_N is the $N \times N$ identity matrix, $\text{diag}(\mathbf{\Delta}_q) = \mathbf{0}$ and $\mathbf{\Delta}_q$ is found such that $J_1(\mathbf{B}_{q+1})$ is minimized at each step. The LU-based NOJD algorithm considers $\mathbf{\Delta}_q$ with only one non-zero element and refers to $\mathbf{I}_n + \mathbf{\Delta}_q$ as a triangular Jacobi matrix. Then, the equation (4.47) can be further written as

$$\mathbf{B}_{q+1} = \mathbf{L}_q \mathbf{U}_q \mathbf{B}_q \quad (4.48)$$

where \mathbf{L}_q and \mathbf{U}_q are $N \times N$ unit lower and upper triangular matrices, respectively. Here a unit triangular matrix is a triangular matrix with diagonal elements of one. Unit lower and upper triangular matrices of dimension $N \times N$ form Lie groups (definition found in Appendix D.3) denoted by $\mathcal{L}(N)$ and $\mathcal{U}(N)$, respectively. This fact simplifies the minimization process significantly. Now we can find \mathbf{L}_q and \mathbf{U}_q separately in the LU form to minimize J_1 at each step.

4.3.3.2.2. STJOCO transformation for real-valued symmetric matrices

Given a set of N symmetric real-valued matrices $\{\mathbf{C}_i\}_{i=1}^N$ of dimension $n \times n$, it is the goal of the STJOCO transformation to jointly minimize the magnitude of the off-diagonal elements in the i th row and the i th column of the matrix $\mathbf{C}_i \in \mathbb{R}^{n \times n}$ with the sought matrix \mathbf{B} . The scaling-invariant cost function for the STJOCO transformation is given by

$$J_2(\mathbf{B}) = \sum_{i=1}^N \left\| \mathbf{C}_i - \mathbf{B}^{-1} [\text{diag}(\mathbf{B} \mathbf{C}_i \mathbf{B}^H) + \mathbf{G}_i] \mathbf{B}^{-H} \right\|_F^2, \quad (4.49)$$

where $\mathbf{G}_i = \mathbf{B} \mathbf{C}_i \mathbf{B}^H$, except for the diagonal elements and the off-diagonal elements of the i th row and the i th column which are zero. We also introduce an LU-based algorithm using triangular Jacobi matrices for the minimization of J_2 . The matrix \mathbf{B} is updated iteratively in the following manner

$$\begin{aligned} \mathbf{B}_{q+1} &= (\mathbf{I}_N + \mathbf{\Delta}_q) \mathbf{B}_q \\ &= \mathbf{L}_q \mathbf{U}_q \mathbf{B}_q. \end{aligned} \quad (4.50)$$

The triangular matrices \mathbf{L}_q and \mathbf{U}_q are found separately such that $J_2(\mathbf{B}_{q+1})$ is reduced at each step.

Let us define $\mathbf{L}_{l,k}(a)$ as a unit lower triangular matrix with parameter $a \in \mathbb{R}$ corresponding to the position $(l, k), l > k$ and the rest of its off-diagonal entries are zero. The matrix $\mathbf{L}_{l,k}(a)$ is an element of $\mathcal{L}(N)$. Similarly, we define an unit upper triangular Jacobi matrix with parameter a corresponding to the position $(l, k), l < k$ as $\mathbf{U}_{l,k}(a)$ which is an element of $\mathcal{U}(N)$.

Any element of $\mathcal{L}(N)$ or $\mathcal{U}(N)$ can be represented as a product of lower or upper triangular Jacobi matrices. Then the $\frac{N(N-1)}{2}$ dimensional minimization problem of finding a \mathbf{L}_q or \mathbf{U}_q to minimize J_2 can be replaced by a sequence of one dimensional problems of finding the parameter a of a triangular Jacobi matrix $\mathbf{L}_{l,k}(a)$ or $\mathbf{U}_{l,k}(a)$ for minimizing J_2 . We propose a simple lemma to solve the one-dimensional problem.

Lemma 4.3.4. *For $\mathbf{L}_{l,k}(a)$ with $l > k$, we have $J_2(\mathbf{L}_{l,k}(a)) = b_4 a^4 + b_3 a^3 + b_2 a^2 + b_1 a + b_0$. For $\mathbf{U}_{l,k}(a)$ with $l < k$, we have $J_2(\mathbf{U}_{l,k}(a)) = b_4 a^4 + b_3 a^3 + b_2 a^2 + b_1 a + b_0$, where*

$$\begin{aligned} b_4 &= 4 \sum_{i \in \{l,k\}} \mathbf{C}_i(k,k)^2 \\ b_3 &= 8 \sum_{i \in \{l,k\}} \mathbf{C}_i(k,k) \mathbf{C}_i(k,l) \\ b_2 &= \sum_{i \in \{l,k\}} \sum_{\substack{p=1 \\ p \neq l,k}}^N [2 \mathbf{C}_i(k,k)^2 + 4 \mathbf{C}_i(k,l)^2 + 2 \mathbf{C}_i(k,p)^2] \\ b_1 &= 4 \sum_{i \in \{l,k\}} \mathbf{C}_i(k,l) \mathbf{C}_i(k,k) + 4 \sum_{\substack{p=1 \\ p \neq l,k}}^N \mathbf{C}_l(l,p) \mathbf{C}_l(k,p) \\ b_0 &= 2 \sum_{i=1}^N \sum_{\substack{m=1 \\ m \neq i}}^N \mathbf{C}_i(i,m)^2. \end{aligned}$$

Here we use $\mathbf{C}_i(\text{index}_1, \text{index}_2)$ to represent the entry of the matrix \mathbf{C}_i at the position $(\text{index}_1, \text{index}_2)$.

Notice that $J_2(\mathbf{L}_{l,k}(a))$ and $J_2(\mathbf{U}_{l,k}(a))$ are fourth-order polynomials in a and are always non-negative. For a small a (i.e., $|a| < 1$) J_2 is convex on \mathbb{R} and we can always find a global minimum by solving the cubic polynomial $\frac{\partial J_2(\mathbf{L}_{l,k}(a))}{\partial a} = 0$ or $\frac{\partial J_2(\mathbf{U}_{l,k}(a))}{\partial a} = 0$. As a result, the value of the cost function J_2 is reduced at each step. Note that a only depends on the elements of the matrices \mathbf{C}_l and \mathbf{C}_k for the minimization of $J_2(\mathbf{L}_{l,k}(a))$ or $J_2(\mathbf{U}_{l,k}(a))$. The STJOCO transformation is summarized below.

1. Set $\mathbf{B}_0 = \mathbf{I}_N$, a threshold ϵ , and the iteration index $q = 0$.

2. For $l = 1, \dots, N$ and $k = 1, \dots, N$

* Upper triangular part ($l < k$): set $\mathbf{U} = \mathbf{I}_N$

– find a such that $J_2(\mathbf{U}_{l,k}(a))$ is minimized according to Lemma 4.3.4.

- Update \mathbf{C}_i and \mathbf{U} by setting

$$\begin{aligned}\mathbf{C}_i &\leftarrow \mathbf{U}_{l,k}(a)\mathbf{C}_i\mathbf{U}_{l,k}(a)^T \\ \mathbf{U} &\leftarrow \mathbf{U}_{l,k}(a)\mathbf{U} .\end{aligned}$$

- * Lower triangular part ($l > k$): set $\mathbf{L} = \mathbf{I}_N$

- find a such that $J_2(\mathbf{L}_{l,k}(a))$ is minimized according to Lemma 4.3.4.

- Update \mathbf{C}_i and \mathbf{L} by setting

$$\begin{aligned}\mathbf{C}_i &\leftarrow \mathbf{L}_{l,k}(a)\mathbf{C}_i\mathbf{L}_{l,k}(a)^T \\ \mathbf{L} &\leftarrow \mathbf{L}_{l,k}(a)\mathbf{L} .\end{aligned}$$

3. Update \mathbf{B} by setting $\mathbf{B}_{q+1} \leftarrow \mathbf{L}\mathbf{U}\mathbf{B}_q$ and $q \leftarrow q+1$. If $\frac{J_2(\mathbf{B}_q) - J_2(\mathbf{B}_{q+1})}{J_2(\mathbf{B}_q)} > \epsilon$, then go to step 2. Otherwise, the procedure has ended.

We can also use other stopping criteria such as tracking the changes in \mathbf{B} (e.g., $\|\mathbf{L}\mathbf{U} - \mathbf{I}_N\|_F$).

4.3.3.2.3. STJOCO transformation for complex-valued Hermitian matrices

The STJOCO transformation can be extended to complex-valued Hermitian matrices $\mathbf{C}_i \in \mathbb{C}^{n \times n}$ ($i = 1, 2, \dots, N$). In this case, the cost function is the same as in equation (4.49). We still apply an LU-based algorithm using complex triangular Jacobi matrices to update the matrix \mathbf{B} and minimize the cost function J_2 .

To this end, we define $\mathbf{L}_{l,k}(a \cdot \exp(j\varphi))$ as a unit lower triangular matrix with $a \cdot \exp(j\varphi)$ at the position (l, k) for $l > k$, the remaining off-diagonal entries of $\mathbf{L}_{l,k}(a \cdot \exp(j\varphi))$ are zero. The parameters a and φ are real-valued and $a > 0$. In a similar fashion we define a unit upper triangular matrix $\mathbf{U}_{l,k}(a \cdot \exp(j\varphi))$ for $l < k$. A product of lower or upper triangular Jacobi matrices forms an element of $\mathcal{L}(N)$ or $\mathcal{U}(N)$. Then we can still use a sequence of one dimensional minimization problems to replace the $\frac{n(n-1)}{2}$ dimensional minimization problem. However, in contrast to the real-valued case, two parameters a and φ have to be determined. We propose Lemma 4.3.5 to solve the complex one-dimensional problem.

Lemma 4.3.5. For $\mathbf{L}_{l,k}(a \cdot \exp(j\varphi))$ with $l > k$, $J_2(\mathbf{L}_{l,k}(a \cdot \exp(j\varphi))) = c_4 a^4 + c_3(\varphi) a^3 + c_2(\varphi) a^2 + c_1(\varphi) a + c_0$. For $\mathbf{U}_{l,k}(a \cdot \exp(j\varphi))$ with $l < k$, we have $J_2(\mathbf{U}_{l,k}(a \cdot \exp(j\varphi))) =$

$c_4 a^4 + c'_3(\varphi) a^3 + c_2(\varphi) a^2 + c'_1(\varphi) a + c_0$, where

$$\begin{aligned}
 c_4 &= 4 \sum_{i \in \{l, k\}} \mathbf{C}_i(k, k)^2 \\
 c_3(\varphi) &= 8 \sum_{i \in \{l, k\}} \mathbf{C}_i(k, k) \cdot \text{Re}\{\mathbf{C}_i(l, k) \exp(j\varphi)\} \\
 c_2(\varphi) &= 2 \sum_{i \in \{l, k\}} \sum_{\substack{p=1 \\ p \neq l, k}}^N \left[\mathbf{C}_i(k, k)^2 + |\mathbf{C}_i(l, k)|^2 + |\mathbf{C}_i(k, p)|^2 + \text{Re}\{\mathbf{C}_i(k, l)^2 \exp(2j\varphi)\} \right] \\
 c_1(\varphi) &= 4 \sum_{i \in \{l, k\}} \mathbf{C}_i(k, k) \cdot \text{Re}\{\mathbf{C}_i(k, l) \exp(j\varphi)\} + 4 \sum_{\substack{p=1 \\ p \neq l, k}}^N \text{Re}\{\mathbf{C}_l(p, l) \mathbf{C}_l(k, p) \exp(j\varphi)\} \\
 c_0 &= 2 \sum_{i=1}^N \sum_{\substack{m=1 \\ m \neq i}}^N |\mathbf{C}_i(i, m)|^2,
 \end{aligned}$$

and $c'_3(\varphi)$ is the same as $c_3(\varphi)$, except that $\exp(j\varphi)$ is replaced by $\exp(-j\varphi)$. The same change happens to $c'_1(\varphi)$.

The cost function $J_2(\mathbf{L}_{l,k}(a \cdot \exp(j\varphi)))$ or $J_2(\mathbf{U}_{l,k}(a \cdot \exp(j\varphi)))$ is a fourth-order polynomial in a and a second-order polynomial in $\cos(\varphi)$. We can compute the algebraic solutions for $\frac{\partial J_2(\mathbf{L}_{l,k}(a \cdot \exp(j\varphi)))}{\partial a} = 0$ and $\frac{\partial J_2(\mathbf{L}_{l,k}(a \cdot \exp(j\varphi)))}{\partial \varphi} = 0$. However, the expressions are very complicated. Alternatively, we can employ numerical nonlinear convex optimization methods (e.g., the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method [BV04] with cubic line search) to find the optimal point (a, φ) which minimizes the cost function $J_2(\mathbf{L}_{l,k}(a \cdot \exp(j\varphi)))$, since J_2 is convex for small a and φ (i.e., $0 < a < 1$ and $|\varphi| < \frac{\pi}{4}$). In our case these two methods reach the same optimal point for (a, φ) . Lemma 4.3.5 shows us that the minimizations of $J_2(\mathbf{L}_{l,k}(a \cdot \exp(j\varphi)))$ and $J_2(\mathbf{U}_{l,k}(a \cdot \exp(j\varphi)))$ only depends on the elements of the matrices \mathbf{C}_l and \mathbf{C}_k .

The procedure of the STJOCO transformation for the complex-valued case is the same as for the real-valued case, except for step 2. In step 2 we find a and φ for the minimization of $J_2(\mathbf{L}_{l,k}(a \cdot \exp(j\varphi)))$ or $J_2(\mathbf{U}_{l,k}(a \cdot \exp(j\varphi)))$ according to Lemma 4.3.5.

4.3.4. Convergence behavior

We evaluate the convergence of the proposed NCG and STJOCO solutions of the SeDJoCo problem in the terms of the logarithm of the residual root-mean-squares (RMS, $\sqrt{J_2(B)}$ in equation (4.49)) error versus the iteration number. The target matrices are arbitrary, randomly generated, symmetric, positive-definite, real-valued matrices. Since the results for complex-

valued matrices are very similar to the real-valued matrices, we do not present them here. The convergence is compared to the iterative relaxation (IR) algorithm proposed in [DZ04], but not to the solution proposed in [PG97], since the latter only converges for “nearly jointly diagonalizable” matrices instead of the arbitrary target matrices. For all three solutions we initialize the sought matrix \mathbf{B} to the identity matrix (i.e., $\mathbf{B} = \mathbf{I}_N$), except for the NCG algorithm with large values of N , because as N grows, the NCG algorithm becomes more sensitive to the initialization. Therefore, for NCG with $N = 10$ we initialize \mathbf{B} to the output of the IR algorithm obtained as soon as the RMS error falls below 10^{-5} . For fair comparison, we continue the iteration count from the respective IR iteration number.

Figures 4.3 and 4.4 show typical convergence pattern of the three iterative algorithms (IR, NCG, and STJOCO) for several independent trials. Note that the numbers on the y axis are \log_{10} of the RMS, and are not given in dB. The lower “saturation line” reflects an average residual error of about 10^{-30} , which means that a convergence pattern reaching that line attains the *exact* solution. Here, we define “convergence to an existing solution” as the state where the residual RMS error drops to the machine-accuracy around 10^{-30} .

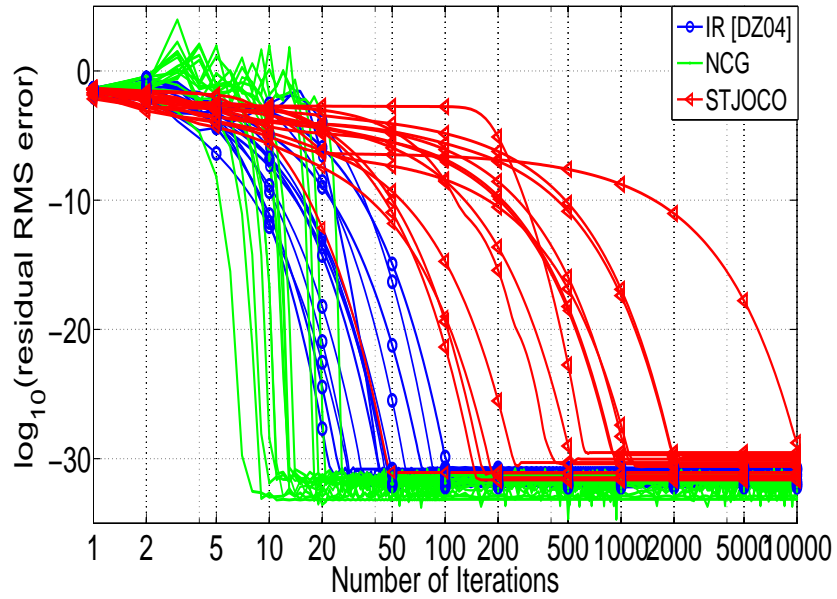


Figure 4.3.: RMS error for IR, NCG, and STJOCO with arbitrary positive-definite real-valued matrices for $N = n = 3$.

It is evident that the NCG algorithm significantly accelerates the convergence. For $N = 3$, in

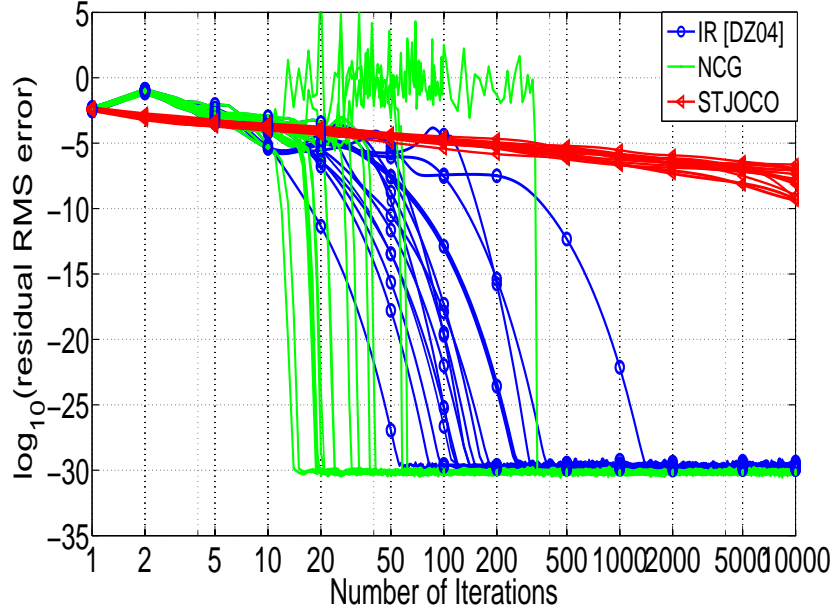


Figure 4.4.: RMS error for IR, NCG, and STJOCO with arbitrary positive-definite real-valued matrices for $N = n = 10$.

100 independent trials the median number of iterations until convergence to an exact solution was 42 for IR, 13 for NCG and 635 for STJOCO. As evident from Figure 4.4, STJOCO did not converge to a machine-accuracy solution for $N = 10$ with the maximal tested number of 10 000 iterations, but still attained very reasonable solutions with small residual errors (way below 10^{-5} after more than 500 iterations), which are probably local minima of its respective cost function.

The accelerated convergence of NCG is obtained at the cost of only a moderate increase in the computational complexity per iteration. The complexity per iteration is $\mathcal{O}(N^4)$ (i.e., the order of the multiplication is N^4 [Knu98]) for the IR algorithm and approximated $\mathcal{O}(N^5)$ (i.e., the order of the multiplication is N^5) for the NCG algorithm. The STJOCO algorithm has a comparable computational complexity $\mathcal{O}(N^4)$ per iteration, but occasionally converges to local non-zero minima of the cost-function. As we will show in the context of CBF that this apparent disadvantage of STJOCO is generally compensated by a higher effective signal-to-interference-plus-noise ratio (SINR) when a common threshold ϵ is defined for NCG, STJOCO, and IR algorithms (e.g., $\epsilon = 10^{-7}$). Since the SINR is given by the ratio between the square of the diagonal element on the “drilled” row or column (i.e., the (i,i) th element of the “drilled”

i th row or column) and the sum of the squares of the off-diagonal elements in the same row or column, this effective SINR is not effected by the scaling constraint on the sought matrix \mathbf{B} (discussed in Section 4.3.2).

4.3.5. Achievable sum rate of SeDJoCo-based CBF

With the application of the SeDJoCo transformation the combined transmit beamformer \mathbf{F} defined in equation (4.12) can be directly obtained as \mathbf{B}^H . The number of users K is the same as the number of target matrices, and the number of transmit antennas M_T corresponds to the parameter N (namely, to the dimensions of the target matrices). Usually, N would equal the number of users K . If N is smaller than K , then the number of matrices exceeds their dimensions, and generally SeDJoCo does not have a solution in such case. Conversely, if N is larger than K , then the system contains inherent redundancy. We can either add more users or transmit multiple data streams to some of the users. For example, if the sequence of K target matrices is augmented from $\{\mathbf{C}_1, \dots, \mathbf{C}_K\}$ to a sequence of $N > K$ target matrices $\{\mathbf{C}_1, \dots, \mathbf{C}_K, \mathbf{C}_K, \dots, \mathbf{C}_K\}$, such that \mathbf{C}_K is repeated $N - K + 1$ times. Then, following the SeDJoCo solution the K th user is able to receive $N - K + 1$ data streams with zero MUI. In this work, we concentrate in the sequel on the case $N = K$.

The receive beamformer of each user is matched to the user's effective channel (i.e., $\mathbf{w}_i = \mathbf{H}_i \mathbf{f}_i$). In a system where dedicated pilots are used for the downlink, each user can estimate its own receive beamformer. For the systems which use only a common pilot channel for all users (e.g., 3GPP long term evolution), there is no way to estimate the effective channel gain at the receiver. To solve this problem, a limited feedforward method can be utilized to inform the receivers about the post-processing vectors [CMIH08a].

The performance of the SeDJoCo-based closed-form CBF in terms of the achievable sum rate of a multi-user MIMO system is investigated. We also compare our results to the IR solution [DZ04]. In addition, we compare them to the iterative coordinated beamforming (CBF) algorithms [CMIH08a], the suboptimal coordinated BD algorithm [SSH04], and regularized block diagonalization (RBD) linear precoding, since RBD can still be applied under the condition that the system has a smaller number of transmit antennas M_T than the total number of receive antennas M_R . In the simulations, we have $M_T = K$ and transmit one data stream to each user. For simplicity, an equal power allocation is employed among the users. The achievable sum rate is calculated as $R = \sum_{i=1}^K \log_2(1 + \text{SINR}_i)$, where SINR_i indicates the signal-to-interference-plus-noise ratio at the user i . Dirty paper coding (DPC) has been shown to achieve the capacity region of Gaussian MIMO broadcast channels. Therefore, we use the

achievable sum rate of DPC as a benchmark.

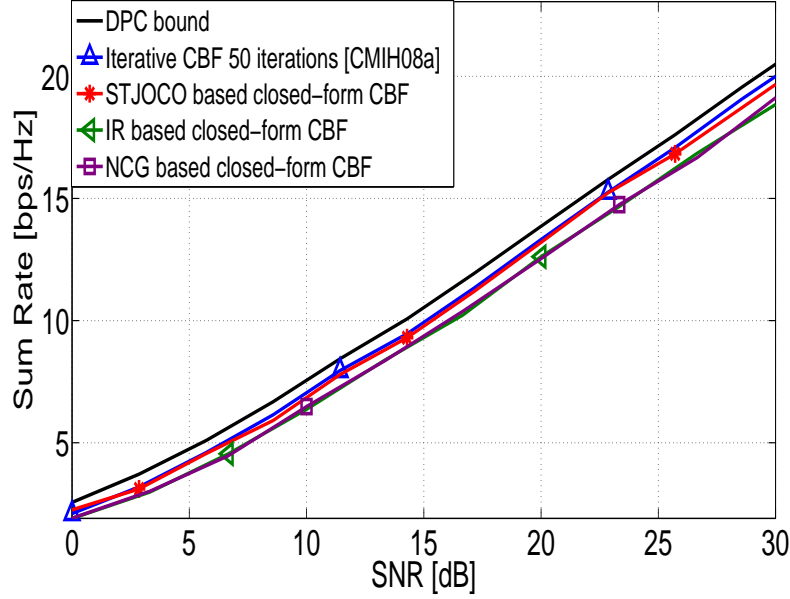


Figure 4.5.: Achievable sum rate comparison for $M_T = M_{R_i} = K = 2$.

Figure 4.5 shows the comparisons of the iterative coordinated beamforming (CBF) algorithms [CMIH08a], the proposed SeDJoCo-based, as well as IR-based closed-form CBF algorithms, when the system has two transmit antennas with two users and each user is equipped with two receive antennas. It is observed that the STJOCO based closed-form CBF almost achieves the same sum rate performance as the iterative CBF and performs better than NCG-based and IR-based closed-form CBF. After rescaling each column of the combined transmit beamforming matrix \mathbf{B}^H to have unit norm, the STJOCO solution tends to yield larger effective SINRs (compared to the IR and NCG solutions). Since the SINR is given by the ratio between the square of the diagonal element on the “drilled” row or column (i.e., the (i,i) th element of the “drilled” i th row or column) and the sum of the squares of the off-diagonal elements in the same row or column, this effective SINR is not effected by the scaling constraint on the sought matrix \mathbf{B} (discussed in Section 4.3.2). The larger SINRs lead to higher achievable sum rates, even in cases where STJOCO does not attain an exact solution and some residual MUI is present. This has been consistently observed in our simulations.

In Figures 4.6 and 4.7, the comparisons of the iterative CBF algorithm [CMIH08a], the proposed STJOCO- and NCG- and IR-based closed-form CBF algorithms, the sub-optimal coordinated BD algorithm [SSH04], and RBD precoding are presented. Here, the system has

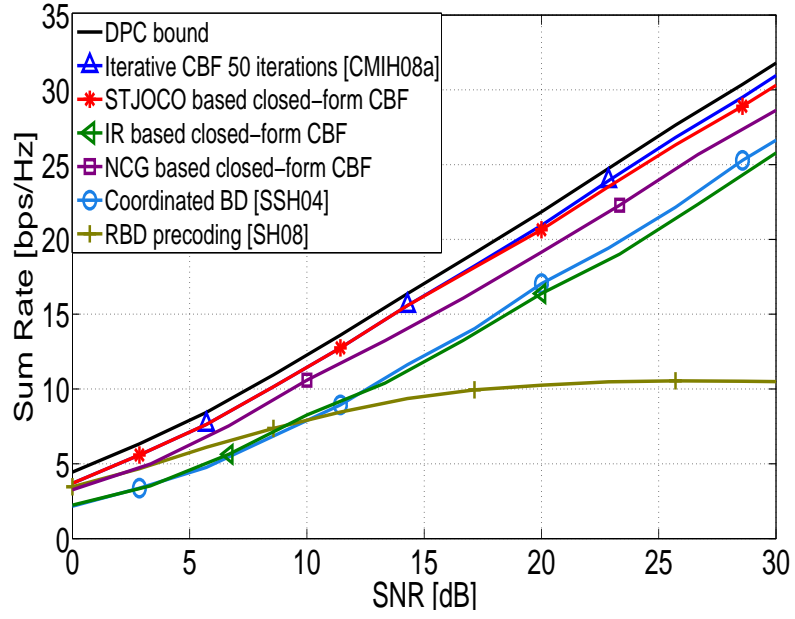


Figure 4.6.: Achievable sum rate comparison for $M_T = M_{R_i} = K = 3$.

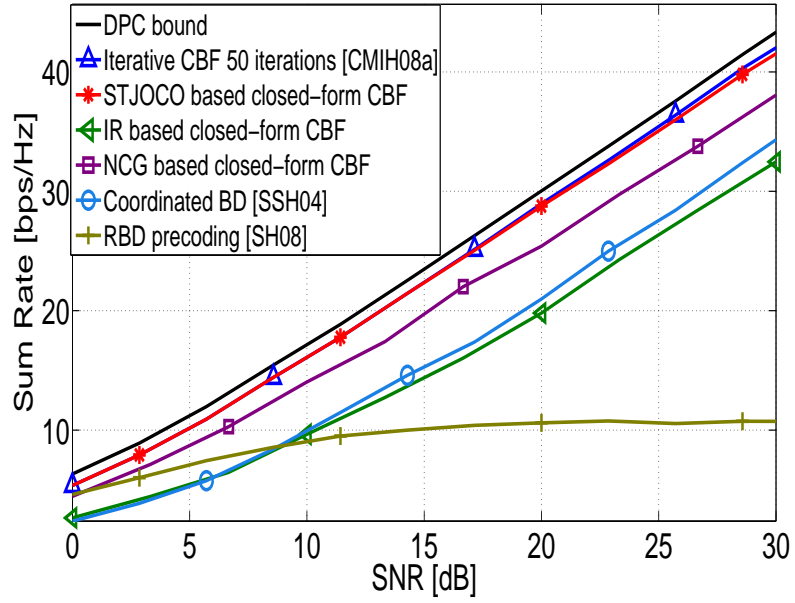


Figure 4.7.: Achievable sum rate comparison for $M_T = M_{R_i} = K = 4$.

$M_T = M_{R_i} = K = 3$ and $M_T = M_{R_i} = K = 4$ for Figures 4.6 and 4.7, respectively. We can see that the STJOCO-based closed-form CBF performs much better than the NCG- and the IR-based closed-form CBF algorithms as well as the sub-optimal coordinated BD, by achieving almost the same performance as the iterative CBF. The performance of RBD is heavily degraded when the system has a much larger number of total receive antennas than the number of transmit antennas.

4.4. Flexible coordinated beamforming (FlexCoBF)

In the above section, we have discussed our proposed closed-form CBF algorithm which provides an efficient solution for the open problem of CBF. However, our investigation focuses on supporting one data stream per user (i.e., $M_T = K$) and the receive beamforming strategy is fixed to the MRC (i.e., $\mathbf{W}_i = \mathbf{H}_i \mathbf{F}_i$) receiver.

In this section we propose an iterative coordinated beamforming algorithm for the multi-user MIMO downlink channels, which is named as flexible coordinated beamforming (FlexCoBF) [SRH10b]. Providing a high flexibility, the FlexCoBF algorithm efficiently solves the dimensionality problem which constrains the applications of some linear precoding techniques. Compared to the previous iterative coordinated beamforming [SSH04, ZHV08, SH08, CMIH08a], the main advantages of FlexCoBF are as follows.

- * It supports the transmission of multiple data streams to each user.
- * FlexCoBF does not require special designs of the transmit-receive beamforming weights contrast to the existing CBF algorithms, because FlexCoBF provides freedom in the choice of the linear transmit and receive beamforming strategies.
 - For the transmit beamforming, any existing linear precoding technique can be applied (e.g., ZF, BD, and RBD).
 - The receive beamforming strategy can be chosen flexibly (e.g., MRC or MMSE receivers).
- * The convergence behavior of FlexCoBF is investigated. The simulation results demonstrate that the sum rate performance of FlexCoBF approaches the sum capacity of the MIMO broadcast channel as the algorithm in [CMIH08a], while requiring significantly fewer iterations.

Furthermore, the extension of the FlexCoBF algorithm to a cellular scenario is considered. In a cellular system, the transmission in each cell acts as interference to the other

cells, and the entire system is essentially interference-limited. Conventional approaches to mitigate multi-cell interference, such as sectoring and frequency reuse, have unneglectable drawbacks [ACJH07]. By introducing the concept of coordinated multipoint (CoMP) transmission proposed for future wireless communication systems, full cooperation or limited cooperation can be envisioned between BSs to combat the multi-cell interference. In a multi-cell MIMO system considering the full cooperation between transmitters, BSs are linked to a central processor via ideal backhaul links (error free and unlimited capacity). Not only channel state information, but also the full data signals of the users can be shared among BSs. Consequently, the multi-cell system is transformed into a multi-user MIMO system. To this end, the existing linear transmit-receive strategies for single cell such as DPC, ZF, BD, RBD, CBF can be directly applied. However, some issues like the complexity of the joint processing across all the BSs, the difficulty of CSI acquisition of all the users at each BS, and the time or phase synchronization render full cooperation extremely difficult for realistic cellular systems. Therefore, several coordinated beamforming algorithms for downlink multi-cell system have been explored when limited cooperations between BSs are considered [KLL⁺09, CHHT13, CKH09, GMK10, TGR09, DY10, VPW10]. The method in [KLL⁺09] is an extension of the coordinated beamforming algorithm in [CMIH08a] for a multi-cell system. A single data stream to each user is considered and the receive beamforming strategy is fixed to MRC matched filtering. The authors in [CHHT13, CKH09, GMK10, TGR09] consider the scenario that one user intends to receive a desired data stream only from the desired BS, which refers to a K -user MIMO interference channel. The coordinated beamforming algorithms with interference alignment are valid only for some configurations of the number of transmit-receive pairs and the number of antennas at the BS and users such as a two cell system with one transmit-receive pair per cell in [CHHT13] and a three cell system with one transmit-receive pair per cell in [CKH09] where the number of transmit antenna per BS is 1 (i.e., $M_T = 1$). In [DY10, VPW10], the design criteria of the coordinated beamforming algorithms are the minimization of the total weighted transmit power subject to signal-to-interference-and-noise-ratio (SINR) constraints at the users or the maximization of the instantaneous weighted sum-rate subject to per-base-station power constraints. The optimal beamformers are found as the solutions of the nonconvex problems. However, each user is equipped only with one antenna in [DY10, VPW10] and the computational complexity could be an issue in practice.

A clustered cooperation strategy between BSs has been proposed [Ven07, BH07, ZCA08] and has drawn a significant amount of interest recently. This strategy considers the full cooperation in a cluster of N cells and limited cooperation between adjacent clusters. In this case, the users are classified into two groups: cluster interior users and cluster edge users. Only the

cluster edge users suffer from the inter-cluster interference (ICI). With full cooperation within one cluster, the available spatial degrees of freedom are greatly increased, which can be utilized to mitigate the intra-cluster interference and enhance the sum rate. Meanwhile, the limited cooperation between adjacent cluster is used to reduce the interference for the cluster edge users. In this way, the intra-cluster and inter-cluster interferences are efficiently mitigated. Moreover, compared to the full cooperation within a cellular network, the system complexity and users' CSI and data requirements are significantly reduced due to the definition of clusters.

We extend FlexCoBF for this clustered cellular scenario in Section 4.4.2. Instead of only considering inter-cluster interference reduction by limited cooperation between adjacent clusters, we assume that the adjacent clusters involved in the limited cooperation transmit the same data streams to the edge user to assist the transmission instead of acting as interference. To this end, the available degrees of freedom are fully utilized.

4.4.1. FlexCoBF in a single cell

We consider a multi-user MIMO downlink system with a single base station (BS) and K users, where the BS is equipped with M_T transmit antennas and the user i has M_{R_i} receive antennas. The total number of receive antennas is denoted by M_R , i.e., $M_R = \sum_{i=1}^K M_{R_i}$. In this section we focus on the case $M_R > M_T$. Notice that if $M_R \leq M_T$, any existing linear precoding method or non-linear precoding method can be directly applied. We represent a quasi-static block-fading MIMO channel between the BS and the i th user by $\mathbf{H}_i \in \mathbb{C}^{M_{R_i} \times M_T}$. In each fading block, the channel \mathbf{H}_i is considered constant. Let $\mathbf{s}_i \in \mathbb{C}^{r_i}$ denote the transmitted signal for the i th user and $\mathbf{F}_i \in \mathbb{C}^{M_T \times r_i}$ indicate the transmit beamforming matrix of user i . The receive beamforming matrix for user i is denoted by $\mathbf{W}_i \in \mathbb{C}^{M_{R_i} \times r_i}$. The variable r_i represents the number of data streams to user i . We use the term r to indicate the total number of data streams for all users (i.e., $r = \sum_{i=1}^K r_i$) and we have $r \leq M_T$. The i th receiver observes zero mean circularly symmetric complex Gaussian white noise $\mathbf{n}_i \in \mathbb{C}^{M_{R_i}}$ with variance σ_n^2 . Then, the received signal of the i th user after receive combining is expressed as

$$\mathbf{y}_i = \mathbf{W}_i^H \mathbf{H}_i \mathbf{F}_i \mathbf{s}_i + \mathbf{W}_i^H \mathbf{H}_i \sum_{\ell=1, \ell \neq i}^K \mathbf{F}_\ell \mathbf{s}_\ell + \mathbf{W}_i^H \mathbf{n}_i, \quad i = 1, \dots, K. \quad (4.51)$$

The first term on the right-hand-side (RHS) of equation (4.51) is the desired signal for user i and the second term represents the MUI at the user i caused by the other users in the system. With a joint design of the transmit-receive beamformers, FlexCoBF can enforce zero MUI at each user and achieve a sum rate close to the sum capacity of the MIMO broadcast channel.

We define an equivalent multi-user channel matrix $\mathbf{H}_e \in \mathbb{C}^{r \times M_T}$ and a combined transmit

beamforming matrix $\mathbf{F} \in \mathbb{C}^{M_T \times r}$ as

$$\mathbf{H}_e = \begin{bmatrix} \mathbf{W}_1^H \mathbf{H}_1 \\ \mathbf{W}_2^H \mathbf{H}_2 \\ \vdots \\ \mathbf{W}_K^H \mathbf{H}_K \end{bmatrix} \quad (4.52)$$

and

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \dots & \mathbf{F}_K \end{bmatrix}. \quad (4.53)$$

Note that r should satisfy $r \leq M_T$. The receive beamforming matrices \mathbf{W}_i can be chosen flexibly, e.g., MRC or MMSE receivers. The transmit beamforming matrix \mathbf{F}_i is found by applying an arbitrary linear precoding technique on the matrix \mathbf{H}_e . In this section, we use ZF, BD, and RBD precoding as examples. The transmit-receive beamformers are updated iteratively until the stopping criterion is satisfied.

The FlexCoBF algorithm is summarized as follows.

1. Initialize $\mathbf{W}_i^{(0)}$ ($i = 1, \dots, K$) to some random matrices, the iteration index p to zero, and set the threshold ϵ .
2. Set $p \leftarrow p + 1$ and compute $\mathbf{H}_e^{(p)}$ as

$$\mathbf{H}_e^{(p)} = \begin{bmatrix} \mathbf{W}_1^{(p-1)H} \mathbf{H}_1 \\ \mathbf{W}_2^{(p-1)H} \mathbf{H}_2 \\ \vdots \\ \mathbf{W}_K^{(p-1)H} \mathbf{H}_K \end{bmatrix}$$

3. Apply the desired linear precoding algorithm on the matrix $\mathbf{H}_e^{(p)}$ to obtain the transmit beamforming matrices $\mathbf{F}_i^{(p)}$ for all users ($i = 1, \dots, K$).
4. Compute the receive beamformers for the p th iteration according to the desired strategy using the precoded channel $\mathbf{H}_i \mathbf{F}_i^{(p)}$. For instance,

* MRC receiver:

$$\mathbf{W}_i^{(p)} = \mathbf{H}_i \mathbf{F}_i^{(p)}$$

* MMSE receiver²:

$$\mathbf{W}_i^{(p)} = (\mathbf{H}_i \mathbf{F}_i^{(p)} \mathbf{F}_i^{(p)\text{H}} \mathbf{H}_i^{\text{H}} + \sigma_n^2 \mathbf{I}_{M_{\text{R}_i}})^{-1} \mathbf{H}_i \mathbf{F}_i^{(p)}$$

5. Check the stopping criterion. If it is not satisfied, go back to step 2. Otherwise, convergence is achieved and the procedure has ended.

* **Stopping criterion I** for the chosen linear precoding which enforces zero MUI (e.g., ZF and BD): track the residual MUI which is given as

$$\text{MUI}(\mathbf{H}_e^{(p+1)} \mathbf{F}^{(p)}) = \left\| \text{off}(\mathbf{H}_e^{(p+1)} \mathbf{F}^{(p)}) \right\|_F^2 < \epsilon,$$

where $\mathbf{F}^{(p)} = [\mathbf{F}_1^{(p)}, \mathbf{F}_2^{(p)}, \dots, \mathbf{F}_K^{(p)}]$ and $\text{off}(\cdot)$ indicates all off-diagonal elements of the matrix $\mathbf{H}_e^{(p+1)} \mathbf{F}^{(p)}$.

* **Stopping criterion II** for the chosen linear precoding which allows some MUI (e.g., RBD): track the changes of the transmit beamformer which is $\left\| \mathbf{F}^{(p+1)} - \mathbf{F}^{(p)} \right\|_F^2 < \epsilon$.

Compared to the previous iterative CBF algorithms [SSH04, ZHV08, SH08, CMIH08a], FlexCoBF provides freedom in the choice of the transmit-receive beamforming strategies. The receive beamforming strategy can be chosen flexibly (e.g., MRC or MMSE receivers) and any existing linear precoding techniques (e.g., ZF, BD, and RBD) can be applied as the transmit beamforming strategy. The complexity of the FlexCoBF algorithm mainly depends on the complexity of the chosen transmit beamforming strategy. For example, if ZF precoding is chosen as the transmit beamforming strategy for the FlexCoBF algorithm, we obtain the same low complexity per iteration as the CBF algorithm in [CMIH08a], while achieving the same sum rate performance as the CBF algorithm in [CMIH08a] with significantly fewer iterations. This is demonstrated numerically in the simulation results.

In order to achieve a fast convergence, we propose two stopping criteria for FlexCoBF depending on the chosen linear precoding scheme. For instance, ZF and BD precoding techniques lead to zero MUI. Tracking residual MUI is more efficient than tracking the changes of the transmit beamformer. However, RBD precoding allows some MUI and the amount of MUI changes with the SNR. If we use stopping criterion I to track the residual MUI for FlexCoBF with RBD precoding, the threshold assigned to this stopping criterion should be adaptive with the SNR to ensure the convergence. This adaptive process requires an accurate prediction

²In the MMSE receiver structure, $\sigma_n^2 \mathbf{I}_{M_{\text{R}_i}}$ can be replaced by an interference plus noise covariance matrix if it can be estimated in the system. Otherwise σ_n^2 is the estimated background noise level.

of the MUI for different SNRs which is difficult to achieve. Consequently, we find that the stopping criterion II is a better solution for FlexCoBF with linear precoding techniques which allow some MUI. The efficiency of stopping criterion II is demonstrated numerically in the simulation results.

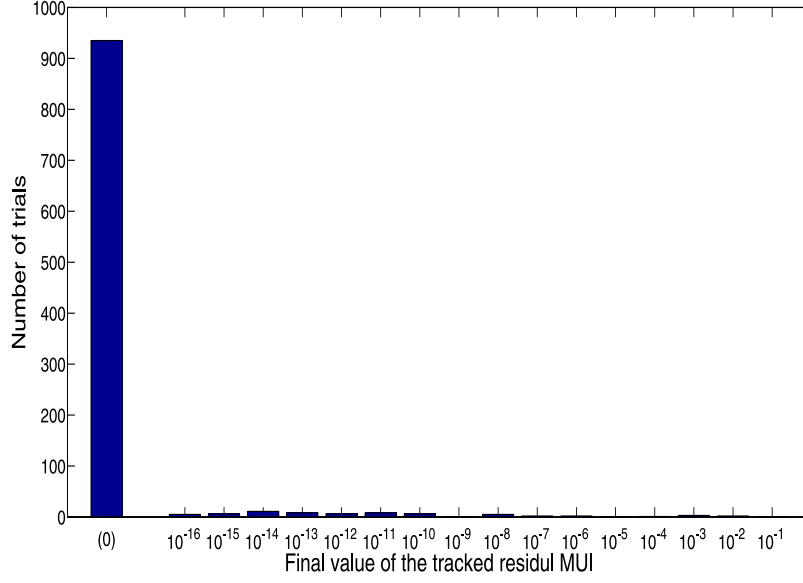


Figure 4.8.: The tracked residual MUI after 50 iterations for FlexCoBF with BD precoding where stopping criterion I is applied and $M_T = K = 3$, $M_{R_i} = 2$, and SNR = 10 dB

FlexCoBF with stopping criterion I and II converges in most cases. Figure 4.8 is shown as an example, where 1000 trials have been investigated and all trials stop after the 50th iteration. We consider a multi-user MIMO system with 3 users and each user has two receive antennas. The BS is equipped with 3 transmit antennas. The tracked residual MUI of the tested trials can reach zero after 50 iterations. But there are rare trial cases where the convergence is not guaranteed. In these cases the behavior is generally that a certain equilibrium point is reached where the tracked residual MUI or the changes of the transmit beamformers converges to a value above the threshold, and does not decrease with further iterations. Then, the best available solutions for these cases are the results after a predetermined number of iterations. Table 4.1 is shown as an example. Here, the threshold is set to 10^{-5} . The trials, which did not reach the threshold, have been manually stopped after 150 iterations.

Table 4.1.: Convergence investigation for $M_T = K = 3$, $M_{R_i} = 2$, and SNR = 10 dB

Precoding methods	Nr. of investigated trials	Nr. of manually stopped trials	Reached values of the manually stopped trials
FlexCoBF + ZF (stopping criterion I)	5000	11 (0.2%)	0.0884 1.5425 1.2704 0.0422 2.5802 0.0396 0.2376 0.0483 1.3408 1.4471 0.2376
FlexCoBF + BD (stopping criterion I)	5000	8 (0.16%)	0.5633 0.0008 0.0884 0.1026 1.5683 0.3126 0.0323 1.6592
FlexCoBF + RBD (stopping criterion II)	5000	17 (0.34%)	2.6740 0.0007 1.2611 1.3410 0.0009 1.3844 1.3333 0.00004 2.5937 0.0009 1.0156 2.6808 2.6989 1.4073 1.2091 0.00008 0.0006

4.4.1.1. Simulation results

To show the improvement of FlexCoBF, we compare it with the iterative CBF in [CMIH08a] and the coordinated BD in [SSH04]. We assume that the BS and each user know the channel state information (CSI) perfectly. Spatial correlation of the channels is considered and the spatial correlated channel $\mathbf{H} \in \mathbb{C}^{M_R \times M_T}$ is generated by the Kronecker model introduced in Chapter 2. We have

$$\mathbf{H} = \mathbf{R}_r^{1/2} \mathbf{H}_w (\mathbf{R}_t^{1/2})^H, \quad (4.54)$$

where \mathbf{H}_w is a spatially white unit variance flat fading MIMO channel of dimension $M_R \times M_T$. The matrices $\mathbf{R}_r \in \mathbb{C}^{M_R \times M_R}$ and $\mathbf{R}_t \in \mathbb{C}^{M_T \times M_T}$ are the receive and transmit correlation matrices with $\text{tr}(\mathbf{R}_r) = M_R$ and $\text{tr}(\mathbf{R}_t) = M_T$, respectively. In the simulation, we assume that

$$\mathbf{R}_r = \begin{bmatrix} \rho_r^0 & (\rho_r^*)^1 & \dots & (\rho_r^*)^{M_R-1} \\ \rho_r^1 & \rho_r^0 & \dots & (\rho_r^*)^{M_R-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_r^{M_R-1} & \rho_r^{M_R-2} & \dots & \rho_r^0 \end{bmatrix} \quad (4.55)$$

and

$$\mathbf{R}_t = \begin{bmatrix} \rho_t^0 & (\rho_t^*)^1 & \dots & (\rho_t^*)^{M_T-1} \\ \rho_t^1 & \rho_t^0 & \dots & (\rho_t^*)^{M_T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_t^{M_T-1} & \rho_t^{M_T-2} & \dots & \rho_t^0 \end{bmatrix}, \quad (4.56)$$

where ρ_r and ρ_t ($0 \leq |\rho_r|, |\rho_t| \leq 1$) are the correlation coefficients at the receiver and transmitter side, respectively. In the simulations, we assume $\rho_r = \rho_t = \rho$.

The threshold ϵ for the stopping criterion is set to 10^{-5} in all simulations and the maximum number of iterations is limited to 50. The total transmit power P_T is equally allocated among users. The received signal-to-noise ratio is defined as $\text{SNR} = P_T/\sigma_n^2$. The sum rate performance is estimated by averaging over 500 channel realizations.

4.4.1.1.1. Single Data Stream Transmission

In order to compare the proposed FlexCoBF algorithm with the CBF algorithm in [CMIH08a], we consider the transmission of a single data stream per user first.

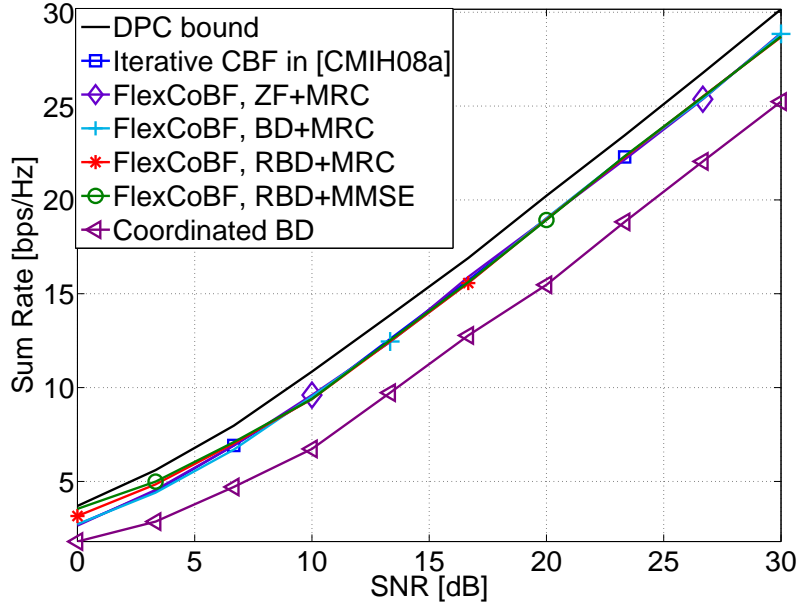


Figure 4.9.: Achievable sum rate for $M_T = K = 3$ and $M_{R_i} = 2$.

In Figures 4.9-4.11, we assume a MIMO downlink system with three users in uncorrelated Rayleigh fading (i.e., $\rho = 0$). Each user is equipped with $M_{R_i} = 2$ receive antennas and the base station has $M_T = 3$ transmit antennas. Figure 4.9 shows the sum rate performance comparisons among the CBF algorithm in [CMIH08a], the proposed FlexCoBF algorithm applying different linear precoding techniques (e.g., ZF, BD, and RBD) as the transmit beamforming strategy, and the coordinated BD in [SSH04]. We observe that coordinated BD has the worst sum rate performance due to the missing joint optimization between the transmit and receive

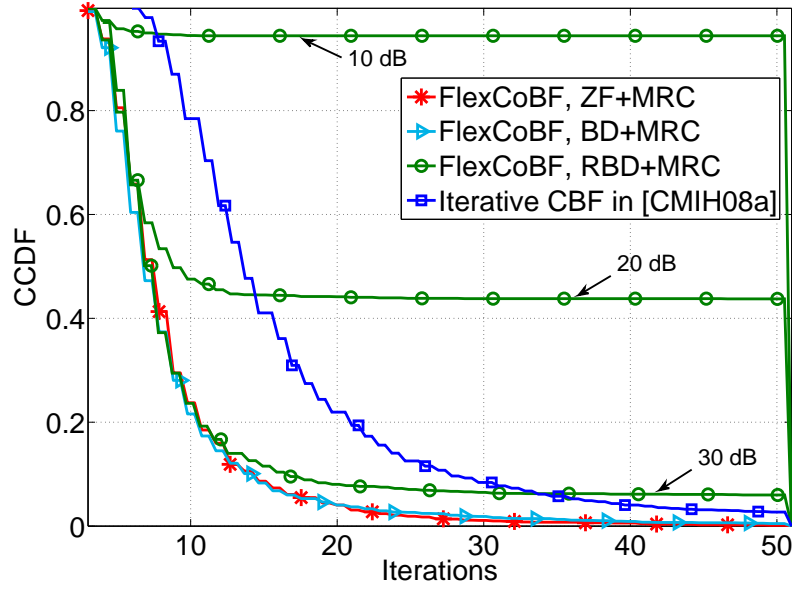


Figure 4.10.: CCDF of required number of iterations for $M_T = K = 3$, $M_{R_i} = 2$ and all SNRs (i.e., 0 dB – 30 dB), with stopping criterion I.

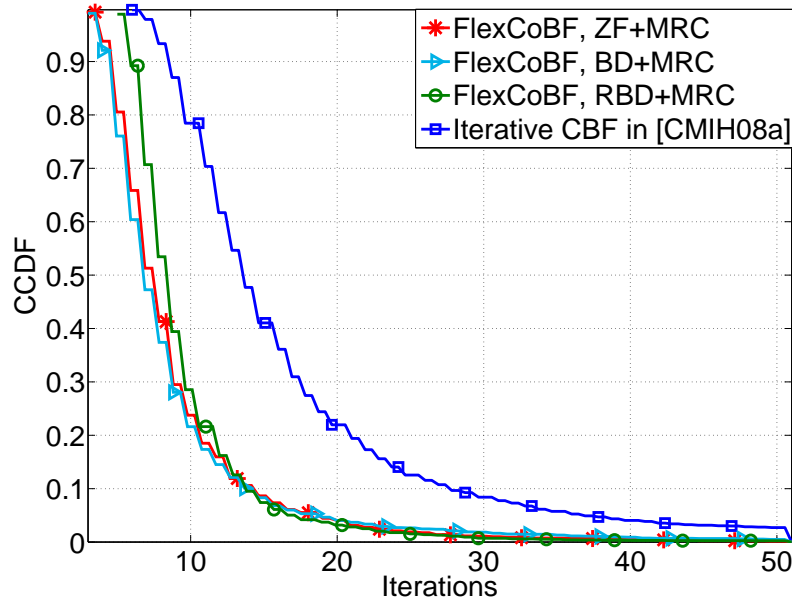


Figure 4.11.: CCDF of required number of iterations for $M_T = K = 3$, $M_{R_i} = 2$ and all SNRs (i.e., 0 dB – 30 dB), stopping criterion I for FlexCoBF with ZF and BD, stopping criterion II for FlexCoBF with RBD.

beamformers. The FlexCoBF algorithm can achieve a noticeably increased sum rate performance, as the CBF algorithm in [CMIH08a]. It is noticed that FlexCoBF with RBD precoding as the transmit beamforming strategy can achieve the best sum rate performance at low SNRs.

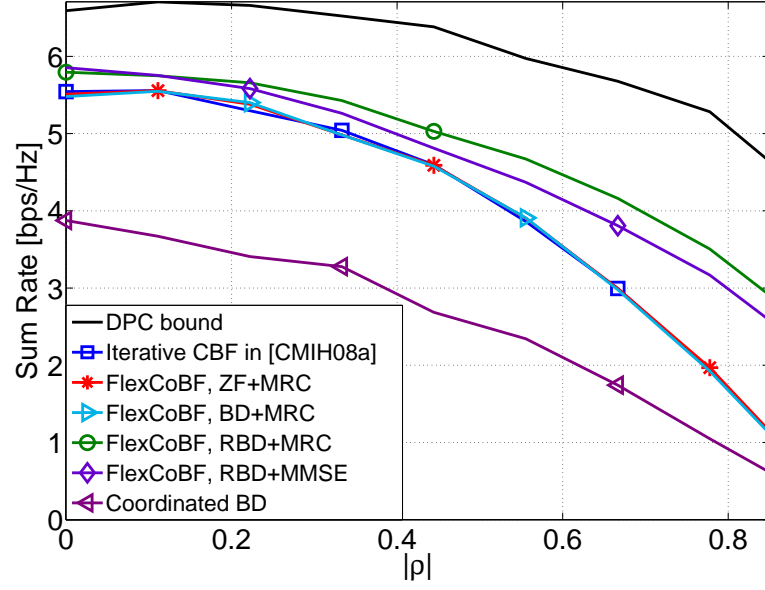
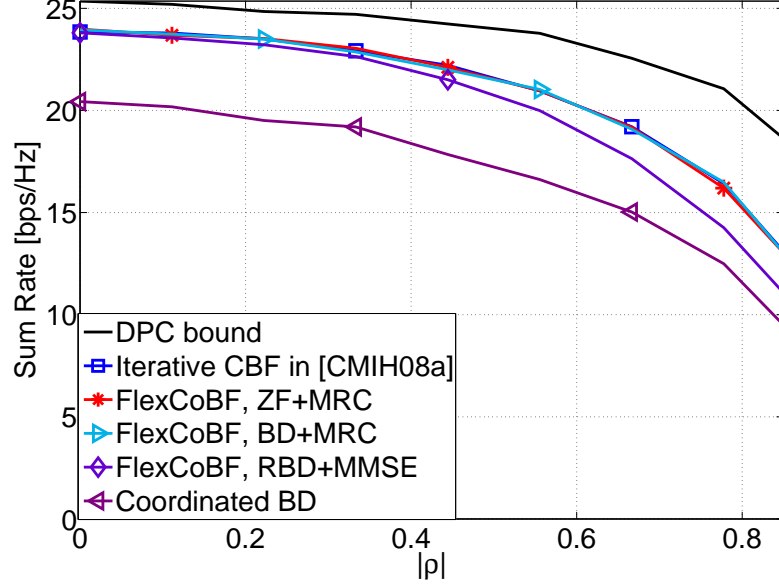
Figure 4.10 displays the complementary cumulative distribution function (CCDF) of the required number of iterations of the FlexCoBF algorithm with stopping criterion I. We find that FlexCoBF applying ZF and BD requires significantly fewer iterations compared to the CBF in [CMIH08a], while having a comparable computational complexity per iteration as the CBF in [CMIH08a]. Furthermore, in 80 percent of the trial cases FlexCoBF with ZF and BD precoding converged in less than 10 iterations. However, the required number of iterations of FlexCoBF applying RBD precoding is affected by the SNR, because of the property of RBD precoding that a small amount of MUI is allowed except for very high SNRs. A fixed threshold $\epsilon = 10^{-5}$ does not work well for this case. In order to ensure the convergence for all SNRs, we have to adapt the threshold ϵ with the SNR which requires a prediction of the remaining MUI and is difficult to solve.

Therefore, we investigate stopping criterion II for FlexCoBF with RBD precoding to avoid the adaptation. Instead of tracking the residual MUI, we take the changes of the transmit beamformers as a stopping criterion (i.e., $\|\mathbf{F}^{(p+1)} - \mathbf{F}^{(p)}\|_F^2 < \epsilon$). It is found that FlexCoBF with RBD requires significantly less iterations with stopping criterion II, while achieving the same sum rate performance as FlexCoBF with RBD that tracks the residual MUI. Figure 4.11 shows that the convergence of FlexCoBF with RBD is ensured by stopping criterion II for all SNRs (i.e., 0 dB – 30 dB).

Figures 4.12 and 4.13 show the effect of the spatial correlation on the sum rate performance. We still consider the system with $M_T = K = 3$ and $M_{R_i} = 2$. The sum rate performance is found to degrade with the increasing spatial correlation coefficient ρ . Furthermore, FlexCoBF with RBD as the transmit beamforming strategy and an MMSE receiver as the receive beamforming strategy is more sensitive to the spatial correlation compared to FlexCoBF with RBD as the transmit beamforming strategy and an MRC receiver as the receive beamforming strategy.

4.4.1.1.2. Multiple Data Streams Transmission

In Figures 4.14-4.16, we consider multiple data streams transmission per user. Equal power allocation is employed among the different data streams of all users. It is observed that the proposed FlexCoBF algorithm has a significantly increased sum rate performance compared to the coordinated BD proposed in [SSH04], when the base station simultaneously transmits multiple data streams to each user in the case where the total number of receive antennas M_R exceeds the number of transmit antennas M_T . At low SNRs, FlexCoBF with RBD precoding as


 Figure 4.12.: Achievable sum rate for $M_T = K = 3$ and $M_{R_i} = 2$ at SNR = 5 dB vs. ρ .

 Figure 4.13.: Achievable sum rate for $M_T = K = 3$ and $M_{R_i} = 2$ at SNR = 25 dB vs. ρ .

the transmit beamforming strategy achieves best sum rate performance compared to the other methods. It is also found that FlexCoBF with RBD as the transmit beamforming strategy and an MMSE receiver as the receive beamforming strategy is more sensitive to the spatial correlation compared to FlexCoBF with RBD as the transmit beamforming strategy and an MRC receiver as the receive beamforming strategy.

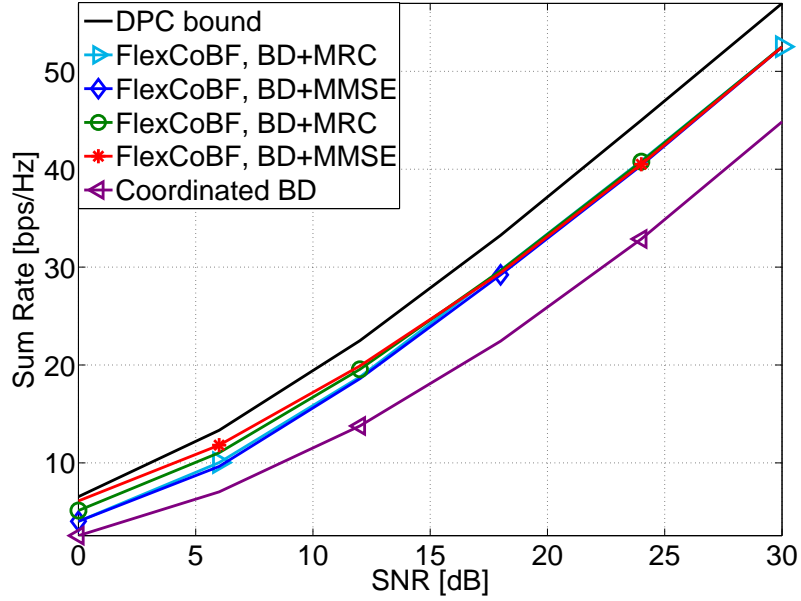


Figure 4.14.: Achievable sum rate for $M_T = 6$, $K = M_{R_i} = 3$ and $\rho = 0$. The number of data streams per user is $\{2, 2, 2\}$.

4.4.2. FlexCoBF in clustered cellular MIMO network

4.4.2.1. Clustered cellular MIMO network

We consider a cluster based cellular MIMO network where the universal frequency reuse is applied. The network is divided into a number of disjoint clusters. Each cluster contains a group of adjacent cells as shown in Figure 4.17 where each cluster consists of 7 cells.

We assume that the BSs within a cluster can fully share CSI and data of all the users in this cluster and are perfectly synchronized in time and frequency. The BSs in different clusters can exchange information such as user data or CSI. Therefore, a full cooperation is included within a cluster and limited cooperation is introduced between adjacent clusters. To efficiently accommodate all the users, we group them into two classes:

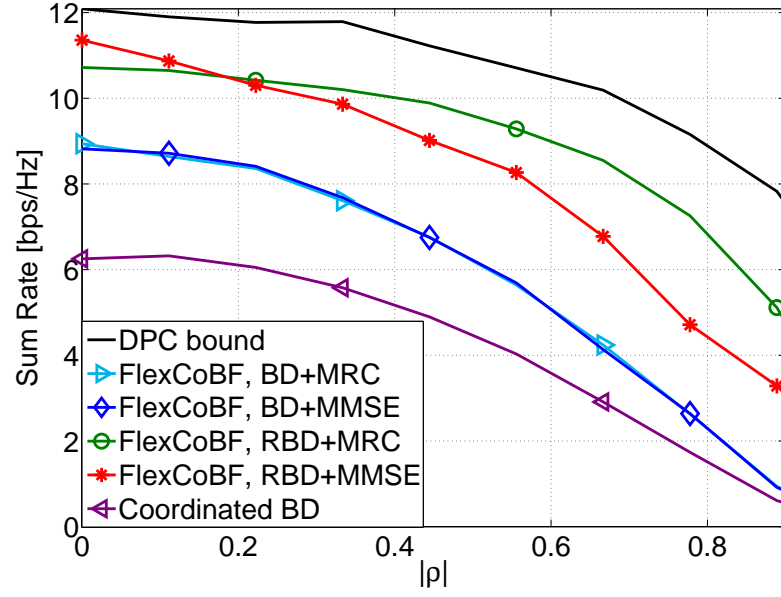


Figure 4.15.: Achievable sum rate for $M_T = 6$ and $K = M_{R_i} = 3$ at $\text{SNR} = 5$ dB v.s. ρ . The number of data streams per user is $\{2, 2, 2\}$.

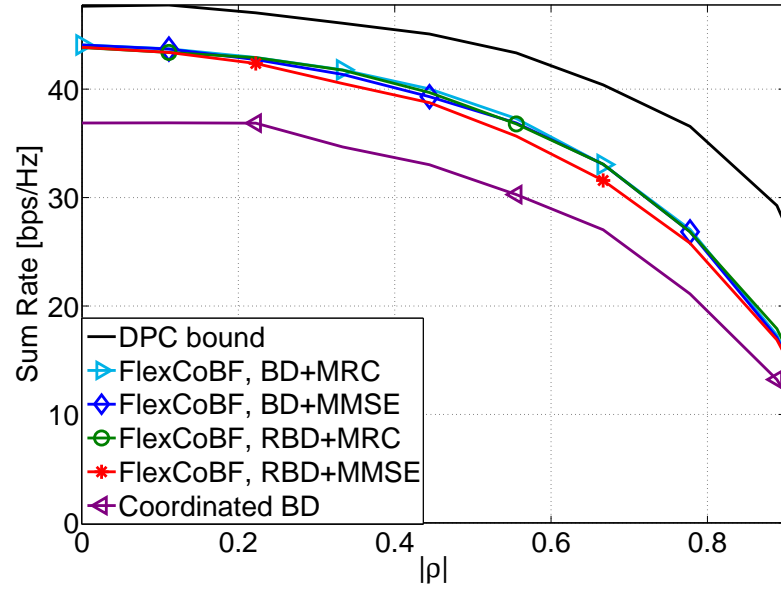


Figure 4.16.: Achievable sum rate for $M_T = 6$ and $K = M_{R_i} = 3$ at $\text{SNR} = 25$ dB v.s. ρ . The number of data streams per user is $\{2, 2, 2\}$.

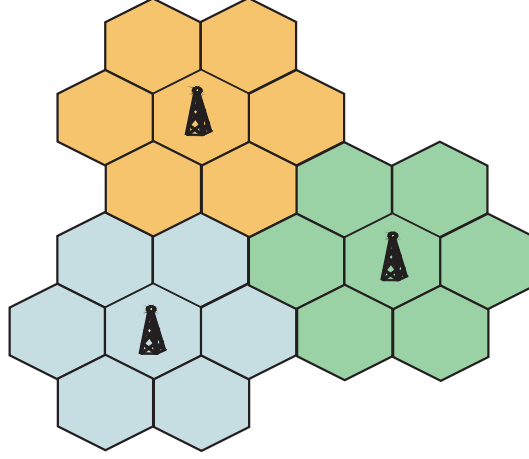


Figure 4.17.: Clustered cellular scenario. In each cluster there is a virtual controller due to the full BSs coordination within each cluster.

- * *Cluster interior users*: they are in the center of the cluster and well protected from the inter-cluster interference by path loss. The BSs within the same cluster work together as a (virtual) super BS to serve the interior users with FlexCoBF. Therefore, the intra-cluster interference is efficiently mitigated for these users.
- * *Cluster edge users*: they are located at the cluster boundary and experience strong inter-cluster interference. The cluster which the edge users belong to is named *home cluster*. If a cluster edge user involves multiple neighboring clusters in a cooperation, the data and CSI information of this user are shared between these clusters which are named as *help clusters*. These help clusters transmit the same data streams to this edge user in order to assist the transmission and take this user into consideration when implementing FlexCoBF. In this way, the intra-cluster interference caused by the home cluster of this edge user and the inter-cluster interference from the neighboring help clusters will be efficiently suppressed.

For simplicity, we use a two-cluster configuration as an example to show the system and receive signal models, which is also considered in the following section to simplify the explanation of the extended FlexCoBF algorithm.

In this example, there are two adjacent clusters. Each of them has two users. Due to the full cooperation within a cluster, each cluster is equivalent to a multi-user MIMO system with M_T transmit antennas. There is a total of four users u_i ($i = 1, \dots, 4$), where the i th user has M_{R_i} receive antennas. As shown in Figure 4.18, the users u_1 and u_2 belong to cluster 1 and

the users u_3 and u_4 are in cluster 2. We assume that the users u_1 and u_4 are cluster interior users. The users u_2 and u_3 are located at the cluster edge and experience a low signal-to-interference ratio (SIR) due to the strong ICI. Limited inter-cluster cooperation is introduced between the adjacent clusters. Let $\mathbf{s}_i \in \mathbb{C}^{r_i}$ denote the transmitted signal for the i th user.

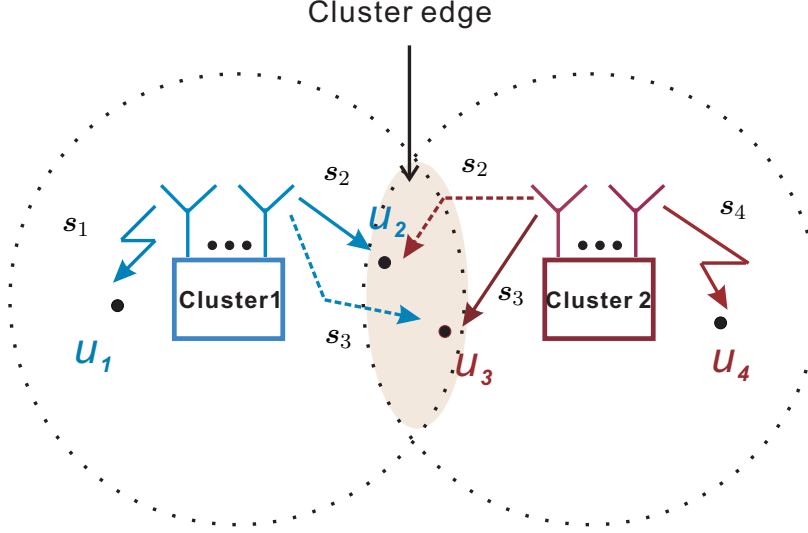


Figure 4.18.: Two adjacent clusters with limited cooperation

The two clusters simultaneously transmit the same signal \mathbf{s}_i to the cluster edge user u_i jointly, while transmitting different signals for cluster interior user independently. Since each cluster serves its own interior user and all the cluster edge users simultaneously, it is reasonable to assume that the number of transmit antennas M_T is smaller than the total number of receive antennas of all users served by this cluster. We use $\mathbf{F}_{j,i} \in \mathbb{C}^{M_T \times r_i}$ to indicate the transmit beamforming matrix within the j th cluster for user i . The receive beamforming matrix for user i is denoted by $\mathbf{W}_i \in \mathbb{C}^{M_{R_i} \times r_i}$. Here, the variable r_i represents the number of data streams to user i . We denote the flat fading MIMO channel between the j th cluster and the i th user by $\mathbf{H}_{j,i} \in \mathbb{C}^{M_{R_i} \times M_T}$ and each user observes zero mean circularly symmetric complex Gaussian white noise $\mathbf{n}_i \in \mathbb{C}^{M_{R_i}}$ with variance σ_n^2 . Then, as an example, the received signals of the cluster interior user u_1 and the cluster edge user u_2 are given as follows: ³

$$\mathbf{y}_1 = \mathbf{W}_1^H \mathbf{H}_{1,1} \mathbf{F}_{1,1} \mathbf{s}_1 + \mathbf{W}_1^H \mathbf{H}_{1,1} \sum_{\ell=2,3} \mathbf{F}_{1,\ell} \mathbf{s}_\ell + \mathbf{W}_1^H \mathbf{n}_1 \quad (4.57)$$

³We do not consider the interference caused by the cluster interior user u_4 to u_1 , since it is neglectable compared to the interference caused by users u_2 and u_3 . Furthermore, we consider that the cluster edge users receive the same data streams (not different data streams) from the help clusters due to the dimensionality constraint on $\mathbf{H}_e \in \mathbb{C}^{r \times M_T}$ in equation (4.52) where $r \leq M_T$.

and

$$\mathbf{y}_2 = \mathbf{W}_2^H \left(\sum_{j=1,2} \mathbf{H}_{j,2} \mathbf{F}_{j,2} \right) \mathbf{s}_2 + \mathbf{W}_2^H \mathbf{H}_{1,2} \sum_{\ell=1,3} \mathbf{F}_{1,\ell} \mathbf{s}_\ell + \mathbf{W}_2^H \mathbf{H}_{2,2} \sum_{k=3,4} \mathbf{F}_{2,k} \mathbf{s}_k + \mathbf{W}_2^H \mathbf{n}_2. \quad (4.58)$$

The cluster interior user u_1 only experiences intra-cell interference caused by the other users served simultaneously by the first cluster. The cluster edge user u_2 suffers interference from both clusters. The first term of the right-hand-side (RHS) of equation (4.58) is the desired signals for the user u_2 . The terms $\mathbf{W}_2^H \mathbf{H}_{1,2} \sum_{\ell=1,3} \mathbf{F}_{1,\ell} \mathbf{s}_\ell$ and $\mathbf{W}_2^H \mathbf{H}_{2,2} \sum_{k=3,4} \mathbf{F}_{2,k} \mathbf{s}_k$ indicate the interference signals. To remove all interference in the scenario, we introduce the extended FlexCoBF algorithm which is described in the following section.

4.4.2.2. The extended FlexCoBF algorithm

Compared to the original FlexCoBF algorithm, we change the updating process of the cluster edge users' receive beamformers at each iteration. Here, we explain the algorithm for the case of two adjacent clusters as depicted in Figure 4.18 for simplicity. Actually, it can be utilized for an arbitrary number of adjacent clusters. The extended FlexCoBF algorithm includes the joint design of the transmit-receive beamformers of all users in their home clusters by using FlexCoBF in order to mitigate the intra-cluster interference, while updating the receive beamformers of the cluster edge users with limited cooperations between the help clusters to mitigate the inter-cluster interference. There is no direct interaction among the transmit beamformers of the home clusters and the help clusters. An example of the extended FlexCoBF algorithm at the p th iteration is illustrated in Figure 4.19, where the matrix $\mathbf{F}_j^{(p)}$ is a combined transmit beamforming matrix for all users in the j th cluster. The terms $\mathbf{W}_1^{(p)}$ and $\mathbf{W}_4^{(p)}$ denote the receive beamformers of the cluster interior users u_1 and u_4 , respectively, which are obtained directly from FlexCoBF at the p th iteration. For cluster edge users, we use $\mathbf{W}_{j,i}^{(p)}$ to indicate the precoded channel $\mathbf{H}_{j,i} \mathbf{F}_{j,i}^{(p)}$ from cluster j to user i (i.e., $\mathbf{W}_{j,i}^{(p)} = \mathbf{H}_{j,i} \mathbf{F}_{j,i}^{(p)}$). After information exchange between the cooperating clusters, the equivalent precoded channel of the i th cluster edge user $\bar{\mathbf{H}}_i^{(p)}$ is computed as the sum of the terms $\mathbf{W}_{j,i}^{(p)}$ ($j = 1, 2$), since the i th cluster edge user receives the same signal from the adjacent clusters. Then, the final receive beamformers of the cluster edge users at the p th iteration are obtained according to the desired receiver strategy using the equivalent precoded channel. Consequently, the receive beamformers of all users at the p th iteration are incorporated into the $(p + 1)$ th iteration of FlexCoBF for both clusters. After convergence has been achieved, all the interference in the scenario is reduced.

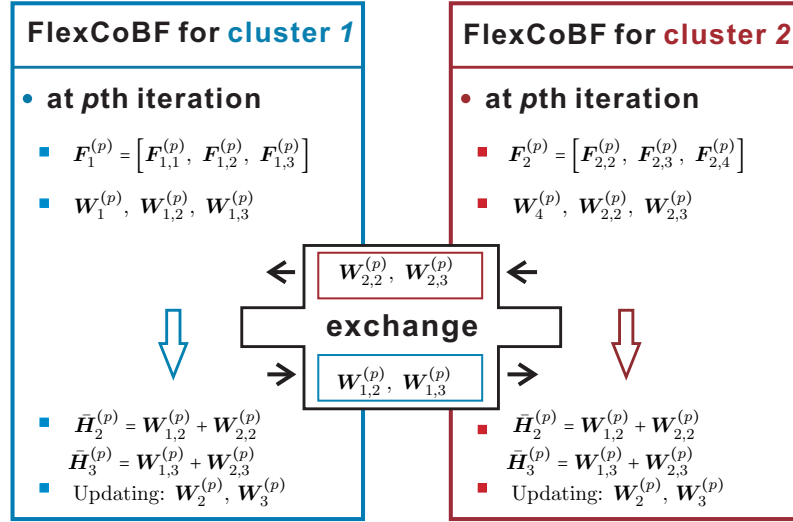


Figure 4.19.: Block diagram of extended FlexCoBF for two adjacent clusters.

Based on the above description, we summarize the extended FlexCoBF algorithm for the clustered multi-cell scenario as follows. We assume K users in each cluster where L ($L < K$) users are cluster interior users and the remaining $K - L$ users are cluster edge users. There are N_c adjacent clusters, which experience interference from each other and are involved for limited cooperation.

Initialization:

For the j th cluster, set $\mathbf{W}_i^{(0)}$ ($i = 1, \dots, K$) to some random matrices, the iteration index p to zero, and set the threshold ϵ .

Iteration:

1. Set $p \leftarrow p + 1$ and compute $\mathbf{H}_e^{(p)}$ as

$$\mathbf{H}_e^{(p)} = \begin{bmatrix} \mathbf{W}_1^{(p-1)\text{H}} \mathbf{H}_{j,1} \\ \mathbf{W}_2^{(p-1)\text{H}} \mathbf{H}_{j,2} \\ \vdots \\ \mathbf{W}_K^{(p-1)\text{H}} \mathbf{H}_{j,K} \end{bmatrix}$$

2. Apply the desired linear precoding algorithm on the matrix $\mathbf{H}_e^{(p)}$ to obtain the transmit beamforming matrices $\mathbf{F}_{j,i}^{(p)}$ for all users (i.e., $i = 1, \dots, K$).
3. Compute the receive beamformers for the cluster interior users according to the desired receive strategy. For instance,

- MRC receiver:

$$\mathbf{W}_i^{(p)} = \mathbf{H}_{j,i} \mathbf{F}_{j,i}^{(p)}, \quad i = 1, \dots, L$$

- MMSE receiver⁴:

$$\mathbf{W}_i^{(p)} = (\mathbf{H}_{j,i} \mathbf{F}_{j,i}^{(p)} \mathbf{F}_{j,i}^{(p)\text{H}} \mathbf{H}_{j,i}^{\text{H}} + \sigma_n^2 \mathbf{I}_{M_{R_i}})^{-1} \mathbf{H}_{j,i} \mathbf{F}_{j,i}^{(p)}, \quad i = 1, \dots, L$$

4. Compute the receive beamformers for the cluster edge users.

- a) Assign the receive beamformers for cluster edge users as the precoded channel $\mathbf{H}_{j,i} \mathbf{F}_{j,i}^{(p)}$

$$\mathbf{W}_{j,i}^{(p)} = \mathbf{H}_{j,i} \mathbf{F}_{j,i}^{(p)}, \quad i = L + 1, \dots, K$$

- b) Compute the equivalent precoded channel of the cluster edge users by considering adjacent interfering clusters.

$$\begin{aligned} \bar{\mathbf{H}}_i^{(p)} &= \sum_{\ell=1}^{N_c} \mathbf{W}_{\ell,i}^{(p)} \\ &= \sum_{\ell=1}^{N_c} \mathbf{H}_{\ell,i} \mathbf{F}_{\ell,i}^{(p)}, \quad i = L + 1, \dots, K \end{aligned}$$

- c) Update the receive beamformers for the p th iteration according to the desired strategy using the equivalent precoded channel.

⁴In the MMSE receiver structure, $\sigma_n^2 \mathbf{I}_{M_{R_i}}$ can be replaced by an interference plus noise covariance matrix if it can be estimated in the system. Otherwise σ_n^2 is the estimated background noise level.

- MRC receiver:

$$\mathbf{W}_i^{(p)} = \bar{\mathbf{H}}_i^{(p)}, \quad i = L + 1, \dots, K$$

- MMSE receiver⁴:

$$\mathbf{W}_i^{(p)} = (\bar{\mathbf{H}}_i^{(p)} \bar{\mathbf{H}}_i^{(p)\text{H}} + \sigma_n^2 \mathbf{I}_{M_{R_i}})^{-1} \bar{\mathbf{H}}_i^{(p)}, \quad i = L + 1, \dots, K$$

5. Compute the stopping criterion of FlexCoBF. If it is not satisfied, go back to step 2. Otherwise, convergence is achieved and the procedure has ended.

4.4.2.3. User grouping

In a home cluster, users can be grouped into cluster interior and cluster edge users according to the user locations, the path-loss, and the average signal strength. For example, a coordination distance D_c is proposed in [ZCA08]. If the distance of the user to the cluster edge is not larger than D_c , this user is classified as a cluster edge user. Otherwise, it is a cluster interior user.

We group the users based on the user position. Let us define d_1 as the distance of a user to the center of its home cluster, while using d_2 to indicate the distance of this user to one of its adjacent cluster. If the ratio of these two distances (i.e., $\frac{d_2}{d_1}$) is smaller than a threshold ξ , this user is classified as a cluster edge user. Meanwhile, the adjacent cluster related to the distance d_2 is considered for the limited cooperation and becomes a help cluster. In general, a user achieves large ratios of $\frac{d_2}{d_1}$, if this user is allocated close to the center of its home cluster (i.e., small d_1). In contrast, a user, near to the cluster edge, always can find a help cluster which satisfies $\frac{d_2}{d_1} < \xi$. Compared to the user grouping algorithm in [ZCA08], the method we proposed can simultaneously classify the users and determine the help clusters.

Actually there is a trade-off when choosing the threshold ξ . If ξ is too small (e.g. $\xi = 1$), all users in the home cluster will be treated as cluster interior users, which is equivalent to the case that no cooperation is employed between the adjacent clusters. If ξ is too high (e.g., $\xi = 3$), even the users closed to the center of the home cluster will be treated as cluster edge users and enjoy the limited cooperation with the adjacent clusters. Thus, the total number of active users will be reduced.

In order to properly choose ξ , we investigate a utility function of ξ (i.e., $U(\xi)$), which consists of the mean minimum signal to interference plus noise ratio (SINR) (i.e., $\overline{\text{SINR}}_{\min}(\xi)$) and the effective sum rate (i.e., $R_{\text{eff}}(\xi)$).

* *Mean minimum SINR*: For a given ξ and one realization of user location, the minimum SINR among all users in the home cluster is denoted as $\text{SINR}_{\min}(\xi)$. The mean minimum

SINR is the average value of $\text{SINR}_{\min}(\xi)$ over different channel realizations. $\overline{\text{SINR}}_{\min}(\xi)$ is mainly determined by the cluster edge users and increases as ξ increases.

- * *Effective sum rate*: For a given ξ and given user locations, if there are N_c clusters serving user i simultaneously by transmitting the same data streams, the effective rate of this user is given by R_i/N_c , where R_i is the throughput achieved by user i and calculated by

$$R_i = \sum_{\ell=1}^{r_i} \log_2(1 + \text{SINR}_{\ell}). \quad (4.59)$$

The term r_i indicates the number of data streams for user i . Then, the effective sum rate of the home cluster is defined as

$$R_{\text{eff}}(\xi) = \sum_{i=1}^K \frac{R_i}{N_c(\xi)}. \quad (4.60)$$

Here, $N_c(\xi)$ indicates that the number of the clusters involved the limited cooperation for the user i and it changes with the parameter ξ . The effective sum rate $R_{\text{eff}}(\xi)$ decreases with the increase of ξ , since more users are treated as cluster edge user with increasing ξ .

The mean minimum SINR and the effective sum rate lead to the opposing objectives with respective to ξ . Therefore, we can further introduce a variable α to reflect the design objective and define the utility function as

$$U(\xi) = \alpha \frac{\overline{\text{SINR}}_{\min}(\xi)}{\max_{\xi} \overline{\text{SINR}}_{\min}(\xi)} + (1 - \alpha) \frac{R_{\text{eff}}(\xi)}{\max_{\xi} R_{\text{eff}}(\xi)}, \quad 0 \leq \alpha \leq 1. \quad (4.61)$$

If edge users are more valuable to care about, α can set close to 1. If the sum rate is more important, α is picked close to 0. This utility function can be interpreted as that the mean minimum SINR (i.e., $\overline{\text{SINR}}_{\min}(\xi)$) and the effective sum rate (i.e., $R_{\text{eff}}(\xi)$) are computed with a given ξ , given user locations, and different channel realizations. Then, they are further normalized by the maximum values of $\overline{\text{SINR}}_{\min}(\xi)$ and $R_{\text{eff}}(\xi)$ over all possible values of ξ (e.g., $1 \leq \xi \leq 3$ in simulations), respectively. Therefore, the utility function is greater than 0 and smaller or equal to 1.

4.4.2.4. Cluster size

The size of the cluster has an impact on the sum rates, the degrees of freedom of the system, and the complexity of the full cooperation within one cluster. For example, with a fixed ratio

ξ , a small cluster size results in too many cluster edge users which will consume lots of degrees of freedom and result in low effective sum rates. Therefore, it is important to select a suitable cluster size for a practical system. In this work, we investigate the suitable cluster size by considering the sum rates for different cluster sizes. It will be demonstrated in the simulation results in Section 4.4.2.8.2 that there is a diminishing gain with the increase of the cluster size.

4.4.2.5. Exchange mechanism for clusters with limited cooperation

Once the users in one cluster have been classified into cluster interior and cluster edge users, the help clusters which should be considered for limited cooperation are determined as well. Here, we assume that N_c clusters have been involved for limited cooperation. We propose two alternative exchange mechanisms for the implementation of the extended FlexCoBF algorithm for one fading block where the channel is considered constant.

1. Online exchange mechanism:

All N_c clusters implement FlexCoBF simultaneously. But each cluster implements FlexCoBF only for its own, while requiring information about $\mathbf{W}_{j,i}^{(p)}$ from the other $(N_c - 1)$ clusters for the cluster edge user i at the p th iteration of FlexCoBF. Here, j indicates the help clusters (i.e., $j = 1, \dots, N_c - 1$) and p enumerates the iterations. To this end, all N_c clusters should exchange $\mathbf{W}_{j,i}^{(p)}$ at each iteration of FlexCoBF. Once the stopping criterion of FlexCoBF is fulfilled in one cluster, the information exchange can end as well. In this case, the transmission delay between the N_c clusters is the predominant factor limiting the time required to calculate the beamforming weights.

2. Offline exchange mechanism:

This mechanism is proposed to avoid the information exchange among the N_c clusters at each iteration of FlexCoBF and to combat the effect of the transmission delay. Instead of exchanging information about $\mathbf{W}_{j,i}^{(p)}$ for cluster edge users, the CSI knowledge of all the users in the other $(N_c - 1)$ clusters is required for each cluster. With the users' CSI, each cluster can implement FlexCoBF not only for its own, but also for the other $(N_c - 1)$ clusters simultaneously. Consequently, the information about $\mathbf{W}_{j,i}^{(p)}$ is acquired within each cluster without information exchange among the N_c clusters at each iteration of FlexCoBF. In this case, the implementation complexity of each cluster is increased by the number of clusters involved in the limited cooperation.

4.4.2.6. Detection method for receive beamforming

For FlexCoBF in a single cell or the extended FlexCoBF in clustered multiple cells, the precoded channel information (i.e., $\mathbf{H}_i \mathbf{F}_i$) is required for each user to calculate the receive beamformers. In a system where dedicated pilots per user are used for the downlink, each user can estimate its own precoded channel information using least squares (LS) or MMSE channel estimation. For systems which use only a common pilot channel for all users (e.g., 3GPP long term evolution), there is no way to estimate the precoded channel at the receiver. To solve this problem, in [CMIH08a] and [CMIH08b] a limited feedforward method has been proposed to inform the users about the receive beamformers. In our work, we assume in the simulations that each user has perfect knowledge of its precoded channel for simplicity.

4.4.2.7. Simulation results for two adjacent clusters

We consider two adjacent clusters shown in Figure 4.18 firstly for simplicity and evaluate the performance of the extended FlexCoBF algorithm in terms of the system sum rate. The achievable system sum rate is calculated as $R = \sum_{\ell=1}^r \log_2(1 + \text{SINR}_\ell)$, where SINR_ℓ indicates the signal-to-interference-plus-noise ratio of the ℓ th data stream and r is the total number of different data streams. We take each cluster as a virtual multi-user MIMO broadcast channel and assume that the transmit power of each cluster is P_T . The system signal-to-noise ratio is defined as $\text{SNR} = P_T/\sigma_n^2$. There are two users in each cluster in which one user is assumed to be an interior user and another is a cluster edge user. The exact position of the users are not known, we only assume that the path-loss of the cluster edge user is 10 times larger than the path-loss of the cluster interior users. The threshold ϵ for the stopping criterion is set to 10^{-5} and the maximum number of iterations is limited by 50. For comparison, we consider cellular scenarios with three different cooperation strategies.

1. FlexCoBF without cooperation between clusters

Each cluster only serves the users on its own and does not consider any cooperation with adjacent clusters. As a result, the cluster edge users suffer from strong inter-cluster interference. The intra-cluster interference among the users in one cluster is mitigated by FlexCoBF performed for one cluster.

2. FlexCoBF with full cooperation in cellular scenarios

Full cooperation among all BSs in cellular scenarios is considered. The channel state information (CSI) and transmit data of all users are shared among the cooperating BSs. Consequently, the cellular scenario acts as a virtual single cell multi-user MIMO system.

Therefore, the interference among users can be canceled by FlexCoBF.

3. Extended FlexCoBF with limited Cooperation between Clusters

Only the transmit data of the cluster edge users are shared among the home cluster and the help clusters due to the limited cooperation. Online or offline exchange mechanisms can be used to perform the extended FlexCoBF algorithm. We assume that each cluster acquires perfect knowledge of the exchanged information (i.e., CSIs of all users or pre-coded channels between cluster edge users and the involved help clusters). As a result, all the significant interference can be canceled.

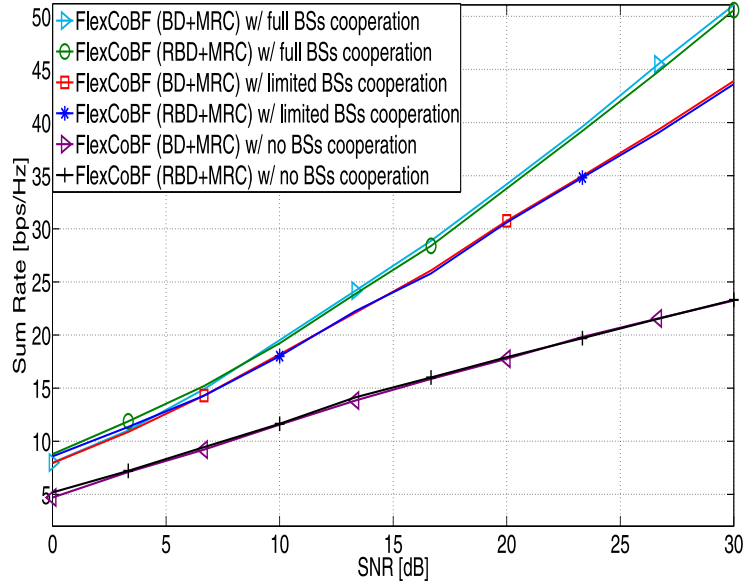


Figure 4.20.: Achievable sum rate for two-cluster scenario with $M_T = 3$, $M_{R_i} = 2$, and $\rho = 0$.

Figure 4.20 shows the sum rate performance of this two-cluster scenario with different cooperation strategies. Each cluster has three transmit antennas and each user is equipped with two receive antennas. To utilize all spatial degrees of the freedom supported by one cluster, each cluster simultaneously transmits three data streams. For FlexCoBF without cooperation between clusters and FlexCoBF with full cooperation in cellular scenarios, two data streams are intended to each cluster edge user and one data stream is for each cluster interior user. For extended FlexCoBF with limited cooperation between clusters, each cluster serves three users simultaneously and each user has one data stream. Notice that in this case each cluster edge user jointly receives the same data stream from two adjacent clusters. Therefore, there is a total of 6 data streams, but only 4 different data streams. Obviously, FlexCoBF

with full cooperation in cellular scenario has the best sum rate performance, because of the sharing of all users' transmit data and CSI. FlexCoBF without cooperation between clusters offers the worst sum rate performance because of the strong inter-cluster interference at the cluster edge users. There is a performance gap between the extended FlexCoBF with limited cooperation and FlexCoBF with full cooperation. This performance gap is caused by the fact that only the information about the receive beamformers $\mathbf{W}_{j,i}^{(p)}$ (also called precoded channels since $\mathbf{W}_{j,i}^{(p)} = \mathbf{H}_{j,i}\mathbf{F}_{j,i}^{(p)}$) of the cluster edge users are exchanged between BSs due to the limited cooperation. The precoding matrices corresponding to the joint transmission to the cluster edge users are not fully coordinated with the receive beamformers. However, the achievable sum rate of the extended FlexCoBF with limited cooperation approaches FlexCoBF with full cooperation at low SNRs due to the introduced multi-cluster diversity, which comes from a situation where several clusters transmit the same signal to a cluster edge user. In Figure

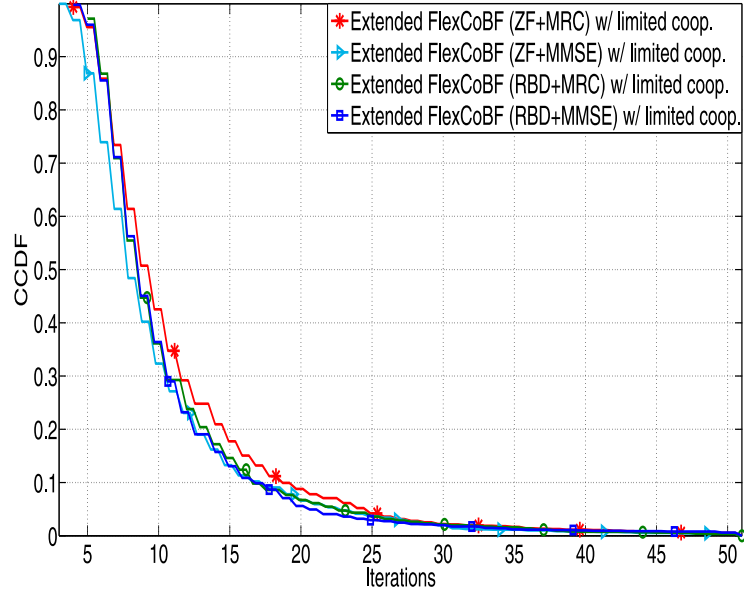


Figure 4.21.: CCDF of required number of iterations for extend FlexCoBF in two-cluster scenario over all SNRs (i.e., 0 dB - 30 dB), stopping criterion I for extended FlexCoBF with ZF and stopping criterion II for extended FlexCoBF with RBD.

4.21, we evaluate the convergence of extended FlexCoBF for the two-cluster scenario. We have observed that the extended FlexCoBF algorithm always converges in the simulations and most of the trials require less than 10 iterations to achieve the convergence.

4.4.2.8. Simulation results with exact user positions

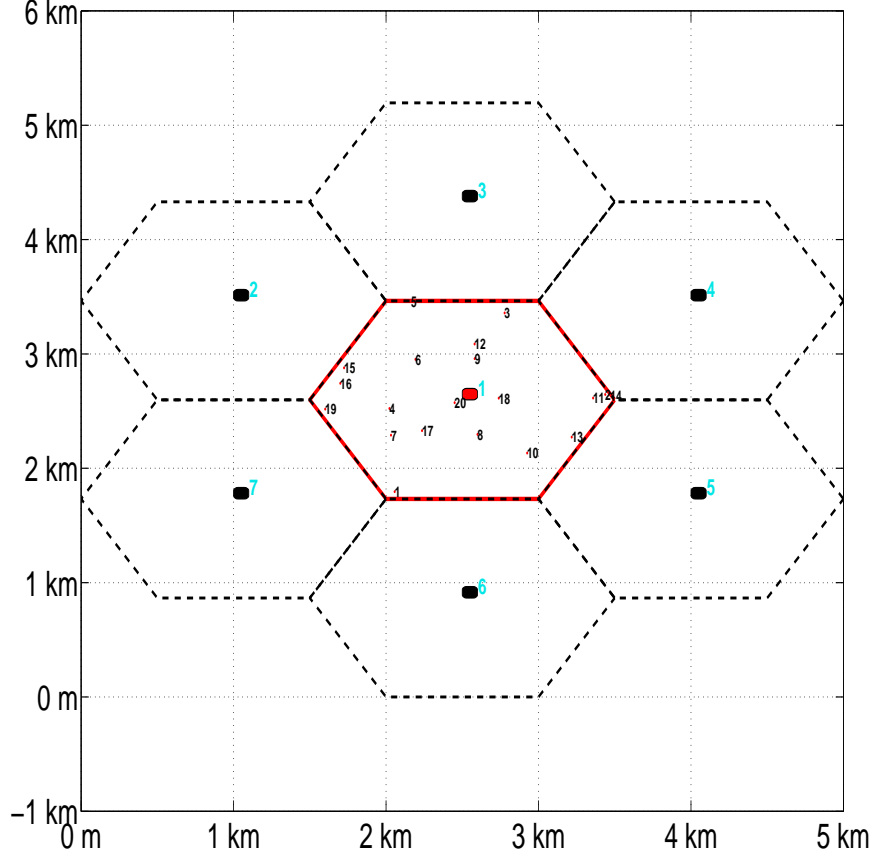


Figure 4.22.: A clustered cellular network with 7 adjacent clusters where each cluster has one cell as an example.

We consider a more realistic scenario as shown in Figure 4.22. The cluster 1 is assumed to be the home cluster. There are 6 adjacent clusters around. Users are uniformly distributed in each cluster and each user has 2 receive antennas. Here, one realization of user locations in cluster 1 is shown as an example. In the middle of each cluster is the virtual controller, which perform the full cooperation within each cluster. There is N_r cells in each cluster, where each cell is represented by a hexagon structure with side length of 1000 m. There is one BS per cell and each BS is equipped with 4 transmit antennas. The maximum transmit power of each BS is $P_{BS} = 1$ dBm. Thus, the total transmit power of one cluster is $P_T = N_r \cdot P_{BS}$. The channel

models of the users include Rayleigh fading, path loss (corresponding to a carrier frequency of 2 GHz), and shadowing. The standard deviation of the shadowing is 8 dB, the path loss exponent is 4, and the reference distance d_0 for the antenna far field is 100 m. The extended FlexCoBF employing ZF and MRC is considered due to its low complexity.

4.4.2.8.1. User grouping with ξ

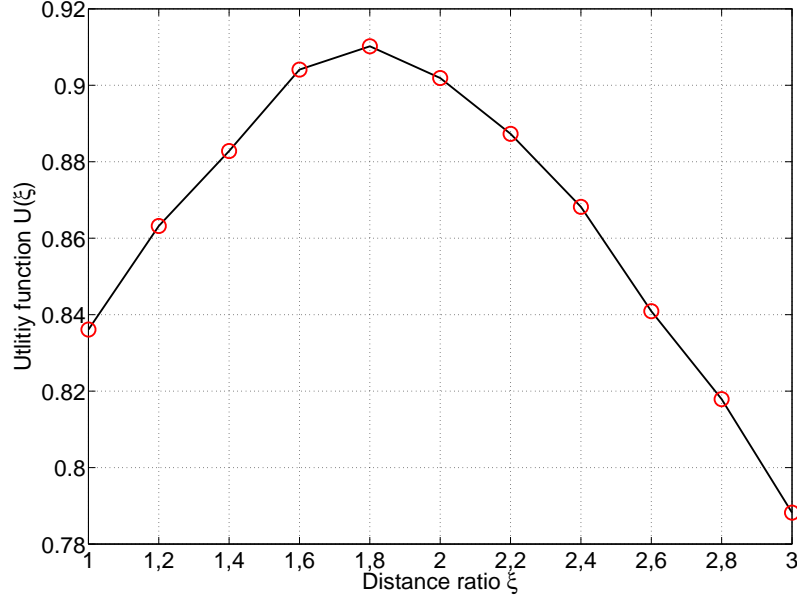


Figure 4.23.: Utility function $U(\xi)$ for different ξ .

Figure 4.23 shows the resulting values for the utility function $U(\xi)$. We assume that the cluster has the size of $N_r = 3$. There are 500 realizations of the user locations. For each user realization $K = 12$ and 500 independent channel realizations are simulated. We treat the changes of $\overline{\text{SINR}}_{\min}(\xi)$ and $R_{\text{eff}}(\xi)$ as of equal value, which means $\alpha = 0.5$. It is found that the maximum value of $U(\xi)$ is achieved around $\xi = 1.8$, which is a proper choice for the other simulations in this subsection.

4.4.2.8.2. Cluster size

With the fixed $\xi = 1.8$ and $K = 50$ per cluster, we have investigated the sum rate performance of the home cluster for different cluster sizes (i.e., $N_r = 1, 3, 5, 7, 10$) by considering limited cooperation. The round-robin user scheduling is applied. There are 1000 trials for each cluster

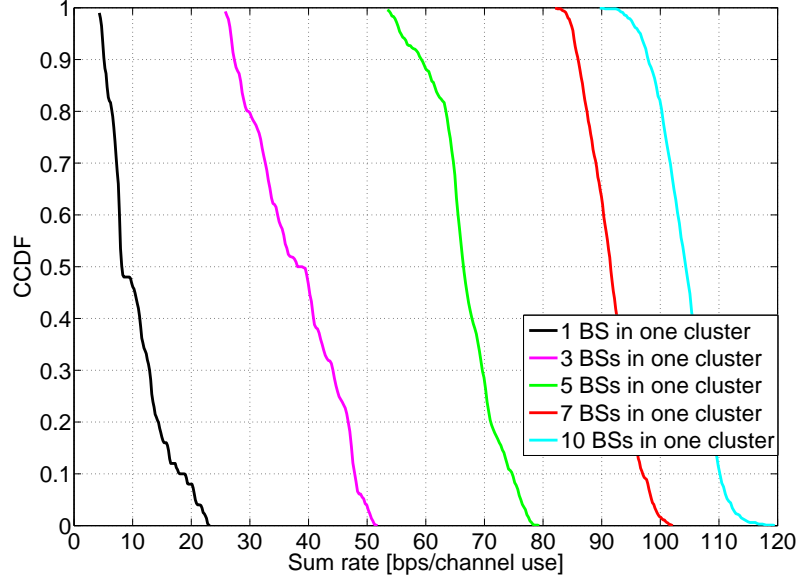


Figure 4.24.: Sum rates of the home cluster with different cluster size.

size. For each trial 500 independent channel realization are simulated. Figure 4.24 shows that there is a diminishing gain with the increasing cluster size. The achievable sum rate improvement for the 3-cell cluster is around 28 bps/channel relative to the one cell cluster. Similarly, around 27 bps/channel improvement is achieved for the 5-cell cluster compared to the 3-cell cluster. From the 5-cell cluster to the 7-cell cluster and the 7-cell cluster to the 10-cell cluster the achievable sum rate improvements are around 19 bps/channel and 13 bps/channel, respectively. Therefore, we can conclude that until $N_r = 7$ a significant sum rate performance gain can be achieved for the clustered cellular scenario.

4.4.2.8.3. Sum rates performance

With the fixed $\xi = 1.8$ and the cluster size $N_r = 3$, we investigate the sum rate performance of the home cluster which is around by 6 adjacent clusters as shown in Figure 4.22. Three different cooperation strategies are considered. They are FlexCoBF without cooperation between clusters, FlexCoBF with full cooperation in cellular scenarios, and the extended FlexCoBF with the limited cooperation between clusters defined in Section 4.4.2.7. There are $K = 12$ users in the scenario. Based on the exact user locations, 12 users are grouped into cluster interior users and cluster edge users by applying the user grouping strategy proposed in Section 4.4.2.3. Table 4.2 shows the grouping results and the involved help clusters with respect to

the cluster edge users. 1000 trials are simulated for each cooperation strategy. For each trial 100 independent channel realizations are run.

Table 4.2.: User grouping and the involved help clusters

User index	Cluster interior user	Cluster edge user	Help clusters
1		x	cluster 4 and cluster 5
2		x	cluster 5
3		x	cluster 2 and cluster 7
4	x		
5	x		
6		x	cluster 2 and cluster 7
7		x	cluster 2 and cluster 7
8	x		
9	x		
10	x		
11	x		
12	x		

Figure 4.25 shows the sum rate performance of the home cluster. Obviously, FlexCoBF without cooperation between clusters provides the worst sum rate due to the strong inter-cluster interference at the cluster edge users. The performance gain introduced by the extended FlexCoBF with limited cooperation is significant. The performance gap between the extended FlexCoBF with limited cooperation and FlexCoBF with full cooperation has been caused by the fact that only the information about the receive beamformers $\mathbf{W}_{j,i}^{(p)}$ (also namely precoded channels since $\mathbf{W}_{j,i}^{(p)} = \mathbf{H}_{j,i} \mathbf{F}_{j,i}^{(p)}$) of the cluster edge users are exchanged between BSs due to the limited cooperation. The precoding matrices corresponding to the joint transmission of the cluster edge users are not fully coordinated with the receive beamformers.

that the calculations of the receiver beamformers for the cluster edge users are based on the equivalent precoded channels (see Section 4.4.2.2) instead of their own precoded channels due to the limited cooperation.

The significant performance gains by introducing the extended FlexCoBF with limited cooperation are evidently observed from Figures 4.26 and 4.27 where only the throughputs of the cluster edge users are shown.

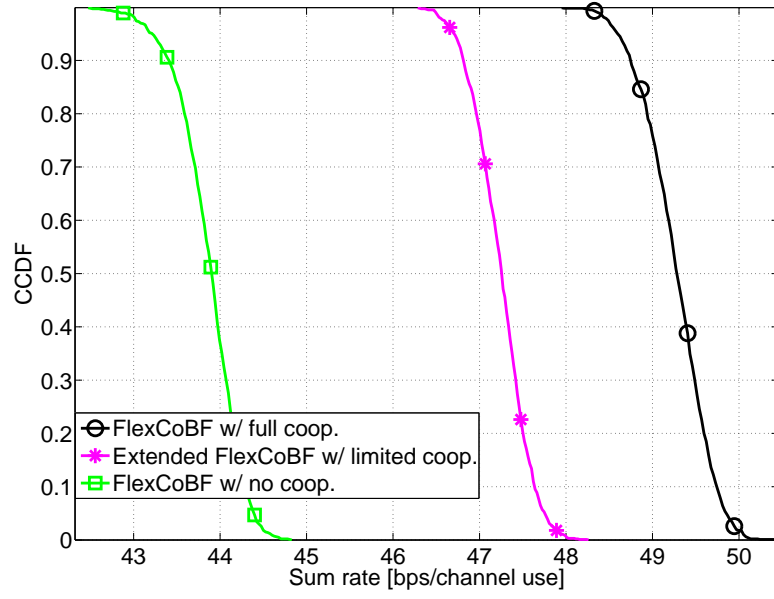


Figure 4.25.: Sum rates of the home cluster with different cooperation strategies.

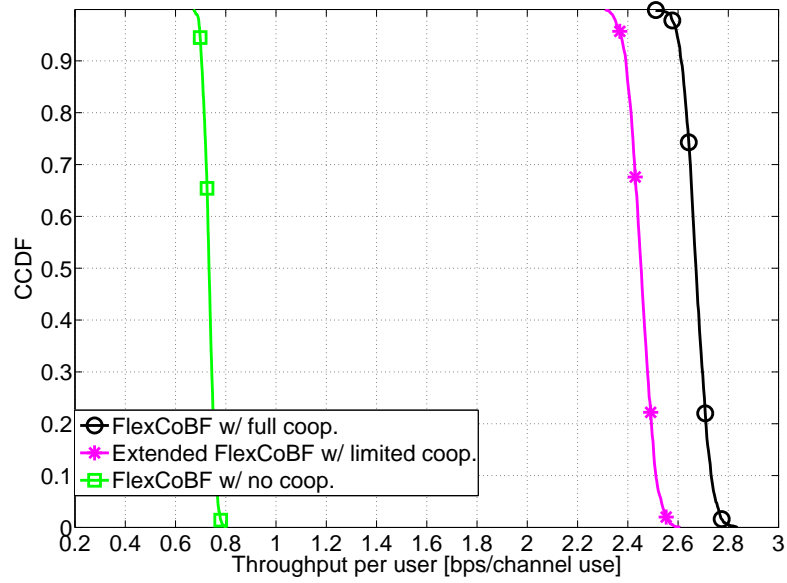


Figure 4.26.: An example of the throughput of user 3 with different cooperation strategies.

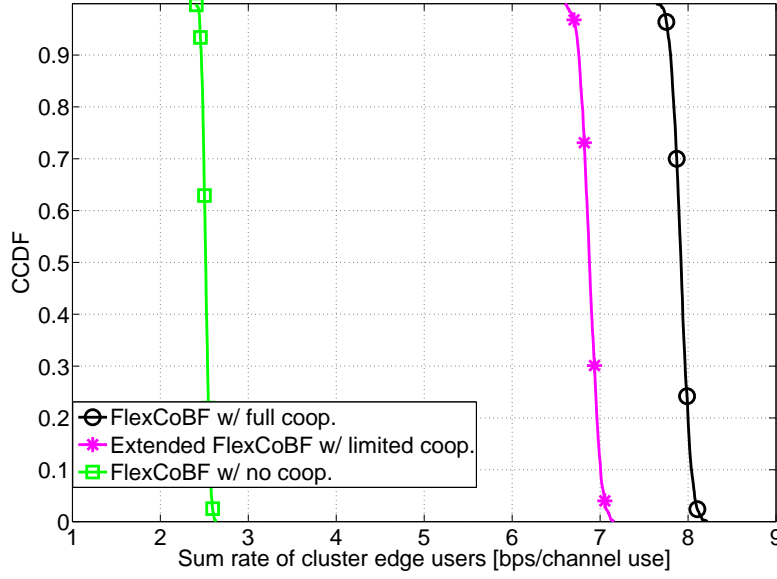


Figure 4.27.: Sum rates of the cluster edge users with different cooperation strategies.

4.5. Summary and Conclusions

In this chapter of the thesis, we have discussed coordinated beamforming (CBF) techniques for the case that the total number of the receive antennas is greater than the number of the transmit antennas in a multi-user MIMO broadcast channel. The SeDJoCo-based closed-form CBF has been proposed by us [SH09b, YSRH12]. Furthermore, we have designed an iterative CBF (FlexCoBF) for the single cell and the clustered multiple cells scenarios. The main novel contributions are:

- * The proposed SeDJoCo transformation solves the problem that seeks a matrix \mathbf{B} for a set of symmetric, positive definite target matrices $\{\mathbf{C}_i\}_{i=1}^N$, such that the i th row and the i th column of the transformed i th target matrix $\mathbf{B}\mathbf{C}_i\mathbf{B}^H$ would be all-zeros, except for the diagonal (i,i) th element. To the best of our knowledge, the solutions of this problem have rarely been addressed in the literature. We have further proposed two solutions for the SeDJoCo transformation. Both of them are based on formulating the problem as different optimization problems and taking different approaches in the maximization or minimization of the associated cost functions. Both approaches are provided in both a real-valued and a complex-valued version.

- An approach based on Newton's method (NCG): Normally, the application of New-

ton's method would require the inversion of an $N^2 \times N^2$ Hessian matrix in each iteration, which might be computationally expensive when N is large. However, by identifying and exploiting the sparsity of the Hessian, we are able to apply the conjugate-gradient method and enjoy the fast (quadratic) convergence of Newton's method at a moderate computational cost per iteration.

- An approach (i.e., STJOCO) based on successive unitary transformations involving multiplications by parameterized lower and upper diagonal matrices: This method offers linear convergence at a reduced computational load per iteration.
- Presentation of different equivalent formulations of the problem and their association with the joint diagonalization problem
- A proof of existence of a solution for positive-definite target-matrices, both in the real-valued and complex-valued cases

- * Based on SeDJoCo, a closed-form CBF has been proposed to avoid iterative updating between the transmit and receive beamformers, while achieving the same sum rate performance as the iterative CBF. Unlike the closed-form CBF in [CMJH08] that is only valid for a downlink system with two users and two transmit antennas, the proposed SeDJoCo-based closed-form CBF can support a multi-user MIMO downlink system with an arbitrary number of users and transmit antennas.
- * The applications of SeDJoCo are not only restricted to the CBF, it can also be utilized for maximum likelihood (ML) blind (or semi-blind) source separation [YSRH12].
- * The proposed iterative CBF (FlexCoBF) does not require special designs of the transmit-receive beamforming weights like the existing CBF algorithms, because FlexCoBF provides freedom in the choice of the linear transmit and receive beamforming strategies.
 - For the transmit beamforming, any existing linear precoding technique can be applied (e.g., ZF, BD, and RBD).
 - The receive beamforming strategy can be chosen flexibly (e.g., MRC or MMSE receivers).
- * FlexCoBF supports the transmission of multiple data streams to each user.
- * Two alternative stopping criteria have been designed for FlexCoBF to achieve a fast convergence. The number of required iterations is significantly smaller than those for the CBF in [CMIH08a], while achieving the same sum rate performance.

- * The extended FlexCoBF algorithm can efficiently mitigate the intra-cluster and inter-cluster interference and significantly increase the average throughput of the cluster and the user throughputs of the cluster edge users by introducing limited cooperations in a clustered cellular scenario.
- * Based on the location of the users, a user grouping strategy has been proposed and investigated to classify the users and determine the help clusters.
- * It has been demonstrated by the simulation results that a significant gain can be realized by constructing small size clusters.

Meanwhile, there are still some shortcomings and challenges of the proposed algorithms.

- * Shortcomings of the solutions of SeDJoCo transformation (i.e., NCG and STJOCO)
 - The NCG algorithm significantly accelerates the convergence, but its initialization is sensitive in the case of a large number of target matrices, which often prohibits proper convergence. Furthermore, NCG-based closed-form CBF leads to a lower sum rate performance compared to STJOCO-based closed-form CBF.
 - STJOCO-based closed-form CBF achieves the same sum rate performance as CBF in [CMIH08a], but the STJOCO algorithm requires a larger number of iterations until convergence to an exact solution compared to NCG algorithm. For a large number of target matrices, the STJOCO algorithm might only converge to a local minimum of its respective cost function.
- * Challenges of performing the limited cooperation in the clustered cellular scenario (i.e., CoMP)
 - *Synchronization*: Downlink MIMO cooperation between multiple BSs requires tight synchronization in frequency and in time to avoid inter-carrier interference and inter-symbol interference. It is feasible today by using commercial GPS (global position system) satellite signals for outdoor BSs and a precisely timed network protocol for indoor BSs. However, the cost of the synchronization should be further reduced.
 - *Channel estimation*: Sufficient resources must be allocated to pilot signals to ensure reliable channel estimation. With the increase of the cooperation size, the expense of additional pilot overhead is increased as well. This problem is more critical for the estimation of channels at the BS in FDD networks than in TDD networks. In TDD networks, the channel estimation on the uplink can be used for downlink

transmission due to the reciprocity of the uplink and downlink channels. In FDD networks, the channel estimates obtained at the receiver must be conveyed to the BS over a bandwidth limited uplink feedback channel. Although quantized channel estimates have been proposed to reduce the feedback bits, the size of codewords will still be increased in a cooperative network. Thus, additional feedback bits might be required to maintain a given rate loss.

- *Backhaul*: The ideal backhaul with unlimited capacity and zero-delay is not feasible. In a practical cooperation network, a backhaul with a high bandwidth and a low latency is required to connect the BSs with each other or with a central processor.
- *Clustering*: In practice, only a limited number of BSs can cooperate due to the overhead management. Therefore, the cell cluster should be set up adaptively according to the RF channel measurements and user positions in order to exploit the cooperative gain at a limited complexity.

5. Channel State Information Acquisition

5.1. Introduction and Motivation

In a multi-user MIMO downlink channel, the goal of the linear transmit and receive strategies (e.g., linear precoding techniques and coordinated beamforming techniques) is to exploit the spatial diversity of the mobile radio channel by using a large number of antennas to serve a large number of users at the same time and on the same frequency resources. To achieve this goal, the base station (BS) requires the channel state information (CSI) of the MIMO channels between all BS antennas and all of the user antennas. In general, the CSI can be acquired at the BS by either invoking the reciprocity principle or using feedback from the users.

The reciprocity principle [JVG01] in wireless communication states that the channel from terminal A to another terminal B is identical to the transpose of the channel from terminal B to terminal A, if both forward and reverse links occur at the same time, the same frequency, the same antenna locations, and the same transmit and receive radio frequency (RF) chains. Therefore, the BS might obtain the downlink channel from the uplink channel measurements by applying the reciprocity principle as illustrated in Figure 5.1. However, the downlink and uplink channels cannot use all identical time, frequency, spatial instance, and RF chains. Transmit-receive chain calibration and equalization can transfer the transmit and receive RF chains to be identical [BCC⁺07]. If the difference in time, frequency, and spatial instance is relatively small compared to the channel variation across the referenced dimension, the reciprocity principle may still hold. For example, in the time dimension this condition implies that any time offset $\Delta\tau$ between the uplink and downlink transmission must be smaller than the channel coherence time T_c (i.e., $\Delta\tau \ll T_c$). Similarly, any frequency offset Δf must be smaller than the channel coherence bandwidth B_c (i.e., $\Delta f \ll B_c$), and the antenna location differences Δd on the both sides must be smaller than the channel coherence distance D_c (i.e., $\Delta d \ll D_c$).

The reciprocity principle is applicable in time division duplex (TDD) system, if the same frequency resources are used for the down- and uplink transmissions. Therefore, the downlink CSI can be extracted from the uplink transmissions of the user terminals. To achieve this, each user needs to transmit dedicated pilots per antenna on the uplink. Generally, this overhead is comparably small if the user terminals are only equipped with one or two antennas. From each pilot transmission, the BS can acquire the channel coefficients to all its antennas at the

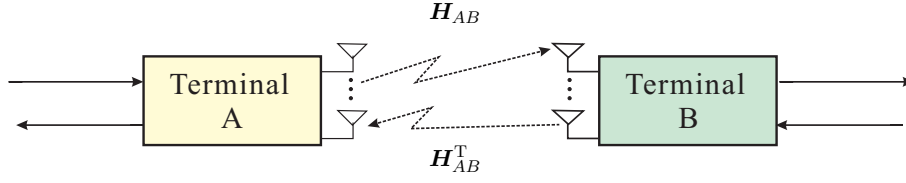


Figure 5.1.: Reciprocity for channel state information (CSI) acquisition.

same time. This channel state information is then used for the downlink transmissions. It should be noted that the available CSI is usually erroneous due to channel estimation errors, channel instationarity (outdated CSI), and RF impairments (e.g., in-phase quadrature-phase imbalance and carrier frequency offset). The spatial reuse scheme should be able to operate also in presence of imperfect CSI. It has been shown in the literatures [SHGJ06, ZdLH13a] that the linear precoding and decoding techniques are significantly more robust against CSI impairments compared to the non-linear precoding techniques such as THP, although a performance loss does occur due to the imperfect CSI [SH09c].

Another method for obtaining CSI is to use feedback from the receiver of the forward links depicted in Figure 5.2. It is generally utilized in frequency division duplex (FDD) system. Let the user terminals estimate their own channels during the downlink transmissions and use feedback to signal this information back to the base station. Such feedback channels do exist in current systems (e.g., for power control), but the required rate for the feedback is clearly an important quantity for the system design. The practically motivated finite rate feedback model was first considered for the point-to-point MIMO channels in [LHS03, MSEA03], where the transmitter utilizes such feedback to more accurately direct its transmit power towards the receiver. It is found in [LHSH04] that even a small number of bits per antenna can be quite beneficial. Furthermore, the level of CSI available at the transmitter only effects the SNR-offset in point-to-point MIMO channels. It does not affect the slope of the capacity vs. SNR curve (i.e., the multiplexing gain). In contrast, the multiplexing gain of the multi-user MIMO downlink channels is critically affected by the level of CSI available at the BS. Channel feedback therefore is considerably more important for the multi-user MIMO downlink channels than for the point-to-point channels.

The authors in [Jin06, RJ07] have discussed limited feedback strategies for multi-user MIMO downlink channels employing the ZF and BD precoding, where the knowledge of the user channels' direction is considered to be more useful than the knowledge of only the channel magnitude. Therefore, a quantized version of the channel direction of each user is fed back to the BS. The number of feedback bits per user must be increased linearly with the SNR (in dB)

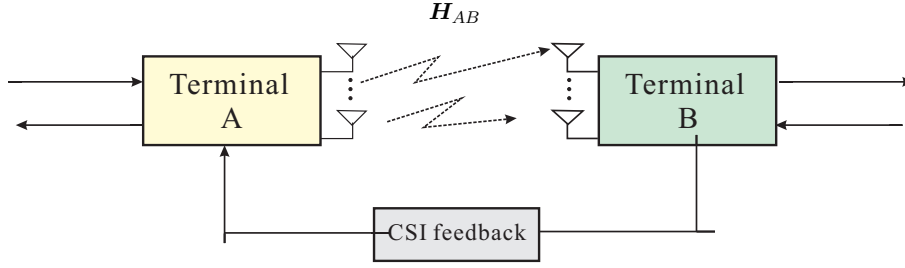


Figure 5.2.: channel state information (CSI) acquisition using feedback.

in order to achieve the full multiplexing gain. However, it has been found by us in [SRH08] that both the channel direction and the channel magnitude information are important for the BS to perform the RBD precoding. Motivated by it, we have proposed a new limited feedback scheme in [SRH08] to provide the quantized version of the channel direction and magnitude to the BS.

Exact channel knowledge at each time instance is known as perfect CSI, which is ideal and usually difficult to acquire in a time-variant channel. If perfect CSI is available at the BS, the multi-user interference (MUI) can be efficiently eliminated by employing the linear transmit and receive strategies such as linear precoding and CBF. If the channel varies too fast to obtain instantaneous CSI, the spatial channel correlation (namely long-term CSI) can be used alternatively to reduce the MUI and improve the system performance [SHMK09, KTS⁺10] by applying the same linear transmit and receive strategies. Furthermore, if there is no CSI available at the BS, the BS has to give up the transmit strategies like linear precoding and CBF and serves only one user at a time. In this case, only the strategies (e.g., space-time coding) which do not require CSI at the BS can be utilized. In [SHMK09], we have proposed a new scheme (namely rank-one approximated long-term CSI (ROLT-CSI)) to exploit the second order statistics of the channels at the BS.

As we have mentioned before, the CSI can be acquired from the known pilot symbols inserted in the transmission at the expense of a reduced bandwidth efficiency. To avoid this drawback, several efficient blind and semi-blind methods of channel estimation have been proposed in various channel contexts, for example, for the SIMO and MIMO fading channels [TXK91, MDCM95, Nef04], the space-time coded MIMO channels [ZMG02, VS08], and the MIMO-OFDM channels [MdCDB99, YP03]. The basic idea of the blind channel estimation (BCE) is to derive the channel characteristics from the received signal only. Depending on the different ways to extract this information from the received signal, BCE schemes can be distinguished in two classes: the moment-based BCE and the maximum likelihood (ML) based BCE. The main

advantages and drawbacks of them have been briefly described in Chapter 1. The moment-based BCE can be further divided into the second-order statistics (SOS) based BCE and the higher-order statistics (HOS) based BCE. The authors of [BKP98] have compared the performance of a HOS based BCE and a pilot based non-blind channel estimation scheme in a TDMA mobile communication system. It has been shown that the BCE algorithm leads to an SNR loss of 1.2 – 1.3 dB only compared to the non-blind channel estimation, while it saves a 22 % overhead caused by the transmission of the training sequences.

The SOS based blind channel estimation method in [MDCM95] has been viewed as a pioneering work which provides a subspace method for the blind estimation of the single-input multiple-output (SIMO) channels. This subspace-based blind channel estimation algorithm performs a singular value decomposition (SVD) of the correlation matrix of the measurement data to separate the observed space into two orthogonal subspaces, namely the signal subspace and the noise subspace. Then, the channel can be estimated up to a scalar factor by exploiting the orthogonality property between the signal and noise subspaces. The concept behind the subspace method is quite useful, since it is easily extendible to the blind estimation of the MIMO channels, the orthogonal space-time coded MIMO channels, and the MIMO OFDM channels.

However, in the existing subspace-based approaches to blind channel estimation, the measurement data is stored in one highly structured vector by a stacking operation. As a result, the structure inherent in the measurement data is not considered in the subspace estimation step. A more natural approach to store and exploit the inherent structure of the measurement data is given by tensors.

Tensor-based signal processing has become increasingly popular in many different areas of signal processing. This is due to the fact that it offers several fundamental advantages compared to matrix-based techniques, which have been introduced in Chapter 1. First of all, multilinear decompositions are essentially unique without additional constraints and allow to separate more components compared to the bilinear (matrix) approaches, which renders them attractive for component separation tasks [KB09]. Moreover, since the structure of the data is preserved, structured denoising can be applied, which leads to an improved tensor-based signal subspace estimate and can enhance any subspace-based parameter estimation scheme [HRD08].

In [SRH10a] and [SRH13b], we have proposed tensor-based approaches for blind channel estimation of SIMO and MIMO channels. Furthermore, inspired by these work, a tensor-based semi-blind channel estimation for Orthogonal STBC coded MIMO systems has been proposed by us in [RSS⁺11] for arbitrary OSTBCs.

This chapter is dedicated to the CSI acquisition. We start with an introduction of our new scheme ROLT-CSI in Section 5.2. The quantization based limited feedback model is discussed in Section 5.3. In Section 5.4, our tensor-based approaches for blind channel estimation are studied. Finally, a summary is provided in Section 5.5.

5.2. Linear precoding with long-term channel state information

In this section we consider the multi-user MIMO downlink and assume that the channel is correlated and varies too rapidly to obtain the instantaneous CSI (namely short-term CSI). Instead, the knowledge of the spatial correlation at the BS is exploited (namely long-term CSI), which allows us to use the existing precoding techniques (e.g., BD and RBD) designed for the perfect CSI at the BS.

5.2.1. System and data model

We model the multi-user MIMO downlink channel as a perfectly tuned orthogonal frequency division multiplexing (OFDM) channel without any inter subcarrier interference (the basics, perspectives, and challenges of MIMO-OFDM are found in [Bol06]). In this system (as we have shown in Figure 1.2), there are K users. The BS is equipped with M_T transmit antennas and the i th user has M_{R_i} receive antennas. The total number of receive antennas of all users is denoted by M_R (i.e., $M_R = \sum_{i=1}^K M_{R_i}$). We use $\mathbf{H}_i(N_f, N_t) \in \mathbb{C}^{M_{R_i} \times M_T}$ to denote the propagation channel between the BS and the user i at subcarrier N_f and OFDM symbol N_t . Then the combined MIMO channel matrix of all users can be defined as

$$\mathbf{H}(N_f, N_t) = [\mathbf{H}_1^T(N_f, N_t) \ \mathbf{H}_2^T(N_f, N_t) \ \dots \ \mathbf{H}_K^T(N_f, N_t)]^T \in \mathbb{C}^{M_R \times M_T}. \quad (5.1)$$

We assume that it is not possible to track the fast variations of the users' channels, but information about the spatial correlation of the channels can be obtained.

The downlink input output data model with linear precoding matrix \mathbf{F} can be expressed as

$$\mathbf{y} = \mathbf{H}(N_f, N_t) \mathbf{F} \mathbf{s} + \mathbf{n}, \quad (5.2)$$

where the vectors \mathbf{s} , \mathbf{y} , and \mathbf{n} represent the vectors of transmitted symbols, received signals at all users, and additive noise at the receive antennas, respectively. The matrix $\mathbf{F} = [\mathbf{F}_1, \dots, \mathbf{F}_K]$ denotes the combined precoding matrix for all users which is used to mitigate MUI. The dimensions r_i and r denote the number of data streams at the i th user terminal and the total number of data streams (i.e., $r = \sum_{i=1}^K r_i$), respectively.

Here, we introduce a chunk as the basic resource element, which was first proposed by the Wireless World Initiative New Radio (WINNER) project [ISTrg]. A chunk contains N_T consecutive OFDM symbols in the time direction and N_F subcarriers in the frequency direction as shown in Figure 5.3. Therefore, the number of $N_{\text{chunk}} = N_F \cdot N_T$ symbols are available within each chunk. The linear precoding matrix \mathbf{F} is kept constant in one chunk (called chunk-wise linear precoding).

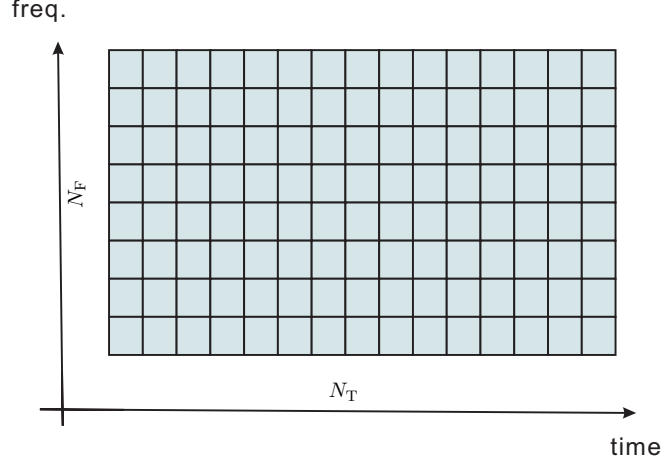


Figure 5.3.: One chunk structure in time and frequency domain.

5.2.2. Previous long-term CSI method

The authors in [SH05b, RFH08] introduce a method to exploit the long-term CSI for the multi-user MIMO downlink employing linear precoding techniques. They also use a chunk as the smallest time-frequency resource allocation unit and define the spatial correlation matrix estimate $\hat{\mathbf{R}}_{i,b}$ for user i at the chunk b as [SH05b, RFH08]

$$\hat{\mathbf{R}}_{i,b} = \frac{1}{N_{\text{chunk}}} \sum_{N_f=1}^{N_F} \sum_{N_t=1}^{N_T} \hat{\mathbf{H}}_i^H(N_f, N_t) \hat{\mathbf{H}}_i(N_f, N_t) \in \mathbb{C}^{M_T \times M_T}, \quad (5.3)$$

where $\hat{\mathbf{H}}_i(N_f, N_t)$ is the estimate for the channel of the i th user at subcarrier N_f and OFDM symbol N_t . If several pilots per chunk are available, there are two options to obtain the estimate $\hat{\mathbf{H}}_i(N_f, N_t)$. Either compute one estimate for the channels in chunk b from all pilots jointly. Alternatively, one can compute one estimate per pilot and then interpolate between these estimates for every symbol in chunk b .

The singular value decomposition (SVD) of $\hat{\mathbf{R}}_{i,b}$ is

$$\hat{\mathbf{R}}_{i,b} = \mathbf{V}_{i,b} \mathbf{\Lambda}_{i,b} \mathbf{V}_{i,b}^H . \quad (5.4)$$

Then, the equivalent channel is defined based on the spatial correlation matrix estimate $\hat{\mathbf{R}}_{i,b}$ as follows [SH05b, RFH08]

$$\hat{\mathbf{H}}_{i,b} = \mathbf{\Lambda}_{i,b}^{1/2} \mathbf{V}_{i,b}^H \in \mathbb{C}^{M_{R_i} \times M_T} . \quad (5.5)$$

Notice that if $\hat{\mathbf{R}}_{i,b}$ has a rank of M_{R_i} , the matrix $\hat{\mathbf{H}}_{i,b}$ can be directly obtained by the SVD of $\hat{\mathbf{R}}_{i,b}$ in equation (5.4). If $\hat{\mathbf{R}}_{i,b}$ has a full rank of M_T , the truncated SVD has to be applied to $\hat{\mathbf{R}}_{i,b}$ to obtain $\hat{\mathbf{H}}_{i,b}$ (illustrations of various SVDs are found in Table 5.4 of Section 5.4). The matrices $\hat{\mathbf{H}}_{i,b}$, $i = 1, \dots, K$ contain all information about the long-term subspace of each user available at the BS. We can now use the matrix $\hat{\mathbf{H}}_{i,b}$ as a long-term equivalent channel and perform the precoding on this matrix as if it represented the actual channel.

5.2.3. ROLT-CSI

We have proposed a new method (i.e., ROLT-CSI) to effectively represent the channel by exploiting the knowledge of the estimated channel spatial correlation. Compared to the previous long-term CSI introduced above, our new method is more efficient, especially for the case that the user has a low rank spatial correlation matrix. The performance improvement achieved by our method will be demonstrated by the simulation results in the following section.

In our method, we consider the spatial correlation matrix for each receive antenna per user. Then, for user i the spatial correlation matrix associated to the ℓ th receive antenna is expressed as

$$\mathbf{R}_{i,\ell}(N_f, N_t) = \mathbb{E} \{ \hat{\mathbf{h}}_{i,\ell}(N_f, N_t) \hat{\mathbf{h}}_{i,\ell}^H(N_f, N_t) \} \in \mathbb{C}^{M_T \times M_T} . \quad (5.6)$$

Here $\hat{\mathbf{h}}_{i,\ell}^H(N_f, N_t)$ denotes the ℓ th row of the channel matrix estimate $\hat{\mathbf{H}}_i(N_f, N_t) \in \mathbb{C}^{M_{R_i} \times M_T}$. The index ℓ indicates the ℓ th receive antenna of user i . The spatial correlation matrix of the ℓ th receive antenna of user i can be estimated by averaging over one chunk. Let $\hat{\mathbf{R}}_{i,b,\ell}$ denote the estimated spatial correlation matrix of user i , chunk b , and receive antenna ℓ . Then we have

$$\hat{\mathbf{R}}_{i,b,\ell} = \frac{1}{N_{\text{chunk}}} \sum_{N_f=1}^{N_F} \sum_{N_t=1}^{N_T} \hat{\mathbf{h}}_{i,\ell}(N_f, N_t) \hat{\mathbf{h}}_{i,\ell}^H(N_f, N_t) \quad (5.7)$$

and its SVD as

$$\hat{\mathbf{R}}_{i,b,\ell} = \mathbf{V}_{i,b,\ell} \mathbf{\Lambda}_{i,b,\ell} \mathbf{V}_{i,b,\ell}^H, \ell = 1, \dots, M_{R_i} . \quad (5.8)$$

It has been found in [JG05a, JB06, BO02] that if only second-order channel statistics are

available at the transmitter, the optimum transmission strategy is to transmit along the dominant eigenmode of its spatial correlation matrix. Inspired by it, we define the equivalent channel matrix of user i in chunk b as

$$\hat{\mathbf{H}}_{i,b} = \mathbf{A}_{i,b} \mathbf{B}_{i,b} \in \mathbb{C}^{M_{R_i} \times M_T}, \quad (5.9)$$

where

$$\mathbf{A}_{i,b} = \begin{bmatrix} \sqrt{\boldsymbol{\Lambda}_{i,b,1}(1,1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \sqrt{\boldsymbol{\Lambda}_{i,b,2}(1,1)} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \sqrt{\boldsymbol{\Lambda}_{i,b,M_{R_i}}(1,1)} \end{bmatrix} \in \mathbb{C}^{M_{R_i} \times M_T}$$

and

$$\mathbf{B}_{i,b} = \begin{bmatrix} \mathbf{V}_{i,b,1}^H(:,1) \\ \mathbf{V}_{i,b,2}^H(:,1) \\ \vdots \\ \mathbf{V}_{i,b,M_{R_i}}^H(:,1) \end{bmatrix} \in \mathbb{C}^{M_T \times M_T}.$$

Here $\boldsymbol{\Lambda}_{i,b,l}(1,1)$ indicates the largest eigenvalue of $\hat{\mathbf{R}}_{i,b,l}$ and $\mathbf{V}_{i,b,l}^H(:,1)$ denotes the corresponding eigenvector of $\hat{\mathbf{R}}_{i,b,l}$.

The multi-user MIMO precoding can now be performed on the equivalent channel as defined in equation (5.9). Clearly, the rank-one approximation in equation (5.9) can effectively represent the channel if its spatial correlation matrix in equation (5.7) also has a low rank.

Utilizing BD precoding described in Chapter 3 as an example, if there is only long-term CSI available at the BS, we use the equivalent channel in equation (5.9) from the ROLT-CSI approach instead of the exact channel \mathbf{H}_i in equations (3.9) and (3.11).

5.2.4. Simulation Results

We evaluate the throughput performance of the BD and RBD precoding techniques, when only the long-term CSI is available at the BS. We consider a MIMO downlink system with 3 users. The simulation scenario is illustrated in Figure 5.4. The channels between each user and the BS are generated by a geometry-based channel model called *IlmProp*, which has been developed at Ilmenau University of Technology [DHS03] and is capable of dealing with time variant frequency selective scenarios.

There are 8 transmit antennas at the BS and each user is equipped with 2 receive antennas. The BS simultaneously transmits two data streams per user. User 1 and user 2 always have non-line of sight (NLOS) channels and user 3 always has a line of sight (LOS) channel. The

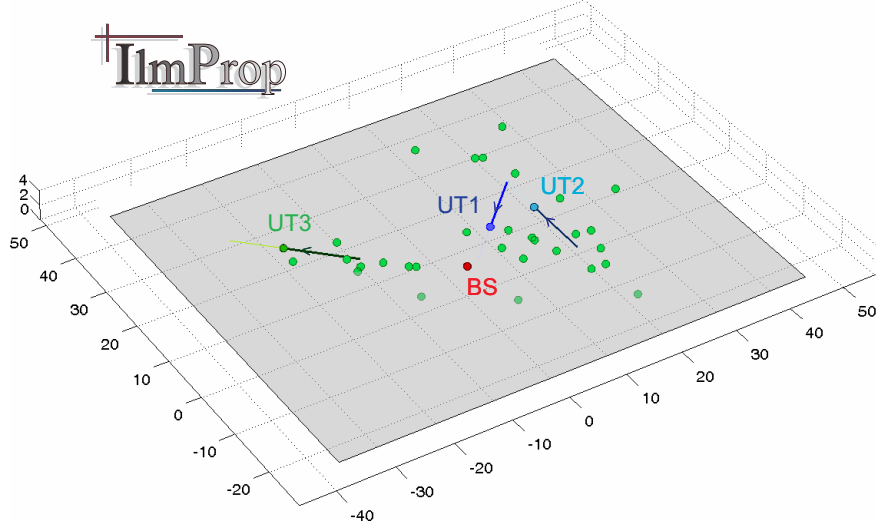


Figure 5.4.: The geometrical representation of the simulation scenario. Each green point represents a fixed scatterer. The channel impulse responses (CIR) are generated as a sum of propagation rays. The channel is computed from the superposition of the LOS component and a number of rays which represent the multi-path components. User 1 and user 2 always have NLOS channels and user 3 always has a LOS channel.

velocities of the three users are 10 km/h. In Table 5.1, the important OFDM parameters are listed.

Table 5.1.: OFDM Parameters

Parameters	Values
Carrier Frequency	5 GHz
Subcarrier Spacing	0.50196 MHz
Useful Symbol Duration	1.9922 μ s
System Bandwidth	128.5 MHz
Used Subcarriers	$[-128 : +128]$, 0 not used
Chunk Size	8 subcarriers, 15 OFDM symbols
Duplexing Mode	TDD

Since we assume time division duplexing (TDD), the channel reciprocity can be exploited. We can therefore extract downlink CSI from the uplink transmission of the user terminal. To achieve this, each user requires *uplink dedicated pilots* for the estimation of the channel between the user terminal and all BS antennas. For each chunk, if there are several pilots available (e.g., four pilots per chunk in a rectangular pattern have been defined for the TDD mode in the WINNER project [ISTrg]), we can get the estimate of the channel $\hat{\mathbf{H}}_i(N_f, N_t)$ in

two ways: either compute one channel estimate per pilot and then interpolate between these estimates for every symbol in the chunk or compute one estimate from all pilots jointly and use it as a representative for the entire chunk. In our simulation, we do not consider specific interpolation algorithms or joint estimation from different pilots for simplicity. We assume that $\hat{\mathbf{H}}_i(N_f, N_t)$ is already known exactly¹. Then, we can calculate the equivalent channel of the chunk with equation (5.9) for the ROLT-CSI approach and with equation (5.5) for the long-term CSI method of [SH05b], respectively. Consequently, the BS can compute the precoding matrix \mathbf{F} for each chunk. The linear precoding schemes used in the simulation are BD precoding and RBD precoding.

In Figures 5.5, 5.6, and 5.7 we assume that the channel $\hat{\mathbf{H}}_i(N_f, N_t)$ is known perfectly. In Figure 5.5, we compare the throughput of the system with precoding based on ROLT-CSI to the throughput based on the state of the art long-term CSI method in [SH05b]. We can see that RBD precoding can achieve a higher data rate than BD precoding. When linear precoding is performed based on long-term CSI, a significant performance gain can be achieved by our new approach relative to the previous long-term CSI method.

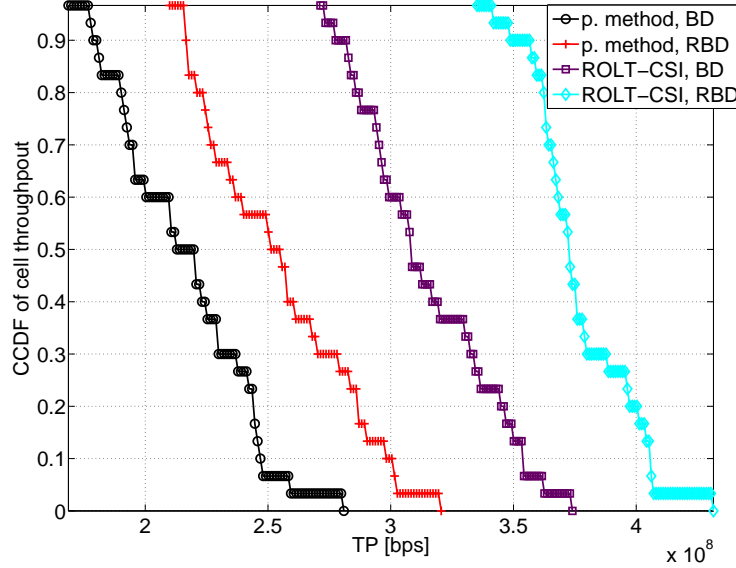


Figure 5.5.: CCDF of the sum rates with BD and RBD precoding based on long-term CSI at the BS, respectively. p. method indicates the previous long-term CSI method.

In Figures 5.6 and 5.7 the individual user throughputs based on ROLT-CSI and the previous

¹This assumption is reasonable, because the comparison of our ROLT-CSI method and the previous long-term CSI method does not depend on the way to get $\hat{\mathbf{H}}_i(N_f, N_t)$

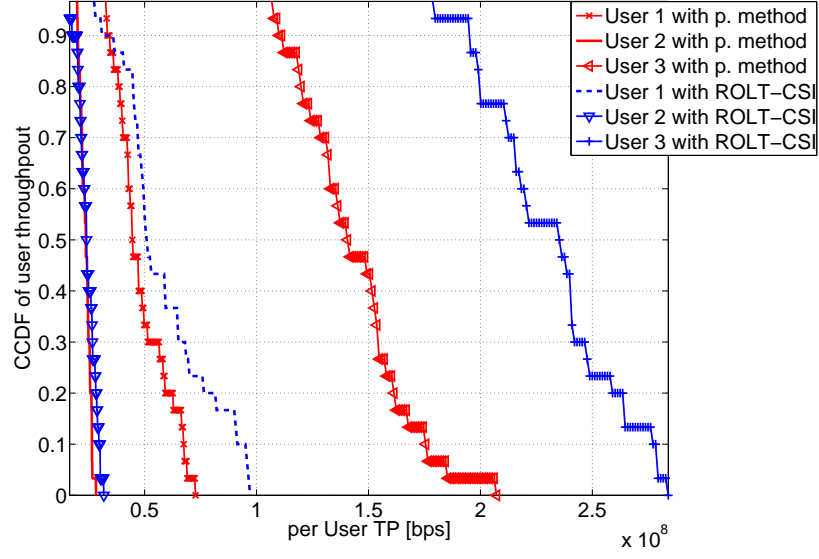


Figure 5.6.: CCDF of the individual user throughput with BD precoding based on long-term CSI at the BS, p. method indicates the previous long-term CSI method.

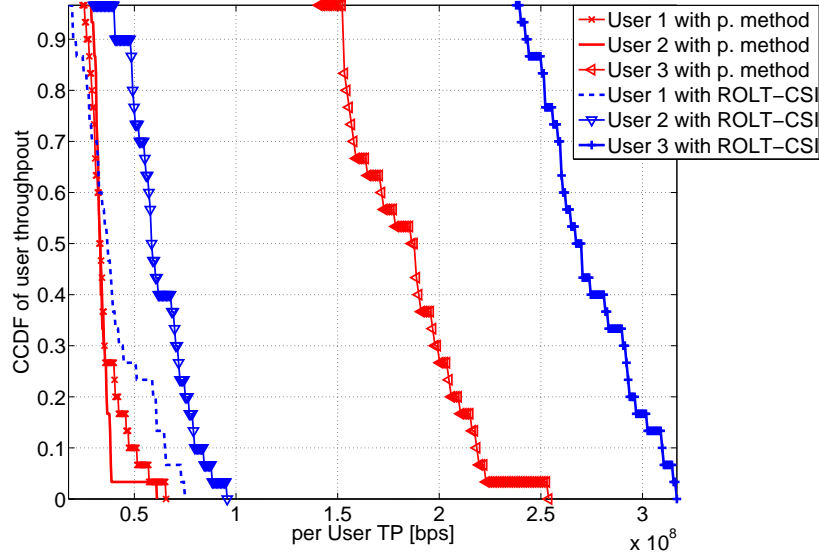


Figure 5.7.: CCDF of the individual user throughput with RBD precoding based on long-term CSI at the BS, p. method indicates the previous long-term CSI method.

long-term CSI approach are compared. It is shown that the ROLT-CSI approach is particularly efficient for the user who has the LOS channel. Even for the users who only have NLOS channels, which means that the spatial correlation matrix of these user channels have a high rank, relative to the previous long-term CSI method there are still some performance gains available for the presented ROLT-CSI approach.

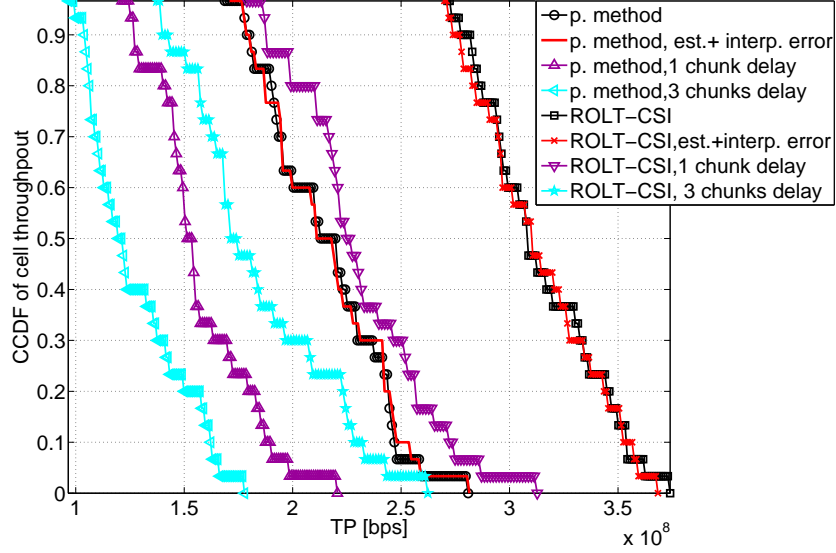


Figure 5.8.: CCDF of the sum rates with BD precoding based on long-term CSI at the BS, p. method indicates the previous long-term CSI method.

Taking into account realistic channel propagation conditions, for Figure 5.8 and 5.9 we assume that the channel $\hat{\mathbf{H}}_i(N_f, N_t)$ is known imperfectly. We consider a channel estimation error, a channel interpolation error, and the delay resulting from the fact that the available CSI of chunk k will be used to optimize the transmission over the channel realization of chunk $(k+n)$. One chunk and three chunks delay are considered separately in the simulation. According to Table 5.1, the duration of one chunk is equal to the duration of 15 OFDM symbols.

For the CSI imperfection, the channel estimation error and interpolation error are modeled according to [ISTrg], but we increase the interpolation error variance to -20 dB. Compared to the previous long-term CSI method, the performance gains achieved by our proposed ROLT-CSI method are significant. It is also found that the delay is the predominant cause of a performance degradation in a precoded multi-user MIMO system with long-term CSI.

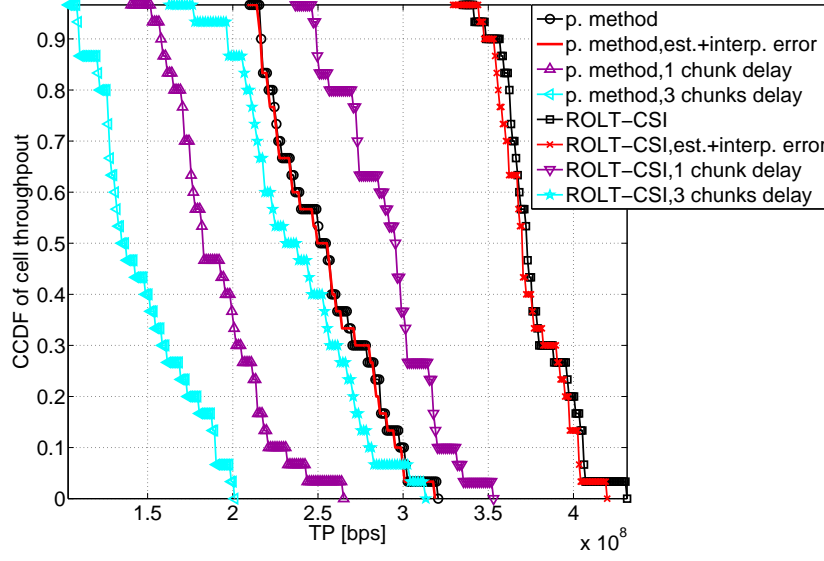


Figure 5.9.: CCDF of the sum rates with RBD precoding based on long-term CSI at the BS, p. method the indicates previous long-term CSI method.

5.3. Quantized finite rate feedback for multi-user MIMO broadcast channels

This section is focused on multi-user MIMO broadcast channels employing linear precoding techniques and finite rate feedback strategies. Figure 5.10 shows us the basic system model. We consider a multi-user MIMO system with a single base station (BS) and K users, where the BS has M_T transmit antennas and the i th user has M_{R_i} receive antennas. Employing linear precoding, the received signal after the decoding of the i th user is expressed as

$$\mathbf{y}_i = \mathbf{W}_i^H \left(\mathbf{H}_i \sum_{k=1}^K \mathbf{F}_k \mathbf{s}_k + \mathbf{n}_i \right) \quad (5.10)$$

where the vector $\mathbf{s}_k \in \mathbb{C}^{r_k \times 1}$ contains the data symbols and r_k represents the number of data streams for user k , $k = 1, 2, \dots, K$. The matrix $\mathbf{F}_k \in \mathbb{C}^{M_T \times r_k}$ denotes the precoding matrix. Here the matrix $\mathbf{H}_i \in \mathbb{C}^{M_{R_i} \times M_T}$ is the channel matrix from the BS to the i th user. The channel is assumed to be i.i.d. block Rayleigh fading. The vector $\mathbf{n}_i \in \mathbb{C}^{M_{R_i} \times 1}$ represents the complex Gaussian noise vector with unit variance (i.e., $\sigma_n^2 = 1$), which is independent of \mathbf{s}_k . The matrix $\mathbf{W}_i \in \mathbb{C}^{M_{R_i} \times r_i}$ denotes the decoding matrix and $\mathbf{y}_i \in \mathbb{C}^{r_i \times 1}$ is the receive signal vector of user i .

We assume that the total transmit power of all users is constrained by P_T , i.e., $\mathbb{E}\left\{\left|\sum_{k=1}^K \mathbf{F}_k \mathbf{s}_k\right|^2\right\} \leq P_T$ ($P_T \geq 0$). Let us define $\tilde{P}_T = \frac{P_T}{\sigma_n^2}$ and $\text{SNR} = 10 \log_{10} \tilde{P}_T$. Furthermore, we assume that each user can perfectly estimate its own channel and quantize it to B bits. The BS acquires the B bits of each user through an error free and zero delay feedback channel.

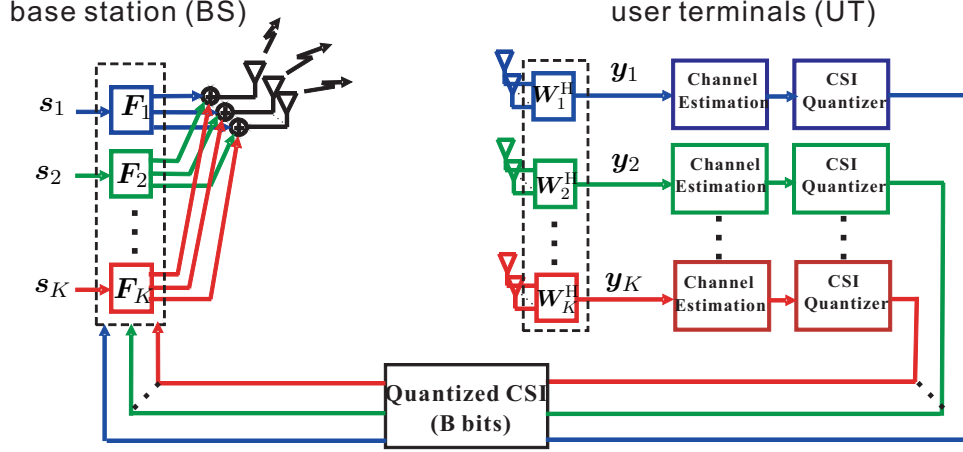


Figure 5.10.: Limited feedback system model.

There are several ways to quantize the channel matrices. A straightforward method is to view the channel matrix as a set of complex numbers and to encode every complex number individually. If we assume that N_b bits are required to represent one floating point number, then there are $B = 2N_b \cdot M_{R_i} \cdot M_T$ bits in total for the user i and $2N_b \cdot M_R \cdot M_T$ for all users ($M_R = \sum_{i=1}^K M_{R_i}$). This costs too much. The alternative is to quantize every individual matrix by looking up a predefined codebook. If we allocate B bits for each codebook, we only need B bits for one user and KB bits for all users. Therefore, this channel quantization method is more efficient regarding the required feedback bits. It is also used in this section.

In the following part, we first overview the existing limited feedback strategies [Jin06, RJ07] proposed for the multi-user MIMO broadcast channels employing ZF and BD. Then, we study a new quantization scheme proposed by us in [SRH08] for the multi-user MIMO downlink applying the RBD precoding.

5.3.1. Quantization scheme for ZF based system

In [Jin06] a simple downlink transmission scheme that uses ZF precoding in conjunction with finite rate feedback has been studied. Since each user has only one receive antenna, the channel between the i th user and the BS is a vector $\mathbf{h}_i \in \mathbb{C}^{M_T \times 1}$. It is assumed that the number of users is equal to the number of transmit antennas (i.e., $K = M_T = M_R$).

5.3.1.1. Finite rate feedback model

The quantization is performed by using a vector quantization codebook which is known at the BS and the user. In order to prevent the multiple users from quantizing their channel to the same quantization vector, each user has a different codebook. A quantization codebook \mathcal{C} consists of 2^B M_T -dimensional unit norm vectors, i.e., $\mathcal{C} = \{\mathbf{w}_1, \dots, \mathbf{w}_{2^B}\}$. Here, B indicates the number of feedback bits per user. Each user quantizes its channel to the codeword that is closest to its channel vector. The closeness is measured in terms of the principal angle θ between two vectors or equivalently the inner product of these two vectors. Thus, user i chooses a quantization index α_i according to

$$\begin{aligned} \alpha_i &= \arg \max_{j=1, \dots, 2^B} |\mathbf{h}_i^H \mathbf{w}_j|^2 \\ &= \arg \min_{j=1, \dots, 2^B} \sin^2 \left(\angle(\mathbf{h}_i, \mathbf{w}_j) \right), \\ &\quad \text{where } \theta_i = \angle(\mathbf{h}_i, \mathbf{w}_j) \in [0, \pi/2] \text{ and } \cos \theta_i = |\mathbf{h}_i^H \mathbf{w}_j| \end{aligned} \quad (5.11)$$

and feeds this index back to the BS. It is noted that only the direction of the channel vector is quantized, the channel magnitude information is not conveyed to the BS.

Clearly, the choice of the vector quantization codebook significantly affects the quality of the quantized CSI. Random vector quantization (RVQ) has been used in [Jin06] because the optimal vector quantizer for this problem is not known in general.

5.3.1.2. Random vector quantization (RVQ)

Random vector quantization was first used to analyze the performance of CDMA and point-to-point MIMO systems with finite rate feedback. It has been shown that RVQ is asymptotically optimal in the large system limit (e.g., very large number of users) [SH05a, SH04a]. In [AYL07], RVQ has been also utilized to quantize the beamforming vector for a point-to-point MISO system.

Let $\hat{\mathbf{h}}_i$ denote the quantization of the channel direction $\tilde{\mathbf{h}}_i$ (i.e., $\tilde{\mathbf{h}}_i = \frac{\mathbf{h}_i}{\|\mathbf{h}_i\|}$). The quantization error can be expressed as

$$\begin{aligned} D(\mathcal{C}) &\triangleq 1 - |\tilde{\mathbf{h}}_i^H \hat{\mathbf{h}}_i|^2 \\ &\triangleq \sin^2 \left(\angle(\tilde{\mathbf{h}}_i, \hat{\mathbf{h}}_i) \right), \end{aligned} \quad (5.12)$$

which is the minimum of 2^B independent random variables. The expectation of the quantization error averaged over all possible random quantization codebooks has been derived in a

closed-form in Corollary 1 of [AYL07] as

$$\mathbb{E}[D(\mathcal{C})] = 2^B \cdot \beta(2^B, \frac{M_T}{M_T - 1}). \quad (5.13)$$

Here, the beta function is defined in terms of the gamma function as $\beta(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$. The gamma function is the extension of the factorial function to non-integers, and satisfies the fundamental properties $\Gamma(n) = (n-1)!$ for positive integers (see Appendix C.2). Furthermore, a simple extension of inequalities given in [AYL07] provides a strict upper bound of the expected quantization error.

Lemma 5.3.1. *The expected quantization error can be upper bounded as [Jin06]*

$$\mathbb{E}[D(\mathcal{C})] < 2^{-\frac{B}{M_T-1}}. \quad (5.14)$$

Proof: see [Jin06]. The numerical results in [Jin06] indicate that this bound is tight and indeed sufficiently precise to accurately characterize the throughput degradation due to finite rate feedback.

5.3.1.3. Throughput analysis

Let the matrix $\hat{\mathbf{H}} \in \mathbb{C}^{M_T \times M_T}$ indicate the quantized version of the channels all users (i.e., $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \dots, \hat{\mathbf{h}}_K]$, notice that $K = M_T = M_R$). The ZF precoding is performed upon it. If equal power allocation is used for the transmission, the receive signal to interference plus noise ratio (SINR) at the user i is given by

$$\text{SINR}_i = \frac{\frac{\tilde{P}_T}{M_T} |\mathbf{h}_i^T \mathbf{f}_i|^2}{1 + \sum_{j \neq i} \frac{\tilde{P}_T}{M_T} |\mathbf{h}_i^T \mathbf{f}_j|^2}. \quad (5.15)$$

Since the beamforming vectors $\mathbf{f}_i \in \mathbb{C}^{M_T \times 1}$ are chosen orthogonal to the quantized channel and not the actual channel realizations, the interference term in the dominator of the SINR expression are not zero. Therefore, a throughput loss is introduced due to the residual interference term. Considering the throughput averaged over the fading distribution, the performance degradation can be quantified as a function of the feedback rate. Let us define the rate gap $\Delta R(P_T)$ per user to be the difference between the throughput achieved by perfect CSI based ZF and finite rate feedback based ZF. We have

$$\Delta R(\tilde{P}_T) = [R(\tilde{P}_T) - R_{\text{FB}}(\tilde{P}_T)], \quad (5.16)$$

Theorem 5.3.2. *Finite rate feedback with B feedback bits per user incurs a throughput loss relative to ZF with perfect CSI upper bounded by [Jin06]*

$$\Delta R(\tilde{P}_T) < \log_2(1 + \tilde{P}_T \cdot 2^{-\frac{B}{M_T-1}}). \quad (5.17)$$

Proof: see [Jin06].

5.3.1.3.1. Fixed feedback bits

It is noted from Theorem 5.3.2 that the rate loss is an increasing function of \tilde{P}_T where $\tilde{P}_T = \frac{P_T}{\sigma_n^2}$. When the number of feedback bits B is fixed, the finite rate feedback system is interference limited at high SNRs because interference and signal power both scale linearly with \tilde{P}_T . This motivates the following result.

Theorem 5.3.3. *For a finite rate feedback system with ZF and a fixed number of feedback bits B per user, the throughput is bounded as \tilde{P}_T is taken to infinity.*

$$R_{FB}(\tilde{P}_T) \leq M_T \left(1 + \frac{B + \log_2 e}{M_T - 1} + \log_2(M_T - 2) + \log_2 e \right). \quad (5.18)$$

Proof: see [Jin06].

Although the upper bound in Theorem 5.3.3 is not quite tight in general, it does correctly predict the roughly linear dependence of the limiting throughput and the number of feedback bits B . Figure 5.11 shows the performance of a 5 transmit antennas, 5 users MIMO system with 10, 15, and 20 feedback bits per user. When the SNR is quite low, the limited feedback performances approach to ZF with perfect CSI. However, with the increased SNR, the limited feedback system becomes interference limited and the throughputs converge to an upper limit.

5.3.1.3.2. Increased feedback bits

In order to avoid the interference-limited behavior experienced in the case of a fixed number of feedback bits, the number of feedback bits per user B should be scaled linearly with \tilde{P}_T . In fact, if B is scaled at an appropriate rate, the full multiplexing gain of M_T can be achievable. It is also desired to maintain a throughput loss $\Delta R(\tilde{P}_T)$ that is not larger than a given bound $\log_2 b$ per user. The following theorem specifies a sufficient scaling of feedback bits B to maintain the given bound $\log_2 b$.

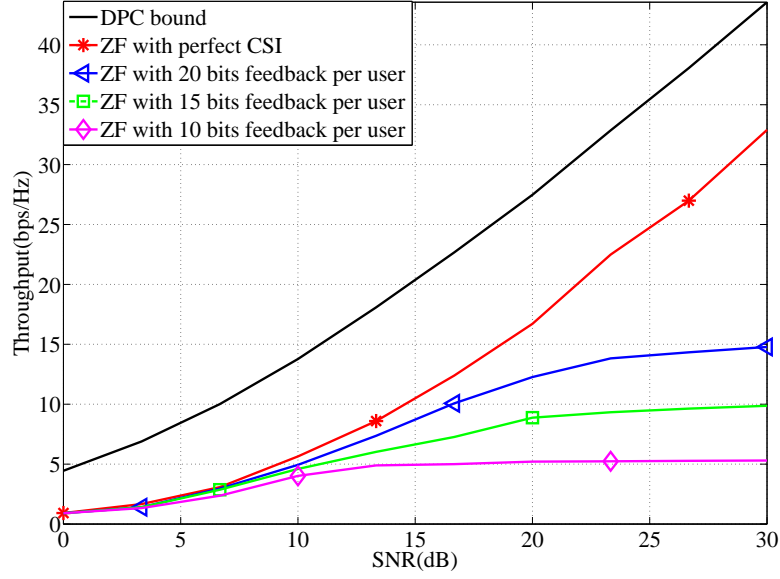


Figure 5.11.: Achievable throughputs for ZF $M_T = K = 5$ with fixed feedback bits per user.

Theorem 5.3.4. *In order to maintain a throughput loss that is not larger than $\log_2 b$ per user between ZF with perfect CSI and quantized CSI, it is sufficient to scale the number of feedback bits per user according to [Jin06]*

$$\begin{aligned}
 B &= (M_T - 1) \log_2 \tilde{P}_T - (M_T - 1) \log_2 (b - 1) \\
 &\approx \frac{M_T - 1}{3} \text{SNR} - (M_T - 1) \log_2 (b - 1).
 \end{aligned} \tag{5.19}$$

Proof: see [Jin06].

The throughput loss of $\log_2 b$ per user can easily be translated into a power offset, which is a more useful metric from the design perspective. As we have known from Chapter 3, the ZF precoding has a slope of M_T bps/Hz/3 dB at asymptotically high SNR (i.e., multiplexing gain, and note that $M_T = M_R$ in our assumption of the ZF based system model). Therefore, a throughput loss of $\log_2 b$ bps/Hz per user, or equivalently $M_T \log_2 b$ bps/Hz in the sum rate of all users, corresponds to a power offset of $3 \log_2 b$ dB (regarding equation (3.35) in Chapter 3). Thus, $b = 2$ indicates a 1 bps/Hz throughput loss per user or equivalently a 3 dB power

offset. From Theorem 5.3.4, the scaling of bits for a 3 dB offset is

$$B = \frac{M_T - 1}{3} \text{SNR} \quad \text{bits/user}, \quad (5.20)$$

which is a quite simple form.

Figure 5.12 shows the throughputs for a 5 transmit antennas and 5 users MIMO system employing the ZF precoding based on perfect CSI and quantized CSI, respectively. The number of feedback bits per user B is scaled according to equation (5.20) in order to maintain a 3 dB power offset. Notice that the actual power offset is smaller than 3 dB due to the use of Jensen's inequality (definition in Appendix E.1) in deriving the throughput loss upper bound in Theorem 5.3.2.

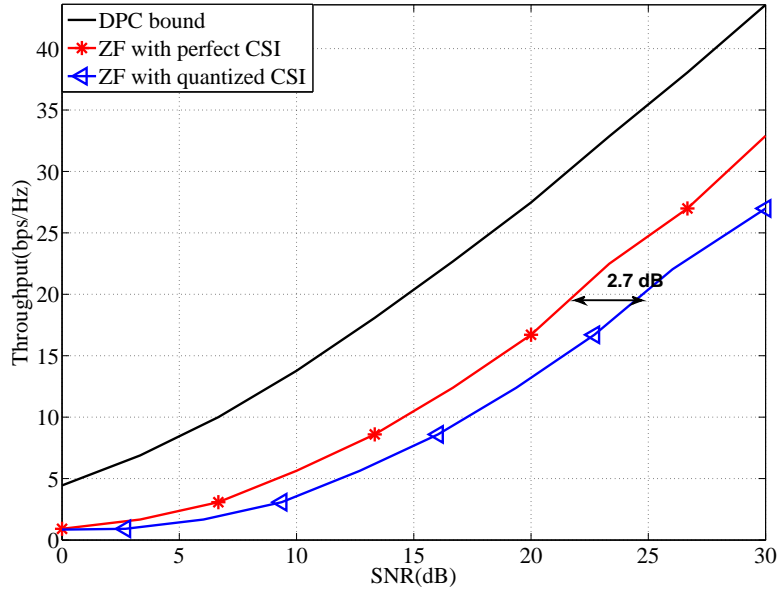


Figure 5.12.: Achievable throughputs for ZF $M_T = K = 5$ with increased feedback bits per user.

5.3.2. Quantization scheme for BD based system

In [RJ07] a limited feedback system employing the BD precoding at the BS has been considered, where each user is assumed to have M_r receive antennas (i.e., $M_{R_i} = M_r$ for $\forall i$), the number of users $K > 1$, and $K = \frac{M_T}{M_r}$ which implies that the aggregate number of receive antennas equals the number of transmit antennas (i.e., $M_R = M_T$). In order to perform BD, the BS only needs

to know the spatial direction of each user's channel, i.e., the subspace spanned by user channel $\mathbf{H}_i^T \in \mathbb{C}^{M_T \times M_r}$, $i = 1, \dots, K$. Thus, the feedback only conveys this information.

5.3.2.1. Finite rate feedback model

Each user has its own quantization codebook which is fixed beforehand and is known to the BS and each user. A quantization codebook consists of 2^B matrices, i.e., $\mathcal{C} = \{\mathbf{W}_1, \dots, \mathbf{W}_{2^B}\}$. Here, B is the number of feedback bits per user and each codeword has the dimension $M_T \times M_r$. The quantization of a channel spatial direction matrix $\tilde{\mathbf{H}}_i^T$ (i.e., $\tilde{\mathbf{H}}_i^T = \frac{\mathbf{H}_i^T}{\|\mathbf{H}_i^T\|_F}$) is chosen from the codebook \mathcal{C} according to [RJ07]

$$\hat{\mathbf{H}}_i^T = \arg \min_{\mathbf{W} \in \mathcal{C}} d_c^2(\tilde{\mathbf{H}}_i^T, \mathbf{W}), \quad \mathbf{W} \in \mathbb{C}^{M_T \times M_r}. \quad (5.21)$$

Here, $d(\tilde{\mathbf{H}}_i^T, \mathbf{W})$ is the distance metric which uses the chordal distance (details are found in Appendix E.2). That is

$$d_c(\tilde{\mathbf{H}}_i^T, \mathbf{W}) = \sqrt{\sum_{i=1}^{M_T} \sin^2 \theta_i}, \quad (5.22)$$

where the θ_i for $i = 1, \dots, M_T$ are the principal angles (i.e., $\theta_1, \dots, \theta_{M_T} \in [0, \pi/2]$) between the two subspaces spanned by the columns of the matrices. Since the principal angles only depend on the subspace spanned by the columns of the matrices, each codeword can be chosen as a unitary matrix. No channel magnitude information is fed back to the BS.

5.3.2.2. Random vector quantization (RVQ) codebook

Since it is very difficult to design optimal quantization codebooks for the given distance metric, the random vector quantization codebooks are studied instead in [RJ07]. Each codeword of the RVQ codebook is chosen independently and uniformly distributed over the Grassmannian manifold $\mathbb{G}_{M_T, M_r}(\mathbb{C})$ which is the set of all M_r -dimensional subspaces in an M_T -dimensional Euclidean space (the definition is found in Appendix E.2). Points in $\mathbb{G}_{M_T, M_r}(\mathbb{C})$ are equivalence classes of orthonormal matrices with the dimension $M_T \times M_r$. Each codeword therefore can be assumed to have unitary columns. Let us analyze the performance averaged over all possible random quantization codebooks. The expectation of the quantization error is defined as

$$\begin{aligned} D(\mathcal{C}) &= \mathbb{E} [d_c^2(\tilde{\mathbf{H}}_i^T, \hat{\mathbf{H}}_i^T)] \\ &= \mathbb{E} [d_c^2(\tilde{\mathbf{H}}_i^T, \mathbf{W})]. \end{aligned} \quad (5.23)$$

It has been shown in [DLR05] that the quantization error $D(\mathcal{C})$ has an upper bound for a codebook of size 2^B as follows

$$\begin{aligned} D(\mathcal{C}) &\leq \bar{D}(\mathcal{C}) \\ &= \frac{\Gamma(\frac{1}{M})}{M} (C_{M_T, M_r, \beta})^{-\frac{1}{M}} 2^{-\frac{B}{M}} + M_r \exp \left[-(2^B C_{M_T, M_r, \beta})^{(1-a)} \right], \end{aligned} \quad (5.24)$$

where $\Gamma(\frac{1}{M})$ is the gamma function (definition is found in Appendix C.2). The term $M = M_r(M_T - M_r)$ and $a \in (0, 1)$ is a real number and chosen such that $(2^B C_{M_T, M_r, \beta})^{-\frac{a}{M}} \leq 1$. The term $C_{M_T, M_r, \beta}$ is given by

$$C_{M_T, M_r, \beta} = \begin{cases} \frac{1}{M!} \prod_{i=1}^{M_r} \frac{\Gamma(\frac{\beta}{2}(M_T - i + 1))}{\Gamma(\frac{\beta}{2}(M_r - i + 1))} & \text{if } 0 < M_r \leq \frac{M_T}{2} \\ \frac{1}{M!} \prod_{i=1}^{M_T - M_r} \frac{\Gamma(\frac{\beta}{2}(M_T - i + 1))}{\Gamma(\frac{\beta}{2}(M_T - M_r - i + 1))} & \text{if } \frac{M_T}{2} \leq M_r \leq M_T \end{cases} \quad (5.25)$$

If each codeword is complex valued, β then is equal to 2. Otherwise, $\beta = 1$ for the real valued case. The second exponential term in equation (5.24) can be neglected for large B .

5.3.2.3. Throughput analysis

In the case of perfect CSI based BD precoding, the BS has the ability to mitigate all multi-user interference (MUI). Thus, the achievable throughput of user i is given by

$$R(\tilde{P}_T) = \mathbb{E} \left\{ \log_2 \left| \mathbf{I}_{M_r} + \frac{\tilde{P}_T}{K} \mathbf{H}_i \mathbf{F}_i \mathbf{F}_i^H \mathbf{H}_i^H \right| \right\}, \quad (5.26)$$

where the expectation is carried out over all channel realizations. The transmit power is assumed to be equally allocated between K users. In the case of quantized CSI based BD precoding, the MUI cannot be perfectly canceled and leads to a throughput loss compared to equation (5.26). Therefore, the achievable throughput of user i is given by

$$R_{\text{FB}}(\tilde{P}_T) = \mathbb{E} \left\{ \log_2 \left| \mathbf{I}_{M_r} + \frac{\tilde{P}_T}{K} \sum_{j=1}^K \mathbf{H}_i \hat{\mathbf{F}}_j \hat{\mathbf{F}}_j^H \mathbf{H}_i^H \right| \right\} - \mathbb{E} \left\{ \log_2 \left| \mathbf{I}_{M_r} + \frac{\tilde{P}_T}{K} \sum_{j=1, j \neq i}^K \mathbf{H}_i \hat{\mathbf{F}}_j \hat{\mathbf{F}}_j^H \mathbf{H}_i^H \right| \right\}, \quad (5.27)$$

where $\hat{\mathbf{F}}_j$ indicates the BD precoding matrix calculated through the quantized CSI. Let us define the rate gap $\Delta R(\tilde{P}_T)$ per user to be the difference between equations (5.26) and (5.27).

Theorem 5.3.5. *Finite rate feedback with B feedback bits per user incurs a throughput loss*

relative to BD with perfect CSI upper bounded by [RJ07]

$$\begin{aligned}\Delta R(\tilde{P}_T) &= R(\tilde{P}_T) - R_{FB}(\tilde{P}_T) \\ &\leq M_r \log_2(1 + \tilde{P}_T D(\mathcal{C})).\end{aligned}\tag{5.28}$$

Proof: see [RJ07].

5.3.2.3.1. Increased feedback bits

With the fixed B , the finite rate system eventually becomes interference limited at high SNRs (similar to the ZF case shown in Section 5.3.1.3.1). To avoid the interference-limited behavior, the number of feedback bits per user B should be scaled linearly with the SNR. If we consider to maintain that the throughput loss $\Delta R(\tilde{P}_T)$ should be no larger than a given bound $\log_2 b$ per user, we have the following theorem.

Theorem 5.3.6. *In order to maintain a throughput loss no larger than $\log_2 b$ per user between BD with perfect CSI and quantized CSI, it is sufficient to scale the number of feedback bits per user according to*

$$\begin{aligned}B \approx & \frac{M_r(M_T - M_r)}{3} \cdot \text{SNR} - M_r(M_T - M_r) \log_2(b^{\frac{1}{M_r}} - 1) \\ & + M_r(M_T - M_r) \log_2 \left(\frac{\Gamma(\frac{1}{M_r(M_T - M_r)})}{M_r(M_T - M_r)} \right) - \log_2 C_{M_T, M_r, \beta}\end{aligned}\tag{5.29}$$

Proof: see [RJ07].

Furthermore, if we maintain a system throughput loss of K bps/Hz (i.e., $b = 2$) which corresponds to a power offset of no more than 3 dB with respect to BD with perfect CSI, it is sufficient to scale the number of feedback bits per user [RJ07]

$$B \approx \frac{M_r(M_T - M_r)}{3} \cdot \text{SNR} - \log_2 C_{M_T, M_r, \beta}.\tag{5.30}$$

The number of feedback bits can grow very large for MIMO broadcast channels. Simulation therefore becomes a computationally complex task. However, MIMO systems with a small number of antennas can be simulated in a reasonable amount of time. Figures 5.13 and 5.14 show the achievable throughputs for multi-user MIMO broadcast channels employing BD precoding based on perfect CSI and quantized CSI. The number of feedback bits per user B is scaled according to equation (5.30) (the exact number of B can be found in Table 5.2). The actual power offsets is found to be smaller than 3 dB due to the use of Jensen's inequality

(definition in Appendix E.1) in deriving the throughput loss upper bound in Theorem 5.3.5. The simulations also suggest that the fixed number of feedback bits per user will result in an increasing rate loss with increased SNR.

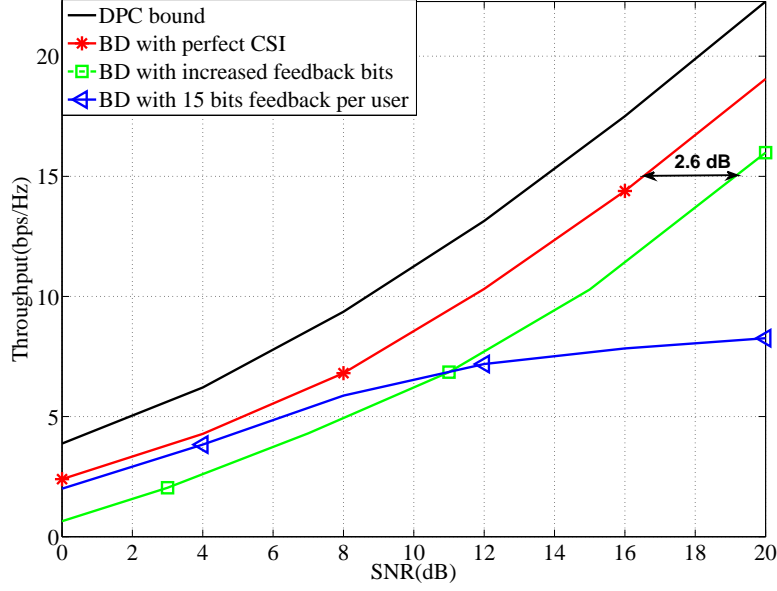


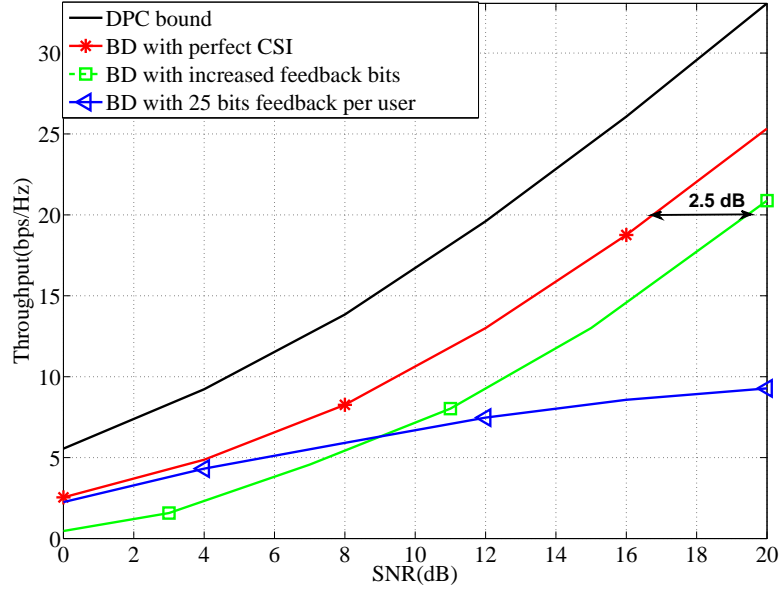
Figure 5.13.: Achievable throughputs for BD $M_T = 4$, $K = 2$, and $M_r = 2$

Table 5.2.: The number of feedback bits B for the case of BD with increased feedback bits. The calculation of B is based on equation (5.30).

SNR	0 dB	4 dB	8 dB	12 dB	16 dB	20 dB
B for $M_T = 4$ and $M_r = 2$	1 bit	6 bits	12 bits	17 bits	22 bits	28 bits
B for $M_T = 6$ and $M_r = 2$	4 bits	15 bits	25 bits	36 bits	47 bits	57 bits

5.3.3. Quantization scheme for RBD based system

In order to correctly perform RBD, the BS requires not only the channel direction information, but also the channel magnitude information which is used to avoid the noise enhancement and improve the diversity. Therefore, the quantization methods proposed in [Jin06, RJ07] cannot be applied to the multi-user MIMO systems employing RBD. In [SRH08], we have presented a new efficient channel quantization scheme to provide both channel direction and magnitude


 Figure 5.14.: Achievable throughputs for BD $M_T = 6$, $K = 3$, and $M_r = 2$

information to the BS.

5.3.3.1. Finite rate feedback model

The new channel quantization scheme is based on the equation

$$\text{vec}\{\hat{\mathbf{H}}_i\} = \arg \min_{\mathbf{w}_j \in \mathcal{C}} d_c^2(\text{vec}\{\mathbf{H}_i\}, \mathbf{w}_j) \quad (5.31)$$

where the vector $\mathbf{w}_j \in \mathbb{C}^{M_{R_i} \cdot M_T \times 1}$ is one codeword of the quantization codebook \mathcal{C} used at the i th user, which is fixed beforehand and known to the transmitter and user i . Here, $\text{vec}\{\mathbf{H}_i\}$ denotes the stacked vector of the channel matrix \mathbf{H}_i and $d_c(\text{vec}\{\mathbf{H}_i\}, \mathbf{w}_j)$ is the distance metric. Here, we consider the chordal distance (details are found in Appendix E.2). It is

$$d_c(\text{vec}\{\mathbf{H}_i\}, \mathbf{w}_j) = \sqrt{\sin^2 \theta_i}, \quad (5.32)$$

where θ_i is the principal angle between the two vectors (i.e., $\theta_i \in [0, \pi/2]$). The matrix $\hat{\mathbf{H}}_i \in \mathbb{C}^{M_{R_i} \times M_T}$ is the quantized version of the channel matrix \mathbf{H}_i . Instead of directly quantizing the channel matrix, we first quantize the stacked vector of the channel matrix \mathbf{H}_i according to equation (5.31), i.e., the codeword which is closest to $\text{vec}\{\mathbf{H}_i\}$ is chosen as $\text{vec}\{\hat{\mathbf{H}}_i\}$. Then

we reshape $\text{vec}\{\hat{\mathbf{H}}_i\}$ to get $\hat{\mathbf{H}}_i$. The advantages of the proposed channel quantization scheme are:

- * To perform quantization, instead of calculating the minimum value of the sum of the principal angles spanned by the columns of the channel matrices, we calculate the minimum value of the angle spanned by the two vectors.
- * By quantizing the stacked vector of the channel matrix, we keep the relative magnitude information for the columns of the channel matrix and avoid the loss of the channel magnitude information which is caused by quantizing the channel matrix as a unitary codeword matrix directly.

The quantization codebook \mathcal{C} consists of 2^B unit norm vectors ($\mathcal{C} = \{\mathbf{w}_1, \dots, \mathbf{w}_{2^B}\}$). Each of the codewords has the dimension $M_{R_i} \cdot M_T \times 1$. Here B is the number of feedback bits per user. Clearly, the choice of the codebook significantly affects the quality of the CSI provided to the BS. Therefore, we study two quantization codebook designs in this part.

5.3.3.2. Random vector quantization (RVQ) codebook

In the RVQ codebook design, we choose 2^B codewords independently and uniformly from the Grassmannian manifold $\mathbb{G}_{n,p}(\mathbb{C})$, which is the set of all p -dimensional planes in the n -dimensional Euclidean space (the definition is found in Appendix E.2). Here, $p = 1$ and $n = M_{R_i} \cdot M_T$. One random codebook is generated for each user. We analyze the performance averaged over all possible random codebooks.

The quantization error associated with all possible random codebooks \mathcal{C} for the quantization of \mathbf{H}_i is defined as

$$D(\mathcal{C}) = \mathbb{E} \left[\min_{\mathbf{w} \in \mathcal{C}} d_c^2(\text{vec}\{\mathbf{H}_i\}, \mathbf{w}) \right] \quad (5.33)$$

In [DLR05] an upper bound to the quantization error is given by

$$\begin{aligned} D(\mathcal{C}) &\leq \bar{D}(\mathcal{C}) \\ &= \frac{\Gamma(\frac{1}{n-1})}{n-1} \cdot (C_{n,p,\beta})^{-\frac{1}{n-1}} \cdot 2^{-\frac{B}{n-1}} + \exp \left[-(2^B \cdot C_{n,p,\beta})^{(1-a)} \right] \end{aligned} \quad (5.34)$$

where $C_{n,p,\beta} = \frac{1}{(n-1)!} \prod_{i=1}^p \frac{\Gamma(\frac{\beta}{2}(n-i+1))}{\Gamma(\frac{\beta}{2}(p-i+1))} = 1$, if complex value is assumed for each codeword (i.e., $\beta = 2$) and $p = 1$. Here, $\Gamma(\cdot)$ represents the Gamma function. Since the second exponential term in equation (5.34) can be neglected for large B [RJ07], the upper bound therefore can be

simplified to

$$\bar{D}(\mathcal{C}) = \frac{\Gamma(\frac{1}{n-1})}{n-1} \cdot 2^{-\frac{B}{n-1}} \quad (5.35)$$

5.3.3.3. Throughput analysis associated with random quantization codebook

A limited feedback of B bits per user ultimately leads to a throughput loss. Let us define the throughput loss $\Delta R_i(\tilde{P}_T)$ to be the difference between the per user throughput achieved by perfect CSI for RBD and quantized CSI for RBD. We have

$$\Delta R_i(\tilde{P}_T) = R_i(\tilde{P}_T) - R_{i,\text{FB}}(\tilde{P}_T), \quad (5.36)$$

where $R_i(\tilde{P}_T)$ is the throughput per user achieved by perfect CSI based RBD and $R_{i,\text{FB}}(\tilde{P}_T)$ refers to the throughput per user achieved by the quantized CSI based RBD. We have derived an upper bound for $\Delta R_i(\tilde{P}_T)$ (this result can also be found in our publication [SRH08]).

Theorem 5.3.7. *The throughput loss per user incurred due to finite rate feedback relative to RBD with perfect CSI can be upper bounded by*

$$\Delta R_i(P_T) \leq \log_2(1 + \Delta I_i + \tilde{P}_T \cdot (\Delta I_i + D(\mathcal{C}))) \quad (5.37)$$

where ΔI_i is a residual MUI relative to the system SNR for the desired user i by using RBD precoding with perfect CSI.

Proof: see Appendix E.3.

5.3.3.3.1. Increased feedback bits

For a fixed number of feedback bits per user B , the finite rate feedback system employing RBD is interference limited at high SNRs like the finite rate feedback systems utilizing ZF and BD precodings. In order to maintain a constant throughput loss, the number of feedback bits per user B should be scaled linearly with the SNR.

Theorem 5.3.8. *In order to maintain a throughput loss no larger than $\log_2 b$ per user between RBD with perfect CSI and quantized CSI, it is sufficient to scale the number of feedback bits per user according to*

$$B = (n-1) \cdot \log_2 \tilde{P}_T - (n-1) \cdot \log_2 (b-1 - \Delta I_i) + (n-1) \cdot \log_2 \left(\frac{\Gamma(\frac{1}{n-1})}{n-1} \right)$$

$$\approx \frac{n-1}{3} \text{SNR} - (n-1) \cdot \log_2(b-1-\Delta I_i) \quad (5.38)$$

This expression can be found by equating the upper bound to the throughput loss in equation (5.37) with $\log_2 b$. Instead of $(\Delta I_i + D(\mathcal{C}))$ in equation (5.37), we use the expression of the quantization error upper bound $\bar{D}(\mathcal{C})$ from equation (5.35). Then we solve the equation for B as a function of \tilde{P}_T and b .

For RBD precoding with perfect CSI, ΔI_i is approximately equal to zero at high SNRs and equal to a small number at low SNRs. In our simulations, a two users MIMO system with $M_T = 4$, $M_{R_i} = 2$ is considered. We average the MUI over the system SNR (i.e., 0 dB - 20 dB) and obtain an experimental value of ΔI_i as 0.25. The bound of the throughput loss can be adjusted by the value of b . Here we set $b = 2$ which means 1 bps/Hz rate offset per user and refers to a bound of 3 dB power offset. In Figure 5.15, we compare the performance of the finite rate feedback system based on our new channel quantization scheme with the channel quantization scheme proposed in [RJ07], which can only provide channel direction information. We use the same number of feedback bits for both channel quantization schemes. It is noticed that the number of feedback bits B can grow very large at high SNRs for MIMO broadcast channels (the exact number of B can be found in Table 5.3). Simulations therefore become a computationally complex task. However, if we consider a MIMO system with a small number of antennas and a wide range of meaningful SNR values, the simulation is finished in a reasonable amount of time.

Table 5.3.: The number of feedback bits B for the case of RBD with increased feedback bits. The calculation of B is based on equation (5.38).

SNR	0 dB	4 dB	8 dB	12 dB	16 dB	20 dB
B for $M_T = 4$ and $M_r = 2$	2 bit	11 bits	21 bits	30 bits	39 bits	48 bits

5.3.3.4. DE-LBG vector quantization codebook

Instead of the RVQ codebook design, we consider another efficient codebook design based on the Linde-Buzo-Gray (LBG) vector quantization algorithm in order to reduce the number of feedback bits. The LBG vector quantization algorithm [LBG80] is an iterative algorithm based on the Lloyd algorithm which is known to provide an alternative systematic approach for the Grassmannian subspace packing problem. Since the LBG vector quantization algorithm was developed for applications on either known probabilistic source descriptions or on a long training sequence of data, the codebooks obtained by the LBG vector quantization algorithm

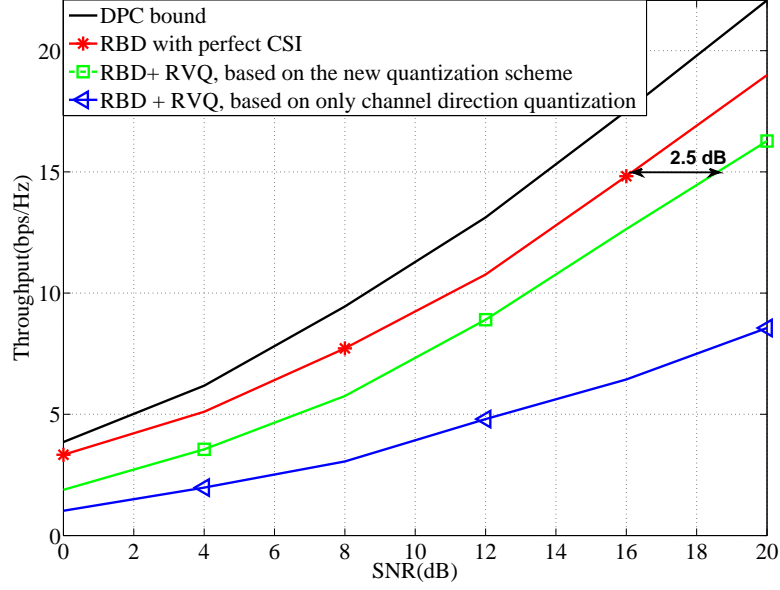


Figure 5.15.: Achievable throughputs for RBD $M_T = 4$, $K = 2$, and $M_r = 2$ with increased B .

can better capture the statistics of the channel than RVQ by using a certain number of channel realizations as a training sequence for the LBG algorithm.

The LBG vector quantization algorithm based codebook design problem can be stated as that we want to find maximally spaced 2^B points in the Grassmannian manifold $\mathbb{G}_{n,p}(\mathbb{C})$ with the given training sequence. We assume the chordal distance as the distortion measure (the definition of the chordal distance is found in Appendix E.2). Then, the minimum chordal distance of the codebook is given by

$$d_{c,\min}(\mathcal{C}) = \min d_c(\mathbf{c}_i, \mathbf{c}_j) \text{ for } \mathbf{c}_i, \mathbf{c}_j \in \mathcal{C}, \forall i \neq j. \quad (5.39)$$

Our aim is to find a codebook of which the minimum chordal distance is maximized. Thus, the codebook design problem can be expressed as follows. With the given training sequence \mathcal{T} and the number of codewords 2^B , find a codebook \mathcal{C} and the encoding region \mathbb{S} such that the $d_{c,\min}(\mathcal{C})$ is maximized.

Compared to the original LBG vector quantization algorithm, we modify the optimality criteria in [LBG80] as follows.

* Nearest Neighbor Condition:

This condition states that the encoding region \mathbb{S}_n should consist of all vectors that are closer to \mathbf{c}_n than any of the other code vectors in the chordal distance sense.

* Centroid Condition:

In contrast to the original LBG algorithm, where the centroid of one encoding region is determined by the arithmetic average of all the training vectors in this region, we modify this condition by requiring the code vector \mathbf{c}_n of the encoding region \mathbb{S}_n to be equal to the dominant eigenvector of the covariance matrix \mathbf{R}_n of all training vectors in this encoding region. This scheme efficiently captures the statistics of the training vectors in the encoding region. Hence we name this modified LBG algorithm as dominant eigenvector (DE)-LBG vector quantization algorithm.

The DE-LBG VQ algorithm, first published in our paper [SRH08], is an iterative algorithm which satisfies the above two optimality criteria. The algorithm requires an initial codebook $\mathcal{C}^{(0)}$ which is obtained by the splitting method. In this method an initial code vector is split into two code vectors. The iterative algorithm is run with these two vectors as the initial codebook. At the end of this step the two code vectors are split into four and the iterative algorithm is run again. The process is repeated until the desired number of code vectors is obtained. The number of iterations depends on the number of required codewords. The algorithm is summarized below.

1. Generate a training sequence \mathcal{T} which captures the statistical properties of the stacked vectors $\text{vec}\{\mathbf{H}_i\} \in \mathbb{C}^{M_{\text{R}_i} \cdot M_{\text{T}} \times 1}$ of channel matrix samples. Channel matrix samples are generated by Monte-Carlo simulations in this work.

$$\mathcal{T} = \{\text{vec}\{\mathbf{H}_1\}, \text{vec}\{\mathbf{H}_2\}, \dots, \text{vec}\{\mathbf{H}_M\}\} \quad (5.40)$$

Here M is the number of channel samples.

2. Generate the initial code vector $\check{\mathbf{c}}_1$ by choosing it as the dominant eigenvector of the covariance matrix \mathbf{R} of the entire training sequence. Set $N = 1$.

$$\mathbf{R} = \frac{1}{M} \sum_{m=1}^M \text{vec}\{\mathbf{H}_m\} \text{vec}\{\mathbf{H}_m\}^H \quad (5.41)$$

3. Splitting: for $i = 1, 2, \dots, N$, set $\mathbf{c}_i^{(0)} = (1 + \epsilon)\check{\mathbf{c}}_i$, $\mathbf{c}_{N+i}^{(0)} = (1 - \epsilon)\check{\mathbf{c}}_i$, here $\epsilon > 0$ is a very small number and we choose $\epsilon = 0.002$. Then set $N \leftarrow 2N$.

4. Codeword update: set the iteration index $k = 0$ and calculate the minimum chordal distance of the initial codebook

$$d_{c,\min}^{(0)}(\mathcal{C}) = \min d_c(\mathbf{c}_i, \mathbf{c}_j) \quad (5.42)$$

- (a) Assign the source vector $\text{vec}\{\mathbf{H}_m\}$ to the N encoding regions by finding

$$\tilde{n} = \arg \min_{n \in 1, 2, \dots, N} d_c(\text{vec}\{\mathbf{H}_m\}, \mathbf{c}_n^{(k)}) \text{ for } m = 1, 2, \dots, M.$$

- (b) Update the code vector in each region by using the centroid condition.

- (c) Set $k \leftarrow k + 1$

- (d) Calculate $d_{c,\min}^{(k)}(\mathcal{C})$, if $d_{c,\min}^{(k)}(\mathcal{C}) > d_{c,\min}^{(k-1)}(\mathcal{C})$ go back to step (a). Otherwise, go to step (e).

- (e) Set $\check{\mathbf{c}}_n = \mathbf{c}_n^{(k-1)}$ as the final code vectors.

5. Repeat steps 3 and 4 until the desired number of code vectors is obtained.

In Figure 5.16 we can see the system performance improvement of RBD with DE-LBG comparing it with RVQ for a fixed number of feedback bits, where uncorrelated flat fading channels are considered.

The DE-LBG vector quantization can be applied to an OFDM system. We assume a microcellular scenario based on the Manhattan grid where users with fixed velocities ($|v| \leq 10$ km/h) are randomly distributed in the streets. An OFDM channel with 128 subcarriers and 30 symbols is considered. The total bandwidth is 5.86 MHz and the carrier frequency $f_c = 3.95$ GHz. RBD precoding is performed per chunk which contains 8 subcarriers and 15 symbols. The channel model is the WINNER B1 channel [WIN06] (i.e., urban micro-cell scenario). The chunk equivalent channel (i.e., the average of all channels corresponding to all the symbols in one chunk, the calculation is based on equation (5.5).) is quantized by the DE-LBG vector quantization codebook.

The result is shown in Figure 5.17. The complementary CDF of the cell throughput shows us that with the DE-LBG codebook design, which can be adapted to the statistics of the channel matrices, the system performance is still not significantly degraded for only 7 feedback bits per chunk for one user compared to the case that the transmitter has perfect CSI.

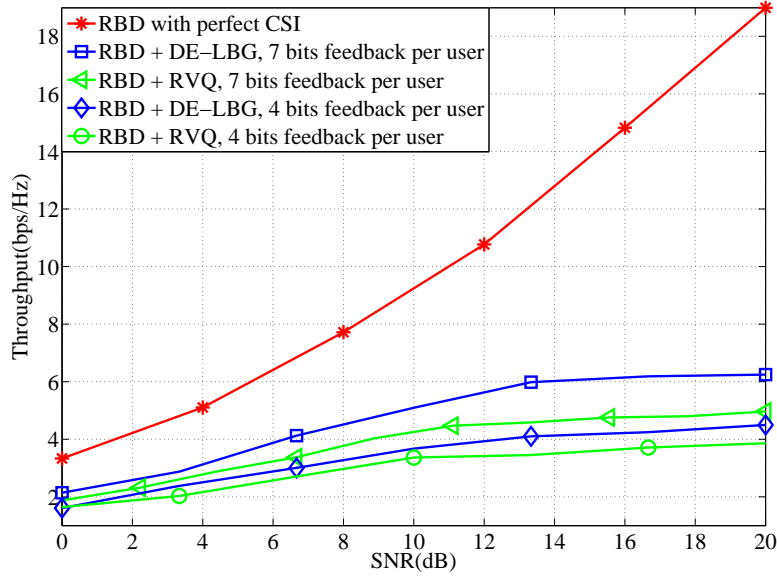


Figure 5.16.: Achievable throughputs for RBD $M_T = 4$, $K = 2$, and $M_r = 2$ with fixed B in uncorrelated flat fading channels

5.4. Blind channel estimation using a tensor-based subspace method

Multi-dimension data analysis has attracted much attention in several scientific fields, such as psychometrics [CC70], chemometrics [AB03], array signal processing [SBG00], communications [MHS04, JLL09], etc. The main reason for its popularity is that the tensor-based signal processing features some significant tensor gains compared to the matrix-based counterparts. In this section we use tensors to solve blind channel estimation problems for SIMO and MIMO multipath fading channels. At the beginning of this section we review the fundamentals of tensor algebra that are used throughout this section such as the notations and operations of tensors and one important tensor decomposition (namely the Higher-Order SVD).

5.4.1. Notations and operations of tensor

In signal processing, a tensor is defined as a multi-dimensional array. More formally, an R -way or R th-order tensor is an element of the tensor product of R vector spaces, each of which has its own coordinate system. In this aspect, an R -D tensor is defined as a collection of numbers

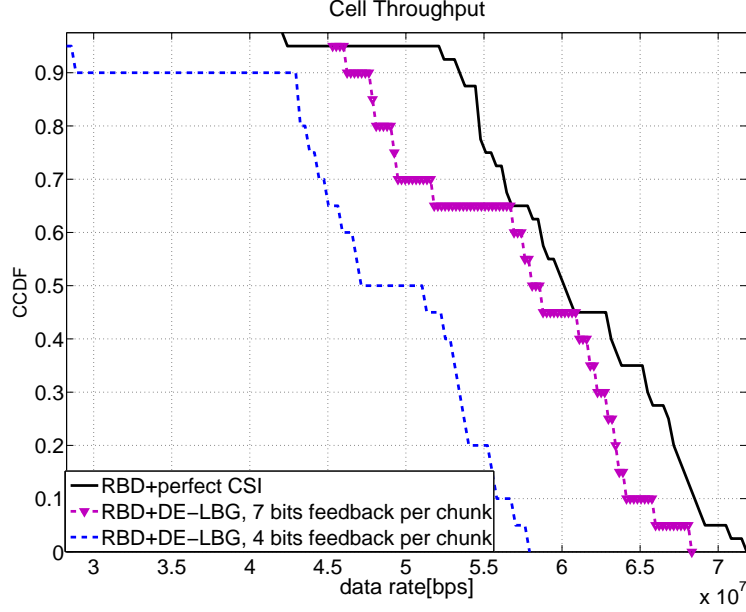


Figure 5.17.: Achievable throughputs for RBD with DE-LBG codebooks for $M_T = 8$, $K = 4$, and $M_{R_i} = 2$ for $\forall i$, OFDM, WINNER channel model B1.

referenced by R indices. Up to $R = 2$, tensors are the same as matrices. For $R > 2$, new notations and operations are required.

An R -D tensor $\mathcal{A} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R}$ is an R -way array which has the size M_r along mode r for $r = 1, 2, \dots, R$. The total number of elements of the tensor \mathcal{A} is denoted as $M = \prod_{r=1}^R M_r$. The tensor operations we use in this section are consistent with [dLdMV00a].

- **The r -mode vectors** of a tensor \mathcal{A} are obtained by varying the r -th index, while keeping all other indices fixed. They represent the generalization of row vectors and column vectors of matrices.
- **The r -mode subspace** or short **r -space** of \mathcal{A} is the vector space spanned by the r -mode vectors.
- **The r -mode unfolding of \mathcal{A}** is obtained by collecting all r -mode vectors as the columns of a matrix with the dimension $M_r \times \frac{M}{M_r}$ and represented by $[\mathcal{A}]_{(r)}$. The ordering of the columns in the r -mode unfolding defines how to arrange the remaining $(R - 1)$ indices. Any permutation of these is a valid unfolding, as long as it is utilized consistently. There are several popular choices:

1. Forward cyclical: Start with the $(r + 1)$ -th index, then the $(r + 2)$ -th index, all the way to the R -th index, then start from the first and proceed forward up to the $(r - 1)$ -th index. This convention was proposed by [dLdMV00a] and has been popularly used as the standard in the signal processing community. Therefore, we choose this column ordering in this section. In Figure 5.18, we visualize this unfolding for a 3-D tensor as an example.
 2. Reverse cyclical: Like forward cyclical, starting with the $(r - 1)$ -th index and proceeding backward up to the $(r + 1)$ -th index.
 3. Forward column ordering: Start with the first index, then the second, up to the $(r - 1)$ -th index, then continue with $(r + 1)$ -th all the way up to the R -th index. It is somehow the most natural way to collect the columns, since it coincides with the standard way to store multi-dimensional data in the memory.
 4. Reverse column ordering: Like forward but in reverse ordering, the columns are collected by starting with the R -th index and proceeding backwards to the first.
- **The r -rank of \mathcal{A}** is defined as the rank of $[\mathcal{A}]_{(r)}$. Note that in general, all the r -ranks of a tensor \mathcal{A} can be different.
 - **The r -mode product** of a tensor \mathcal{A} and a matrix $U_r \in \mathbb{C}^{J_r \times M_r}$ is denoted as $\mathcal{B} = \mathcal{A} \times_r U_r$ which is an $(M_1 \times M_2 \times \cdots \times M_{(r-1)} \times J_r \times M_{(r+1)} \times \cdots \times M_R)$ tensor. Such transformation can be expressed by multiplying all r -mode vectors of \mathcal{A} from the left-hand side by the matrix U_r , i.e., $[\mathcal{B}]_{(r)} = U_r [\mathcal{A}]_{(r)}$. Figure 5.19 visualizes the equation $\mathcal{B} = \mathcal{A} \times_1 U_1 \times_2 U_2 \times_3 U_3$ for the 3-way tensors $\mathcal{A} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$ and $\mathcal{B} \in \mathbb{C}^{J_1 \times J_2 \times J_3}$. Visualization schemes like Figure 5.19 have proven to be very useful to gain insight into tensor techniques.
 - The **concatenation** of two tensors along the r -th mode is symbolized via $[\mathcal{A} \sqcup_r \mathcal{B}]$. The r -mode vectors of the resulting tensor are given by the r -mode vectors of \mathcal{A} stacked on top of the r -mode vectors of \mathcal{B} . In other words, we have

$$[\mathcal{A} \sqcup_r \mathcal{B}]_{(r)} = \begin{bmatrix} [\mathcal{A}]_{(r)} \\ [\mathcal{B}]_{(r)} \end{bmatrix}. \quad (5.43)$$

5.4.2. Higher-Order SVD decomposition

The Higher-Order SVD (HOSVD) can be view as a special case of a Tucker3 decomposition, which has been known since [Tuc66]. Tucker3 is basically a 3-mode Principle Component

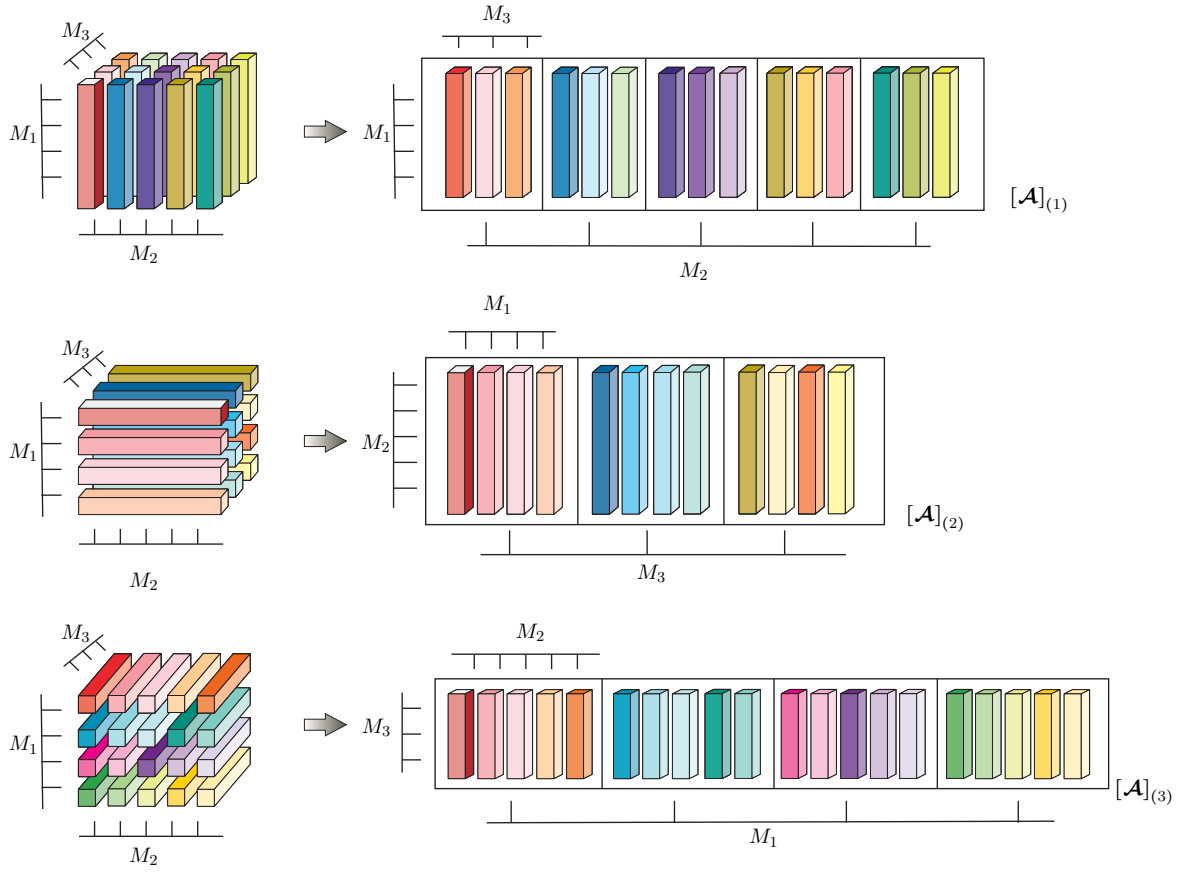


Figure 5.18.: Unfolding of the tensor $\mathcal{A} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$ to the matrix $[\mathcal{A}]_{(1)} \in \mathbb{C}^{M_1 \times M_2 M_3}$, the matrix $[\mathcal{A}]_{(2)} \in \mathbb{C}^{M_2 \times M_3 M_1}$, and the matrix $[\mathcal{A}]_{(3)} \in \mathbb{C}^{M_3 \times M_1 M_2}$. Here, $M_1 = 4$, $M_2 = 5$, and $M_3 = 3$.

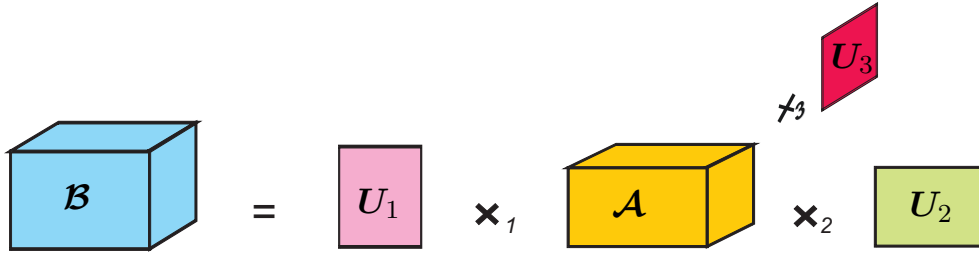


Figure 5.19.: Multiplication of a 3-way tensor $\mathcal{A} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$ with matrices $\mathbf{U}_1 \in \mathbb{C}^{J_1 \times M_1}$, $\mathbf{U}_2 \in \mathbb{C}^{J_2 \times M_2}$, and $\mathbf{U}_3 \in \mathbb{C}^{J_3 \times M_3}$.

Analysis (PCA) and can be expressed as

$$\mathcal{A} = \mathcal{G} \times_1 \mathbf{V}_1 \times_2 \mathbf{V}_2 \times_3 \mathbf{V}_3, \quad (5.44)$$

where the tensor \mathcal{A} has the dimension $M_1 \times M_2 \times M_3$, the matrices $\mathbf{V}_r \in \mathbb{C}^{M_r \times p_r}$ are the factor matrices of the decomposition, and the tensor $\mathcal{G} \in \mathbb{C}^{p_1 \times p_2 \times p_3}$ is denoted as the core tensor. The term p_r refers to the r -rank of \mathcal{A} . If the matrices \mathbf{V}_r are chosen to contain a basis for the r -space of \mathcal{A} , the decomposition is exact. Typically, the factor matrices \mathbf{V}_r are chosen to be unitary but this is not required. Therefore, the decomposition is not unique, because every factor matrix can be post-multiplied with a square invertible matrix \mathbf{T}_r when the inverse matrix \mathbf{T}_r^{-1} are absorbed into the new core tensor.

The HOSVD is introduced in [dLdMV00a] by simplifying the Tucker3 decomposition. In the HOSVD decomposition, the factor matrices \mathbf{V}_r are chosen as the p_r dominant left singular vectors of the r -mode unfolding $[\mathcal{A}]_{(r)}$. To distinguish this special Tucker3 decomposition from the general case shown in equation (5.44) we use the notations \mathbf{U}_r and \mathcal{S} to indicate the factor matrices and the core tensor for HOSVD, respectively. Thus, the HOSVD for the general R -D tensor case can be written as

$$\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \cdots \times_R \mathbf{U}_R, \quad (5.45)$$

where $\mathbf{U}_r \in \mathbb{C}^{M_r \times p_r}$ and $\mathcal{S} \in \mathbb{C}^{p_1 \times p_2 \times \cdots \times p_R}$. The core tensor \mathcal{S} satisfies the so-called “all-orthogonality” condition, which means that all the unfoldings of \mathcal{S} are row-orthogonal matrices. It can be written as

$$[\mathcal{S}]_{(r)} \cdot [\mathcal{S}]_{(r)}^H = \text{diag} \left\{ \left[\sigma_1^{(r)^2}, \sigma_2^{(r)^2}, \dots, \sigma_{p_r}^{(r)^2} \right] \right\}. \quad (5.46)$$

Moreover, $\sigma_n^{(r)}$ indicate the r -mode singular values for $n = 1, \dots, p_r$ and $r = 1, \dots, R$, which are Euclidean norms of these rows. Like for the SVD of the matrix, they appear ordered by magnitude, i.e., $\sigma_1^{(r)} \geq \sigma_2^{(r)} \geq \dots \sigma_{p_r}^{(r)}$ for $r = 1, \dots, R$. The R sets of r -mode singular values are in general different in the tensor case. It is noticed that the orthogonality of the factor matrices \mathbf{U}_r and the all-orthogonality of the core tensor \mathcal{S} are the basic assumptions of the HOSVD.

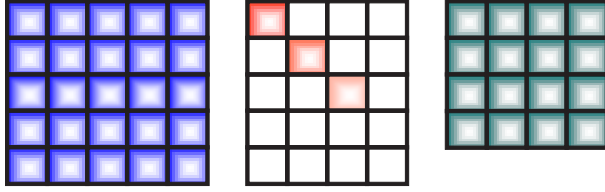
The HOSVD has numerous applications in the area of signal processing. The famous applications are face recognition [VT02], data mining with application to network modeling [STF06], social network analysis [ACKY05], multi-dimensional harmonic retrieval [HRD08], and pattern recognition [SE07]. Generally, these applications apply the HOSVD to perform rank reduction in individual modes and then utilize parts of the decomposition such as the core tensor, the factor matrices, or combinations thereof.

The comparison of the matrix SVD in Table 5.4 and the tensor HOSVD in Table 5.5 shows a clear analogy between the two cases. The essential difference is that the core tensor \mathcal{S} is in

Table 5.4.: Variations of the matrix SVD

Singular Value Decomposition (SVD) of the matrix $A \in \mathbb{C}^{M \times N}$, $\text{rank}(A) = r$

“Full SVD”: $A = U \cdot \Sigma \cdot V^H$

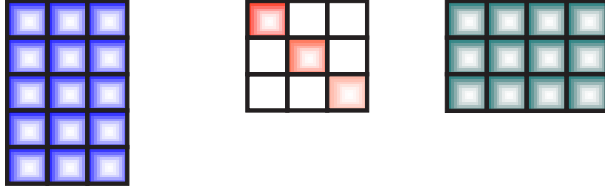


$$U \in \mathbb{C}^{M \times M}$$

$$\Sigma \in \mathbb{C}^{M \times N}$$

$$V^H \in \mathbb{C}^{N \times N}$$

“Economy size SVD”: $A = U_s \cdot \Sigma_s \cdot V_s^H$

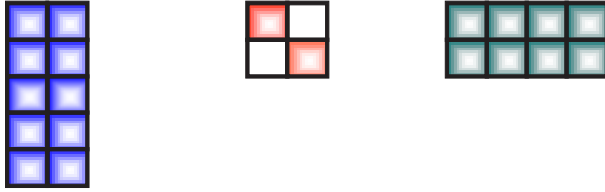


$$U_s \in \mathbb{C}^{M \times r}$$

$$\Sigma_s \in \mathbb{C}^{r \times r}$$

$$V_s^H \in \mathbb{C}^{r \times N}$$

“Low-rank approximation (truncated SVD)”: $A \approx U_s \cdot \Sigma_s \cdot V_s^H$

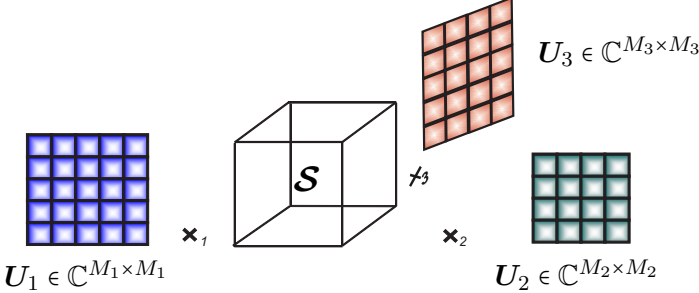
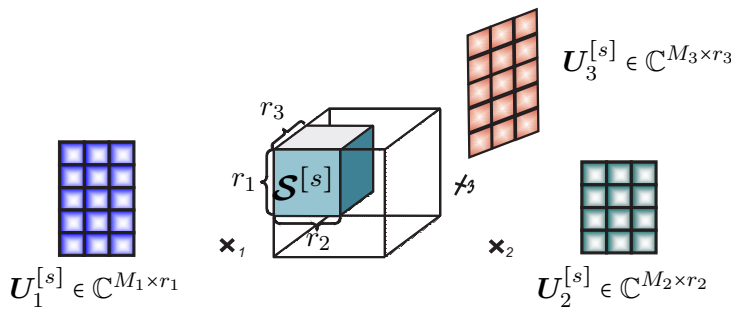
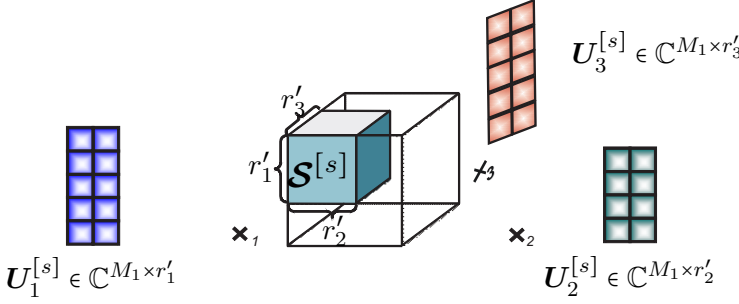


$$U_s \in \mathbb{C}^{M \times r'}$$

$$\Sigma_s \in \mathbb{C}^{r' \times r'}$$

$$V_s^H \in \mathbb{C}^{r' \times N}$$

Table 5.5.: Variations of the tensor HOSVD

Higher-Order SVD of the tensor $\mathcal{A} \in \mathbb{C}^{M_1 \times M_2 \times M_3}$, $n\text{-rank}(\mathcal{A}) = r_n$, $n = 1, 2, 3$	
“Full HOSVD”: $\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3$	
 <p> $\mathbf{U}_1 \in \mathbb{C}^{M_1 \times M_1}$ $\mathbf{U}_2 \in \mathbb{C}^{M_2 \times M_2}$ $\mathbf{U}_3 \in \mathbb{C}^{M_3 \times M_3}$ </p>	
“Economy size HOSVD”: $\mathcal{A} = \mathcal{S}^{[s]} \times_1 \mathbf{U}_1^{[s]} \times_2 \mathbf{U}_2^{[s]} \times_3 \mathbf{U}_3^{[s]}$	
 <p> $\mathbf{U}_1^{[s]} \in \mathbb{C}^{M_1 \times r_1}$ $\mathbf{U}_2^{[s]} \in \mathbb{C}^{M_2 \times r_2}$ $\mathbf{U}_3^{[s]} \in \mathbb{C}^{M_3 \times r_3}$ </p>	
“Low-rank approximation (truncated HOSVD)”: $\mathcal{A} \approx \mathcal{S}^{[s]} \times_1 \mathbf{U}_1^{[s]} \times_2 \mathbf{U}_2^{[s]} \times_3 \mathbf{U}_3^{[s]}$	
 <p> $\mathbf{U}_1^{[s]} \in \mathbb{C}^{M_1 \times r'_1}$ $\mathbf{U}_2^{[s]} \in \mathbb{C}^{M_2 \times r'_2}$ $\mathbf{U}_3^{[s]} \in \mathbb{C}^{M_3 \times r'_3}$ </p>	

general a full tensor, instead of being pseudo-diagonal (i.e., nonzero elements only occur when the indices are same) in the matrix case. The r -mode singular values are positive and real like in the matrix case, on the other hand the entries of \mathcal{S} are not necessarily positive in general. They can be complex, when the tensor \mathcal{A} is a complex-valued tensor.

The Eckart-Young theorem [EY36] shows that the truncated SVD provides the best low-rank approximation of the given matrix in the Frobenius sense. Unfortunately, such a theorem does not exist for the HOSVD. In fact, the truncated HOSVD does provide a good low-rank approximation which is though not optimal in general. In the application of denoising a low-rank desired signal from a noisy observation, the truncated HOSVD is asymptotically optimal in the high signal-to-noise ratio (SNR) regime [dLdMV00b]. For this reason, we only consider the truncated HOSVD for subspace estimation. In matrix-based subspace estimation strategies, a stacking operation is applied to the observed data such that each column of the measurement matrix represents a stacked version of one snapshot of the multi-dimensional signal. However, such a representation does not account for the structure inherent in the data. Therefore, we can improve the subspace estimation accuracy by performing a “structured” subspace estimation based on the HOSVD. In the following we provide a short introduction of the matrix-based subspace estimation and the tensor-based subspace estimation under the observation of the N snapshots of a low-rank data of rank d .

5.4.2.1. Matrix-based subspace estimation

We consider an observation of the N snapshots of low-rank data of rank d on an R -dimensional grid of size $M_1 \times M_2 \times \dots \times M_R$ under additive noise.

Let us define an array steering matrix $\mathbf{A}^{(r)}$ for $r = 1, 2, \dots, R$, which is given by

$$\mathbf{A}^{(r)} = [\mathbf{a}_1^{(r)}, \mathbf{a}_2^{(r)}, \dots, \mathbf{a}_d^{(r)}] \in \mathbb{C}^{M_r \times d}. \quad (5.47)$$

In the classical matrix method, a meaningful definition of an R -D measurement matrix \mathbf{X} is obtained by stacking all steering matrices $\mathbf{A}^{(r)}$ along the rows and align the snapshots along the columns. This stacking operation allows us to write the observation in matrix form

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S} + \mathbf{N}. \quad (5.48)$$

Here, the matrix $\mathbf{A} \in \mathbb{C}^{M \times d}$ is the stacked version of all steering matrices in the form of $\mathbf{A} = \mathbf{A}^{(1)} \diamond \mathbf{A}^{(2)} \diamond \dots \diamond \mathbf{A}^{(R)}$ where \diamond indicates the Khatri-Rao product and $M = \prod_{r=1}^R M_r$. The definition of the matrix $\mathbf{S} \in \mathbb{C}^{d \times N}$ depends on the applications. The matrix $\mathbf{X} \in \mathbb{C}^{M \times N}$ is composed of the stacked measurements. The noise samples are collected in the matrix

$\mathbf{N} \in \mathbb{C}^{M \times N}$. It is obvious that the stacking operation does not capture the structure inherent in the lattice that is used to sample the data.

Let us define $\mathbf{X}_0 = \mathbf{A}\mathbf{S}$ to be the noise-free part of the data, which is at most rank- d . Therefore, we can perform a rank- d -approximation of \mathbf{X} via the truncated SVD. The column space of the matrix \mathbf{X} spanned by the d dominant left singular vectors is an estimate of the true signal subspace. It is also identical to the space spanned by the columns of \mathbf{A} . Note that the true signal subspace is spanned by the d dominant left singular vectors of \mathbf{X}_0 , i.e., all columns of \mathbf{X}_0 lie in the signal subspace. The SVDs of \mathbf{X}_0 and \mathbf{X} are expressed as

$$\mathbf{X}_0 = \begin{bmatrix} \underbrace{\mathbf{U}_s}_{M \times d} & \underbrace{\mathbf{U}_n}_{M \times (M-d)} \end{bmatrix} \cdot \begin{bmatrix} \underbrace{\boldsymbol{\Sigma}_s}_{d \times d} & \mathbf{0}_{d \times (N-d)} \\ \mathbf{0}_{(M-d) \times d} & \mathbf{0}_{(M-d) \times (N-d)} \end{bmatrix} \cdot [\mathbf{V}_s \quad \mathbf{V}_n]^H \quad (5.49)$$

$$\mathbf{X} = \begin{bmatrix} \underbrace{\hat{\mathbf{U}}_s}_{M \times d} & \underbrace{\hat{\mathbf{U}}_n}_{M \times (M-d)} \end{bmatrix} \cdot \begin{bmatrix} \underbrace{\hat{\boldsymbol{\Sigma}}_s}_{d \times d} & \mathbf{0}_{d \times (N-d)} \\ \mathbf{0}_{(M-d) \times d} & \underbrace{\hat{\boldsymbol{\Sigma}}_n}_{(M-d) \times (N-d)} \end{bmatrix} \cdot [\hat{\mathbf{V}}_s \quad \hat{\mathbf{V}}_n]^H, \quad (5.50)$$

where \mathbf{U}_s and \mathbf{U}_n are orthonormal basis for the signal subspace and the noise subspace, and $\hat{\mathbf{U}}_s$ and $\hat{\mathbf{U}}_n$ are their estimates obtained from \mathbf{X} . Furthermore, $\boldsymbol{\Sigma} = \text{diag} \{[\sigma_1, \sigma_2, \dots, \sigma_d]\}$ is composed of the d non-zero real singular values on its main diagonal.

5.4.2.2. Tensor-based subspace estimation

In order to capture the natural structure inherent in the observation, we employ the HOSVD for the subspace estimation. As discussed before, the HOSVD is easily computed via SVDs of the tensor unfoldings. Moreover, the truncated HOSVD allows for multilinear low-rank approximation in a manner similar to the truncated SVD.

By replacing the measurement matrix \mathbf{X} with a measurement tensor $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R \times N}$, the observation can be modeled as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{A}} \times_{R+1} \mathbf{S}^T + \boldsymbol{\mathcal{N}}, \quad (5.51)$$

where the matrix \mathbf{S} is the same as in equation (5.48), and the tensor $\boldsymbol{\mathcal{N}}$ contains the noise samples. The array steering tensor $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R \times d}$ is constructed in the following manner

$$\boldsymbol{\mathcal{A}} = [\boldsymbol{\mathcal{A}}_1 \sqcup_{R+1} \boldsymbol{\mathcal{A}}_2 \dots \sqcup_{R+1} \boldsymbol{\mathcal{A}}_d]. \quad (5.52)$$

Here, \mathcal{A}_i denotes the R -dimensional array steering tensor of the i th signal which is given by

$$\mathcal{A}_i = \mathbf{a}^{(1)}(\mu_i^{(1)}) \circ \mathbf{a}^{(2)}(\mu_i^{(2)}) \circ \dots \circ \mathbf{a}^{(R)}(\mu_i^{(R)}), \quad (5.53)$$

where \circ indicates the outer product operator and $i = 1, 2, \dots, d$.

Compared to the matrix data model, we have the following identities

$$\begin{aligned} \mathbf{A} &= [\mathcal{A}]_{(R+1)}^T \\ \mathbf{X} &= [\mathcal{X}]_{(R+1)}^T \\ \mathbf{N} &= [\mathcal{N}]_{(R+1)}^T. \end{aligned} \quad (5.54)$$

Let \mathcal{X}_0 be the noise-free observation, such that $\mathcal{X} = \mathcal{X}_0 + \mathcal{N}$. The tensor \mathcal{X}_0 is of rank d and, therefore, all the r -ranks are at most equal to d [dLdMV00a]. Thus, we can express \mathcal{X}_0 and \mathcal{X} in terms of the truncated HOSVD as

$$\mathcal{X}_0 = \mathcal{S}^{[s]} \times_1 \mathbf{U}_1^{[s]} \times_2 \mathbf{U}_2^{[s]} \dots \times_{R+1} \mathbf{U}_{R+1}^{[s]} \quad (5.55)$$

$$\mathcal{X} \approx \hat{\mathcal{S}}^{[s]} \times_1 \hat{\mathbf{U}}_1^{[s]} \times_2 \hat{\mathbf{U}}_2^{[s]} \dots \times_{R+1} \hat{\mathbf{U}}_{R+1}^{[s]}. \quad (5.56)$$

Here, the matrices \mathbf{U}_r and $\hat{\mathbf{U}}_r$ for $r = 1, 2, \dots, R+1$ are obtained from the SVD of the r -th unfolding of \mathcal{X}_0 and \mathcal{X} which read as

$$[\mathcal{X}_0]_{(r)} = \begin{bmatrix} \underbrace{\mathbf{U}_r^{[s]}}_{M_r \times p_r} & \underbrace{\mathbf{U}_r^{[n]}}_{M_r \times (M_r - p_r)} \end{bmatrix} \cdot \begin{bmatrix} \underbrace{\boldsymbol{\Sigma}_r^{[s]}}_{d \times d} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot [\mathbf{V}_r^{[s]} \quad \mathbf{V}_r^{[n]}]^H \quad (5.57)$$

$$[\mathcal{X}]_{(r)} = \begin{bmatrix} \underbrace{\hat{\mathbf{U}}_r^{[s]}}_{M_r \times p_r} & \underbrace{\hat{\mathbf{U}}_r^{[n]}}_{M_r \times (M_r - p_r)} \end{bmatrix} \cdot \begin{bmatrix} \underbrace{\hat{\boldsymbol{\Sigma}}_r^{[s]}}_{d \times d} & \mathbf{0} \\ \mathbf{0} & \underbrace{\hat{\boldsymbol{\Sigma}}_r^{[n]}}_{(M-d) \times (N-d)} \end{bmatrix} \cdot [\hat{\mathbf{V}}_r^{[s]} \quad \hat{\mathbf{V}}_r^{[n]}]^H, \quad (5.58)$$

where $\mathbf{U}_r^{[s]}$ and $\mathbf{U}_r^{[n]}$ are the basis for the r -space and its orthogonal complement, respectively. The term p_r denotes the r -rank of \mathcal{X}_0 which can be individually estimated via a model order selection scheme operating on all unfoldings individually. Alternatively, the tensor-based model order selection schemes [dCRHdSJ11] can be applied to estimate d and then decide $p_r = \min(M_r, d)$. To compare the truncated HOSVD of \mathcal{X} with the truncated SVD of \mathbf{X} , it is obvious that the HOSVD performs low-rank approximation in all $R+1$ modes. Hence, the multilinear structure is exploited to perform more efficient denoising.

In the matrix case, $\hat{\mathbf{U}}_s \in \mathbb{C}^{M \times d}$ represents the basis of the estimated signal subspace. As a multidimensional extension of $\hat{\mathbf{U}}_s$ we define a *signal subspace tensor* $\hat{\mathbf{u}}^{[s]} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R \times d}$ in the following manner [HRD08]

$$\hat{\mathbf{u}}^{[s]} = \hat{\mathbf{S}}^{[s]} \times_1 \hat{\mathbf{U}}_1^{[s]} \times_2 \hat{\mathbf{U}}_2^{[s]} \dots \times_R \hat{\mathbf{U}}_R^{[s]}. \quad (5.59)$$

The relationship between $\hat{\mathbf{U}}_s$ and $\hat{\mathbf{u}}^{[s]}$ is given by the following theorem:

Theorem 5.4.1. *The HOSVD-based signal subspace estimate $\left[\hat{\mathbf{u}}^{[s]}\right]_{(R+1)}^T$ can be expressed by the SVD-based subspace estimate $\hat{\mathbf{U}}_s$ as [RBHW09]*

$$\left[\hat{\mathbf{u}}^{[s]}\right]_{(R+1)}^T = (\hat{\mathbf{T}}_1 \otimes \hat{\mathbf{T}}_2 \dots \otimes \hat{\mathbf{T}}_R) \cdot \hat{\mathbf{U}}_s, \quad (5.60)$$

where \otimes indicates the Kronecker product and the matrices $\hat{\mathbf{T}}_r \in \mathbb{C}^{M_r \times M_r}$ are the estimates of the projection matrices onto the r -spaces of \mathbf{X}_0 , which are computed as $\hat{\mathbf{T}}_r = \hat{\mathbf{U}}_r^{[s]} \hat{\mathbf{U}}_r^{[s]H}$.

Proof: see [RBHW09]

The relation (5.60) shows that the HOSVD-based subspace estimate can be seen as the projection of the matrix-based subspace estimate onto the Kronecker structure inherent in the data. In the presence of noise, if the rank of signal d is strictly less than the number of sensors M_r in at least one of the R modes, the estimated signal subspace given by $\left[\hat{\mathbf{u}}^{[s]}\right]_{(R+1)}^T$ is improved and differs from the estimated signal subspace given by $\hat{\mathbf{U}}_s$ due to the denoising which is performed by filtering out the part of the noise that does not obey the required Kronecker structure. On the other hand, if $d \geq M_r$ for any mode r , we have $\hat{\mathbf{T}}_r = \mathbf{I}_r$ and hence no performance improvement can be obtained in this particular mode r . Moreover, if $d \geq \max_{r=1,2,\dots,R} (M_r)$, we have $\left[\hat{\mathbf{u}}^{[s]}\right]_{(R+1)}^T = \hat{\mathbf{U}}_s$, i.e., if the number of signals d is greater than or equal to the number of sensors in all R modes, we cannot achieve a performance improvement in terms of the subspace estimation accuracy from the HOSVD-based subspace estimate.

The truncated HOSVD based subspace estimation has been applied to various areas in signal processing. For instance, multi-dimensional model order selection [dCRHdSJ11], direction of arrival (DOA) estimation [HRD08, THG09], and multi-dimensional channel estimation with training data [RH10]. In the following subsections, we discuss one additional application for the blind channel estimation. The achieved tensor gain is demonstrated numerically.

5.4.3. Tensor-based blind estimation of SIMO channels

The blind channel estimation scheme in [MDCM95] has been viewed as a pioneering work which provides an estimation of SIMO channels based on the second-order statistics of the measurement data. In this subsection, we extend this matrix-based blind channel estimation to the tensor case [SRH10a]. We use a 3-way tensor to model the measurement data and employ the truncated HOSVD to obtain the improved signal subspace estimate.

We consider a SIMO system where the receiver is equipped with M_R receive antennas. The channel between each transmit and receive antenna pair is modeled as a finite impulse response (FIR) filter with $L + 1$ taps. Let $s[k]$ denote the symbol emitted over the transmit antenna at time kT . Here T indicates the symbol duration. The discrete-time signal experiences the unknown channel which is assumed to be time-invariant during the observation interval. Then, the received signal vector at time kT can be obtained by the convolution of the transmit signals and the channel impulse responses

$$\mathbf{y}[k] = \sum_{\ell=0}^L \mathbf{h}_\ell s[k - \ell] + \mathbf{n}[k] \in \mathbb{C}^{M_R}. \quad (5.61)$$

Here, $\mathbf{h}_\ell \in \mathbb{C}^{M_R}$ contains the coefficients of the channel impulse responses corresponding to the ℓ th channel tap. The elements of the noise vector $\mathbf{n}[k]$ are circularly symmetric complex Gaussian distributed with variance σ_n^2 and assumed mutually uncorrelated in space and time.

5.4.3.1. Matrix-based subspace method for SIMO channels

In practice, the measurement data is observed during consecutive data windows over all receive antennas. We use W to indicate the length of the observed data window. The matrix-based subspace method for the blind estimation of SIMO channels stacks the dimensions of M_R receive antennas and the data window length W into one highly structured $M_R \cdot W \times 1$ vector. Thus, the measurement data with respect to the n th observation data window is given by

$$\mathbf{y}_n = \mathbf{H}_T \mathbf{s}_n + \mathbf{n}, \quad (5.62)$$

where the vector $\mathbf{s}_n = [s[nW], s[nW-1], \dots, s[nW-W-L+1]]^T$ denotes the input data sequence with dimension $(W+L) \times 1$. The vector \mathbf{y}_n has the structure

$$\mathbf{y}_n = \begin{bmatrix} \mathbf{y}_n^{(1)} \\ \mathbf{y}_n^{(2)} \\ \vdots \\ \mathbf{y}_n^{(M_R)} \end{bmatrix} \in \mathbb{C}^{M_R \cdot W \times 1}, \quad (5.63)$$

where $\mathbf{y}_n^{(j)}$ for $j = 1, 2, \dots, M_R$ has the dimension $W \times 1$. The matrix \mathbf{H}_T is the filtering matrix and structured as

$$\mathbf{H}_T = \begin{bmatrix} \mathbf{H}_T^{(1)} \\ \mathbf{H}_T^{(2)} \\ \vdots \\ \mathbf{H}_T^{(M_R)} \end{bmatrix} \in \mathbb{C}^{M_R \cdot W \times (W+L)}. \quad (5.64)$$

Here, the matrices $\mathbf{H}_T^{(j)}$ indicate the banded Toeplitz matrix associated to the j th receive antenna's impulse response $\mathbf{h}^{(j)}$. The vector $\mathbf{h}^{(j)}$ is defined as

$$\begin{aligned} \mathbf{h}^{(j)} &\triangleq [h_0^{(j)}, h_1^{(j)}, \dots, h_L^{(j)}]^T \\ &\triangleq [h^{(j)}[t_0], h^{(j)}[t_0+T], \dots, h^{(j)}[t_0+LT]]^T, \end{aligned} \quad (5.65)$$

where $h^{(j)}[t_0]$ indicates the channel impulse response with respect to the j th receive antenna at time t_0 . Then, we have

$$\mathbf{H}_T^{(j)} = \begin{bmatrix} h_0^{(j)} & \dots & h_L^{(j)} & 0 & \dots & \dots & 0 \\ 0 & h_0^{(j)} & \dots & h_L^{(j)} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & h_0^{(j)} & \dots & h_L^{(j)} \end{bmatrix} \in \mathbb{C}^{W \times (W+L)}. \quad (5.66)$$

5.4.3.1.1. Signal subspace estimation

The space-time correlation matrix $\mathbf{R}_{yy} \in \mathbb{C}^{M_R \cdot W \times M_R \cdot W}$ of the measurement data is calculated as

$$\mathbf{R}_{yy} = \mathbb{E} \{ \mathbf{y}_n \mathbf{y}_n^H \} = \mathbf{H}_T \mathbf{R}_{ss} \mathbf{H}_T^H + \sigma_n^2 \mathbf{I}_{M_R \cdot W} \quad (5.67)$$

where the matrix $\mathbf{R}_{ss} = \mathbb{E} \{ \mathbf{s}_n \mathbf{s}_n^H \}$ denotes the correlation matrix of the transmit data with dimension $(W+L) \times (W+L)$. The matrix \mathbf{R}_{ss} is assumed to be full-rank, the noise-free part of \mathbf{R}_{yy} (i.e., $\mathbf{H}_T \mathbf{R}_{ss} \mathbf{H}_T^H$) therefore has the rank of $W+L$. With the observation of N consecutive

data windows at the receiver, the estimate of the space-time correlation matrix \mathbf{R}_{yy} is obtained as

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}_n \mathbf{y}_n^H. \quad (5.68)$$

To calculate an SVD of the estimated correlation matrix $\hat{\mathbf{R}}_{yy}$, we have

$$\hat{\mathbf{R}}_{yy} = \begin{bmatrix} \underbrace{\hat{\mathbf{U}}_s}_{M_R \cdot W \times (W+L)} & \underbrace{\hat{\mathbf{U}}_n}_{M_R \cdot W \times (M_R \cdot W - W - L)} \end{bmatrix} \begin{bmatrix} \underbrace{\hat{\Sigma}_s}_{(W+L) \times (W+L)} & \mathbf{0} \\ \mathbf{0} & \underbrace{\hat{\Sigma}_n}_{(M_R \cdot W - W - L) \times (M_R \cdot W - W - L)} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_s & \hat{\mathbf{U}}_n \end{bmatrix}^H \quad (5.69)$$

Here, the columns of $\hat{\mathbf{U}}_s$ span the signal subspace, which is also the linear space spanned by the columns of the filtering matrix \mathbf{H}_T .

5.4.3.1.2. Signal subspace based parameter estimation

Considering the orthogonality between the estimated signal subspace $\hat{\mathbf{U}}_s$ and the estimated noise subspace $\hat{\mathbf{U}}_n$, the unknown SIMO channel coefficients incorporated in the filtering matrix \mathbf{H}_T can be identified up to a scalar factor by solving the problem 1 or problem 2 as defined below. In problem 1, we minimize the following quadratic form by using the estimated noise subspace $\hat{\mathbf{U}}_n$

$$q(\mathbf{H}) \triangleq \sum_{i=1}^{M_R \cdot W - W - L} \|\hat{\mathbf{U}}_n(:, i)^H \mathbf{H}_T\|_2^2, \quad (5.70)$$

where $\hat{\mathbf{U}}_n(:, i)$ represents the i th column of the matrix $\hat{\mathbf{U}}_n$. Alternatively, in problem 2 we maximize the following quadratic form by using the estimated signal subspace $\hat{\mathbf{U}}_s$

$$q(\mathbf{H}) \triangleq \sum_{i=1}^{W+L} \|\hat{\mathbf{U}}_s(:, i)^H \mathbf{H}_T\|_2^2, \quad (5.71)$$

Here, the matrix $\mathbf{H} \in \mathbb{C}^{(L+1) \times M_R}$ is a combined channel matrix for all M_R subchannels denoted as $\mathbf{H} = [\mathbf{h}^{(1)}, \mathbf{h}^{(2)}, \dots, \mathbf{h}^{(M_R)}]$.

Let us solve the problem 2 as an example. In order to specify the quadratic dependence of $q(\mathbf{H})$ on the matrix \mathbf{H} rather than on the associated filtering matrix \mathbf{H}_T found in (5.71), we apply Lemma 5.4.2

Lemma 5.4.2. [MDCM95] If $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots, \mathbf{u}^{(M_R)}$ are M_R arbitrary $W \times 1$ vectors, which have the structure $\mathbf{u}^{(j)} = [u_0^{(j)}, u_1^{(j)}, \dots, u_{W-1}^{(j)}]^T$ for $j = 1, \dots, M_R$. The matrices $\mathbf{G}^{(1)}, \mathbf{G}^{(2)}, \dots, \mathbf{G}^{(M_R)}$

are the corresponding filtering matrices which have the banded Toeplitz structure

$$\mathbf{G}^{(j)} = \begin{bmatrix} u_0^{(j)} & \cdots & u_{W-1}^{(j)} & 0 & \cdots & \cdots & 0 \\ 0 & u_0^{(j)} & \cdots & u_{W-1}^{(j)} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & u_0^{(j)} & \cdots & u_{W-1}^{(j)} \end{bmatrix} \in \mathbb{C}^{(L+1) \times (W+L)}. \quad (5.72)$$

Let us define $\mathbf{u} = [\mathbf{u}^{(1)\top}, \mathbf{u}^{(2)\top}, \dots, \mathbf{u}^{(M_R)\top}]^\top \in \mathbb{C}^{M_R \cdot W \times 1}$ and $\mathbf{G} = [\mathbf{G}^{(1)\top}, \mathbf{G}^{(2)\top}, \dots, \mathbf{G}^{(M_R)\top}]^\top \in \mathbb{C}^{M_R \cdot (L+1) \times (L+W)}$. Then, the following structured relation holds:

$$\mathbf{u}^H \mathbf{H}_T = \text{vec}(\mathbf{H})^H \mathbf{G}, \quad (5.73)$$

where $\text{vec}(\mathbf{H})$ indicates the stacked vector of the channel matrix \mathbf{H} .

Proof: see [MDCM95]

Note that equation (5.73) corresponds to commutativity of the convolution operation. Applying Lemma 5.4.2, the following commutativity is reached

$$\hat{\mathbf{U}}_s(:, i)^H \mathbf{H}_T = \text{vec}(\mathbf{H})^H \mathbf{G}_i, \text{ for } i = 1, \dots, W + L, \quad (5.74)$$

where the vector $\hat{\mathbf{U}}_s(:, i)^H$ and the matrix \mathbf{G}_i correspond to the vector \mathbf{u}^H and the matrix \mathbf{G} in equation (5.73), respectively. The vector $\hat{\mathbf{U}}_s(:, i) \in \mathbb{C}^{M_R \cdot W \times 1}$ denotes the i th column of $\hat{\mathbf{U}}_s$ which is splitted into M_R subvectors $\hat{\mathbf{U}}_s^{(j)}(:, i) \in \mathbb{C}^{W \times 1}$, i.e.,

$$\hat{\mathbf{U}}_s(:, i) = \left[\hat{\mathbf{U}}_s^{(1)\top}(:, i), \hat{\mathbf{U}}_s^{(2)\top}(:, i), \dots, \hat{\mathbf{U}}_s^{(M_R)\top}(:, i) \right]^\top. \quad (5.75)$$

These subvectors $\hat{\mathbf{U}}_s^{(j)}(:, i)$ can be used to construct the associated filtering matrix $\mathbf{G}_i \in \mathbb{C}^{M_R \cdot (L+1) \times (L+W)}$. Thus, the maximization problem in (5.71) can be transformed to the maximization of the following quadratic form

$$\begin{aligned} q(\mathbf{H}) &\triangleq \text{vec}(\mathbf{H})^H \cdot \left(\sum_{i=1}^{W+L} \mathbf{G}_i \mathbf{G}_i^H \right) \cdot \text{vec}(\mathbf{H}) \\ &\triangleq \text{vec}(\mathbf{H})^H \cdot \mathbf{G} \cdot \text{vec}(\mathbf{H}). \end{aligned} \quad (5.76)$$

The solution of this maximization problem is the eigenvector associated to the largest eigenvalue of the matrix \mathbf{G} . To ensure the channel identifiability, there are several conditions which have to hold [MDCM95]:

1. The correlation matrix \mathbf{R}_{ss} is full-rank but otherwise unknown, which requires $N \geq (W + L)$.
2. The matrix \mathbf{H}_T has a full column rank.
3. The observed data window length is greater than the channel order L (i.e., $W > L$).
4. The number of channel taps $L + 1$ has been correctly estimated before.
5. The noise samples are uncorrelated with the input data.

5.4.3.2. Tensor-based subspace method for SIMO channels

Instead of the stacking operation employed in the definition of \mathbf{y}_n in equation (5.62), we introduce a 3-way tensor $\mathcal{Y} \in \mathbb{C}^{M_R \times W \times N}$ to model the measurement data. The three dimensions of the tensor \mathcal{Y} represent the number of receive antennas, the observed data window length, and the number of data windows, respectively. The corresponding input output data model can be expressed as

$$\mathcal{Y} = \mathcal{H} \times_3 \mathbf{S}^T + \mathcal{N}. \quad (5.77)$$

The matrix $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N]$ has the dimension $(W + L) \times N$ and contains input data sequences associated to N sequentially observed data windows at the receiver. Each input data sequence $\mathbf{s}_n = [s[nW], s[nW - 1], \dots, s[nW - W - L + 1]]^T$ has the dimension $(W + L) \times 1$ for $n = 1, 2, \dots, N$. The filtering tensor $\mathcal{H} \in \mathbb{C}^{M_R \times W \times (W + L)}$ is constructed by aligning the banded Toeplitz matrices of the M_R subchannels along the first dimension as shown in Figure 5.20. The tensor \mathcal{N} contains noise samples and has the same size as the tensor \mathcal{Y} .

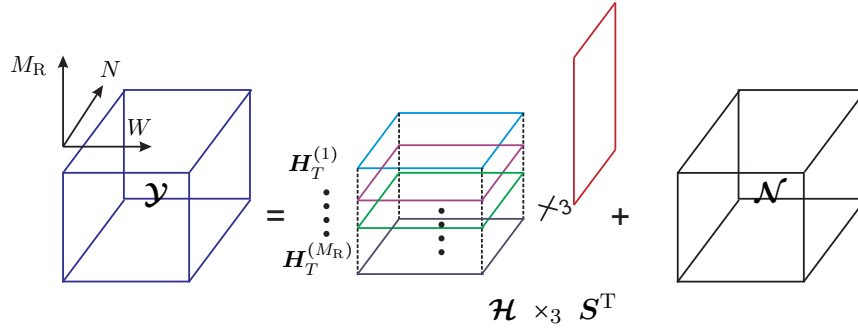


Figure 5.20.: Block diagram of the tensor based data model in equation (5.77)

5.4.3.2.1. Signal subspace estimation

Instead of computing the SVD of the estimated space-time correlation matrix $\hat{\mathbf{R}}_{yy}$ found in (5.96), we employ a truncated HOSVD of the measurement tensor \mathcal{Y} to obtain an enhanced estimate of the signal subspace. We have

$$\mathcal{Y} \approx \mathcal{S}^{[s]} \times_1 \mathbf{U}_1^{[s]} \times_2 \mathbf{U}_2^{[s]} \times_3 \mathbf{U}_3^{[s]}, \quad (5.78)$$

where $\mathcal{S}^{[s]} \in \mathbb{C}^{r_1 \times r_2 \times r_3}$, $\mathbf{U}_1^{[s]} \in \mathbb{C}^{M_R \times r_1}$, $\mathbf{U}_2^{[s]} \in \mathbb{C}^{W \times r_2}$, and $\mathbf{U}_3^{[s]} \in \mathbb{C}^{N \times r_3}$. Here, r_n ($n = 1, 2, 3$) denotes the n -rank of the noiseless tensor $\tilde{\mathcal{Y}}$ (i.e., $\tilde{\mathcal{Y}} = \mathcal{H} \times_3 \mathbf{S}^T$). In this application, we have $r_1 = \min(M_R, L + 1)$, $r_2 = \min(W, N \cdot M_R)$, and $r_3 = \min(N, W + L)$. Due to the assumption of $N \geq W + L$, we can conclude that r_2 is the same as the observed data window length W and r_3 is equal to $W + L$.

From equation (5.78), we define the estimated signal subspace tensor $\hat{\mathcal{U}}^{[s]} \in \mathbb{C}^{M_R \times W \times r_3}$ as

$$\hat{\mathcal{U}}^{[s]} = \mathcal{S}^{[s]} \times_1 \mathbf{U}_1^{[s]} \times_2 \mathbf{U}_2^{[s]}. \quad (5.79)$$

The columns of $\left[\hat{\mathcal{U}}^{[s]} \right]_{(3)}^T \in \mathbb{C}^{M_R \cdot W \times r_3}$ span the estimated signal subspace. According to Theorem 5.4.1, $\left[\hat{\mathcal{U}}^{[s]} \right]_{(3)}^T$ provides a more accurate estimate than $\hat{\mathbf{U}}_s$ from the matrix case under the conditions that the measurement tensor $\tilde{\mathcal{Y}}$ is rank-deficient in the first or second mode (i.e., $M_R > r_1$ or $W > r_2$). Otherwise, both the tensor-based and the matrix-based signal subspace estimation yield exactly the same accuracy. Since r_2 is equal to W for our model, $\left[\hat{\mathcal{U}}^{[s]} \right]_{(3)}^T$ can achieve a better estimate under the condition $M_R > L + 1$.

Computational complexity: We compare the computational complexity of the truncated SVD and the truncated HOSVD in terms of the number of required multiplications for the computation of the signal subspace. There is a large variety of methods to compute the SVD with different complexities. [GL96] shows an efficient solution employing the method of orthogonal iterations which has a complexity in terms of the required number of multiplications of $k_t \cdot M \cdot N \cdot r$ for an $M \times N$ matrix truncated to rank r , where k_t is a constant that depends on the design of the algorithm. In the matrix case, a single SVD of the estimated correlation matrix $\hat{\mathbf{R}}_{yy}$ truncated to rank $W + L$ is computed to obtain $\hat{\mathbf{U}}_s$. In the tensor case, the truncated HOSVD of the measurement tensor \mathcal{Y} is computed to obtain the estimated signal subspace $\left[\hat{\mathcal{U}}^{[s]} \right]_{(3)}^T$, which is equivalent to truncated SVDs of all its unfolding. Moreover, additional multiplications are required to compute the core tensor $\mathcal{S}^{[s]}$ and the signal subspace tensor

Table 5.6.: Comparison of the required number of multiplication for signal subspace estimation in case of the matrix-based method and the tensor-based method, $r_2 = W$ and $r_3 = W + L$.

	Matrix-based method	Tensor-based method
$r_1 < M_R$:	$k_t \cdot M_R^2 \cdot W \cdot r_2 \cdot r_3$	$(M_R + W) \cdot r_1 \cdot r_2 \cdot r_3 + (k_t + 1) \cdot M_R \cdot W \cdot N \cdot (r_1 + r_2 + r_3)$
$r_1 = M_R$:	$k_t \cdot M_R \cdot W \cdot r_1 \cdot r_2 \cdot r_3$	$(M_R + W) \cdot r_1 \cdot r_2 \cdot r_3 + (k_t + 1) \cdot M_R \cdot W \cdot N \cdot (r_1 + r_2 + r_3)$

$\hat{\mathbf{U}}^{[s]}$. The total number of required multiplications is compared in Table 5.6. It indicates that the computational complexity of the tensor method is higher than the matrix method but of the same order. However, the performance improvement demonstrated in Section 5.4.3.2.4 justifies this increase of the computational complexity.

5.4.3.2.2. Signal subspace based parameter estimation

Since the column spaces of $\left[\hat{\mathbf{U}}^{[s]}\right]_{(3)}^T$ and $[\mathbf{H}]_{(3)}^T$ approximately coincide, the unknown SIMO channel coefficients incorporated in the filtering tensor \mathbf{H} can be identified up to a scalar factor by solving the maximization of the following quadratic form

$$q(\mathbf{H}) \triangleq \sum_{i=1}^{W+L} \left\| \hat{\mathbf{U}}_{s_T}(:, i)^H [\mathbf{H}]_{(3)}^T \right\|_2^2. \quad (5.80)$$

Here, we use $\hat{\mathbf{U}}_{s_T}$ to indicate the estimated signal subspace of the tensor case for notational simplicity (i.e., $\hat{\mathbf{U}}_{s_T} = \left[\hat{\mathbf{U}}^{[s]}\right]_{(3)}^T$). The channel parameter estimation scheme obeys the exact same procedure as the scheme mentioned in Section 5.4.3.1.2, except for replacing $\hat{\mathbf{U}}_s(:, i)$ and \mathbf{H}_T by $\hat{\mathbf{U}}_{s_T}(:, i)$ and $[\mathbf{H}]_{(3)}^T$, respectively.

The channel identification of the tensor-based subspace method requires the following necessary conditions

1. The matrix \mathbf{S} of the transmit signal has the rank $W + L$, which requires $N \geq (W + L)$.
2. the 3-mode unfolding of the filtering tensor \mathbf{H} has a full row rank $W + L$.
3. The observed data window length is greater than the channel order L (i.e., $W > L$).
4. The number of channel taps $L + 1$ has been correctly estimated before.
5. The noise samples are uncorrelated with the input data.

5.4.3.2.3. Oversampled antenna array

As mentioned above, the performance benefit of the tensor-based subspace method is achieved under the condition $M_R > L + 1$. To maintain the performance benefit of the tensor model for the case $M_R \leq L + 1$, we introduce an oversampling of the received signals by a factor $P = T/\Delta$. Then, for the received signal of the j th receive antenna $\mathbf{y}_n^{(j)}$ in the n th observed data window, a set of P sequences are constructed according to

$$\mathbf{y}_n^{(j,m)} = \begin{bmatrix} y_n^{(j)} \left[k + \frac{m}{P} \right] \\ y_n^{(j)} \left[k + 1 + \frac{m}{P} \right] \\ \vdots \\ y_n^{(j)} \left[k + W - 1 + \frac{m}{P} \right] \end{bmatrix} \in \mathbb{C}^{W \times 1}, \quad (5.81)$$

where $m = 0, 1, \dots, P - 1$. Notice that the vector $\mathbf{y}_n^{(j,0)}$ corresponds to the vector $\mathbf{y}_n^{(j)}$. Each sequence $\mathbf{y}_n^{(j,m)}$ depends on the oversampled discrete-time impulse responses $\mathbf{h}^{(j,m)}$. We have

$$\begin{aligned} \mathbf{h}^{(j,m)} &\triangleq [h_0^{(j,m)}, h_1^{(j,m)}, \dots, h_L^{(j,m)}] \\ &\triangleq [h^{(j)}[t_0 + m\Delta], h^{(j)}[t_0 + m\Delta + T], \dots, h^{(j)}[t_0 + m\Delta + LT]], \end{aligned} \quad (5.82)$$

where the vector $\mathbf{h}^{(j,0)}$ is equal to the vector $\mathbf{h}^{(j)}$ in equation (5.65). The filtering matrix in this case is constructed as

$$\mathbf{H}_{TP} = \begin{bmatrix} \mathbf{H}_T^{(1,0)} \\ \mathbf{H}_T^{(1,1)} \\ \vdots \\ \mathbf{H}_T^{(1,P-1)} \\ \mathbf{H}_T^{(2,0)} \\ \vdots \\ \mathbf{H}_T^{(M_R, P-1)} \end{bmatrix} \in \mathbb{C}^{P \cdot M_R \cdot W \times (W+L)}, \quad (5.83)$$

where m changes from 0 to $P - 1$ for each j and the matrix $\mathbf{H}_T^{(j,m)}$ is a banded Toeplitz matrix associated to the discrete-time impulse responses $\mathbf{h}^{(j,m)}$ and constructed as

$$\mathbf{H}_T^{(j,m)} = \begin{bmatrix} h_0^{(j,m)} & \dots & h_L^{(j,m)} & 0 & \dots & \dots & 0 \\ 0 & h_0^{(j,m)} & \dots & h_L^{(j,m)} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & h_0^{(j,m)} & \dots & h_L^{(j,m)} \end{bmatrix} \in \mathbb{C}^{W \times (W+L)}. \quad (5.84)$$

We still can use a 3-way tensor $\mathcal{Y}_P \in \mathbb{C}^{M_R \cdot P \times W \times N}$ to model the oversampled received signals. Compared to the tensor \mathcal{Y} , only the first dimension changes due to the oversampling. We describe the corresponding input output data model as

$$\mathcal{Y}_P = \mathcal{H}_P \times_3 \mathbf{S}^T + \mathcal{N}_P. \quad (5.85)$$

The filtering tensor \mathcal{H}_P has the dimension $M_R \cdot P \times W \times (W + L)$ and is organized by stacking the slices of the block matrices defined in equation (5.83) along the first dimension of the tensor \mathcal{H}_P as shown in Figure 5.21. The noise tensor \mathcal{N}_P has the same size as the tensor \mathcal{Y}_P . Notice that the noise samples are not necessarily temporally uncorrelated due to the oversampling.

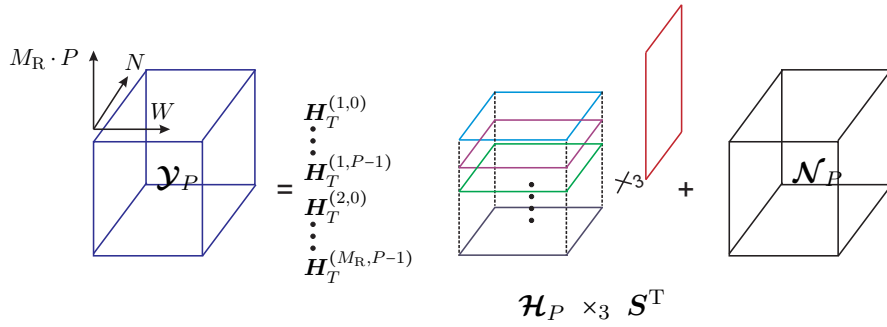


Figure 5.21.: Block diagram of the tensor based data model with oversampling.

By computing the truncated HOSVD of the measurement tensor \mathcal{Y}_P , the signal subspace tensor $\hat{\mathcal{U}}_P^{[s]} \in \mathbb{C}^{M_R \cdot P \times W \times r_{3P}}$ can be estimated as

$$\hat{\mathcal{U}}_P^{[s]} = \mathcal{S}_P^{[s]} \times_1 \mathbf{U}_{1_P}^{[s]} \times_2 \mathbf{U}_{2_P}^{[s]}, \quad (5.86)$$

where $\mathcal{S}_P^{[s]} \in \mathbb{C}^{r_{1P} \times r_{2P} \times r_{3P}}$, $\mathbf{U}_{1_P}^{[s]} \in \mathbb{C}^{M_R \cdot P \times r_{1P}}$ and $\mathbf{U}_{2_P}^{[s]} \in \mathbb{C}^{W \times r_{2P}}$. The terms r_{n_P} ($n = 1, 2, 3$) denote the n -rank of the noiseless tensor $\tilde{\mathcal{Y}}_P$ (i.e., $\tilde{\mathcal{Y}}_P = \mathcal{H}_P \times_3 \mathbf{S}^T$). It is found that $r_{1_P} = \min(M_R \cdot P, L + 1)$. The parameter r_{2_P} is equal to W and r_{3_P} is the same as $W + L$. In this case, the condition for achieving a more accurate signal subspace estimate in the tensor case as compared to the matrix case is loosened to $M_R \cdot P > L + 1$.

5.4.3.2.4. Simulation results

By comparing the matrix-based subspace method and the tensor-based subspace method, we demonstrate the performance improvement of the tensor-based subspace method for the blind estimation of SIMO channels. The comparisons are shown in terms of the root mean square

error (RMSE) of the estimated normalized channels. This RMSE is defined as

$$\text{RMSE} = \frac{1}{P} \sqrt{\mathbb{E} \left\{ \left\| \hat{\mathbf{H}}a - \mathbf{H} \right\|_F^2 \right\}}. \quad (5.87)$$

Here, $a = \frac{\text{vec}(\hat{\mathbf{H}})^H \text{vec}(\mathbf{H})}{|\text{vec}(\hat{\mathbf{H}})^H \text{vec}(\hat{\mathbf{H}})|}$ is a scalar factor due to the fact that the unknown SIMO channels are only estimated up to a multiplication by a scalar. The channel matrix \mathbf{H} is normalized to unit Frobenius norm. The RMSE is averaged over 500 channel realizations.

Monte Carlo simulations have been conducted where the successive symbols are generated statistically independent and emitted in 4-QAM format. The signal to noise ratio (SNR) is defined as $10 \log_{10} \frac{\mathbb{E}\{\|\mathbf{s}(k)\|_2^2\}}{\mathbb{E}\{\|\mathbf{n}(k)\|_2^2\}}$. To simulate a multipath environment, we adopt a commonly used model [TXK91] to construct an $(L + 1)$ -ray multipath continuous-time channel $h^{(j)}(t)$ between the j th receive antenna and the transmit antenna using a root raised cosine (RRC) pulse shaping filter $g_c(t - \gamma_\ell, \beta)$. We have

$$h^{(j)}(t) = \sum_{\ell=0}^L \alpha_\ell^{(j)} g_c(t - \gamma_\ell^{(j)}, \beta), \quad (5.88)$$

where the roll-off factor β is set to 0.5 in the simulations and $\alpha_\ell^{(j)}$ are complex valued Gaussian random variables. The magnitudes of $\alpha_\ell^{(j)}$ (i.e., $\text{abs}(\alpha_\ell^{(j)})$) have different mean values corresponding to different ℓ . We set $\text{abs}(\alpha_0^{(j)})$ to be real valued Gaussian random variable with mean value 1 and standard deviation 0.5. Furthermore, we decrease the mean values of $\text{abs}(\alpha_\ell^{(j)})$ with a step of size 0.2 (i.e., $\text{abs}(\alpha_1^{(j)})$ with the mean value of 0.8, $\text{abs}(\alpha_2^{(j)})$ with the mean value of 0.6, etc.) Note that the maximum number of L is smaller than 5 in our simulations. The term $\gamma_\ell^{(j)}$ indicates the delay of the ℓ th path, which is set to $\gamma_0^{(j)} = 0$ and $\gamma_\ell^{(j)} = \ell \cdot T + \vartheta$. Here, ϑ is randomly chosen from $0.25 T$, $0.5 T$, and $0.75 T$. The discrete-time channel is obtained by sampling $h^{(j)}(t)$ at a rate of T/P . We observe the measurement data with a smoothing window. The length of the observed data window is set to $W = 10$. The noise samples are correlated in case of introducing oversampling. The correlated noise samples can be obtained from the output of the RRC pulse shaping filter when additive white Gaussian noise is introduced as the input.

* **Smoothing Window with $\eta = W$**

The smoothing parameter is denoted by η which indicates the number of the new measurements in the next observed data window. First, we consider the case $\eta = W$, where the adjacent observed data windows do not overlap with each other as shown in Figure

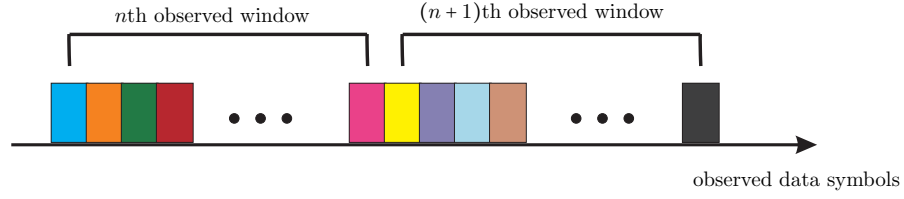


Figure 5.22.: The smoothing window with the smoothing parameter $\eta = W$.

5.22. Under this condition, Figures 5.23 and 5.24 show the comparison between the proposed tensor method and the matrix-based subspace method. Under the condition $M_R > L + 1$, an enhanced estimate has been achieved by the proposed tensor-based subspace method, especially for a small number of observed data symbols. Much larger differences between M_R and $L + 1$ lead to a more significant improvement.

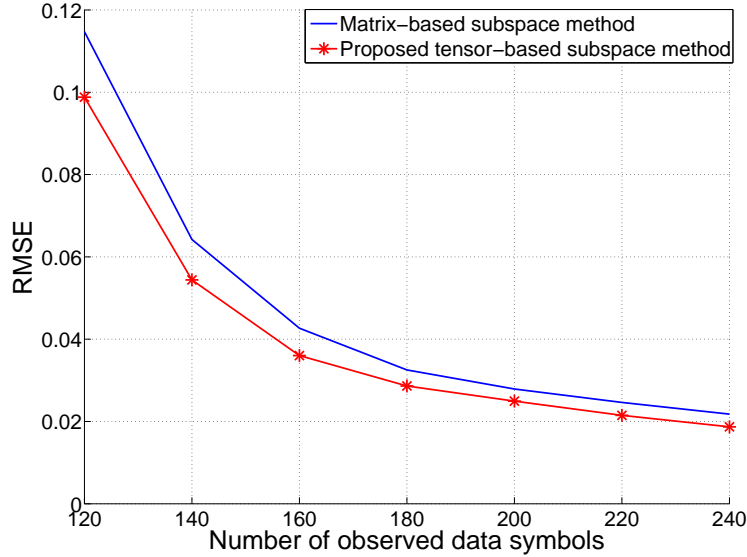


Figure 5.23.: RMSE for blind estimation of SIMO channels with $M_R = 4$ and $r_1 = L + 1 = 3$ at SNR = 20 dB

Figure 5.25 shows the case $M_R < L + 1$, where both methods achieve the same performance. In order to maintain the benefit of the tensor method, we use oversampling at the receiver. For a fair comparison, the oversampling is utilized for both the matrix-based method and the proposed tensor-based method. There is a significant performance improvement achieved by the tensor-based method, since the condition for achieving an

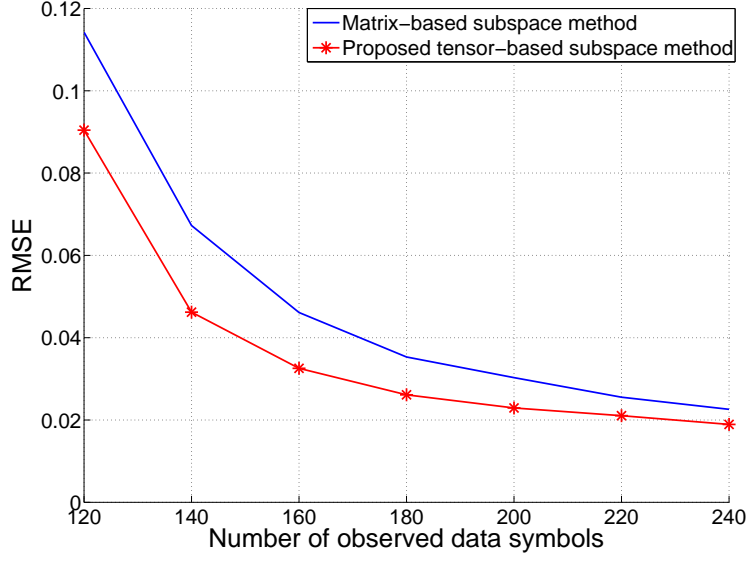


Figure 5.24.: RMSE for blind estimation of SIMO channels with $M_R = 6$ and $r_1 = L + 1 = 3$ at SNR = 20 dB

improved estimate is loosened to $M_R \cdot P > L + 1$.

* **Smoothing Window with $1 \leq \eta < W$**

Next, we investigate the overlap of the observed data windows as shown in Figure 5.26. We evaluate the RMSE performance of the matrix-based subspace method and the proposed tensor-based method with different smoothing parameters η . Due to the assumption $N \geq W + L$, the cases with various η have different requirements for the minimum number of observed data symbols. It is observed in Figure 5.27 that the accuracy of the estimate improves with the decrease of the parameter η . But notice that, with the same number of observed data symbols, the smaller parameter η leads to a larger number of observed data windows which results in an increased computation time. The proposed tensor-based method always outperforms the matrix-based subspace method for any η .

Moreover, we show that the performance improvement of the tensor method in terms of RMSE leads to a better performance in term of bit error rate (BER). Since the blind estimation method of the SIMO channel identifies the unknown SIMO channel coefficients up to a scalar a , a few pilots are therefore required to further identify a in order to decode the received signals. Then, this blind channel estimation scheme changes to a few pilots aided

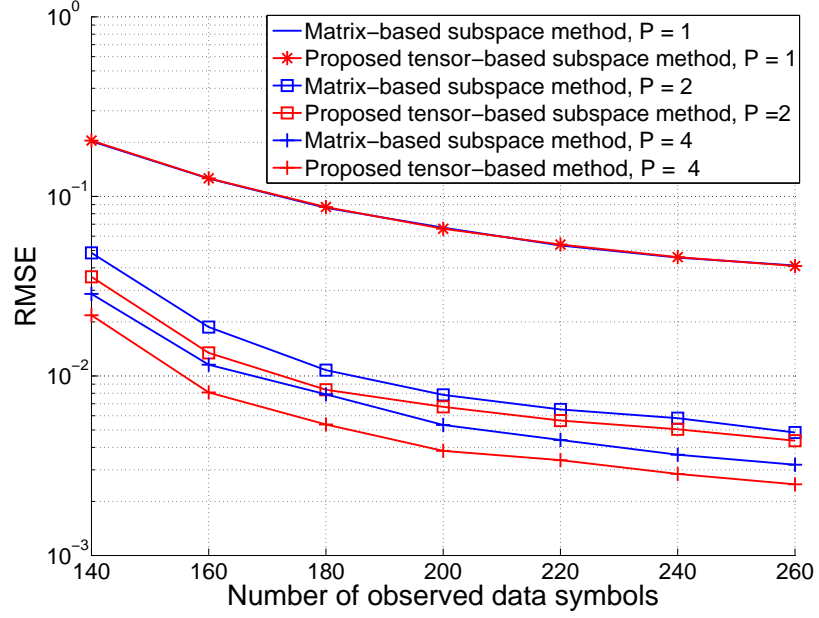


Figure 5.25.: RMSE for blind estimation of SIMO channels with $M_R = 4$, $L + 1 = 5$ at SNR = 20 dB.

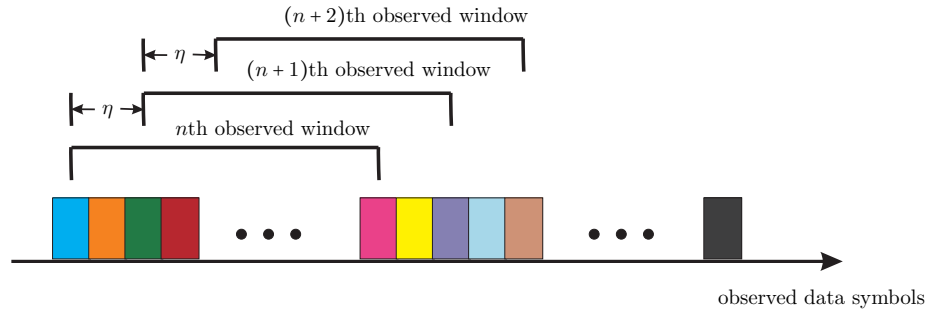


Figure 5.26.: The smoothing window with the smoothing parameter $1 \leq \eta < W$.

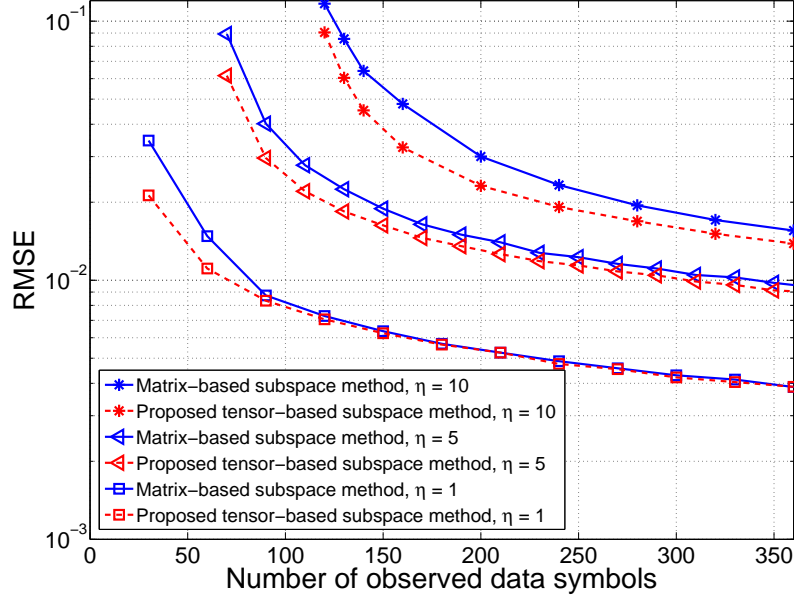


Figure 5.27.: RMSE for blind estimation of SIMO channels with varied smoothing parameter η . $M_R = 6$, $r_1 = L + 1 = 3$, $P = 1$, and SNR = 20 dB.

blind estimation (namely Semi-Blind Channel Estimation (SBCE)). In our simulations, we first use training-based channel estimation (e.g., the MMSE estimator [WZ07]) to obtain an initial estimate of the channel matrix \mathbf{H} and indicate this estimate as $\widetilde{\mathbf{H}}$. Then, using the channel estimate from the blind scheme (i.e., $\hat{\mathbf{H}}$), we can identify a by solving the minimization problem $\min_a \|\text{vec}(\hat{\mathbf{H}}) \cdot \alpha - \text{vec}(\widetilde{\mathbf{H}})\|_2^2$. The solution is $a = \frac{\text{vec}(\hat{\mathbf{H}})^H \text{vec}(\widetilde{\mathbf{H}})}{|\text{vec}(\hat{\mathbf{H}})^H \text{vec}(\hat{\mathbf{H}})|}$.

We compare the BER performances of the semi-blind channel estimation based on the matrix case and the tensor case. Furthermore, the BER performance of training-based non-blind channel estimation is considered for comparison.

We consider $M_R = 5$ and $L + 1 = 3$ which promises an enhanced performance of the tensor method in terms of the RMSE. The number of observed data symbols is 14. Figure 5.28 shows that this performance improvement still can be obtained by considering the BER performance.

The performance comparison between the semi-blind channel estimation and the training-based channel estimation with the same pilot overhead are shown in Figure 5.29. With the same pilot overhead, the semi-blind channel estimation reaches a better BER performance. Here, the training-based channel estimation employs the MMSE estimator [WZ07].

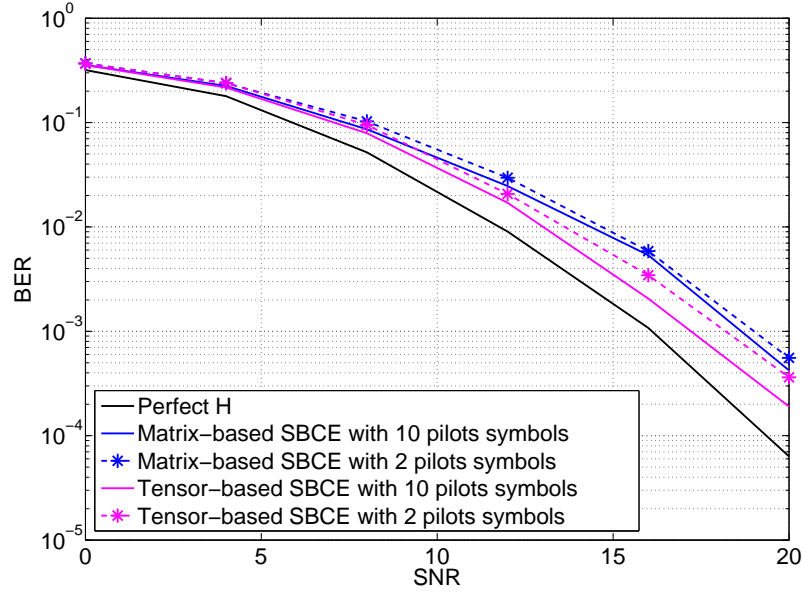


Figure 5.28.: BER for semi-blind estimation of SIMO channels with $M_R = 5$, $r_1 = L + 1 = 3$.

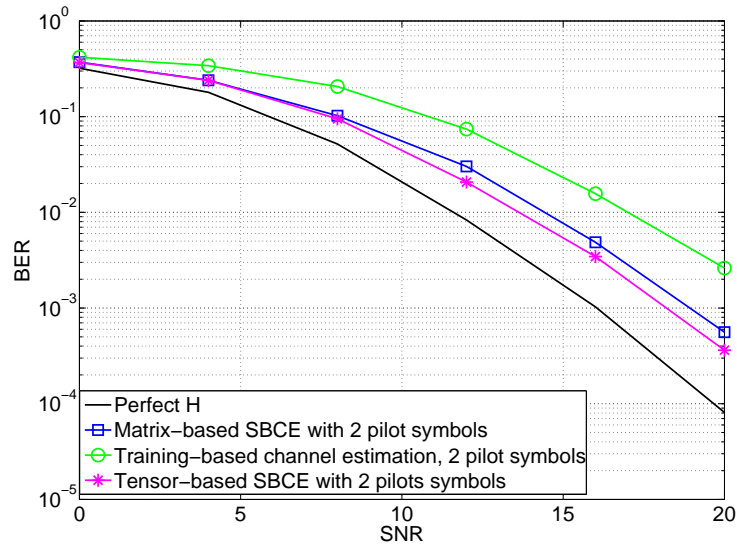


Figure 5.29.: BER performance comparison for semi-blind estimation and training-based channel estimation of SIMO channels with $M_R = 5$, $r_1 = L + 1 = 3$.

5.4.4. Tensor-based blind estimation of MIMO channels

The blind estimation scheme of SIMO channels discussed previously can be naturally extended to the MIMO channel case [SRH13b]. We consider a MIMO system where the transmitter has M_T transmit antennas and receiver is equipped with M_R receive antennas. The channel between each transmit and receive antenna pair is modeled as an FIR filter with a maximum of $L + 1$ taps. Let $\mathbf{s}[k] = [s_1[k], s_2[k], \dots, s_{M_T}[k]]^T$ denote the symbol vector emitted over M_T transmit antennas at time kT . Here T is the symbol duration. This discrete-time signal experiences an unknown communication channel which is assumed to be time-invariant during the observation interval. Then, the received signal at time kT is formulated as

$$\mathbf{y}[k] = \sum_{\ell=0}^L \mathbf{H}_\ell \mathbf{s}[k-\ell] + \mathbf{n}[k] \in \mathbb{C}^{M_R}, \quad (5.89)$$

where $\mathbf{H}_\ell \in \mathbb{C}^{M_R \times M_T}$ contains the coefficients of the channel impulse responses corresponding to lag ℓ . We assume that all subchannels have the same length $L + 1$ for simplicity. The combined channel matrix for all $L + 1$ subchannels is denoted as $\mathbf{H} = [\mathbf{H}_0^T, \dots, \mathbf{H}_L^T]^T \in \mathbb{C}^{M_R \cdot (L+1) \times M_T}$. The elements of the noise vector $\mathbf{n}[k]$ are circularly symmetric complex Gaussian distributed with variance σ_n^2 and assumed mutually uncorrelated in space and time.

5.4.4.1. Matrix-based data model and signal subspace estimation

The measurement data is observed by consecutive data windows over all receive antennas. Each window has the length W . The dimensions of M_R receive antennas and the data window length W are stacked into one highly structured vector. We denote the measurement data with respect to the n th observed data window by $\mathbf{y}_n \in \mathbb{C}^{M_R \cdot W \times 1}$, which is given by

$$\mathbf{y}_n = \mathbf{H}_T \mathbf{s}_n + \mathbf{n}. \quad (5.90)$$

The term \mathbf{s}_n is the stacked vector of the input data sequences and has the structure

$$\mathbf{s}_n = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_{M_T} \end{bmatrix} \in \mathbb{C}^{M_T \cdot (W+L) \times 1}, \quad (5.91)$$

where $\mathbf{s}_i = [s_i[nW], s_i[nW-1], \dots, s_i[nW-W-L+1]]^T$ is $(W+L) \times 1$ dimensional input data sequence on i th transmit antenna for $i = 1, \dots, M_T$. The matrix $\mathbf{H}_T \in \mathbb{C}^{M_R \cdot W \times M_T \cdot (W+L)}$

is the filtering matrix and structured as

$$\mathbf{H}_T = \begin{bmatrix} \mathbf{H}_T^{(1,1)} & \mathbf{H}_T^{(1,2)} & \dots & \mathbf{H}_T^{(1,M_T)} \\ \mathbf{H}_T^{(2,1)} & \mathbf{H}_T^{(2,2)} & \dots & \mathbf{H}_T^{(2,M_T)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_T^{(M_R,1)} & \mathbf{H}_T^{(M_R,2)} & \dots & \mathbf{H}_T^{(M_R,M_T)} \end{bmatrix}. \quad (5.92)$$

Here, the matrix $\mathbf{H}_T^{(j,i)} \in \mathbb{C}^{W \times (W+L)}$ denotes a banded Toeplitz matrix associated to the channel impulse response $\mathbf{h}^{(j,i)}$ between the j th receive antenna and i th transmit antenna. The vector $\mathbf{h}^{(j,i)}$ is defined as

$$\begin{aligned} \mathbf{h}^{(j,i)} &\triangleq [h_0^{(j,i)}, h_1^{(j,i)}, \dots, h_L^{(j,i)}]^T \\ &\triangleq [h^{(j,i)}[t_0], h^{(j,i)}[t_0 + T], \dots, h^{(j,i)}[t_0 + LT]]^T. \end{aligned} \quad (5.93)$$

Then, we have

$$\mathbf{H}_T^{(j,i)} = \begin{bmatrix} h_0^{(j,i)} & \dots & h_L^{(j,i)} & 0 & \dots & \dots & 0 \\ 0 & h_0^{(j,i)} & \dots & h_L^{(j,i)} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & h_0^{(j,i)} & \dots & h_L^{(j,i)} \end{bmatrix}. \quad (5.94)$$

The space-time correlation matrix $\mathbf{R}_{yy} \in \mathbb{C}^{M_R \cdot W \times M_R \cdot W}$ of the measurement data is calculated as

$$\mathbf{R}_{yy} = \mathbb{E} \{ \mathbf{y}_n \mathbf{y}_n^H \} = \mathbf{H}_T \mathbf{R}_{ss} \mathbf{H}_T^H + \sigma^2 \mathbf{I}_{M_R \cdot W} \quad (5.95)$$

where the matrix $\mathbf{R}_{ss} = \mathbb{E} \{ \mathbf{s}_n \mathbf{s}_n^H \}$ indicates the correlation matrix of the input data with dimension $M_T \cdot (W + L) \times M_T \cdot (W + L)$. The matrix \mathbf{R}_{ss} is assumed to be full-rank, the noise-free part of \mathbf{R}_{yy} (i.e., $\mathbf{H}_T \mathbf{R}_{ss} \mathbf{H}_T^H$) therefore has the rank of $r = M_T \cdot (W + L)$. With the observation of N consecutive data windows at the receiver, the estimate of the space-time correlation matrix \mathbf{R}_{yy} is obtained as

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}_n \mathbf{y}_n^H. \quad (5.96)$$

To calculate an SVD of the estimated correlation matrix $\hat{\mathbf{R}}_{yy}$, we have

$$\hat{\mathbf{R}}_{yy} = \begin{bmatrix} \hat{\mathbf{U}}_s & \hat{\mathbf{U}}_n \\ \underbrace{M_R \cdot W \times r}_{M_R \cdot W \times r} & \underbrace{M_R \cdot W \times (M_R \cdot W - r)}_{M_R \cdot W \times (M_R \cdot W - r)} \end{bmatrix} \cdot \begin{bmatrix} \hat{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \hat{\Sigma}_n \\ \underbrace{r \times r}_{r \times r} & \underbrace{(M_R \cdot W - r) \times (M_R \cdot W - r)}_{(M_R \cdot W - r) \times (M_R \cdot W - r)} \end{bmatrix} \cdot [\hat{\mathbf{U}}_s \ \hat{\mathbf{U}}_n]^H \quad (5.97)$$

The signal subspace estimate is spanned by the first $r = M_T \cdot (W + L)$ dominant left singular vectors.

5.4.4.2. Tensor-based data model and signal subspace estimation

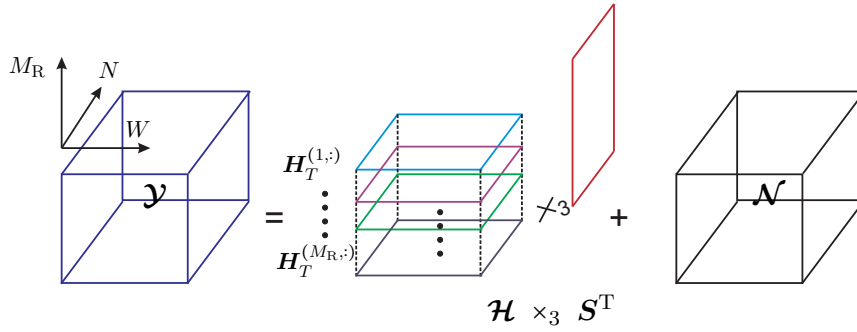


Figure 5.30.: Block diagram of the tensor based data model.

The dimension stacking operation employed in the definition of \mathbf{y}_n in equation (5.90) does not account for the structure inherent in the measurement data. Therefore, we introduce a 3-way tensor to model the measurement data. The three dimensions of the tensor $\mathcal{Y} \in \mathbb{C}^{M_R \times W \times N}$ represent receive antennas, observed data window length, and the number of data windows, respectively. The corresponding input output data model can be expressed as

$$\mathcal{Y} = \mathcal{H} \times_3 \mathbf{S}^T + \mathcal{N}. \quad (5.98)$$

The matrix $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N] \in \mathbb{C}^{M_T \cdot (W+L) \times N}$ contains the input data sequences corresponding to the N sequentially observed data windows at the receiver. Each column of \mathbf{S} has the same definition as \mathbf{s}_n in equation (5.91). The filtering tensor $\mathcal{H} \in \mathbb{C}^{M_R \times W \times M_T \cdot (W+L)}$ is constructed by aligning the slices of the block matrices \mathbf{H}_T in (5.92) along the first dimension. The slices of \mathcal{H} are defined as $\mathbf{H}_T^{(j,:)}$ for $j = 1, 2, \dots, M_R$ and depicted in Figure 5.31. We compute the 3-mode product of the tensor \mathcal{H} with the matrix \mathbf{S}^T as shown in Figure 5.30. The tensor \mathcal{N} contains noise samples and has the same size as the tensor \mathcal{Y} .

$$\mathbf{H}_T = \begin{bmatrix} \boxed{\mathbf{H}_T^{(1,1)} \quad \dots \quad \mathbf{H}_T^{(1,2)} \quad \mathbf{H}_T^{(1,M_T)}} & \Rightarrow \mathbf{H}_T^{(1,:)} \\ \boxed{\mathbf{H}_T^{(2,1)} \quad \mathbf{H}_T^{(2,2)} \quad \dots \quad \mathbf{H}_T^{(2,M_T)}} & \Rightarrow \mathbf{H}_T^{(2,:)} \\ \vdots & \\ \boxed{\mathbf{H}_T^{(M_R,1)} \quad \mathbf{H}_T^{(M_R,2)} \quad \dots \quad \mathbf{H}_T^{(M_R,M_T)}} & \Rightarrow \mathbf{H}_T^{(M_R,:)} \end{bmatrix}$$

 Figure 5.31.: Construction of the slices of \mathbf{H}_T .

By computing the truncated HOSVD of the measurement tensor \mathbf{Y} , we have

$$\mathbf{Y} \approx \mathbf{S}^{[s]} \times_1 \mathbf{U}_1^{[s]} \times_2 \mathbf{U}_2^{[s]} \times_3 \mathbf{U}_3^{[s]}, \quad (5.99)$$

where $\mathbf{S}^{[s]} \in \mathbb{C}^{r_1 \times r_2 \times r_3}$, $\mathbf{U}_1^{[s]} \in \mathbb{C}^{M_R \times r_1}$, $\mathbf{U}_2^{[s]} \in \mathbb{C}^{W \times r_2}$, and $\mathbf{U}_3^{[s]} \in \mathbb{C}^{N \times r_3}$. Here, r_n ($n = 1, 2, 3$) denotes the n -rank of the noiseless tensor $\tilde{\mathbf{Y}}$ (i.e., $\tilde{\mathbf{Y}} = \mathbf{H} \times_3 \mathbf{S}^T$). In MIMO application, we have $r_1 = \min(M_R, M_T \cdot (L + 1))$, $r_2 = \min(W, N \cdot M_R)$, and $r_3 = \min(N, M_T \cdot (W + L))$. According to the assumption of $N \geq M_T \cdot (W + L)$, the r_2 and r_3 can be simplified to $r_2 = W$ and $r_3 = M_T \cdot (W + L)$, respectively.

From equation (5.78), the estimated signal subspace tensor $\hat{\mathbf{U}}^{[s]} \in \mathbb{C}^{M_R \times W \times r_3}$ is defined as

$$\hat{\mathbf{U}}^{[s]} = \mathbf{S}^{[s]} \times_1 \mathbf{U}_1^{[s]} \times_2 \mathbf{U}_2^{[s]}. \quad (5.100)$$

Then, the estimated signal subspace is spanned by the columns of $\left[\hat{\mathbf{U}}^{[s]} \right]_{(3)}^T \in \mathbb{C}^{M_R \cdot W \times r_3}$. By exploiting the inherent structure in the subspace estimation step, $\left[\hat{\mathbf{U}}^{[s]} \right]_{(3)}^T$ can provide a more accurate estimate than $\hat{\mathbf{U}}_s$ from the matrix-based method under the conditions that the measurement tensor $\tilde{\mathbf{Y}}$ is rank-deficient in the first or second mode (i.e., $M_R > r_1$ or $W > r_2$). Otherwise, both the tensor-based and matrix-based signal subspace estimation yield exactly the same estimate. Since r_2 is always equal to W for our model, a benefit of the tensor-based signal subspace estimation is achieved under the condition $M_R > r_1 = M_T \cdot (L + 1)$.

Comparing the computational complexities of the matrix-based and tensor-based signal subspace estimations, it is found to obey the same expressions shown in Table 5.6 but with

$r_1 = \min(M_R, M_T \cdot (L+1))$, $r_2 = W$, and $r_3 = M_T \cdot (W+L)$. We can conclude that the complexity of the tensor method is higher than the matrix method but of the same order.

5.4.4.3. Oversampled antenna array

Similar to the SIMO channel case, we can introduce an oversampling of the receive signals with a factor $P = T/\Delta$ to maintain the performance benefit for the case $M_R \leq M_T \cdot (L+1)$. With oversampling a set of P sequences are constructed from the received signal of the j th receive antenna in the n th observed data window $\mathbf{y}_n^{(j)}$ as

$$\mathbf{y}_n^{(j,m)} = \begin{bmatrix} y_n^{(j)} \left[k + \frac{m}{P} \right] \\ y_n^{(j)} \left[k + 1 + \frac{m}{P} \right] \\ \vdots \\ y_n^{(j)} \left[k + W - 1 + \frac{m}{P} \right] \end{bmatrix} \in \mathbb{C}^{W \times 1}, \quad (5.101)$$

where $m = 0, 1, \dots, P-1$ for each j . Each sequence $\mathbf{y}_n^{(j,m)}$ depends on the discrete-time impulse responses $\mathbf{h}^{(j,i,m)}$ for $i = 1, \dots, M_T$. We have

$$\begin{aligned} \mathbf{h}^{(j,i,m)} &\triangleq [h_0^{(j,i,m)}, h_1^{(j,i,m)}, \dots, h_L^{(j,i,m)}] \\ &\triangleq [h^{(j,i)}[t_0 + m\Delta], h^{(j,i)}[t_0 + m\Delta + T], \dots, h^{(j,i)}[t_0 + m\Delta + LT]]. \end{aligned} \quad (5.102)$$

To this end, the filtering matrix associated to the j th receive antenna is defined as

$$\mathbf{H}_T^{(j,:,\cdot)} = \begin{bmatrix} \mathbf{H}_T^{(j,1,0)} & \mathbf{H}_T^{(j,2,0)} & \dots & \mathbf{H}_T^{(j,M_T,0)} \\ \mathbf{H}_T^{(j,1,1)} & \mathbf{H}_T^{(j,2,1)} & \dots & \mathbf{H}_T^{(j,M_T,1)} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{H}_T^{(j,1,P-1)} & \mathbf{H}_T^{(j,2,P-1)} & \dots & \mathbf{H}_T^{(j,M_T,P-1)} \end{bmatrix} \in \mathbb{C}^{P \cdot W \times M_T \cdot (W+L)}, \quad (5.103)$$

where each matrix $\mathbf{H}_T^{(j,i,m)} \in \mathbb{C}^{W \times (W+L)}$ has a banded Toeplitz structure associated to the channel impulse response $\mathbf{h}^{(j,i,m)}$ as

$$\mathbf{H}_T^{(j,i,m)} = \begin{bmatrix} h_0^{(j,i,m)} & \dots & h_L^{(j,i,m)} & 0 & \dots & \dots & 0 \\ 0 & h_0^{(j,i,m)} & \dots & h_L^{(j,i,m)} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & h_0^{(j,i,m)} & \dots & h_L^{(j,i,m)} \end{bmatrix}. \quad (5.104)$$

The filtering matrix associated to all receive antennas \mathbf{H}_{T_P} is a accumulation of $\mathbf{H}_T^{(j, :, :)}$ for $j = 1, \dots, M_R$, we have

$$\mathbf{H}_{T_P} = \begin{bmatrix} \mathbf{H}_T^{(1, :, :)} \\ \mathbf{H}_T^{(2, :, :)} \\ \vdots \\ \mathbf{H}_T^{(M_R, :, :)} \end{bmatrix} \in \mathbb{C}^{M_R \cdot P \cdot W \times M_T \cdot (W+L)}. \quad (5.105)$$

We still can use a 3-way tensor $\mathbf{Y}_P \in \mathbb{C}^{M_R \cdot P \times W \times N}$ to model the oversampled received signals. Only the size of the first dimension changes due to the oversampling compared to the previous tensor \mathbf{Y} . The corresponding input output data model is given by

$$\mathbf{Y}_P = \mathbf{H}_P \times_3 \mathbf{S}^T + \mathbf{N}_P. \quad (5.106)$$

As shown in Figure 5.32, the filtering tensor $\mathbf{H}_P \in \mathbb{C}^{M_R \cdot P \times W \times M_T \cdot (W+L)}$ is organized by aligning the slices of the block matrices in equation (5.103) along the first dimension for all receive antennas. The noise tensor \mathbf{N}_P has the same size as the tensor \mathbf{Y}_P . Notice that the noise samples are not necessarily temporally uncorrelated due to the oversampling.

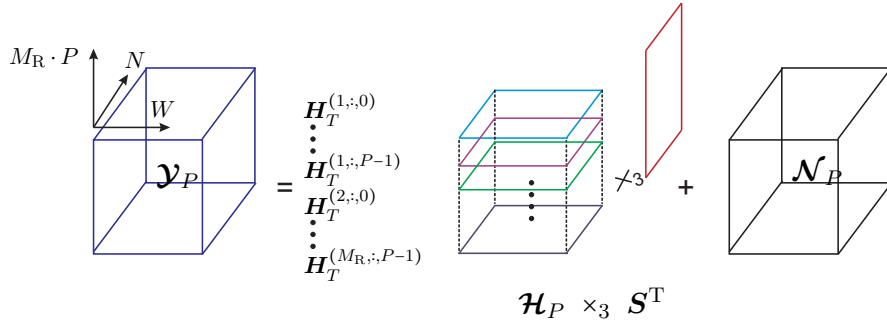


Figure 5.32.: Block diagram of the tensor based data model with oversampling.

By computing the truncated HOSVD of the measurement tensor \mathbf{Y}_P , the signal subspace tensor $\hat{\mathbf{U}}_P^{[s]} \in \mathbb{C}^{M_R \cdot P \times W \times r_3}$ can be estimated as

$$\hat{\mathbf{U}}_P^{[s]} = \mathbf{S}^{[s]} \times_1 \mathbf{U}_{1_P}^{[s]} \times_2 \mathbf{U}_2^{[s]}. \quad (5.107)$$

Here, the ranks of the second and third modes (i.e., r_2 and r_3) are the same as before. Only the rank of the first mode changes to $r_{1_P} = \min(M_R \cdot P, M_T \cdot (L+1))$. Now, the rank-deficient condition for achieving the benefit of the tensor case is loosened to $M_R \cdot P > r_{1_P} = M_T \cdot (L+1)$

due to the oversampling.

5.4.4.4. Signal subspace based parameter estimation

With the estimated signal subspace, the matrix-based method and tensor-based method follow the same parameter estimation procedure as for the SIMO channel. Using the tensor case as an example, we give a short description.

Since the column spaces of $\left[\hat{\mathbf{u}}^{[s]}\right]_{(3)}^T$ and $[\mathbf{H}]_{(3)}^T$ coincide, the unknown MIMO channel coefficients incorporated in the filtering tensor \mathbf{H} can be identified up to a right multiplication by an invertible matrix \mathbf{A} by solving the maximization of the quadratic form $q(\mathbf{H}) \stackrel{\text{def}}{=} \sum_{i=1}^{M_T \cdot (W+L)} \left\| \hat{\mathbf{U}}_{s_T}(:, i)^H [\mathbf{H}]_{(3)}^T \right\|_2^2$. Here, we use $\hat{\mathbf{U}}_{s_T}$ to indicate the estimated signal subspace of the tensor case for notational simplicity (i.e., $\hat{\mathbf{U}}_{s_T} = \left[\hat{\mathbf{u}}^{[s]}\right]_{(3)}^T$). The maximization problem can be solved by utilizing Lemma 5.4.2. The solutions of the maximization problem are the eigenvectors associated to the M_T largest eigenvalues of the matrix \mathbf{G} . The matrix \mathbf{G} obeys the same formulation as in equation (5.76), but has a different size of $M_R \cdot (L+1) \times M_T \cdot (L+W)$ compared to the matrix \mathbf{G} in (5.76). The further determination of the unknown invertible matrix \mathbf{A} can be done by introducing semi-blind channel estimation (SBCE). The basic idea of SBCE is to assign a few pilots on each transmit antenna to acquire a rough estimate of the channel (denoted as $\widetilde{\mathbf{H}}$) which is utilized to identify the unsolved part after the BCE. Assuming that $\hat{\mathbf{H}}$ is the channel estimate from the blind method, the matrix \mathbf{A} can be identified by solving a minimization problem, i.e., $\min_{\mathbf{A}} \left\| \hat{\mathbf{H}} \mathbf{A} - \widetilde{\mathbf{H}} \right\|_{\mathcal{F}}^2$. The solution² is $\mathbf{A} = \hat{\mathbf{H}}^+ \widetilde{\mathbf{H}}$. Alternatively, the authors of [MLM97, GL97] have proposed a blind identification of the unknown matrix \mathbf{A} , which employs a blind source separation algorithm (e.g., joint approximate diagonalization of eigen-matrices (JADE) source separation procedure [CC96]) and only works for the channel estimation in the frequency domain.

The necessary conditions for the channel identifiability are listed as follows.

1. The correlation matrix \mathbf{R}_{ss} has full-rank but is otherwise unknown, which requires $N \geq M_T \cdot (W + L)$.
2. The matrix \mathbf{H}_T has a full column rank.
3. The number of transmit antennas M_T is strictly less than the number of receive antennas M_R .
4. The observed data window length is greater than the channel order L (i.e., $W > L$).

²This solution only exists under the condition $M_R > M_T$, otherwise the matrix \mathbf{A} cannot be identified.

5. The number of channel taps $L + 1$ has been correctly estimated before.
6. The noise samples are uncorrelated with the input data.

5.4.4.5. Simulation results

Compared to the matrix-based subspace method, the new tensor-based subspace method for the blind estimation of MIMO channels shows a performance improvement. The evaluation is performed in terms of the root mean square error (RMSE) of the estimated normalized channels. This RMSE is defined as

$$\text{RMSE} = \frac{1}{P} \sqrt{\text{E} \left\{ \left\| \hat{\mathbf{H}} \mathbf{A} - \mathbf{H} \right\|_F^2 \right\}}, \quad (5.108)$$

where \mathbf{A} is the invertible matrix and is computed as $\mathbf{A} = \hat{\mathbf{H}}^+ \mathbf{H}$. The channel matrix \mathbf{H} is normalized to unit Frobenius norm and the RMSE is averaged over 500 channel realization. The emitted signal is in 4-QAM format. The signal to noise ratio (SNR) is defined as $10 \log_{10} \frac{\text{E} \{ \|\mathbf{s}(k)\|^2 \}}{\text{E} \{ \|\mathbf{n}(k)\|^2 \}}$. To simulate a multipath environment, we adopt a commonly used model [TXK91] to construct an $(L + 1)$ -ray multipath continuous-time channel $h^{(j,i)}(t)$ between the j th receive antenna and i th transmit antenna from a raised cosine pulse shaping filter $g_c(t - \gamma_\ell, \beta)$. We have

$$h^{(j,i)}(t) = \sum_{\ell=0}^L \alpha_\ell^{(j,i)} g_c(t - \gamma_\ell^{(j,i)}, \beta), \quad (5.109)$$

where the roll-off factor β is set to 0.5 for simulations and $\alpha_\ell^{(j,i)}$ are complex valued Gaussian random variables. The term $\gamma_\ell^{(j,i)}$ indicates the delay of the ℓ th path. We use the same way to define $\alpha_\ell^{(j,i)}$ and $\gamma_\ell^{(j,i)}$ as what we have done for the SIMO channel case. The discrete-time channel is obtained by sampling $h^{(j,i)}(t)$ at a rate of T/P . The length of the observed data window is $W = 10$. We introduce a smoothing window to observe the measurement data with a smoothing parameter η as defined for SIMO channel case. We also use the same way to construct the noise samples for the oversampling case as that for SIMO channel.

First, we evaluate the case that satisfies the condition $M_R > r_1$ and $\eta = W$. We consider a MIMO channel consisting of 2 taps and a transmitter with 2 transmit antennas. The first mode rank of the measurement tensor is $M_T \cdot (L + 1)$ (i.e., $r_1 = 4$). The number of receive antennas is greater than r_1 . In this case, we do not employ oversampling at the receiver, the value P is set to 1. White Gaussian noise is added to the output. The performance improvement introduced by the tensor method is observed in Figure 5.33. It is noticed that the performance

improvement increases with the larger difference between M_R and r_1 .

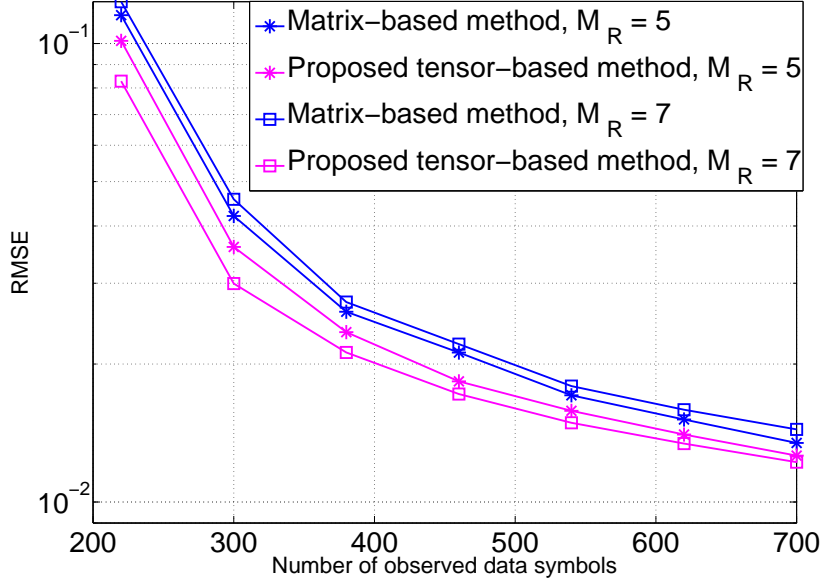


Figure 5.33.: RMSE for MIMO channels: $M_T = 2$, $L + 1 = 2$, $r_1 = 4$, SNR = 20 dB.

Then, we consider the case $M_R \leq r_1$ with $\eta = W$. We assume a MIMO channel consisting of 4 taps. The transmitter is equipped with 2 transmit antennas and the receiver has 5 receive antennas, which leads to $M_R < M_T \cdot (L + 1)$. Under this condition, both the tensor-based subspace method and the matrix-based subspace method achieve the same performance as shown in Figure 5.34. In order to maintain the benefit of the tensor method, we introduce oversampling at the receiver. For fair comparisons, the oversampling is utilized for both matrix-based and tensor-based methods. It is observed that a performance improvement is achieved by the tensor-based method with the oversampling factor $P > 1$, since the benefit condition of the tensor method is loosened to $M_R \cdot P > M_T \cdot (L + 1)$. Larger values of P lead to more significant improvements.

Furthermore, we vary the smoothing parameter within $1 \leq \eta < W$ for both the matrix-based and the tensor-based subspace methods. In Figure 5.35, it is shown that the accuracy of the estimate improves with the decrease of the parameter η . The proposed tensor-based method always outperforms the current matrix-based subspace method for different η . But notice that for the same number of observed data symbols, the smaller η results in an increased number of observed data windows.

In Figure 5.36, we investigate the BER performance of a MIMO system where $M_T = 2$,

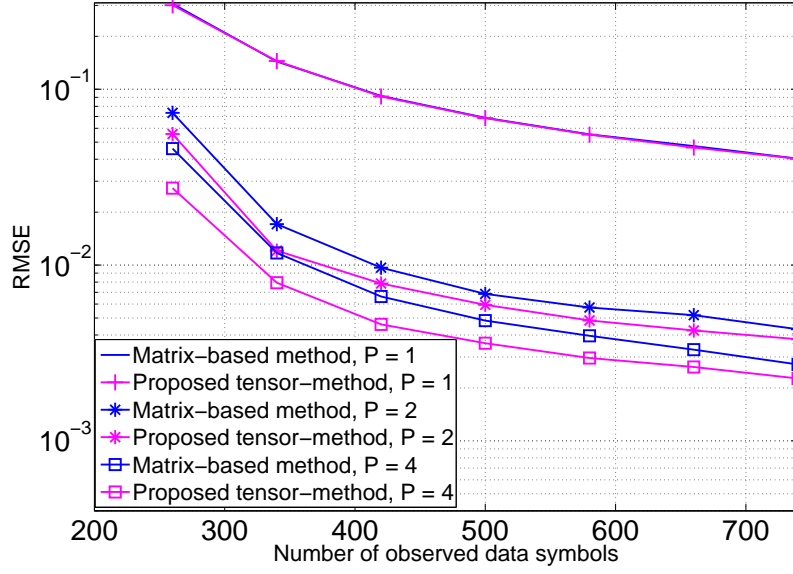


Figure 5.34.: RMSE for MIMO channels: $M_R = 5$, $M_T = 2$, $L + 1 = 4$, SNR = 20 dB.

$M_R = 10$, $L + 1 = 3$, $P = 1$, and $\eta = 10$. We introduce a few pilots per transmit antenna to identify the unknown matrix \mathbf{A} due to the fact that the unknown MIMO channel coefficients can be identified up to an invertible matrix \mathbf{A} by BCE techniques (namely semi-blind channel estimation (SBCE)). Since $M_R > M_T \cdot (L + 1)$, the tensor based BCE can achieve an improved channel estimate compared to the matrix method in terms of the RMSE. Figure 5.36 shows that this improvement leads to a better BER performance of the tensor method compared to the matrix case. Here, we observe 27 data symbols and the training-based channel estimation employs the MMSE estimator [WZ07].

5.5. Summary and Conclusions

In this chapter of the thesis, we have discussed several new approaches for CSI acquisition under different channel conditions.

* Time-varying correlated channel model

When the channel varies too fast to capture the instantaneous CSI (namely short-term CSI), the long-term CSI based on second-order channel statistics is considered alternatively. The proposed rank-one approximated long-term CSI (ROLT-CSI) approach represents the channel by exploiting the knowledge of the estimated spatial correlation per

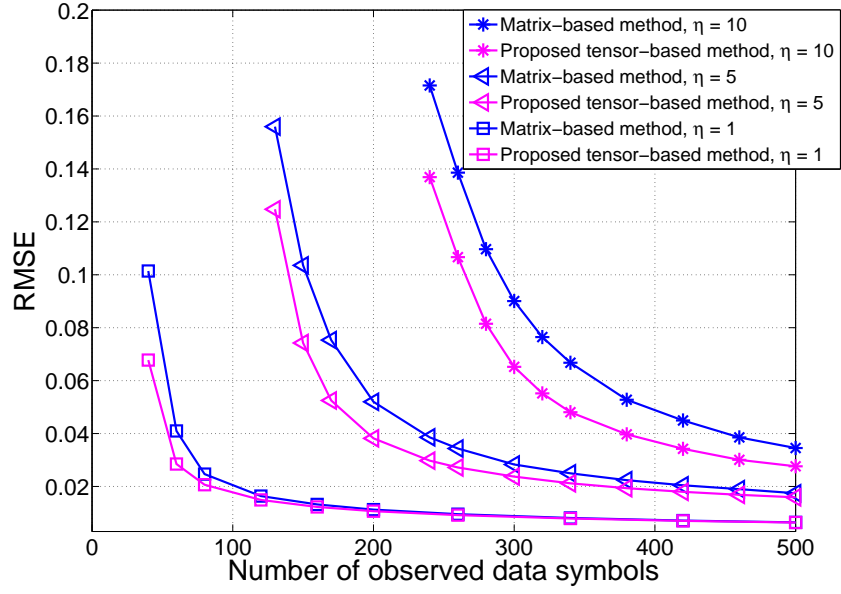


Figure 5.35.: RMSE for blind estimation of MIMO channels with varied smoothing parameter η . $M_R = 10$, $M_T = 2$, $L + 1 = 3$, $r_1 = M_T \cdot (L + 1) = 6$, and SNR = 20 dB.

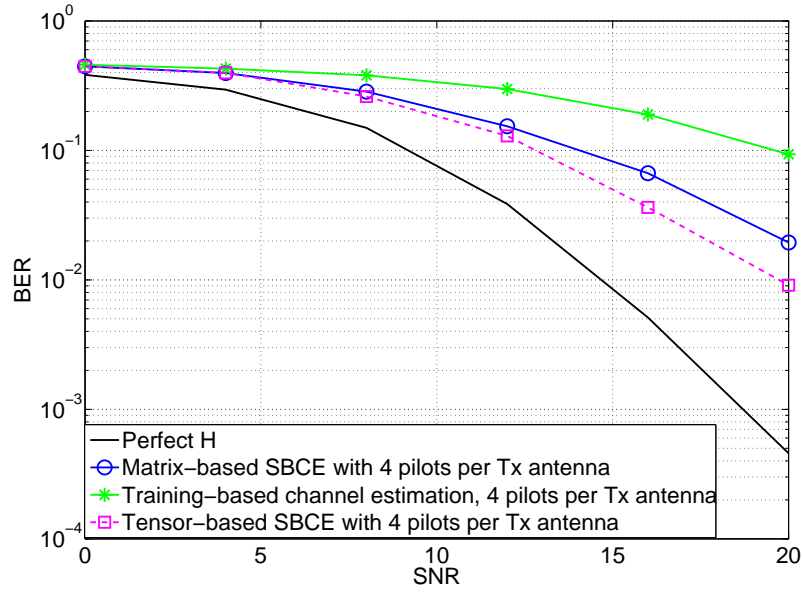


Figure 5.36.: BER performance comparison for semi-blind estimation and training-based channel estimation of MIMO channels with $M_R = 10$, $M_T = 2$, $L + 1 = 3$, $P = 1$, and $\eta = 10$.

receive antenna and transmitting then along the dominant eigenmode of the exploited spatial correlations. Therefore, ROLT-CSI is more efficient compared to the previous long-term method if the channels have a high spatial correlation (e.g., LOS channels). Even for the case that only NLOS channels are considered, the presented ROLT-CSI still achieves some performance gain relative to the previous long-term CSI method.

* **Limited feedback channel model**

In FDD system, CSI is in general obtained by using feedback from the receiver of the forward links. In order to reduce the feedback overhead, a finite rate feedback model is considered, which requires that the channel matrix is quantized first before it is fed back. Most channel quantization schemes for multi-user MIMO downlink employing precoding only consider the directions of the channel [Jin06, RJ07], or they quantize the channel directions and magnitudes separately [KZH08]. The quantization scheme we presented in this chapter stacks the vectors of the channel matrix to maintain the relative magnitude information for the columns of the channels and further quantizes the stacked vector. Moreover, it is only required to calculate the minimum value of the angle spanned by the two vectors instead of calculating the minimum value of the sum of the principal angles spanned by the columns of the channel.

The choice of the codebook significantly affects the quality of the quantized CSI. If the conventional random vector quantization (RVQ) codebook is considered, the limited feedback multi-user MIMO system becomes interference limited with the increased SNR when the number of feedback bits is fixed. On the other hand, the throughput loss can be maintained with the increased SNR when the number of feedback bits per user is scaled linearly with the SNR. However, the number of feedback bits can grow very large to result in a high computational complexity in simulations. Therefore, only MIMO systems with a small number of antennas have been simulated in this chapter. If more efficient codebooks are considered (e.g., the DE-LBG vector quantization codebook), the quantized channel has therefore an enhanced quality. However, a training sequence is in general required for efficient codebook design.

* **Channel estimation with blind techniques**

Blind or semi-blind channel estimation has been well studied for various channel contexts due to its bandwidth efficiency. The second-order statistics (SOS) based blind channel estimation requires the estimate of the signal subspace. A more accurate estimate of the signal subspace leads to a higher quality estimate of the unknown channels. Inspired by it, we use tensors to model the measurement data in order to acquire an enhanced

estimate of the signal subspace compared to the matrix-based data model. The truncated HOSVD instead of the truncated SVD increases the computational complexity slightly, but both have the same order for the number of required multiplications. Notice that the benefit of the tensor gain is only obtained when the structured tensor is rank-deficient in at least one of the R modes. In this chapter, only the tensor-based blind channel estimation of SIMO and MIMO channels is discussed. Actually, we have also extended it to the semi-blind estimation of MIMO systems employing arbitrary orthogonal space-time block codes and per-antenna power loading [RSS⁺11]. It can be expected that many useful applications of the tensor-based signal subspace estimation will be exploited in the future.

6. Conclusions and Outlook

With the development of mobile communications, the demand for reliable high data rates has increased extremely. Multiple-input and multiple-output (MIMO) communication systems are known to provide theoretically attractive and technically feasible solutions that fulfill the aforementioned requirements. Appropriate linear transmit-receive strategies can efficiently exploit the benefits of MIMO technology (e.g., the spatial multiplexing gain and the spatial diversity gain) to promise maximum data rate or diversity order.

In this thesis, we have proposed new designs of linear transmit-receive schemes (e.g., SeDJoCo-based closed-form coordinated beamforming and FlexCoBF) to solve open problems in the area of closed-form coordinated beamforming and improve the performance of the existing algorithms for iterative coordinated beamforming. Furthermore, we have investigated different aspects regarding channel state information (CSI) acquisition by taking into account some practical scenarios.

6.1. Conclusions

In Chapter 2 we go through the fundamentals of MIMO technology. We start with the major achievable benefits of MIMO techniques. Then, we introduce the existing MIMO channel models from the physical perspective and the analytical perspective. Finally, we present the existing capacity results for the single-user MIMO case, the multi-user MIMO case, and the multi-cell MIMO case. They do provide an insight into the capacity limits of the systems with MIMO and the implications of the limits for the practical system designs in spite of the fact that many problems are still unsolved for capacities of MIMO systems.

In Chapter 3 we study the existing well-known linear precoding schemes (i.e., ZF, BD, and RBD) and their throughput losses relative to the DPC scheme. Inspired by the previous works for the performance analysis of ZF and BD, we further propose a method to quantify the average rate and power offsets between RBD, DPC, and BD at high SNRs for two cases. The first case considers a multi-user MIMO broadcast channel where the aggregative number of receive antennas of all users is smaller than or equal to the number of transmit antennas (i.e., $M_T \geq M_R$). We show that the achievable multiplexing gain of RBD is the same as DPC at high SNRs and derive the bounds of the average rate and power offsets between RBD, DPC, and BD as a function of the system parameters. The second case assumes $M_T < M_R$ in a

multi-user MIMO broadcast channel. We find that the achievable multiplexing gain of RBD drops to 1 at high SNRs. Therefore, we recommend to utilize RBD only for the low or medium SNR regime when $M_T < M_R$.

Next, two novel coordinated beamforming algorithms have been developed in Chapter 4. One is proposed as SeDJoCo-based closed-form CBF, which provides a proper solution for the open problem of the closed-form CBF and supports more than two users in the MIMO broadcast channel. The SeDJoCo transformation is a particular form of the classical approximate joint diagonalization (AJD) problem and has the goal of finding a single matrix \mathbf{B} for a set of target matrices $\mathbf{C}_1, \dots, \mathbf{C}_N$ that enforces the i th row and the i th column of the transformed i th target matrix $\mathbf{B}\mathbf{C}_i\mathbf{B}^H$ to approach zero, except for the diagonal element. We prove that the solutions of SeDJoCo exist, but they may not be unique. In addition, we propose two methods to solve SeDJoCo problem (i.e., NCG and STJOCO). It is noticeable that the SeDJoCo transformation can be used not only for the problem of the closed-form CBF but also the problems of blind source separation (BSS) and independent component analysis (ICA). Another novel beamforming algorithm is dedicated to the iterative CBF and named as flexible coordinated beamforming (FlexCoBF). Compared to the existing iterative CBF algorithms, FlexCoBF is designed with the benefits of a high flexibility on the transmit-receive beamforming, a sum rate close to the sum capacity of the MIMO broadcast channel, and a good convergence. We have originally developed FlexCoBF for the multi-user MIMO broadcast channel and we have further extended it to clustered multiple cells. By introducing limited cooperation between clusters (i.e., the concept of coordinated multi-point (CoMP)), the provided numerical results have demonstrated a performance improvement of the proposed algorithm with respect to the cluster throughput and the individual user throughput.

Finally in Chapter 5, three novel approaches associated to channel state information (CSI) acquisition have been proposed for three different channel conditions, respectively. In a multi-user MIMO downlink channel, the available CSI, which is an important issue for the implementation of the linear transmit-receive strategies, can only be obtained at the BS by invoking the reciprocity principle or using feedback from the users. The perfect instantaneous CSI is usually difficult to acquire in practice. Therefore, for the first approach, we show an efficient method to exploit the second-order statistics of a time-varying correlated MIMO channel. This estimated spatial correlation of the channels (also named long-term CSI) can assist BSs to perform linear precoding if perfect CSI is not available. In the second approach, we consider that the BS acquires CSI through a feedback channel from each user. In order to reduce the feedback overhead, we propose a limited feedback strategy which is based on the quantization of the individual channel matrix with a predefined codebook. Instead of directly quantizing

the channel matrix, we stack the vectors of the channel matrix and quantize it in a vector manner. Thus, the relative magnitude information for the columns of the channel matrix is preserved. With the second approach, we can provide the BS with the knowledge of the channel direction and channel magnitude jointly, which is instantly applicable to the multi-user MIMO downlink employing RBD precoding or other linear precoding techniques (e.g., ZF and BD) with power allocations. In the third approach, we introduce the concept of tensors to solve blind channel estimation problems which have been originally constructed and solved by matrix computations. It is motivated by the remarkable advantages of tensor-based signal processing compared to their matrix-based counterparts, which have been described in detail in Chapter 1. The performance improvements by tensor-based processing have been demonstrated in the simulation results.

Overall, the thesis is dedicated to the design of linear transmit-receive strategies and the associated CSI acquisition. We benefit from such designs in multiple ways, e.g., proper solutions for some open issues (as for SeDJoCo-based closed-form CBF and the quantization scheme for RBD based systems), improved performance and flexible extension (as for FlexCoBF, ROLT-CSI, and tensor-based blind channel estimation), or profound analytical results for the system design (as for throughput approximations for linear precoding schemes at high SNRs).

6.2. Future Works

The thesis has addressed a broad spectrum of topics associated with multi-user MIMO techniques, CoMP, and linear transmit-receive strategies. Since they are also the hot topics in 4G and 5G system design, many exciting directions for future research can be opened up.

Let us start with the coordinated beamforming (CBF) discussed in Chapter 4. Our proposed SeDJoCo-based closed-form CBF can effectively solve the open problem existing in the CBF techniques and support a multi-user MIMO downlink system with an arbitrary number of users and transmit antennas. However, there are some shortcomings and challenges. Firstly, only one data stream transmission per user has been considered and demonstrated for the SeDJoCo-based closed-form CBF. From a theoretical perspective, it can be directly extended to the multiple data streams transmission (e.g., r_i data streams for the i th user) for the user i by repeating the target matrix (i.e., \mathbf{C}_i) of this user r_i times as we have mentioned in Section 4.3.5. But it has not been demonstrated. It will be more interesting to find a new transformation which is similar to SeDJoCo but can simultaneously “drill” the off-diagonal elements of the r_i rows and columns of the target matrix \mathbf{C}_i to zeros. However, it is a really challenging task. Another challenging task is associated with the two proposed solutions of the SeDJoCo

transformation (i.e., NCG and STJOCO). We have discussed their advantages and shortcomings in Section 4.5. The further investigations are expected to enhance them, especially for the case with a large number of target matrices. Finally, all of the existing solutions for the SeDJoCo transformation (i.e., NCG, STJOCO, IR proposed in [DZ04], and multiplicative update method in [PG97]) are iterative algorithms. It would be desirable to find a closed-form solution, although the closed-form solution might not exist. Considering the proposed FlexCoBF algorithm for the single cell and the clustered multiple cell scenarios, there are some challenges and interesting directions for future works. We have shown that the existing linear precoding techniques can be applied as the transmit beamforming strategy for FlexCoBF. Actually, some non-linear precoding schemes such as Tomlinson-Harashima precoding (THP) also suffer from the dimensionality constraint, i.e., the number of transmit antennas is equal to or greater than the aggregate number of receive antennas. Therefore, applying the FlexCoBF approach with non-linear precoding techniques as the transmit strategy is a promising path as well. Compared to FlexCoBF employing linear precoding as the transmit strategy, a potential for sum rate and bit error rate (BER) improvement is expected. Actually, we have started to investigate the FlexCoBF algorithm employing THP. Our preliminary results show a remarkable performance gain in terms of the BER and the sum rate compared to the FlexCoBF algorithm employing linear precoding techniques [ZSHdL14a]. Another direction is to introduce lattice reduction (LR) techniques to further improve the performance of FlexCoBF, which is named as LR-aided FlexCoBF. LR techniques have the potential to transform a set of non-orthogonal matrices to be nearly orthogonal [LLL82, ZdLH13b, ZdL12, SKMG05]. As a result, the LR-aided FlexCoBF algorithm can achieve the maximum diversity order (i.e., M_T) and the maximum spatial multiplexing gain (i.e., M_T) at high SNRs [ZSHdL14b]. In particular, the diversity-multiplexing tradeoff changes as compared to the standard FlexCoBF as the LR-aided approach obtains a higher diversity order and a higher spatial multiplexing gain. It is a really promising observation. A third direction is to investigate the potential to implement the extended FlexCoBF algorithm by considering the challenges of CoMP (the details of these challenges are described in Section 4.5). In particular, instead of the clusters with a fixed size, the cluster could be set up adaptively [MF11] according to factors such as the user positions and the RF channel measurements. Finally, applying all these methods to an OFDM system or a filter bank multicarrier (FBMC) system [SS96, SK00, RSFBB10] is an interesting direction as well. Instead of a straightforward extension, we can consider the OFDM or FBMC system with a specific resource allocation (e.g., chunk as a basic resource element), a defined pilot structure, and channel prediction and interpolation.

Concerning the throughput approximation for the linear precoding discussed in Chapter 3,

there are interesting directions for future works as well. We have quantified the average rate and power offsets between linear precoding techniques (i.e., ZF, BD, and RBD) and DPC, which are really interesting and useful for system designs. However, similar analytical results for CBF techniques are still missing. It would be a direction to utilize the capacity approximation framework introduced in Chapter 3 to solve the analytical performance assessment problem of FlexCoBF. Furthermore, non-linear precoding techniques have been demonstrated to achieve a better BER and sum rate performance compared to the linear precoding techniques [ZdLH12, ZdLH13a]. It would also be desirable to quantify the average rate and power offsets between non-linear precoding and linear precoding as a function of the system parameters, although the task is challenging.

Finally, considering the channel state information (CSI) acquisition strategies in Chapter 5, some challenges remain as well. In our proposed upper bound for the throughput loss of RBD caused by the finite rate feedback, the residual multi-user interference (MUI) of RBD (i.e., ΔI_i in equation (5.37)) has not been specified. We only use an experimental value of ΔI_i in our simulations. An exploitation of the distribution and dependence of ΔI_i on the SNR is an interesting open area. Moreover, a random vector quantization (RVQ) codebook can be easily constructed, but has a poor efficiency, especially for a system with a large number of users and antennas where the number of feedback bits can grow very large at high SNRs in order to maintain a constant throughput loss. An efficient codebook design is of significant practical interest. Also, we have shown blind channel estimation as one successful application of tensors. The consideration of applying the same idea to different applications, such as carrier offset estimation techniques, orthogonal space-time block coding schemes, and multi-user detection is a promising path as well.

Appendix A.

Glossary of Acronyms, Symbols and Notation

A.1. Acronyms

1G	First-generation
2G	Second-generation
3G	Third-generation
4G	Fourth-generation
5G	Fifth-generation
AJD	Approximated Joint Diagonalization
AP	Access Point
AR	Auto Regressive
AWGN	Additive White Gaussian Noise
BC	Broadcast
BCE	Blind Channel Estimation
BD	Block Diagonalization
BS	Base Station
BSS	Blind Source Separation
CBF	Coordinated Beamforming
CCI	Channel Covariance Information
CCDF	Complementary Culmulated Distribution Function
CDI	Channel Distribution Information
CMI	Channel Mean Information
CoMP	Coordinated Multipoint
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSIT	Channel State Information at the Transmitter
DFE	Decision-Feedback Equalization
DOA	Direction of Arrival
DOD	Direction of Departure
DPC	Dirty Paper Coding
FBMC	Filter Bank Multicarrier
FDD	Frequency Division Duplex
FIR	Finite Impulse Response
ICA	Independent Component Analysis
IMT-2000	International Mobile Telecommunications-2000
IR	Iterative Relaxation
ISI	Inter-Symbol Interference
LR	Lattice Reducution

MAC	Multiple Access Channel
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MRC	Maximal Ratio Combining
MUI	Multi-User Interference
NCG	Newton with Conjugate Gradient
OFDM	Orthogonal Frequency Division Multiplexing
OSTBC	Orthogonal Space Time Block Coding
PD	Positive Definite
PSDs	Power Spectral Densities
QML	Quasi Maximum Likelihood
QoS	quality of service
RBD	Regularized Block Diagonalization
RF	Radio Frequency
RHS	Right Hand Side
SBCE	Semi-Blind Channel Estimation
SDMA	Space Division Multiple Access
SeDJoCo	Sequentially Drilled Joint Congruence
SINR	Signal to Interference plus Noise Ratio
SISO	Single Input Single Output
SNR	Signal to Noise Ratio
STC	Space Time Coding
STBC	Space Time Block Coding
STJOCO	Structured Joint Congruence
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
THP	Tomlinson-Harashima Precoding
WSSUSH	Wide-Sense Stationary Uncorrelated Scattering Homogeneous
ZF	Zero-Forcing
ZMSW	Zero-Mean Spatially White

A.2. Symbols and Notation

\mathbb{R}	Set of real numbers
\mathbb{C}	Set of complex numbers
\triangleq	Definition
\leftarrow	update to
a, b, c	scalars
$\mathbf{a}, \mathbf{b}, \mathbf{c}$	column vectors
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	matrices
$\mathcal{A}, \mathcal{B}, \mathcal{C}$	tensors

$\mathbf{0}_{M \times N}$	Matrix of zeros of size $M \times N$
\mathbf{I}_M	Identity matrix of size $M \times M$
$\mathcal{I}_{R,d}$	R -way identity tensor of size $d \times d \dots \times d$
$\mathbf{A}_{(i,j)}$	the (i,j) -element of the matrix \mathbf{A}
$\mathbf{A}(i,j)$	also indicated the (i,j) -element of the matrix \mathbf{A}
$(\cdot)^T$	matrix transpose
$(\cdot)^H$	Hermitian transpose
$\text{abs}(a)$	magnitude of the complex valued variable a
$ \cdot $	determinant of a matrix (product of eigenvalues)
$\ \cdot\ _2$	Euclidean (two-) norm
$\ \cdot\ _{\mathcal{F}}$	Frobenius norm
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product between $\mathbf{A} \in \mathbb{C}^{M \times N}$ and $\mathbf{B} \in \mathbb{C}^{P \times Q}$ defined as $\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{1,1} \cdot \mathbf{B} & a_{1,2} \cdot \mathbf{B} & \cdots & a_{1,N} \cdot \mathbf{B} \\ a_{2,1} \cdot \mathbf{B} & a_{2,2} \cdot \mathbf{B} & \cdots & a_{2,N} \cdot \mathbf{B} \\ \vdots & \vdots & \vdots & \vdots \\ a_{M,1} \cdot \mathbf{B} & a_{M,2} \cdot \mathbf{B} & \cdots & a_{M,N} \cdot \mathbf{B} \end{bmatrix}.$
$\mathbf{a} \circ \mathbf{b}$	outer product of two vectors, which can be regarded as a special case of the Kronecker product of matrices.
$\mathbf{A} \diamond \mathbf{B}$	Khatri-Rao (column-wise Kronecker) product between $\mathbf{A} \in \mathbb{C}^{M \times N}$ and $\mathbf{B} \in \mathbb{C}^{P \times N}$ and $\mathbf{B} \in \mathbb{C}^{M \times N}$ and $\mathbf{B} \in \mathbb{C}^{M \times N}$
$\text{vec}\{\cdot\}$	vec-operator: stack elements of a matrix/tensor into a column vector, begin with first (row) index, then proceed to second (column), third, etc.
$\text{unvec}_{I \times J}\{\cdot\}$	inverse vec-operator: reshape elements of a vector back into a matrix/tensor of indicated size
$\text{diag}\{\cdot\}$	transforms a vector into a square diagonal matrix or extract main diagonal of a square matrix and place elements into a vector
$\text{Bdiag}\{\cdot\}$	creates a block-diagonal matrix from its matrix arguments
$\text{tr}(\cdot)$	trace of a matrix (sum of diagonal elements = sum of eigenvalues)
$\det\{\cdot\}$	determinant of a matrix (product of eigenvalues)
$\text{rank}\{\cdot\}$	rank of a matrix
\mathbf{A}^+	the same space as the columns of $\mathbf{A} \in \mathbb{C}^{M \times r}$ (assuming $r \leq M$) Moore-Penrose pseudo inverse of a matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$, which we can compute via

	$\mathbf{A}^+ = \mathbf{V}_s \cdot \boldsymbol{\Sigma}_s^{-1} \cdot \mathbf{U}_s^H$, where $\mathbf{A} = \mathbf{U}_s \cdot \boldsymbol{\Sigma}_s \cdot \mathbf{V}_s^H$ represents the economy-size SVD of \mathbf{A} (cf. Section 5.4).
	$\mathbf{A}^+ = (\mathbf{A}^H \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^H$ if $\text{rank}\{\mathbf{A}\} = N$ (full column rank)
	$\mathbf{A}^+ = \mathbf{A}^H \cdot (\mathbf{A} \cdot \mathbf{A}^H)^{-1}$ if $\text{rank}\{\mathbf{A}\} = M$ (full row rank).
$[\boldsymbol{\mathcal{X}}]_{(n)}$	n -mode unfolding of tensor $\boldsymbol{\mathcal{X}}$ in reverse cyclical column ordering
$\boldsymbol{\mathcal{X}} \times_n \mathbf{U}$	n -mode product between tensor $\boldsymbol{\mathcal{X}}$ and matrix \mathbf{U}
$\boldsymbol{\mathcal{X}} \times_{r=1}^R \mathbf{U}_r$	repeated n -mode products, short-hand notation for $\boldsymbol{\mathcal{X}} \times_1 \mathbf{U}_1 \dots \times_R \mathbf{U}_R$
$[\mathbf{A} \sqcup_n \mathbf{B}]$	n -mode concatenation of tensors \mathbf{A} and \mathbf{B}
$\mathbb{E}\{X\}$	Expectation operator, i.e., mean of the random variable X
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ , variance σ^2
$\mathcal{CN}(\mu, \sigma^2)$	circularly symmetric complex Gaussian distribution

Appendix B.

Proofs and derivations for Chapter 2

B.1. Proof of equation (2.14)

Applying a SVD of the channel matrix \mathbf{H} , the MIMO channel can be converted into r parallel, non-interfering SISO channels ($r \leq \min(M_R, M_T)$) by premultiplying the input by $\mathbf{V}_s \in \mathbb{C}^{M_T \times r}$ and post-multiplying the output by the matrix $\mathbf{U}_s^H \in \mathbb{C}^{r \times M_R}$. Thus, the received signal in equation (2.11) can be rewritten as

$$\tilde{\mathbf{y}} = \mathbf{\Sigma} \mathbf{d} + \tilde{\mathbf{n}}, \quad (\text{B.1})$$

where $\tilde{\mathbf{y}} = \mathbf{U}_s^H \mathbf{y}$, $\mathbf{x} = \mathbf{V}_s \mathbf{d}$, and $\tilde{\mathbf{n}} = \mathbf{U}_s^H \mathbf{n}$. Note that the unitary rows of \mathbf{U}^H do not change the statistics of the noise \mathbf{n} . With water-filling indicated in equation (2.13), we have

$$\mathbf{Q} = \mathbb{E} \{ \mathbf{x} \mathbf{x}^H \} = \mathbf{V}_s \mathbf{P} \mathbf{V}_s^H, \quad (\text{B.2})$$

where the matrix $\mathbf{P} \in \mathbb{C}^{r \times r}$ (i.e., $\mathbf{P} = \mathbb{E} \{ \mathbf{d} \mathbf{d}^H \}$) is equal to $\text{diag}(P_1, \dots, P_r)$. Thereby, the capacity in equation (2.12) can be further calculated as

$$\begin{aligned} C_{\text{SU}} &= \sum_{i=1}^r \log_2(1 + \sigma_i^2 P_i) \\ &= \sum_{i=1}^r \log_2 \left(1 + \sigma_i^2 \left(\mu - \frac{1}{\sigma_i^2} \right)_+ \right) \\ &= \sum_{i=1}^r \log_2 (1 + (\sigma_i^2 \mu - 1)_+) \\ &= \sum_{i=1}^r \left(\log_2(\mu \sigma_i^2) \right)_+. \end{aligned} \quad (\text{B.3})$$

B.2. Schur-convex and Schur-concave function

The Schur-convex function was first introduced in [Sch23] and has been widely used in the study of majorization.

Definition B.2.1. For two vectors $\mathbf{a} \in \mathbb{R}^d$ and $\mathbf{b} \in \mathbb{R}^d$, if $\sum_{i=1}^d a_i = \sum_{i=1}^d b_i$ and $\sum_{i=1}^k a_i^\downarrow \geq \sum_{i=1}^k b_i^\downarrow$ for

$k = 1, \dots, d$, it is called that \mathbf{a} majorizes \mathbf{b} (i.e., $\mathbf{a} > \mathbf{b}$).

Here, a_i^\downarrow and b_i^\downarrow indicate the elements of \mathbf{a} and \mathbf{b} which are sorted in decreasing order, respectively.

Definition B.2.2. For a function f and two vectors (i.e., $\forall \mathbf{a}, \mathbf{b} \in \mathbb{R}^d$), we have \mathbf{a} majorizes \mathbf{b} (i.e., $\mathbf{a} > \mathbf{b}$). If $f(\mathbf{a}) \geq f(\mathbf{b})$, the function f is called Schur-convex. If $f(\mathbf{a}) \leq f(\mathbf{b})$, the function f is called Schur-concave.

Simple examples:

If f is a convex function defined on a real interval, then $\sum_{i=1}^n f(x_i)$ is Schur-convex.

The Shannon entropy function $\sum_{i=1}^d P_i \cdot \log_2 \frac{1}{P_i}$ is Schur-concave.

Appendix C.

Proofs and derivations for Chapter 3

C.1. Proof of equations (3.24) to (3.25)

For high SNRs, from equation (3.23) we get

$$C_{\text{DPC}}(\mathbf{H}, \tilde{P}_{\text{T}}) \cong \log_2 \left| \mathbf{I}_{M_{\text{R}}} + \frac{\tilde{P}_{\text{T}}}{M_{\text{R}}} \mathbf{H} \mathbf{H}^{\text{H}} \right|. \quad (\text{C.1})$$

Due to the consideration of high SNRs, the equation (C.1) can be further simplified to

$$C_{\text{DPC}}(\mathbf{H}, \tilde{P}_{\text{T}}) \cong \log_2 \left| \frac{\tilde{P}_{\text{T}}}{M_{\text{R}}} \mathbf{H} \mathbf{H}^{\text{H}} \right|. \quad (\text{C.2})$$

To utilize the property of the matrix determinants which is

$$\det(c\mathbf{A}) = c^n \det(\mathbf{A}), \text{ for } \mathbf{A} \in \mathbb{C}^{n \times n} \quad (\text{C.3})$$

and the fact that $\mathbf{H} \mathbf{H}^{\text{H}}$ has the dimension $M_{\text{R}} \times M_{\text{R}}$, we have

$$C_{\text{DPC}}(\mathbf{H}, \tilde{P}_{\text{T}}) \cong \log_2 \left(\left(\frac{\tilde{P}_{\text{T}}}{M_{\text{R}}} \right)^{M_{\text{R}}} \cdot |\mathbf{H} \mathbf{H}^{\text{H}}| \right). \quad (\text{C.4})$$

Thus, we can reach the equation (3.24). That is

$$C_{\text{DPC}}(\mathbf{H}, \tilde{P}_{\text{T}}) \cong M_{\text{R}} \log_2 \tilde{P}_{\text{T}} - M_{\text{R}} \log_2 M_{\text{R}} + \log_2 |\mathbf{H} \mathbf{H}^{\text{H}}|. \quad (\text{C.5})$$

According to the capacity approximation framework (i.e., equations (3.19) to (3.21)), we can calculate the multiplexing gain \mathcal{S}_{∞} by computing the first derivative of equation (C.5). Finally, we reach the equations (3.25) and (3.26).

C.2. Notes on several fundamental functions and distributions

C.2.1. Gamma function

Definition C.2.1. The Gamma function is an extension of the factorial function for real and complex numbers. For example, if n is a positive integer, $\Gamma(n) = (n - 1)!$. If m is a complex number with a positive real part, $\Gamma(m) = \int_0^\infty x^{m-1} e^{-x} dx$ [AAR01].

Figure C.1 from [AAR01] is shown here as an example of the Gamma function.

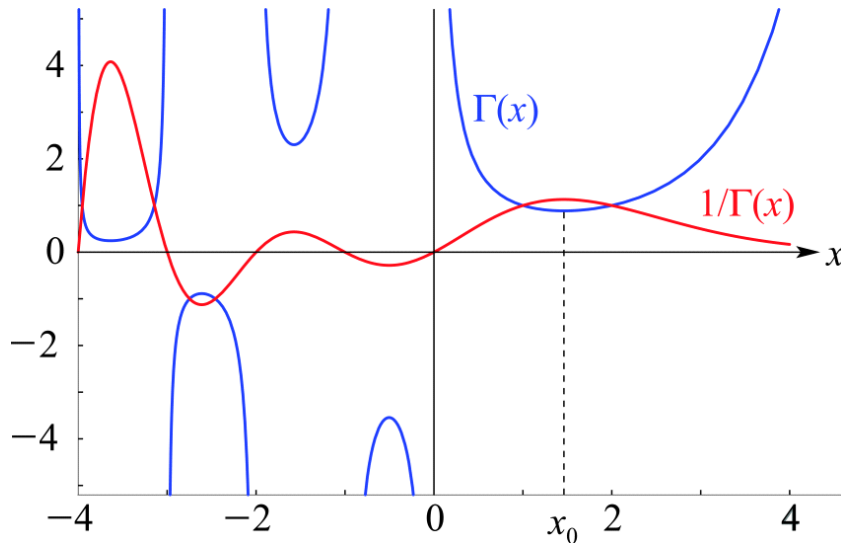


Figure C.1.: Example of $\Gamma(x)$ and $\frac{1}{\Gamma(x)}$ for real valued number, where $x_0 = 1.46163$.

Some particular values of the gamma function are shown as follows:

$$\begin{aligned}
 \Gamma\left(-\frac{3}{2}\right) &= \frac{4}{3}\sqrt{\pi} \\
 \Gamma(-1) &= \infty \\
 \Gamma\left(-\frac{1}{2}\right) &= -2\sqrt{\pi} \\
 \Gamma(0) &= \infty \\
 \Gamma\left(\frac{1}{2}\right) &= \sqrt{\pi} \\
 \Gamma(1) &= 1
 \end{aligned}$$

(C.6)

C.2.2. Digamma function

Definition C.2.2. The digamma function is defined as the logarithmic derivative of the gamma function $\Gamma(x)$ [AAR01], i.e.,

$$\varphi(x) = \frac{d}{dx} \ln \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}. \quad (\text{C.7})$$

Some fundamental properties of digamma function:

The digamma function satisfies the recurrence relation, i.e., $\varphi(x+1) = \varphi(x) + \frac{1}{x}$.

Applying the recurrence relation of the digamma function, the digamma function $\varphi(x)$ can be expressed as

$$\varphi(x) = \sum_{\ell=1}^{x-1} \frac{1}{\ell} - \gamma, \quad (\text{C.8})$$

where γ is the Euler-Mascheroni constant (i.e., $\gamma = 0.577215$).

Equation (C.8) implies $\varphi(1) = -\gamma$.

C.2.3. Wishart distribution

Definition C.2.3. If the columns of the matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ are zero-mean independent real/complex Gaussian vectors with covariance matrix $\mathbf{\Sigma}$, the random matrix $\mathbf{W} = \mathbf{A}\mathbf{A}^H \in \mathbb{C}^{m \times m}$ is a real/complex Wishart matrix with n degrees of freedom and covariance matrix $\mathbf{\Sigma}$, i.e., $\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{\Sigma})$. In other word, the matrix \mathbf{W} has a Wishart distribution. The probability density function of the Wishart matrix \mathbf{W} for $n \geq m$ is

$$f(\mathbf{W}) = \frac{\pi^{-m(m-1)/2}}{\det(\mathbf{\Sigma})^n \prod_{i=1}^m (n-i)!} \exp(-\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{W})) \det(\mathbf{W})^{n-m}. \quad (\text{C.9})$$

Some fundamental properties of a Wishart matrix $\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{I}_m)$ [TV04]:

$$\mathbb{E}\{\text{tr}(\mathbf{W})\} = mn, \text{ for } n \geq m$$

$$\mathbb{E}\{\text{tr}(\mathbf{W}^2)\} = mn(m+n), \text{ for } n \geq m$$

$$\mathbb{E}\{\text{tr}^2(\mathbf{W})\} = mn(mn+1), \text{ for } n \geq m$$

$$\mathbb{E}\{\text{tr}(\mathbf{W}^{-1})\} = \frac{m}{n-m}, \text{ for } n > m$$

$$\mathbb{E}\{\text{tr}(\mathbf{W}^{-2})\} = \frac{mn}{(n-m)^3 - n + m}, \text{ for } n > m + 1$$

Let $\mathbf{B} \in \mathbb{C}^{p \times m}$ be a complex Gaussian matrix with zero-mean unit-variance entries. For $p \leq m \leq n$, we have

$$\mathbb{E} \left\{ \log_e \det(\mathbf{B} \mathbf{W}^{-1} \mathbf{B}^H) \right\} = \sum_{\ell=0}^{p-1} \left(\varphi(m-\ell) - \varphi(p+n-m-\ell) \right). \quad (\text{C.10})$$

C.2.4. Chi-square distribution

Definition C.2.4. If a_1, \dots, a_K are independent Gaussian random variables, the sum of their squares $A = \sum_{i=1}^K a_i^2$ has a chi-square distribution with K degrees of freedom, i.e., $A \sim \mathcal{X}_K^2$. The probability density function of the chi-squared distribution is [AS65]

$$f(A) = \begin{cases} \frac{A^{(K/2-1)} \exp(-A/2)}{2^{K/2} (\frac{K}{2}-1)!}, & A \geq 0; \\ 0, & \text{otherwise.} \end{cases}$$

C.3. Proof of Theorem 3.3.4

Substituting equations (3.60) and (3.38) into equation (3.62), we get

$$\begin{aligned} \bar{\Delta}_{\text{RBD-BD}} &= \mathbb{E} \left\{ \log_2 \prod_{i=1}^K \left(1 + \frac{\mu_i M_R}{\tilde{P}_T} \right) \right\} \\ &= K \mathbb{E} \left\{ \log_2 \left(1 + \frac{\mu_i M_R}{\tilde{P}_T} \right) \right\} \\ &\leq K \log_2 \left(1 + \frac{M_R}{\tilde{P}_T} \mathbb{E} \{ \mu_i \} \right), \end{aligned} \quad (\text{C.11})$$

where

$$0 \leq \mathbb{E} \{ \mu_i \} \leq \overbrace{\mathbb{E} \left\{ \text{tr} \left((\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H})^{-1} \right) \right\}}^{(a)} \overbrace{\mathbb{E} \left\{ \text{tr} \left(\mathbf{G}_i^{(1)} \mathbf{G}_i^{(1)H} \right) \right\}}^{(b)}. \quad (\text{C.12})$$

Note that $\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H} \in \mathbb{C}^{M_r \times M_r}$ is Wishart distributed with $M_T - M_R + M_r$ degrees of freedom and $\mathbf{G}_i^{(1)} \mathbf{G}_i^{(1)H} \in \mathbb{C}^{M_r \times M_r}$ is Wishart distributed with $M_R - M_r$ degrees of freedom.

To utilize the properties of Wishart matrices in [TV04]

$$\mathbb{E} \{ \text{tr}(\mathbf{W}) \} = mn, \text{ for } n \geq m \quad (\text{C.13})$$

$$\mathbb{E} \{ \text{tr}(\mathbf{W}^2) \} = mn(m+n), \text{ for } n \geq m \quad (\text{C.14})$$

and

$$\mathbb{E} \left\{ \text{tr}(\mathbf{W}^{-1}) \right\} = \frac{m}{n-m}, \text{ for } n > m, \quad (\text{C.15})$$

where the matrix $\mathbf{W} \in \mathbb{C}^{m \times m}$ is a complex Wishart matrix with n degrees of freedom, the term (a) and term (b) are calculated as $\frac{M_r}{M_T - M_R}$ and $M_r(M_R - M_r)$ for the case $M_T > M_R$ (i.e., $n > m$), respectively. Thereby, if $M_T > M_R$, it leads to $0 \leq \mathbb{E} \{ \mu_i \} \leq \frac{M_r^2(M_R - M_r)}{M_T - M_R}$.

For the case that $M_T = M_R$ (i.e., $n = m$), the term (a) can be evaluated by Theorem 1 in [BG96], where the lower bound for $\text{tr}(\mathbf{W}^{-1})$ has been utilized here. That is

$$\mathbb{E} \left\{ \text{tr}(\mathbf{W}^{-1}) \right\} \geq [\mathbb{E} \left\{ \text{tr}(\mathbf{W}) \right\} \quad n] \begin{bmatrix} \mathbb{E} \left\{ \text{tr}(\mathbf{W}^2) \right\} & \mathbb{E} \left\{ \text{tr}(\mathbf{W}) \right\} \\ \xi^2 & \xi \end{bmatrix}^{-1} \begin{bmatrix} n \\ 1 \end{bmatrix}, \quad (\text{C.16})$$

where the matrix $\mathbf{W} \in \mathbb{C}^{n \times n}$ is a complex Wishart matrix with n degrees of freedom and ξ denotes the average of the largest eigenvalues of the matrix \mathbf{W} . Then, we can reach Theorem 3.3.4.

C.4. Proof of Theorem 3.3.5

Substituting equations (3.24) and (3.60) into equation (3.66), we get

$$\begin{aligned} \bar{\Delta}_{\text{DPC-RBD}} &= \mathbb{E} \left\{ \log_2 |\mathbf{H}\mathbf{H}^H| \right\} - \mathbb{E} \left\{ \log_2 \prod_{i=1}^K \left(1 + \frac{\mu_i M_R}{\tilde{P}_T} \right) \right\} - \mathbb{E} \left\{ \log_2 \prod_{i=1}^K |\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H}| \right\} \\ &= \mathbb{E} \left\{ \log_2 |\mathbf{H}\mathbf{H}^H| \right\} - K \mathbb{E} \left\{ \log_2 \left(1 + \frac{\mu_i M_R}{\tilde{P}_T} \right) \right\} - K \mathbb{E} \left\{ \log_2 |\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H}| \right\} \\ &\geq \mathbb{E} \left\{ \log_2 |\mathbf{H}\mathbf{H}^H| \right\} - K \log_2 \left(1 + \frac{M_R}{\tilde{P}_T} \mathbb{E} \{ \mu_i \} \right) - K \mathbb{E} \left\{ \log_2 |\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H}| \right\} \\ &\geq \mathbb{E} \left\{ \log_2 |\mathbf{H}\mathbf{H}^H| \right\} - K \log_2 \left(1 + \frac{\mu M_R}{\tilde{P}_T} \right) - K \mathbb{E} \left\{ \log_2 |\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H}| \right\} \end{aligned} \quad (\text{C.17})$$

Note that $\mathbf{H}\mathbf{H}^H$ is Wishart distributed with M_T degrees of freedom and $\mathbf{G}_i^{(0)} \mathbf{G}_i^{(0)H}$ is Wishart distributed with $M_T - M_R + M_r$ degrees of freedom. According to the property of Wishart matrices in equation (2.12) of [TV04]

$$\mathbb{E} \left\{ \log_e \det \mathbf{W} \right\} = \sum_{\ell=0}^{m-1} \varphi(n - \ell), \quad (\text{C.18})$$

where the matrix $\mathbf{W} \in \mathbb{C}^{m \times m}$ is a complex Wishart matrix with the freedom n ($n \geq m$), and $\varphi(\cdot)$ is a digamma function which for natural arguments can be expressed as

$$\varphi(m) = \varphi(1) + \sum_{\ell=1}^{m-1} \frac{1}{\ell}, \quad (\text{C.19})$$

where $\varphi(1)$ is the Euler-Mascheroni constant ($-\varphi(1) = 0.577215$), we can reach Theorem 3.3.5.

Appendix D.

Proofs and derivations for Chapter 4

D.1. Proof of Theorem (4.2.1)

Equation (4.8) implies that $\mathbf{R}_1 \mathbf{f}_1$ is orthogonal to the vector \mathbf{f}_2 . Similarly, equation (4.9) implies that $\mathbf{R}_2 \mathbf{f}_1$ is orthogonal to the vector \mathbf{f}_2 . Since \mathbf{f}_2 is a one-dimensional non-zero vector, its null space is one-dimensional. Therefore, $\mathbf{R}_1 \mathbf{f}_1$ and $\mathbf{R}_2 \mathbf{f}_1$ must be co-linear. Then, we have

$$\mathbf{R}_1 \mathbf{f}_1 = \lambda_1 \mathbf{R}_2 \mathbf{f}_1, \quad (\text{D.1})$$

where λ_1 is a constant scalar. If we take the Hermitian transpose of equations (4.8) and (4.9), and apply the same argument to $\mathbf{R}_1 \mathbf{f}_2$ and $\mathbf{R}_2 \mathbf{f}_2$, we get equation (D.2) as follows

$$\mathbf{R}_1 \mathbf{f}_2 = \lambda_2 \mathbf{R}_2 \mathbf{f}_2, \quad (\text{D.2})$$

where λ_2 is a constant scalar. Equations (D.1) and (D.2) show that \mathbf{f}_1 and \mathbf{f}_2 are the generalized eigenvectors of the channel correlation matrices $(\mathbf{R}_1, \mathbf{R}_2)$. It is also equivalent to \mathbf{f}_1 and \mathbf{f}_2 being eigenvectors of $\mathbf{R}_1^{-1} \mathbf{R}_2$ as well as $\mathbf{R}_2^{-1} \mathbf{R}_1$.

D.2. Notes on inverse discrete time Fourier transform (IDTFT)

Given a sequence $x[n]$ corresponding to the samples of a continuous time function $x(t)$ at the discrete time $t = nT_s$ where T_s is the sampling interval, the discrete time Fourier transform (DTFT) of $x[n]$ is calculated as

$$X_{\frac{1}{T_s}}(f) = \sum_{n=-\infty}^{\infty} x[n] e^{-j2\pi f T_s n} \quad (\text{D.3})$$

To recover the discrete data sequence from the DTFT function, we have to calculate the inverse DTFT. It is given by

$$x[n] = T_s \int_{\frac{1}{T_s}} X_{\frac{1}{T_s}}(f) \cdot e^{j2\pi f n T_s} df, \quad (\text{D.4})$$

where $X_{\frac{1}{T_s}}(f)$ is periodic in frequency with period $f_p = \frac{1}{T_s}$. Equation (D.4) is an integral over any interval of length f_p . Applying the above equation (D.4) and substituting $v = fT_s$, $n = t$, and $X_{\frac{1}{T_s}}(f) = \frac{1}{h_n(v)}$, we reach equation (4.16).

D.3. Notes on Lie group

Definition D.3.1. A real Lie group is a group of a finite-dimensional real smooth manifold where the multiplication and inversion are smooth maps. A morphism of Lie group is a smooth map which preserves the group operation: $f(gh) = f(g)f(h)$.

A complex Lie group is defined in a similar way.

Appendix E.

Proofs and derivations for Chapter 5

E.1. Notes on Jensen's inequality

In probability theory, Jensen's inequality is generally stated as [Haz01]

Definition E.1.1. If $f(\cdot)$ is a convex function and x is a random variable, we have

$$f\left(\mathbb{E}\{x\}\right) \leq \mathbb{E}\{f(x)\}. \quad (\text{E.1})$$

Conversely, if $f(\cdot)$ is a concave function and x is a random variable, we have

$$f\left(\mathbb{E}\{x\}\right) \geq \mathbb{E}\{f(x)\}. \quad (\text{E.2})$$

E.2. Notes on Euclidean space and Grassmannian space

E.2.1. Euclidean space

Definition E.2.1. Let \mathbb{C}^n be an n -dimensional complex vector space. Then, $\mathbb{M} = (\mathbb{C}^n, d)$ is a metric space called Euclidean n -space. The metric d is the Euclidean metric. The general Euclidean metrics are defined as [Haz01]:

For two complex vectors $\mathbf{x} \in \mathbb{C}^n$ and $\mathbf{y} \in \mathbb{C}^n$,

$$\begin{aligned} - d_1(\mathbf{x}, \mathbf{y}) &= \sum_{i=1}^n |x_i - y_i| \\ - d_r(\mathbf{x}, \mathbf{y}) &= \left(\sum_{i=1}^n |x_i - y_i|^r \right)^{\frac{1}{r}}, \quad r \in \mathbb{R} \text{ and } r \geq 1 \\ - d_\infty(\mathbf{x}, \mathbf{y}) &= \max_{i=1}^n |x_i - y_i| \end{aligned}$$

E.2.2. Grassmannian space

Definition E.2.2. The Grassmannian space $\mathbb{G}_{m,n}(\mathbb{C})$ is the set of all n -dimensional subspaces of the m -dimensional Euclidean space.

In general, there are 3 ways to define the distance between two planes $\mathbf{P}, \mathbf{Q} \in \mathbb{G}_{m,n}(\mathbb{C})$. Let us first define the principal vectors and principal angles. We assume that there are two sets of n vectors $\mathbf{u}_i \in \mathbb{C}^m, i = 1, \dots, n$ and $\mathbf{v}_i \in \mathbb{C}^m, i = 1, \dots, n$ corresponding to the planes \mathbf{P} and \mathbf{Q} , respectively. If they satisfy $\mathbf{u}_i^H \mathbf{u}_i = \mathbf{v}_i^H \mathbf{v}_i = 1$ and $\mathbf{u}_i^H \mathbf{u}_j = \mathbf{v}_i^H \mathbf{v}_j = 0$, for $i \neq j$, then the vectors \mathbf{u}_i and \mathbf{v}_i are the principal vectors of the planes \mathbf{P} and \mathbf{Q} , respectively. The angles calculated from the inner product of \mathbf{u}_i and \mathbf{v}_i are the principal angles $\theta_1, \dots, \theta_n \in [0, \pi/2]$ (i.e., $\cos \theta_i = |\mathbf{u}_i^H \mathbf{v}_i|$) between these planes.

$$\text{Geodesic distance: } d_g(\mathbf{P}, \mathbf{Q}) = \sqrt{\theta_1^2 + \dots + \theta_n^2}$$

$$\text{Chordal distance: } d_c(\mathbf{P}, \mathbf{Q}) = \sqrt{\sin^2 \theta_1 + \dots + \sin^2 \theta_n} = \frac{1}{\sqrt{2}} \|\mathbf{P}\mathbf{P}^H - \mathbf{Q}\mathbf{Q}^H\|_{\mathcal{F}}$$

$$\text{Maximum geodesic distance: } d_m(\mathbf{P}, \mathbf{Q}) = \max_{i=1, \dots, n} \theta_i$$

The geodesic distance and the maximum geodesic distance have a common drawback that they are not everywhere differentiable. In contrast, the chordal distance has the expected differentiability (details are found in [CHS96]).

E.3. Proof of Theorem 5.3.7

Considering the averaged individual throughput over the fading distribution, the rate $R_i = \mathbb{E}\{\log_2(1 + \text{SINR}_i)\}$ can be transmitted to user i if Gaussian inputs are used. Thus, the throughput loss $\Delta R_i(P_T)$ associated to user i can be written as

$$\begin{aligned} \Delta R_i(\tilde{P}_T) &= R_i(\tilde{P}_T) - R_{i,\text{LF}}(\tilde{P}_T) & (\text{E.3}) \\ &= \mathbb{E}\left\{\log_2\left(1 + \frac{\frac{\tilde{P}_T}{K} \|\mathbf{H}_i \cdot \mathbf{F}_i\|_{\mathcal{F}}^2}{1 + \Delta I_i}\right)\right\} - \mathbb{E}\left\{\log_2\left(1 + \frac{\frac{\tilde{P}_T}{K} \|\mathbf{H}_i \cdot \hat{\mathbf{F}}_i\|_{\mathcal{F}}^2}{1 + \Delta I_i + \sum_{j \neq i} \frac{\tilde{P}_T}{K} \|\mathbf{H}_i \cdot \hat{\mathbf{F}}_j\|_{\mathcal{F}}^2}\right)\right\} \\ &\stackrel{(1)}{=} \mathbb{E}\left\{\log_2\left(1 + \Delta I_i + \frac{\tilde{P}_T}{K} \|\mathbf{H}_i \cdot \mathbf{F}_i\|_{\mathcal{F}}^2\right)\right\} - \underbrace{\mathbb{E}\{\log_2(1 + \Delta I_i)\}}_{(a)} \\ &\quad - \mathbb{E}\left\{\log_2\left(1 + \Delta I_i + \frac{\tilde{P}_T}{K} \|\mathbf{H}_i \cdot \hat{\mathbf{F}}_i\|_{\mathcal{F}}^2 + \underbrace{\sum_{j \neq i} \frac{\tilde{P}_T}{K} \|\mathbf{H}_i \cdot \hat{\mathbf{F}}_j\|_{\mathcal{F}}^2}_{(b)}\right)\right\} \\ &\quad + \mathbb{E}\left\{\log_2\left(1 + \Delta I_i + \sum_{j \neq i} \frac{\tilde{P}_T}{K} \|\mathbf{H}_i \cdot \hat{\mathbf{F}}_j\|_{\mathcal{F}}^2\right)\right\} \\ &\stackrel{(2)}{\leq} \mathbb{E}\left\{\log_2\left(1 + \Delta I_i + \sum_{j \neq i} \frac{\tilde{P}_T}{K} \|\mathbf{H}_i \cdot \hat{\mathbf{F}}_j\|_{\mathcal{F}}^2\right)\right\} \end{aligned}$$

$$\begin{aligned}
 & \stackrel{(3)}{\leq} \log_2(1 + \Delta I_i + \frac{\tilde{P}_T}{K} \sum_{j \neq i} \mathbb{E} \{ \|\mathbf{H}_i \cdot \hat{\mathbf{F}}_j\|_{\mathcal{F}}^2 \}) \\
 & \stackrel{(4)}{\approx} \log_2(1 + \Delta I_i + \frac{\tilde{P}_T(K-1)}{K} (\Delta I_i + D(\mathcal{C}))) \\
 & \stackrel{(5)}{\leq} \log_2(1 + \Delta I_i + \tilde{P}_T \cdot (\Delta I_i + D(\mathcal{C})))
 \end{aligned}$$

Here, the matrix \mathbf{F}_i is the RBD precoding matrix for user i which is calculated from the perfect CSI and $\hat{\mathbf{F}}_i$ denotes the RBD precoding matrix of user i calculated from the quantized CSI. The term ΔI_i is the residual MUI introduced by the RBD precoding (i.e., MUI cannot be entirely eliminated by RBD). After (1) we first neglect the positive terms (a) and (b). Then, regarding the property of $\|\mathbf{A}\|_{\mathcal{F}}^2 = \text{tr}(\mathbf{A}\mathbf{A}^H)$ and the fact that both \mathbf{F}_i and $\hat{\mathbf{F}}_i$ are unitary matrices, we can get $\|\mathbf{H}_i \cdot \mathbf{F}_i\|_{\mathcal{F}}^2 = \|\mathbf{H}_i \cdot \hat{\mathbf{F}}_i\|_{\mathcal{F}}^2$, which leads to bound (2). Then we use Jensen's inequality to get bound (3). Note that $\mathbb{E} \{ \|\mathbf{H}_i \cdot \hat{\mathbf{F}}_j\|_{\mathcal{F}}^2 \}$ is the MUI caused by the fact that the precoding matrix $\hat{\mathbf{F}}_i$ is calculated from the quantized CSI. It can be upper bounded as follows

$$\begin{aligned}
 \mathbb{E} \{ \|\mathbf{H}_i \cdot \hat{\mathbf{F}}_j\|_{\mathcal{F}}^2 \} &= \mathbb{E} \{ \|(\hat{\mathbf{H}}_i + \boldsymbol{\epsilon}_i) \cdot \hat{\mathbf{F}}_j\|_{\mathcal{F}}^2 \} \\
 &= \mathbb{E} \{ \|(\hat{\mathbf{H}}_i \cdot \hat{\mathbf{F}}_j + \boldsymbol{\epsilon}_i \cdot \hat{\mathbf{F}}_j)\|_{\mathcal{F}}^2 \} \\
 &\leq \mathbb{E} \{ \|\hat{\mathbf{H}}_i \cdot \hat{\mathbf{F}}_j\|_{\mathcal{F}}^2 + \|\boldsymbol{\epsilon}_i \cdot \hat{\mathbf{F}}_j\|_{\mathcal{F}}^2 \} \\
 &\approx (\Delta I_i + D(\mathcal{C}))
 \end{aligned}$$

where $\boldsymbol{\epsilon}_i$ is the error introduced by the quantization of \mathbf{H}_i . At step (4), we neglect $\frac{K-1}{K}$ and reach bound (5).

Bibliography

Patent

- [DCL⁺13] M. Dong, Y. Cheng, S. Li, F. Roemer, J. Zhang, B. Song, and M. Haardt, “Linear precoding method and device for communication system,” Patent WO/2013/029 561, 03 07, 2013. [Online]. Available: <http://patentscope.wipo.int/search/en/WO2013029561>

Publications as Co- or First Author

- [CLZ⁺13] Y. Cheng, S. Li, J. Zhang, F. Roemer, B. Song, M. Haardt, Y. Zhou, and M. Dong, “An efficient and flexible transmission strategy for the multi-carrier multi-user MIMO downlink,” *IEEE Transactions on Vehicular Technology*, vol. 93, pp. 2462–2473, 2013.
- [KTS⁺10] P. Komulainen, A. Toelli, B. Song, F. Roemer, E. Bjoernson, and M. Bengtsson, “CSI acquisition concepts for advanced antenna schemes in the winner + project,” in *Proc. Future Network and Mobile Summit 2010*, Florence, Italy, 2010.
- [LCZ⁺12] S. Li, Y. Cheng, J. Zhang, F. Roemer, B. Song, M. Haardt, Y. Zhou, and M. Dong, “Efficient spatial scheduling and precoding for MC MU MIMO system,” in *Proc. 9th International Symposium on Wireless Communications Systems (ISWCS 2012)*, Paris, France, Aug. 2012.
- [RSS⁺11] F. Roemer, N. Sarmadi, B. Song, M. Haardt, M. Pesavento, and A. B. Gershman, “Tensor-based semi-blind channel estimation for MIMO OSTBC-coded systems,” in *Proc. of the 45th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2011, pp. 449–453.
- [SH09a] B. Song and M. Haardt, “Achievable throughput approximation for RBD precoding,” in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 2821–2824.
- [SH09b] —, “Closed-form coordinated beamforming for multi-user MIMO broadcast channels,” in *Proc. of the 23th Meeting of the Wireless World Research Forum (WWRF)*, Beijing, China, Oct. 2009.
- [SH09c] —, “Effects of imperfect channel state information on achievable rates of precoded multi-user MIMO broadcast channels with limited feedback,” in *Proc. IEEE Int. Conference on Communication (ICC’09)*, Dresden, Germany, 2009, pp. 1–5.

- [SHMK09] B. Song, M. Haardt, T. F. Maciel, and A. Klein, “Multi-user MIMO downlink precoding for time-variant correlated channels,” in *Proc. International ITG Workshop on Smart Antennas (WSA ’09)*, Berlin, Germany, Feb. 2009.
- [SRH08] B. Song, F. Roemer, and M. Haardt, “Efficient channel quantization scheme for multi-user MIMO broadcast channels with rbd precoding,” in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Las Vegas, NV, Apr. 2008, pp. 2389–2392.
- [SRH10a] —, “Blind estimation of simo channels using a tensor-based subspace method,” in *Proc. of the 44th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2010, pp. 8–12.
- [SRH10b] —, “Flexible coordinated beamforming (FlexCoBF) algorithm for the downlink of multi-user MIMO systems,” in *Proc. International ITG Workshop on Smart Antennas (WSA ’10)*, Bremen, Germany, Feb. 2010.
- [SRH10c] —, “Using a new structured joint congruence (STJOCO) transformation of hermitian matrices for precoding in multi-user MIMO systems,” in *Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, TX, Mar. 2010.
- [SRH13a] —, “Flexible coordinated beamforming (FlexCoBF) for the downlink of multi-user mimo systems in single and clustered multiple cells,” *Elsevier Signal Processing*, vol. 93, pp. 2462–2473, 2013.
- [SRH13b] —, “A tensor-based subspace method for blind estimation of MIMO channels,” in *Proc. of the 10th International Symposium on Wireless Communications Systems (ISWCS 2013)*, Ilmenau, Germany, Aug. 2013.
- [YSRH12] A. Yeredor, B. Song, F. Roemer, and M. Haardt, “A sequentially drilled joint congruence (SeDJoCo) transformation with application in blind source separation and multi-user MIMO systems,” *IEEE Trans. on Signal Processing*, vol. 60, pp. 2744–2757, 2012.
- [ZSHdL14a] K. Zu, B. Song, M. Haardt, and R. C. de Lamare, “Iterative coordinate Tomlinson-Harashima precoding design for dimensionality constraint multi-user MIMO downlink,” *submitted to Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, 2014.
- [ZSHdL14b] —, “Lattice reduction aided flexible coordinated beamforming (FlexCoBF) for the multi-user MIMO downlink,” *submitted to Proc. International ITG Workshop on Smart Antennas (WSA 2014)*, 2014.

References by Other Authors

- [AAR01] G. Andrews, R. Askey, and R. Roy, Eds., *Special Functions*. Cambridge University Press, 2001.
- [AB03] C. M. Anderson and R. Bro, “Practical aspects of PARAFAC modeling of fluorescence excitation-emission data,” *Journal of Chemometrics*, vol. 17, pp. 200–215, 2003.
- [ACJH07] J. G. Andrews, W. Choi, N. Jindal, and R. W. Heath Jr., “Overcoming interference in spatial multiplexing MIMO cellular networks,” *IEEE Wireless Communications Magazine*, vol. 14, pp. 95–104, Dec. 2007.
- [ACKY05] E. Acar, S. A. Camtepe, M. S. Krishnamoorthy, and B. Yener, “Modeling and multiway analysis of chatroom tensors,” *Lecture Notes in Computer Science*, vol. 3495, pp. 256–268, 2005.
- [Afs06] B. Afsari, “Simple LU and QR based non-orthogonal matrix joint diagonalization,” *ICA 2006, Springer LNCS 3889*, pp. 1–7, 2006.
- [AHB07] E. E. Abdallah, A. B. Hamza, and P. Bhattacharya, “MPEG video watermarking using tensor singular value decomposition.” in *Proc. of the International Conference on Image Analysis and Recognition (ICIAR’07)*, Lecture Notes in Computer Science, 2007, pp. 772–783.
- [And05] J. G. Andrews, “Interference cancellation for cellular systems: A contemporary overview,” *IEEE Trans. on Wireless Communications*, vol. 12, pp. 19–29, 2005.
- [AS65] M. Abramowitz and I. A. Stegun, Eds., *Handbook of mathematical function with formulas, graphs, and mathematical tables*. New York: Dover, 1965.
- [AV11] V. S. Annapureddy and V. V. Veeravalli, “Sum capacity of MIMO interference channels in the low interference regime,” *IEEE Trans. Information Theory*, vol. 57, pp. 2565–2581, 2011.
- [AYL07] C. K. Au-Yeung and D. J. Love, “On the performance of random vector quantization limited feedback beamforming in a MISO system,” *IEEE Trans. Wireless Communication*, vol. 6, pp. 458–462, Feb. 2007.
- [BCC⁺07] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, “MIMO wireless communications,” *Cambridge University Press*, 2007.
- [BG96] Z. Bai and G. H. Golub, “Bounds for the trace of the inverse and the determinant of symmetric positive definite matrices,” *Baltzer Journals*, Apr. 1996.

- [BH07] F. Boccardi and H. Huang, "Limited downlink network coordination in cellular network," in *Proc. IEEE Int. Symp. on Personal Indoor and Mobile Radio Comm.*, Athens, Greece, 2007.
- [BJ03a] H. Boche and E. A. Jorswieck, "Optimal power allocation and complete characterization of the impact of correlation on the capacity of MISO systems with different CSI at the transmitter," in *Proc. IEEE Int. Symp. Information Theory*, 2003, p. 353.
- [BJ03b] —, "Optimal power allocation for MISO systems and complete characterization of the impact of correlation on capacity," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Hong Kong, Hong Kong, 2003, pp. 373–376.
- [BKP98] D. Boss, K. D. Kammeyer, and T. Petermann, "Is blind channel estimation feasible in mobile communication systems? A study based on GSM," *IEEE Journal on Selected Areas on Communications*, vol. 16, pp. 1479–1492, 1998.
- [BM86] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Audio, Speech, Signal processing*, vol. 34, pp. 1081–1089, 1986.
- [BMWT00] P. W. Baier, M. Meurer, T. Weber, and H. Troger, "Joint transmission (JT), an alternative rationale for the downlink of time division CDMA using multi-element transmission antennas," in *Proc. IEEE Int. Symp. on Spread Spectrum Techniques and Application (ISSSTA'00)*, 2000.
- [BO02] M. Bengtsson and B. Ottersten, "Optimum and suboptimum transmit beamforming," *Handbook of antennas in wireless communications (L. C. Godara, eds.)*, CRC Press, 2002.
- [Bol06] H. Bolcskei, "MIMO-OFDM wireless system: basics, perspectives, and challenges," *IEEE Wireless Communication*, vol. 13, pp. 31–37, 2006.
- [Boy08] R. Boyer, "Decoupled root-MUSIC algorithm for multidimensional harmonic retrieval," in *Proc. of the 9th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Recife, Brazil, Jul. 2008.
- [Bra83] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *Communications, Radar and Signal Processing, IEE Proceedings F*, vol. 130, no. 1, pp. 11–16, 1983.
- [BV04] S. Boyd and L. Vandenberghe, "Convex optimization," *Cambridge University Press*, Mar. 2004.
- [CC70] J. D. Carroll and J. J. Chang, "Analysis of individual differences in multidimensional scaling via an n -way generalization of Eckart-Young decomposition," *Psychometrika*, vol. 35, pp. 283–319, 1970.

-
- [CC96] J.-F. Cardoso and P. Comon, "Equivariant adaptive source separation," *IEEE Transactions on Signal Processing*, vol. 45, pp. 434–444, 1996.
 - [CC00] G. Cinis and J. Cioffi, "A multi-user precoding scheme achieving crosstalk cancellation with application to DSL systems," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Nov. 2000, pp. 1627–1637.
 - [CDG00] S. Catreux, P. Driessen, and L. Greenstein, "Simulation results for an interference-limited multiple-input multiple-output cellular system," *IEEE Communication Letters*, vol. 4, pp. 334–336, 2000.
 - [CG87] M. Costa and A. E. Gamal, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Trans. Information Theory*, vol. 33, pp. 710–711, 1987.
 - [CHA07] R. Chen, R. W. Heath Jr., and J. G. Andrews, "Transmit selection diversity for unitary precoded multiuser spatial multiplexing systems with linear receivers," *IEEE Transaction Signal Processing*, vol. 55, pp. 1159–1171, Mar. 2007.
 - [CHHT13] C.-B. Chae, T. Hwang, R. W. Heath Jr., and V. Tarokh, "Interference aware-coordinated beamforming system in a two-cell environment," *IEEE Journal on Selected Areas in Communications*, vol. 27, Nov. 2013.
 - [CHM06] C.-B. Chae, R. W. Heath Jr., and D. Marzzarese, "Achievable sum rate bounds of zero-forcing based linear multi-user MIMO systems," in *Proc. Allerton conf. on Comm. Control and Comp.*, Sep. 2006.
 - [CHS96] J. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian Space," *Experimental Mathematics*, vol. 5, pp. 139–159, 1996.
 - [CJ10] P. Comon and C. Jutten, Eds., *Handbook of Blind Source Separation*. Academic Press, 2010.
 - [CKH09] C.-B. Chae, S. Kim, and R. W. Heath Jr., "Network coordinated beamforming for cell-boundary users: Linear and non-linear approaches," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, pp. 1094–1105, Dec. 2009.
 - [CMIH08a] C.-B. Chae, D. Marzzarese, T. Inoue, and R. W. Heath Jr., "Coordinated beamforming for the multiuser MIMO broadcast channel with limited feedforward," *IEEE Trans. on Signal Processing*, vol. 56, pp. 6044–6056, 2008.
 - [CMIH08b] —, "Non-iterative multiuser MIMO systems with limited feedforward," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing (ICASSP)*, Las Vegas, NV, Apr. 2008.
-

- [CMJH08] C.-B. Chae, D. Marzzarese, N. Jindal, and R. W. Heath Jr., “Coordinated beamforming with limited feedback in MIMO broadcast channel,” *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 1505–1515, 2008.
- [Cor01] L. M. Correia, “Wireless flexible personalized communications-COST 259: European co-operation in mobile radio research,” *New York: John Wiley and Sons*, 2001.
- [Cos83] M. Costa, “Writing on dirty paper,” *IEEE Trans. on Inform. Theory*, vol. 29, pp. 439–441, May 1983.
- [CS00] G. Caire and S. Shamai, “On achievable rates in a multi-antenna broadcast channel,” in *Proc. 38th Annual Allerton Conf. Communications, Control, Computing*, Oct. 2000, pp. 1188–1193.
- [CTKV02] C. Chuah, D. Tse, J. Kahn, and R. Valenzuela, “Capacity scaling in MIMO wireless systems under correlated fading,” *IEEE Transactions on Information Theory*, vol. 48, pp. 637–650, 2002.
- [CV93] R. Cheng and S. Verdu, “Gaussian multi-access channels with ISI: Capacity region and multi-user water-filling,” *IEEE Trans. on Inform. Theory*, vol. 39, pp. 773–785, May 1993.
- [dCRHdSJ11] J. P. C. L. da Costa, F. Roemer, M. Haardt, and R. T. de Sousa Jr., “Multi-dimensional model order selection,” *EURASIP Journal on advances in Signal Processing*, vol. 26, 2011.
- [DHS03] G. Del Galdo, M. Haardt, and C. Schneider, “Geometry-based channel modelling of MIMO channels in comparison with channel sounder measurements,” *Advances in Radio Science - Kleinheubacher Berichte*, pp. 117–126, more information on the model, as well as the source code and some exemplary scenarios can be found at <http://tu-ilmenau.de/ilmprop>, Oct. 2003.
- [dL11] R. C. de Lamare, “Joint power allocation and interference mitigation techniques for cooperative spread spectrum systems with multiple relays,” *Springer/ACM Mobile Networks and Applications*, vol. 16, pp. 629–639, 2011.
- [dLdMV00a] L. de Lathauwer, B. de Moor, and J. Vanderwalle, “A multilinear singular value decomposition,” *SIAM Journal Matrix Analysis Application*, vol. 21, pp. 1253–1278, 2000.
- [dLdMV00b] —, “On the best rank-1 and rank- (r_1, r_2, \dots, r_n) approximation of higher-order tensors,” *SIAM Journal Matrix Analysis Application*, vol. 21, pp. 1324–1342, 2000.
- [dLHSN08] R. C. de Lamare, M. Haardt, and R. Sampaio-Neto, “Blind adaptive constrained reduced-rank parameter estimation based on constant modulus design

-
- for CDMA interference suppression,” *IEEE Transactions on Signal Processing*, vol. 856, pp. 2470–2482, 2008.
- [DLR05] W. Dai, Y. Liu, and B. Rider, “Quantization bounds on Grassmann manifolds of arbitrary dimensions and MIMO communication with feedback,” in *Proc. IEEE GLOBECOM*, 2005.
- [dLSNH07] R. C. de Lamare, R. Sampaio-Neto, and A. Hjørungnes, “Joint iterative interference cancellation and parameter estimation for cdma systems,” *IEEE Communication Letters*, vol. 114, pp. 916–918, 2007.
- [DY10] H. Dahrouj and W. Yu, “Coordinated beamforming for the multicell multi-antenna wireless system,” *IEEE Wireless Communications*, vol. 9, pp. 1748–1759, May 2010.
- [Dyr04] M. Dyrholm, Ed., *Some matrix results*. Website, Aug. 2004.
- [DZ04] S. Dégerine and A. Zaïdi, “Separation of an instantaneous mixture of Gaussian autoregressive sources by the exact maximum likelihood approach,” *IEEE Trans. on Signal Processing*, vol. 52, no. 6, pp. 1492–1512, 2004.
- [Ett76] W. V. Etten, “Maximum-likelihood receiver for multiple channel transmission systems,” *IEEE Trans. on Communication*, vol. 24, no. 2, pp. 276–283, 1976.
- [EY36] C. Eckart and G. Young, “The approximation of a matrix by another of lower rank,” *Psychometrika*, vol. 1, pp. 211–218, 1936.
- [EZ04] U. Erez and R. Zamir, “Achieving $1/2\log(1+\text{SNR})$ on the AWGN channel with lattice encoding and decoding,” *IEEE Trans. on Inform. Theory*, vol. 50, no. 10, pp. 2293–2314, 2004.
- [FG98] 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, pp. 311–335, 1998.
- [FGH07] M. Fuchs, G. D. Galdo, and M. Haardt, “Low-complexity space-time-frequency scheduling for MIMO systems with SDMA,” *IEEE Trans. on Vehicular Technology*, vol. 56, pp. 2775–2784, 2007.
- [FWLH02] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, “Space-time transmission using Tomlinson-Harashima precoding,” in *Proc. 4th Int. ITG Conference on Source and Channel Coding*, Jan. 2002, pp. 139–147.
- [FZJJ06] G. Fettweis, E. Zimmermann, V. Jungnickel, and E. Jorswieck, “Challenges in future short range wireless systems,” *IEEE Vehicular Technology Magazine*, vol. 1, pp. 24–31, 2006.
-

- [GHW⁺11] X. Ge, H. Hung, C. X. Wang, X. Hong, and X. Yang, "Capacity analysis of a multi-cell multi-antenna cooperative cellular network with co-channel interference," *IEEE Transaction on Wireless Communication*, vol. 10, pp. 3298–3309, 2011.
- [GJJV03] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Fundamental capacity of MIMO channels," *IEEE Journal on Selected Areas in Communications, Special Issue on MIMO systems*, vol. 21, 2003.
- [GKH⁺07] D. Gesbert, M. Kountouris, R. W. Heath Jr., C. B. Chae, and T. Saelzer, "Shifting the MIMO paradigm," *IEEE Signal Processing Magazine*, vol. 24, pp. 36–46, 2007.
- [GL96] G. H. Golub and C. F. V. Loan, "Matrix computations," 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
- [GL97] A. Gorokhov and P. Loubaton, "Subspace-based techniques for blind separation of convolutive mixtures with temporally correlated sources," *IEEE Trans. on Circuits and Systems*, vol. 44, pp. 813–820, 1997.
- [GMK10] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "Interference alignment for the K user MIMO interference channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2010.
- [Gol05] A. Goldsmith, "Wireless communications," Cambridge University Press, 2005.
- [Haz01] M. Hazewinkel, Ed., *Encyclopedia of Mathematics*. Springer, 2001.
- [HN89] D. Hatzinakos and C. Nikias, "Estimation of multipath channel response in frequency selective channels," *IEEE Journal on Selected Areas in Communications*, vol. 7, pp. 12–19, 1989.
- [HN95] M. Haardt and J. A. Nossek, "Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden," *IEEE Transactions on Signal Processing*, vol. 43, pp. 1232–1242, 1995.
- [HRD08] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD based subspace estimation to improve the parameter estimation accuracy in multi-dimensional harmonic retrieval problems," *IEEE Trans. on Signal Processing*, vol. 56, pp. 3198–3213, 2008.
- [HTR04] M. Haardt, R. S. Thomae, and A. Richter, "Multidimensional high-resolution parameter estimation with application to channel sounding," in *High-Resolution and Robust Signal Processing*, Y. Hua, A. Gershman and Q. Chen, Eds., Marcel Dekker, New York, NY, Chapter 5, 2004, pp. 255–338.
- [Hua96] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Processing*, vol. 44, pp. 661–672, 1996.

-
- [HV04] H. Huang and S. Venkatesan, “Asymptotic downlink capacity of coordinated cellular networks,” in *Proc. of Asilomar Conference on Signals, Systems, and Computing*, 2004, pp. 850–855.
- [HW93] S. V. Hanly and P. A. Whiting, “Information-theoretic capacity of multi-receiver networks,” *Telecommunications Systems*, vol. 1, pp. 1–42, 1993.
- [ISTrg] W. I. IST-4-027756, “D6.13.7, WINNER II test scenarios and calibration cases issue 2,” Framework Programme 6, Tech. Rep. v1.0, 2007. [online]. Available: <https://www.ist-winner.org/>.
- [JB04] E. Jorswieck and H. Boche, “Optimal transmission strategies and impact of correlation in multi-antenna systems with different types of channel state information,” *IEEE Transactions on Signal Processing*, vol. 52, pp. 3440–3453, 2004.
- [JB06] —, “Performance analysis of MIMO systems in spatially correlated fading using matrix-monotone functions,” *IEEE Transactions on Fundamentals*, pp. 1454–1472, 2006.
- [JBU04] M. Joham, J. Brehmer, and W. Utschick, “MMSE approaches to multiuser spatio-temporal Tomlinson-Harashima precoding,” in *Proc. 5th Int. ITG Conference on Source and Channel Coding*, Jan. 2004, pp. 387–394.
- [JG01] S. A. Jafar and A. Goldsmith, “On optimality of beamforming for multiple antenna systems with imperfect feedback,” in *Proc. IEEE Int. Symp. on Information Theory (ISIT)*, 2001, pp. 321–321.
- [JG05a] S. Jafar and A. Goldsmith, “Multiple antenna capacity in correlated Rayleigh fading with channel covariance information,” *IEEE Transactions on Wireless Communications*, vol. 4, pp. 990–997, 2005.
- [JG05b] N. Jindal and A. Goldsmith, “Dirty-paper coding versus TDMA for MIMO broadcast channels,” *IEEE Trans. on Inf. Theory*, vol. 51, no. 5, pp. 1783–1794, 2005.
- [Jin05] N. Jindal, “High SNR analysis of MIMO broadcast channels,” in *Proc. IEEE Int. Symp. on Information Theory (ISIT)*, Adelaide, Australia, Sep. 2005.
- [Jin06] —, “MIMO broadcast channels with finite rate feedback,” *IEEE Trans. Information Theory*, vol. 52, pp. 5045–5059, Nov. 2006.
- [JLL09] M. Jin, G. Liao, and J. Li, “Joint DOD and DOA estimation for bistatic MIMO Radar,” *Signal Processing*, vol. 89, pp. 244–251, 2009.
- [JRV⁺05] N. Jindal, W. Rhee, S. Vishwanath, S. Jafar, and A. Goldsmith, “Sum power iterative water-filling for multi-antenna Gaussian broadcast channels,” *IEEE Transaction on Information Theory*, vol. 51, Apr. 2005.

- [JSO08] E. Jorswieck, P. Svedman, and B. Ottersten, "On the performance of TDMA and SDMA based opportunistic beamforming," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 4058–4063, 2008.
- [JVG01] N. Jindal, S. Vishanath, and A. Goldsmith, "On the duality of multiple-access and broadcast channels," in *Proc. Allerton Conf. Communications, Control, and Computing*, Oct. 2001.
- [KB09] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, Sep. 2009.
- [KLL⁺09] K. Kim, J. Lee, C. Lee, N. Jeon, and S. Kim, "Coordinated beamforming with limited BS cooperation for multicell multiuser MIMO broadcast channel," in *Proc. IEEE Vehicular Technology Conference (VTC) Spring*, Barcelona, Spain, Apr. 2009.
- [Knu98] D. Knuth, "Teach calculus with big O," *Notices of the American Mathematical Society*, vol. 45, p. 687, 1998.
- [KS01] T. Karp and G. Schuller, "Joint transmitter/receiver design for multicarrier data transmission with low latency time," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Salt Lake, USA, May 2001, pp. 7–11.
- [KZH08] J. H. Kim, W. Zirwas, and M. Haardt, "Efficient feedback via subspace based channel quantization for distributed cooperative antenna systems with temporally-correlated channels," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, pp. 1–13, 2008.
- [LBG80] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on Communications*, vol. 28, pp. 84–95, Jan. 1980.
- [LC08] L. D. Lathauwer and J. Castaing, "Blind identification of underdetermined mixtures by simultaneous matrix diagonalization," *IEEE Trans. Signal Processing*, vol. 56, pp. 1096–1105, 2008.
- [LHS03] D. Love, R. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Information Theory*, vol. 49, pp. 2735–2747, Oct. 2003.
- [LHSH04] D. Love, R. Heath, W. Santipach, and M. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Communication Magazine*, vol. 42, pp. 54–59, Oct. 2004.
- [LJ07] J. Lee and N. Jindal, "High SNR analysis for MIMO broadcast channels: Dirty paper coding vs. linear precoding," *IEEE Trans. on Inf. Theory*, vol. 53, pp. 4787–4792, 2007.

-
- [LJS03] D. J. Love, R. W. H. Jr., and T. Strohmer, “Grassmannian beamforming for multiple-input multiple-output wireless systems,” *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2735–2747, 2003.
- [LLL82] A. K. Lenstra, H. W. Lenstra, and L. Lovász, “Factoring polynomials with rational coefficients,” *Math. Ann.*, vol. 261, pp. 515–534, 1982.
- [LT02] A. Lozano and A. Tulino, “Capacity of multiple-transmit multiple-receive antenna architectures,” *IEEE Trans. Information Theory*, vol. 48, pp. 3117–3128, 2002.
- [LTV05] A. Lozano, A. M. Tulino, and S. Verdú, “High-SNR power offset in multiantenna communications,” *IEEE Transactions on Information Theory*, vol. 51, pp. 4134–4150, 2005.
- [MdCDB99] B. Muquet, M. de Courville, P. Duhamel, and V. Buzenac, “A subspace based blind and semi-blind channel identification method for OFDM systems,” in *Second IEEE Workshop on SPAWC*, Annapolis, MD, USA, 1999, pp. 170–173.
- [MDCM95] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, “Subspace methods for the blind identification of multichannel FIR filters,” *IEEE Trans. on Signal Processing*, vol. 43, pp. 516–525, 1995.
- [MF11] P. Marsch and G. P. Fettweis, “Coordinated multi-point in mobile communications from theory to practice,” *Cambridge University Press*, 2011.
- [MHS04] M. Milojević, M. Haardt, and V. Stanković, “A subspace-based channel model for frequency selective time variant MIMO channels,” in *Proc. of 15th Int. Symp. on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Barcelona, Spain, 2004.
- [MK09] A. S. Motahari and A. K. Khandani, “Capacity bounds for the Gaussian interference channel,” *IEEE Trans. Information Theory*, vol. 55, pp. 620–643, 2009.
- [MKF06] P. Marsch, S. Khattak, and G. Fettweis, “A framework for determining realistic capacity bounds for distributed antenna systems,” in *IEEE Inf. Theory Workshop (ITW’06)*, Punta del Este, Uruguay, 2006.
- [MLM97] K. A. Meraim, P. Loubaton, and E. Moulines, “A subspace algorithm for certain blind identification problems,” *IEEE Transactions on Information Theory*, vol. 43, pp. 499–511, 1997.
- [MRS⁺12] P. Marsch, B. Raaf, A. Szufarska, P. Mogensen, H. Guan, M. Färber, S. Redana, K. Pedersen, and T. Kolding, “Future mobile communication networks: Challenges in the design and operation,” *IEEE Vehicular Technology Magazine*, vol. 7, pp. 16–23, 2012.
-

- [MSEA03] K. Muekkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Information Theory*, vol. 49, pp. 2562–2579, Oct. 2003.
- [Nef04] N. Nefedov, "On subspace channel estimation in multipath SIMO and MIMO channels," in *11th International Conference on Telecommunications (ICT)*, Fortaleza, Brazil, 2004, pp. 532–540.
- [NS10] D. Nion and N. D. Sidiropoulos, "Tensor algebra and multilinear harmonic retrieval signal processing for MIMO Radar," *IEEE Transactions on Signal Processing*, vol. 58, pp. 5693–5705, 2010.
- [NST⁺07] M. Narandžić, C. Schneider, R. Thomä, T. Jämsä, P. Kyösti, and X. Zhao, "Comparison of SCM, SCME, and WINNER channel models," in *Proc. Vehicular Technology Conference VTC 2007-Spring*, Dublin, Apr. 2007, pp. 413–417.
- [OC10] C. Oestges and B. Clerckx, "MIMO wireless communications from real-world propagation to space-time code design," *Academic Press*, 2010.
- [OHW⁺03] H. Ozcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of "Kronecker" MIMO radio channel model," *Electron. Letter*, vol. 39, pp. 1209–1210, 2003.
- [OYN00] S. Ohmori, Y. Yamao, and N. Nakajima, "The future generations of mobile communications based on broadband access technologies," *IEEE Communication Magazine*, vol. 38, pp. 134–142, 2000.
- [PDF⁺08] S. Parkvall, E. Dahlman, A. Furuskar, Y. Jading, M. Olsson, S. Wanstedt, and K. Zangi, "LTE-advanced - evolving LTE towards IMT-advanced," in *Proc. IEEE 68th Vehicular Technology Conference (VTC Fall 2008)*, Calgary, AB, Canada, Sep. 2008.
- [PF91] B. Porat and B. Friedlander, "Blind equalization of digital communication channels using high-order moments," *IEEE Trans. Signal Processing*, vol. 39, pp. 522–526, 1991.
- [PG97] D.-T. Pham and P. Garat, "Blind separation of mixture of independent sources through a quasi-maximum likelihood approach," *IEEE trans. on Signal Processing*, vol. 45, no. 7, pp. 1712–1725, 1997.
- [PNG03] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*, Cambridge University Press 2003.
- [RBHW09] F. Roemer, H. Becher, M. Haardt, and M. Weis, "Analytical performance evaluation for HOSVD-based parameter estimation schemes," in *Proc. IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2009)*, Aruba, Dutch Antilles, Dec. 2009, pp. 77–80.

-
- [RFH08] F. Roemer, M. Fuchs, and M. Haardt, "Distributed MIMO systems with spatial reuse for high-speed-indoor mobile radio access," in *Proc. of the 20-th Meeting of the Wireless World Research Forum (WWRF)*, Ottawa, ON, Canada, Apr. 2008.
- [rGPPG03] 3rd Generation Partnership Project (3GPP), "Spatial channel model for MIMO simulations," *3GPP TR 25.996 V6.1.0*, Available: <http://www.3gpp.org>, 2003.
- [RH10] F. Roemer and M. Haardt, "Tensor-based channel estimation (TENCE) and iterative refinements for two-way relaying with multiple antennas and spatial reuse," *IEEE Transactions on Signal Processing*, vol. 58, pp. 5720–5735, 2010.
- [RJ07] N. Ravindran and N. Jindal, "MIMO broadcast channels with block diagonalization and finite rate feedback," in *Proc. IEEE ICASSP*, Apr. 2007.
- [RSFBB10] M. Renfors, P. Siohan, B. F.-Boroujeny, and F. Bader, "Filter banks for next generation multicarrier wireless communications," *EURASIP Journal on Advances in Signal Processing*, 2010.
- [SBG00] N. Sidiropoulos, R. Bro, and G. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Processing*, vol. 48, pp. 2377–2388, 2000.
- [SCA⁺06] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Sum capacity of multiuser MIMO broadcast channels with block diagonalization," in *Proc. IEEE Int. Symp. on Information Theory (ISIT)*, Jul. 2006.
- [Sch23] I. Schur, "Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie," *Sitzungsber. Berlin. Math. Ges.*, vol. 22, pp. 9–20, 1923.
- [SE07] B. Savas and L. Elden, "Handwritten digit classification using higher order singular decomposition," *Pattern Recognition*, vol. 40, pp. 993–1003, 2007.
- [SH04a] W. Santipach and M. Honig, "Asymptotic capacity of beamforming with limited feedback," in *Proc. of Int. Symp. Inform. Theory*, 2004, p. 290.
- [SH04b] V. Stankovic and M. Haardt, "Multi-user MIMO downlink precoding for users with multiple antennas," in *Proc. of the 12th Meeting of the Wireless World Research Forum (WWRF)*, Toronto, ON, Canada, Nov. 2004.
- [SH05a] W. Santipach and M. Honig, "Signature optimization for CDMA with limited feedback," *IEEE Trans. Information Theory*, vol. 51, pp. 3475–3492, Oct. 2005.
- [SH05b] V. Stanković and M. Haardt, "Multi-user MIMO downlink beamforming over correlated MIMO channels," in *Proc. International ITG/IEEE Workshop on Smart Antennas (WSA'05)*, Apr. 2005.
-

- [SH05c] V. Stankovic and M. Haardt, "Successive optimization Tomlinson-Harashima precoding (SO THP) for multi-user MIMO systems," in *Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Philadelphia, PA, Mar. 2005, pp. 1117–1120.
- [SH08] —, "Generalized design of multi-user MIMO precoding matrices," *IEEE Trans. on Wireless Communications*, vol. 7, pp. 953–961, 2008.
- [SHGJ06] V. Stankovic, M. Haardt, S. Gale, and A. Jeffries, "Linear and non-linear multi-user MIMO downlink precoding," in *Proc. of the 17th Meeting of the Wireless World Research Forum (WWRF)*, Heideberg, Germany, Nov. 2006.
- [SK00] G. Schuller and T. Karp, "Modulated filter banks with arbitrary system delay: efficient implementations and the time-vary case," *IEEE Transactions on Signal Processing*, vol. 48, pp. 737–748, 2000.
- [SKMG05] G. Schuller, J. Kovačević, F. Masson, and V. K. Goyal, "Robust low-delay audio coding using multiple description," *IEEE Transactions on Speech and Audio Processing*, vol. 13, pp. 1014–1024, 2005.
- [SPSH04] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Communications Magazine, special issue on MIMO Systems*, pp. 60–67, 2004.
- [SS96] G. D. T. Schuller and M. J. T. Smith, "New framework for modulated perfect reconstruction filter banks," *IEEE Transactions on Signal Processing*, vol. 44, pp. 1941–1954, 1996.
- [SSH04] Q. H. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multi-user MIMO channels," *IEEE Trans. Signal Processing*, vol. 52, pp. 461–471, 2004.
- [SSP01] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. on Communication*, vol. 49, no. 12, pp. 2198–2206, 2001.
- [STF06] J. Sun, D. Tao, and C. Faloutsos, "Beyond streams and graphs: Dynamic tensor analysis," in *Proc. of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'06)*, Philadelphia, PA: ACM Press, Aug. 2006.
- [SV01] S. Shamai and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. on Inf. Theory*, vol. 47, May 2001.
- [SW97] S. Shamai and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels: Part I," *IEEE Trans. Information Theory*, vol. 43, pp. 1877–1894, 1997.

-
- [SXLK98] B. Suard, G. Xu, H. Liu, and T. Kailath, "Uplink channel capacity of space-division-multiple-access schemes," *IEEE Transactions on Information Theory*, vol. 44, pp. 1468–1476, 1998.
- [SZ01] S. Shamai and B. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmit end," in *Proc. IEEE Vehicular Tech. Conf. (VTC)*, 2001, pp. 1745–1749.
- [Tel95] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Bell Labs Technical Journal*, 1995.
- [TGR09] R. Tresch, M. Guillaud, and E. Riegler, "On the achievability of interference alignment in the K user constant MIMO interference channel," in *Proc. IEEE Workshop Stat. Signal Processing*, Cardiff, U.K., Sep. 2009.
- [THG09] A. Thakre, M. Haardt, and K. Giridhar, "Single snapshot spatial smoothing with improved effective array aperture," *IEEE Signal Processing Letters*, vol. 16, pp. 505–508, 2009.
- [TJC99] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block code from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, 1999.
- [TP98] L. Tong and S. Perreau, "Multichannel blind identification: from sunspace to maximum likelihood methods," *Proceeding of the IEEE*, vol. 86, pp. 1951–1968, 1998.
- [TUBN05] P. Tejera, W. Utschick, G. Bauch, and J. Nossek, "A novel decomposition technique for multi-user MIMO," in *Proc. 6th International ITG Conference on Source and Channel Coding (ITG SCC'05)*, Apr. 2005.
- [Tuc66] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, pp. 279–311, 1966.
- [Tug95] J. K. Tugnait, "Blind equalization and estimation of digital communication FIR channels using cumulant matching," *IEEE Trans. Signal Processing*, vol. 45, pp. 1240–1245, 1995.
- [TV04] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communication," *now Publishers, Hanover, MA*, 2004.
- [TVP94] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple co-channel digital signals using an antenna array," *IEEE Signal Processing Letter*, vol. 1, pp. 29–31, 1994.
- [TXK91] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *Proc. 25th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 1991, pp. 856–860.

- [TY09] P. Tichavský and A. Yeredor, “Fast approximate joint diagonalization incorporating weight matrices,” *IEEE Trans. on Signal Processing*, vol. 57, no. 3, pp. 878–891, 2009.
- [vdB94] A. van den Bos, “Complex Gradient and Hessian,” *Proc. IEE Vision, Image and Signal Processing*, vol. 141, no. 6, pp. 380–383, 1994.
- [vdV01] A.-J. van der Veen, “Joint diagonalization via subspace fitting techniques,” in *Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP’01)*, 2001.
- [Ven07] S. Venkatesan, “Coordinating base stations for greater uplink spectral efficiency in a cellular network,” in *Proc. IEEE Int. Symp. on Personal Indoor and Mobile Radio Comm.*, Athens, Greece, 2007.
- [VJ98] B. R. Vojoic and W. M. Jang, “Transmitter precoding in synchronous multiuser communications,” *IEEE Transactions on Communications*, vol. 46, pp. 1346–1355, 1998.
- [VJG02] S. Vishwanath, N. Jindal, and A. Goldsmith, “On the capacity of multiple input multiple output broadcast channels,” in *Proc. Int. Conf. Communications*, Apr. 2002, pp. 1444–1450.
- [VJG03] —, “Duality, achievable rates, and sum-rate capacity of MIMO broadcast channels,” *IEEE Trans. on Inf. Theory*, vol. 49, pp. 2658–2668, 2003.
- [VPW10] L. Venturino, N. Prasad, and X. Wang, “Coordinated linear beamforming in downlink multi-cell wireless networks,” *IEEE Wireless Communications*, vol. 9, pp. 1451–1461, Apr. 2010.
- [VS08] J. Via and L. Santamaria, “On the blind identifiability of orthogonal space-time block codes from second-order statistics,” *IEEE Trans. on Information Theory*, vol. 54, pp. 709–722, 2008.
- [VT02] M. A. O. Vasilescu and D. Terzopoulos, “Multilinear analysis of image ensembles: Tensor-faces,” *Lecture Notes in Computer Science*, vol. 2350, pp. 447–460, 2002.
- [VT03] —, “Multilinear subspace analysis of image ensembles,” in *Proc. of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’03)*, IEEE Computer Society, 2003, pp. 93–99.
- [Win87] J. Winters, “On the capacity of radio communication systems with diversity in a Rayleigh fading environment,” *IEEE J. Select. Areas Commun.*, vol. 5, pp. 871–878, 1987.
- [WIN05] “IST-2003-507581 WINNER: D5.4: Final report on link level and system level channel models,” 2005, <http://www.ist-winner.org>.

-
- [WIN06] “IST-4-027756 WINNER II: D1.1.1 WINNER II interim channel models,” Nov. 2006, <https://www.ist-winner.org/deliverables.html>.
- [WIN07] “IST-4-027756 WINNER II: D1.1.2 WINNER II channel models,” Dec. 2007, <https://www.ist-winner.org/deliverables.html>.
- [WJG⁺10] M. Weis, D. Jannek, T. Guenther, P. Husar, F. Roemer, and M. Haardt, “Temporally resolved multi-way component analysis of dynamic sources in event-related eeg data using parafac2,” in *Proc. 18-th European Signal Processing Conference (EUSIPCO 2010)*, Aalborg, Denmark, Aug. 2010, pp. 696–700.
- [Wyn94] A. D. Wyner, “Shannon-theoretic approach to a Gaussian channels,” *IEEE Trans. on Information Theory*, vol. 40, pp. 1713–1727, 1994.
- [WZ07] O. Weikert and U. Zölzer, “Efficient MIMO channel estimation with optimal training sequences,” in *Proc. of the 1st Workshop on Commercial MIMO-Components and -Systems (CMCS 2007)*, Duisburg, Germany, 2007.
- [WZZ⁺05] Y. Wu, J. Zhang, H. Zheng, X. Xu, and S. Zhou, “Receive antenna selection in the downlink of multiuser MIMO systems,” in *Proc. Vehicular Technology Conference (VTC)*, 2005, pp. 477–481.
- [YC01] W. Yu and J. Cioffi, “Trellis precoding for broadcast channel,” in *Proc. Global Communications Conf.*, Oct. 2001, pp. 1344–1348.
- [Yer00] A. Yeredor, “Blind separation of Gaussian sources via second-order statistics with asymptotically optimal weighting,” *IEEE Signal Processing Letters*, vol. 7, no. 7, pp. 197–200, 2000.
- [Yer09] —, “On hybrid exact-approximate joint diagonalization,” in *Proc. IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP’09)*, Aruba, Dutch Antilles, Dec. 2009, pp. 312–315.
- [Yer10] —, “Blind separation of Gaussian sources with general covariance structures: Bounds and optimal estimation,” *IEEE Transactions on Signal Processing*, vol. 58, pp. 5057–5068, 2010.
- [Yer12] —, “Performance analysis of the strong uncorrelating transformation in blind separation of complex-valued sources,” *IEEE Trans. on Signal Processing*, vol. 60, pp. 2744–2757, 2012.
- [YG06] T. Yoo and A. Goldsmith, “On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming,” *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 528–541, 2006.
- [YP03] S. Yatawatta and A. P. Petropulu, “Blind channel estimation in MIMO OFDM systems,” in *Proc. IEEE Workshop Statistical Signal Processing*, St. Louis, MO, 2003, pp. 363–366.
-

- [YRBC01] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative water-filling for vector multiple access channels," in *Proc. of IEEE Int. Symp. Information Theory (ISIT)*, 2001.
- [YRBC04] X. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative water-filling for Gaussian vector multiple access channels," *IEEE Trans. on Inf. Theory*, vol. 50, pp. 145–152, 2004.
- [YRC01] W. Yu, W. Rhee, and J. Cioffi, "Optimal power control in multiple access fading channels with multiple antennas," in *Proc. of IEEE Int. Conf. Communications*, 2001, pp. 575–579.
- [ZCA08] J. Zhang, R. Chen, and J. G. Andrews, "Networked MIMO with clustered linear precoding," *IEEE Trans. on Wireless Communication*, vol. 8, pp. 1910–1921, 2008.
- [ZdL12] K. Zu and R. C. de Lamare, "Low-complexity lattice reduction-aided regularized block diagonalization for MU-MIMO systems," *IEEE Communications Letters*, vol. 16, pp. 925–928, 2012.
- [ZdLH12] K. Zu, R. C. de Lamare, and M. Haardt, "Multi-branch Tomlinson-Harashima precoding for single-user MIMO systems," in *Proc. ITG/IEEE Workshop on Smart Antennas*, Dresden, Germany, Mar. 2012.
- [ZdLH13a] ———, "Multi-branch Tomlinson-Harashima precoding design for MU-MIMO systems: Theory and algorithms," *accepted by IEEE Trans. on Communication*, 2013.
- [ZdLH13b] K. Zu, R. C. de Lamare, and M. Haart, "Generalized design of low-complexity block diagonalization type precoding algorithms for multiuser MIMO systems," *accepted by IEEE Trans. on Communications*, 2013.
- [ZHV08] Z. Zhou, W. Hardjawana, and B. Vucetic, "Iterative multiple beamforming algorithm for MIMO broadcast channels," *IEEE Communication Letters*, vol. 12, pp. 743–745, Oct. 2008.
- [ZMG02] S. Zhou, B. Muquet, and G. Giannakis, "Subspace-based (semi)blind channel estimation for block precoded space-time OFDM," *IEEE Trans. on Signal Processing*, vol. 50, pp. 1215–1228, 2002.
- [ZSE02] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. on Inform. Theory*, vol. 48, no. 6, pp. 1220–1230, 2002.
- [ZT03] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental trade-off in multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 49, pp. 1073–1096, 2003.